Available online at www.sciencedirect.com
ScienceDirect
Cognition 107 (2008) 433–459
COGNITION
www.elsevier.com/locate/COGNIT

Ordered short-term memory differs in signers and speakers: Implications for models of short-term memory

Daphne Bavelier *, Elissa L. Newport, Matt Hall, Ted Supalla, Mrim Boutla

Department of Brain and Cognitive Sciences, University of Rochester, Meliora Hall Box 270268, Rochester, NY 14627-0268, USA

Received 31 January 2007; revised 2 October 2007; accepted 28 October 2007

Abstract

Capacity limits in linguistic short-term memory (STM) are typically measured with forward span tasks in which participants are asked to recall lists of words in the order presented. Using such tasks, native signers of American Sign Language (ASL) exhibit smaller spans than native speakers ([Boutla, M., Supalla, T., Newport, E. L., & Bavelier, D. (2004). Short-term memory span: Insights from sign language. *Nature Neuroscience, 7*(9), 997–1002]). Here, we test the hypothesis that this population difference reflects differences in the way speakers and signers maintain temporal order information in short-term memory. We show that native signers differ from speakers on measures of short-term memory that require maintenance of temporal order of the tested materials, but not on those in which temporal order is not required. In addition, we show that, in a recall task with free order, bilingual subjects are more likely to recall in temporal order when using English than ASL. We conclude that speakers and signers do share common short-term memory processes. However, whereas short-term memory for spoken English is predominantly organized in terms of temporal order, we argue that this dimension does not play as great a role in signers' short-term memory. Other factors that may affect STM processes in signers are discussed.

* Corresponding author. Tel.: +1 585 275 8714; fax: 1 585 442 9216.
E-mail address: daphne@cvs.rochester.edu (D. Bavelier).

0010-0277/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.cognition.2007.10.012

Keywords: Short-term memory; Deafness; American Sign Language

1. Introduction

Linguistic short-term memory (STM) refers to the ability to maintain and manipulate linguistic information over short periods of time. A central feature of linguistic STM is its limited capacity, whereby only a handful of items can be temporarily maintained and manipulated. One of the most common measures of capacity limits in STM is the digit span task, where subjects must repeat lists of digits in the same order as they are presented (i.e. forward serial recall). The number of to-be-recalled digits is progressively increased, and the span is defined as the longest sequence recalled correctly. Recently, using such a measure of the span, we have found that users of American Sign Language (ASL – the natural gestural language used by Deaf[1] people in the United States and parts of Canada) have a shorter span than that of English speakers. Given prior evidence establishing that speakers and signers rely on formational/phonological encoding and use a sub-articulatory mechanism to rehearse signs in STM (for a review, see Wilson & Emmorey, 2000), this earlier work focused on showing that the advantage of speech over signed information in serial recall span tasks could not be explained by factors such as phonological similarity, phonological complexity or duration of articulation (Bavelier, Newport, Hall, Supalla, & Boutla, 2006; Boutla, Supalla, Newport, & Bavelier, 2004). A similar argument has been put forward by Ronnberg, Rudner, and Ingvar (2004), who observed no correlation between recall rate and performance on a serial STM task in Swedish/Swedish Sign Language bilinguals. In the present paper, we test the hypothesis that the span difference noted in earlier studies between English and ASL reflects a difference in the retention of the relative temporal order of the items, and not a generalized difference in short-term memory for linguistic materials.

The advantage of speech over signed information in serial span tasks may arise because signers and speakers encode order information in different manners. Speakers might rely predominantly on temporal encoding, while signers might use more varied encoding, including spatial encoding. Previous literature supports this view (Carey & Blake, 1974; Hanson, 1990). For example, O'Connor and Hermelin (1972) presented deaf and hearing children with three digits, one after the other on a horizontal line, but with the temporal order of presentation incongruent with the spatial left-right order in which the digits were arranged. Subjects were asked to report back the digits in writing. Hearing subjects recalled the digits in their temporal order of presentation, whereas deaf individuals reproduced them as they appeared from left to right. In a follow-up experiment, these authors showed that this behavior did not just reflect a biased strategy for responding; performance on

[1] While the lowercase deaf refers to the audiological condition of not hearing, the uppercase *Deaf* refers to a particular group of deaf people who share a culture and communicate primarily using ASL (Padden & Humphries, 1988).

a recognition task was also better for temporal order in hearing and for spatial order in the deaf children. In a later study, these authors documented higher backward span in deaf children than in hearing children, when scoring order recall based on both correct spatial location and temporal order (O'Connor and Hermelin, 1976). This finding again supports the view that spatial encoding may be favored in the deaf population. Interestingly, when item recall was scored independently of order (spatial and temporal), deaf and hearing children displayed equivalent performance. If, as suggested by these studies, hearing speakers rely more readily on temporal order but deaf signers on spatial order, the difference in standard serial span tasks between these populations might arise, at least in part, from the span task requirement of recalling items in serial order, combined with the use of stimuli with a clear temporal pattern but little, if any, spatial patterning (see Wilson, 2001 for review).

The proposal that the STM difference between signers and speakers reflects a difference in temporal order processing, rather than a more general difference in short-term memory processes, is further supported by the finding that signers and speakers exhibit comparable memory capacity limits in tasks that do not require serial order recall. For example, using free instead of ordered recall, Hanson and collaborators have shown equivalent short-term memory capacities across deaf signers and hearing speakers (Hanson, 1982, 1990). More recently, we have shown that signers and speakers exhibit equivalent performance on a linguistic STM task which requires active maintenance and on-line manipulation of linguistic information, but no maintenance of serial order. In this STM task, participants were presented with a list of words and were asked to recall each of the words in a separate, self-generated sentence (Boutla et al., 2004). Importantly, order of recall was free. Hearing speakers and deaf signers performed equivalently. A similar effect has been described in bilinguals of Swedish and Swedish Sign Language, who demonstrated equivalent performance in the two languages when asked to perform semantic retrieval of information, a task that depends heavily on executive function and the frontal lobes (Thompson-Schill, 2003), but still displayed lower serial STM span in Swedish Sign Language than in Swedish (Ronnberg et al., 2004). This pattern of results indicates that the advantage of spoken over signed information in STM is not present in all STM tasks. Rather, we have proposed that it is restricted to tasks requiring encoding, maintenance and/or recall of temporal order information.

Although the view that signers and speakers differ in their ability to perform tasks that call for temporal order encoding and maintenance is consistent with a number of previous studies (Bellugi & Siple, 1974; Bonvillain, AlthausRea, Orlansky, & Slade, 1987; Hachinski et al., 1975; Krakow & Hanson, 1985; Wilson, 2001), it does stand in contrast to some reports in the literature. First, Wilson, Bettger, Niculae, and Klima (1997) compared English speakers and Deaf native ASL signers on a backward span task using digits. The backward span task is identical to the forward span task described above, except that participants are asked to recall the presented list in the reverse order. The backward span task therefore requires both maintenance and manipulation of temporal order information. Wilson et al. (1997) reported a significantly higher backward span for native ASL signers than for native

English speakers, casting doubt on the proposal that maintenance of temporal order differs between signers and speakers. Wilson and Emmorey (2006a,b) recently proposed that the difference between signers and speakers on the serial order recall task we have reported may be due to the use of digits to test speakers, but letters to test signers (a stimulus difference whose purpose was to equate phonological similarity among the materials across the two languages). According to Wilson & Emmorey's view, digits as a semantic category have a preferential treatment in short-term memory and thus lead to higher spans.

The goal of the present paper is to revisit the issue of span and temporal order maintenance in signers versus speakers (i) by considering the impact of stimulus choice when performing cross-linguistic comparisons, (ii) by assessing forward and backward spans in native English speakers and native ASL signers, and (iii) by contrasting ordered versus item recall in signers and speakers.

2. Experiment 1: Digit spans – deaf signers versus hearing speakers

Experiment 1 focused on the forward and backward spans of adult native English speakers and adult Deaf native ASL signers, using the identical digit span task as that used in the Wilson et al. (1997) study, which reported greater backward span for signers than for speakers. It is important to note that, unlike Wilson et al. (1997) in which the participants were children of about 10 years of age, this study focuses on adult participants. Indeed, the findings of Wilson et al. might reflect differences in developmental stages between Deaf signers and Hearing speakers. The present study compares forward and backward spans in adult native English speakers and adult native ASL signers. In addition, Experiment 1 asks whether the use of digits, rather than letters as in our past experiments, may increase the span of signers and thus reduce the population difference previously noted in Boutla et al. (2004).

2.1. Methods

2.1.1. Participants

Twelve adult Deaf native signers were recruited from the Rochester, NY area (National Technical Institute of the Deaf at the Rochester Institute of Technology and University of Rochester) and at Gallaudet University, Washington, D.C. (mean age = 22 years old). All participants were exposed to ASL from birth (through their Deaf parents or older siblings who knew ASL). All participants considered ASL to be their primary language and used ASL daily.

Twenty adult native English speakers (mean age = 20 years old) were also included in the present experiment. The English speakers were recruited from Monroe Community College and from the University of Rochester (Rochester, NY). None of the English speakers were familiar with ASL. Informed consent was obtained from each participant. All participants were paid for their participation.

2.1.2. Stimuli

2.1.2.1. ASL digit span. The digit span stimuli included two sets of digit lists, one used for the forward sequences and another for the backward sequences. All sets were taken from the WAIS (Wechsler, 1955). The lists increased in length from two to nine digits in the forward condition, and from two to eight in the backward condition. As in the standard procedure, two trials per length were presented in both forward and backward conditions. All of the lists were produced by a hearing native ASL/English bilingual and videotaped. The signer signed the digits maintaining a neutral facial expression throughout list presentation. An auditory metronome was used to ensure an accurate presentation rate of approximately 1 item/s (mean stimulus presentation rate = 1.2 items/s). Some mouth movements were produced, but without any voice. The last item was marked by a longer hold. Based on the intuition of fluent signers, this longer hold is analogous to the drop of voice required in the spoken version of the digit span task.

2.1.2.2. English digit span. The lists were similar to those used in the ASL condition and enunciated by the same native ASL/English bilingual. All lists were videotaped. The items were enunciated at the standard rate of 1 item per second, and a visual metronome was used to ensure a precise timing of enunciation (mean stimulus presentation rate = 1.1 items/s).

2.1.3. Design and procedure

Each participant was tested individually. All stimuli were displayed on a Macintosh G3 Powerbook with a 14-in. monitor running Psyscope 1.2.5 (Cohen, MacWhinney, Flatt, & Provost, 1993). At the beginning of the session, participants watched a movie in which a native ASL/English bilingual gave the instructions for the experiment (in ASL for the signers and in English for the speakers). Following the standard WAIS procedure (Wechsler, 1955), all participants started with the forward recall condition first. Each participant saw a series of trials in which a list of digits was presented via videotape (in spoken English for the English speakers, in ASL for the signers), and the participant was to repeat back the sequence of digits in the same order as presented. If they could not remember every item, they were asked to sign 'zero' at the position of the forgotten digit (or say 'zero' for the English speakers). The ASL signers were also given two practice trials of two items, in order to make sure that they understood the task.

In both the ASL and English conditions, the participant initiated the trial by pressing the space bar. After the end of each sequence presentation, the participant recalled the target sequence. As in the standard WAIS version of the digit span, the participant was first tested on two trials of two-item sequences, then on two trials of three-item sequences, and so forth. Testing continued until the participant failed both trials of the same length.

Immediately after having completed the forward version, each participant was then tested on the backward portion of the experiment. In this condition, each participant saw a series of trials in which a list of digits was presented via videotape, and then had to repeat the sequence of digits in the reverse order as that presented. At the beginning of the condition a movie was shown explaining the backward recall requirement. Then an example of a sequence of two items was given to insure that the participant understood the instructions.

In both forward and backward conditions, the participant's performance was videotaped, using a Sony TVR-900 DV camera, to allow for offline coding of the performance. As in the standard digit span task, the forward and backward spans of ASL signers and English speakers were scored as the longest list length at which a correct recall in one or both of the recall attempts at that length was observed.

2.2. Results

2.2.1. Impact of language modality and direction of recall on span

The forward and backward digit spans of the English native speakers and Deaf native ASL signers were defined as the last list length tested before participants incorrectly recalled both trials (Fig. 1). Although this measure does not allow direct comparison with the published WAIS score tables (which score the total number of items correctly recalled), it better captures the notion of capacity limit by providing a measurement of the list length subjects can recall in each population. In our previous work, we have measured both span and WAIS score and found that they are highly correlated and provide a similar picture of between-group differences. Therefore, all analyses reported in this paper will be based on span measurements.

A 2×2 analysis of variance (ANOVA) on spans, with language (English vs. ASL) as a between-subject factor and direction of recall (forward vs. backward) as a within-subject factor, revealed a significant main effect of language $[F(1, 30) =$

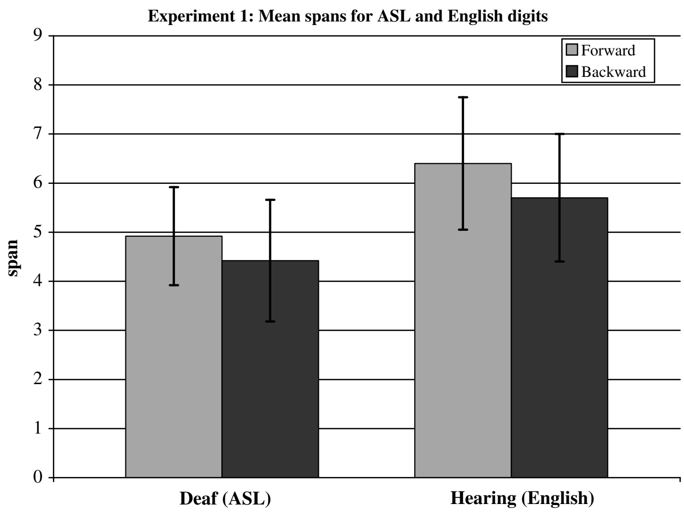

Fig. 1. Forward and backward digit spans (and standard deviations) in deaf signers and hearing non-signers.

11.92, $p < .005$, $p\eta^2 = .284$]. This result is consistent with previous findings documenting higher serial recall spans in English speakers than in deaf ASL signers (Bellugi, Klima, & Siple, 1974-1975; Boutla et al., 2004; Marschark & Mayer, 1998). The significant main effect of direction of recall [$F(1,30) = 7.17$, $p < .02$, $p\eta^2 = .193$] indicated that forward serial spans were higher than backward serial spans, as expected. The interaction between language modality and direction of recall did not reach significance [$F(1,30) < 1$, $p\eta^2 = .007$], establishing that the difference between forward and backward span is comparable across speakers and signers.

2.2.2. Impact of language modality

2.2.2.1. Forward span across ASL and English. The mean forward span of signers (mean = 4.92, $SD = 1.00$) was significantly shorter than that of the speakers (mean = 6.40, $SD = 1.35$) [$F(1,30) = 10.83$, $p < .003$, $p\eta^2 = .265$].

2.2.2.2. Backward span across ASL and English. To be directly comparable to Wilson et al. (1997), we further investigated the impact of language modality on the backward span by contrasting the mean backward span of signers (mean = 4.42, $SD = 1.24$) to that of the speakers (mean = 5.70, $SD = 1.3$). In contrast with Wilson et al. (1997), native adult Deaf ASL signers had a lower backward digit span than native adult English speakers [$F(1,30) = 7.55$, $p < .02$, $p\eta^2 = .201$].

2.3. Discussion

Experiment 1 compared the forward and backward serial span in signers and speakers using digits as stimuli. As in our previous studies, the forward serial span was found to be significantly smaller in signers than in speakers by about 2 items. Importantly, this group difference was observed not only for the forward span, but also for the backward span. Thus the difference in span size between English and ASL users does not appear limited to the forward serial span, but also generalizes to the backward span. This finding is consistent with the proposal that signers and speakers differ on short-term memory tasks which require temporal order recall.

The finding of a lower backward span in ASL than in English stands in contrast to the earlier observation of Wilson et al. (1997) in 10-year-old children. It is possible that developmental differences may be present between Deaf signers and hearing speakers in the learning of order reversal, accounting for this differing pattern of results between children and adults. This should be an interesting avenue of research for future studies.

The results of Experiment 1 establish that using identical digit stimuli across populations does not alleviate the population difference in serial spans. Native signers were still observed to have lower spans than native speakers. Although on the surface it would seem ideal to use stimuli which are literal translations from one language to the other when doing cross-linguistic comparisons, doing so generally overlooks a major complication: measures of short-term memory span are highly dependent on phonological factors such as phonological similarity across items and overall phonological complexity of each item. These factors need to be tightly controlled when comparing across languages. Translating items from one language to the other does not guarantee appropriate matching along these well-characterized phonological properties (and is often impossible to achieve under these circumstances) (see Bavelier et al., 2006). In fact, this problem in the use of digits to compare capacity across English speakers and ASL signers may explain, at least in part, the lower span of signers. The signs representing digits in ASL are more phonologically similar than are the words representing digits in English (see Fig. 2). Since span size is known to be reduced for lists of words that are phonologically similar both in spoken (Baddeley, Lewis, & Vallar, 1984) and in signed STM tasks (Wilson & Emmorey, 1997), the lower performance of deaf signers in Experiment 1 may have been due to the use of digits in signers. Experiment 2 addresses this issue. In addition, Experiment 2 controls for the possibility that the lower span noted in Experiment 1 may be due to reduced mnemonic abilities in the deaf population, by contrasting ASL and English spans in the same individuals – hearing adult native ASL/English bilinguals.

3. Experiment 2: ASL letter span versus English digit span – hearing, adult native ASL/English bilinguals

Experiment 2 was designed to measure the forward and backward span in adult Deaf native ASL signers while controlling for the possible impact of the higher phonological similarity among ASL digits than among English digits. To control for phonological similarity, a set of 9 ASL finger-spelled letters were selected to minimize phonological similarity (see Fig. 2). This set of 9 finger-spelled letters was used instead of the 1–9 digits in the ASL span tasks. ASL letters were chosen because their enunciation time is short and of comparable duration to that of English digits (Boutla et al., 2004). This point is important, as duration of enunciation has been shown to be a determinant in span tasks across spoken languages (Elliott, 1992; Ellis & Hennelly, 1980; Hoosain & Salili, 1987). Accordingly, the stimuli were chosen to elicit equivalent recall rate in ASL and in English. We acknowledge that this choice of stimuli does not address the issue raised by Wilson and Emmorey, that is, that digits

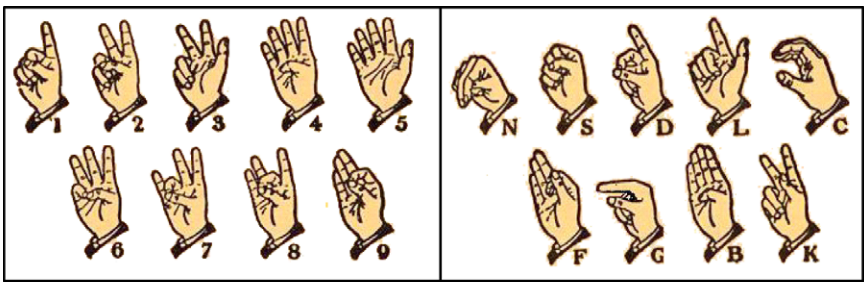

Fig. 2. Hand-shapes of ASL digits used in experiment 1 (high phonological similarity) and ASL finger-spelled letters used in experiment 2 (low phonological similarity).

may be better remembered during short-term tasks than any other materials. We will turn to this issue in Experiment 3.

Experiment 2 controlled for possible effects of hearing status on span size by assessing the forward and backward spans of hearing adult native ASL/English bilinguals (CoDAs). All participants had Deaf signing parents from whom they learned ASL from birth, and none were trained as professional interpreters, as the effects of interpreting on span size remain largely unknown. The forward span part of this experiment was reported in Boutla et al. (2004).

3.1. Methods

3.1.1. Participants

Twenty hearing native ASL/English bilinguals were recruited among hearing Children of Deaf Adults (CoDAs. $N = 20$, 16 females, mean age = 40). All participants were exposed to ASL through their Deaf parents and to English through their interactions with the hearing population. All participants attended public or private schools for hearing children, and none of the participants had been trained in interpreting.

3.1.2. Stimuli, design and procedure

For English, each participant was tested using the same materials and procedure as in the English digit span described in Experiment 1.

For ASL, new ASL letter span tasks were created using a set of nine ASL fingerspelled letters that were chosen to minimize phonological similarity in ASL (B, C, D, F, G, K, L, N and S; see Fig. 2). Each letter was arbitrarily paired with a number between '1' and '9' to convert the stimuli of the English digit span (Wechsler, 1955) into an ASL letter span task. In order to avoid chunking, sequences containing meaningful acronyms were modified to produce sequences of meaningless letter strings. All other aspects of the stimuli were identical to Experiment 1. To create the stimuli, a visual and auditory metronome was used in order for a hearing native ASL signer to produce the sequences at the desired rate of 1 item per second (mean stimulus presentation rate = 1.2 items/s). In all conditions, the items were produced without any co-articulation of neighboring items. This absence of co-articulation is important in preventing chunking of the items during encoding and rehearsal in STM.

CoDA participants were tested on two conditions of the ASL span task, only one of which is reported here. One of the span tasks used a rate of 1 item/s and is described below. The other used a faster rate of presentation (3.6 item/s) and led to similar results as the 1 item/s condition but will not be described further (Boutla, 2003).

The order of language tested and the order of tasks within ASL were counterbalanced across participants; none of these effects were significant [in all cases, $Fs(1, 18) < 1.81$, $ps > .19$]. In order to control for interference across memory tasks in different languages, participants were asked to fill out a series of background questionnaires (15–20 min) once they completed all tests within a language and before starting testing in the second language.

In all other respects the procedure was identical to that of Experiment 1.

3.2. Results

3.2.1. Impact of language modality and direction of recall on span

The mean forward and backward spans for ASL letters and English digits obtained for the native ASL/English bilinguals are presented in Fig. 3. A 2×2 ANOVA, with language (English vs. ASL) and direction of recall (forward vs. backward) as within-subject factors, revealed a significant main effect of language modality on the serial span [$F(1,19) = 38.87$, $p < .001$, $p\eta^2 = .672$]. Consistent with our previous findings on forward spans (Boutla et al., 2004), the present results show that, within a native ASL/English bilingual, serial spans are significantly larger for speech than for signs. Similar to Experiment 1, the significant main effect of direction of recall [$F(1,19) = 64.40$, $p < .001$, $p\eta^2 = .772$] indicates that forward spans were larger than backward spans. The lack of interaction between language modality and direction of recall [$F(1,19) = 1.62$, $p > .2$, $p\eta^2 = .079$] suggests that the advantage of spoken language is comparable for forward and backward span.

3.2.2. Impact of language modality

3.2.2.1. Forward span across ASL and English. For completeness we report here the forward span results (already presented in Boutla et al. (2004) Experiment 2). The signed STM span [mean = 5.2, $SD = .83$] was significantly shorter than the spoken one [mean = 7.05, $SD = 1.19$] in ASL/English bilinguals [$F(1,19) = 37.6$, $p < .001$, $p\eta^2 = .665$].

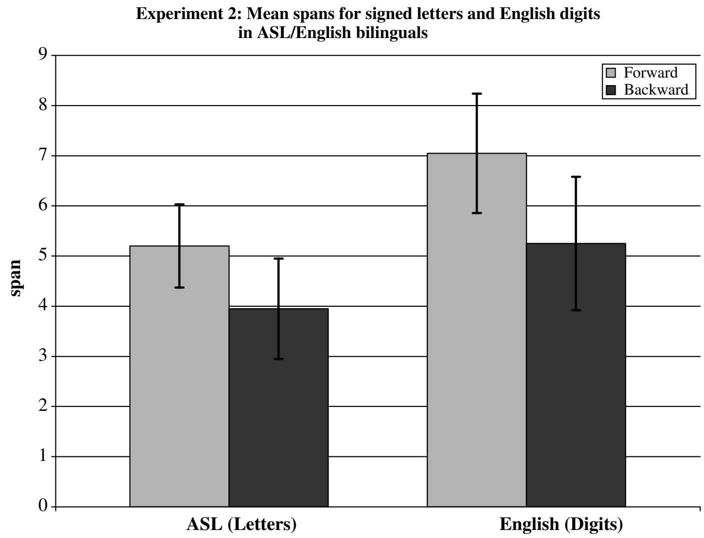

Fig. 3. Forward and backward spans (and standard deviations) in American Sign Language/English bilinguals.

3.2.2.2. Backward span across ASL and English. A direct comparison of English and ASL backward spans in hearing native ASL/English bilinguals showed that the mean English backward span [mean = 5.25, $SD = 1.33$] was significantly higher than the mean ASL backward span [mean = 3.95, $SD = 1$, $F(1,19) = 21.27$, $p < .001$, $p\eta^2 = .528$]. Therefore the present experiment shows that, *within the same individual*, significantly larger backward spans are observed for English than for ASL. These results indicate that the smaller ASL backward span observed in Experiment 1 in Deaf signers compared with hearing speakers is not due to reduced mnemonic abilities in Deaf individuals per se.

3.2.3. Span measures across subject populations

A recurring issue when studying bilingual populations is the extent to which their performance is comparable to that of monolingual native users in each language.[2] To address this question, we compared the performance of our bilingual group in ASL to that of Deaf signers from Experiment 1, and their performance in English to that of hearing speakers from Experiment 1.

3.2.3.1. ASL. A 2×2 ANOVA, with population (Bilingual-ASL vs. Deaf-Experiment 1) and direction of recall (forward vs. backward), was performed on the ASL span measures. A main effect of direction of recall was observed [$F(1,30) = 13.58$, $p < .002$, $p\eta^2 = .312$], indicating greater forward than backward spans across populations. Importantly, however, no effect of population and no interaction with population was observed [Fs $< 1, p\eta^2 < .08$], showing that there was comparable ASL span performance for hearing native ASL/English bilinguals and Deaf native ASL signers.

3.2.3.2. English. A 2x2 ANOVA, with population (Bilingual-English vs. hearing-Experiment 1) and direction of recall (forward vs. backward), was also performed on the English span measures. A main effect of direction of recall [$F(1,38) = 31.49$; $p < .001$, $p\eta^2 = .453$], but no effect of population [$F(1,38) < 1$, $p\eta^2 = .002$], were observed. A significant population by direction of recall interaction [$F(1,38) = 6.09$, $p < .02$, $p\eta^2 = .138$], indicated that the difference between forward and backward span was bigger in bilinguals than in the monolingual hearing participants of Experiment 1. Importantly, separate analyses of the forward and backward spans revealed no effect of population on the forward span [$F(1,38) = 2.60$, $p > .1$, $p\eta^2 = .064$] or on the backward span [$F(1,38) = 1.17$, $p > .2$, $p\eta^2 = .03$], showing that

[2] Poorer performance in ASL than in English in hearing bilinguals may result from lower ASL skills rather than reflect a difference between ASL and English in STM as we propose. If so, individuals with the poorest ASL skills should have the shortest memory span. Post-hoc analyses of self-rating scores for comprehension and production in ASL revealed that participants divided into 3 different groups: strong, intermediate and weak ASL users. When entered in a $3 \times 2 \times 2$ ANOVA with ASL skills, language tested and forward/backward order as factors, the ASL skill factor was not significant and did not interact with any other factors (all ps $> .25$). Thus it is not the case that participants with the weakest ASL skills have the shortest memory span. Although self-rating is not an ideal measure of language use, these results support the view that it is the type of language, and not fluency differences among the groups, which affects STM capacity.

there were also comparable English spans for hearing ASL/English bilinguals and hearing monolingual English speakers.

3.2.4. Control for phonological factors

Recall performance in serial order tasks is known to be inversely related to phonological similarity, phonological complexity and articulatory duration (Baddeley, Thomson, & Buchanan, 1975). To confirm that the materials were phonologically comparable across the two languages, we compared recall rate across languages. Recall rate was defined as the number of items enunciated per second during the recall phase of the forward part of the STM task. We acknowledge that this is not the only way to measure recall rate. In our past work we have also used another measure, known as speeded reading, in which one measures the rate at which participants read a page with the same items as those used in the STM task (Boutla, 2003). Although speeded reading is more common in the literature, recall rate has the advantage of measuring articulatory speed while participants are actually performing the STM task, and thus may be more appropriate to compare rehearsal rate across populations (see for further discussion of this point, Mueller, Seymour, Kieras, & Meyer, 2003). Importantly, in all cases, recall rate and span measures were obtained in the very same participants.

The recall rate was comparable for sign (mean $\pm$ SEM $= 2.9 \pm .18$ items/s) and for English materials (mean $\pm$ SEM $= 2.55 \pm .11$ items/s; $F(1,19) = 3.47$, $p > .07$, $p\eta^2 = .15$). Clearly, the span differences observed above cannot be explained by slower recall rate (that is, slower articulation rate) in sign than in English. If anything, sign enunciation was slightly faster than English enunciation for these materials.

3.3. Discussion

Experiment 2 establishes that both forward and backward spans in ASL are significantly shorter than the corresponding English spans, even when tested within the same bilingual individuals. These results unambiguously show that the span difference across languages is due to language modality differences, and not to reduced mnemonic abilities in Deaf individuals.

It has recently been argued by Wilson and Emmorey (2006a) that the span differences we report here and in our previous work (Boutla, 2003; Boutla et al., 2004) may stem from stimulus selection in each language rather than a true cross-linguistic difference. In particular, Wilson and Emmorey (2006a) have argued that comparing digits in English to letters in ASL may give an unfair advantage to speakers, as higher spans are typically found for digits than for letters in native English speakers. It is indeed the case that digits often lead to higher spans than letters in speakers (for a review see Cavanaugh, 1972; Warrington, Kinsbourne, & James, 1966); however, this finding has typically been attributed to the fact that letters are phonologically similar to one another in English (e.g, *bee*, *dee*, *ee*, *gee*, etc), leading to lists containing high phonological similarity and therefore longer enunciation duration. This phonological difference, associated with longer enunciation duration, has been hypothesized to produce the shorter span. If this were the correct account of the typical difference between digits and letters, then producing materials with comparable phonological properties (and showing comparable articulation rate/enunciation duration) should resolve this problem; remaining differences in span must arise from other sources, not from this difference in materials. However, in contrast to this literature, Wilson and Emmorey (2006a) have proposed that the difference between digits and letters arises from semantic factors, not just phonological properties.

In order to address all of these considerations – both phonological and semantic – we must ask whether span differences can still be observed between signers and speakers while using stimuli that belong to the same semantic category in the two languages, and also while adequately matching the phonological properties of the items across the two languages. To this end, Experiment 3 compares the span of signers to that of speakers using letters that are carefully chosen to be comparable in terms of phonological factors across the two languages.

4. Experiment 3: Letter span – deaf native signers versus hearing speakers

Experiment 3 was designed to compare the forward and backward span in signers and speakers while using items that were matched semantically and phonologically across languages. Although a semantic match can be satisfactorily achieved through translation from one language to another, this method does not insure proper phonological matching. Indeed, given the structure of natural languages, it is unlikely that lists of items that are direct translations of each other will have the same phonological complexity in the two languages. To design such a set, we restricted ourselves to consonants, and through bilingual informants and preliminary studies, selected two sets of letters which appear to give rise to comparable enunciation durations in the two languages. Our final set consisted of the same set of 9 finger-spelled letters used as in Experiment 2, and the following set of 9 English letters: M,Y,S,L,R,K,H,G,P. (Although G and P are phonologically similar, they did not appear in the same list until list length 9. There are not enough letters in English that have no phonological overlap to find a set of 9 letters that are unrelated.) To demonstrate that these two sets were equivalent in terms of phonological factors across languages, duration of articulation was computed for each group.

If, as we have proposed above, the higher English spans observed in our previous experiments was due to an advantage of spoken information over signed information, one would expect higher spans for speakers than for signers in both forward and backward span conditions. The forward span from speakers has already been presented in Bavelier et al. (2006).

4.1. Methods

4.1.1. Participants

Twenty Deaf participants (10 females, mean age = 21 years of age) were recruited from the Deaf population at the National Technical Institute for the

Deaf (Rochester, N.Y.) and at Gallaudet University (Washington, D.C.). All participants were congenitally Deaf and were exposed to ASL from birth through having at least one Deaf parent. All participants considered ASL to be their primary language and used ASL daily. Four of these Deaf native signers had participated in Experiment 1 (digit STM task) several months earlier. Participants were tested on two conditions of the ASL span task, only one of which is reported here. One of the span tasks used a rate of 1 item/s and is described below. The other used a faster rate of presentation (3.6 item/s) and led to similar results as the 1 item/s condition, and therefore will not be described further (Boutla, 2003). Condition order was counterbalanced across participants, and the two conditions were separated by at least 30 min during which participants filled out various questionnaires and participated in unrelated experiments. ANOVAs including condition order as a within-subject variable were performed on the span scores and indicated that there was no impact of task order (effect of presentation rate order: $F(1,18) < 1$).

The hearing group consisted of 20 adult native English speakers (10 females, mean age = 18.9 years of age). The control English speakers were recruited from the campus of the University of Rochester (Rochester, NY). None of the English speakers were familiar with ASL.

Informed consent was obtained from each participant. All participants were paid for their participation.

4.1.2. Stimuli

The measure of ASL forward and backward spans was obtained by using the same materials and procedure used in the ASL letter span tasks of Experiment 2.

The English letter span task was created using a set nine English letters that were chosen to minimize phonological similarity in English (M,Y,S,L,R,K,H,G,P). Two lists from list length 2 up to 9 were constructed following the English digit span structure (Wechsler, 1955) with the constraint that G and P did not appear in the same sequence before length 9. As with the ASL stimuli, sequences forming familiar units (acronyms, words, expressions like “YR”, consecutive alphabetical order) were modified to produce sequences of meaningless letter strings. The English stimuli differed from the ASL materials and the WAIS in only one respect: at list length 9, the ASL and WAIS materials contained one sequence with three repeated letters, whereas such repetitions were not present in our English materials. Although this stimulus difference was not planned, it is unlikely to be the source of the population effect reported below as (i) item repetition is likely to lighten memory load and so increase span size – we are therefore putting speakers in a more demanding situation than signers; (ii) in practice, no signers reached list length 9.

For both languages, the items were enunciated at the standard rate of 1 item per second, and a visual metronome was used to ensure a precise timing of enunciation (mean stimulus presentation rate of 1.1 items/s for English and 1.2 item/s for ASL).

4.2. Results

4.2.1. Impact of language modality and direction of recall on span

The mean forward and backward spans for ASL letters and English letters are presented in Fig. 4. A 2×2 ANOVA with language (English vs. ASL) and direction of recall (forward vs. backward) as between subject factors revealed a significant main effect of language modality on the serial span [$F(1,38) = 23.46$, $p < .001$, $p\eta^2 = .382$]. Consistent with our previous findings, the present results showed that serial spans are significantly longer for speech than for signs even when letters are used in both languages. As in Experiments 1 and 2, the significant main effect of direction of recall [$F(1,38) = 45.08$, $p < .001$, $p\eta^2 = .543$] indicates that forward spans were bigger than backward spans. A significant language × direction interaction [$F(1,38) = 7.81$, $p < .005$, $p\eta^2 = .197$] also indicated that while backward recall reduced the span in both populations, it did so to a greater extent in speakers than in signers.

4.2.2. Impact of language modality

4.2.2.1. Forward span for ASL and English. The signed letter span in deaf individuals [mean = 4.85, $SD = 1.23$] was significantly shorter than the spoken letter span in hearing controls [mean = 6.8, $SD = 1.06$; $F(1,38) = 29.04$, $p < .001$, $p\eta^2 > .5$].

4.2.2.2. Backward span for ASL and English. A direct comparison of English and ASL backward spans for English speakers and deaf signers showed that the mean English backward span [mean = 4.80, $SD = 1.06$] was significantly higher than the

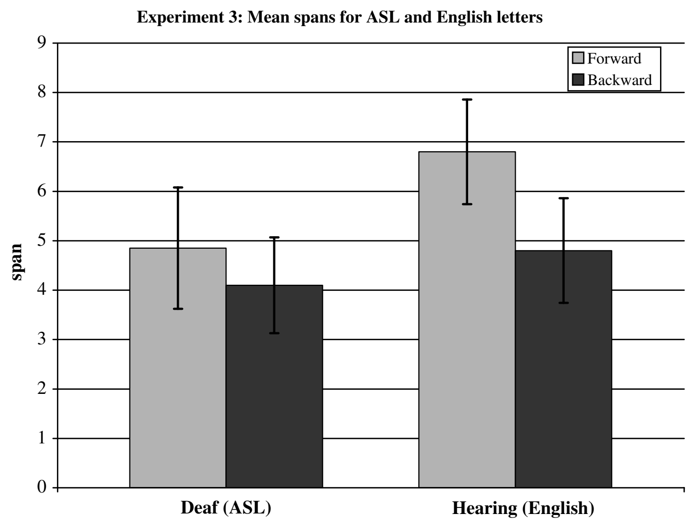

Fig. 4. Forward and backward letter spans (and standard deviation) in deaf native signers and hearing non-signers.

mean ASL backward span [mean = 4.10, $SD = .97$, $F(1,38) = 4.77$, $p < .038$, $p\eta^2 = .112$]. Thus, the present experiment confirms that a significantly longer backward span is observed for English than for ASL.

4.2.3. Controlling for phonological factors

To confirm that the span differences observed between sign and speech are not due to faster articulation rate in English speakers than in Deaf signers, we compared articulation rate in the two languages. We measured recall rate (the number of items enunciated per second) during the recall phase of all correct trials in the forward span task, as in Experiment 2. The recall rate was significantly faster for ASL deaf signers [mean $\pm$ SEM = 3.52 $\pm$ 0.24 items/s] than for English speakers [mean $\pm$ SEM = 2.69 $\pm$ .18 items/s; $F(1,38) = 7.60$, $p < .01$, $p\eta^2 = .167$], unambiguously ruling out a faster recall rate in speakers as the source of their higher span.

4.3. Discussion

The use of stimuli matched in terms of phonological factors and semantic category across languages did not remove the difference in performance between speakers and signers on the serial STM span task. Signers still exhibited a shorter forward span than speakers. Unlike what was observed in Experiments 1 and 2, switching from forward to backward recall led to greater span reduction in speakers than in signers. However, even in the presence of this effect, the backward span of signers was still reliably shorter than that of speakers. Overall, Experiment 3 confirms that ASL signers have significantly shorter forward as well as backward spans as compared to English speakers.

Together Experiments 1–3 support the view that signers and speakers differ in linguistic short-term memory tasks that require temporal order recall. This result was observed for both the forward and backward span tasks, which both require maintenance of temporal order. In addition, this finding was observed in Deaf as well as in hearing signers, alleviating the concern that the span difference could be due to different mnemonic abilities in the Deaf. Rather, the span difference observed in serial order recall tasks appears to emerge from a true cross-linguistic difference. It remains, however, to be established whether differences between signers and speakers are predominantly found when using ordered recall, and therefore might stem from a different reliance on temporal order in the two language modalities.

A review of the available literature does provide support for this view. Serial recall tasks using ASL materials lead to lower performance than those using spoken or printed materials (Bavelier et al., 2006; Bellugi et al., 1974-1975; Bonvillain et al., 1987; Boutla et al., 2004; Hanson, 1982; Krakow & Hanson, 1985). In contrast, the few memory studies that have used a recall paradigm with free order indicate equivalent performance for Deaf signers and hearing speakers (Boutla et al., 2004; Hanson, 1982). In addition, a number of studies indicate that a speech-based code may be especially useful when processing temporal order information (Kanabus, Szelag, Rojek, & Poeppel, 2002; Krakow & Hanson, 1985; Wilson, 2001). Based on a similar argument, Krakow and Hanson (1985) argued that "speech-based code may facilitate serial-order recall in a way that alternative coding mechanisms, including sign-based coding of ASL signs, do not." The goal of Experiment 4 was to further examine the relative role of temporal order during recall of spoken versus signed material, by contrasting item recall and temporal order recall in a recall paradigm with free order.

5. Experiment 4: Maintenance of temporal order in ASL versus English – hearing English/ASL bilinguals

Experiment 4 compares item and temporal order recall in an immediate recall task as a function of the language used. The proposal that signers and speakers differ in their tendency to maintain temporal order predicts equivalent item recall when using sign or speech, but greater spontaneous conservation of temporal order information when recalling in speech than in sign. To be in a better position to assess conservation of temporal order, longer, supra-span lists were used, as in Hanson's work, rather than the short lists of Experiments 1–3 (Hanson, 1982; Krakow & Hanson, 1985). Several other factors conspire to make the WAIS-like tasks used in Experiments 1–3 poor choices to compare maintenance of temporal order in English and ASL. First, the use of a closed set of items (9 items) favors order errors over item errors. Second and more importantly, the tests in Experiments 1–3 were given in the standard WAIS way, that is, testing stopped once the subjects failed to recall in the proper order on both trials tested at each list length. Since testing was stopped based on order error, it is difficult, if not impossible, to recover the span corresponding to item scoring for each population. Indeed, the span corresponding to item scoring is typically higher than the span for order scoring.

Previous research on recall with free order in signers has focused on the serial position curve, which quantifies how well an item is recalled as a function of its serial position in the presented list. This measure does not directly index, however, whether temporal order information is maintained at the time of recall. Positional recall is often used in the STM literature to address the issue of order (Ronnberg and Nilsson, 1987; Franklin & Mewhart, 2002; Beaman & Jones, 1998). For example, Ronnberg and Nilsson (1987) used an input-output order analysis in which one computes the probability that an item at a given input position occurred at a given output position. Although this and related measures of serial order nicely index the recall probability for a particular serial position, they do not directly reflect whether the relative temporal order of the items has been maintained during recall. Indeed, relative temporal order could be maintained with little to no maintenance of positional recall. Of interest here is whether item 4 tends to be recalled before item 5, independently of their exact serial position. To assess this issue, maintenance of relative temporal order was compared when recalling in sign versus in speech. Finally, to ensure that language use can be identified as the source of this temporal order difference, this study was carried out within-subject by testing English/ASL bilinguals.

5.1. Methods

5.1.1. Participants

Seventeen hearing native ASL/English bilinguals were recruited from hearing Children of Deaf Adults (CoDAs, 10 females, mean age = 23.1). All participants were exposed to ASL from their Deaf parents (15 had two deaf parents) and to English through their interactions with the hearing population. All participants had attended public or private schools for hearing children and none of the participants had been trained in interpreting. They were recruited through contacts in Rochester, NY, Washington, DC, and a summer camp in Old Forge, NY.

5.1.2. Stimuli, design and procedure

Materials for the supra-span recall task were selected from a pool of high frequency nouns that were rated as highly concrete, familiar, and imageable in English according to the MRC Psycholinguistics database (mean score > 500, Wilson, 1988). From these words, 32 test items were selected according to the following criteria: (1) that there was a standardized ASL sign for the word; (2) that the sign did not include classifiers or fingerspelling and was not a compound; (3) that the sign was unambiguously a noun; (4) that the word's frequency, concreteness, familiarity, or imageability was judged to be equivalent in ASL and in English by our informants (unfortunately, published ratings on these dimensions for ASL are not available).

These 32 items were then arranged into two lists of 16 items each, and both lists were filmed in ASL and in English by a native speaker of each language. Each subject was presented with all 32 words, but the lists were counterbalanced such that across all subjects, each list was presented equally often in ASL and in English. Presentation was accomplished by displaying a movie clip every 5 s, showing the stimulus model articulating one item. Presentation rate was controlled by using a Matlab script to calculate the length of each stimulus clip, thereby determining the length of the ISI (blank screen) before the next word was presented.

Participants were instructed to watch each 16-word list carefully and to recall as many of the items as they could at the end of each presentation. Their recall cue was a line of asterisks on the screen, signifying the end of presentation. Participants were told at the start of the experiment that there was no penalty for guessing or repetition, that they could recall the items in any order, and that there was no time limit for recall. When they could not recall any more items, they were to inform the experimenter that they were done. Rarely, a subject would remember another item after they had said they were done. In these cases, responses were included as long as the next condition in the experiment had not yet begun (this only occurred on 2 occasions).

5.1.3. Scoring

Two types of scoring measures were used: item scoring and relative order scoring. In either scoring scheme, repetitions and intrusions were first removed from the output list. Under item scoring, an item was counted as correct if it appeared at any point in the subject's recall sequence. This score represents the total number of unique items that the subject was able to recall from the list (with a maximum value of 16).

Under relative order scoring, we scored the number of adjacent response pairs that were in the correct relative order with respect to the presentation list, divided by the number of possible pairs in the subject's responses (see Asch & Ebenholtz, 1962). Such relative order scoring provides a reliable measure of the degree to which subjects are spontaneously using temporal order information, and has been argued to provide a robust method for performing between-study comparisons (Addis & Kahana, 2004; Waugh, 1963). For example, consider the list A B C D E, and a subject's response A C E D A X. After discarding the repetition of A and the intrusion of X, the output list consists of A C E D. Item scoring gives then a score of 4: 1 point for each unique item (A, C, D, and E). For relative order scoring, the two pairs A–C and C–E are in the correct relative temporal order, but the pair E-D is not. Therefore, the relative temporal order score equals 2 correct pairs divided by 3 possible pairs, for a score of 66.67% (out of a maximum value of 100%). Because each subject's relative order score is an indirect function of that subject's overall item performance (the denominator is essentially the item score minus one), this measure has the advantage of correcting for possible differences in item scores when comparing groups.

5.2. Results

5.2.1. Item scoring

The mean number of items recalled was comparable between ASL [mean = 10.47, $SD = 2.29$] and English [mean = 9.53, $SD = 1.84$) – $t(16) = 1.474$, $p = .16$, $p\eta^2 = .12$]. Importantly, what little difference there is indicates, if anything, a trend for better performance in ASL than in English. This stands in contrast to Experiments 1–3, which repeatedly showed lower performance for ASL than for English.

5.2.2. Relative temporal order scoring

Unlike item scoring, relative temporal order scoring revealed a different pattern across the two languages. The mean relative order score for ASL was 60.6%, which was significantly less than the score of 70.9% in English. A Wilcoxon signed-rank test verified that relative temporal order is more likely maintained in English than in ASL (Fig. 5; $W+ = 22$, $W- = 131$, $p < .008$). Thus even though subjects recalled just as many items in ASL as in English, the same subjects relied on temporal order to a greater degree during English recall than during ASL recall.

5.2.3. Serial position curves

For the sake of comparison with previous studies, we also report the serial position curve for each language (Fig. 6). No gross difference can be seen between languages. Calculations of serial order position effects for both languages were made by collapsing the recall scores across four consecutive serial positions: 1–4, 5–8, 9–12, 13–16. The first interval provides a measure of the primacy effect and the last

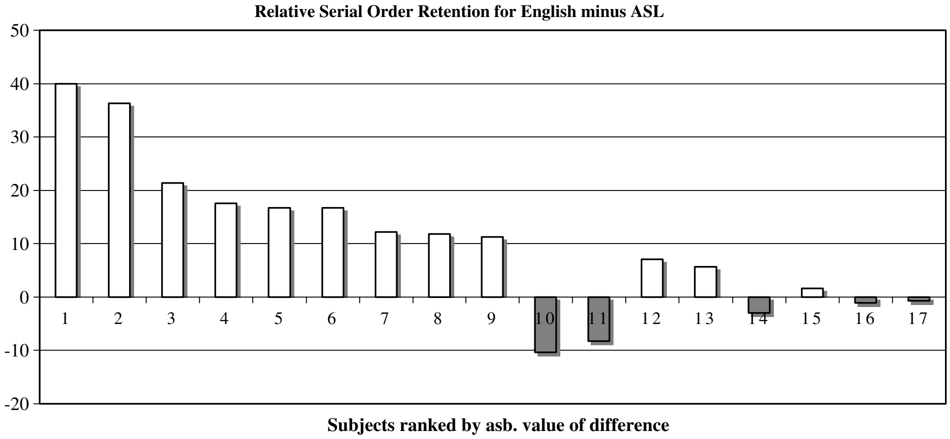

Fig. 5. Retention of relative temporal order was computed for English and ASL lists separately for each of the 17 English/American Sign Language bilinguals. For each subject, this figure plots the difference between the percent order retention in English and that in ASL. A positive score means that temporal order was better maintained in English, whereas a negative score means that temporal order was better maintained in ASL.

one of the recency effect. A 2×4 ANOVA with language (English/ASL) and serial position (1-4/5-8/9-12/13-16) as factors revealed no main effect of language [$p > .16$, $p\eta^2 = .117$], a main effect of serial order [$F(3,48) = 25.16$, $p < .0001$, $p\eta^2 = .614$] and no interaction [$p > .9$, $p\eta^2 = .007$]. The lack of interaction indicates equivalent primacy and recency effects across language modality.[3]

5.3. Discussion

Experiment 4 establishes that when order of recall is free, ASL use leads to comparable memory capacity as English use, supporting the view that the lower spans observed in Experiments 1–3 do not reflect generalized lower memory abilities in signers, but rather differences in the processing or maintenance of temporal order information. This view is further confirmed by the analysis of relative temporal order scoring, which shows that the use of English favors the maintenance of relative temporal order to a greater extent than the use of ASL. This was the case even though subjects were merely instructed to recall as many items as they could, without any mention of item order. The fact that these findings were observed in the very same subjects when asked to perform the task either in ASL or in English clearly demon-

[3] As in Experiment 2, the role of ASL skill on these results was controlled for by grouping subjects according to their self-rating level. In this experiment, participants fell into two main categories: strong and intermediate ASL users. 2×2 ANOVAs with language and ASL skill were carried out on item and relative temporal order scoring separately. Effects of ASL skill were far from significant (all ps $> .55$), reinforcing the view that ASL skill is not a main determinant of performance in these experiments.

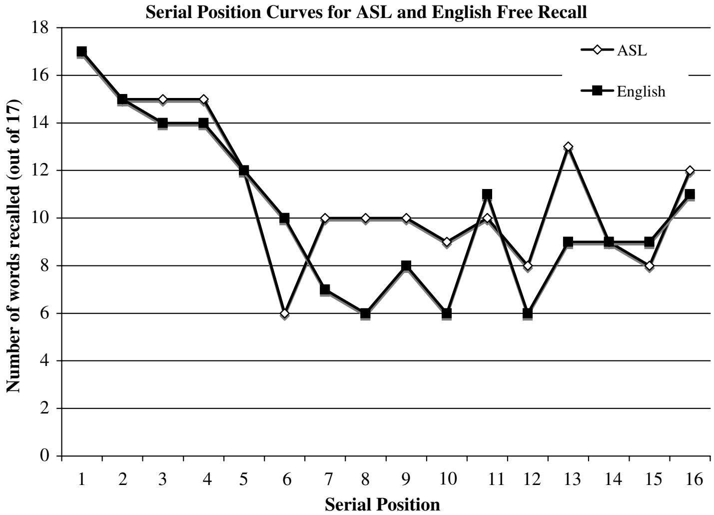

Fig. 6. Serial position curve for English and ASL lists in English/American Sign Language bilinguals ($N = 17$).

strates that it is the nature of the language modality which differentially engages temporal order maintenance.

Finally, serial position curves were equivalent for sign and speech, establishing similar primacy and recency effects across languages. This finding is in line with previous reports in the literature that deaf signers exhibit similar primacy and recency effects as hearing controls (Krakow & Hanson, 1985; Shand, 1980).

6. General discussion

The present studies investigated the hypothesis that signers and speakers differ when asked to recall linguistic stimuli in temporal order, but display equivalent performance on tasks devoid of a temporal order requirement. In Experiments 1–3, smaller ASL than English spans were observed for both forward and backward spans, confirming the previously documented difference in serial spans between signers and speakers. These findings held whether ASL stimuli consisted of digits or letters, and whether the populations tested were Deaf native signers or hearing ASL/English bilinguals. The latter demonstrates that the smaller forward and backward signed spans are not due to lower mnemonic abilities in Deaf native signers, but rather are attributable to an effect of language modality. These results indicate that spoken information leads to higher serial spans than signed information in serial recall tasks regardless of the direction of recall.

This set of studies also shows that in adults the backward span is shorter in ASL than in English. This finding contrasts with a previous study in children which reported similar backward spans in signers and speakers at 10 years of age (Wilson et al., 1997). It is possible that order reversal, a demanding process performed at recall time by hearing individuals, matures relatively slowly, allowing signers, who may rely on a different strategy, to display some advantage early during development. This pattern of findings, if confirmed, would provide additional evidence that the strategies underlying serial recall in signers and speakers are rather different. Clearly, future studies are needed to determine the developmental time course of STM maintenance and manipulation as a function of language experience (Altom & Weill, 1977).

Taken together, the results of Experiments 1–3 are consistent with the view that at least one of the factors relevant to the difference between adult signers and speakers in immediate memory tasks is the requirement to encode and store temporal order information. The speech-based representations used by speakers in linguistic STM are likely to build upon auditory-based representations (Baddeley, 2003), hypothesized to be particularly well suited to maintaining temporal order (Kanabus et al., 2002). As forward and backward serial recall tasks require accurate encoding and rehearsal of temporal order, speech-based representations may be advantaged in these tasks. Speakers have been reported to rely preferentially on speech-based representations in both forward and backward span tasks, reversing order in the latter task only at recall (Rosen & Engle, 1997; Sakai & Passingham, 2003). Additionally, temporal order has been shown to be more efficiently maintained in auditory-based representations than in visually-based representations in STM (Paivio & Csapo, 1971; Watkins, LeCompte, Elliott, & Fish, 1992; Watkins & Watkins, 1980). However, while the visual system is more limited in its ability to retain temporal order information, it may be more efficient in retaining other types of information, such as spatial structure, (O'Connor and Hermelin, 1972). Building on a similar argument, Wilson (2001) has proposed that signers and speakers may encode order information in quite different manners, with speakers relying predominantly on temporal encoding and signers predominantly on spatial encoding.

An advantage of speech-based over sign-based representations when encoding and maintaining temporal order predicts that speakers and signers should not differ in all aspects of linguistic STM, but rather only on those tasks that involve temporal order recall. The present study confirms that signers and speakers differ when using linguistic serial span tasks, but exhibit similar performance for immediate memory tasks that do not require temporal order. In accord with the few studies that have compared signers to speakers in linguistic STM tasks that do not require temporal order (production span – Boutla et al., 2004; e.g. recall with free order – Hanson, 1982; Hanson, 1990), we observed equivalent recall performance whether speech or sign is used. In addition, by analyzing the extent to which the original order of presentation is maintained during recall (even though temporal order is irrelevant to the task), Experiment 4 established that by default subjects maintain original temporal order more readily when recalling in English as compared to recalling in ASL. This result confirms different biases in the two language modalities with respect to temporal order maintenance.

The flip side of this advantage for temporal order in speakers appears to be an advantage for spatial order in signers. Early studies by O'Connor and Hermelin

(O'Connor and Hermelin, 1972, 1976, 1973) document that deaf children more readily maintain spatial organization, whereas hearing children more readily maintain the temporal organization of to-be-remembered items. More recently, performance on the Corsi block task, a spatial memory span task, was found to be better for deaf signers than hearing non-signers (Wilson et al., 1997). Supporting the view that sign language is the source of this effect, deaf children who have no exposure to sign language perform as well as, but not better than, hearing children on the Corsi Block Task (Parasnis, Samar, Bettger, & Sathe, 1996), and hearing children (age 6 years) who attended a year-long course in Italian Sign Language (Lingua Italiana Dei Segni/LIS) exhibited increased spatial memory span on the Corsi Block Task compared to peers who were not in the course (Capirci, Cattani, Rossini, & Volterra, 1998). Although suggestive, the spatial advantage for signers is less well documented in the literature than the temporal order advantage for speakers (Beck, Beck, & Gironella, 1977; Das, 1983 for failure to find a spatial advantage in the deaf). It will be important for future research to establish the extent to which signing favors spatial organization in memory, and the type of visuo-spatial information that may be so fostered.

In light of previous research linking performance on span tasks to other cognitive abilities, in particular language acquisition and processing (Baddeley, Gathercole, & Papagno, 1998; Baddeley & Hitch, 1974; Case, Kurland, & Goldberg, 1982; Daneman & Carpenter, 1980; Gathercole & Baddeley, 1993), one may wonder about the practical implications of a shorter serial span when using linguistic materials in signers. Although the link between serial span task performance and vocabulary development is often stressed in developmental studies, there are reasons to doubt the usefulness of the serial span performance in predicting other cognitive abilities (Gathercole, 1999). In a seminal study, Daneman and Carpenter (1980) established that a storage and manipulation measure of short-term memory, the reading span, predicted the accuracy of text comprehension better than traditional serial span measures (see also Turner & Engle, 1989). Building on this early work, Cowan et al. (2005) have recently argued that the best predictors of cognitive skills and language ability do not appear to be tests such as the serial span task, but rather tasks that combine storage and manipulation of information, such as the O-Span or the reading span task. This view holds that short-term memory tasks which best predict aptitude measures are those which, unlike serial span tasks, do not rely on rehearsal and chunking. Thus, by weighing so heavily on rehearsal and chunking, serial recall of a list of unrelated items may highlight capacities that are highly specialized for ordered recall of speech-based representations, but have little predictive power for cognitive or language skills more generally. This view is further supported by research on short-term memory for materials other than lists of unrelated items. It has been argued that for materials that provide a conceptual frame, such as scenes or sentences, short-term memory performance is based on the gist, or content-addressable memory structures, and not on ordered phonological representations like those used in the recall of random lists (Potter, 1999; Potter & Lombardi, 1990). Supporting this view, recent work indicates that the type of order information necessary during sentence processing is different from the slow, temporal order processing that mediates list recall (Lewis, Vasishth, & Van Dyke, 2006; McElree, Foraker, & Dyer, 2003). Reports that a semantic span measure is a better predictor of reading comprehension than the traditional serial span measures also highlights the greater role of conceptual representations in short-term memory than is suggested by research on serial span (Haarman, Davelaar, & Usher, 2003). Thus, the shorter serial span in signers may not have much significance when it comes to language processing, as the serial span does not engage the same type of order processing and representations as those involved in language comprehension and production. In summary, performance on the serial span task appears to reveal the capacity of a highly specialized system for rote ordered recall of unrelated items (which depends heavily on phonological representations, rehearsal and chunking), but performance of this system seems to have less bearing on other cognitive or language aptitudes than originally thought.

The findings of Experiments 1–4 raise important issues for clinical and educational settings, where serial span tasks remain the most commonly used measures of STM processing (Pisoni, 2001). This work illustrates the importance of developing language specific norms for serial span tasks such as the digit span task. In addition, these results show that caution may be needed in interpreting results from studies using serial spans as predictors of cognitive abilities. For instance, it has been recently proposed that backward span measures might be a reliable predictor of attentional and cognitive abilities (Hale, Hoeppner, & Fiorello, 2002; Ramsay & Reynolds, 1995). Similarly, span tasks (especially the digit span task) are often part of batteries used to pinpoint mild cognitive impairment and to diagnose dementia in the elderly (see, for example, Karlawish & Clark, 2003). The finding of significantly different serial spans within the same individual as a function of language tested (as shown in Experiment 2) clearly illustrates that serial span tasks do not provide valid measures of STM processing skills across spoken and signed languages.

Acknowledgements

We wish to thank the students and the staff of the National Technical Institute for the Deaf, Rochester, NY, and of Gallaudet University, Washington, DC, as well as the participants and organizing committee of CoDAWay 2002. We are grateful to P. Clark, Dr. P. Hauser, N. Fernandez, O. Pouliot, J. Cohen, D. Tuladhar and N. Schaeffer for their help with stimuli and subject recruitment. This research was supported by the National Institutes of Health (DC04418) and the James S. McDonnell Foundation to D.B., and by the National Institutes of Health (DC00167) to E.N. and T.S.

References

Addis, K. M., & Kahana, M. J. (2004). Decomposing serial learning: What is missing from the learning curve? *Psychonomic Bulletin & Review, 11*(1), 118–124.

Altom, M. W., & Weill, J. (1977). Young children's use of temporal and spatial order information in short-term memory. *Journal of Experimental Child Psychology, 24*, 147–163.

Asch, S. E., & Ebenholtz, S. M. (1962). The process of free recall: Evidence for non-associative factors in acquisition and retention. *The Journal of Psychology, 54*, 3–31.

Baddeley, A. (2003). Working memory and language: An overview. *J. Commun. Disord., 36*, 189–208.

Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105*, 158–173.

Baddeley, A., & Hitch, G. J. (1974). Working Memory. In G. Bower (Ed.). *Recent advances in learning and motivation* (Vol. VIII). New York: Academic Press.

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

Baddeley, A. D., 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., Newport, E. L., Hall, M. L., Supalla, T., & Boutla, M. (2006). Persistent difference in short-term memory span between sign and speech: Implications for cross-linguistic comparisons. *Psychological Science, 17*(12), 1090–1092.

Beaman, C. P., & Jones, D. (1998). Irrelevant sound disrupts order information in free recall as in serial recall. *Quarterly Journal of Experimental Psychology A, 51*, 615–636.

Beck, K., Beck, G., & Gironella, O. (1977). Rehearsal and recall strategies of deaf and hearing individuals. *American Annals of the Deaf, 122*(26), 544–552.

Bellugi, U., Klima, E., & Siple, P. (1974-1975). Remembering in signs. *Cognition, 3*(2), 93–125.

Bellugi, U., & Siple, P. (1974). Remembering with and without words. In F. Bresson (Ed.), *Current problems in psycholinguistics* (pp. 215–236). Paris: Centre National de la Recherche Scientifique.

Bonvillain, J. D., AlthausRea, C., Orlansky, M. D., & Slade, L. A. (1987). The effect of sign language rehearsal on deaf's subjects immediate and delayed recall of English word lists. *Applied Psycholinguistics, 8*, 33–54.

Boutla, M. (2003). *Cognitive Mechanisms Underlying Capacity Limits in Working Memory: Insights from American Sign Language*. Rochester, NY: University of Rochester.

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

Capirci, O., Cattani, A., Rossini, P., & Volterra, V. (1998). Teaching sign language to hearing children as a possible factor in cognitive enhancement. *Journal of Deaf Studies and Deaf Education, 3*, 135–142.

Carey, P., & Blake, J. (1974). Visual short-term memory in the hearing and the deaf. *Canadian Journal of Psychology, 28*, 1–14.

Case, R., Kurland, D. M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology, 33*, 386–404.

Cavanaugh, J. (1972). Relation between the immediate memory span and the memory search rate. *Psychological Review, 79*(6), 525–530.

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.

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.

Daneman, M., & Carpenter, P. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behaviour, 19*, 450–466.

Das, J. P. (1983). Memory for spatial and temporal order in deaf children. *American Annals of the Deaf, 128*, 894–899.

Elliott, J. M. (1992). Forward digit span and articulation speed for Malay, English and two Chinese dialects. *Percept Mot Skills, 74*, 291–295.

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.

Franklin, D. R. J., & Mewhort, D. J. K. (2002). An analysis of immediate memory: The free-recall task. In N. J. Dimpoloulos & K. F. Li (Eds.), *High Performance Computing Systems and Applications 2000*. New York: Kluwer.

Gathercole, S. E. (1999). Cognitive approaches to the development of short-term memory. *Trends in Cognitive Science, 3*(11), 410–419.

Gathercole, S. E., & Baddeley, A. D. (1993). *Working memory and language*. Hillsdale, N.J.: Lawrence Erlbaum Associate Inc.

Haarman, H. J., Davelaar, E. J., & Usher, M. (2003). Individual differences in semantic short-term memory capacity and reading comprehension. *Journal of Memory and Language, 48*, 320–345.

Hachinski, V., ILiff, L., Zilhka, E., DuBoulay, G., McAllister, V., & Marshall, J. (1975). Cerebral blood flow in dementia. *Archives of Neurology, 32*, 632–637.

Hale, J. B., Hoeppner, J.-A. B., & Fiorello, C. A. (2002). Analyzing digit span components for assessment of attention processes. *Journal of Psychoeducational Assessment, 20*, 128–145.

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. (1990). Recall of order information by deaf signers: Phonetic coding in temporal order recall. *Memory and Cognition, 18*(6), 604–610.

Hoosain, R., & Salili, F. (1987). Language differences in pronunciation speed for numbers, digit span and mathematical ability. *Psychologia, 30*, 34–38.

Kanabus, M., Szelag, E., Rojek, E., & Poeppel, E. (2002). Temporal order judgement for auditory and visual stimuli. *Acta Neurobiologiae Experimentalis, 62*, 263–270.

Karlawish, J. H. T., & Clark, C. M. (2003). Diagnostic evaluation of elderly patients with mild memory problems. *Annals in internal medicine, 138*, 411–419.

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

Lewis, R. L., Vasishth, S., & Van Dyke, J. A. (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Sciences, 10*(10), 447–454.

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.

Mueller, S. T., Seymour, T. L., Kieras, D. E., & Meyer, D. E. (2003). Theoretical implications of articulatory duration, phonological similarity, and phonological complexity in verbal Working Memory. *Journal of Experimenal Psychology. Learning Memory and Cognition, 29*(6), 1353–1380.

O'Connor, N., & Hermelin, B. (1972). Seeing and hearing and space and time. *Perception and Psychophysics, 11*(1A), 46–48.

O'Connor, N., & Hermelin, B. (1976). Backward and forward recall by Deaf and Hearing children. *Quarterly Journal of Experimental Psychology, 28*, 83–92.

O'Connor, N., & Hermelin, B. M. (1973). The spatial or temporal organization of short-term memory. *Quarterly Journal of Experimental Psychology, 25*, 335–343.

Padden, C., & Humphries, T. (1988). *Deaf in America: voices from a culture*. Cambridge (MA): Harvard University Press.

Paivio, A., & Csapo, K. (1971). Short-term sequential memory for pictures and words. *Psychonomic Science, 24*(2), 50–51.

Parasnis, I., Samar, V. J., Bettger, J. G., & Sathe, K. (1996). Does deafness lead to enhancement of visual spatial cognition in children? *Journal of Deaf Studies and Deaf Education, 1*(2), 145–152.

Pisoni, D. B. (2001). Cognitive factors and cochlear implants: Some thoughts on perception, learning and memory in speech perception. *Ear and Hearing, 21*(1), 70–78.

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

Potter, M. C., & Lombardi, L. (1990). Regeneration in the Short-Term Recall of Sentences. *Journal of Memory and Language, 29*, 633–653.

Ramsay, M. C., & Reynolds, C. R. (1995). Separate digit tests: A brief history, a literature review, and a reexamination of the factor structure of the test of memory and learning (TOMAL). *Neuropsychology Review, 5*(3), 151–171.

Ronnberg, J., & Nilson, L. G. (1987). The modality effect, sensory handicap and compensatory functions. *Acta Psychologica, 65*(3), 263–283.

Ronnberg, J., Rudner, M., & Ingvar, M. (2004). Neural correlates of working memory for sign language. *Cognitive Brain Research, 20*(2), 165–182.

Rosen, V. M., & Engle, R. W. (1997). Forward and backward serial recall. *Intelligence, 25*(1), 37–47.

Sakai, K., & Passingham, R. E. (2003). Prefrontal interactions reflect future task operations. *Nature Neuroscience, 6*(1), 75–81.

Shand, M. A. (1980). Short-term coding processes in congenitally deaf signers of ASL: Natural language considerations. [41.4, 41.4]..

Thompson-Schill, S. L. (2003). Neuroimaging studies of semantic memory: Inferring 'how' from 'where'. *Neuropsychologia, 41*, 280–292.

Turner & Engle (1989). Is working memory capacity-dependent? *Journal of Memory and Language, 28*, 127–154.

Warrington, E., Kinsbourne, M., & James, M. (1966). Uncertainty and transitional probability in the span of apprehension. *British Journal of Psychology, 57*(1 and 2), 7–16.

Watkins, M. J., LeCompte, D. C., Elliott, M. N., & Fish, S. B. (1992). Short-term memory for the timing of auditory and visual signals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(5), 931–937.

Watkins, O. C., & Watkins, M. J. (1980). The modality effect and echoic persistence. *Journal of Experimental Psychology: General, 109*(3), 251–278.

Waugh, N. C. (1963). Two methods for testing serial memorization. *Journal of Experimental Psychology, 65*(2), 215–216.

Wechsler, D. (1955). *Wechsler Adult Intelligence Scale (WAIS)*. New York: The Psychological Corporation.

Wilson, M. (2001). The case for sensorimotor coding in working memory. *Psychonomic Bulletin and Review, 8*(1), 44–57.

Wilson, M., Bettger, J. G., Niculae, I., & Klima, E. S. (1997). Modality of language shapes working memory: Evidence from digit span and spatial span in ASL signers. *Journal of Deaf Studies. Deaf Education, 2*(3), 150–160.

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

Wilson, M., & Emmorey, K. (2000). When does modality matter? Evidence from ASL on the nature of working memory. In K. Emmorey & H. Lane (Eds.), *The signs of language revisited. An anthology to honor Ursula Bellugi and Edward Klima* (pp. 135–142). Mahwah, NJ: Lawrence Erlbaum Associates.

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

Wilson, M., & Emmorey, K. (2006b). No difference in short-term memory span between sign and speech. *Psychological Science, 17*(12), 1093.

Wilson, M. D. (1988). The MRC psycholoinguistic database; Machine Readable Dictionary, Version 2. *Behavioral Research Methods: Instruments and Computers, 20* (6–11).