(2001) In V. Dively, M. Metzger, S. Taub, + A.M. Baer, Signed Languages: Discoveries From International Research. Washington DC: Gallaudet University Press

Functional Consequences of Modality: Spatial Coding in Working Memory for Signs

Margaret Wilson and Karen Emmorey

Traditional approaches to psychology and psycholinguistics treat mental representations and mental events as phenomena that can be divorced from the body. The mind, of course, cannot be separated from the brain, but we traditionally view the mind-brain unit as being largely divorced from the physical body that interacts with the world. Indeed, philosophically inclined theorists invite us to imagine surgically extracting the brain from the body and placing it in a vat in which it has complete life support, with the expectation that it would still be able to "think." If you are trying to solve an algebra problem, for example, it should not matter that you have two hands, that your eyes are at the front of your head (and not at the sides like a bird), or that you cannot process sonar echolocation like a bat. In short, the physical form of the body is not supposed to affect thinking.

The difficulty with this logic, however, is that it does not fit with what we know about the evolution of the brain. Neurologically simpler creatures, presumably similar to those from which we evolved, have brains that are devoted primarily to processing sensory input and producing motor output—in other words, devoted to the body's communication with the world. These neural resources are the raw material that evolution would have to work with to build a human brain.

Recently, though, interest in a perspective that is driven by precisely this point has been generated. The perspective of embodied cognition suggests that so-called central cognitive processes may, in fact, be parasitic to more peripheral mechanisms of sensory input and motor output. According to this idea, our minds make use of the fact that we are bodies that move around in the world and take information in from the world. We use the mechanisms we already have for those purposes to perform more abstract, or what appear to be more abstract, cognitive tasks. It is from this perspective that we will consider the structure of working memory.

THE CLASSIC MODEL OF WORKING MEMORY

Working memory, sometimes called short-term memory, is a set of mental resources or mechanisms that people use for retaining information temporarily—for example, to remember a phone number long enough to cross the room and dial it or to think about how various locations are arranged on a map.

We have long known that working memory exhibits effects that suggest it makes use of perceptual and motoric processes. The model of working memory developed by Baddeley and Hitch (1974; see also Baddeley 1986) states that working memory has two major domains. One is verbal and uses the auditory and vocal mechanisms of speech to remember words. The other is visuospatial and is used to remember nonlanguage materials such as visual shapes or the relative locations of items in space. In both cases, information seems to be represented in its surface form—what it sounds like, what it looks like, or how we would produce it with our own bodies.

More specifically, verbal working memory is thought to consist of a mechanism called the "phonological loop." This loop is a two-part system consisting of a buffer that stores information in phonological form and an articulatory rehearsal process that is used to load or refresh the buffer. That is, at least in hearing subjects, verbal information is encoded in terms of speech input and speech output (see Baddeley and Hitch 1994). In contrast, this type of structure does not appear to exist in the visual domain of working memory. It has been argued that inherent differences between audition and vision are responsible for the lack of parallel structure between the two domains (see Logie 1995).

This type of model, then, appears to be very much compatible with the idea of embodied cognition. A difficulty arises, however, if we acknowledge an ambiguity as to how the two domains of working memory are defined. On the one hand, the distinction appears to be between auditory and visual processing. On the other hand, given that the model was developed based entirely on data from users of spoken languages, the distinction could also be one between linguistic and nonlinguistic processing. We hypothesized that, in fact, the differences between verbal working memory and visual working memory had less to do with sensory modality, as had been assumed, and more to do with the structure of the information that had to be maintained in memory. With the appropriate kind of structure—namely, language structure—we hypothesized that a phonological loop for visual information could exist. To test this hypothesis, we began investigating working memory for signed language.

EVIDENCE FOR A PHONOLOGICAL LOOP FOR SIGN LANGUAGE

A set of classic findings are taken as evidence for the structure of working memory for speech (see Baddeley 1986). In the phonological similarity effect, memory performance is worse when people have to remember a list of similar sounding words than when they have to remember a set of diverse words. This finding indicates that the words are being coded in terms of their sound, and when the sounds are too similar, confusion occurs. In the word-length effect, words that take a long time to pronounce are harder to remember than words that are quick to pronounce. This finding suggests that mental processes related to the planning of speech output are involved. In the articulatory suppression effect, memory is disrupted by competing articulatory activity. That is, if subjects are required to do something else with their mouths, such as repeating a nonsense word, their ability to use speech planning as a memory device is disrupted and performance goes down. Finally, in the irrelevant speech effect, performance is disrupted if subjects are required to listen to speech or other structured sounds while they are trying to remember.

To demonstrate that a phonological loop could exist in the visuospatial domain, we tested for the same set of effects using sign language stimuli, with deaf signers as subjects. A sign-based similarity effect in which signs that use the same handshape are more difficult to remember than signs with diverse handshapes has, in fact, been found by a number of investigators (Hanson 1982; Klima and Bellugi 1979; Poizner, Bellugi, and Tweney 1981). We replicated this effect (Wilson and Emmorey 1997a) and found a number of other effects that indicate a phonological loop structure (see Wilson and Emmorey 1997b). When subjects were asked to remember signs that had either long path movement or short local movement, a sign length effect was observed in which the long sign was more difficult to remember than the short signs (Wilson and Emmorey 1998). We also found that when subjects were asked to perform a competing movement with their hands (repeatedly producing a nonsense sign) during stimulus presentation, memory was disrupted, constituting a manual suppression effect (Wilson and Emmorey 1997a, 1998). Finally, we found an irrelevant sign effect: When subjects were required to watch nonsense signs while trying to hold a list of signs in working memory, performance was disrupted (Wilson and Emmorey 2000). These results suggest that deaf people "sign to themselves" just as hearing people "talk to themselves" to maintain information in working memory. This pattern of data from deaf subjects is particularly striking because it has no parallel in visual working memory in hearing people.

These data seem to support the conclusion that the structure of the phonological loop develops in response to the linguistic nature of the input and is not constrained by sensory modality. This position is, in fact, compatible with the idea of embodied cognition in that working memory is making use of sensory and motor devices in both deaf and hearing people. However, the position provides a fairly weak version of embodied cognition because the sensory and motor modalities that are used appear to have no important functional consequences. Thus, having a speech-based phonological loop or a sign-based phonological loop would appear to place essentially no constraints on how information is represented, just as the choice between a chalkboard or a dry-erase whiteboard places few or no constraints on what one writes.

This conclusion seems surprising given the radically different capabilities of audition and vision for representing information. In particular, the ability of the visual modality to form rich and detailed representations of space and spatial relationships might be expected to have functional consequences for the cognitive system. Thus, we set out to ask whether sign-based working memory encodes space as opposed to being restricted to more abstract phonological representations and, if space is encoded, whether it has functional consequences for working memory.

LOCATION AS A PHONOLOGICAL PARAMETER

One of the most basic ways that signed language uses spatial locations is as a phonological parameter. Previous demonstrations of a phonological similarity effect for signed language have usually tested only the parameter of handshape. (One study reported in Klima and Bellugi [1979] examined the location parameter, but methodological considerations prevent a clear interpretation of those data.) To test for a location similarity effect, we asked subjects to remember and repeat sequences of signs that either shared a common location (the chin) or used a variety of locations (chin, base hand, chest, neutral space, etc.). Signs were matched as closely as possible for handshape and movement so that the sequences differed systematically only in terms of location (e.g., ORANGE, CAFETERIA, LIGHT, PIG versus MILK, COMMITTEE, PUMPKIN, PANTS). In the same-location condition, mean recall of signs in the correct serial position was 74 percent whereas, in the varied-location condition, mean recall was 83 percent. That is, a similarity effect for spatial location did occur in which groups of signs with diverse locations led to fewer confusions and better recall.

We must consider two points, however, regarding these data. First, whether the location similarity effect actually reflects spatial representation is not clear. Are locations such as the forehead and the chin represented in terms of their spatial relationships to one another, or are they represented more abstractly, simply as phonological parameters? This point raises the interesting question of how the phonology of signed language is mentally represented, a question that remains unanswered for the moment. However, a second and more serious concern regarding these data requires consideration, at least for answering the question of whether language modality plays a role in shaping cognition. The concern is that the representation of spatial location as a phonological parameter is unlikely to have functional consequences because it is simply part of the encoding of the sign itself. It serves merely to help identify the lexical item that is to be retrieved. In this sense, it is no different from handshape similarity or from sound similarity for speech.

However, space is used in multiple ways in ASL, giving us other ways to approach this question. Prominent among these other uses of space are the use of space for grammatical functions, for the representation of space itself, and for the representation of time or serial order. These uses of space allow information to be represented in fundamentally different ways than they are represented in spoken language. If the embodied cognition perspective is correct, we might expect this unique feature of signed languages to have functional consequences for working memory.

USING SPACE TO ENCODE SERIAL ORDER

Results from our earlier research contained hints that deaf subjects were using space to encode serial order. First, we noticed that some of our subjects spontaneously reported the to-be-remembered items in a sequence of spatial locations, usually arrayed left to right. Further, this spatial ordering appeared to be playing a functional role in memory. Some subjects used the spatial ordering as a mechanism for indexing the serial position of specific items, for example, by returning to a location to make a correction.

A further suggestion that space is used to encode serial order comes from our previous research in which subjects were asked to report a sequence of words either in the same order they were presented or in backwards order. We found that deaf children who are native signers of ASL perform equally well on backward report as on forward report (Wilson et al. 1997; for a similar result with deaf adults, see also Mayberry and Eichen 1991). This finding strikingly contrasts with the standard finding for hearing subjects, for whom backward report is a considerably more difficult task than forward report. Indeed, the deaf subjects outperformed hearing subjects on backward report despite greater forward span in the hearing subjects. This finding indicates that the equal performance in the two conditions by the deaf subjects is not a floor effect, and it does not result from a failure to retain serial order information. To be able to perform backward report at levels above that of hearing subjects, these deaf subjects must be retaining serial order information—and retaining it in some form that is amenable to the task of reversing the order. In fact, spatial ordering, which is not only physically possible in signed languages but also actually incorporated into the grammar, could provide exactly such a form of reversible serial ordering. Items that are arrayed across space, unlike items arrayed across time, do not have a necessary directionality.

To test the hypothesis that deaf subjects can use space as a tool for maintaining serial order information, we tested subjects under conditions that were designed to encourage or discourage spatial encoding. First, we encouraged the use of spatial encoding by actually presenting the signs in various locations. We divided the video screen into four quadrants, and sequences of signs to be remembered were shown in a predictable sequence of locations: upper left, upper right, lower left, lower right. We compared this condition to one in which all the signs were presented in the same location. Second, we discouraged the use of spatial encoding by presenting signs that use a fixed location on the body (e.g., LEMON, METAL). The location of these signs cannot be varied without changing the meaning of the sign. These were compared to signs that take place in neutral space (e.g., LIBRARY, TEXAS). Unlike the fixed-location signs, these signs can be performed in various spatial locations, allowing the subject to mentally rehearse the signs in a spatial sequence. Both variables were manipulated in the same experiment, in a 2 × 2 design. Our hypothesis was that, if showing signs in varied locations induces a spatial rehearsal strategy, then we should see better performance in the varied-location condition than in the same-location condition, but only when the signs themselves are capable of being varied spatially. Similarly, we should see better performance for the neutral-space signs that can be rehearsed spatially than for the fixed-location signs, at least in the varied-location condition.

The results showed that neutral-space signs were, in fact, easier to remember than fixed-location signs, just as we would expect if spatial encoding boosts memory for serial order (see table 1). Interestingly, however, the presentation of spatially varied signs had no effect, and no interaction occurred between the two variables.

Table 1. Deaf and hearing subjects' ability to remember neutral-space vs. fixed-location signs

	Mean Percentage Correct for Deaf Subjects		Mean Percentage Correct for Hearing Subjects	
	Same location	Varied location	Same location	Varied location
Fixed Signs	71%	72%	47%	39%
Neutral Signs	77%	76%	48%	47%

One possible explanation for why varying location had no effect is that we did not vary the locations of the signs within signing space. Instead, we showed the signer's whole body at various spatial locations. It may be that if we were to show the signer taking up the whole video screen but performing the signs in a sequence of locations within signing space, then spatial representations within the linguistic system of ASL would be activated and performance would improve. Another possibility, however, is that deaf subjects may already be spontaneously using a spatial strategy. The attempt to induce such a strategy by presenting the stimuli differently may have been ineffective precisely because the strategy was already being used. This second possibility is strongly suggested by the fact that we did find an effect of neutral- versus fixed-location signs, even when varied spatial presentation was not used. This finding suggests that subjects are using a spatial strategy when the signs allow it. This explanation is also compatible with our earlier results, which showed evidence of spontaneous spatial encoding without any spatial variation in the stimuli.

IS A SPATIAL STRATEGY UNIQUE TO SIGNED LANGUAGE?

One question we must ask is whether spatial rehearsal is unique to signed language or whether it is a strategy that users of a spoken language might also be able to use. The idea that anybody could benefit by mentally associating each item to be remembered with a location in space has, at least, surface plausibility. The "method of loci" is a famous strategy for storing lists in long-term memory by associating each item on the list with a familiar location. Perhaps, working memory also can benefit from such a strategy. If this possibility were the case, then hearing subjects, too, might show evidence of spatial encoding if they were encouraged to use such a strategy. But in fact, recent work by Li and Lewandowski (1993, 1995) and by Serra and Jonas (1996) shows that associating words with spatial locations does not help hearing subjects in a standard working memory task. That is, using space to encode serial order appears not to be an effective strategy with speech. It could still be argued, though, that hearing subjects simply have less practice with spatial tasks and spatial encoding of materials than deaf signers do. That is, the spatial effects that have been observed in deaf subjects but not in hearing subjects may merely reflect, because of language modality, a preferred cognitive strategy rather than structural differences in working memory.

To test this hypothesis, we conducted a further experiment in which hearing subjects were taught the sixteen signs used in the previous experiment. They practiced the signs until they were able to produce them all from memory, but they were not told the meanings of the signs. These hearing subjects were then tested in the same four conditions in which the deaf subjects had been tested.

Results are shown in the second half of table 1. If we take performance with the fixed-location signs as a baseline, we find that the varied-location condition is more difficult for hearing subjects than the single-location condition. This finding may reflect the fact that the stimulus is more visually complex, and unpracticed hearing subjects may have difficulty sorting out how to encode the relevant visual information. Indeed, in the absence of a well-practiced rehearsal mechanism, hearing subjects may be simply trying to match their rehearsal exactly to what they saw, location and all, even when that is impossible. However, this added difficulty of spatially varied presentation vanishes when the stimuli themselves allow spatial rehearsal (the neutral-space signs). This pattern of results contrasts with the case of remembering printed words, where spatial variation neither helped nor hurt but was simply irrelevant (Li and Lewandowski 1993, 1995; Serra and Jonas 1996).

In short, the results with deaf subjects and the results with hearing subjects converge on the same conclusion: Space is used as an encoding device in working memory when, and only when, the physical structure of the stimulus allows the body to enact movements in space. And although quite an artificial task must be created to demonstrate this point with hearing subjects, the use of spatial rehearsal by deaf signers can be observed in the task of remembering language and, therefore, has potentially widespread cognitive consequences.

Our conclusion, then, is that modality matters. The structure and functioning of working memory is dependent in part on the modality of one's language. In contrast to our original findings, which suggested that the abstract properties of language were responsible for the structure of working memory, it appears that the structure is also shaped by the physical realities of sensory and motor modalities. How our language is embodied has functional consequences for how the cognitive system uses it.

REFERENCES

Baddeley, A. 1986. *Working memory.* Oxford: Oxford University Press.

Baddeley, A., and G. Hitch. 1974. Working memory. In *Recent advances in learning and motivation,* Vol. 8, ed. G. Bower, 647–67. Hillsdale, N.J.: Lawrence Erlbaum.

———. 1994. Developments in the concept of working memory. *Neuropsychology* 8(4):485–93.

Hanson, V. L. 1982. Short-term recall by deaf signers: Phonetic coding in temporal order recall. *Memory and Cognition* 18:604–10.

Klima, E. S., and U. Bellugi. 1979. *The signs of language.* Cambridge, Mass.: Harvard University Press.