sisted of synthetic peptide conjugated to bovine thyroglobulin with the use of glutaraldehyde. The antiserum recognized a protein band of 80 to 90 kD on protein immunoblots of membranes prepared from cells transfected with the rat SPR [S. R. Vigna *et al.*, *J. Neurosci.* **14**, 834 (1994)]. The cells could be immunostained with the antiserum, and the staining could be blocked by preabsorbing the antiserum with $SPR_{393-407}$. After the incubation with the primary antibody, the tissue sections were washed for 30 min at 22°C in PBS (pH 7.4) and then incubated in the secondary antibody solution (pH 7.4) for 2 hours at 22°C. This secondary antibody solution was identical to the primary antibody solution, except that cyanine (Cy3)–conjugated donkey anti–rabbit immunoglobulin G (no. 711-165-152; Jackson ImmunoResearch Labs, West Grove, PA) was present (at a concentration of 1:600) instead of the SPR antibody. Finally, the tissue sections were washed for 20 min in PBS (pH 7.4, 22°C), mounted on gelatin-coated slides, and coverslipped with PBS-glycerine containing 1.0% *p*-phenylenediamine to reduce photobleaching.

12. To examine the sites of internalization within the cell, sections were examined with an MRC-600 Confocal Imaging System (Bio-Rad, Boston, MA) and with an Olympus BH-2 microscope equipped for epifluorescence (Olympus America, Lake Success, NY). Both of the microscopes were set up as previously described [T. C. Brelje, D. W. Scharp, R. L. Sorenson, *Diabetes* **38**, 808 (1989); P. W. Mantyh *et al.*, *J. Neurosci.* **15**, 152 (1995)]. To determine the percentage of neurons showing intense internalization, sections were examined with a Leitz Orthoplan II microscope (Ernst Leitz Wetzler, Germany) equipped for fluorescence. To determine the total number of cell bodies showing SPR-positive internalization, immunostained sections were viewed through a 1.0-cm^2 eyepiece grid that was divided into 100 1-mm by 1-mm units, and the total number of SPR-positive cell bodies and the number of SPR-positive cell bodies showing significant SPR-positive internalization were counted. SP-induced internalization has been shown to be blocked by nonpeptide SPR antagonists in transfected cells [A. M. Garland, E. F. Grady, D. G. Payan, S. R. Vigna, N. W. Bunnett, *Biochem. J.* **303**, 177 (1994)] and in cultured spinal neurons (*19*).
13. To determine whether the capsaicin-induced SPR internalization in lamina I neurons was dose dependent, 0.1, 1.0, 100, and 300 μg of capsaicin was injected into one hindpaw of rats. The percentage of neurons showing significant internalization (>20 endosomes per cell body) was examined 1 min later. The percentage of L4 lamina I neurons showing significant internalization was 26, 54, 55, and 70% for 0.1, 1.0, 100, and 300 μg of capsaicin, respectively.
14. We have consistently observed an increase in total apparent SPR immunoreactivity per cell or per field after either exogenous injection of SP in the spinal cord (*19*) or after increased release of endogenous SP as observed here. Although this observation may reflect a change in the accessibility of the antiserum to the SPR or a change in SPR immunoreactivity related to possible changes in receptor conformation after agonist binding, both of these possibilities seem unlikely. The most plausible explanation is that the increase in total SPR-positive immunoreactivity observed after SP binding simply reflects the concentration of SPRs into small membrane-bound endosomes, which are present in confocal slices in greater amounts than is plasma membrane, in which the SPR is located before peptide binding. This interpretation is supported by the observation that an early step in agonist-induced endocytosis is the migration of receptors in the plane of the plasma membrane to pits (*2*), thereby effectively concentrating the receptors in the membranes destined to become endosomal.
15. J. Jack, in *The Neurosciences: Fourth Study Program*, F. O. Schmitt and F. G. Worden, Eds. (MIT Press, Cambridge, MA, 1979), pp. 423–437; B. Katz, *Nerve, Muscle and Synapse* (McGraw-Hill, New York, 1966), pp. 12–51.
16. M. N. Castel, J. Woulfe, X. Wang, P. M. Laduron, A. Beaudet, *Neuroscience* **50**, 269 (1992).
17. C. J. Woolf, *Nature* **306**, 686 (1983); R. Dubner, in *Proceedings of VI World Congress on Pain, Pain Research and Clinical Management*, M. Bond, E. Charlton and C. J. Woolf, Eds. (Elsevier, Amsterdam, 1991), vol. 5, pp. 263–270; D. A. Simone *et al.*, *J. Neurophysiol.* **66**, 228 (1991).
18. T. Mitchison and M. Kirschner, *Neuron* **1**, 761 (1988); G. J. Brewer and C. W. Cotman, *Neurosci. Lett.* **99**, 268 (1989); D. Bigot and S. P. Hunt, *ibid.* **111**, 275 (1990); *ibid.* **131**, 21 (1991); D. Bigot, A. Matus, S. P. Hunt, *Eur. J. Neurosci.* **3**, 551 (1991).
19. P. W. Mantyh, unpublished observations.
20. We thank A. Georgopoulos for helpful comments, R. Elde and the Biomedical Imaging Processing Laboratory for use of the confocal microscope, and C. Garret for a gift of RP-67580. Supported by NIH grants NS23970, NS14627, NS21445, DA08973, NS31223, and GM15904 and by a Veterans Administration Merit Review.

30 September 1994; accepted 2 March 1995

Integration of Visual and Linguistic Information in Spoken Language Comprehension

Michael K. Tanenhaus,* Michael J. Spivey-Knowlton, Kathleen M. Eberhard, Julie C. Sedivy

Psycholinguists have commonly assumed that as a spoken linguistic message unfolds over time, it is initially structured by a syntactic processing module that is encapsulated from information provided by other perceptual and cognitive systems. To test the effects of relevant visual context on the rapid mental processes that accompany spoken language comprehension, eye movements were recorded with a head-mounted eye-tracking system while subjects followed instructions to manipulate real objects. Visual context influenced spoken word recognition and mediated syntactic processing, even during the earliest moments of language processing.

The two essential properties of language are that it refers to things in the world and that its grammatical structure can be characterized independently of meaning or reference (*1*). The autonomy of grammatical structure has led to a long tradition in psycholinguistics according to which it is assumed that the brain mechanisms responsible for the rapid syntactic structuring of continuous linguistic input are "encapsulated" from other cognitive and perceptual systems (*2*), much as early visual processing often is claimed to be structured by autonomous processing modules (*3*). This contrasts with a second tradition by which language processing is inextricably tied to reference and relevant behavioral context (*4*). The primary empirical evidence that syntactic processing is modular is that brief syntactic ambiguities, which arise because language unfolds over time, appear to be initially resolved independently of prior context. Unfortunately, it has been impossible to perform the crucial test to determine whether strongly constraining nonlinguistic information can influence the earliest moments of syntactic processing, because experimental techniques that provide fine-grained temporal information about spoken language comprehension could not be used in natural contexts. However, by recording eye movements (*5*) as participants followed instructions to move objects (for example, "Put the apple that's on the towel in the box"), we were able to monitor the ongoing comprehension process on a millisecond time scale. This enabled us to observe the rapid mental processes that accompany spoken language comprehension in natural behavioral contexts in which the language had clear real-world referents.

Our initial experiments demonstrated that individuals processed the instructions incrementally, making saccadic eye movements to objects immediately after hearing relevant words in the instruction. Thus the eye movements provided insight into the mental processes that accompany language comprehension. For example, when asked to touch one of four blocks that differed in marking, color, or shape, with instructions such as "Touch the starred yellow square," a person made an eye movement to the target block an average of 250 ms after the end of the word that uniquely specified the target with respect to the visual alternatives (for example, after "starred" if only one of the blocks was starred, and after "square" if there were two starred yellow blocks). With more complex instructions, individuals made informative sequences of eye movements that were closely time-locked to words in the instruction that were relevant to establishing reference. In one experiment, subjects were given a complex instruction such as "Put the five of hearts that is below the eight of clubs above the three of diamonds," with a display composed of seven miniature playing cards, including two fives of hearts. As the person heard "the five of hearts," she looked at each of the two potential referents successively. After hearing "below the," she immediately

M. K. Tanenhaus, M. J. Spivey-Knowlton, K. M. Eberhard, Department of Brain and Cognitive Sciences, University of Rochester, Rochester, NY 14627, USA.
J. C. Sedivy, Department of Linguistics, University of Rochester, Rochester, NY 14627, USA.

*To whom correspondence should be addressed.

looked at a ten of clubs, which was above the five of hearts on which she had been fixating. By the time she heard the end of the word "clubs," her eyes moved to interrogate the card above the other five of hearts, which was the eight of clubs; thus she identified that five as the target. The eye immediately shifted to the target card and remained there until the hand reached for it.

We also found that the visual context affected the resolution of temporary ambiguities within individual words. For example, halfway through the spoken word "candy," the auditory input is consistent with both "candy" and "candle." Subjects were presented with a display of everyday objects that sometimes included two objects with initially similar names (for example, candy and candle) and were instructed to move them around ("Pick up the candy. Now put it above the fork"). The mean time to initiate an eye movement to the correct object (candy) was 145 ms from the end of the word when there was not another object with a similar name, compared with 230 ms when an object with a similar name was included in the display (6). Because it takes about 200 ms to program a saccadic eye movement (7), these results demonstrate that the individual identified the object before hearing the end of the word, when none of the other objects had a similar name.

The compelling evidence for the rapid and nearly seamless integration of visual and linguistic information that emerged from these experiments led us to test whether information provided by the visual context would affect the syntactic processing of an instruction. The strongest evidence for the modularity of syntactic processing has come from studies of sentences with brief syntactic ambiguities, in which readers have clear preferences for particular interpretations that persist momentarily even when preceding linguistic context supports the alternative interpretation (8). However, under these conditions, the context may not be immediately accessible, because it has to be represented in memory. We reasoned that a relevant visual context that was available for the listener to interrogate as the linguistic input unfolded might influence its initial syntactic analysis. If so, this would provide definitive evidence against syntactic modularity.

We used instructions containing the temporary syntactic ambiguity with perhaps the strongest syntactic preference in English, as illustrated by the examples, "Put the apple on the towel in the box," and "Put the apple that's on the towel in the box."

In the first sentence, the first prepositional phrase, "on the towel," is ambiguous as to whether it modifies the noun phrase ("the apple"), thus specifying the location of the object to be picked up, or whether it denotes the destination, that is, the location where the apple is to be put. Those who experiment with this type of ambiguity consistently find that readers and listeners interpret the first prepositional phrase as specifying the destination, which results in momentary confusion when they encounter the second preposition ("in") (9). Modular models attribute this preference to syntactic simplicity and to the syntactic requirements of the verb "put" (10). In the second sentence, the word "that's" disambiguates the phrase as a modifier and serves as an unambiguous control condition.

Six people who had not performed these tasks before were presented with three instances of each of the four conditions created by pairing the two types of instructions (ambiguous and unambiguous), as illustrated in the example above, with a one-referent visual context that supported the destination interpretation and a two-referent context that supported the modification interpretation (11). In the one-referent context for that example, the workspace contained a towel with an apple on it, another towel without an apple, a box, and a pencil. Upon hearing the phrase "the apple," individuals can immediately identify the object to be moved because there is only one apple, and thus they are likely to assume that "on the towel" is specifying the destination. In the two-referent context, the pencil was replaced by a second apple that was on a napkin. Thus "the apple" could refer to either of the two apples, and the phrase "on the towel" provides modifying information that specifies which apple is the correct referent (12). However, if initial syntactic processing is encapsulated, as modular theories claim, then people should still initially interpret "on the towel" as the destination.

In fact, strikingly different fixation patterns between the two visual contexts revealed that the ambiguous phrase "on the towel" was initially interpreted as a destination in the one-referent context, but as a modifier in the two-referent context. In the one-referent context with the ambiguous instruction, participants first looked at the target object (the apple) 500 ms after hearing "apple," but then they looked at the incorrect destination (the irrelevant towel) 55% of the time, shortly after hearing "towel"; this indicated that they had initially interpreted "on the towel" as specifying the destination. The participants then looked back

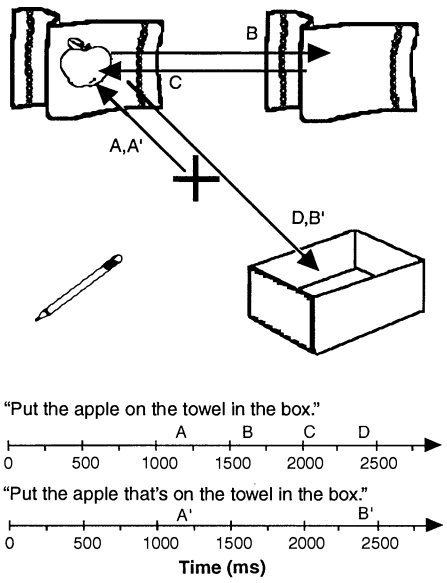

Fig. 1. Typical sequence of eye movements in the one-referent context for the ambiguous and unambiguous instructions. Letters show when in the instruction each eye movement occurred, as determined by the mean latency for that type of eye movement (A′ and B′ correspond to the unambiguous instruction).

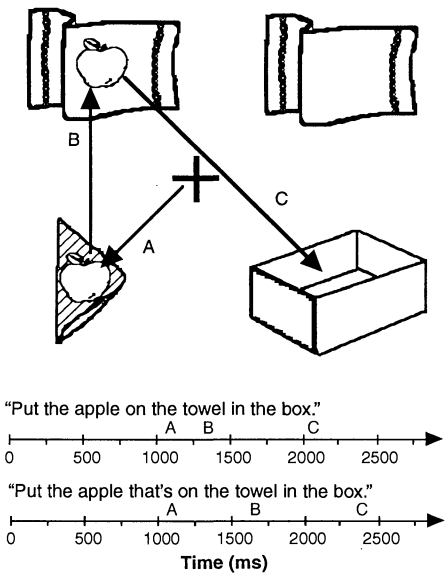

Fig. 2. Typical sequence of eye movements in the two-referent context. Note that the sequence and the timing of eye movements, relative to the nouns in the speech stream, did not differ for the ambiguous and unambiguous instructions.

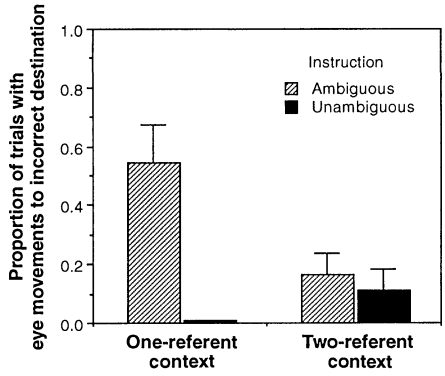

Fig. 3. Proportion of trials in which participants looked at the incorrect destination.

at the apple to pick it up, and finally at the box for placement. When the unambiguous instruction was presented in the one-referent context, participants never looked at the incorrect destination (*13*) (Fig. 1).

In the two-referent context, participants often looked at both apples shortly after hearing "the apple," which reflected the fact that reference could not be established on the basis of just that input. Participants looked at the incorrect referent during 42% of the unambiguous trials and during 61% of the ambiguous trials. [In contrast, in the one-referent context, in which reference could be established given just "the apple," individuals rarely looked at the incorrect object (pencil); this occurred during 0 and 6% of the trials for the ambiguous and unambiguous instructions, respectively.] The time it took participants to establish reference correctly in the two-referent context did not differ for the ambiguous and unambiguous instructions, which indicates that "on the towel" was immediately interpreted as a modifier, not as a destination. Individuals then typically looked directly to the box for object placement without looking at the incorrect destination (Fig. 2). In contrast with the one-referent context, ambiguity in the instruction did not affect the proportion of eye movements to the incorrect destination in the two-referent context (*14*) (Fig. 3).

Our results demonstrate that in natural contexts, people seek to establish reference with respect to their behavioral goals during the earliest moments of linguistic processing. Moreover, referentially relevant nonlinguistic information immediately affects the manner in which the linguistic input is initially structured. Given these results, approaches to language comprehension that assign a central role to encapsulated linguistic subsystems are unlikely to prove fruitful. More promising are theories by which grammatical constraints are integrated into processing systems that coordinate linguistic and nonlinguistic information as the linguistic input is processed (*10, 15*). Finally, our results show that with well-defined tasks, eye movements can be used to observe under natural conditions the rapid mental processes that underlie spoken language comprehension. This paradigm can be extended to explore questions on topics ranging from recognition of spoken words to conversational interactions during cooperative problem solving.

REFERENCES AND NOTES

1. N. Chomsky, *Aspects of the Theory of Syntax* (MIT Press, Cambridge, MA, 1965); S. Pinker, *Science* **253**, 530 (1991); *The Language Instinct* (Morrow, New York, 1994).
2. J. A. Fodor, *Modularity of Mind* (MIT Press, Cambridge, MA, 1983).
3. Early stages of visual information processing appear to segregate different features of visual input, such as form, color, motion, and depth, both anatomically and functionally, presumably to increase speed and efficiency in early computation [M. Livingstone and D. Hubel, *Science* **240**, 740 (1988)].
4. H. Clark, *Arenas of Language Use* (Univ. of Chicago Press, Chicago, 1994); W. D. Marslen-Wilson, *Nature* **244**, 522 (1973); *Science* **189**, 226 (1975).
5. We monitored eye movements with an Applied Scientific Laboratories camera that was mounted on a lightweight helmet. The camera provides an infrared image of the eye at 60 Hz. The center of the pupil and the corneal reflection are tracked to determine the orbit of the eye relative to the head. Accuracy is better than 1 degree of arc, with virtually unrestricted head and body movements. For details, see D. Ballard, M. Hayhoe, J. Pelz, *J. Cog. Neurosci.* **7**, 66 (1995). Instructions were spoken into a microphone connected to a Hi-8 VCR that also recorded the field of view and eye position of the participant.
6. Eight objects were on a table with a center fixation cross. Each trial began with the instruction, "Look at the cross." The eye-movement latency difference between the conditions with and without objects with similar names was reliable [$t(7) = 3.04$, $P < 0.02$].
7. E. Matin, K. Shao, K. Boff, *Percept. Psychophys.* **53**, 372 (1993).
8. For review, see L. Frazier, in *Attention & Performance XII*, M. Coltheart, Ed. (Lawrence Erlbaum, Hove, UK, 1987), pp. 559–586.
9. F. Ferreira and C. Clifton, *J. Mem. Lang.* **25**, 348 (1986); M. Britt, *ibid.* **33**, 251 (1994).
10. For review, see M. Spivey-Knowlton and J. Sedivy, *Cognition*, in press.
11. The 12 critical instructions were embedded among 90 filler instructions. Each trial began with the command, "Look at the cross."
12. S. Crain and M. Steedman [in *Natural Language Parsing*, D. Dowty, L. Kartunnen, H. Zwicky, Eds. (Cambridge Univ. Press, Cambridge, 1985), pp. 320–358] and G. Altmann and M. Steedman [*Cognition* **30**, 191 (1988)] have developed a theory of syntactic ambiguity resolution in which referential context is central.
13. This difference between ambiguous and unambiguous instructions was reliable by a planned comparison [$t(5) = 4.11$, $P < 0.01$].
14. The interaction between context and ambiguity for eye movements to the incorrect destination was reliable [$F(1,5) = 8.24$, $P < 0.05$]. Also, a three-way interaction between context, ambiguity, and type of incorrect eye movement (to object or to destination) revealed the bias toward a destination interpretation in the one-referent context and toward a modification interpretation in the two-referent context [$F(1,5) = 18.41$, $P < 0.01$].
15. J. McClelland, in *Attention and Performance XII*, M. Coltheart, Ed. (Lawrence Erlbaum, Hove, UK, 1987), pp. 3–36; R. Jackendoff, *Languages of the Mind* (Bradford, Cambridge, MA, 1992); C. Pollard and I. Sag, *Head-Driven Phrase Structure Grammar* (Univ. of Chicago Press, Chicago, 1993); M. MacDonald, N. Pearlmutter, M. Seidenberg, *Psychol. Rev.* **101**, 676 (1994); M. Tanenhaus and J. Trueswell, in *Handbook of Cognition and Perception*, J. Miller and P. Eimas, Eds. (Academic Press, San Diego, CA, in press).

16. We thank D. Ballard and M. Hayhoe for encouraging us to use their laboratory (National Resource Laboratory for the Study of Brain and Behavior) and for advice on the manuscript, P. Lennie and R. Jacobs for helpful comments, J. Pelz for teaching us how to use the equipment, and K. Kobashi for assistance in data collection. Supported by NIH resource grant 1-P41-RR09283, NIH HD27206 (M.K.T.), an NSF graduate fellowship (M.J.S.-K.), and a Social Sciences and Humanities Research Council of Canada fellowship (J.C.S.). All participants gave informed consent.

9 January 1995; accepted 4 April 1995

TECHNICAL COMMENTS

Origins of Fullerenes in Rocks

Naturally occurring fullerenes have been found in rock samples that were subject to singular geologic events such as lightning strokes (*1*), wildfires at the K-T boundary (*2*), and meteoritic impacts (*3*). These findings are expected, as fullerenes form normally under highly energetic conditions. However, P. R. Buseck *et al.* (*4*) reported the presence of C_{60} in a carbon-rich rock sample from Shunga, in Karelia, Russia, in which the host geologic unit was highly metamorphosed and there was no evidence of exposure to extreme conditions. If fullerenes did form naturally in such an environment, we would expect them to be widely present elsewhere, and there would be many ramifications. For example, the presence of fullerenes in the earliest times would have implications for the evolution of life (that is, as an early source of large molecules).

We studied the occurrence and distribution of fullerenes in carbon-rich rocks, including samples of shungite from the deposit in Shunga. To avoid sources of contamination by fullerenes, our samples were prepared in laboratories where there had been no previous work done on fullerenes. The outer 2- to 4-mm portion of the shungite samples was removed, and only the core material was gently crushed and ground before mass spectrometry (MS) analysis was carried out directly on the rock powder. Laser Fourier-transform MS and thermal desorption negative ion MS methods were used. In the thermal desorption MS, the temperature was scanned up to 450°C, at which C_{60} and C_{70} are fully volatilized. One sample was purposely contaminated with 100 ppm of commercial fullerenes as a control and to check the sensitivity of the analysis. The result of this reference test indicated that we could detect fullerenes at 10 ppm, or less, without difficulty.

The three samples from the Shunga locality (*5*) had a variable carbon content of about 100, 90, and 10% by weight. These samples were hosted by about 2-billion-year-old metamorphosed volcanic and sedimentary rocks of the Karelian terrain, which extends northwest through Finland and into Finnmark (northern Norway). We also analyzed one carbon-rich sample from the Bidjovagge mine near Kautokeino, Finnmark, from rocks with broadly similar age, provenance, and metamorphic history as those of Shunga.