# Research Article

## SYNCHRONIZING VISUAL AND LANGUAGE PROCESSING: An Effect of Object Name Length on Eye Movements

Gregory J. Zelinsky[1,2] and Gregory L. Murphy[2]

*[1]State University of New York at Stony Brook and [2]Beckman Institute, University of Illinois at Urbana-Champaign*

**Abstract**—*Are visual and verbal processing systems functionally independent? Two experiments (one using line drawings of common objects, the other using faces) explored the relationship between the number of syllables in an object's name (one or three) and the visual inspection of that object. The tasks were short-term recognition and visual search. Results indicated more fixations and longer gaze durations on objects having three-syllable names when the task encouraged a verbal encoding of the objects (i.e., recognition). No effects of syllable length on eye movements were found when implicit naming demands were minimal (i.e., visual search). These findings suggest that implicitly naming a pictorial object constrains the oculomotor inspection of that object, and that the visual and verbal encoding of an object are synchronized so that the faster process must wait for the slower to be completed before gaze shifts to another object. Both findings imply a tight coupling between visual and linguistic processing, and highlight the utility of an oculomotor methodology to understand this coupling.*

People have at their disposal two mechanisms helping them to remember objects in scenes over brief intervals: a verbal mechanism by which they semantically label objects, then subvocally rehearse their names in a continuous articulatory loop (Baddeley, 1986; Baddeley, Lewis, & Vallar, 1984; Baddeley, Thomson, & Buchanan, 1975; Zhang & Simon, 1985), and a "scratchpad" onto which they encode the visual properties of these objects and their spatial interrelationships (Baddeley, 1986; Brooks, 1967; Hatano & Osawa, 1983; Logie, 1986). Given that these two working memory mechanisms have very different representational formats and time courses, how are verbal and visual encoding synchronized to particular objects in a scene? One possibility is that they are not. If these processes are modular and lack coordination, observers may independently represent the visual features of one object while encoding the name of another object into an articulatory buffer. The second possibility, that these two mechanisms are synchronized to a given object, raises the problem of coordinating behaviors having different time courses. What happens if the visual encoding of an object finishes before the verbal representation? If visual and verbal encoding mechanisms are synchronized, the time spent by an observer looking at a particular object may therefore depend on the name attached to that object. If encoding is unsynchronized, no oculomotor dependencies on linguistic structure would be expected.

In this article, we report evidence for synchrony, and in so doing demonstrate an important linguistic constraint on visual behavior in a free-viewing memory task. The specific linguistic variable considered in this study is word length. Earlier work has shown that naming times to words (Eriksen, Pollock, & Montague, 1970; Spoehr & Smith, 1973) or pictures (Klapp, Anderson, & Berrian, 1973) increases with the number of syllables—even if a word is generated only implicitly and not actually produced. There is also a large literature on reading demonstrating a relationship between a word's length and the gaze duration on that word (see Rayner, 1998, for a review). We extend these findings, showing that the number of syllables in a pictorial object's name can affect the time spent looking at that object, even when the task does not explicitly require naming.

Although several earlier studies have described relationships between oculomotor variables and either the visual properties of pictorial objects and scenes (Groner, Walder, & Groner, 1984; Kowler, 1990; Mackworth & Morandi, 1967; Yarbus, 1967) or the linguistic properties of words and sentence structure (Balota, Pollatsek, & Rayner, 1985; Dobkins, Morris, & Rayner, 1992; Ferreira & Henderson, 1990; Just & Carpenter, 1980), only a handful of studies have bridged these two domains by investigating the constraints imposed by language on the viewing of pictorial objects. Tanenhaus, Spivey-Knowlton, Eberhard, and Sedivy (1995) monitored subjects' eye movements as they heard an instruction to manipulate real-world objects. They found that the linguistic referents to objects from a spoken instruction influenced the moment-by-moment oculomotor inspection of the actual objects (see also Cooper, 1974, and Keysar, Barr, & Horton, 1998). More recently, Meyer, Sleiderink, and Levelt (1998) found that subjects looked longer at objects with low-frequency names than those with high-frequency names in a task requiring the naming of two depicted objects. This relationship between an object's name and its viewing time, although clearly implying synchrony, may be specific to the explicit verbal response required in this study. If subjects are not instructed to produce the name of an object, then its name may not be retrieved before gaze shifts away.

In a study similar to our own, Noizet and Pynte (1976; see also Pynte, 1974) investigated whether this relationship between object naming and viewing behavior is obligatory. They did this by instructing subjects to serially shift their gaze to a set of three objects and to silently identify each in turn. The objects had either four- and five-syllable names (in French) or one-syllable names (e.g., *hélicoptère* vs. *main*). No verbal response was required, and subjects were assured that they would not be tested on these objects later. Despite these minimal task requirements, and the distinct absence of instructions to verbalize the object names, Noizet and Pynte found that subjects looked about 207 ms longer at the multisyllable items—an effect they attributed to a spontaneous implicit labeling of the pictorial objects (p. 219) and a relationship between labeling and oculomotor behavior.

The current study provides further evidence for visual-verbal synchronization by both establishing boundary conditions on when such synchrony might be expected and addressing problems in Noizet and Pynte's (1976) study that cloud the interpretation of this earlier work.

Although intriguing, Noizet and Pynte's (1976) study had several shortcomings. First, because their subjects were not assigned a task, it is unclear whether and to what extent subjects were identifying the objects. Furthermore, if Noizet and Pynte were correct in suggesting

Address correspondence to Greg Zelinsky, Department of Psychology, Psych B Bldg., Room 240, SUNY Stony Brook, Stony Brook, NY 11794-2500; e-mail: gzelinsky@notes.cc.sunysb.edu.

  

that objects are named spontaneously during identification, then the effect of name length should not vary as a function of task. Therefore, we examined the effect in two tasks that differed in their demands on verbal encoding. Second, Noizet and Pynte used only 7 items in each length condition (14 total), and 2 items had to be discarded because their results suggested problems with the name or picture. However, they did not establish that their other stimuli were free of such problems. In our study, we selected pictures that had very high name agreement and that were equated in a categorization pretest—thereby eliminating most variables potentially confounded with name length. Third, because one cannot dissociate a picture of a common object from its name, finding an effect of name always leaves open the possibility that some uncontrolled visual or conceptual variable was actually responsible for the result. To address this concern, we sought evidence for visual-verbal synchronization in both common objects (Experiment 1) and arbitrarily paired face and surname stimuli (Experiment 2). By comparing different encoding conditions, and controlling potential confounds, these experiments provide important new information about the relationship between linguistic structure and visual behavior.

## EXPERIMENT 1

This experiment assessed the effect of name length on the viewing of common pictures during the study period of a short-term memory recognition task. The pictures were divided into two groups: objects having one-syllable names (e.g., *ball*, *harp*) and objects having at least three-syllable names (e.g., *elephant*, *bicycle*). We refer to these as one- and three-syllable objects, respectively. Each study display depicted two one-syllable objects and two three-syllable objects (Fig. 1a), and the observers' task was to study these objects in preparation for a recognition test (Fig. 1c). Eye position data were collected during the study period (Fig. 1b). Verbal-visual synchronization in this task would be revealed if the oculomotor inspection of the study objects varied with the number of syllables in their names. Note that although this task did not require the use of object naming, it is nevertheless likely that subjects preparing for a short-term memory test would have used a verbal encoding strategy (Conrad, 1964).

### Method

#### *Subjects*

Six students from the University of Illinois at Urbana-Champaign were paid for their participation in the eye-tracking experiment. Ten different students performed the pretest. All subjects were naive with regard to the questions under investigation and had normal or corrected-to-normal visual acuity.

#### *Materials*

Twenty one-syllable objects and 20 three-syllable objects were selected from the Snodgrass and Vanderwart (1980) norms. Two selection criteria in addition to name length were followed. First, the selected items averaged 93% name agreement in both length conditions (Snodgrass & Vanderwart, 1980), eliminating one of the problems Noizet and Pynte (1976) discovered in their stimuli. Second, the items selected were equated for ease of identification using a standard categorization pretest based on the procedure followed by Murphy and Brownell (1985). In this pretest, subjects saw a category name prior to a picture and had to make a speeded judgment as to their agreement. They therefore had to identify the picture, but did not have to retrieve its category name, because this was provided. For the object groups used in this experiment, the one-syllable items were categorized in 522 ms and the three-syllable items in 523 ms—thus providing a control for visual-complexity and typicality differences that might otherwise have affected categorization times (other potential problems identified by Noizet and Pynte).

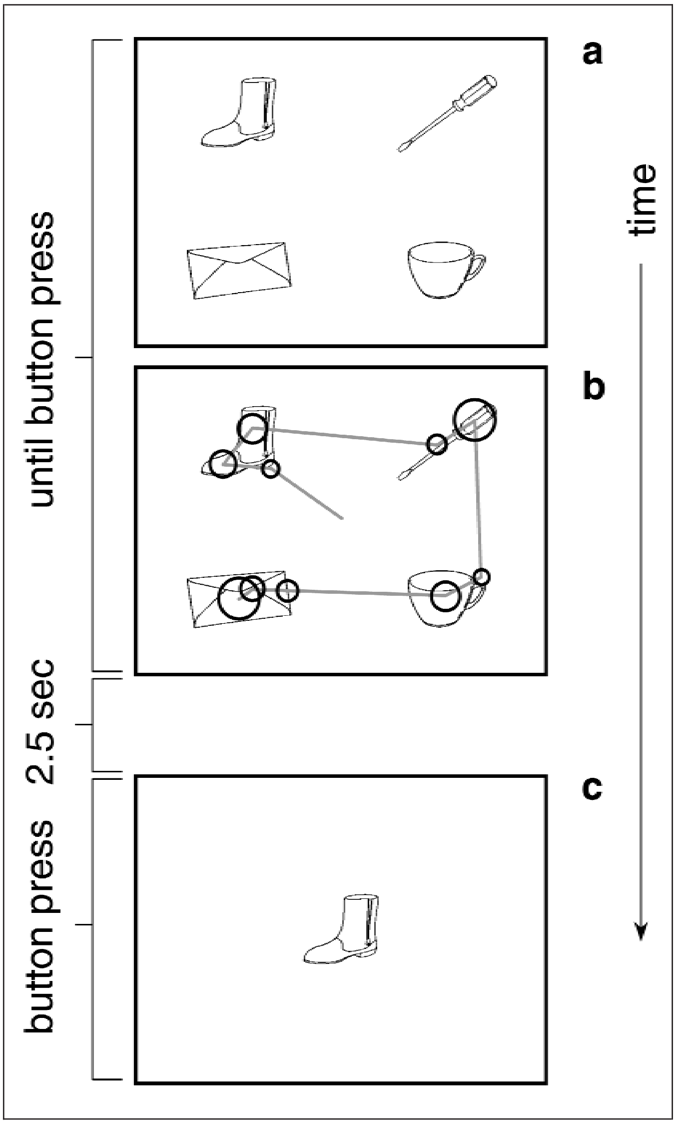


**Fig. 1.** Displays and eye movements in a representative trial in Experiment 1. A study display (a) was presented, and the observer's eye movements during this study period were recorded (b). The test display (c) appeared 2.5 s after the observer terminated the study display. In (b), eye movements are shown as gray lines, and fixations are shown as black circles. Circle diameter indicates relative fixation duration.

#### *Procedure and apparatus*

Each of the observer-initiated trials began with the presentation of four common objects (Fig. 1). The identity and location of objects in

**Table 1.** *Eye movements as a function of task and number of syllables in object name*

| | Fixation number | | Gaze duration | | Total fixation number | | Total fixation time | | Initial fixation duration | |
|---|---|---|---|---|---|---|---|---|---|---|
| Task | 1-syllable name | 3-syllable name | 1-syllable name | 3-syllable name | 1-syllable name | 3-syllable name | 1-syllable name | 3-syllable name | 1-syllable name | 3-syllable name |
| Object recognition | 2.21 | 2.54 | 567 | 700 | 4.22 | 4.78 | 1,028 | 1,235 | 252 | 246 |
| Face recognition | 2.34 | 2.66 | 833 | 1,013 | 4.27 | 4.72 | 1,512 | 1,770 | 285 | 288 |
| Face search | 1.21 | 1.18 | 225 | 224 | 2.12 | 2.25 | 450 | 547 | 194 | 196 |

*Note.* Fixation number and gaze duration refer to eye movement during the initial viewing of an object (before looking at another object); total fixation number and total fixation time include gaze shifts back to a previously viewed object.

these displays were random, with the constraints that each display depicted 2 one-syllable and 2 three-syllable objects and that each of the 40 objects appeared three times in each of the four display positions. The 6 observers freely viewed these study objects for as long as they wished, with the intention of remembering their identities; when they were done studying the objects, they pressed a button terminating the display. Following a 2.5-s blank interval, a single centrally positioned object appeared, and the observers had to indicate whether it had been presented in the previous display by pressing one of two hand-held buttons. There were 20 practice trials and 120 test trials. Trials were evenly divided into randomly interleaved target-present and target-absent conditions. No feedback was provided.

Each object in the study display subtended at most 6.9° of visual angle horizontally and 5.0° vertically, was separated from its nearest neighbor by 8° (center to center), and had an initial visual eccentricity of 5.6°. Eye position was recorded throughout the study display using a Fourward Technologies Generation 5 Dual-Purkinje-image eye-tracker sampling at 1000 Hz. A roughly 12.5°/s velocity-based algorithm was used to extract fixations off-line. A fixation was attributed to an object if it fell within the 6.9° × 5.0° bounding box enclosing that object.

## Results and Discussion

Results are reported for five dependent measures, including (a) *fixation number*, the number of initial fixations on an object before gaze shifted to another object; (b) *gaze duration*, the summed duration of all fixations during the initial viewing of an object; (c) *total fixation number*, the total number of fixations on an object, including revisitations of that object; and (d) *total fixation time*, the summed duration of all the fixations. We included both initial and total gaze measures to determine whether a syllable effect extends beyond initial viewing. We also report (e) the *initial fixation durations* on objects, although we did not expect this measure to be revealing.[1]

Evidence for an effect of number of syllables on object inspection would suggest visual-verbal synchronization during the encoding of objects into working memory. The results (Table 1, top row) indicate such a relationship. The data were analyzed by both subjects ($t_1$) and items ($t_2$). Observers made an average of 0.33 more fixations on objects having three-syllable names than on objects having one-syllable names in their initial viewing of these items, $t_1(5) = 2.54$, $p = .052$; $t_2(37) = 4.09$, $p < .001$. Consistent with these additional fixations, gaze duration on the three-syllable objects was longer by 133 ms, $t_1(5) = 2.96$, $p = .032$; $t_2(37) = 5.64$, $p < .001$. Both of these effects were more pronounced when the total viewing behavior was considered. Observers devoted an average of 0.56 more fixations to the study of the three-syllable objects than to the study of one-syllable objects, $t_1(5) = 3.04$, $p = .029$, and $t_2(37) = 4.48$, $p < .001$, resulting in a mean additional inspection of 207 ms, $t_1(5) = 3.05$, $p = .028$; $t_2(37) = 5.31$, $p < .001$. As expected, no effect of number of syllables was found for initial fixation duration, $t_1(5) = 0.80$, $p = .462$. Given that the pictures were equated for categorization difficulty in the pretest, these differences suggest an effect of object names on oculomotor inspection.

Because our argument for verbal-visual synchronization rests on the assumption that subjects were requiring more time to implicitly name the multisyllable objects, we derived a direct measure of this variable. To estimate the spoken name durations, one of us spoke each of the names in the context of a carrier sentence, which was digitized at 22.255 kHz using SoundEdit™ 16 speech analysis software. The other author then used this same software to isolate the name from the speech waveform and obtain its duration.[2] As expected, an analysis comparing the durations of the one- and three-syllable spoken names revealed a pronounced 187-ms effect of the number of syllables, $t(37) = 7.77$, $p < .001$. However, despite this difference, the number of syllables in a name is not a perfect predictor of name duration because some syllables take longer to pronounce than others. To relate individual gaze durations directly to speech, we correlated the spoken name durations of the 40 objects with the corresponding mean object-gaze durations (Fig. 2). This analysis yielded a Pearson's coefficient of .73, meaning that name duration accounted for 53% of the variability in gaze times. Separate correlations on the one- and three-syllable data yielded coefficients of .24 and .59, respec-

1. In pilot testing, we observed that many of the initial saccades to an object landed on or near the object's leading edge, and were followed by a corrective saccade bringing gaze to the object's interior. We therefore suspect that noninitial and initial object fixations may serve different functions in this task, with the former being true opportunities for object processing and the latter being a sort of stepping stone in the process of accurate object fixation.

2. It is unlikely that the author's knowledge of the hypotheses under investigation introduced a bias into the name-duration estimates. Although any rater would have expected the one-syllable names to be shorter than the three-syllable names, neither of us had available the exact gaze-duration values during the recording or segmenting of the speech patterns, thereby making the artifactual generation of the observed correlations impossible.

tively.[3] Note that the correlation within the three-syllable items is particularly important because it suggests a direct relationship between gaze duration and inner speech beyond a relationship based strictly on syllable number.

## EXPERIMENT 2

Although the effects of syllable number on gaze in the first experiment suggest a verbal encoding of objects during study, this viewing behavior might also reflect other differences between the objects or their depictions. For example, if the objects having three-syllable names also happened to be more detailed or interesting, then visual factors, not the linguistic structure of the object names, might have determined the viewing behavior. Or perhaps people just prefer to look at elephants more than they do at boots. Although the close relationship between gaze and naming duration (Fig. 2) makes such explanations implausible, they cannot be entirely ruled out in a correlational design. To eliminate such possibilities while replicating the earlier pattern of results, we conducted a similar memory experiment, but using faces as stimuli. Unlike common objects, a given face can be arbitrarily assigned a name. Arbitrary assignment meant that face-name pairings could be counterbalanced for number of syllables, thereby eliminating the possibility of visual properties confounding the linguistic manipulation.

Experiment 2 also addressed the encoding conditions needed for visual-verbal synchronization. If the synchrony observed in Experiment 1 was caused by the implicit naming of objects in the short-term memory task, then the effect would disappear if the selected task no longer involved naming. To test this possibility, Experiment 2 used both a recognition and a visual search task. Unlike a recognition task in which subjects have to encode the study items in preparation for a memory probe, in visual search the target item is presented first, followed by the search array. The encoding demands of these two tasks are therefore quite different (Palmer, 1990; Zelinsky, 1999). Subjects engaged in search need only encode the single target item into memory, then compare this item with each element appearing in the following search array. Because this comparison process need not involve naming items in the search array, we did not expect to find an effect of number of syllables in the search condition. However, if Noizet and Pynte (1976) were correct in suggesting that object naming is a fairly automatic response to a familiar object, then recognition and search should show similar evidence for an effect of number of syllables.

### Method

#### *Subjects*

Eight different University of Illinois students were paid for their participation.

3. The correlation in the three-syllable data was significantly different from 0 ($p = .007$). The correlation in the one-syllable duration data failed to reach significance ($p = .317$), partly because of the lesser variability in the data for that condition ($SD$s $= 65.7$ ms for gaze and 70.9 ms for spoken words) relative to the three-syllable data ($SD$s $= 79.3$ ms for gaze and 80.4 ms for spoken words).

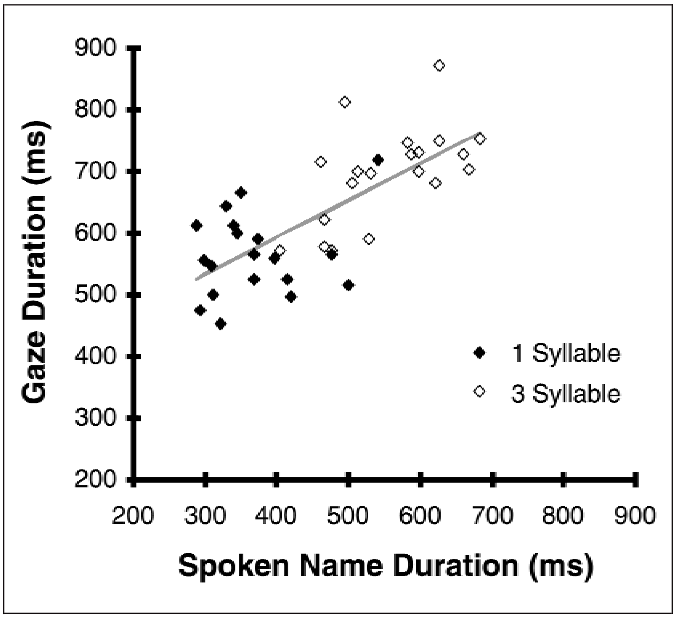


**Fig. 2.** Relationship between mean gaze duration on an object and the spoken duration of the object's name in Experiment 1. The gray line shows a linear fit to the combined data for the one- and three-syllable names.

#### *Materials*

The stimuli consisted of eight faces and eight names. The faces were randomly selected from a publicly available face database (Samaria & Harter, 1994). The eight names chosen (four one-syllable and four three-syllable names) had similar familiarity scores (all about 2.8 on a 7-point scale as rated by 10 judges). The face stimuli used in this experiment were smaller than the object stimuli used in Experiment 1 (subtending 2.2° horizontally and 2.7° vertically), although the center-to-center interobject distance and the initial visual eccentricity were the same. So that eye movements on faces and objects could be directly compared, the region of allowable face fixation was expanded to the dimension of the Experiment 1 objects, meaning that a fixation might be counted as "on" a face even if it fell slightly off the image.

#### *Procedure*

Subjects were trained in two phases to associate faces with names. Each subject first studied a static display showing the faces and their corresponding names (Fig. 3a or 3d), and then was tested by viewing each face and having to produce its name within 3 s. If the subject failed to begin producing the name within this period, the name would be displayed below the face. Training continued until the subject could correctly name all of the faces over three consecutive repetitions of the eight-face set, a process requiring about 20 min.

A recognition task and a visual search task followed the training phase. The procedure for the recognition task was identical to that described for Experiment 1 (Figs. 3b and 3c). The subject first studied a display depicting four faces (two with one-syllable names, two with three-syllable names), and then indicated whether a following probe face had appeared among this study set. The visual search task was created by reversing the presentation of the study and probe displays in each memory trial (Figs. 3e and 3f). Subjects first saw a target face and then had to indicate whether it was present in or absent from the

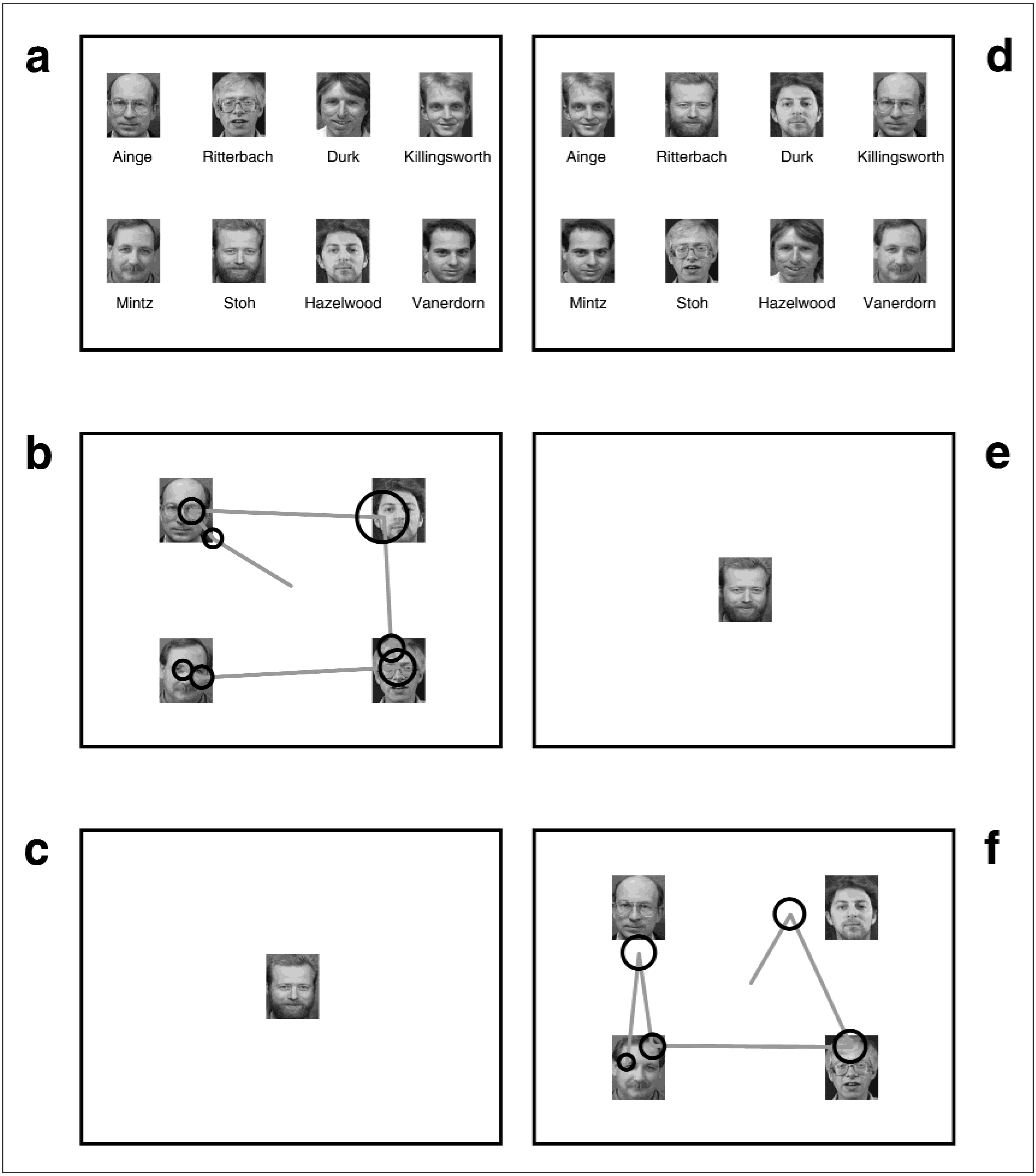


**Fig. 3.** Examples of training displays (a, d) and displays used in the recognition (b, c) and search (e, f) tasks in Experiment 2. Prior to both tasks, subjects were trained to associate names with eight faces (a, d). Except for the fact that face stimuli were used, the presentation of the study (b) and probe (c) displays in the face recognition task was unchanged from the description provided in Figure 1. The search task was also identical to the memory task except for the presentation order of the target (e) and search (f) displays. In (b) and (f), eye movements are shown as gray lines, and fixations are shown as black circles. Circle diameter indicates relative fixation duration.

four faces appearing in the following search display. Except for this different ordering of the two displays, the stimuli were identical in the two tasks. Note that the names never appeared in the study or test displays for either task, nor was naming required in order to make an accurate response. Ordering of the memory and search tasks was counterbalanced across observers, as was the assignment of one- and three-syllable name pairs to a given face. There were 80 trials per observer per task, although only the target-absent data were analyzed in the search task so as not to introduce oculomotor biases associated with a visible target (Zelinsky, 1996).

### Results and Discussion

The pattern of viewing behavior observed in Experiment 1 for common objects was replicated for face recognition (Table 1, middle row). Subjects made 0.32 more fixations during their initial viewing of faces associated with three-syllable names than during their initial viewing of faces associated with one-syllable names, $t_1(7) = 4.49$, $p = .003$, and $t_2(6) = 3.72$, $p = .010$, resulting in a 180-ms increase in mean gaze duration, $t_1(7) = 6.29$, $p < .001$; $t_2(6) = 3.39$, $p = .015$. Both effects were again more pronounced when the analyses included repeated object inspection, with 0.45 more total fixations and a 258-ms increase in total fixation time for the faces associated with three-syllable names, $t_1(7) \geq 3.58$, $p \leq .01$; $t_2(6) \geq 3.11$, $p \leq .021$. As before, no effect of name length was found for initial fixation duration, $t_1(7) = 0.34$, $p = .741$. Because faces and names were counterbalanced in this experiment, these findings cannot be attributed to visual or conceptual differences in the stimuli.

In stark contrast to performance in the memory task, visual search showed no effects of syllable number on viewing behavior (Table 1, bottom row), all *p*s > .15. This task dependency is consistent with an explanation for the syllable number effect that relies on inner speech during memory encoding, as well as previous reports of oculomotor task dependencies between reading and search (Rayner & Fischer, 1996; Rayner & Raney, 1996). Because observers in the search task were not attempting to remember the faces in the search display, the names associated with the faces were neither implicitly articulated nor encoded into working memory—resulting in no effect of linguistic structure on oculomotor behavior.[4]

## GENERAL DISCUSSION

When people attempt to remember objects in a multi-item display, the time they spend looking at each object depends on the number of syllables in its name, and more generally, on the time required to subvocalize this name. This finding has important implications for the encoding of verbal and visual information into working memory. Not only are verbal and visual encoding synchronized to a given object in a scene, but the faster of these two processes waits for the slower process to be completed. In the recognition tasks reported here, the observed syllable dependency suggests that the visual encoding of an object required less time than the subvocal generation of its name, a discrepancy resulting in the faster visual encoding device (the eye) waiting for the slower speech production system. This tight visual-verbal coupling imposed by working memory is likely designed to minimize interference. If visual processing were directed to a tree while the verbal system was attempting to encode the word *hippopotamus* into working memory, visual or semantic properties of the tree might interfere with the verbal encoding of *hippopotamus*. Apparently, observers avoid this conflict by delaying movement of gaze to a new object until verbal processing of the fixated object is complete.

The current findings also have implications for the study of visual perception and speech production. For example, scene and object representation are often described strictly in terms of visual processes. An external object is encoded as two-dimensional (Bülthoff & Edelman, 1992; Tarr & Bülthoff, 1998) or three-dimensional (Biederman, 1987; Biederman & Gerhardstein, 1993) visual primitives, then compared with similarly structured internal representations—with the degree of match determining the goodness of recognition. Finding an effect of syllable number on visual behavior suggests that this "pure vision" perspective may be too narrow to account for the complexities of memory encoding during free viewing. Depending on the task, people may elect to supplement their visual representations with verbal encoding, with a by-product of this dual-encoding scheme being a linguistic constraint imposed on the visual processing. Such synchronization is consistent with recent studies showing a dependence between speech production and the moment-by-moment behavior of the oculomotor system in linguistic tasks (Meyer et al., 1998; Tanenhaus et al., 1995). Our study extends this evidence for visual-verbal synchrony beyond the purview of an explicitly linguistic task, showing that linguistic structure affects both the time course of information encoding into working memory and the manner in which people inspect the visual world.

**Acknowledgments—**We thank Kathryn Bock, Gary Dell, David Irwin, and Keith Rayner for helpful comments, and George McConkie and Gary Wolverton for eye-tracking resources. This work was supported by a Beckman Institute Fellowship to G.J.Z. and National Institute of Mental Health Grant MH41704 to G.L.M.

4. The 97-ms difference in total fixation time for visual search (Table 1, bottom row), which appears to suggest an effect of number of syllables, can be attributed almost entirely to the behavior of 1 subject (the difference drops to 12 ms when this participant is removed). Upon postexperiment questioning, this observer reported subvocalizing the names of the faces in the search array. He suspected a future memory test and used this opportunity to practice the associations. No other observer reported using a naming strategy in the search task.

## REFERENCES

Baddeley, A.D. (1986). *Working memory*. Oxford, England: Oxford University Press.

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

Balota, D.A., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. *Cognitive Psychology*, *17*, 364–390.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, *94*, 115–147.

Biederman, I., & Gerhardstein, P.C. (1993). Recognizing depth-rotated objects: Evidence and conditions for three-dimensional viewpoint invariance. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 1162–1182.

Brooks, L. (1967). The suppression of visualization by reading. *Quarterly Journal of Experimental Psychology*, *19*, 289–299.

Bülthoff, H.H., & Edelman, S. (1992). Psychophysical support for a two-dimensional view interpolation theory of object recognition. *Proceedings of the National Academy of Sciences, USA*, *89*, 60–64.

Conrad, R. (1964). Acoustic confusion in immediate memory. *British Journal of Psychology*, *55*, 75–84.
Cooper, R.M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology*, *6*, 84–107.
Dobkins, R.S., Morris, R.K., & Rayner, K. (1992). Lexical ambiguity and eye fixations in reading: A test of competing models of lexical ambiguity resolution. *Journal of Memory and Language*, *31*, 461–476.
Eriksen, C.W., Pollock, M.D., & Montague, W.E. (1970). Implicit speech: Mechanisms in perceptual encoding? *Journal of Experimental Psychology*, *84*, 502–507.
Ferreira, F., & Henderson, J.M. (1990). The use of verb information in syntactic parsing: Evidence from eye movements and word-by-word self-paced reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 555–568.
Groner, R., Walder, F., & Groner, M. (1984). Looking at faces: Local and global aspects of scanpaths. In A. Gale & F. Johnson (Eds.), *Theoretical and applied aspects of eye movement research* (pp. 523–533). Amsterdam: North Holland.
Hatano, G., & Osawa, K. (1983). Digit memory of grand experts in abacus-derived mental calculation. *Cognition*, *15*, 95–110.
Just, M.A., & Carpenter, P.A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, *87*, 329–354.
Keysar, B., Barr, D.J., & Horton, W.S. (1998). The egocentric basis of language use: Insights from a processing approach. *Current Directions in Psychological Science*, *7*, 46–50.
Klapp, S.T., Anderson, W.G., & Berrian, R.W. (1973). Implicit speech in reading, reconsidered. *Journal of Experimental Psychology*, *100*, 368–374.
Kowler, E. (1990). The role of visual and cognitive processes in the control of eye movement. In E. Kowler (Ed.), *Eye movements and their role in visual and cognitive processes* (pp. 1–70). Amsterdam: Elsevier.
Logie, R.H. (1986). Visuo-spatial processes in working memory. *Quarterly Journal of Experimental Psychology*, *38A*, 229–247.
Mackworth, N.A., & Morandi, A.J. (1967). The gaze selects informative details within pictures. *Perception & Psychophysics*, *2*, 547–552.
Meyer, A.S., Sleiderink, A.M., & Levelt, W.J.M. (1998). Viewing and naming objects: Eye movements during noun phrase production. *Cognition*, *66*, B25-B33.
Murphy, G.L., & Brownell, H.H. (1985). Category differentiation in object recognition: Typicality constraints on the basic category advantage. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*, 70–84.
Noizet, G., & Pynte, J. (1976). Implicit labelling and readiness for pronunciation during the perceptual process. *Perception*, *5*, 217–223.
Palmer, J. (1990). Attentional limits on the perception and memory of visual information. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 332–350.
Pynte, J. (1974). Readiness of pronunciation during the reading process. *Perception & Psychophysics*, *16*, 110–112.
Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*, 372–422.
Rayner, K., & Fischer, M.H. (1996). Mindless reading revisited: Eye movements during reading and scanning are different. *Perception & Psychophysics*, *58*, 734–747.
Rayner, K., & Raney, G.E. (1996). Eye movement control in reading and visual search: Effects of word frequency. *Psychonomic Bulletin & Review*, *3*, 245–248.
Samaria, F., & Harter, A. (1994, December). *Parameterization of a stochastic model for human face identification*. Paper presented at the 2nd IEEE Workshop on Applications of Computer Vision, Sarasota, FL.
Snodgrass, J.G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 174–215.
Spoehr, K.T., & Smith, E.E. (1973). The role of syllables in perceptual processing. *Cognitive Psychology*, *5*, 71–89.
Tanenhaus, M.K., Spivey-Knowlton, M.J., Eberhard, K., & Sedivy, J.C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*, 1632–1634.
Tarr, M.J., & Bülthoff, H.H. (1998). Image-based object recognition in man, monkey and machine. *Cognition*, *67*, 1–20.
Yarbus, A.L. (1967). *Eye movements and vision*. New York: Plenum Press.
Zelinsky, G.J. (1996). Using eye saccades to assess the selectivity of search movements. *Vision Research*, *36*, 2177–2187.
Zelinsky, G.J. (1999). Precueing target location in a variable set size "nonsearch" task: Dissociating search-based and interference-based explanations for set size effects. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 875–903.
Zhang, G., & Simon, H.A. (1985). STM capacity for Chinese words and idioms: Chunking and acoustical loop hypotheses. *Memory & Cognition*, *13*, 193–201.

(RECEIVED 3/24/99; ACCEPTED 7/27/99)

 