Tracking the Continuity of Language Comprehension: Computer Mouse Trajectories Suggest Parallel Syntactic Processing

QUERY SHEET

Q1: Au: Still in press? Provide update here, is possible.
Q2: Au: Still in press? Provide updated entry here if now published.
Q3: Au: Spell out full name of journal here.

Cognitive Science 31 (2007) 1–21

Tracking the Continuity of Language Comprehension: Computer Mouse Trajectories Suggest Parallel Syntactic Processing

Thomas A. Farmer[a], Sarah A. Cargill[a], Nicholas C. Hindy[a], Rick Dale[b], Michael J. Spivey[a]

[a]*Department of Psychology, Cornell University*
[b]*University of Memphis*

Received 29 May 2006; received in revised form 22 April 2007; accepted 22 April 2007

Abstract

Although several theories of online syntactic processing assume the parallel activation of multiple syntactic representations, evidence supporting simultaneous activation has been inconclusive. Here, the continuous and non-ballistic properties of computer mouse movements are exploited, by recording their streaming x, $y$coordinates to procure evidence regarding parallel versus serial processing. Participants heard structurally ambiguous sentences while viewing scenes with properties either supporting or not supporting the difficult modifier interpretation. The curvatures of the elicited trajectories revealed both an effect of visual context and graded competition between simultaneously active syntactic representations. The results are discussed in the context of 3 major groups of theories within the domain of sentence processing.

Keywords: Mouse movements; Language comprehension; Syntactic ambiguity; Continuity

---

1. Introduction

Sentences such as, "The adolescent hurried through the door tripped," are difficult to process because, at least temporarily, multiple possible structural representations exist (see Bever, 1970). In this example, *hurried* could either signal the onset of a reduced relative clause, equivalent in meaning to "*The adolescent* who was *hurried through the door . . .* "; or, *hurried* could be interpreted as the main verb of the sentence, such that the adolescent is the entity that willfully hurried. If *hurried* is initially interpreted as the main verb, then processing difficulty

Correspondence should be addressed to Thomas A. Farmer, Department of Psychology, Uris Hall, Cornell University, Ithaca, NY 14853. E-mail: taf22@cornell.edu

is experienced upon encountering the word *tripped* because it requires the less- or non-active reduced relative clause interpretation. This kind of processing difficulty is classically referred to as the garden-path effect.

Contemporary accounts of how the comprehension system processes such syntactic ambiguity can be distinguished based on (a) the degree to which they rely on the activation of one versus multiple syntactic representations at any one time during the comprehension process, and (b) the time frame in which non-syntactic information can constrain interpretation. Syntax-first models (e.g., Ferreira & Clifton, 1986; Frazier & Clifton, 1996) have traditionally proposed that, at a point of syntactic ambiguity, syntactic heuristics alone select a single structure to pursue, and recovery from a misanalysis is achieved via a separate reanalysis mechanism that uses semantic and contextual information. Thus, these models propose that only one representation is active at any given time and that non-syntactic information only influences interpretation at a later reanalysis stage.

Multiple constraint-based theories (e.g., Green & Mitchell, 2006; MacDonald, Pearlmutter, & Seidenberg, 1994; McRae, Spivey-Knowlton, & Tanenhaus, 1998; Trueswell, Tanenhaus, & Garnsey, 1994), on the other hand, describe language comprehension as an interactive process whereby all possible syntactic representations are simultaneously partially active and competing for more activation across time. Unlike the syntax-first models, multiple sources of information, be they syntactic or non-syntactic, integrate *immediately* to determine the amount of activation provided to each of the competing alternatives. In this framework, what feel like garden-path effects are due to the incorrect syntactic alternative winning much of the competition during the early portion of the sentence, and then nonconforming information from the latter portion of the sentence inducing a laborious reversal of that activation pattern. More important, the degree to which the incorrect alternative had been winning the competition early on affects the degree to which the reversal of that activation pattern will be protracted and difficult. As a result, one can expect that some garden-path events may be very mild, some moderate, and some extreme such that a wide variety of sentence readings should all belong to one population of events with a relatively continuous distribution.

Recently, a sort of hybrid account has emerged that combines certain aspects of each of these theories. The Unrestricted Race model (Traxler, Pickering, & Clifton, 1998; van Gompel, Pickering, Pearson, & Liversedge, 2005; van Gompel, Pickering, & Traxler, 2001) follows in the footsteps of constraint-based models in proposing simultaneous integration of multiple constraints from statistical, semantic, and contextual sources. However, rather than ambiguity resolution being based on a temporally dynamic competition process, the Unrestricted Race model posits an instantaneous probabilistic selection among the weighted alternatives of an ambiguity. Therefore, much like the syntax-first models, it must hypothesize a separate reanalysis mechanism that is responsible for garden-path effects when the initial selected alternative turns out to be syntactically or semantically inappropriate. Thus, the Unrestricted Race model predicts that sentences with garden-paths and sentences without garden-paths are two separate populations of events (either reanalysis is needed or it is not). In other words, in conditions where mean performance is expected to exhibit a garden-path effect, there should exist one of two possible patterns: (a) a bimodal distribution of some substantial garden-path responses and some non-garden-path responses, or (b) practically all trials exhibiting substantial garden-path effects. A graded pattern involving some minimal garden paths, some moderate garden paths, and some substantial garden paths is not predicted by the Unrestricted Race model.

One source of evidence often used to distinguish between syntax-first and multiple constraint-based accounts of online language comprehension comes from eye movements recorded during the comprehension of syntactically ambiguous sentences (like 1a of the following list) that are presented auditorily while participants are looking at a relevant visual display:

1a. Put the apple on the towel in the box.
1b. Put the apple that's on the towel in the box.

In example 1a, the prepositional phrase (PP) *on the towel* creates a syntactic ambiguity in that it could be initially interpreted as a destination (or goal) for *the apple*, thus attaching to the verb phrase *Put*; or it could be interpreted as a modifier of *the apple* and thus syntactically attached to that noun phrase. Although corpus analyses have shown that PP attachment ambiguities are in general more frequently noun-phrase attached than verb-phrase attached (Hindle & Rooth, 1993), in the case of the verb *put* and the ambiguous preposition *with*, there exists a reliable lexically motivated bias for verb-phrase attachment (Britt, 1994; Spivey-Knowlton & Sedivy, 1995).

When ambiguous sentences like 1a are heard in the presence of visual scenes where only one possible referent is present (an apple already on a towel), along with an incorrect destination (an empty towel), and a correct destination (a box), as in the top portion of Fig. 1, about 50% of the time participants fixate the incorrect destination after hearing the first PP. After the second disambiguating PP is heard, eye movements tend to be redirected to the correct referent and then to the correct destination. When the unambiguous version of the sentence is heard (1b), participants do not look at the incorrect destination (e.g., the empty towel). The tendency in this one-referent context to look at the incorrect destination until the disambiguating second PP is heard provides evidence of the garden-path effect and is indicative of initially attaching the ambiguous PP to the verb phrase.

This garden-path effect can, however, be modulated by contextual information contained within the visual scene (Snedeker & Trueswell, 2004; Spivey, Tanenhaus, Eberhard, & Sedivy, 2002; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995; Trueswell, Sekerina, Hill, & Logrip, 1999; see also Knoeferle & Crocker, 2006). When two possible referents (say, an apple on a towel and another apple on a napkin) are present (Fig. 1, bottom panel) along with an ambiguous sentence like 1a, participants tend to look at the correct referent (the apple on the towel) and move it to the correct destination while rarely, if ever, looking at the incorrect destination. In accordance with previous studies of referential context (e.g., Altmann & Steedman, 1988; Spivey & Tanenhaus, 1998; van Berkum, Brown, & Hagoort, 1999), then, it seems that when two possible referents are present, an expectation is created that they will be discriminated amongst, thus forcing a modifier interpretation of the ambiguous PP. The attenuation of looks to the incorrect destination by the presence of two possible referents, then, is evidence for an early influence of non-syntactic (even non-linguistic) information on the parsing process and is problematic for traditional syntax-first accounts discussed earlier.

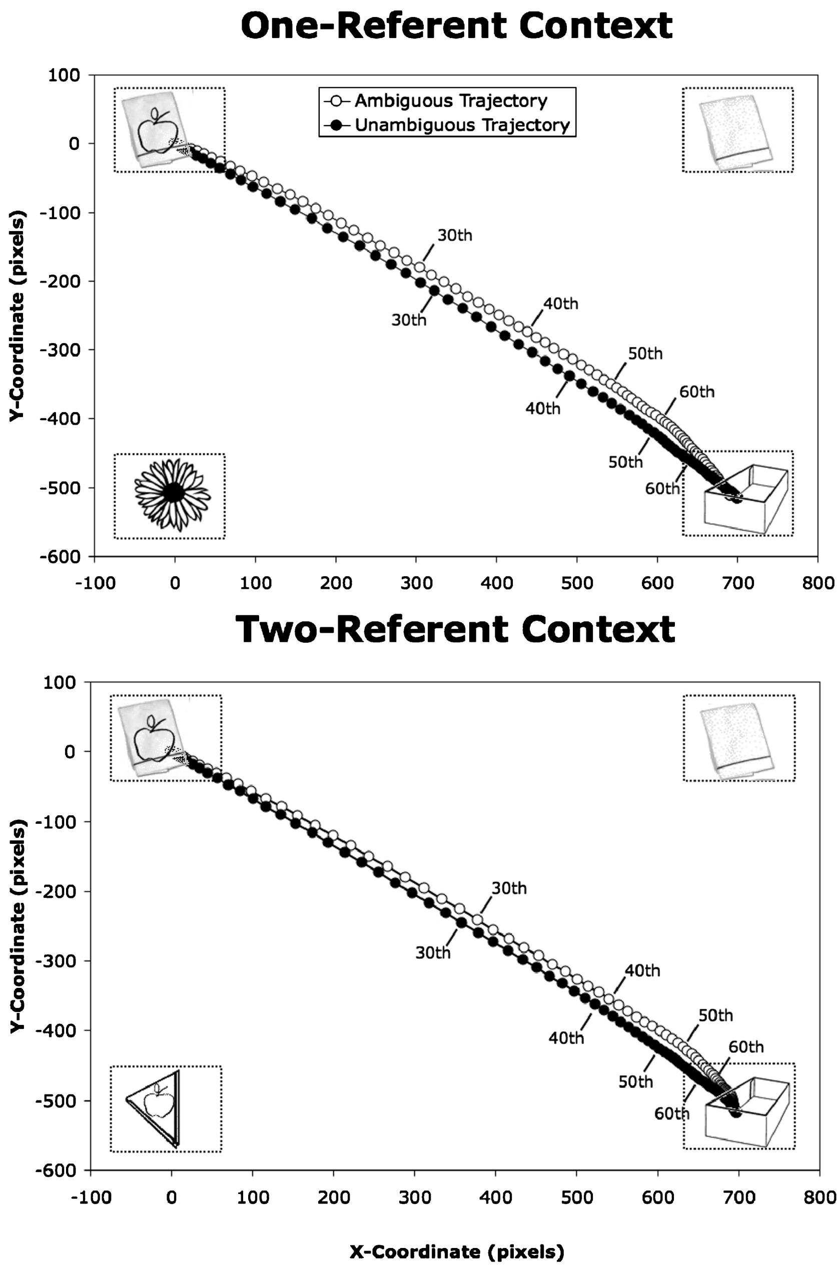

Fig. 1. An example of a one-referent (top) and a two-referent (bottom) display for the instruction, "Put the apple (that's) on the towel in the box." *Note*: The trajectories plotted are the averaged trajectories, per condition, elicited in each context; and the numbers "30th" through "60th" denote a point's timestep. Due to the horizontally elongated shape of the overall display, differences in x coordinates of the mouse movements are somewhat more indicative of velocity differences, and differences in the $y$coordinates are more indicative of genuine spatial attraction toward the incorrect destination in the upper right corner. Substantial statistically reliable$x$- and y-coordinate divergence existed between the two sentence conditions in the one-referent context, but both the x and the y coordinates for the ambiguous- and unambiguous-sentence trajectories were statistically indistinguishable in the two-referent context.

Although early contextual effects elicited in these and similar visual-world experiments strongly support constraint-based models of human sentence processing over syntax-first models, eye-movement data do not readily afford a clear discrimination between constraint-based and unrestricted race accounts of the data. Within the one-referent context, one might expect that if both possible representations of the ambiguous PP were simultaneously active (as predicted by the constraint-based approaches), participants might, as frequently observed (Spivey et al., 2002; Tanenhaus et al., 1995), look back and forth between the competitor objects. However, because saccadic eye movements are generally ballistic, they either send the eyes to fixate an object associated with a garden-path interpretation or they do not. The evidence from this paradigm, therefore, is also consistent with the Unrestricted Race model, where the various constraints are combined immediately, but on any given trial only one syntactic representation is initially pursued—that is, across experimental trials, distributions of eye-movement patterns are almost always bimodal because the fixations are coded as binomial. There are saccades to locations on the display corresponding to either one of the possible representations, but almost never to a blank region in between those two potential targets. In the following experiment, we examined the dynamics of hand movement in the same sentence comprehension scenario with the goal of determining whether the non-ballistic, continuous nature of computer mouse trajectories can serve to tease apart these two remaining theoretical accounts.

2. Experiment 1

Recently, it has been demonstrated that continuous nonlinear trajectories recorded from the streaming x, y coordinates of computer mouse movements can serve as an informative indicator of the cognitive processes underlying spoken-word recognition (Spivey, Grosjean, & Knoblich, 2005), categorization (Dale, Kehoe, & Spivey, 2007), and referential communication (Brennan, 2005). Although individual saccadic eye movements can occasionally show some curvature (Doyle & Walker, 2001; Port & Wurtz, 2003) and some informative variation in landing position (Gold & Shadlen, 2000; Sheliga, Riggio, & Rizzolatti, 1994), individual movements of the arm and hand can show quite dramatic curvature (Goodale, Pélisson, & Prablanc, 1986; Song & Nakayama, 2006; Tipper, Howard, & Jackson, 1997), which can be interpreted as the dynamic blending of two mutually exclusive motor commands (Cisek & Kalaska, 2005; Tipper, Howard, & Houghton, 2000). In addition, whereas self-paced reading affords 2 to 3 data points (button presses) per second, and eye-movement data allow for approximately 3 to 4 data points (saccades) per second, "mouse tracking" yields somewhere between 30 and 60 data points per second, depending on the sampling rate of the software used. In light of the ability to record many data points per second, and in light of their ability to curve mid-flight as a result of competition between multiple potential targets, mouse movements have the ability to convey the continuity of processing.

The context and garden-path effects reported in the visual world paradigm are highly replicable when tracking eye movements (Snedeker & Trueswell, 2004; Spivey et al., 2002; Tanenhaus et al., 1995; Trueswell et al., 1999). As such, recording mouse movements in the visual world paradigm can serve as a strong test case by which to evaluate the efficacy of the mouse-tracking procedure for the study of language processing in real time. If the mouse-tracking technique can produce results from the visual world paradigm commensurate with those obtained by tracking eye movements, we would predict that:

Averaged trajectories recorded in response to ambiguous sentences in the one-referent context should show significantly more curvature toward the incorrect destination than the averaged trajectories elicited by unambiguous sentences—a pattern corresponding to the garden-path effect.

The curvature of averaged trajectories in the two-referent condition should not differ statistically between ambiguous and unambiguous sentences, thus demonstrating an influence of referential context on the garden-path effect.

If the influence of referential context is observed, it would provide further evidence against the traditional syntax-first models, but would be consistent with either the constraint-based or the unrestricted race accounts of syntactic processing. The second purpose of this study, then, was to exploit the continuity of the mouse-movement trajectories to discriminate between these two remaining theoretical accounts. To do so, a measure of curvature magnitude was used to determine the amount of spatial attraction toward the incorrect destination that was exhibited by the ambiguous- and unambiguous-sentence trajectories in the one-referent context. If only one representation were active at any one time, as the unrestricted race account predicts, then the trial-by-trial distribution of trajectory curvatures in the ambiguous-sentence condition should be either (a) bimodal—comprised of highly curved garden-path movements and non-curved, correct-interpretation movements; or (b) uniformly in the more extreme curved range, indicating that almost every trial exhibited a garden-path effect. In contrast, as predicted by the constraint-based approach, if both representations were active and competing simultaneously, one should expect to see a unimodal distribution with a continuous range of non-, somewhat-, and highly curved trajectories—that is, a gradation of "garden pathing."

2.1. Method

2.1.1. Participants

Forty right-handed, native English-speaking undergraduates from Cornell University participated in the study for extra credit in psychology courses. We used only right-handed individuals to avoid variability associated with subtle kinematic differences in leftward and rightward movement of the left versus the right arms.

2.1.2. Materials and procedures

Sixteen experimental items, along with 102 filler sentences, were adapted from Spivey et al. (2002) and digitally recorded. The unambiguous version (1b) of each of the 16 experimental items was recorded first, and then the "that" was removed to produce the ambiguous (1a) sentence condition (see Spivey et al., 2002 for details). Each visual context corresponding to the 16 experimental items was varied to produce a one- and two-referent condition. The one-referent visual context (illustrated in Fig. 1, top) contained the target referent (an apple on a towel), an incorrect destination (a second towel), the correct destination (a box), and a distracter object (a flower). In the two-referent context, all items were the same except that the distracter object was replaced with a second possible referent (such as an apple on a napkin). Twenty-four filler scenes, designed to accompany filler sentences, were also constructed.

Spoken instructions with a single male voice were recorded using Mac-based digital audio recording software. At the beginning of each sound file for every item (consisting of a set of 3 instructions), participants first heard, "Place the cursor at the center of the cross." Then, for the sound files accompanying scenes that were to be paired with experimental items, the experimental sentence always occurred second, followed by two additional unambiguous filler instructions. For the filler-item scenes corresponding to items without any experimental manipulation, participants heard three scene-appropriate unambiguous instructions. In all cases, 2 sec separated the offset of one sentence from the onset of the next sentence within each item.

In critical trials for both the one- and two-referent conditions, the target referent always appeared in the top left corner of the screen, the incorrect destination always appeared in the top right corner of the screen, and the correct destination was always located at the bottom right portion of the screen. The distracter object in the one-referent trials and the second referent in the two-referent trials always appeared in the bottom left corner of the screen. Given that the scene layout was held constant across all items in each experimental condition, a left-to-right movement was always necessary. Although there could exist a systematic bias toward specific locations in the display when moving rightward, this was viewed as unproblematic given that the bias would be held constant across both the ambiguous and unambiguous sentences, which were directly compared in all statistical analyses, for each context. The filler sentences were constructed to prevent participants from detecting any statistical regularities created by the object placements in the experimental trials. In addition to the movement used in the experimental instructions, 11 distinct movements were possible in the visual scene across trials, and an approximately equal number of filler sentences (either 8 or 10) were assigned to each of these movements. Therefore, 10 sentences required an object in the upper left-hand corner of the display be moved to the upper right-hand corner of the display, 8 sentences required an object in the upper left-hand corner of the display be moved to the bottom left-hand corner of the display, and so on.

In each scene, participants saw four to six color images, depending on how many objects were needed for the scene. The images were constructed from pictures of real objects taken by a digital camera and edited in Adobe Photoshop. The visual stimuli subtended an average of $5.96^\circ \times 4.35^\circ$ of visual angle and were positioned 14.38° diagonally from the central cross. The mouse movements were recorded at an average sampling rate of 40 Hz.

The experimental items were counterbalanced across four presentation lists. Each list contained four instances of each possible condition but only one version of each sentence frame and corresponding visual context. Two filler sentences were included with the experimental items as described earlier, and three filler sentences were included with each of 24 distracter scenes. The presentation order was randomized for each participant. Participants were randomly assigned to one of the four presentation lists.

2.2. Results

2.2.1. Data screening and coding

Mouse movements were recorded during the grab-click, transferal, and drop-click of the referent object in the experimental trials. As a result of the large number of possible trajectory

Table 1
The errors causing for a trial to be excluded from all analyses, per condition

Error Type	One Referent, Ambiguous	One Referent, Unambiguous	Two Referent, Ambiguous	Two Referent, Unambiguous
Target referent moved to incorrect destination	6	2	1	1
Incorrect referent moved to incorrect destination	2	0	2	0
Picture representing a destination was moved	0	0	5	0
Erratic movement yielding an uninterpretable trajectory	5	1	2	0

shapes, the x, y coordinates for each trajectory from each experimental trial were plotted to detect the presence of any aberrant movements. A trajectory was considered valid and submitted to further analysis if it was initiated at the top left quadrant of the display and terminated in the bottom right quadrant, indicating that the correct referent had been picked up and then placed at the correct destination. This screening procedure resulted in 27 deleted trials, accounting for less than 5% of all experimental trials.

The types of errors that resulted in the exclusion of a trial, along with their frequency of occurrence per condition, are presented in Table 1. The most frequent error involved placing the correct referent on the incorrect destination, with no evidence of a corrective movement toward the intended destination. In addition, errors classified as "erratic" typically contained aberrant movements of the correct referent that can be characterized best as oscillating between rightward movement and leftward movement, with the correct referent either making it eventually to the correct destination or not. A 2 (Context) $\times$2 (Ambiguity) analysis of variance (ANOVA) on the number of included trials per condition yielded no significant main effect of context, $F(1, 39) = 1.20$, *ns*; or two-way interaction, $F(1, 39) = 0.01$, *ns*. There was, however, a significant main effect of ambiguity, $F(1, 39) = 9.78$, $p = .003$, mean square error (*MSE*) $= .134$, with more trajectories included in the unambiguous ($M = 7.9$, $SD = .38$) than in the ambiguous ($M = 7.42$, $SD = .98$) conditions. The fact that more trials were excluded in the ambiguous conditions is not surprising in light of the increased difficulty associated with the processing of these sentences and is consistent with error rates in eye-tracking experiments of this type where there are more movement-related errors on ambiguous than on unambiguous trials (Trueswell et al., 1999).

To make sure that trajectories in one condition were not initiated (or that objects were not grabbed) at a systematically different region of the display than in the other conditions, we conducted two 2 (Context) $\times$ 2 (Ambiguity) ANOVAs on the x and y coordinates, separately. There was no significant main effect or interaction for either the x or the y coordinates (all ps were nonsignificant) indicating that, across conditions, the trajectories were initiated at approximately the same location of the display. Subsequently, all analyzable trajectories were "time normalized" to 101 timesteps by a procedure described in Spivey et al. (2005) and Dale et al. (2007). All trajectories were spatially aligned so that their first recorded point corresponded to x, ycoordinates of (0, 0). Although the time-normalized data mirror the general trends evident in raw x- and y-coordinate analyses (see the following), they are much more detailed and fine grained, thus affording more precise information about hand location across time.

2.2.2. *Context and garden-path effects*

The mean trajectories from ambiguous and unambiguous sentences in the one-referent context, illustrated in Fig. 1 (top), demonstrate that the average ambiguous-sentence trajectory was more curved toward the incorrect destination than the average trajectory elicited by the unambiguous sentences. The point-labels "30th" through "60th" denote a data point's corresponding normalized timestep; and they reveal that, in the one-referent context, the average trajectory for the unambiguous sentences traveled to the correct destination much more quickly than did the average trajectory elicited by the ambiguous sentence. Both of these observations support the notion that participants were garden pathed by the syntactic ambiguity manipulation.

In our initial analysis, we conducted a series of t tests to discern whether the divergences observed across the ambiguous- and unambiguous-sentence trajectories in the one-referent context were statistically reliable and to determine whether any statistically reliable divergence existed in the two-referent context. Due to the horizontally elongated shape of the overall display, differences in x coordinates of the mouse movements are somewhat more indicative of velocity differences, and differences in the y coordinates are more indicative of genuine spatial attraction toward the incorrect destination in the upper right corner. As such, the t tests were conducted across the x coordinates of each sentence condition and the y coordinates of each sentence condition, separately, at each of the 101 timesteps. To avoid the increased probability of a Type-1 error associated with multiple t tests, and in keeping with Bootstrap simulations of such multiple t tests on mouse trajectories (Dale et al., 2007), an observed divergence was not considered significant unless the coordinates between the ambiguous- and unambiguous-sentence trajectories elicited p values $< .05$ for at least eight consecutive timesteps.

In the one-referent context, two significant divergences were found when comparing the x coordinates from the ambiguous- and unambiguous-sentence trajectories at each timestep. The comparisons between sentence conditions from Timestep 41 to Timestep 54 all elicited p values $< .05$ (all ts > 2.057, average effect size $d = .348$). There were also significant differences (ps $< .05$) in x coordinates from Timesteps 64 to 79 (all ts > 2.05, average effect size $d = .347$). The y coordinates at each timestep were compared in the same manner for the ambiguous- and unambiguous-sentence trajectories in the one-referent context. The t tests revealed differences in y coordinates from Timesteps 29 through 82 (all ps $< .05$, all ts > 2.068, average effect size $d = .433$).[1]

In the two-referent context, the same analyses were conducted on the x and y coordinates from the ambiguous- and unambiguous-sentence trajectories at each timestep. For both the x-coordinate and y-coordinate comparisons, it is important to note that no t test yielded a p value $< .05$ at any of the 101 timesteps.

To address concerns associated with multiple comparisons in the previous $t$tests, and to assess directly the statistical reliability of the Context × Ambiguity interaction, we conducted

Table 2
Means (and standard errors) for the middle segment analyses of variance

Set	Context	Sentence Type	Mean Coordinate (SE)
x	One referent	Ambiguous	527.02 (22.47)
		Unambiguous	575.95 (18.26)
	Two referent	Ambiguous	613.15 (11.70)
		Unambiguous	592.14 (14.01)
y	One referent	Ambiguous	−340.06 (19.79)
		Unambiguous	−406.12 (13.81)
	Two referent	Ambiguous	−416.47 (11.13)
		Unambiguous	−419.95 (9.84)

two separate 2 $\times$2$\times$ 3 ANOVAs: one for x coordinates and one for y coordinates. Based on normalized timesteps, x and y coordinates were grouped into three time bins: 1 to 33, 34 to 67, and 68 to 101, yielding the third independent variable of time segment. The three-way interaction was significant for the x coordinates, $F(2, 78) = 5.06$, $p = .009$; and for the y coordinates, $F(2, 78) = 48.75$, $p < .0005$.[2] As can be observed in Fig. 1, and as demonstrated by the t tests above, the effect is especially prevalent among the points comprising Time Segment 2. As such, only the Context $\times$ Ambiguity interaction at Time Segment 2 is considered in further detail here.

In this middle time segment, the Context $\times$ Ambiguity interaction was significant for both the x coordinates, $F(1, 39) = 7.15$, $p = .011$, $MSE = 6,844$; and the y coordinates, $F(1, 39) = 8.13$, $p = .007$, $MSE = 4,819$. The means and standard errors for all possible combinations of the independent variables in these x- and y-coordinate analyses appear in Table 2. To assess the context effect, we compared each point in the one-referent context to its commensurate point in the two-referent context. For the x coordinates, there was no difference between coordinates in the one-referent context versus the two-referent context for the unambiguous sentences, $t(39) = 0.99$, *ns*; but there was for the ambiguous sentences, $t(39) = 4.14$, $p < .0005$, $d = .655$; with the x coordinates for the two-referent context being closer to the correct destination. Likewise, for the y coordinates, there was no difference in average screen location for the unambiguous sentences in the one- versus two-referent context, $t(39) = 1.26$, *ns*; but there was for the ambiguous sentences, $t(39) = 3.71$, $p = .001$, $d = .586$; with the y coordinates in the one-referent condition being closer to the top of the display.

In relation to the ambiguity effect for the x coordinates in this middle time segment, there was no significant difference between ambiguous- and unambiguous-sentence trajectories in the two-referent context, $t(39) = 1.65$, *ns*; but there was in the one-referent context, $t(39) = 2.17$, $p = .036$, $d = .343$; with x coordinates from the unambiguous-sentence trajectories being closer to the right of the display. For the y coordinates, there was no significant difference in location between ambiguous- and unambiguous-sentence trajectories in the two-referent context, $t(39) = .31$, *ns*. However, in the one-referent context, the y coordinates for the ambiguous-sentence trajectories were significantly closer to the incorrect destination than were the y coordinates for the unambiguous-sentence trajectories, $t(39) = 3.13$, $p = .003$, $d = .495$.

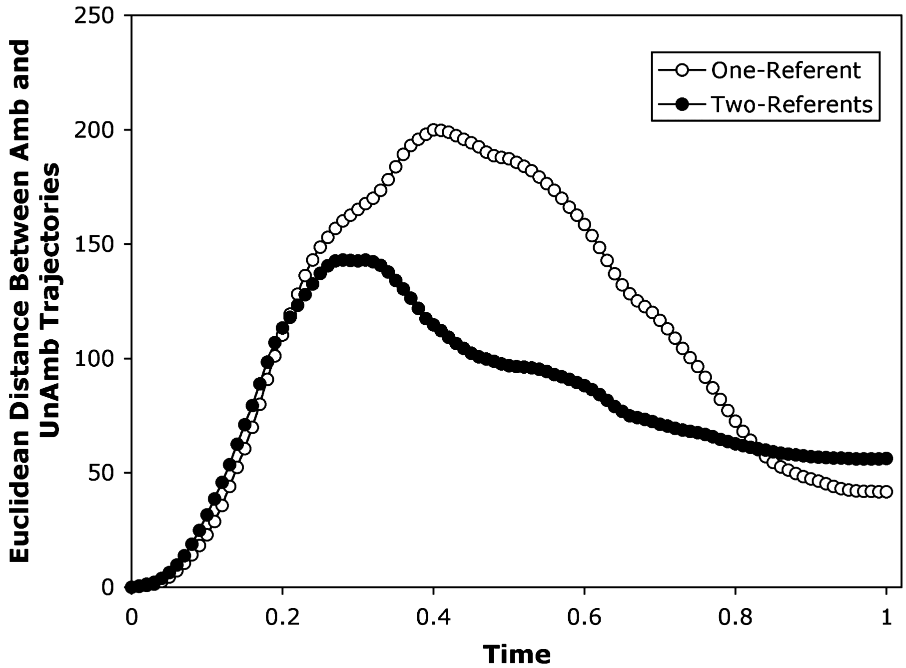

Fig. 2. The Euclidean distance between the ambiguous- and unambiguous-sentence conditions, per context.

To account for both the x and y coordinates in one analysis, we computed the average Euclidean distance at each timestep between corresponding timesteps in the ambiguous- and unambiguous-sentence conditions, per context. Figure 2 illustrates that the distance between the ambiguous and unambiguous trajectories in both contexts is similar during the beginning of the trial but then diverges such that the distance between the conditions is considerably larger in the one-referent than in the two-referent context.

Paired-samples t tests, conducted at each timestep as those above, revealed differences in the Euclidean distance between ambiguous and unambiguous sentences in the one- versus two-referent context from Timesteps 37 through 73, all ps $< .05$ (all ts > 2.11, average effect size $d = .459$). In Fig. 1, the averaged ambiguous-sentence trajectory in the one-referent condition is numerically closer to the incorrect destination than its corresponding unambiguous-sentence trajectory across all timesteps. Thus, in the presence of the garden-path effect, it seems clear that there exists more spatial attraction toward the incorrect destination for the ambiguous sentences. It should be noted that the Euclidean distance measure includes both the velocity and spatial attraction effects that cannot be readily delineated given the properties of the scene layout used here. Therefore, in the analyses of the two-referent context, although the ambiguous- and unambiguous-sentence trajectories are statistically indistinguishable when analyzing x (more indicative of velocity) and y (more indicative of spatial attraction toward the competitor) coordinates separately, their combined effects do produce some small coordinate differences between the two sentence conditions. These small coordinate differences in the two-referent condition are, however, largely due to the trajectory in the *ambiguous* condition being faster—perhaps due to the fact that the unambiguous sentence has a slight delay introduced by the word “that’s.”

Although analyses of the time-normalized trajectories reveal significant attraction to the incorrect destination in the one-referent ambiguous-sentence condition, two potential criticisms remain. First, it could be argued that the trajectories were initiated, and divergence observed, well after the completion of the spoken sentence, rendering the trajectories, essentially, offline. In addition, in light of the velocity difference seen in the one-referent context in Fig. 1 in which the correct object arrives at the correct destination faster in the unambiguous sentence condition, it could be argued that velocity differences, and not spatial attraction, are driving the statistical significance of the divergence.

To address these concerns, we returned to the raw timestamps in the trajectories (and their correspondence with portions of the spoken sentences) by examining the average x and y coordinates at each of eight different time bins. The first time bin was composed of the time between the onset of the second (disambiguating) PP up to 250 msec past the onset of that second PP. Each of the following time bins consisted of consecutive incremental 250 msec intervals, ending with 1,750 to 2,000 msec after the onset of disambiguation.[3] As illustrated in Fig. 3, the trajectories in the ambiguous-sentence condition always lag behind the unambiguous-sentence trajectories in the one-referent condition (x coordinates) and are always closer to the incorrect destination (y coordinates). To assess the statistical reliability of these divergence trends, we conducted a t test between the average ambiguous- and unambiguous-sentence trajectories at each of the eight time bins for x and y coordinates, separately. To correct for multiple comparisons, the Bonferroni adjustment was used, yielding an adjusted alpha cutoff value of $.05/8 = .00625$.

For the x coordinates recorded in the one-referent context, average unambiguous- sentence trajectories diverged significantly from average ambiguous-sentence trajectories at Time bin 4 (750–1,000 msec), $t(32) = 3.58$, $p = .001$, $d = .624$; and Time bin 6 (1,250–1,500 msec), $t(38) = 2.95$, $p = .005$, $d = .47$; and marginally significant at Time bin 5, $t(37) = 2.76$, $p = .009$. Thus, we see that in this context, ambiguous-sentence trajectories took significantly longer to reach the correct destination than their unambiguous counterparts. More important for the goals of this study, however, we see that there was also significant spatial attraction to the competing incorrect destination. Corresponding analyses of the y coordinates recorded in the one-referent condition reveal substantial attraction toward the incorrect destination from Time bins 4 through 8 (all ts > 3.20, all ps $< .003$, average effect size $d = .63$). Figure 3 (bottom panel) illustrates that average y coordinates from the ambiguous-sentence condition were indeed closer to the top of the screen (y-pixel values closer to zero) than were those of the unambiguous-condition trajectories. In addition, in line with the time-normalized analyses presented above, none of the eight time bins in the two-referent context showed the ambiguous- and unambiguous-sentence trajectories significantly diverging for either the x or the y coordinates.

2.2.3. Serial versus parallel activation

We examined response distributions in the garden-path condition to determine whether one or both syntactic representations were active (see Gibson & Pearlmutter, 2000; Lewis, 2000). As an initial attempt to assess whether the distribution of trajectory curvatures in the one-

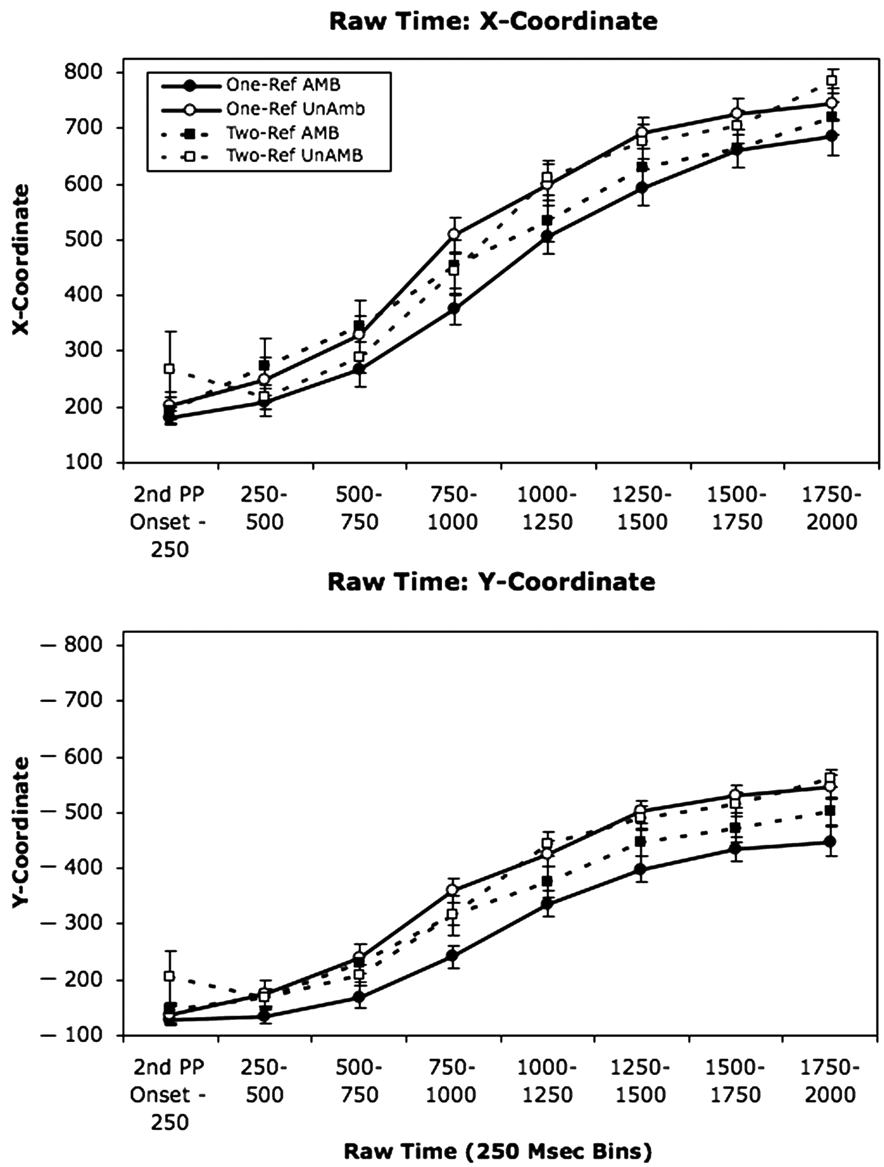

Fig. 3. Raw time *x* and *y* coordinates. *Note*: In the one-referent context (solid bars), raw non-normalized time bins show *x* pixels and *y* pixels converging more directly on the correct destination when the instruction is unambiguous than when it is ambiguous. In the two-referent context (dashed bars), this difference between ambiguous and unambiguous instructions is not significant. (Greater positive *x* values indicate rightward movement, and negative *y* values indicate downward movement.)

referent ambiguous (garden-path) condition was bimodal (thus indicating only discrete garden paths and discrete non-garden paths), we plotted together each of the 146 time-normalized trajectories in that condition, along with a time-normalized reference line from (0, 0) to (700, –500). Figure 4 (top panel) illustrates that although there were some extreme garden-path trials

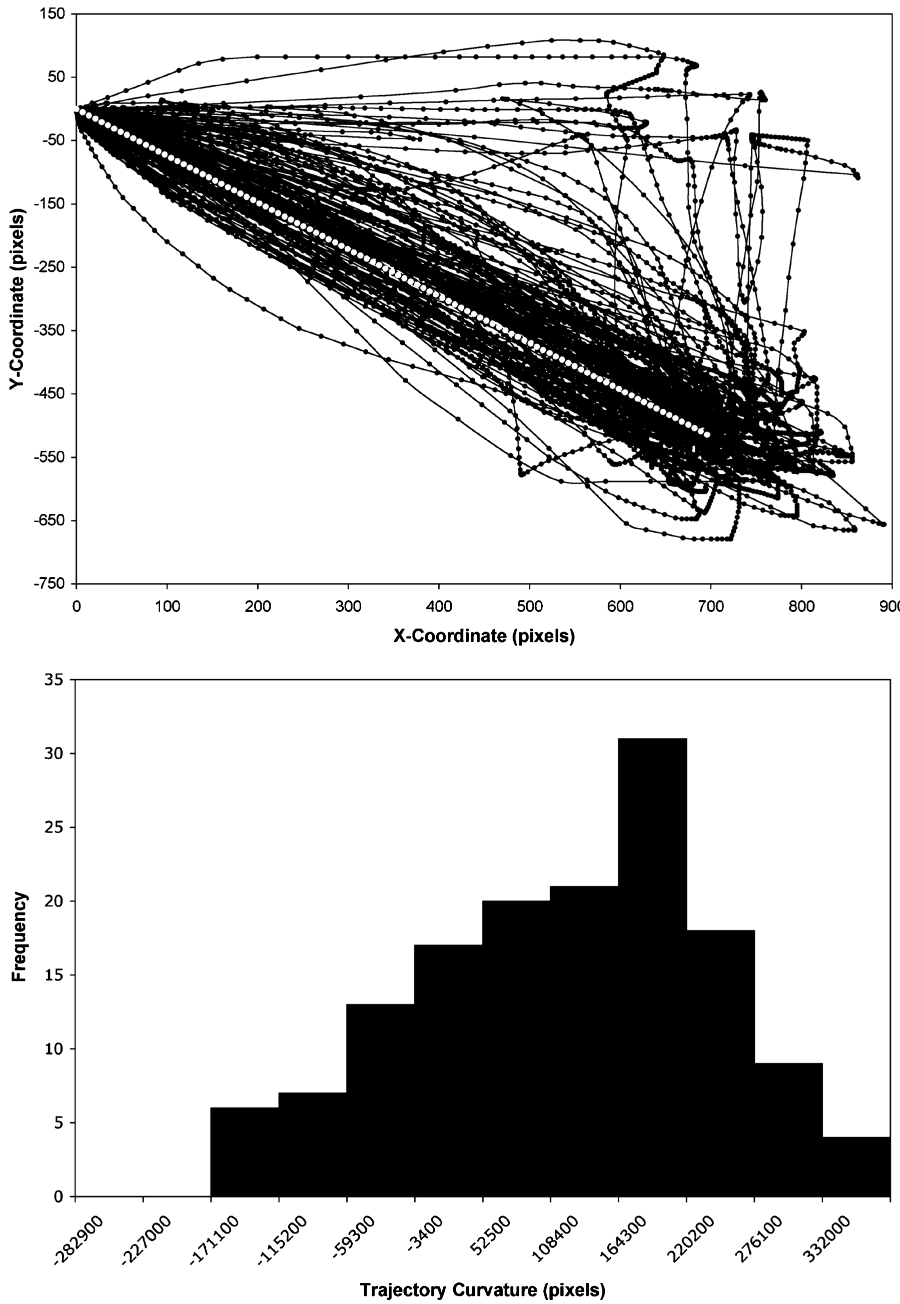

Fig. 4. Distributions of trajectory curvature in the one-referent ambiguous sentence condition. Note: The top panel illustrates, graphically, that most trajectories curved above a time-normalized reference line (the line of white points) thus illustrating, trial-by-trial, the garden-path effect. The bottom panel illustrates that the distribution of trajectory curvatures is indeed unimodal.

Table 3
Statistics necessary for assessing the bimodality of a distribution

Condition	n	Variance	Skewness	Kurtosis	Bimodality (*b*)
One referent, ambiguous	147	1.477E + 10	−.289	−.535	.429
One referent, unambiguous	157	1.699E + 10	−.126	−1.141	.529
Two referents, ambiguous	150	1.629E + 10	−.387	−.731	.493
Two referents, unambiguous	159	1.647E + 10	−.545	−.533	.514

and some non-garden-path trials, the majority of the trajectories elicited in this condition fell somewhere in between those two extremes, forming a single population of non-, somewhat-, and highly curved responses.

To determine whether any bimodality is present in the distribution of responses, we computed the area under the curve on a trial-by-trial basis. First, the straight line from the starting to the ending coordinates of each observed trajectory was normalized to 101 timesteps. Then the total area (in pixels) between that straight line and the observed trajectory was calculated, resulting in an index of trajectory curvature. Area subtending toward the incorrect destination was coded as positive area, and area subtending in the opposite direction from the straight line was coded as negative area. Area of curvature is positively correlated with an alternative measure of curvature, maximum deviation (Atkeson & Hollerbach, 1985), but steady increases in curvature will result in much steeper increases of area than in maximum deviation. Thus, with a much greater range of values in the area measure, the opportunity to observe bimodality in the distribution of curvatures is optimized.

Figure 4 (bottom panel) illustrates the shape of the distribution of trajectory curvatures for the one-referent, ambiguous-sentence trials. As an index of bimodality, we calculated the bimodality coefficient *b* (SAS Institute, 1989, based on work by Darlington, 1970—see DeCarlo, 1997, for a discussion), which has a standard cutoff value of $b = .555$; with values greater than .555 indicating the presence of bimodality.[4] Although we focus on the one-referent ambiguous response distribution here, Table 3 presents the descriptive statistics for each condition's distribution, along with its corresponding bimodality statistic value. The *b* value for each distribution is less than .555, indicating no presence of bimodality within the distributions. Notably, with regard to the distribution of responses in the one-referent, ambiguous-sentence condition, $b < .555$ indicates that the graded spatial attraction effects elicited in this condition came not from two different types of trials but from a single population of trials.

To explore further the modality of the distribution, we compared the area-under-the-curve values in the one-referent, ambiguous-sentence condition (where garden pathing was observed) to the one-referent, unambiguous-sentence condition (where no garden paths were predicted by any of the theories outlined in the introduction) and observed very similar distributional properties. The means are, of course, different, but the standard deviations are nearly identical

P1: TPA
HCOG_06_252943 HCOG.cls August 4, 2007 16:39

($SD = 121,500$ and $SD = 130,300$ for the ambiguous- and unambiguous-sentence conditions, respectively), as are the interquartile ranges (178,110 and 221,470). In fact, when the shapes of the two distributions are compared directly through the Kolmogorov–Smirnov goodness-of-fit test, we find that they are not statistically different, $p > .10$. Distributional characteristics of a population of trials that every theory expects would have a unimodal distribution with no garden pathing (the unambiguous-sentence condition) and those of a population of trials that should have *substantial* garden pathing are, in fact, not distinguishable. This suggests that there is no greater evidence of bimodality in the garden-path condition (where certain theories predict it) than in the unambiguous control condition (where no theory predicts it).

Finally, one might argue that bimodality was not detected (thus, $b < .555$) in the crucial one-referent, ambiguous-sentence condition due to a lack of statistical power resulting from the relatively small number of trials in the garden-path distribution. To address this concern, we created an artificial distribution with a sample size almost identical to our crucial garden-path distribution by randomly sampling 50% of the trials from the one-referent, ambiguous-sentence condition (where garden pathing was observed) and 50% of the trials from the one-referent, unambiguous-sentence condition. This "combination" distribution should produce the response distribution that the unrestricted race account predicts for equibiased syntactically ambiguous sentences—one in which a garden path would either occur due to the discrete selection of the ultimately incorrect representation or would not occur, due to the discrete selection of the ultimately correct alternative.

By examining the distributional properties of the area-under-the-curve values produced by the garden-path and non-garden-path trials together, we can thus determine whether the bimodality statistic (b) we used to assess the bimodality of the garden-path distribution (above) is capable of detecting bimodality in a case where the response distribution should clearly be bimodal. Indeed, the bimodality coefficient elicited by this combination distribution ($n = 151$, skew $= -.266$, kurtosis $= -1.19$) was $b = .572$. The fact that this bimodal "combination" distribution did elicit a b value above the absolute cutoff of .555 illustrates that with the sample size used in this study, the bimodality coefficient is capable of detecting bimodality when it should be present (see also Farmer, Cargill, & Spivey, in press , for additional experimental work showing that the mouse-tracking technique can produce bimodal distributions of curvature when they are expected and that the statistical methods employed here will detect that bimodality).

Q1

3. General discussion

Converging evidence from the foregoing analyses illustrates that the effects traditionally associated with the visual-world paradigm (Spivey et al., 2002; Tanenhaus et al., 1995) are replicable with the mouse-tracking methodology (see also Magnuson, 2005; Spivey et al., 2005). In the one-referent context, participants' mouse movements in response to the ambiguous sentences curved significantly closer to the top right of the screen (toward the incorrect destination) than in response to unambiguous sentences. Thus, it would seem that when only one referent was present, the incorrect destination (e.g., the towel) was partially considered relevant, until disambiguating information was processed—a trend corresponding to the garden-path effect associated with this condition. More important, any statistically detectable divergence between the *x* and *y* coordinates of the trajectories in the ambiguous- and unambiguous-sentence conditions was completely absent in the two-referent context, demonstrating that visual context can prevent the syntactic garden path. The fact that most mouse trajectories began while the speech file was still being heard suggests that the effect of visual context modulating the garden path took place during early moments of processing the linguistic input, not during a second stage of syntactic reanalysis. Indeed, the timeframe in which significant divergence was observed in the one-referent condition—within 1 sec of the onset of the disambiguating PP—is within the same period of time (relative to the spoken sentence) as when many of the critical fixations of competing objects occur in the visual-world paradigm (Chambers, Tanenhaus, & Magnuson, 2004; Spivey et al,. 2002; Tanenhaus et al., 1995; Trueswell et al., 1999).

In addition, by capitalizing on the continuous, non-linear, and non-ballistic properties of trajectories produced by computer mouse movements, mouse tracking has the potential to answer questions that have been difficult to answer with more traditional methodologies. The context effect in the two-referent condition is problematic for syntax-first models of sentence processing, but does not distinguish between constraint-based and unrestricted race accounts. What does distinguish between these two accounts is the gradiency observed in the curvature of the trajectories in the garden-path condition. If the Unrestricted Race model posits that only one syntactic representation is pursued at any one time, then it must predict that mouse movements in a garden-path condition should initially move either in the direction of the correct destination or in the direction of the incorrect destination (producing either a bimodal distribution or an all-curved distribution). In contrast, because the constraint-based account posits simultaneous graded activation of multiple syntactic alternatives, it predicts that mouse movements can move in directions that are dynamically weighted combinations of the two competing destinations (producing a unimodal distribution of moderate curvatures).

Figure 4 shows that although approximately 5% of the trajectories moved all the way to the incorrect destination before changing direction, the vast majority of the trajectories responsible for the mean curvature were unmistakably graded in their degree of spatial attraction toward the incorrect destination. The lack of bimodality in the distribution of trial-by-trial trajectory curvatures suggests that the garden-path effect so frequently associated with this manipulation is not an all-or-none phenomenon—that is, the activation of one structural representation does not forbid simultaneous activation of other possible representations. Instead, the garden-path effect is graded, meaning that although sometimes one syntactic alternative may have greater activation than another, it is also the case that, until disambiguating information is presented, both can be considered in parallel, and the simultaneously active representations may compete for activation over time. Tabor and Hutchins (2004) recently offered evidence of this interpretation. By increasing the length of the region that introduces a garden path, they showed an increase in the time required to reverse the activation of an incorrect interpretation. Results reveal the gradual commitment to one syntactic interpretation, rather than a discrete selection of one with the immediate dismissal of the others. Their findings, along with the results presented here, appear to strongly support constraint-based accounts of syntactic processing as outlined in the introduction.

More broadly, these results demonstrate that the mouse-tracking technique can be used with tasks that involve complex and interactive displays. We believe that mouse tracking is a viable method for examining online language processing in a wide array of cognitive tasks and across a relatively large age range. Through a large-scale survey of children's computer use, for example, Calvert, Rideout, Woolard, Barr, and Strouse (2005) found that the mean age at which a child was able to point and click a computer mouse was 3.5 years, and that the mean age of the onset of autonomous computer use was 3.7 years. This observation suggests that experiments employing the mouse-tracking procedure could be feasible with children as young as 3.5 to 4 years of age, a population for which real-time measures of cognitive processing are often hard to find. In addition, in light of its accessible, portable, and inexpensive nature, and in light of the replicability of results across the eye- and mouse-tracking methodologies, we believe mouse tracking can serve as "the poor man's eye tracker," providing detailed indices of cognitive processing to laboratories that cannot afford expensive eye-tracking equipment. Finally, it is important to note that we do not advocate, or foresee, the usurping of eye-tracking methods in lieu of the advantages of mouse tracking enumerated here. Instead, we believe that the two techniques can be used in a complementary (even simultaneous) fashion to more fully unlock the nature of the complex interactions associated with high-level cognitive processes.

Notes

1. After examining the trial-by-trial distribution of trajectory curvatures in the one-referent, ambiguous-sentence condition (Fig. 4), one might be concerned that the significant divergences reported are an artifact of the trials in which an extreme garden path occurred (as indicated by movements all the way to the far upper right corner of the display). To address this concern, we excluded all trials in the one-referent, ambiguous-sentence condition in which the trajectories passed over the incorrect destination before ultimately terminating at the correct destination. Even with these most extreme 5.1% of one-referent trajectories excluded, we still observed significant x-coordinate divergence between the ambiguous- and unambiguous-sentence trajectories from Timesteps 39 to 57 (all ts > 2.02, all ps $< .05$, average $d = .36$) and 63 to 82 (all ts > 2.03, all ps $< .05$, average $d = .34$), and significant y-coordinate divergence from Timesteps 39 to 55 (all ts > 2.06, all ps $< .05$, average $d = .35$) and from 67 to 79 (all ts > 2.02, all ps $< .05$, average $d = .33$).
2. As per the previous t-test analyses (see also Note 1), after excluding the extreme garden-path trials in the one-referent, ambiguous-sentence condition, we still observe a significant three-way interaction for both the x coordinates, $F(2, 78) = 5.07$, $p = .009$, $MSE = 2{,}286$; and y coordinates, $F(2, 78) = 3.44$, $p = .037$, $MSE = 1,291$. In addition, the Context $\times$ Ambiguity interaction at Segment 2 was significant for both the x coordinates, $F(1, 39) = 7.64$, $p = .009$, $MSE = 7,616$; and marginally for the y coordinates, $F(1, 39) = 3.88$, $p = .056$, $MSE = 4,987$.
3. Not all trajectories were initiated before the end of the sentence. A participant was included in the analysis if average x and y coordinates could be calculated at the time bin of interest. By Time bin 4, notably, most participants were included in the analyses

(i.e., they had initiated at least 1 trajectory in that condition during the 750–1,000 msec time bin).

4. Caution is warranted when interpreting this cutoff value. A bimodality coefficient b =.555 signals the presence of a uniform distribution whereby all values of X within the distribution have an equal probability of occurring; that is, when the distribution is rectangular, b =.555. More important, b does not operate like a p value, such that values approaching $p = .05$ are informally treated as indicating the existence of a less statistically reliable result than values much lower than $p = .05$. Instead, the value for the bimodality coefficient b, typically, must surpass $b = .555$ before one may infer the presence of *any* noteworthy bimodality.

Acknowledgments

The work presented here was supported by National Institute of Mental Health Grant R01–63961 to Michael J. Spivey and by a Dolores Zohrab Liebmann Fellowship awarded to Thomas A. Farmer.

We thank three anonymous reviewers for their constructive comments on previous versions of this manuscript.

References

Altmann, G. T. M., & Steedman, M. J. (1988). Interaction with context during human sentence processing. *Cognition, 30,* 191–238.

Atkeson, C. G., & Hollerbach, J. M. (1985). Kinematic features of unrestrained vertical arm movements. *Journal of Neuroscience, 5,*2318–2330.

Bever, T. G. (1970). The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and the growth of cognition* (pp. 279–362). New York: Wiley.

Brennan, S. E. (2005). How conversation is shaped by visual and spoken evidence. In J. Trueswell & M. Tanenhaus (Eds.), *Approaches to studying world-situated language use: Bridging the language-as-product and language-action traditions* (pp. 95–129). Cambridge, MA: MIT Press.

Britt, M. A. (1994). The interaction of referential ambiguity and argument structure in the parsing of prepositional phrases. *Journal of Memory and Language, 33,* 251–283.

Calvert, S., Rideout, V., Woolard, J., Barr, R., & Strouse, G. (2005). Age, ethnicity, and socioeconomic patterns in early computer use: A national survey. *American Behavioral Scientist, 48,* 590–607.

Chambers, C., Tanenhaus, M. K., & Magnuson, J. (2004). Actions and affordances in syntactic ambiguity resolution. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 30,* 687–696.

Cisek, P., & Kalaska, J. (2005). Neural correlates of reaching decisions in dorsal premotor cortex: Specification of multiple direction choices and final selection of action. *Neuron, 45,* 801–814.

Dale, R., Kehoe, C., & Spivey, M. J. (2007). Graded motor responses in the time course of categorizing exemplars. *Memory and Cognition, 35,* 15–28.

Darlington, R. B. (1970). Is kurtosis really "peakedness"? *American Statistician, 24,* 19–22.

DeCarlo, L. T. (1997). On the meaning of kurtosis. *Psychological Methods, 2,* 292–307.

Doyle, M., & Walker, R. (2001). Curved saccade trajectories: Voluntary and reflexive saccades curve away from irrelevant distractors. *Experimental Brain Research, 139,* 333–344.

Farmer, T. A., Cargill, S. A., & Spivey, M. J. (in press). Gradiency and visual context in syntactic garden-paths. *Journal of Memory and Language*.

Ferreira, F., & Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory and Language, 25,* 248–368.

Frazier, L., & Clifton, C. (1996). *Construal*. Cambridge, MA: MIT Press.

Gibson, E., & Pearlmutter, N. J. (2000). Distinguishing serial and parallel parsing. *Journal of Psycholinguistic Research, 29,* 231–240.

Gold, J., & Shadlen, M. (2000). Representation of a perceptual decision in developing oculomotor commands. *Nature, 404,* 390–394.

Goodale, M., Pélisson, D., & Prablanc, C. (1986). Large adjustments in visually guided reaching do not depend on vision of the hand or perception of target displacement. *Nature, 320,* 748–750.

Green, M., & Mitchell, D. (2006). Absence of real evidence against competition during syntactic ambiguity resolution. *Journal of Memory and Language, 55,* 1–17.

Hindle, D., & Rooth, M. (1993). Structural ambiguity and lexical relations. *Computational Linguistics, 19,* 103–120.

Knoeferle, P., & Crocker, M. W. (2006). The coordinated interplay of scene, utterance, and world knowledge: Evidence from eye tracking. *Cognitive Science, 30,* 481–529.

Lewis, R. L. (2000). Falsifying serial and parallel parsing models: Empirical conundrums and an overlooked paradigm. *Journal of Psycholinguistic Research, 29,* 241–248.

MacDonald, M., Pearlmutter, N., & Seidenberg, M. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review, 101,* 676–703.

Magnuson, J. S. (2005). Moving hand reveals dynamics of thought. *Proceedings of the National Academy of Sciences, 102,* 9995–9996.

McRae, K., Spivey-Knowlton, M. J., & Tanenhaus, M. K. (1998). Modeling the effects of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language, 37,* 283–312.

Port, N. L., & Wurtz, R. H. (2003). Sequential activity of simultaneously recorded neurons in the superior colliculus during curved saccades. *Journal of Neurophysiology, 90,* 1887–1903.

SAS Institute. (1989). *SAS/STAT user's guide*. Cary, NC: Author.

Sheliga, B. M., Riggio, L., & Rizzolatti, G. (1994). Orienting of attention and eye movements. *Experimental Brain Research, 98,* 507–522.

Snedeker, J., & Trueswell, J. C. (2004). The developing constraints on parsing decisions: The role of lexical-biases and referential scenes in child and adult sentence processing. *Cognitive Psychology, 49,* 238–299.

Song, J., & Nakayama, K. (2006). Role of focal attention on latencies and trajectories of visually guided manual pointing. *Journal of Vision, 6,* 982–999.

Spivey, M. J., Grosjean, M., & Knoblich, G. (2005). Continuous attraction toward phonological competitors. *Proc. National Academy of Sciences, 29,* 10393–10398.

Spivey, M. J., & Tanenhaus, M. K. (1998). Syntactic ambiguity resolution in discourse: Modeling the effects of referential context and lexical frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 1521–1543.

Spivey, M. J., Tanenhaus, M. K., Eberhard, K. M., & Sedivy, J. C. (2002). Eye movements and spoken language comprehension: Effects of syntactic context on syntactic ambiguity resolution. *Cognitive Psychology, 45,* 447–481.

Spivey-Knowlton, M. J., & Sedivy, J. (1995). Resolving attachment ambiguities with multiple constraints. *Cognition, 55,* 227–267.

Tabor, W., & Hutchins, S. (2004). Evidence for self-organized sentence processing: Digging-in effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 431–450.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268,* 1632–1634.

Tipper, S., Howard, L., & Houghton, G. (2000). Behavioral consequences of selection from population codes. In S. Monsell & J. Driver (Eds.), *Attention and performance XVIII: Control of cognitive processes* (pp. 223–245). Cambridge, MA: MIT Press.

Tipper, S., Howard, L., & Jackson, S. (1997). Selective reaching to grasp: Evidence for distractor interference effects. *Visual Cognition, 4,* 1–38.

Traxler, M. J., Pickering, M. J., & Clifton, C. (1998). Adjunct attachment is not a form of lexical ambiguity resolution. *Journal of Memory and Language, 39,* 558–592.

Trueswell, J. C., Sekerina, I., Hill, N. M., & Logrip, M. L. (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition, 73,* 89–134.

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic disambiguation. *Journal of Memory and Language, 33,* 285–318.

van Berkum, J., Brown, C., & Hagoort, P. (1999). Early referential context effects in sentence processing: Evidence from event-related brain potentials. *Journal of Memory and Language, 41,* 147–182.

van Gompel, R. P. G., Pickering, M. J., Pearson, J., & Liversedge, S. P. (2005). Evidence against competition during syntactic ambiguity resolution. *Journal of Memory and Language, 52,* 284–307.

van Gompel, R. P. G., Pickering, M. J., & Traxler, M. J. (2001). Reanalysis in sentence processing: Evidence against constraint-based and two-stage models. *Journal of Memory and Language, 45,* 225–258.