Continuous attraction toward phonological competitors

Michael J. Spivey*†, Marc Grosjean†‡§, and Günther Knoblich†¶

*Department of Psychology, Cornell University, Ithaca, NY 14853; ‡Department of Psychology, Max Planck Institute for Human Cognitive and Brain Sciences, Amalienstrasse 33, Munich 80799 Germany; and ¶Department of Psychology, Rutgers, The State University of New Jersey, Newark, NJ 07102

Communicated by James L. McClelland, Carnegie Mellon University, Pittsburgh, PA, May 10, 2005 (received for review January 28, 2005)

Certain models of spoken-language processing, like those for many other perceptual and cognitive processes, posit continuous uptake of sensory input and dynamic competition between simultaneously active representations. Here, we provide compelling evidence for this continuity assumption by using a continuous response, hand movements, to track the temporal dynamics of lexical activations during real-time spoken-word recognition in a visual context. By recording the streaming *x*, *y* coordinates of continuous goal-directed hand movement in a spoken-language task, online accrual of acoustic–phonetic input and competition between partially active lexical representations are revealed in the shape of the movement trajectories. This hand-movement paradigm allows one to project the internal processing of spoken-word recognition onto a two-dimensional layout of continuous motor output, providing a concrete visualization of the attractor dynamics involved in language processing.

dynamical systems | psycholinguistics | word recognition

Modular stage-based accounts of language processing have generally assumed that, rather than continuously cascading partial results of information processing to later stages (1, 2), the neural subsystems responsible for perception and cognition each wait until a stable unique representation has been computed before that information is passed on to the next processing stage (3–5). This kind of discrete stage-based theoretical framework has motivated much of cognitive psychology since its inception, and it continues to be a guiding force in various contemporary theories of cognitive processing. However, in the case of spoken-word recognition, a number of judgment-based experimental techniques have provided indirect evidence for partial activation of multiple lexical representations ("cohorts") cascading to later stages of processing even just part of the way through hearing a word (6, 7). Moreover, recent eye-movement data have supported a continuously dynamic and highly interactive account of the real-time integration of information sources during spoken-language processing (8, 9). For example, eye movements to objects with phonologically similar names (e.g., saccades to a *cand*le when instructed to "pick up the *candy*") have been interpreted as evidence for continuous processing of phonological input and parallel activation of temporarily consistent lexical representations in monolingual adults (10, 11), bilingual adults (12), and children (13). Thus, it appears that neural patterns corresponding to multiple lexical representations may signal later stages of processing before the single correct lexical item is identified. However, it is still not entirely clear whether the activations of these lexical representations are updated continuously by the acoustic–phonetic input and constantly cascaded to later stages or whether there are intermediate noncascading stages in spoken-language comprehension.

Completely ruling out discrete-time incremental versions of the modular stage-based account (4) has proven to be difficult because there is, in principle, the possibility that the apparent continuity in recent results may be an artifact of averaging discontinuous or semicontinuous motor outputs (such as button presses and saccades). In this work, we recorded a continuous hand-movement response during comprehension of spoken instructions in a visual context, and we show that it provides an unusually high-fidelity emission of the continuous cognitive dynamics inherent in real-time spoken-language processing.

Several computational models of spoken-word recognition assume relatively continuous input and parallel partial activation of lexical representations (10, 14–17). Corresponding to simulation results from the interactive-activation TRACE model of speech processing (14), the eye-movement data typically show a nonlinear rising curve over time for the probability of fixating the target object (referred to in the speech stream; e.g., "beaker"), and a significant rising-then-falling curve for the probability of fixating an object whose name has phonological overlap with the spoken word (e.g., a *bee*tle, or a sp*eaker*) (18).

The semicontinuous record of eye position, alternating between steady fixations of 300–400 ms and fast, ballistic saccades of 20–40 ms, is a significant improvement over traditional outcome-based experimental methods that record only accuracy and reaction time at the end of a trial. Nonetheless, a disadvantage of the eye-movement evidence for parallel partial activation of lexical alternatives during spoken-word recognition is that it involves averaging "categorical" data (steady fixations of one object or another over time) to produce "continuous" functions. Thus, it can only approximate continuous central tendencies of group data.

Because saccades are largely ballistic (but cf. ref. 19), the experimental trials that contribute to evidence that the cohort lexical item is substantially active are always trials in which the participant briefly fixated directly on the cohort object at some point in the trial and then later fixated the target object before picking it up. In contrast to saccades, many arm movements are nonballistic and can often be smoothly redirected midflight (20). Therefore, by recording continuous arm movements in a similar visual display, one can observe graded effects of a competing object pulling the movement in its direction even on trials in which the hand only ever settles on the correct target object.

Experiment

Methods. Forty-two Cornell University undergraduates participated in the experiment for extra credit in psychology courses. Participants were presented with color images of two objects on a screen (one target and one distractor), and a prerecorded speech file instructed them to click one of them with the mouse. Objects were presented in the upper left and upper right corners of the computer screen (e.g., a *candle* and a *candy*, in the cohort condition, or a *candle* and a *jacket*, in the control condition). Eight target objects were used to make 32 trials in which the distractor object was either a cohort for the target object or a

See Commentary on page 9995.

†To whom correspondence may be addressed. E-mail: spivey@cornell.edu, grosjean@cbs.mpg.de, or knoblich@psychology.rutgers.edu.

§Present address: Institute for Occupational Physiology, University of Dortmund, Ardeystrasse 67, 44139 Dortmund, Germany.

phonologically dissimilar control and in which the target object was either on the left or right side of the display. Participants were instructed to mouse-click a box in the bottom center of the screen when they were ready to begin a trial. At this time, the two object images would appear in the upper left and right, and 500 ms after the onset of the images, a single spoken word (from a speech file on the computer; mean duration, 532 ms) would name the target object. [Imposing this asynchrony between image onset and speech onset grew out of observations from pilot studies in which simultaneous onset caused participants to occasionally wait until the entire word was spoken before beginning their mouse movement. With the spoken word beginning 500 ms after onset of the images, participants usually begin their mouse movement (straight upward) before the onset of the spoken word, which gives distinguishing properties in the acoustic–phonetic input a chance to influence the continuous motor output midflight.]

The x, y screen coordinates of the computer mouse were sampled at ≈36 Hz. At this sampling rate, each trial typically collected ≈30–60 data points indicating where in space the motor output had taken the hand in response to the speech input. Although eye trackers actually sample eye position at 60, 240, or even 1,000 Hz, most eye-movement analyses involve a fixation analysis that, quite sensibly, pools samples during steady fixations, resulting in two to four fixations per s (8, 9). Thus, although saccades are generally initiated earlier than skeletal movements, giving eye-tracking a clear advantage in immediacy, acquiring 30–60 data points per trial from nonballistic hand movements provides a substantially richer signal of continuous processing than two to four data points per trial. Moreover, the statistical distributions of these x, y coordinates at any given time slice tend to be narrower, as well as more normal (as described below), than the statistical distributions of categorical object-fixation data from eye movements.

Results. In the cohort condition (in which the name of the distractor object shared phonological overlap with the spoken target word), participants clicked on the incorrect image on 8.9% of the trials, which was significantly greater ($P < 0.001$) than the 2.5% error rate in the control condition (in which the name of the distractor object did not share phonological overlap with the spoken target word). During debriefing, participants generally reported that when they made an error they attempted to correct it by rapidly clicking the target object. These error trials were excluded from further data analysis.

Not surprisingly, the cohort condition showed longer total response times (from visual onset to correct mouse click) than the control condition (1,812 vs. 1,717 ms; $P < 0.001$). There was no main effect of condition in time to initiate a movement (335 vs. 357 ms; $P > 0.1$), because participants typically initiated the movement before speech onset. This preemptive movement allowed the speech input to influence motor output while the hand was in motion. A significant main effect of condition in the duration of the movement itself (1,477 vs. 1,360 ms; $P < 0.001$) provided evidence that the effects of phonological similarity were indeed taking place while the hand was moving. Moreover, average movement length on the computer screen was longer in the cohort condition than in the control condition (30.36 vs. 26.68 cm; $P < 0.001$).

Fig. 1 shows averaged mouse-movement trajectories, with durations normalized to 101 time slices (0–100%). Before averaging, all trajectories were lined up to a common x, y starting position (0, 0). They were then individually normalized by resampling the time vector at 101 equally time-spaced values and computing, by means of linear interpolation, the corresponding mouse-coordinate values (separately for the x and y coordinate vectors). Last, pixel coordinates were converted into centimeters on the computer screen. Note how the average movement

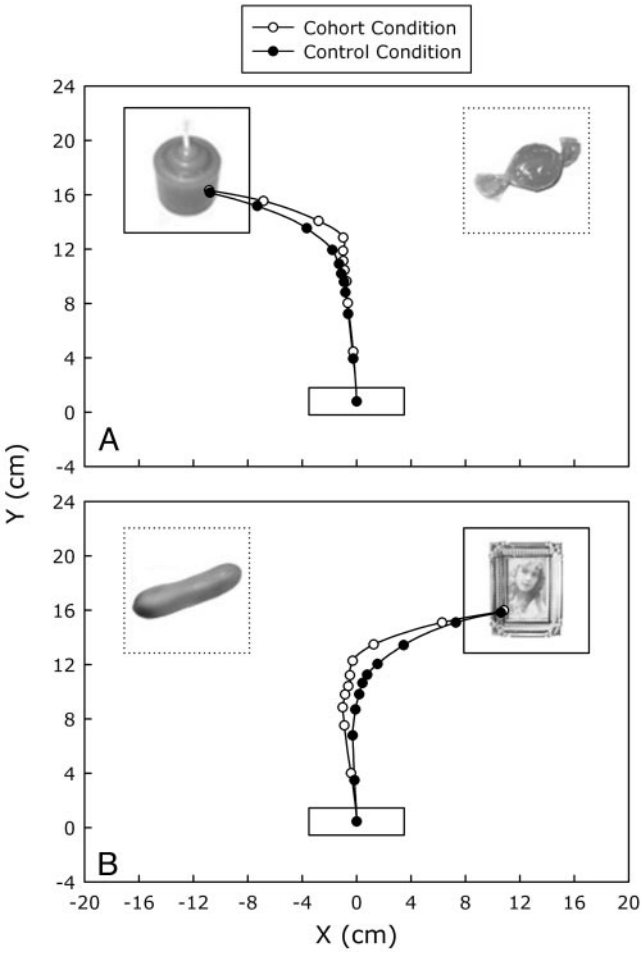

Fig. 1. Mean trajectories for the control (●) and cohort (○) conditions when the target was on the left (*A*) and right (*B*) side. Both object images depict cohort conditions. (*A*) ''Candle'' with a candy distractor. (*B*) ''Picture'' with a pickle distractor. The rectangles surrounding the objects indicate regions in which a mouse click would end the trial. Symbols are plotted every 10 normalized time slices (e.g., at the 10th, 20th, etc., percentile of the mean movement duration) for this and other figures.

trajectory in the cohort condition travels further upward and equidistant from the two objects for a longer period than the trajectory in the control condition. This gravitation of the mouse movement to a region in between the two objects in the cohort condition is remarkably consistent with a description of two nearly equibiased attractors initially pulling the system toward their shared midpoint, compared with one strong attractor (control condition) pulling the system toward its own midpoint (21–23).

In mean trajectories, paired time slices for leftward cohort- and control-condition trajectories (Fig. 1*A*) showed statistically significant differences in the x coordinate ($P < 0.05$) contiguously from the 70th to the 81st time slice, indicating a sustained spatial attraction of the mouse cursor in the direction of the cohort distractor object. A similar comparison for the rightward trajectories (Fig. 1*B*) showed statistically significant differences contiguously from the ninth to the 92nd time slice. (The more robust effect for rightward movements may be due to kinematics of the right arm and positioning of the mouse to the right of the keyboard.)

As done in eye-tracking studies that plot rising and falling curves of probability-of-fixation for target and distractor objects (10, 18), the proximity of the mouse cursor to the target and

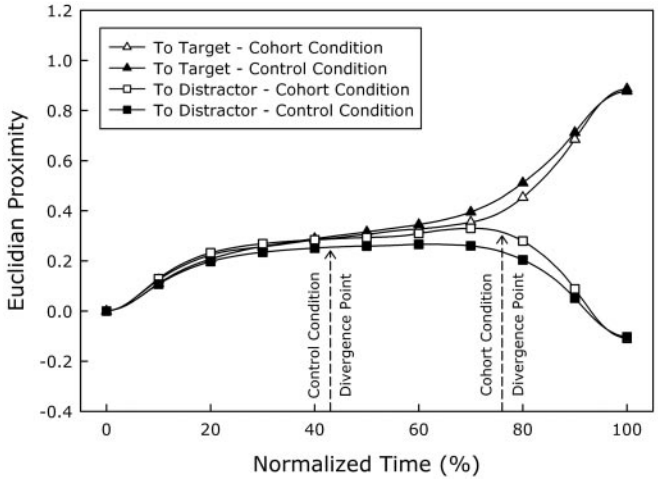

Fig. 2. Proximity of the mouse cursor to target (triangles) and to distractor (squares) over normalized time in the control condition (filled symbols) and in the cohort condition (open symbols). Averaged movement trajectories exhibit a significantly greater proximity to target than to distractor much earlier in the control condition than in the cohort condition.

distractor objects over time can be similarly treated as an indicator of the activation of the competing lexical representations. Fig. 2 shows the proportional Euclidean proximity, 1 − distance/max(distance), to the centers of the target object and the distractor object over normalized time (averaged across leftward and rightward movements). During early portions of the movement trajectory, proximity to the target and distractor are not significantly different from one another, because the movement is largely in the vertical dimension. In the control condition, these two proximities diverge significantly from one another at the 43rd normalized time slice, and they continue to be significantly diverged for the remaining 57% of the movement duration. In the cohort condition, they do not significantly diverge until the 76th normalized time slice. Moreover, showing that the cohort object attracted the movement toward itself, proximity to the distractor in the cohort condition was significantly greater than proximity to the distractor in the control condition ($P < 0.05$) all of the way from the fourth to the 93rd normalized time slices. As evidence that the activation of the spoken lexical item may reach asymptote less quickly when an object with a phonologically competing name is visually present, proximity to the target object in the cohort condition was significantly lower than proximity to the target object in the control condition from the 66th to the 91st normalized time slices.

Recall that mouse movement was initiated, on average, at ≈345 ms after visual onset of the objects, and speech onset was exactly 500 ms after visual onset of the objects. Therefore, by the time the speech input began, participants were, on average, already ≈155 ms into their ≈1,400-ms mouse-movement trajectory. Therefore, the fact that the divergence between distractor proximity in the cohort and control conditions is observed as early as the fourth normalized time step (Fig. 2) indicates that in a proportion of trials in which movement initiation was later than average, there was sufficient distinguishing information in the spoken input to reliably sway the trajectory almost immediately.

To cement the claim that movement trajectories in the cohort condition are all statistically deflected toward the competitor object in a continuously graded fashion, it is necessary to look at the distribution of these deflections across multiple trials. In principle, it could be the case that, as with saccadic eye movements, there are some trials in which the competitor object does not attract the motor output and other trials in which it does. Such a bimodal distribution could be consistent with a stage-based account of spoken-word recognition in which an incorrect interpretation of the spoken word is occasionally briefly instated in discrete symbolic form (thus triggering a motor output toward that competitor object), and then quickly replaced by the correct symbolic lexical form (thus triggering a corrective movement toward the object being referred to). When averaged, this hypothetical data pattern would produce mean movement trajectories that could falsely suggest simultaneous partial activation and competition among multiple lexical representations.

To examine this possibility, we calculated the degree of curvature (toward the distractor object) among the trajectories in the cohort and control trials in terms of the area (in pixels) between each actual trajectory and a straight line connecting its start and endpoint. (Portions of curvature away from the distractor object and away from the straight line resulted in negative area calculations.) With too few trials within a participant to provide an adequate measure of the unimodality or bimodality of the distribution of these trajectory deflections, the values for both cohort and control conditions were together converted into *z* scores within a participant and then pooled across participants. Fig. 3*A* shows the *z* distribution for the cohort trials ($n = 611$; mean, 0.164; variance, 1.043; kurtosis, 0.76; skewness, 0.658), looking quite similar to that of the control trials ($n = 606$; mean, −0.165; variance, 0.889; kurtosis, 1.75; skewness, 0.861). For the cohort *z* distribution, the bimodality coefficient (*b*) was 0.381, and for the control *z* distribution, it was 0.366 (with $b > 0.555$ being the standard cutoff for multimodality). Note that if continuous eye movement scan paths sampled at 60+ Hz (instead of fixation analyses) were subjected to corresponding curvature analyses, there would be a decidedly bimodal pattern in the distribution. Participants in those studies either fixate the competitor object or they do not, on any given trial. They do not make saccades slightly toward the competitor object the way these mouse-movement trajectories show deflections slightly toward the competitor object.

In a Kolmogorov–Smirnov test of normality, the cohort *z* distribution was not significantly different from a normal distribution with the same mean and variance, but high kurtosis (indicating an unusually high proportion of trials near the mean) did make the test marginally significant ($P = 0.054$). In such a test with the control *z* distribution, the even higher kurtosis caused it to be significantly different from a normal distribution with the same mean and variance ($P = 0.023$). In both cases, the deviation from normality is due to high kurtosis, meaning that the distributions are even more sharply singly peaked and further from bimodality than their corresponding normal distributions with matched mean and variance.

We also *z*-scored, within each participant, the area under the trajectory separately for cohort and control trials, and we then pooled across subjects (see Fig. 3*B*). With these two *z* distributions having the same mean (of zero) and the same variance (of 0.966), the Kolmogorov–Smirnov test can evaluate the difference between their respective shapes (e.g., skewness, kurtosis, and multimodality). With no theoretical reason to imbue bimodality in the control *z* distribution, quantitative evidence for high similarity between the control and cohort *z* distributions would substantially allay concerns that some hidden bimodal behavior exists in the cohort condition. When comparing these two *z* distributions, the Kolmogorov–Smirnov test produces a *P* value > 0.9999.

In sum, three tests cast considerable doubt on the possibility of the cohort condition being composed of (*i*) some trials that behave like control trials (indicating no competing lexical activation of the cohort item) and (*ii*) some trials that exhibit uniquely curved trajectories (consistent with a discrete tempo-

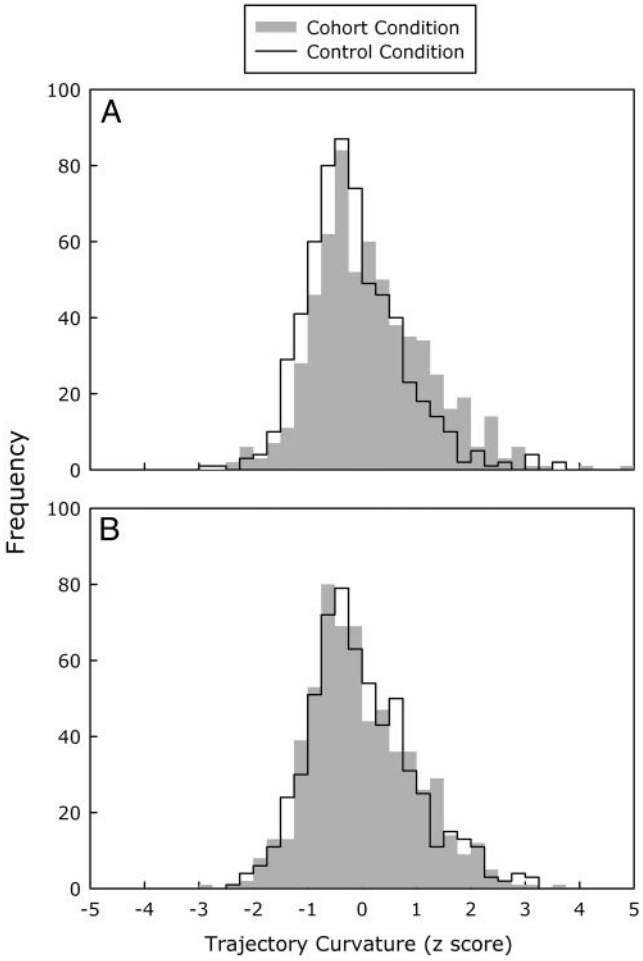

Fig. 3. Distributions of trajectory curvature show no evidence of bimodality. (*A*) Overlaid histograms of trajectory curvature in the cohort and control conditions (*z* scored together within subject) as measured by the area between each trajectory and a straight line connecting its start and endpoint, exhibit unimodal distributions. (*B*) When the cohort and control conditions are *z* scored separately (within-subject), the shapes of their distributions are statistically indistinguishable. (See text for details.)

rary misinterpretation of the spoken word as referring to the cohort object). The bimodality coefficients rule out bimodality in the distributions of trajectory curvatures. The Kolmogorov–Smirnov tests of normality indicate borderline-significant deviation from normality in the direction toward unusual "peakedness" and away from bimodality, for both *z* distributions. Last, the Kolmogorov–Smirnov comparison between equal-mean and equal-variance cohort and control *z* distributions finds absolutely no reason to treat their shapes differently. The results of these analyses are consistent with the interpretation that each mouse movement in the cohort condition is being partially attracted toward the cohort object because the lexical representation associated with that object is simultaneously partially active and competing with the correct lexical representation.

Simulation

Methods. To map a linking hypothesis between hypothesized partially active lexical representations in the brain and the observed hand-movement trajectories across a mouse pad, a computational simulation of the results was implemented by interfacing the TRACE model of spoken-word recognition (14) by using a lexicon of 14 words corresponding to the objects used in the experiment, with a normalized-recurrence attractor network (24). If we used TRACE alone, we would be forced to summarily turn off lexical nodes in TRACE that correspond to objects that are not present in a given visual display, to prevent the simulation results from predicting movements toward objects that are not there. This manipulation would be tantamount to making the unrealistic claim that lexical activation is discretely constrained by visual context, such that if a candle were not in the visual field, then hearing "candy" would cause no activation whatsoever of the lexical representation for "candle." However, because one of the vectors in the attached normalized recurrence network represents objects in the visual scene, this graded constraint allows us to feed all of the lexical activations of TRACE over time into the normalized recurrence network and allow the gradual competitive process of that network to induce motor movement only toward objects that are actually present and actionable.

Activations of lexical nodes over time in TRACE were fed gradually into the lexical vector of a localist attractor network composed of 14 lexical nodes, 14 visual nodes, and 14 integration nodes, where the lexical vector (L) and the visual vector (V) are each normalized so that their n nodes sum to 1.0 before being averaged at the integration vector. In the normalized-recurrence competition algorithm, cumulative feedback to a lexical or visual node is based on multiplying the current activation of that node by the activation of its corresponding integration node (which is itself simply an average of the corresponding visual and lexical nodes) and adding this product to the activation of that lexical or visual node. At the next time step, t, the lexical and visual vectors are normalized again, and the cycle continues.

$$\mathrm{L}_{n,t}(\mathrm{norm}) = (\mathrm{L}_{n,t} + \mathrm{TRACE}_{n,t}) \Big/ \sum_{n} (\mathrm{L}_{n,t} + \mathrm{TRACE}_{n,t})$$

$$\mathrm{V}_{n,t}(\mathrm{norm}) = \mathrm{V}_{n,t} \Big/ \sum_{n} \mathrm{V}_{n,t}$$

$$\mathrm{L}_{n,t+1} = \mathrm{L}_{n,t}(\mathrm{norm}) + \mathrm{L}_{n,t}(\mathrm{norm}) \cdot \left(\frac{\mathrm{L}_{n,t}(\mathrm{norm}) + \mathrm{V}_{n,t}(\mathrm{norm})}{2} \right) \quad [1]$$

$$\mathrm{V}_{n,t+1} = \mathrm{V}_{n,t}(\mathrm{norm}) + \mathrm{V}_{n,t}(\mathrm{norm}) \cdot \left(\frac{\mathrm{L}_{n,t}(\mathrm{norm}) + \mathrm{V}_{n,t}(\mathrm{norm})}{2} \right)$$

Because the feedback is multiplicative in normalized recurrence, only the visual nodes that start out nonzero (because their corresponding objects are present in the display) will exhibit activation. For a simplified linking hypothesis of lexical activation patterns cascading to visual salience patterns that cascade to motor output systems, the visual nodes were stochastically sampled at each time step, based on their probabilistic activation, to produce simulated changes in x, y mouse position.

The x and y increments in simulated mouse position were calculated such that the mouse moved closer to the object corresponding to that visual node. Thus, if the two active visual nodes both elicited movements on a particular time step, then the simulated mouse-position change would exhibit two y-axis increments and zero x-axis increments (because the two opposite horizontal movements would cancel each other out). In this kinematically simplified simulation, the base increment for horizontal (left and/or right) movements was 0.4 cm, and the base increment for vertical (upward) movements was 2 cm. To better mimic the temporal dynamics of the overall movement, these x and y increments were scaled by two factors. Multiplying each x and y change in position by an inverted Gaussian velocity

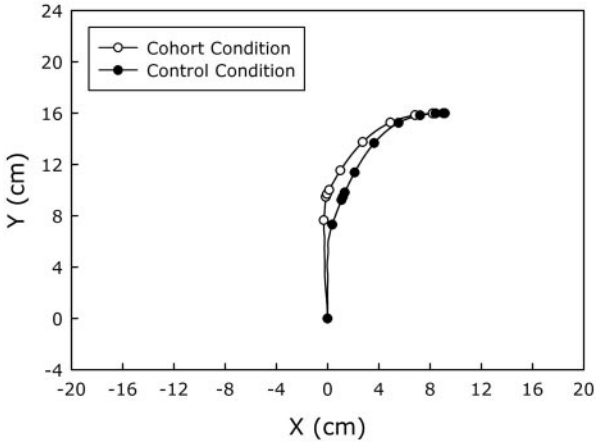

Fig. 4. Simulated movement trajectories resulting from interfacing the TRACE model of spoken-word recognition with a localist attractor network and stochastically sampling *x* and *y* mouse movements from its visual object nodes. When the cohort distractor object is present (○), the simulated movement trajectory gravitates to the region in between the two objects for longer than when the control distractor object is present (●).

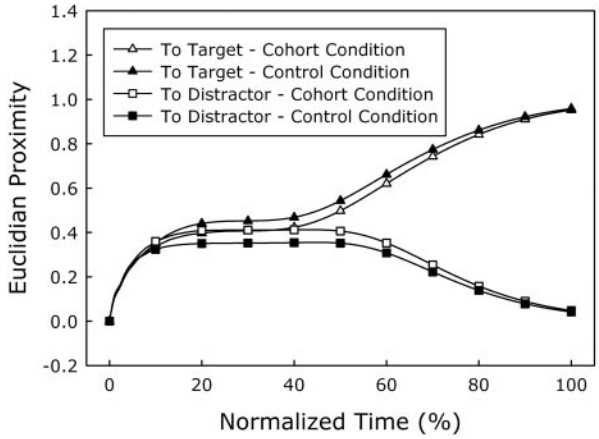

Fig. 5. Proximity to target (triangles) and to distractor (squares) over normalized time in the control condition (filled symbols) and the cohort condition (open symbols) from the simulated mouse movements triggered by graded lexical activation continuously spreading to visual object nodes.

envelope over time, which approximated the observed U-shaped velocity profile in the human data, produced the slow down that occurs in both conditions around the middle of the movement.

$$\Delta\text{pos} = \Delta\text{pos}\cdot\left(1 - \frac{125}{50\sqrt{2\pi}}\, e^{-\frac{(t-50)^2}{2\cdot 50^2}}\right) \qquad [2]$$

Also, multiplying each *x* and *y* increment by the inverse proportion of the current distance from the *x*, *y* position of the goal in cm, $1 - x/10$, and $1 - y/16$, produced the slow down that occurs in both conditions as the goal is reached (see, for example, the smooth final approach to asymptote in Fig. 2). Last, for comparison with the human data, the 170 time steps of *x*, *y* coordinates of the model were normalized to 101 time slices.

Results. Fig. 4 plots the mean *x*, *y* movement trajectories (in cm) from 10 runs of the model in the cohort condition and the control condition. Much as in the human data (Fig. 1), when the lexical representation of the distractor object is initially accruing partial activation at a similar rate to the lexical representation of the target object (cohort condition), the simulated movement trajectory continues upward between, and equidistant from, the two objects for a longer period than when the name of the distractor object exhibits no phonological similarity to the spoken word (control condition).

When proximity to target and distractor over time is plotted from these simulated changes in *x*, *y* position (Fig. 5), the resulting pattern bears considerable similarity to the human data (Fig. 2), although the current simulation does exhibit divergence somewhat earlier than the human data. Comparing the four curves at every normalized time slice in this image to the four curves at every time slice in Fig. 2 produces a root-mean-squared error of 0.0625, and $r^2 = 0.76$ ($P < 0.0001$).

As for variability in trajectory curvature across multiple runs of the simulation, the distributions are, not surprisingly, highly normal. When the area between each trajectory and its straight line was calculated for 600 runs of the model in each condition and the data were *z* scored together, the *z* distribution for the cohort trials (mean, 0.114; variance, 0.997; kurtosis, −0.520; skewness, −0.027) was quite similar to that of the control trials (mean, −0.114; variance, 0.997; kurtosis, −0.488; skewness, 0.129). The distributions are shown in Fig. 6, bearing some resemblance to Fig. 3*A*. Their bimodality coefficients were $b = 0.403$ and $b = 0.404$, respectively. Rather than exhibiting the unusually sharp peak near the mean that was seen in the human data, the distributions of the simulation conformed quite closely to a normal distribution. In the Kolmogorov–Smirnov normality test, both cohort and control *z* distributions did not remotely differ from their corresponding (matched mean and variance) normal distributions (both $P > 0.9999$). When this normality test was used to compare the cohort *z* distribution of the simulation with the cohort *z* distribution of the human data (because their means and variances are quite similar), the difference was not significant ($P = 0.142$). However, when the control *z* distribution of the simulation was compared with the human data control *z* distribution, they were significantly different ($P = 0.007$). The notable differences in kurtosis and skewness between the distributions of the human data and the simulation results remain to be examined.

General Discussion

These results provide powerful support for models of continuous uptake of acoustic–phonetic input during spoken-word recognition. The substantial fit between model simulation and human data provides an encouraging, if simplified, linking hypothesis to

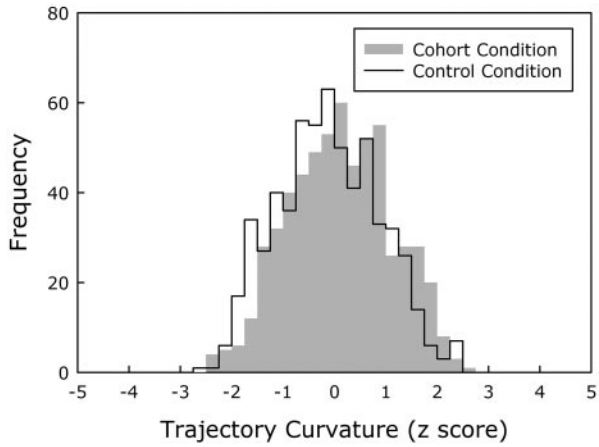

Fig. 6. Overlaid histograms of the trajectory curvature of the simulation in the cohort and control conditions (*z* scored together over all 1,200 runs) reveal normal unimodal distributions with slightly lower kurtosis and skewness than those from the human data shown in Fig. 3*A*.

support the claim that continuous temporal dynamics of lexical activations in the brain are being reflected in the continuous temporal dynamics of motor output. During and soon after the presentation of a spoken word, temporary ambiguity in the reference of the speech stream to visible objects produced competition between two motor output goals, which manifested itself as a graded spatial attraction toward the competing object even when the movement eventually settled into the correct object region.

This tightly coupled relationship among language, vision, and action is seen in other areas as well, such as signed languages (where experience with American Sign Language affects perception of nonlinguistic motor movements; ref. 25), in the following of spoken instructions (where perceived motor affordances have an immediate influence on comprehension; ref. 26), and even in the coupled postural sway of two speakers conversing (27). The present findings demonstrate that the continuous processing of a spoken word is observable in the continuous execution of motor output, consistent with a nonstop cascaded sharing of information among perception, cognition, and action (28, 29).

However, note that it would be a mistake to interpret these mouse-movement data as evidence that nonballistic mouse movements are generally more informative for perception and cognition than ballistic eye movements. The two methodologies have compensatory strengths and weaknesses. Saccades are highly temporally sensitive to the existence of early partially active representations, whereas these mouse movements are highly spatially sensitive to continuous ongoing competition between partially active representations. Combining the two measures at the same time would be useful.

Following from other measures of dynamic motor output revealing temporally continuous perceptual-motor processes (2, 30, 31), our present findings do more than contribute to evidence for a cascaded flow of information from perceptual, to cognitive, to motor systems. Our present findings virtually project the ongoing output of the language comprehension process onto a two-dimensional action space in which the potential goal objects act like attractor points and the manual movement serves as a record of the mental trajectory traversed as a result of the continuously updated interpretation of the linguistic input (21–23). This experimental paradigm promises to facilitate explorations of continuous temporal dynamics in many aspects of real-time language comprehension, categorization, visual search, and other aspects of cognition in general.

We thank Jay McClelland, Mike Tanenhaus, Jim Magnuson, Bob McMurray, Rick Dale, and especially Peter C. Gordon for helpful comments on earlier drafts; Paul Allopenna for providing the lexical activations from TRACE; Jillian Caly and Deborah Birnbaum for assistance with software and data collection; and Arturo Galvan and Dick Darlington for help with data analysis. The Cornell University Committee on Human Subjects approved the experimental protocol. This work was supported by National Institute of Mental Health Grant R01-63961 (to M.J.S.) and the Max Planck Institute for Human Cognitive and Brain Sciences Department of Psychology (G.K.).

1. McClelland, J. L. (1979) *Psychol. Rev.* **86,** 287–330.
2. Coles, M. G. H., Gratton, G., Bashore, T. R., Eriksen, C. W. & Donchin, E. (1985) *J. Exp. Psychol. Hum. Percept. Perform.* **11,** 529–553.
3. Fodor, J. A. (1983) *The Modularity of Mind* (MIT Press, Cambridge, MA).
4. Forster, K. (1979) in *Sentence Processing: Psycholinguistic Studies Presented to Merrill Garrett*, eds. Cooper, W. & Walker, E. (Erlbaum, Hillsdale, NJ).
5. Frazier, L. & Clifton, C. (1996) *Construal* (MIT Press, Cambridge, MA).
6. Grosjean, F. (1980) *Percept. Psychophys.* **28,** 267–283.
7. Marslen-Wilson, W. D. (1987) *Cognition* **25,** 71–102.
8. Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M. & Sedivy, J. C. (1995) *Science* **268,** 1632–1634.
9. Altmann, G. T. M. & Kamide, Y. (1999) *Cognition* **73,** 247–264.
10. Magnuson, J. S., Tanenhaus, M. K., Aslin, R. N. & Dahan, D. (2003) *J. Exp. Psychol.* **132,** 202–227.
11. Spivey-Knowlton, M. J., Tanenhaus, M. K., Eberhard, K. M. & Sedivy, J. C. (1998) in *Representation and Processing of Spatial Expressions*, eds. Olivier, P. & Gapp, K. (Erlbaum, Mahwah, NJ), pp. 201–214.
12. Spivey, M. J. & Marian, V. (1999) *Psychol. Sci.* **10,** 281–284.
13. Fernald, A., Pinto, J. P., Swingley, D., Weinberg, A. & McRoberts, G. W. (1998) *Psychol. Sci.* **9,** 228–231.
14. Gaskell, M. G. & Marslen-Wilson, W. D. (2002) *Cogn. Psychol.* **45,** 220–266.
15. Grossberg, S. & Myers, C. W. (2000) *Psychol. Rev.* **107,** 735–767.
16. Luce, P. A., Goldinger, S. D., Auer, E. T. & Vitevitch, M. S. (2000) *Percept. Psychophys.* **62,** 615–625.
17. McClelland, J. L. & Elman, J. L. (1986) *Cogn. Psychol.* **18,** 1–86.
18. Allopenna, P. D., Magnuson, J. S. & Tanenhaus, M. K. (1998) *J. Mem. Lang.* **38,** 419–439.
19. Doyle, M. & Walker, R. (2001) *Exp. Brain Res.* **139,** 333–344.
20. Goodale, M. A., Pélisson, D. & Prablanc, C. (1986) *Nature* **320,** 748–750.
21. Elman, J. L. (1995) in *Mind as Motion: Explorations in the Dynamics of Cognition*, eds. Port, R. F. & van Gelder, T. (MIT Press, Cambridge, MA), pp. 195–225.
22. Hinton, G. E. & Shallice, T. (1991) *Psychol. Rev.* **98,** 74–95.
23. McRae, K., DeSa, V. R. & Seidenberg, M. S. (1997) *J. Exp. Psychol.* **126,** 99–130.
24. Spivey, M. J. & Tanenhaus, M. K. (1998) *J. Exp. Psychol. Learn. Mem. Cognit.* **24,** 1521–1543.
25. Poizner, H. (1981) *Science* **212,** 691–693.
26. Chambers, C. G., Tanenhaus, M. K. & Magnuson, J. S. (2004) *J. Exp. Psychol. Learn. Mem. Cognit.* **30,** 687–696.
27. Shockley, K., Santana, M. V. & Fowler, C. A. (2003) *J. Exp. Psychol. Hum. Percept. Perform.* **29,** 326–332.
28. Gold, J. I. & Shadlen, M. N. (2001) *Trends Cogn. Sci.* **5,** 10–16.
29. Shin, J. C. & Rosenbaum, D. A. (2002) *J. Exp. Psychol.* **13,** 206–219.
30. Abrams, R. A. & Balota, D. A. (1991) *Psychol. Sci.* **2,** 153–157.
31. Brennan, S. (2004) in *Approaches to Studying World-Situated Language Use: Bridging the Language-as-Product and Language-as-Action Traditions*, eds. Trueswell, J. & Tanenhaus, M. (MIT Press, Cambridge, MA), pp. 95–129.