

Lévy-like diffusion in eye movements during spoken-language comprehension

Damian G. Stephen,¹ Daniel Mirman,^{2,3} James S. Magnuson,^{1,3} and James A. Dixon^{1,3}

¹*Department of Psychology, University of Connecticut, 406 Babbidge Road, Unit 1020, Storrs, Connecticut 06269-1020, USA*

²*Moss Rehabilitation Research Institute, 1200 West Tabor Road, Philadelphia, Pennsylvania 19141, USA*

³*Haskins Laboratories, 300 George Street, New Haven, Connecticut 06511, USA*

(Received 3 November 2008; revised manuscript received 5 February 2009; published 27 May 2009)

This study explores the diffusive properties of human eye movements during a language comprehension task. In this task, adults are given auditory instructions to locate named objects on a computer screen. Although it has been convention to model visual search as standard Brownian diffusion, we find evidence that eye movements are hyperdiffusive. Specifically, we use comparisons of maximum-likelihood fit as well as standard deviation analysis and diffusion entropy analysis to show that visual search during language comprehension exhibits Lévy-like rather than Gaussian diffusion.

DOI: 10.1103/PhysRevE.79.056114

PACS number(s): 89.75.Da, 05.40.Fb, 05.45.Tp, 87.19.lv

I. DIFFUSIVITY OF COGNITIVE BEHAVIORS

Human behavior in cognitive tasks is a dynamical process that evolves over time. One fundamental problem for cognitive science is characterizing the dynamics of this behavior. The variability of behavior in language processing and visual search has motivated numerous accounts of these cognitive phenomena in terms of diffusion [1–5]. Among these accounts, a major divide has arisen over the class of diffusion characterizing behavioral variability. This divide bears on the distinction between standard diffusion and hyperdiffusion. The traditional approach to cognitive science had modeled language comprehension and visual search as ordinary diffusion, that is, ordinary Brownian motion (Bm) [1,2]. Under this approach, cognition is a scale-dependent short-memory diffusion process that propagates as a linear function of time or stimulus-set size [6,7]. More recent research has found that the variability of cognitive behaviors is temporally correlated [3–5,8,9], and these temporal correlations have been interpreted as fractional Gaussian noise (fGn). Evidence of fGn would indicate that cognition is instead a scale-invariant fractal process [4,5,8,9]. Importantly, fGn is just one example of a broader class of *hyperdiffusive* processes that differ from ordinary Brownian diffusion in that they propagate as a nonlinear function of time [10].

The present research deals with hyperdiffusion in the context of human looking behavior during a cognitive task. Visual search has long been an important part of many standard cognitive tasks. For example, previous research in attention and feature integration drew on tasks asking participants to find a target stimulus among several distractor stimuli (e.g., a green “N” among brown “Ns” and green “Xs”) [11]. Visual search was used only as a cognitive exercise whose reaction times indicated greater or lesser computational loads when arriving at a response. Reaction times in such tasks increase as a linear function of stimulus-set size (i.e., the number of targets and distractors present during a trial) [6]. However, it need not be the case that responses of the same reaction time tap into the same cognitive processes. Cognitive scientists have lately begun to consider the exact trajectory of cognitive processing within the time it takes to formulate a response. For this purpose, studying eye movements has

proven to be a helpful strategy for discerning the dynamics of cognition [12], with the visual world paradigm (VWP) [13] playing an important role (see below). Particularly relevant to a discussion of hyperdiffusion is the discovery of temporal correlations in eye movements during cognitive tasks [5,9]. This finding has so far been taken that visual search is a fGn diffusion process.

Cognitive science has lately begun to consider another variety of hyperdiffusion, namely, Lévy diffusion. Often found in animal foraging behavior [14], Lévy-like processes have been found in the cognitive function of searching semantic memory in free recall [15]. Lévy diffusion is a departure from Gaussian statistics that may drive the strange kinetics of chaotic systems [16]. Although it is also a scale-invariant fractal process, Lévy diffusion is notable in that it can be either temporally correlated or temporally uncorrelated [17]. The presence or absence of temporal correlations is neither exhaustive nor conclusive evidence for or against hyperdiffusion, respectively [18]. Therefore, it is important to revisit the problem of visual search and language processing with a view toward distinguishing Lévy and fGn diffusion processes. In this paper, we examine human eye movements during a spoken-language processing task and investigate their diffusive properties using relative likelihood estimation (e.g., [14]) and multiscaling comparative analysis (MSCA) (see [18]).

II. TASK: VISUAL WORLD PARADIGM

We carried out two experiments in the VWP [13]. The VWP was proven to be an effective means of examining the dynamics of language processing as it unfolds in its naturally multisensory context. Because much of language refers to an information-rich visual environment [12], the VWP exploits the fact that people make anticipatory eye movements to named objects, even before the object’s name has been fully pronounced. In a standard trial, a participant wearing headphones sits in front of a computer screen displaying images. Spoken instructions to interact with displayed items (e.g., “click on the beaker”) are presented through the headphones while gaze position is recorded. The target image is the pictorial representation of a target word (e.g., “beaker”); the rest

of the images are distractors or else potential competitors (items that overlap with the target in phonological, semantic, or visual properties). In essence, the VWP task is visual search under direction from an auditory linguistic stimulus, and eye movements to the target and competing images are closely time locked to significant information in the speech signal. For example, the timing and proportions of eye movements to competing images of items with similar words (e.g., beaker, beetle, and speaker) is predicted by phonetic similarity over time [13,19]. That is, eye movements serve as indicators of unfolding cognitive dynamics as participants comprehend the speech stream.

The present research represents one foray in modern cognitive science that addresses the close intertwining of cognition with its perceptual-motor underpinnings [20]. Once dismissed as “jitter” irrelevant and insensitive to abstract mental processes, the fine-grained variability of perceptual-motor systems has been shown to have a powerful relationship to the higher-order functions of the cognitive system. Changes in the fine-grained variability have been shown to predict and, in some cases, to induce changes in linguistic, categorical, and mathematical reasonings [8,9,21]. This fine-grained behavior exhibits task sensitivities that belie any notion that perceptual-motor systems are simply machines that carry out commands from the cognitive system [4,5,8,9]. What is remarkably task insensitive is the presence of fractal processes. Indeed, just as ecologists have studied the optimality of fractal (specifically, Lévy) search patterns as a driving force in evolution [14], it is the view of a growing number of cognitive scientists that the fractal structure may be crucial for the emergence of flexible context-dependent processes that constitute higher-order cognition [4,8,9,22]. For these reasons, we propose to analyze fine-grained behavior within a standard cognitive task (i.e., VWP) and to characterize its diffusive structure with a view to recent developments in scaling estimation [18].

The first study examined looking behaviors when the target word was a homophone and one of the competing images was conceptually related to the alternate meaning of the target word (e.g., *deck of cards* and *deck of a boat*); 18 University of Connecticut undergraduates completed 70 trials for this study. The second study examined looking behaviors when the items represented by competing images were conceptually related to the target word (e.g., the target is a lion and one competitor is a tiger); 37 University of Connecticut undergraduates completed 95 trials for this study. The effects of the manipulations in each study were not of interest for the present paper. The data will here serve our present purpose of discerning the diffusive properties of eye movements during language processing over multiple dimensions of linguistic complexity. The eye movements during the task were recorded using an ASL 6000 eye-tracking device sampling at 60 Hz.

III. ORDINARY DIFFUSION AND HYPERDIFFUSION

Diffusion is typically quantified in terms of a relationship between fluctuation of a variable $x(t)$ and time,

$$x(t) \sim kt^\delta, \quad (1)$$

where k is a constant and δ is the diffusion coefficient. The fluctuations of $x(t)$ give rise to a probability distribution function $p(x,t)$ with scaling form

$$p(x,t) \sim \left(\frac{1}{t^\delta}\right) F\left(\frac{x}{t^\delta}\right). \quad (2)$$

Diffusion may scale linearly with time, leading to ordinary diffusion, or it may scale nonlinearly with time, leading to hyperdiffusion. Hyperdiffusive processes may be classed as Gaussian or Lévy, depending on whether the central limit theorem (CLT) holds. CLT entails ordinary statistical mechanics. That is, it entails a Gaussian form for F in Eq. (2) composing a random walk without temporal correlations (i.e., $\delta=0$). The crucial point is that, under the CLT, the probability distribution function (pdf) $p(x,t)$ describing the probabilities of $x(t)$ has a finite second moment $\langle x^2 \rangle$, and when the second moment diverges, $x(t)$ no longer falls under the CLT and instead indicates that the generalized central limit theorem applies [16]. Failures of CLT are interesting in light of a growing body of evidence that the thermodynamics underlying many physical, biological, and social phenomena exhibit a departure from ordinary statistical mechanics [17,18].

When the CLT applies, the pdf $p(x)$ has a finite second moment $\langle x^2 \rangle$ as in the case of the Gaussian distribution or the inverse power-law distribution

$$p(x) = x^{-\mu}, \quad (3)$$

with $\mu \geq 3$ [10,16]. Properties of Gaussian diffusion may be expressed in terms of the mean squared displacement (MSD) of x and its relation to time:

$$V = \langle |x(t) - x(0)|^2 \rangle = kt^{2H}, \quad (4)$$

where V is MSD and k is a constant. Under the CLT, H is the Hurst exponent generally taken to be an estimate of δ . When $H=0.5$, MSD is linearly proportional to time:

$$V = \langle |x(t) - x(0)|^2 \rangle = kt. \quad (5)$$

Equation (5) exemplifies the ordinary condition of Bm. The derivative of Bm is additive white Gaussian noise. On the other hand, when $H > 0.5$, MSD increases nonlinearly with respect to time, indicative of hyperdiffusion. In particular, the case of $H=1$ gives rise to the relation

$$V = \langle |x(t) - x(0)|^2 \rangle = kt^2, \quad (6)$$

according to which diffusion follows correlated fractional Brownian motion, whose derivative is fGn.

The divergence of the second moment $\langle x^2 \rangle$ indicates non-Gaussian diffusion. In this case, the pdf $p(x,t)$ may instead be a heavy-tailed distribution with no characteristic scale. One such alternative is found in Lévy diffusion wherein $p(x,t)$ takes the shape of an inverse power law [see Eq. (3)] with $1 < \mu < 3$ [10,16]. Lévy diffusion exhibits scale-invariant fractal trajectories characterized by large steps. Gaussian diffusion has a scale-dependent distribution of steps but can be hyperdiffusive when the steps are temporally correlated; in contrast, Lévy diffusion is always hyper-

diffusive, regardless of temporal correlation in the steps.

IV. SCALING METHODS FOR ASSESSING DIFFUSIVITY

A. Finite-variance scaling methods and multiscaling comparative analysis

The relationships described above are often analyzed in empirical data using finite-variance scaling methods (FVSMs). Examples of FVSMs are the standard deviation analysis (SDA) [23], rescaled range analysis [24], and detrended fluctuation analysis [25]. FVSMs capitalize on the relationship in Eq. (4), estimating the diffusion coefficient δ as equivalent to the Hurst exponent H . FVSMs compare root-mean-square fluctuations over successively larger time bins and thus rely on the MSD of the diffusing variable x .

FVSMs fail to provide reliable estimates for D when diffusion is Lévy rather than Gaussian. In the case of Lévy flight, the second moment diverges to infinity, and FVSMs fail completely as estimators of δ , yielding $H=0.5$ despite the hyperdiffusive nature of Lévy statistics. In the case of Lévy walks, the second moment is finite. FVSMs do not fail completely in estimating δ but the relationship between MSD and time is nonetheless misleading. Because finite variance is not the general case and because variance may give only coarse approximation of diffusivity, an analysis based strictly on entropy has been necessary. In order to adequately estimate δ using Shannon entropy rather than MSD [18].

MSCA is a strategy employed to distinguish between Lévy and Gaussian statistics. It involves estimating H using an FVSM and estimating δ using DEA for the same time series. If $H=\delta$, the time series is governed by Gaussian statistics; if $H \neq \delta$, the time series is not governed by Gaussian statistics. In the latter case, it is possible that Lévy statistics are applicable but it is necessary also to check the pdf $p(x, t)$ of the time series to evaluate the fit of a Lévy distribution [18]. We will carry out MSCA for the time series of eye movements, using SDA as a FVSM to compare with DEA and using relative likelihood estimation to test the fit of a Lévy distribution to the pdf of the time series. The next sections review SDA and DEA.

B. Computing SDA and DEA

Given a time series $x(t)$ for $t=1, 2, \dots, N$, SDA and DEA both begin by creating overlapping subtrajectories of progressively longer length. The strategy here is to describe the diffusion of $x(t)$ as it unfolds over time windows of increasing lengths, throughout the broader time course of the process. The analysis will describe the outcome of diffusion after n time steps, for many different values of n (e.g., $1 \leq n \leq N$). Subtrajectories are created for each available sequence of n time steps, each subtrajectory being lagged by one time step from the previous subtrajectory. Hence, for each value of n , there are $N-n+1$ trajectories.

The creation of subtrajectories is an integration of consecutive values in the time series $x(t)$. The i th subtrajectory $z_i(t)$ of length n is calculated for $t=i, i+1, i+2, \dots, i+n-1$ as follows:

$$z_i(t) = \sum_{j=1}^n x(i+j-1), \quad (7)$$

where i is the starting position, in time series $x(t)$, of the subtrajectory and where j increments by one as the subtrajectory continues by each time step in $x(t)$. That is, $z_i(t)$ is the sum of the values in the time series $x(t)$ in the time window $t=i$ to $t=i+n-1$.

For each length n , there are $N-n+1$ subtrajectories and so also $N-n+1$ values of $z_i(t)$, which form a sample of end points $z(n)$ from which to assess how diffusion scales with time. Here, the similarity between SDA and DEA comes to an end, as each analysis brings to bear a different statistic upon the sample of end points.

As a FVSM, SDA relies on the relation in Eq. (4) to estimate H . Specifically, SDA draws on the square root of this relation,

$$\sqrt{\langle |x(t) - x(0)|^2 \rangle} = kt^H. \quad (8)$$

That is, the standard deviation of the diffusion process is related to time scaled to a power of H . For each subtrajectory length n , SDA takes the standard deviation $D(n)$ of the end points

$$D(n) = \sqrt{\frac{\sum_{i=1}^{N-n} [z_i(n) - \langle z(n) \rangle]^2}{N-n}}, \quad (9)$$

where $\langle z(n) \rangle$ is the average end point of the $N-n+1$ subtrajectories. Because n is equivalent to the time allowed for an n -length subtrajectory to run, SDA approximates Eq. (8) as

$$D(n) \propto n^H, \quad (10)$$

and so

$$\frac{\log D(n)}{\log n} \propto H. \quad (11)$$

Hence, the slope of the function $D(n)$ on double-logarithmic axes gives an estimate of H .

DEA takes a different approach, assessing diffusivity by computing the Shannon entropy of the end points. The time dependence of the entropy measure determines the estimate of δ . To do so, DEA constructs a histogram of $z(n)$, that is, the end points of all subtrajectories of length n , with m bins. The width of the bins is held constant over all values of n . DEA proceeds by computing the Shannon entropy $S(n)$ of the resulting histogram:

$$S(n) = -\sum_{i=1}^m p_i \log p_i, \quad (12)$$

where p_i is the probability of an end point populating the i th bin. Contrary to SDA and all other FVSMs, DEA departs from MSD-based interpretations of Eq. (1). Instead, it recasts

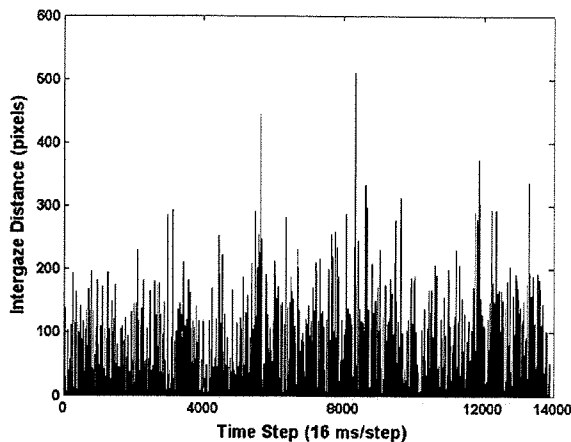


FIG. 1. Time series of intergaze distance for a single participant over the course of an experiment.

Eq. (1) as a relation between entropy $S(n)$ and length of subtrajectory n ,

$$\frac{S(n)}{\log n} \propto \delta. \quad (13)$$

Because $S(n)$ is already logarithmically scaled, there is no need to take its logarithm as SDA does for $D(n)$ in Eq. (11). Hence, the rate of increase in $S(n)$ across a logarithmically scaled axis for n serves as an estimate of δ [16].

V. RESULTS

The data analyzed in the following section are the time series of Euclidean distances between gaze positions at each sample of the eye-tracking device. That is, we analyzed the time series of intergaze distances as they were sampled every 16 ms. Each participant produced one intergaze distance time series across the entire duration of the experiment. Figure 1 shows an example participant's time series over the course of an experiment. Analysis was twofold. First, relative likelihood estimation was used to test the best model fit for the pdf of the intergaze distances. The candidate models tested were power-law, exponential, and gamma, following the recommendation of [26]. Log likelihoods of each candidate model may be used to generate an Akaike weight, an information-theoretic statistic providing a standardized comparison of model fits that can be generalized across samples. The models with higher log likelihoods receive greater Akaike weights. First, the Akaike information criterion (AIC) for each of the three models is computed:

$$A_i = -2L_i + 2P_i, \quad (14)$$

where A_i is the AIC for the i th model, L_i is the log likelihood of model i , and P_i is the number of parameters in the i th model (one for power law and exponential, two for gamma), and second, each model's Akaike weight w is computed by a ranking of the log likelihoods:

TABLE I. Log likelihoods and Akaike weights for intergaze distances by study, which was averaged across participant. Asterisks mark the highest log likelihood and Akaike weight for each model fit.

		Power law	Exponential	Gamma
Study 1	LL	-19 834.38*	-22 937.75	-22 725.64
		(454.17)	(563.32)	(551.36)
	Akaike	1.00*	0.00	0.00
		(0.00)	(0.00)	(0.00)
Study 2	LL	-15 310.14*	-20 136.61	-17 952.20
		(651.15)	(807.64)	(790.49)
	Akaike	1.00*	0.00	0.00
		(0.00)	(0.00)	(0.00)

$$w_i = \frac{\exp\left(-\frac{A_i - A_{\min}}{2}\right)}{\sum_{q=1}^y \exp\left(-\frac{A_q - A_{\min}}{2}\right)}, \quad (15)$$

where w_i is the Akaike weight for the i th model, A_{\min} is the minimum AIC of the three models, and y is the number of models tested (here, $y=3$) [14,27]. Second, MSCA was used to compare the H from SDA and the δ from DEA of the same intergaze distance time series.

In both experiments, the log likelihood of the power-law fit was significantly greater than the log likelihood of the exponential and gamma fits [experiment 1: $F(2,108) = 13.48$, $p < 0.0001$; experiment 2: $F(2,48) = 3.60$, $p < 0.0001$], and the average power-law exponent μ was within the Lévy range [experiment 1: $M = 1.63$, $SE = 0.01$; experiment 2: $M = 1.66$, $SE = 0.02$]. All Akaike weights favored the power-law over the gamma and exponential fits (see Table I). Thus, analysis of the pdf of intergaze distances suggests Lévy-like search patterns in VWP (as in [14]). Figure 2 shows the pdf for the time series in Fig. 1. The solid curve describes the power-law fit ($\mu = 1.62$) of this participant's pdf.

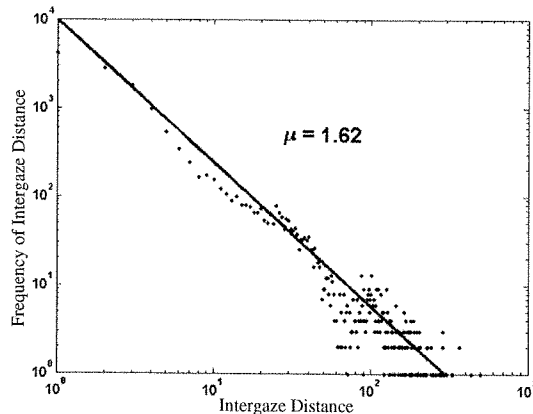


FIG. 2. Pdf of the time series of intergaze distances shown in Fig. 1.

TABLE II. Results of MSCA of intergaze distance time series, showing the estimates of δ from DEA and H from SDA.

	DEA: δ	SDA: H
Study 1	0.74 (0.01)	1.00 (0.00)
Study 2	0.71 (0.01)	1.00 (0.01)

Table II lists the DEA estimates of δ and the SDA estimates of H for both experiments. A paired-sample t test indicated that SDA estimates of H were higher than DEA estimates of δ [experiment 1: $t(35)=33.95$, $p<0.0001$; experiment 2: $t(16)=17.04$, $p<0.0001$]. Figure 3 shows the fluctuation function resulting from SDA of the intergaze distance time series shown in Fig. 1. Figure 4 shows the entropy function from DEA again, for the same time series. The H values are consistent with fGn but the difference between δ and H indicates a departure from Gaussian statistics. The relatively smaller values for δ indicate that, as time elapses in the diffusion process underlying eye movements, the second moment of the diffusion process grows at a faster rate than does the entropy. Whereas the second moment will grow at the same rate as entropy in standard kinetics, this discrepancy indicates a departure from the standard kinetics and suggests strange kinetics instead.

VI. LÉVY-LIKE DIFFUSION SUPPORTS LANGUAGE-DRIVEN VISUAL SEARCH

These results suggest that the perceptual-motor dynamics underlying language-driven visual search are characterized by Lévy-like diffusion. The results are consistent across analyses of intergaze distance time series in two experiments. Thus, the first part of our analysis indicates Lévy-like diffusion, and the second part provides converging evidence using SDA and DEA of the intergaze distance time series to demonstrate the failure of Gaussian statistics.

This finding has a number of implications for understanding the variability of cognitive behavior. First, the Lévy-like diffusive property of human search behavior in the VWP

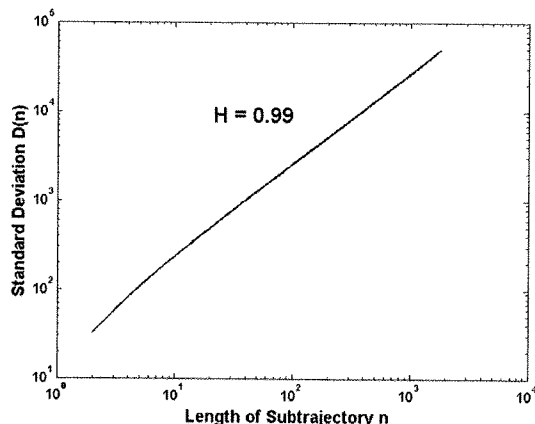


FIG. 3. Fluctuation function from SDA on one participant's intergaze distance time series, $H=0.99$.

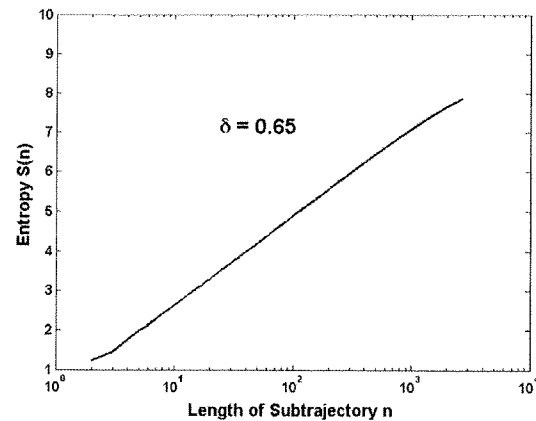


FIG. 4. Entropy function from DEA on the same sample participant's intergaze distance time series, $\delta=0.65$.

mirrors Lévy-like patterns of foraging behavior in many animal species [14], suggesting that the dynamics of language comprehension may be similar to those underlying animal foraging. Second, Lévy-like diffusion indicates that cognitive behaviors are hyperdiffusive. Visual search in spoken-language comprehension thus reflects a scale-invariant fractal process. Looking behavior is not the product of modular components in cognitive architecture (e.g., attention, working memory, etc.) but instead is an emergent property of nonlinear interactions among lower-order biological dynamics, as in Hebbian and self-organizing map algorithms [28].

Third and most importantly, these results suggest that further inquiries into the variability of cognitive behaviors should not take the presence or absence of temporal correlations as the sole diagnostic for assessing diffusive properties. FVSMs will be suitable for testing the fractal nature of cognitive behaviors when the underlying statistics are Gaussian. However, FVSMs will misrepresent the true fractal nature of cognitive behaviors when the underlying statistics appear to be Lévy-like. The recent movement in cognitive science investigating the fractal nature of cognitive behaviors [3–5,8,9,15] will profit from more careful consideration of the strange kinetics [16] driving hyperdiffusion and the multiscaling methods [18] that serve to distinguish them in empirical data. Cognitive behavior can be added to the list of physical phenomena exhibiting strange kinetics. It now remains to pursue better elucidations of the chaotic dynamics that underlie the strange kinetics of cognition. That is, Lévy-like diffusion subsumes a continuum of multiplicativity that may include lognormal as well as power-law diffusion regimes [29]. Future research will be aimed at distinguishing the diffusion regimes in eye movements along this continuum in the VWP task.

ACKNOWLEDGMENTS

Partial support was provided by NSF Grant No. BCS-0643271 to J.D. and NIH Grant No. HD052364 to D.M.

- [1] R. Ratcliff, *Psychol. Rev.* **85**, 59 (1978); R. Ratcliff, P. Gomez, and G. McKoon, *ibid.* **111**, 159 (2004); R. Ratcliff, A. Cheria, and M. Segraves, *J. Neurophysiol.* **90**, 1392 (2003).
- [2] J. Palmer, P. Verghese, and M. Pavel, *Vision Res.* **40**, 1227 (2000); A. Voss and J. Voss, *J. Math. Psychol.* **52**, 1 (2008); V. van den Berg and E. M. van Loon, *Vision Res.* **45**, 1543 (2005).
- [3] D. L. Gilden, *Psychol. Rev.* **108**, 33 (2001); T. L. Thornton and D. L. Gilden, *Psychon. Bull. Rev.* **12**, 409 (2005).
- [4] G. C. Van Orden, J. G. Holden, and M. T. Turvey, *J. Exp. Psychol. Gen.* **132**, 331 (2003); G. C. Van Orden, J. G. Holden, and M. T. Turvey, *ibid.* **134**, 117 (2005); C. T. Kello, G. G. Anderson, J. G. Holden, and G. C. Van Orden, *Cogn. Sci.* **32**, 1217 (2008); J. G. Holden, G. C. Van Orden, and M. T. Turvey, *Psychol. Rev.* (to be published).
- [5] D. J. Aks, G. J. Zelinsky, and J. C. Sprott, *Nonlin., Dyn. Psychol. Life Sci.* **6**, 1 (2002); D. J. Aks, J. C. Sprott, *ibid.* **7**, 159 (2003).
- [6] T. S. Horowitz and J. M. Wolfe, *Nature (London)* **394**, 575 (1998).
- [7] E.-J. Wagenmakers, S. Farrell, and R. Ratcliff, *Psychon. Bull. Rev.* **11**, 579 (2004); E.-J. Wagenmakers, S. Farrell, and R. Ratcliff, *J. Exp. Psychol. Gen.* **134**, 108 (2005); S. Farrell, E.-J. Wagenmakers, and R. Ratcliff, *Psychon. Bull. Rev.* **13**, 737 (2006).
- [8] D. G. Stephen, N. Stepp, J. A. Dixon, and M. T. Turvey, *Physica A* **387**, 5271 (2008); D. G. Stephen, J. A. Dixon, and R. W. Isenhower, *J. Exp. Psychol. Hum. Percept. Perform.* (to be published); D. G. Stephen and J. A. Dixon, in *Studies in Perception and Action IX: Proceedings from the Fourteenth International Conference on Perception and Action*, edited by S. Cummins-Sebree and M. A. Riley (Erlbaum, New York, 2007), p. 172; D. G. Stephen, D. Mirman, and J. S. Magnuson (unpublished).
- [9] D. G. Stephen, R. A. Boncoddo, J. A. Dixon, and J. S. Magnuson, (unpublished).
- [10] D. H. Zanette, *Phys. Rev. E* **55**, 6632 (1997).
- [11] A. M. Treisman and G. Gelade, *Cogn. Psychol.* **12**, 97 (1980).
- [12] R. Engbert, A. Nuthmann, and R. Kliegl, in *Eye Movements: A Window on Mind and Brain*, edited by R. P. G. van Gompel, M. H. Fischer, W. Murray, and R. L. Hill, (Elsevier, Oxford, 2007), pp. 319–337; J. O'Regan, *Can. J. Psychol.* **46**, 461 (1992); D. Ballard, M. Hayhoe, and J. Pelz, *J. Cogn. Neurosci.* **7**, 66 (1995); D. Richardson and M. J. Spivey, *Cognition* **76**, 269 (2000); M. J. Spivey, *The Continuity of Mind* (Oxford University Press, New York, 2007).
- [13] M. K. Tanenhaus, M. J. Spivey-Knowlton, K. M. Eberhard, and J. C. Sedivy, *Science* **268**, 1632 (1995).
- [14] D. W. Sims *et al.*, *Nature (London)* **451**, 1098 (2008); G. M. Viswanathan *et al.*, *ibid.* **381**, 413 (1996); F. Bartumeus, J. Catalan, U. L. Fulco, M. L. Lyra, and G. M. Viswanathan, *Phys. Rev. Lett.* **88**, 097901 (2002); F. Bartumeus, M. G. E. da Luz, G. M. Viswanathan, and J. Catalan, *Ecology* **86**, 3078 (2005); F. Bartumeus, *Fractals* **15**, 151 (2007).
- [15] T. Rhodes and M. T. Turvey, *Physica A* **385**, 255 (2007).
- [16] M. F. Shlesinger, G. M. Zaslavsky, and J. Klafter, *Nature (London)* **363**, 31 (1993); *Lévy Flights and Related Topics in Physics*, edited by M. F. Shlesinger, G. M. Zaslavsky, and U. Frisch (Springer, Berlin, 1995).
- [17] M. F. Shlesinger, B. J. West, and J. Klafter, *Phys. Rev. Lett.* **58**, 1100 (1987); S. V. Buldyrev, A. L. Goldberger, S. Havlin, C.-K. Peng, M. Simons, and H. E. Stanley, *Phys. Rev. E* **47**, 4514 (1993).
- [18] N. Scafetta and P. Grigolini, *Phys. Rev. E* **66**, 036130 (2002); N. Scafetta and B. J. West, *Phys. Rev. Lett.* **92**, 138501 (2004); N. Scafetta and B. J. West, *Complexity* **10**, 51 (2005); N. Scafetta, V. Latora, and P. Grigolini, *Phys. Rev. E* **66**, 031906 (2002).
- [19] M. K. Tanenhaus and J. C. Trueswell, in *Processing World-Situated Language: Bridging the Language-As-Action and Language-As-Product Traditions*, edited by J. C. Trueswell and M. K. Tanenhaus (MIT Press, Cambridge, MA, 2005), p. 3; P. D. Allopenna, J. S. Magnuson, and M. K. Tanenhaus, *J. Mem. Lang.* **38**, 419 (1998); J. S. Magnuson, M. K. Tanenhaus, R. N. Aslin, and D. Dahan, *J. Exp. Psychol. Gen.* **132**, 202 (2003).
- [20] A. Glenberg, *Behav. Brain Sci.* **20**, 1 (1997); A. Clark, *Being There: Putting Brain, Body and World Together Again* (MIT Press, Cambridge, MA, 1997); G. Lakoff and M. Johnson, *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought* (Basic, New York, 1999); L. W. Barsalou, *Behav. Brain Sci.* **22**, 577 (1999); G. Hollis, H. Kloos, and G. C. Van Orden, in *Chaos and Complexity: Recent Advances and Future Direction in the Theory of Nonlinear Dynamical Systems Psychology*, edited by S. Guastello, M. Koopmans, and D. Pincus (Cambridge University Press, Cambridge, 2008), p. 206.
- [21] R. A. Dale, J. Roche, K. Snyder, and R. McCall, *PLoS ONE* **3**, e1728 (2008).
- [22] E. Thelen and L. B. Smith, *A Dynamic Systems Approach to the Development of Cognition and Action* (MIT Press, Cambridge, MA, 1996).
- [23] P. Allegrini, M. Barbi, P. Grigolini, and B. J. West, *Phys. Rev. E* **52**, 5281 (1995).
- [24] H. E. Hurst, R. P. Black, and Y. M. Simaika, *Long-Term Storage: An Experimental Study* (Constable, London, 1965).
- [25] C.-K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, and A. L. Goldberger, *Phys. Rev. E* **49**, 1685 (1994).
- [26] A. M. Edwards *et al.*, *Nature (London)* **449**, 1044 (2007).
- [27] K. P. Burnham and D. R. Anderson, *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach* (Springer, New York, 2002).
- [28] J. L. McClelland, in *Processes of Change in Brain and Cognitive Development: Attention and Performance XXI*, edited by Y. Munakata and M. H. Johnson (Oxford University Press, Oxford, 2006), p. 33; Y. Silberman, S. Bentin, and R. Mikkulainen, *Cogn. Sci.* **3**, 645 (2007); F. G. Ashby, J. M. Ennis, and B. J. Spiering, *Psychol. Rev.* **114**, 632 (2007); T. J. Sullivan and V. R. de Sa, *Neural Networks* **19**, 734 (2006).
- [29] E. W. Montroll and M. F. Shlesinger, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 3380 (1982).