ECOLOGICAL PSYCHOLOGY, *11*(1), 45–79

What Swimming Says About Reading: Coordination, Context, and Homophone Errors

Guy C. Van Orden, John G. Holden,
Michelle N. Podgornik, and Christine S. Aitchison
Cognitive Systems Group
Arizona State University

The pattern of homophone errors (e.g., *BREAK* mistaken for *brake*) may change across different tasks. Categorization, word-identification, and phrase-evaluation experiments were conducted to explore this pattern. Tasks with weak contextual constraints did not yield homophone errors to highly familiar homophones (e.g., *BREAK* is neither falsely identified as *brake* when presented in isolation nor miscategorized as an *object*). However, strong contextual constraints yield homophone errors in phrase-evaluations even to high-frequency homophones (*BREAK*::*part of a car*). The latter result is counterintuitive—the phrase context appears after word identification should have already occurred.

All written languages include systematic relations between spoken words and written words—between phonology and spelling (Mattingly, 1992). This fact suggests a general, if not universal, role for phonology in reading (Perfetti & Zhang, 1995; Perfetti, Zhang, & Berent, 1992). Despite the plausibility of this hypothesis, and a century of research, however, no scientific consensus has been reached concerning the role of phonology in reading (for a contemporary overview, see Frost, 1998). Some laboratory reading tasks produce large, reliable phonology effects, but relatively subtle changes in these tasks reduce or eliminate the effect. As a consequence, the phonology debate has become a debate about tasks in which phonology effects occur or do not occur. The debate about proper task contexts now takes priority and is rapidly expanding. To determine whether, or when, phonology plays a

Requests for reprints should be sent to Guy C. Van Orden, Cognitive Systems Group, Department of Psychology, Arizona State University, Tempe, AZ 85287–1104. E-mail: guy.van.orden@asu.edu

role in reading, we must first determine which laboratory tasks are transparent to the cognitive architecture of reading and which are not. A primary goal in this article is to confront this problem as it appears in the psychology of reading and in cognitive psychology at large.

The development of laboratory reading tasks was guided by a desire to induce context-independent aspects of cognition. With respect to this goal, empirical phenomena become suspect if they depend too much on particular contexts for their expression. Task context is viewed as a source of experimental contamination, not a legitimate source of cognitive phenomena. However, the reliable interplay between phonology effects and task contexts suggests that these carefully contrived laboratory tasks have not created a transparent "context-free" environment within which to induce components of cognition.

The most well-known tool for the discovery of context-independent components of cognition is Sternberg's (1969) additive factors method. Factorial designs allow simultaneous manipulation of several variables. When the effects of two or more factors are strictly additive, the manipulated variables may influence separate, nonoverlapping cognitive components. Alternatively, when interactions are observed, these factors do not satisfy the assumption of selective influence. Factors that interact influence (at least) one common cognitive component. An interaction precludes the assignment of effects to separate cognitive components.[1]

Interaction between phonology factors and task contexts makes it impossible to decide the role of phonology in reading. Interaction could imply that the "common cognitive component" is artifactual, an ad hoc product of participation in the laboratory task. Although we wish that the context debate could be resolved by new empirical findings, the only resolution would come from phonology effects that remain relatively constant across various task contexts. Recent studies that include phonology manipulations have only produced more, and higher order, interactions with task contexts (Azuma & Van Orden, 1997; Berent, 1997; Farrar, 1998; Farrar, Van Orden, & Hamouz, 1998; Gibbs, 1996; Gibbs & Van Orden, 1998; Gottlob, Goldinger, Stone, & Van Orden, 1998; Jared, 1997; Lukatela, Feldman, Turvey, Carello, & Katz, 1989; Lupker, Brown, & Colombo, 1997; Rayner, Sereno, Lesch, & Pollatsek, 1995; Stone & Van Orden, 1993; Strain, Patterson, & Seidenberg, 1995; Taft & van Graan, 1998; Tan & Perfetti, 1997; Xu & Perfetti, 1998; Ziegler, Montant, & Jacobs, 1997).

More, and higher order, interaction effects allow more, not fewer, positions within the context debate. Frost (1998) demonstrated this point in his recent review, although his goal was to resolve these differences. Frost's review illustrated

[1]This discussion is consistent with Sternberg's (1969) original presentation of additive factors method. He suggested, for example, that the $p < .05$ significance level was far too conservative for tests of interactions. In a pursuit of cognitive components, trusting that an interaction effect is not statistically significant amounts to the same thing as accepting the null hypothesis. A more liberal criterion for tests of interactions thus protects the integrity of the theoretical enterprise.

the inherent stalemate in the phonology and context debate. Investigators who emphasize the role of phonology in reading trust task environments that produce reliable phonology effects, but investigators who de-emphasize phonology trust task environments that produce reliable null effects (Van Orden, Aitchison, & Podgornik, 1996). Without empirical resolution, we may choose a different theoretical goal, one that takes interaction effects (which include context) as a basis for cognitive theory. We may adopt a working assumption that does not require phonology effects to be free of context. Task environments determine the likelihood and magnitude of phonology effects in reading, and so task environments are as fundamental to explaining these phenomena as phonology is. Contexts cannot be separated from reading; they are part and parcel of reading.

Some areas of psychology already assume a similar working hypothesis (Saltzman & Kelso, 1987). Consider locomotion, in general, and swimming, in particular. Locomotion always entails an environment, and swimming provides an intuitive picture of how locomotion is causally situated in its environment. We may understand the intertwining of reading performance and task environments by analogy to swimming. No profound insight drives this analogy, other than swimming is just so obviously, inextricably, contextually situated that it makes for an easy illustration. It was used for a similar purpose by Shanon (1993). Swimming naturally motivates a more inclusive scientific perspective, not strictly "what's inside the [swimmer]" but also "what the [swimmer's] inside of" (cf. Mace, 1977, p. 43). We discuss swimming in the next section and then compare swimming to reading. This analogy is used to understand why phonology effects may appear or disappear in different task contexts.

Homophone errors are prototypical phonology effects with regard to the context debate. Homophones, like *rows*, are sometimes mistaken for sound-alike mates (e.g., *rose*). Such homophone errors are quintessential cognitive phenomena that may be observed in everyday reading and writing. Supplying a homophone for its mate is a common spelling error (Bosman & Van Orden, 1997), and homophone substitutions are likely to slip past proofreaders of all ages (Bosman & de Groot, 1996; Van Orden, 1991). Homophone errors can be systematically induced, as when participants mistake *ROWS* for a *flower* in a categorization task. Such homophone errors occur to both familiar (*ROWS*) and unfamiliar (*ROZE*) homophones in a variety of laboratory tasks, in different languages, and by readers at all skill levels (Bosman & de Groot, 1996; V. Coltheart, Avons, Masterson, & Laxon, 1991; V. Coltheart, Avons, & Trollope, 1990; V. Coltheart, Laxon, Rickard, & Elton, 1988; V. Coltheart, Patterson, & Leahy, 1994; Doctor & Coltheart, 1980; Sakuma, Sasanuma, Tatsumi, & Masaki, 1998; Van Orden, 1987; Van Orden, Johnston, & Hale, 1988; Van Orden, Stone, et al., 1992; Wydell, Patterson, & Humphreys, 1993).

ROWS is mistaken for *rose* because homophones share phonology. Readers know the correct spellings. For instance, the same participants who miscategorize *ROWS* as a *flower* perform at ceiling when asked to spell *rows* in another task (Van

Orden, 1987). It is a truly paradoxical effect. The sound or pronunciation of the homophone influences performance in a categorization task in which the word is never heard. The paradoxical character of homophone errors marks them as quintessential cognitive phenomena. The phonology of words is not available in their printed form, which means that homophone errors cannot be attributed exclusively to a source in the environment. Homophone phonology is only implicit in the knowledge that readers bring to reading. Thus, in some fashion, phonology is causally "in between" the presentation of a homophone and the error response. Just such mediating causes motivate a cognitive approach (Mandler, 1985). Mediating causal structures make up the cognitive architecture that cognitive theorists seek to describe. Thus, homophone errors would appear to be ideal candidates for cognitive analysis. No other performance phenomena more transparently point to causes inside the mind, inside the head.

Despite the apparent transparency of homophone errors to cognitive analysis, no consensus has been reached concerning the source of homophone errors or their relation to the cognitive architecture. Some cognitive psychologists accept homophone errors as prima facie evidence that phonology mediates identification of printed words; others interpret homophone errors as artifacts of laboratory tasks (Frost, 1998). This debate is fueled by the fact that some homophones, in some task conditions, do not produce homophone errors. Moreover, the likelihood of homophone errors may change with task demands (V. Coltheart et al., 1994; Jared & Seidenberg, 1991, Peter & Turvey, 1994; Van Orden, 1987) and may also change from task to task (Bosman & de Groot, 1996; Van Orden et al., 1992). Thus we confront a specific instance of the larger context debate: Are task contexts that produce homophone errors more transparent than task contexts that do not? In the remainder of this article, we illustrate an alternative working hypothesis that obviates the contemporary debate over task context. A comparison of swimming to reading may clarify this working hypothesis.

SWIMMING

Swimming is the interplay of a swimmer and water within the larger context of gravity. It can be understood only within the embedding contexts that make buoyancy possible. Swimming exploits a local change in the effect of gravity on the swimmer, as locomotion of any type exploits gravity (Bernstein, 1967; Turvey, 1990a, 1990b). Walking is constrained falling, for example—the center of mass of the body moves outside of its base of support with each step. Swimming is special, precisely because the force of buoyancy against gravity changes the possibilities for locomotion (Thelen, Fisher, & Ridley-Johnson, 1984; Thelen & Smith, 1994). The gravity of the earth does not change in swimming, but the context of its interaction with the body does change.

Other aspects of context change on a faster time scale than gravity. In swimming, a body changes local trajectories of water molecules, which simultaneously affect the trajectory of the body, which simultaneously changes its effect on the water, and so on. Each body movement changes the context for swimming—this is true for the context that moving water provides and for the context that limbs, muscles, and joints provide for each other. At the fast, real-time scale of swimming, it is impossible to measure and control real-time changes in water movement or real-time limb positions (Bernstein, 1967; Thelen, 1995; Turvey, Fitch, & Tuller, 1982). Instead, macroscopic patterns of interaction (e.g., acts of swimming) may serve as objects of analysis. "Coordinations can be described macroscopically, in terms of the patterning of body and limb motions relative to the patterning of environmental objects and events" (Turvey, 1990b, p. 938).

Coordinations are "soft-assembled," consistent with the conceptual framework of emergent, self-organizing phenomena (Kugler & Turvey, 1987). The confluence of constraints associated with bodily movement, fluid dynamics, and the swimmer's goals are coupled via feedback, which allows the online emergence of a unique, efficient, coordination solution. "This type of organization allows the system great flexibility to meet the demands of the task within a continually changing environment, while maintaining a movement category suited to the goal in mind" (Thelen, 1995, p. 81).

Imagine a swimming context, in which a swimmer must swim against a controlled stream of water in a laboratory pool. This model system is a useful simplification of natural swimming (cf. Haken, 1988; Turvey & Carello, 1995). In each trial, the swimmer soft-assembles coordination to remain aligned with a designated position on the edge of the pool for a brief time interval. The trial ends when the swimmer satisfies the time criterion, at which time an X-ray camera is triggered. The X-ray yields a discrete portrait of this complex time-varying system—a skeleton frozen in space, at a fixed point in time (the water being invisible to the X-ray camera). The X-ray reveals the swimmer's stroke, which is the categorical response for that trial.

The swimmer's stroke is the response category, a qualitative macroscopic index of the previously noted, microscopic interactions. Of course, a great deal of swimming goes on between successive X-ray trials. The complex interactions between body and stream occur on the fast, real-time scale of swimming. Stroke categories, on the other hand, are photographed on a slower time scale—the time scale of trial-by-trial observations. Although the static snapshot is opposite in character to the actual continuous dynamics of swimming, the X-ray suffices to distinguish treading water from the breast stroke by recording the relative positions of skull, trunk, and limb bones.

Next, suppose that we change the water pressure that drives the current through the laboratory pool. Flow is first reduced to a trickle, a very weak current. Water molecules do not stop moving, but the water appears to be relatively still; there is no ordered current flow and the swimmer simply treads water to stay in the designated place. The swimmer now maintains the position with a relatively less-ordered coordination of limbs—so much less coordinated that calling treading water a "stroke" gives it a status about which we might disagree.[2]

Increasing the water pressure produces a more coherent current, and this stronger current elicits the breast stroke—a more orderly coordination of limbs that yields a more coherent force against the current. This qualitative change, from treading water to the breast stroke, can be observed at the time scale of trials. But the change is elicited by the quantitative change in the force of the current, complex interactions at the real-time scale of swimming, and the task demands of holding a fixed position until a time criterion. As the current gains force, it elicits a more orderly pattern of coordination to suit task demands.

Now imagine a younger, more frail, participant. The frail swimmer may employ the breast stroke in a weaker current than would the hardy swimmer. The movements of the frail swimmer are more vulnerable to the effects of the flowing water. But if we again reduce the water pressure to a trickle, then the strength of the current may only require treading water for maintaining position, even for the frail swimmer. These qualitative changes, from the breast stroke to treading water, or vice versa, illustrate *bifurcation phenomena:* sudden qualitative changes in stroke induced by continuous changes in the relation between swimmer and current. Changes occur when the force of the stream is too weak or too strong for a previous stroke or when a frail swimmer replaces a hardy swimmer, or vice versa, with respect to task demands.

In this example, each trial outcome, the choice of stroke, is determined by two factors: the force of the current versus the force of the swimmer's strokes. Both factors must be taken into account within their context of observation. There is no context-free point of reference for either factor. In addition, small changes in the relation between the force of the current and the force of the stroke may or may not prompt a change in the type of stroke. A change in stroke is required only when the ratio of current and force of stroke changes outside of a homeorheotic range, the range within which a particular stroke maintains a stable trajectory.

The relation between force of current and force of stroke may provide a succinct account of a swimmer's change in response category, the bifurcation from treading water to the breast stroke. The idealized laboratory situation scales the force of the current against the force of a swimmer's stroke to estimate the bifurcation point. When this ratio (the *control parameter*) becomes greater than 1, the current is too strong for the swimmer to continue treading water, and the bifurcation is observed. "You don't really know you have a control parameter unless its variation causes qualitative change" (Kelso, 1995, p. 45).

Performance in the swimming task highlights the causal interdependence of context and swimming. Neither the stream "numerator" nor the stroke "denomi-

[2]A poll of several friends revealed some disagreement about whether treading water is a stroke, but all agreed that the breast stroke and the crawl are true strokes.

nator" suffices to explain the observed behavior in the task. Their functional ratio is necessary to predict the bifurcation between treading water and the breast stroke, because their effects are causally intertwined in task performance. The force of the current is always strictly entailed in this explanation—even a current that does not induce a change in stroke (e.g., a control parameter close to zero). This account does not isolate a cause of treading water inside the swimmer nor does it isolate a cause of the breast stroke in the movement of water. The terms *cause* and *effect* most usefully refer to macroscopic phenomena, not the morass of microscopic interactions (Stewart & Cohen, 1997). The premonitory cause of the bifurcation is the instability that emerges as current force begins to overwhelm the force of treading water. The effect is the breast stroke, an emergent change in coordination. A roughly parallel example of reading performance is described in the next section.

READING

As we mentioned, Shanon (1993) used swimming to illustrate how cognition is contextually situated. Likewise, discourse is an embedding context for perception of printed words in natural reading (Rueckl & Oden, 1986). Typically, readers exploit discourse in perception, as a swimmer may exploit the flow of a river in swimming downstream. Eye-movement studies illustrate the general intuition. During reading, eyes spend less time on words congruent with discourse. Discourse supplies constraints that quickly cohere with the fixated word and recursively perpetuate the discourse.

Perturbations of the natural scenario add to fixation time and gaze duration (for reviews, see Rayner, 1993; Rayner & Pollatsek, 1989; Reichle, Pollatsek, Fisher, & Rayner, 1998). For example, brief presentation of an unrelated random-letter display, while the eyes are moving toward the target word, will add to the eyes' time spent on the target. Even the brief presentation of a "facilitating" prime word (e.g., *sand* or *beech*), causes the eyes to linger on the target (e.g., *BEACH*), compared to natural reading. The less related a priming word is to the local interaction of context and target, the stronger the perturbation, which further adds to the lag time of the eyes on target (Rayner et al., 1995; Sereno, 1995). Consequently, to facilitate eye movements, the best manipulation is no manipulation; simply leave reading alone.

If we are to measure reading performance, however, we must disturb the natural scenario.[3] It is impossible to manipulate reading without also changing the conditions under which reading occurs. Laboratory studies never selectively observe

[3]Robert Shaw (personal communication, March 10, 1998) at the University of Connecticut made this point more generally: "Anytime you take a measurement you establish a dynamic linkage between two systems."

component processes of natural reading. Rather, observation is limited to the macroscopic products of interaction between readers and laboratory manipulations within the boundary conditions of task demands—just as the X-ray observation is a snapshot that reveals the type of stroke produced in the interaction of swimmer and current force for the swimmer to maintain position in the lab pool.

Measurement is concerned strictly with the interaction of readers and task environments, created by scientists and expressed in performance. Eye-movement studies illustrate how subtle changes in reading context perturb the pattern of a reader's eye movements. In laboratory reading tasks, more dramatic changes may pit the reading context fully against a stimulus-word identity, analogous to pitting the force of the current against the force of the swimmer's stroke. Again, the predicted outcome will depend on a ratio of constraints—the control parameter. Such opposing forces allow for a test of online interactions that are not observable in natural reading, in a manner parallel to pitting the current force against the forces of the swimmer's stroke to reveal the bifurcation point between treading water and the breast stroke, which is not obvious in natural swimming.

Consider again homophone errors in the categorization task. Like the swimming task, the categorization task is a useful model system, a carefully constructed simplification of natural reading (cf. Haken, 1988; Turvey & Carello, 1995). Participants are presented with a category name (*flower*) followed by a target foil such as *ROWS*. If participants identify *ROWS* as *rose*, they will incorrectly respond *yes*, and if they identity *ROWS* as *rows*, they will correctly respond *no*. The target homophone *ROWS* sounds just like a category exemplar of the category *flower*. Category constraints may combine with the phonology of *ROWS* in a resonance, which is the basis of a *yes* response.[4] The context *flower* constrains performance to the homophone *ROWS*, in a manner that strongly favors the identity of the sound-alike exemplar *rose*. If the combination of constraints that favor a *yes* response cohere, then a homophone is misidentified as a category exemplar. Strong category constraints may elicit a *yes* response, loosely analogous to a stronger current eliciting the more ordered breast stroke.

---

[4]Our working hypothesis concerns how context may participate in the self-organization of a categorical response. It combines ideas that have demonstrated utility in "neural" network modeling. Masson (1995) and Farrar (1998) presented models in which contextual constraints cohere with stimulus constraints to speed naming. Masson used *hard-clamping* of context: a select set of node values that cannot change. Farrar used *soft-clamping* of context: changing node values to which constant relatively weak activation is added on each iteration of interactive activation. Lewenstein and Nowak (1989a, 1989b) described models in which recognition performance is soft-assembled (i.e., an emergent product of self-organization in a dynamical system), and they took pains to demonstrate the general character of these models (see also Skarda & Freeman, 1987). *Yes* responses in a recognition task correspond to a more-ordered attractor state; *no* responses correspond to a less-ordered state. Gibbs and Van Orden (1998) and Van Orden and Goldinger (1994) discussed soft-assembly of lexical decision performance in similar terms (cf. Lukatela & Turvey, 1987, 1994).

Categories may be circumscribed broadly (*living thing*) or narrowly (*flower*). A broad category such as *living thing* is a relatively "weak current," whereas a narrow category such as *flower* is a "stronger current." With this in mind, please take note of the following analogies:

1. A broad category supplies only weak constraints in favor of the sound-alike identity. Nevertheless, If *ROWS* is an uncommon, low-frequency homophone, then even the weak context may elicit the false-positive *rose* identity and the homophone error. Compare this to a frail swimmer, who is forced to the breast stroke in both weak and strong currents.
2. If *ROWS* is a high-frequency homophone, then this error is unlikely to occur (Jared & Seidenberg, 1991). A weak context is too weak to force coherence with context upon a high-frequency homophone. Compare this to a hardy swimmer, who is not forced to the breast stroke in a weak current; treading water suffices.
3. On the other hand, when the category is narrow, then homophone errors are made to both high- and low-frequency homophones (Jared & Seidenberg, 1991; Van Orden, 1987; Van Orden et al., 1988). Compare this to both frail and hardy swimmers, who are forced to the breast stroke when the current is strong.

Suppose that *ROWS* is a high-frequency word. We may assume strong stimulus constraints in favor of the correct identity *rows*. For such a hardy swimmer, the context might narrow quite a bit from *thing* to *living thing* or possibly even to *plant*, in the direction of *rose*, but the balance of forces would still favor correct identification of *ROWS*. However, when *flower* precedes *ROWS*, the balance teeters in favor of *rose*, and a *yes* response is produced. Performance seems to indicate two qualitatively different processes, as though a shift had occurred from a data-driven (bottom-up) process to a conceptually driven (top-down) process (as though treading water resides in the swimmer, and the breast stroke resides in the force of the current).

However, it is just as reasonable to interpret the apparent shift from bottom-up to top-down as a bifurcation phenomenon. From this alternative perspective, an interaction of environmental and cognitive constraints is always present. The bifurcation occurs because continuous changes in a single control parameter eventually reconfigure the outcome of this interaction. It favors one or the other response to */roz/*. The interaction occurs at the fast real-time scale of perception and action, but the response outcome is revealed on the slower time scale of laboratory observations. When the ratio of constraints favors the *rows* identity, it does not elicit coherence with the category context. The consequent incoherent dynamics suffice for a *no* response. When the ratio of constraints favors the *rose* identity, it also elicits coherence with the category context. The consequent coherent dynamics suffice for a *yes* response.

In this example, the ratio between category breadth and homophone word-frequency approximates the control parameter, but this description is too simple. Similarity in spelling between the stimulus homophone and the exemplar homophone and the frequency of the exemplar homophone also affect performance (Van Orden, 1987). These factors affect the local resonance between stimulus *ROWS* spelling and top-down activation of *rose* spelling, respectively. This local resonance verifies spelling (cf. Grossberg & Stone, 1986), which also determines whether category context and stimulus will resonate. Matching and mismatching constraints in any part of the system are combined in the ratio of constraints that favor or oppose the exemplar identity. The dissimilarity in spelling between *ROWS* and *rose*, as well as high (*rose*) exemplar frequency, both contribute forces that oppose coherence of the *rose* identity with stimulus constraints (cf. Stone & Van Orden, 1994; Van Orden & Goldinger, 1994) and thus oppose resonance with category constraints. The local resonance contributes to the global ratio of constraints that determine categorization performance.[5] However, our simple description, which was framed in terms of homophone frequency and category breadth, will suffice for the prediction tested here concerning high-frequency homophones.

Given the previous scenario, a sudden change of context to favor the *rose* identity could yield a sudden change in the identity of *ROWS*. Moreover, if the outcome is controlled by a ratio of context and stimulus constraints, rather than by a distinct, time-dependent component of word identification, then the temporal order of context and stimulus should not be crucial. Our radical prediction is that a trailing context may determine homophone identity after word identification should have supplied an incorruptible *rows* identity. This prediction follows directly from treating homophone errors as bifurcation phenomena, so we call this the *identity bifurcation hypothesis*.

The remainder of this article is organized in three parts. First, a categorization experiment is described that replicated the key result from Jared and Seidenberg

[5]At the appropriate level of abstraction, recurrent network equations yield summary control parameters (Farmer, 1990; Saltzman & Munhall, 1992) that determine bifurcation points between behavioral options (cf. Farrar & Van Orden, 1998). Typically, recurrent activation includes matching sources of feedback that promote resonance and mismatching sources of feedback that oppose resonance. We may perpetuate this dynamic with a *control vector*—a fixed vector of activation values corresponding to stimulus and semantic activation (stimulus–context), added to state vector of a model on each iteration. (See Peitgen, Jürgens, & Saupe, 1992, for a general introduction to iterative models, including the use of control vectors.) Subsequently, the network exhibits bifurcation phenomena. The ratio of "matching feedback," which promotes resonance (the numerator), to "mismatching feedback," which opposes resonance (the denominator), is the control parameter, loosely analogous to a generalized Reynolds number, a π-number. (See Turvey, 1990a, for an introduction to the latter concepts.) Below a critical value of the control parameter—the bifurcation point—network dynamics remain relatively incoherent (cf. Lewenstein and Nowak, 1989a, 1989b; Skarda & Freeman, 1987; Turvey, 1990a). This is a soft-assembled *no* response. Above the critical value resonance occurs, the soft-assembled *yes* response.

(1991). This replication established that a broad category context, presented prior to target identification, does not produce an exaggerated rate of homophone errors to high-frequency homophones. Second, a word-identification experiment verified that the presentation conditions of the categorization experiment yield veridical perception of high-frequency homophones' spellings. Third, we report the results of the phrase-evaluation experiment. Targets were evaluated against trailing phrases that either were or were not closely related in meaning. The phrase-evaluation experiment tested our prediction derived from the identity bifurcation hypothesis: A trailing phrase may change the ratio of constraints (the control parameter) of context and stimulus and induce homophone errors to very-high-frequency homophone words, after word identification should already have occurred.

CATEGORIZATION

Performance to homophone words can be strongly affected by a preceding context. In the categorization task, for example, a category name such as *flower* or *part of a car* is presented and then replaced by a target word. The participant indicates whether the target word is an exemplar of the category. In this task, participants make an inordinate number of errors to homophones such as *ROWS* or *BREAK*. They are often incorrectly categorized as exemplars, much more often than control items such as *ROBS* or *BRAVE*. As we noted previously, Jared and Seidenberg (1991) established that category context strongly affects performance to high-frequency homophone words.

We next replicated Jared and Seidenberg (1991). To our knowledge, no replication of their results has been published. It is important for our argument that their key result be reliably established using presentation conditions comparable to the experiments that follow. Additionally, Jared and Seidenberg did not report the means from the conditions of their experiments; they reported only difference scores. Thus, our replication supplies a more detailed rendition of the effect.

The method of our categorization experiment was strictly parallel to that of Jared and Seidenberg (1991) with respect to their key hypothesis. We conducted a categorization task with broad categories (*object*) and relatively high-frequency homophone targets (*BREAK*). We extended their method, however, with the addition of a pattern mask (#######) that replaced the target word after 200 msec. The 200-msec stimulus-onset asynchrony (SOA) set precise control over presentation conditions while allowing participants plenty of time to see the target word. Consequently, we replicated the essential circumstances of Jared and Seidenberg's Experiment 2 and supplied a more precise basis for comparison to our experiments. As in Jared and Seidenberg's experiment, we expected no more errors to homophones (*BREAK*) than to controls (*BRAVE*).

The categorization experiment included a 43-msec SOA condition as well. Pattern masking after brief presentation weakens stimulus constraints and may exaggerate the effect of context, especially a context that precedes the stimulus. If this effect combines with stimulus constraints (including homophone phonology), then we may find an exaggerated rate of homophone errors as compared with controls. Alternatively, the context may be too weak to affect target identification, even when target identification could otherwise benefit from contextual constraints.

Method

Participants. Eighty introductory psychology students from Arizona State University participated for course credit. All were native English speakers with normal or corrected-to-normal vision. Forty participants in the 200-msec condition and 40 participants in the 43-msec condition were assigned at random to each of two stimulus-list conditions, described next.

Stimuli. The key stimuli were 18 high-frequency homophones, such as *BREAK* and their yoked controls (*BRAVE*). The 18 yoked stimulus pairs were grouped into two lists. Each list contained nine homophones and nine controls. No participant was presented with both the homophone and its corresponding control (i.e., if *BREAK* was in the first list, then *BRAVE* was in the second list). The frequency counts of the homophone targets ranged from 27 to 6377 per million (*Mdn* = 98; Kučera & Francis, 1967). Each control (*BRAVE*) was equated to its yoked homophone (*BREAK*) along theoretically meaningful dimensions other than phonologic similarity. The crucial dimension was similarity in spelling to the exemplar homophone (*brake*), estimated using an index of orthographic similarity (*OS*). *OS* varies from 0 to 1, where 1 indicates identical spellings (Van Orden, 1987; cf. Weber, 1970). On average, *OS* was greater for controls (.66, $SE = .02$) than for homophones (.55, $SE = .04$; $t[17] = 3.08, p < .05$).

Procedure. Each participant was seated in a soundproof room facing an IBM-compatible computer. All stimuli were presented on the computer monitor and the participants responded by pressing labeled *yes* and *no* keys on the computer keyboard. In each session, a participant was provided with 20 practice trials followed by 118 experimental trials: nine homophones, nine controls, 50 filler *yes* trials, and 50 filler *no* trials. The practice trials did not contain any homophones. Each participant saw a different random presentation order of the experimental trials with the exception of the first four trials, which were filler trials presented in fixed order. Participants were instructed to respond quickly but accurately.

Each trial began with a 2-sec presentation of a category name (*living thing* or *object*), after which it was replaced by a fixation stimulus (+). The + appeared for

500 msec in the center of the monitor, to help the participant fixate on the location on the screen where the target would appear. Each target word (*BREAK*) was presented for 200 msec before it was replaced by a pattern mask (#######). The pattern mask was displayed until the participant responded, then the next trial was initiated. In the 43-msec SOA condition, SOA between target and mask was decreased gradually in the practice trials from an initial SOA of approximately 100 msec to an SOA of 43 msec (3 ticks at 14.2 msec/tick).[6] The first four trials were conducted with a 100-msec SOA.

Results and Discussion

The significance level for all statistical tests was $p < .05$. The dependent variable was the percentage of categorization errors to homophones and controls. Item means appear in Appendix A.

In the 43-msec SOA condition, neither analyses by participants, $t(39) < 1$, nor by items, $t(17) < 1$, yielded a statistically reliable homophone effect. The percentage of categorization errors to homophones was almost identical to that of controls (homophone M = 25.6%, $SE = 3.0$; control M = 25.0%, $SE = 3.6$). Apparently, broad contexts do not elicit homophone errors to high-frequency homophones, even in a brief exposure condition, when contextual constraints could be of most use.

It is sometimes argued that brief SOA pattern masking induces a strategic reliance on phonology (cf. Carr, Davidson, & Hawkins, 1978; Carr & Pollatsek, 1985; Hawkins, Reicher, & Peterson, 1976; Verstaen, Humphreys, Olson, & D'Ydewalle, 1995; Wydell et al., 1993; but see Berent & Van Orden, 1998; Hooper & Paap, 1997; Xu & Perfetti, 1998; Ziegler, Van Orden, & Jacobs, 1997, for counterarguments). This strategy hypothesis is not supported by the present null homophone effect, but neither is it strongly contradicted. The homophone targets were all relatively high-frequency words, typically higher in frequency (often by an order of magnitude) than their mates. Stimulus constraints, including associations with phonology, would strongly favor correct homophone identities. If phonology were the only basis for word identification, the present high-frequency homophones would still be correctly identified. Thus, failure to see homophone errors to high-frequency homophones is not a test of any existing phonologic mediation hypotheses. There is no basis for predicting homophone errors to high-frequency homophones except with respect to a biasing context.

[6]The categorization experiment was rerun after the phrase-evaluation experiment, using software that allowed more precise control of SOA. However, we had twice previously replicated the outcome of the brief SOA condition of the categorization experiment using the same software as the phrase-evaluation task, but with fewer participants. We reran the experiment on the fastest, 43-msec, setting of the 43–57-msec range, with more participants, to increase the power of the categorization method to detect homophone errors.

In the 200-msec SOA condition, neither the analysis by participants, $t(39) = 1.23, p > .22$, nor by items, $t(17) < 1$, yielded a statistically reliable homophone effect. Only a few more categorization errors were made to homophones than to controls (homophone M = 11.0%, *SE* = 1.9; control M = 9.2%, *SE* = 2.2). This null homophone effect replicates the null finding of Jared and Seidenberg (1991). High-frequency homophones were correctly rejected at a rate comparable to controls. Thus, presentation conditions in which high-frequency homophones appeared for 200 msec, and were categorized with respect to broad categories, yielded correct identification of the homophone words, compared to controls. Categorization performance was not misled by the ambiguous phonology of homophone words.

WORD IDENTIFICATION

The word-identification task was straightforward. Participants simply wrote down the target item that they perceived. Thus, we could directly establish whether targets that appeared for 200 msec are correctly identified. Also, the relative freedom to choose stimuli (i.e., they did not need to sound like exemplars from broad categories) allowed us to use the same target items as appeared in the phrase-evaluation task. Transcription and reading have spelling representations in common (see Rapp & Caramazza, 1997, for an overview). Thus, this experiment had the opportunity to establish a baseline of veridical perception with which to contrast word identification in the phrase-evaluation task.

Method

Participants. Forty Arizona State University introductory psychology students participated for course credit. All were native English speakers with normal or corrected-to-normal vision. As in the categorization experiment, 20 participants were assigned at random to each of two stimulust-list conditions.

Stimuli. The target items were the targets from the phrase-evaluation experiment, discussed later (listed in Appendix B). These targets were frequently encountered homophones such as *BREAK* and their yoked controls (*BRAVE*). The frequency counts of the homophone targets ranged from 54 to 1125 per million (*Mdn* = 116; Kučera & Francis, 1967). Twenty yoked stimulus pairs were grouped into two lists for presentation. Each list contained 10 homophones and 10 spelling controls. No homophone and its corresponding spelling control were presented together within the same session (i.e., if *BREAK* was in one list, then *BRAVE* was in the other list). As in the categorization experiment, each control (*BRAVE*) was equated to its yoked homophone (*BREAK*) along theoretically meaningful dimensions other than phonologic similarity. The key dimension was similarity in spelling to the alternative homophone (*brake*) as estimated by *OS*. On average, *OS* was greater for controls (.66, $SE = .02$) than for homophones (.60, $SE = .02$; $t[19] = 2.42, p < .05$). Additionally, yoked items were matched closely on number of letters and average bigram frequency (Massaro, Taylor, Venezky, Jastrzembski, & Lucas, 1980).

Procedure. Each participant was seated in a soundproof room facing an IBM-compatible computer. Stimuli were presented on the computer monitor. Participants responded by writing the stimulus word on a separate sheet of paper, provided by the experimenter. Each session included 20 practice trials followed by 20 experimental trials. The practice trials did not contain any homophones. Each participant saw a different random presentation order of experimental trials. The participants were instructed to print clearly and respond as accurately as possible.

The trials were self-paced; participants pressed a button to initiate each trial. After a trial was initiated, a fixation stimulus (+) appeared, followed by a target word. The + appeared for 500 msec in the center of the monitor, to help ensure that the participant was fixating the location on the screen where the target would appear. Each target word was presented for 200 msec and then replaced by a pattern mask (#######). The pattern mask was displayed for 100 msec and then the screen was cleared until the next trial was initiated.

Results and Discussion

A total of 800 responses were collected (20 words × 40 participants). A single unambiguous error occurred. The control word *DOER* was presented, but the homophone *dear* was recorded by the participant (*DEAR* was not presented in this participant's session; see Method). In addition, one other control word was recorded ambiguously. The word PILL was presented, and the participant printed a word that looked like the word *pill*, but there was a small mark between the *p* and the *i* that may have been a lowercase *r* or possibly a stray mark; we couldn't tell which. In any case, 798 unambiguously correct responses of 800 possible responses is sufficient to claim performance is at ceiling. The spellings of high-frequency homophones were correctly perceived in these presentation conditions.

PHRASE EVALUATION

The phrase-evaluation task completed our test of the identity bifurcation hypothesis. Participants saw a target followed by a pattern mask, before they were presented with a comparison phrase, and then judged whether the target and the trailing comparison phrase were closely related in meaning. As in the previous experiments, we included homophone (*BREAK*) and control (*BRAVE*) targets and the target-mask SOA condition of 200 msec. One can consider the possibilities for identification of *BREAK*. If we assume that homophones are inherently ambiguous due to their phonology, then the capacity of *BREAK* to maintain a correct identity is inherently weaker than a nonhomophonic word. Strong constraints due to the trailing comparison phrase (*part of a car*) may elicit the identity of *brake*. If so, then we should observe phrase evaluations in which *BREAK* is mistaken for *part of a car*. These conditions test for a postidentification context effect. This context directly opposes correct word identification, but it uses presentation conditions that yield correct categorization performance and correct perception of spelling.

The method was biased against the predicted outcome in several ways. The 200-msec presentation condition was adequate for veridical word perception. By comparison, the amount of time that a skilled reader gazes at a particular word before a saccadic jump is 200 to 250 msec (Rayner & Pollatsek, 1989). Word identification is complete within approximately 200 msec (cf. Swinney, 1979). For example, "We shouldn't be too surprised if word identification took place in as little as 100 msec or as much as 200 msec [on average] after the word is first sensed by the eye" (Rayner & Pollatsek, 1989, p. 68). Sereno and Posner (1995) estimated that it is complete after 150 msec (on average). Our targets were visible for 200 msec. Additionally, the homophone targets were relatively high-frequency words, and high-frequency words are recognized more quickly (Forster & Chambers, 1973; Whaley, 1978). Moreover, the two previous experiments reliably established correct performance to high-frequency homophone words in comparable tasks and presentation conditions. Prior reports of trailing context effects did not satisfy all of these criteria.

Prior demonstrations of trailing context effects used very brief target presentation (Potter, Moryadas, Abrams, & Noel, 1993; Whittlesea & Jacoby, 1990) or presented a facilitating context (Briand, den Heyer, & Dannenbring, 1988; Kiger & Glass, 1983) or both (Berent & Perfetti, 1995; Jacobson & Rhinelander, 1978; Perfetti et al., 1992; Potter, Stiefbold, & Moryadas, 1998; Stone & Van Orden, 1989). Moreover, facilitating contexts were usually presented within the time course of word identification (Berent & Perfetti, 1995; Jacobson & Rhinelander, 1978; Kiger & Glass, 1983, Perfetti et al., 1992; Stone & Van Orden, 1989). Thus, these previous experiments were not experiments in which a word identity was reliably established and then changed by an opposing trailing context.

Except for Potter et al. (1993), the related context in all the cited studies favored the correct identity, not an alternative identity. Potter et al.'s Experiment 5 is most comparable to this phrase-evaluation experiment. In their experiment, sentences were presented using rapid serial visual presentation (RSVP). In this RSVP procedure, each word was replaced after 100 msec by the word that followed it in the sentence at the same location on a computer monitor, as though the words came one on top of the other. The end of each sentence was signaled by a pattern mask (&&&&&&) that replaced the last word. After each sentence was presented, participants recalled the words aloud. The targets were words such as *RACE* presented in a trailing sentence context appropriate to a different word (*rice;* e.g., "In the *RACE* she cooked for supper there were vegetables.") This presentation pits the trailing context against the correct *RACE* identity. For the targets (*RACE*), the trailing context reliably biased recall performance away from the correct identity and toward a contextually appropriate identity such as *rice* (see also Forster & Hall, unpublished, as cited in Forster, 1974).

Potter et al. (1993) successfully demonstrated a dramatic effect of the trailing sentence context. However, the presentation conditions of 100-msec SOA may have been less than the time required for correct word identification. Potter et al.'s experiments produced performance near ceiling in only a few conditions, which would demonstrate adequate conditions for word identification. However, those few conditions were all conditions in which the target meaning was consistent with, and facilitated by, its sentence context. Conditions without facilitating contexts always produced suboptimal baseline performance. Suboptimal baseline performance was appropriate to Potter et al.'s design (and the designs of the other cited experiments), given their hypotheses. Here, however, we wished to test for a context effect under conditions that would otherwise produce ceiling performance. In this phrase-evaluation experiment, a trailing context supplied systematic constraints in favor of an alternative homophone identity, and followed a target presented with a 200-msec SOA. High-frequency homophones presented for 200 msec, followed by a pattern mask, established a sufficient baseline of correct performance. We demonstrated this fact in the previous categorization and word-identification experiments.

The phrase-evaluation experiment included a 50-msec SOA condition, as well. Pattern masking after 50 msec was expected to weaken stimulus constraints. The guiding framework of our predictions assumes an implicit competition between *break* versus *brake*, due to ambiguous phonology. A manipulation that weakens stimulus constraints may weaken constraints that favor *break* and exaggerate the effect of context. We were particularly interested in whether context (*part of a car*) builds on the phonology of the homophone (*/breyk/*) targets. By some accounts, the phonology of high-frequency words (homophones) is activated too slowly to affect word identification (M. Coltheart & Rastle, 1994; Seidenberg, 1985), unless a biasing context preceded presentation (Jared & Seidenberg, 1991).

Method

Participants. Forty Arizona State University introductory psychology students participated for course credit. All were native English speakers with normal or corrected-to-normal vision. Twenty participants, in the 200-msec condition, and

20 participants, in the 50-msec condition, were assigned at random to each of two stimulust-list conditions, described next.

Stimuli. The target homophones and controls were the same as in the previous experiment.

The phrase-evaluation task was like the categorization task. Participants judged the semantic relatedness between two stimulus events (although the phrase-evaluation task included a wider range of semantic relations—e.g., *BREAK::part of a car*, *CELL::trade goods for money*, *SENT::penny*). However, the phrase-evaluation experiment, unlike the previous categorization task, included narrow contexts (e.g., *BREAK::part of a car* rather than *object::BREAK*). This should make false-positive *yes* responses more likely. The categorization task demonstrated that a broad category context was too weak to destabilize the correct word identities of high-frequency homophone words. Presumably, the narrow context of the phrases was a stronger source of constraint.

Each phrase appeared once for each participant. Consequently, no participant saw both a homophone and its yoked control. This guaranteed that comparison phrases could not be used to generate expectations concerning targets on upcoming trials (cf. Jared & Seidenberg, 1991). If participants had seen yoked homophones and controls in the same session, then half of the homophone and control trials (on average) would have been preceded at some point in the experiment by their corresponding comparison phrases (i.e., they would have been preceded by their yoked control or homophone trial, respectively). To prevent this source of contextual bias, we formed two lists of stimuli, each containing 10 homophones and 10 controls. If a homophone target appeared in one list, then its yoked control appeared in the other list. Each participant was presented with only one of the lists of homophones and controls. Counterbalancing ensured that the respective lists contributed equally to each cell of the design.

In addition to 10 homophone and 10 control *no* trials, each session included 50 filler *yes* trials and 30 filler *no* trials, for a total of 100 experimental trials. *Yes* trials presented targets that were semantically related to their comparison phrase, and *no* trials presented targets that were not related to their comparison phrase (e.g., *TRUST::kitchen appliance*).

Procedure. Each participant was seated in a soundproof room before an IBM-compatible computer. Stimuli were presented on the computer monitor, and participants responded by pressing labeled *yes* and *no* keys on the computer keyboard. Each participant was presented with 20 practice trials followed by 100 experimental trials. The practice trials did not contain any homophones, and each participant saw a different, random, presentation order of experimental trials. Participants were instructed to respond quickly but accurately.

Each trial began with the message "PRESS THE SPACE BAR WHEN READY." When the participant pressed the space bar, a fixation stimulus (+) appeared, followed by a target word. The + appeared for 500 msec in the center of the monitor, the location at which the target would appear. The target word (e.g., *BREAK*) was presented for either 50 msec or 200 msec and then replaced by a pattern mask (#######). The pattern mask was displayed for 100 msec and then replaced by the comparison phrase (e.g., *part of a car*). The comparison phrase remained visible until the participant responded. Immediately after a response, participants were signaled that they may initiate the next trial (i.e., "PRESS THE SPACE BAR WHEN READY").

Participants were instructed to press a *yes* key if the target and the comparison phrase were closely related in meaning and a *no* key otherwise. Accuracy was recorded, but response-time data were not collected. Response times would primarily reflect a phrase length effect, and phrase length was not controlled within participants across the manipulation of homophony.

In the 50-msec SOA condition, SOA between target and mask was decreased gradually in the practice trials from an initial SOA of approximately 100 msec to approximately 50 msec (±7 msec). The first four trials were all conducted with a 100-msec SOA (±7 msec). The random variability was within the time span of a single refresh cycle of the computer monitor. This random variability did not differentially affect homophones versus controls (the crucial dimension of control in this experiment). In the 200-msec condition, the SOA between each target and the mask was the same as in the previous experiment and did not change between the practice and the experimental trials. At the end of the practice block, participants were given an opportunity to ask questions and clear up any misunderstandings concerning the task. Following the practice block, participants were presented with the sequence of 100 experimental trials. No homophones or controls appeared in the first 10 (filler) trials.

Results

The dependent variable was the percentage of phrase-evaluation errors to homophone and control target items. Item means appear in Appendix B.

Participants made more phrase-evaluation errors to homophones (homophone M = 23.5%, *SE* = 5.2) than to controls (control M = 5.5%, *SE* = 1.8) in the 200-msec SOA condition. This effect is reliable in both participant, $t(19) = 3.89$, and item analyses, $t(19) = 5.55$. Likewise, reliably more errors were made to homophones (homophone M = 46.5%, *SE* = 5.3) than to controls (control M = 19.5%, *SE* = 4.0) in the 50-msec SOA condition, (participant $t[19] = 5.42$; item $t[19] = 6.37$). The larger homophone effect (27%) in the 50-msec SOA condition versus the 200-msec SOA condition (18%) was reliably larger by items, $F(1,19) = 4.83$, but not by participants, $F(1,38) = 1.75, p < .20$. The apparent interaction suffices to contradict the prediction of slow phonology. If phonology was too slow to constrain identification in the 50-msec SOA condition, then the homophony effect in this condition should have been smaller compared to the 200-msec condition, not larger.

GENERAL DISCUSSION

At the beginning of this article, we drew an analogy between categorical responses in a swimming task (treading water vs. the breast stroke) and categorical *no* and *yes* responses in reading tasks. In both examples, context was idealized as a force in the task environment that may induce a more highly ordered response category—the breast stroke in swimming or the *yes* response in category and phrase-evaluation tasks. This analogy illustrates how context can be integrated into explanations and why it is not useful to ask where stimulus effects start and context effects end. Stimulus effects are always contextually conditioned (Gibbs & Van Orden, 1998; Stone & Van Orden, 1993; Van Orden, Jansen op de Haar, & Bosman, 1997, Van Orden & Paap, 1997). The view that context and stimulus effects are intertwined accommodates the absence of homophone errors in the categorization experiment. In turn, it correctly predicted the reappearance of homophone errors in the phrase-evaluation experiment.

On 23% of the trials, in the 200-msec SOA condition, high-frequency words (*BREAK*) were misidentified as their lower frequency homophone mates (*brake*), and the error rate rose to 46% for an SOA of 50 msec. These results corroborate the prediction derived from the identity bifurcation hypothesis. A trailing context, such as *part of a car*, elicited the false sound-alike identity of a high-frequency homophone (*BREAK*), long after word identification should have occurred. This result is striking because the homophone effect was large and because it occurred to frequently encountered words. Reading common words must rank among the most frequently recurring episodes of (literate) human experience. By many accounts, frequently encountered homophones should not produce homophone errors; they should be correctly interpreted. We first discuss this result and other support for the identity bifurcation hypothesis, and then discuss how conventional accounts might attempt to accommodate our findings.

The outcome of the phrase-evaluation experiment supports our working hypothesis, but it does not strictly demonstrate a bifurcation point. A continuous manipulation of the control parameter could demonstrate the point of bifurcation in performance, but it would require a continuous scale for constraints of stimulus and context. We did not systematically construct a continuous scale for these manipulations. Nevertheless, the distributions of categorization response times described by Van Orden (1987) and Van Orden et al. (1988) for yoked within-participant trials provide additional motivation for the identity bifurcation hypothesis.

In a categorization task with a narrow category context (*flower*::*ROWS*), homophones produced slower mean *yes* and *no* response times, compared to actual exemplars or control foils, respectively. However, these differences in mean response times were produced by nonlinear changes in the shape of homophones' *yes* and *no* response-time distributions. The distribution of homophone error (*yes*) response times differed only in their slow response-time tails from the distribution of yoked correct response times to actual exemplars (*TULIP*). The fast ends of their distributions were identical. Likewise, correct *no* response times to homophones (*ROWS*) differed only in their slow response-time tails from the distribution of yoked correct response times to control foils (*ROBS*). Otherwise, the distributions were identical (these findings are detailed in Van Orden, 1987; Van Orden et al., 1988).

An additive (linear) homophone effect would shift the entire distribution of homophone response times to slower response times; however, the shape of the distribution would be the same as the distribution of response times to yoked control items. Slower mean response times to homophones are often interpreted as linear effects, but they are not. Instead, a nonlinear interaction stretches the slow end of the distribution of homophone response times. The shape of distribution is changed by this stretching, and it is this stretching, and this stretching alone, that produces the slower mean response times. This is a nonlinear skewing effect. We propose the skewing is a consequence of critical slowing, a signature of strongly nonlinear systems, produced by a combination of internal and external perturbations to the system. The next section describes our critical slowing hypothesis (in the terms of dynamical systems theory).

Perturbations and Skewing

Two sources of perturbation exist for a dynamical system: external perturbations and internal perturbations. These sources of variability, in combination, may explain the skewing effect on distributions of homophone response times in the categorization task.

External perturbations to performance in the categorization task (the model system) are exclusively random sources of variability. Any measurement includes random sources of variability. Random external perturbations contribute changes in the dynamics of a system that are damped as the system settles into a response trajectory. Damping takes time, and the time associated with random perturbations accumulates in the variability of response time distributions.

Internal perturbations are produced by small changes in the control parameter(s) of the model system(Latash, 1993). Internal perturbations are also transient but on a slower time scale than external perturbations. External perturbations are damped within a categorization trial, but internal perturbations set the value of a control parameter that may remain constant for the duration of a categorization trial.

Categorization performance emerges from the interaction of participants and the task environment. Thus, the model system of categorization performance strictly couples the participant and the task environment; generic patterns of organism–environment interactions are its least common denominators (see also Flach & Holden, 1998). This implies that unsystematic trial-by-trial changes in category names and target items actually produce internal perturbations. They produce small changes inside the model system and thereby set the value of the control parameter.

Consider trials where exemplars are presented (e.g., a *flower*::*TULIP*). Exemplars entail constraints that favor coherence with the category context, and errors to these trials are unlikely. Likewise, errors are unlikely to nonexemplar controls (e.g., a *flower*::*ROBS*)—on the whole, nonexemplar trials produce constraints that resist coherence with context. Thus, trial-by-trial changes in category breadth, typicality, target frequency, as well as other factors not systematically controlled in the experiment, set the control parameter relatively far from the bifurcation point. By analogy, small, unplanned changes in the controlled flow of current between swimming trials, or small, between-trial changes in the swimmer's stroke, would result in relatively contained sources of variability in the breast stroke or treading water. Response trajectories of exemplar (*TULIP*) and control (*ROBS*) items settle at values of the control parameter far enough away from the bifurcation point. Consequently, their response time distributions are less skewed (compared to homophone trials).

In contrast, responses to homophone trials may entail a more even mix of constraints that favor coherence with context, and constraints that resist coherence with context. So far, we have emphasized the contribution of category breadth and homophone word frequency to this mix (e.g., "a hardy swimmer is not forced to the breast stroke in a weak current"). We also mentioned that matching and mismatching constraints across any dimensions of the system combine in the ratio of constraints that favor or oppose coherence with context.

For example, mismatching constraints attendant on the mismatch in spelling between *rows* and *rose* (and mismatching constraints from other factors) oppose resonance with category context. The configuration and magnitude of matching and mismatching constraints will vary from homophone trial to homophone trial (along with the other previously mentioned factors that vary unsystematically in exemplar and control trials). Thus, homophone trials induce a larger range of change in the control parameter—the ratio of forces that favor coherence with context versus forces that resist coherence with context.

As we suggested previously, the value of the control parameter is sometimes sufficient to produce an incorrect *yes* response (*ROWS* is categorized as a *flower*). On other trials, the value of the control parameter produces a fast, correct *no* response. "In between" these values, homophone trials produce values closer to the bifurcation point than on exemplar (*TULIP*) or control (*ROBS*) trials. When a control parameter is set close to the bifurcation point (on one side or the other), the effect of external perturbations is amplified nonlinearly. The system requires an exaggerated amount of time to settle into a response trajectory.

Strongly nonlinear dynamical systems settle more slowly when they are close to bifurcation points. This inherent feature of nonlinear dynamics is called *critical slowing*. The portion of trials in which critical slowing occurs produces the nonlinear skewing pattern in the distributions of homophone response times.

Conventional Explanations

In addition to the empirical motivation for the identity bifurcation hypothesis, it demonstrates greater utility than conventional explanations of homophone errors. For example, this hypothesis, plus standard assumptions about variability, accommodates the appearance and occasional absence of homophone errors, the probabilistic nature of homophone errors (when they are observed), as well as the skewing effect in distributions of homophone response times. (It also naturally extends to other related interaction effects of context and word identity, that are reported in Farrar, 1998; Farrar et al., 1998; Gottlob et al., 1998; Potter et al., 1993; and Potter et al., 1998) It is hard to imagine how a conventional view of word identification could have anticipated these aspects of homophone errors.

Conventional theories include a strict distinction between the time before and the time after word identification (cf. *prelexical* vs. *postlexical* and *access* vs. *recognition*, respectively), in contrast to this approach that emphasizes continuous interaction. Given the conventional distinction, it does not make sense to suggest that a trailing phrase, which appeared after word identification was achieved, produced its effect prelexically, prior to word identification. The source of homophone errors must therefore be postlexical (based on the a priori assumption that it is useful to distinguish prelexical access processes from postlexical recognition processes). To preserve this distinction, conventional accounts may only elaborate on existing hypotheses or discount a result as a task artifact. Here we consider each type of explanation respectively, a memory hypothesis and an artifact hypothesis.

Perhaps spelling representations and semantic representations of homophones were occasionally lost from short-term memory prior to the appearance of the phrase. About 20% of the time in the 200-msec SOA condition, the uncertainty attributable to maintaining structures in memory (for over 300 msec) caused reaccess of long-term, lexical memory, after the phrase was presented. Reaccess was not based on spelling representations, or else homophones would have produced no more errors than controls. Thus, both spelling and semantic representations must have been lost prior to reaccess, because either of these representations would have ensured a correct response. Reaccess must be based on postlexical representations of phonology, but homophone phonology is ambiguous. Consequently, reaccess may be biased by the phrase context (*part of a car*) toward the false sound-alike identity (*brake*). This explanation in terms of a failure of short-term memory respects the distinction between prelexical and postlexical processes. And, it preserves a conventional no-phonology hypothesis concerning the identification of high-frequency homophones (and all other high-frequency words; Jared & Seidenberg, 1991).

A failure of short-term memory could explain the homophone errors in this phrase-evaluation task, but it is not motivated by memory theory in general. Representations in visual short-term memory are thought to survive up to 5 sec (Irwin, 1991) and are less susceptible to pattern masking than more peripheral visual representations (Carlson-Radvansky & Irwin, 1995). Additionally, there is related evidence for the retention of graphemic information for seconds (Kirsner, 1973; Scarborough, Cortese, & Scarborough, 1977) to minutes (Hock, Throckmorton, Webb, & Rosenthal, 1981) and even up to a year (Kolers, 1977). It is unlikely that spelling representations are lost after only 300 msec. Furthermore, letter representations become more salient when a less-frequent or subordinate meaning for a target word is elicited (Moravcsik & Healy, 1995). Phrase-evaluation errors entail an unusual meaning of a homophone (the meaning of its mate), and letter representations in memory should therefore become more salient, not less salient. Finally, a failure of short-term memory is contradicted by the results of our previous word-identification task. Presumably, participants achieved adequate short-term memory of spelling representations to write down words that are no longer present, and their performance was at ceiling.

Perhaps, as Hock et al. (1981) suggested, the translation of graphemic into phonologic representations makes graphemic representations less available for memory processes. Nevertheless, spelling similarity between stimulus homophones (*BREAK*) and their absent mates (*brake*), and participants' familiarity with spelling of the absent mate (*brake*), both affect the likelihood of homophone errors in categorization experiments (Van Orden, 1987). Why would spelling representations affect homophone identification for categorization, but not for phrase evaluation? The only previous categorization experiments in which spelling effects were not observed used a threshold SOA between target and mask. In these conditions, only phonology effects were observed (Peter & Turvey, 1994; Experiment 3, Van Orden, 1987; cf. Berent & Perfetti, 1995; Berent & Van Orden, 1998). Phonology effects under threshold SOA conditions suggest prelexical phonology. Postlexical phonology is superfluous.

Now we have come full circle. Prelexical phonology cannot be the basis of homophone errors in the phrase-evaluation task. We cannot explain the effects of prelexical homophone phonology and postlexical phrase context as an interaction between component processes. Prelexical components of word identification and postlexical components of phrase evaluation never overlap in time. Instead we are left with a quirky failure of short-term memory in the laboratory context of homophones judged against phrases appropriate to their mates.

We may also construct an idiosyncratic explanation, specific to the phrase-evaluation task. For example, consider Jared and Seidenberg's (1991) explanation of homophone errors to narrow categories presented prior to the homophones (i.e., *part of a car*::*BREAK*; but see Bosman & de Groot, 1996; Lesch & Pollatsek, 1993; Nielson, 1991; Rayner, Pollatsek, & Binder, 1998, for counter- proposals). Narrow categories may provide a narrow basis for expectations concerning the upcoming target word. If these expectations take the form of phonologic representations, then a match between expected *brake* (*/breyk/*) phonology and target *BREAK* phonology could generate homophone errors. Thus, the phonology effect is simply a task artifact. Jared and Seidenberg claimed that broad categories (*object*::*BREAK*) eliminate the artifact, as they do not yield homophone errors. We replicated the latter finding in the categorization experiment reported here.

An equally idiosyncratic explanation could be offered for these results. Perhaps, at some time within the 300 msec between onset of the target and onset of the phrase, activation from the postlexical phonologic representation of the target spreads to a semantic network and activates all semantic associates of *BREAK* and *brake* (compare O'Seaghdha & Marin, 1997). Associates and pseudoassociates may include words that appear in the subsequent phrase (e.g., *part of a car*). About 20% of the time, pseudoassociates (i.e., associates of *brake*) are activated that appear in the trailing phrase. The match between pseudoassociates and the phrases is the artifactual source of homophone errors.

The problem with explanations of this type is that they lack utility outside of specific task contexts. Slightly different task contexts may require entirely different explanations. For example, we require different explanations for the null homophone effect in the 43-msec condition of the categorization task versus the 50-msec condition of the phrase-evaluation task or the effects of spelling similarity between homophones (Van Orden, 1987) or the exemplar-frequency effect of the nonpresented homophone (Van Orden, 1987, 1991; Van Orden et al., 1992), etc. This is not a useful or acceptable outcome for a scientific research program (Lakatos, 1970; Putnam, 1994; Van Orden, Pennington, & Stone, 1998). In the end, we are left with a hodgepodge of idiosyncratic explanations, each dealing with its own brand of homophone error.

Idiosyncratic explanations of homophone errors serve no purpose other than to protect theories that deemphasize phonology in reading. These explanations implicitly accept that the broad categorization task (*object*::*BREAK*) is transparent, but all other relevant contexts are contaminated by task idiosyncrasies. Bradley and Forster (1987) anticipate this outcome. They warn against too freely couching explanations in postaccess, recognition processes, because this type of explanation will protect indefinitely any theory of lexical access. "Any unwelcome facts about language performance can be attributed to recognition" (p. 110). According to Bradley and Forster, this use of the terms *access* (prelexical) versus *recognition* (postlexical) renders the distinction vacuous and justifies that we abandon the distinction altogether.

The results of our experiments, with those of Jared and Seidenberg (1991) and Van Orden (1987), have failed to identify a source of homophone errors that operates independently of context. This is not to say that we have falsified alternative accounts of homophone errors in favor of our account. Critical tests may only falsify hypotheses locally, within a specific research program. They cannot distinguish between incommensurate programs (Kuhn, 1970; Lakatos, 1970). A research program that axiomatically seeks to minimize context effects and isolate temporally ordered components is incommensurate with a research program that emphasizes context effects and continuous real-time interaction. Moreover, although local tests of a hypothesis may fail, one may always resurrect the more fundamental core assumptions behind the hypothesis (Duhem, 1906/1954; Einstein & Infeld, 1938/1966; Lakatos, 1970; Quine, 1953/1961). We illustrated this fact with the previous artifact explanations (cf. Gibbs & Van Orden, 1998; Stone & Van Orden, 1993). What our results question is whether conventional analyses may ever provide useful explanations of homophone errors.

Resolving the Context Debate

The charge of scientific psychology is to provide simplifying explanations that cut across as many laboratory contexts as possible. The conventional answer to this charge has been to seek context-free effects—that is, the same effects regardless of task environment (cf. Jacobs & Grainger, 1994; Sternberg, 1969). This answer relies on the existence of objective (transparent) task environments from which to abstract a simplifying cognitive architecture.[7] This could have been an elegant solution, except that no context-free effects were discovered. Paradoxically, a context-specific explanation is posited for each context-specific effect. As context-specific "strategies" (artifacts) multiply, the number of explanations (representations + processes + strategies) tends to grow as fast as the number of phenomena to be explained (Van Orden et al., 1996). Furthermore, as the number of interactions with context grows, it becomes incredible to insist that some particular task environment provides a transparent view of the role of phonology in reading, or of any other cognitive process for that matter.

Patterns of homophone errors are cited as evidence for, and evidence against, a general role for phonology in reading. The presence of homophone errors is a phonology effect, and the absence of homophone errors is a null phonology effect. As Frost (1998) points out, investigators who emphasize the role of phonology in reading favor results from task environments that produce reliable phonology effects, but investigators who deemphasize phonology favor task environments that produce reliable null effects (see also Van Orden et al., 1996). This debate concerning

[7]The idea of transparent task contexts seems somewhat parallel to the problematic idea of an inertial coordinate system in physics, which was finally obviated by relativity theory (Einstein & Infeld, 1938/1966).

context and homophone errors, with which we began, represents, in microcosm, a vast contemporary debate in all areas of cognitive psychology. All traditional cognitive analyses emphasize the discovery of context-free effects. Only the existence of context-free effects allows for the dissociation of cognition from the contexts of performance. No context-free effects have been discovered, however. Task contexts reliably modulate the effects of all contributing cognitive variables, which forces the question, Which contexts are transparent to underlying cognitive components and which are not?

We speculate that the larger context debate reveals the fundamental limits and consequent failure of traditional cognitive analysis. Context-free effects are never discovered, because they do not exist. It is impracticable to isolate cognition from the contexts in which it is embedded (see also Van Orden et al., 1997; Van Orden et al., 1996; Van Orden et al., 1998; Van Orden & Paap, 1997; cf. Flach & Holden, 1998; Mandler, 1997; Shanon, 1993; Turvey & Carello, 1981, 1995; Turvey, Shaw, Reed, & Mace, 1981; Uttal, 1998; Watkins, 1990). Cognitive systems are causally embedded in their environments and thus always entail their environments with regard to cognitive performance.

The way out of the context debate is to choose a different goal, one that takes interaction effects, and thereby context, as a basis for cognitive theory. This general working assumption led us to the identity bifurcation hypothesis, which we introduced by analogy to swimming. As we apply the general assumption, context plays an integral role in explanation (cf. Shaw & Turvey, 1981). Rather than seeking to avoid context effects and assuming that we may eventually hit on transparent task contexts, we choose to put context to work both empirically and theoretically. To explore the role of phonology in reading, we manipulate context to reveal patterns of interaction.[8] To understand the role of phonology in reading, we propose that context-induced phonology effects and their occasional context-induced absence imply a context-sensitive interactive system (Carello, Turvey, & Lukatela, 1992; Van Orden, Pennington, & Stone, 1990). What swimming says about reading is this: It makes as little sense to speak of word identities and phonology outside of a context of discourse as to speak of swimming outside of a context of water or gravity.

ACKNOWLEDGMENTS

This research was funded by a First Independent Research Support and Transition Award (CMS 5 R29 NS26247) and an Independent Scientist Award, National In-

[8]Roger Schvaneveldt, with David Meyer, discovered semantic priming, perhaps the most well-known context effect in experimental psychology (Meyer & Schvaneveldt, 1971). Schvaneveldt's (personal communication, August 15, 1998) interest in context effects reflects, in part, his realization that context effects can be studied in a true experimental fashion, through manipulation. Many stimulus variables cannot be manipulated, only sampled, as when we try to study word frequency in psycholinguistics. As a result, more certain conclusions can be drawn about context effects.

stitute of Neurological Disorders and Stroke (1 K02 NS 01905), both awarded to Guy Van Orden. The Independent Scientist Award included training at the Center for the Ecological Study of Perception and Action (CESPA), where this article was drafted.

Guy Van Orden thanks his teachers Bruce Kay, Bob Shaw, and especially, Michael Turvey. We also acknowledge Claudia Carello, and faculty and students, for the combination of intellectual rigor and warm hospitality that is the CESPA way. Anny Bosman and Jay Rueckl provided helpful comments, and Marian Jansen op de Haar gave detailed constructive feedback on several versions of this article.

REFERENCES

Azuma, T., & Van Orden, G. C. (1997). Why SAFE is better than FAST: The relatedness of a word's meanings affects lexical decision times. *Journal of Memory and Language, 36,* 484–504.

Berent, I. (1997). Phonological effects in the lexical decision task: Regularity effects are not necessary evidence for assembly. *Journal of Experimental Psychology: Human Perception and Performance, 23,* 1–16.

Berent, I., & Perfetti, C. (1995). A rose is a REEZ: The two cycles model of phonology assembly in reading English. *Psychological Review, 102,* 146–184.

Berent, I., & Van Orden, G. C. (1998). *Homophone dominance modulates the phonemic-masking-effect.* Manuscript submitted for publication.

Bernstein, N. (1967). *The coordination and regulation of movements.* London: Pergamon.

Bosman, A. M. T., & de Groot, A. M. B. (1996). Phonologic mediation is fundamental to reading: Evidence from beginning readers. *Quarterly Journal of Experimental Psychology, 49A,* 715–744.

Bosman, A. M. T., & Van Orden, G. C. (1997). Why spelling is more difficult than reading. In C. A. Perfetti, L. Rieben, & M. Fayol (Eds.), *Learning to spell* (pp. 173–194). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Bradley, D. C., & Forster, K. I. (1987). A reader's view of listening. *Cognition, 25,* 103–134.

Briand, K., den Heyer, K., & Dannenbring, G. L. (1988). Retroactive semantic priming in a lexical decision task. *Quarterly Journal of Experimental Psychology, 40A,* 341–359.

Carello, C., Turvey, M. T., & Lukatela, G. (1992). Can theories of word recognition remain stubbornly nonphonological? In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 211–226). Amsterdam: North-Holland.

Carlson-Radvansky, L. A., & Irwin, D. E. (1995). Memory for structural information across eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1441–1458.

Carr, T., Davidson, B., & Hawkins, H. (1978). Perceptual flexibility in word recognition: Strategies affect orthographic computation but not lexical access. *Journal of Experimental Psychology: Human Perception and Performance, 4,* 647–690.

Carr, T. H., & Pollatsek, A. (1985). Recognizing printed words: A look at current models. In D. Besner, T. G. Waller, & G. E. MacKinnon (Eds.), *Reading research: Advances in theory and practice 5* (pp. 1–82). San Diego, CA: Academic.

Coltheart, M., & Rastle, K. (1994). Serial processing in reading aloud: Evidence for dual-route models of reading. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 1197–1211.

Coltheart, V., Avons, S. E., Masterson, J., & Laxon, V. (1991). The role of assembled phonology in reading comprehension. *Memory & Cognition, 19,* 387–400.

Coltheart, V., Avons, S. E., & Tollope, J. (1990). Articulatory suppression and phonological codes in reading for meaning. *Quarterly Journal of Experimental Psychology, 42A,* 375–399.

Coltheart, V., Laxon, V., Rickard, M., & Elton, C. (1988). Phonological recoding in reading for meaning by adults and children. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 387–397.

Coltheart, V., Patterson, K., & Leahy, J. (1994). When a ROWS is a ROSE: Phonological effects in written word comprehension. *Quarterly Journal of Experimental Psychology, 47A*, 917–955.

Doctor, E. A., & Coltheart, M. (1980). Children's use of phonological encoding when reading for meaning. *Memory & Cognition, 8*, 195–209.

Duhem, P. (1954). *Aim and structure of physical theory*. York, England: Antheneum. (Original work published 1906)

Einstein, A., & Infeld, L. (1966). *The evolution of physics*. New York: Simon & Schuster. (Original work published 1938)

Farmer, J. D. (1990). A Rosetta Stone for connectionism. *Physica D, 42*, 153–187.

Farrar, W. T. (1998). Investigating single-word syntactic primes in naming tasks: A recurrent network approach. Journal of *Experimental Psychology: Human Perception and Performance, 24*, 648–663.

Farrar, W. T., & Van Orden, G. C. (1998). *Nonlinear dynamics of naming errors*. Manuscript submitted for publication.

Farrar, W. T., Van Orden, G. C., & Hamouz, V. (1998). *When SOFA primes TOUCH: Interdependence of spelling, phonology and meaning in "semantically-mediated" priming*. Manuscript submitted for publication.

Flach, J. M., & Holden, J. G. (1998). The reality of experience: Gibson's way. *Presence, 7*, 90–95.

Forster, K. I. (1974). The role of semantic hypotheses in sentence processing. In F. Bresson & J. Mehler (Eds.), *Current problems in psycholinguistics* (pp. 391–408). Paris: Editions du CRNS.

Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior, 12*, 627–635.

Frost, R. (1998). Toward a strong phonological model of reading: True issues and false trails. *Psychological Bulletin, 123*, 71–99.

Gibbs, P. (1996). Strategic control of nonlexical effects in word recognition: Testing the utility of pathway selection. *Dissertation Abstracts International, 57*(07B), 4734.

Gibbs, P., & Van Orden, G. C. (1998). Pathway selection's utility for control of word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 1162–1187.

Gottlob, L. R., Goldinger, S. D., Stone, G. O., & Van Orden, G. C. (in press). Reading homographs: Orthographic, phonologic, and semantic dynamics. *Journal of Experimental Psychology: Human Perception and Performance.*

Grossberg, S., & Stone, G. O. (1986). Neural dynamics of word recognition and recall: Priming, learning, and resonance. *Psychological Review, 93*, 46–74.

Haken, H. (1988). *Information and self-organization*. Berlin: Springer-Verlag.

Hawkins, H., Reicher, G., & Peterson, L. (1976). Flexible coding in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 2*, 380–385.

Hock, H. S., Throckmorton, B., Webb, E., & Rosenthal, A. (1981). The effect of phonemic processing on the retention of graphemic representations for words and nonwords. *Memory & Cognition, 9*, 461–471.

Hooper, D. A., & Paap, K. R. (1997). The use of assembled phonology during performance of a letter recognition task and its dependence on the presence and proportion of word stimuli. *Journal of Memory and Language, 37*, 167–189.

Irwin, D. E. (1991). Information integration across saccadic eye movements. *Cognitive Psychology, 23*, 420–456.

Jacobs, A. M., & Grainger, J. (1994). Models of visual word recognition—Sampling the state of the art. *Journal of Experimental Psychology: Human Perception and Performance, 20*, 1311–1334.

Jacobson, J. Z., & Rhinelander, G. (1978). Geometric and semantic similarity in visual masking. *Journal of Experimental Psychology: Human Perception and Performance, 4*, 224–231.

Jared, D. (1997). Spelling–sound consistency affects the naming of high-frequency words. *Journal of Memory and Language, 36*, 505–529.

Jared, D., & Seidenberg, M. S. (1991). Does word identification proceed from spelling to sound to meaning? *Journal of Experimental Psychology: General, 120*, 358–394.

Kelso, J. A. S. (1995). *Dynamic patterns: The self-organization of brain and behavior*. Cambridge, MA: MIT Press.

Kiger, J. I., & Glass, A. L. (1983). The facilitation of lexical decisions by a prime occurring after the target. *Memory & Cognition, 11*, 356–365.

Kirsner, K. (1973). An analysis of the visual component in recognition memory for verbal stimuli. *Memory & Cognition, 1*, 449–453.

Kolers, P. A. (1977). Pattern-analyzing memory. *Science, 191*, 1280–1281.

Kučera, H., & Francis, W. (1967). *Computational analysis of present-day American English.* Providence, RI: Brown University Press.

Kugler, P. N., & Turvey, M. T. (1987). *Information, natural law, and the self-assembly of rhythmic movement.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Kuhn, T. S. (1970). *The structure of scientific revolutions.* Chicago: University of Chicago Press.

Lakatos, I. (1970). Falsification and the methodology of scientific research programmes. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the growth of knowledge* (pp. 91–195). London: Cambridge University Press.

Latash, M. L. (1993). *Control of human movement.* Champaign, IL: Human Kinetics.

Lesch, M. F., & Pollatsek, A. (1993). Automatic access of semantic information by phonological codes in visual word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 285–294.

Lewenstein, M., & Nowak, A. (1989a). Fully connected neural networks with self-control of noise levels. *Physical Review Letters, 62*, 225–228.

Lewenstein, M., & Nowak, A. (1989b). Recognition with self-control in neural networks. *Physical Review A, 40*, 4652–4664.

Lukatela, G., Feldman, L. B., Turvey, M. T., Carello, C., & Katz, L. (1989). Context effects in bi-alphabetical word perception. *Journal of Memory and Language, 28*, 214–236.

Lukatela, G., & Turvey, M. T. (1987). Loci of phonological effects in the lexical access of words written in a shallow orthography. *Psychological Research, 49*, 139–146.

Lukatela, G., & Turvey, M. T. (1994). Visual lexical access is initially phonological: 2. Evidence from phonological priming by homophones and pseudohomophones. *Journal of Experimental Psychology: General, 123*, 331–353.

Lupker, S. J., Brown, P., & Colombo, L. (1997). Strategic control in a naming task: Changing routes or changing deadlines? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 570–590.

Mace, W. M. (1977). James J. Gibson's strategy for perceiving: Ask not what's inside your head, but what your head's inside of. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting, and knowing* (pp. 43–65). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Mandler, G. (1985). *Cognitive psychology: An essay in cognitive science.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Mandler, G. (1997). *Human nature explored.* New York: Oxford University Press.

Massaro, D. W., Taylor, G. A., Venezky, R. L., Jastrzembski, J. E., & Lucas, P. A. (1980). *Letter and word perception.* Amsterdam: North-Holland.

Masson, M. E. J. (1995). A distributed memory model of semantic priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 509–514.

Mattingly, I. G. (1992). Linguistic awareness and orthographic form. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 11–26). Amsterdam: North-Holland.

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90*, 227–234.

Moravcsik, J. E., & Healy, A. F. (1995). Effect of meaning on letter detection. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 82–95.

Nielson, C. S. (1991). *Phonology affects identification of high frequency printed words.* Unpublished honors thesis, Honors College, Arizona State University, Tempe.

O'Seaghdha, P. G., & Marin, J. W. (1997). Mediated semantic–phonological priming: Calling distant relatives. *Journal of Memory and Language, 36*, 226–252.

Peitgen, H.-O., Jürgens, H., & Saupe, D. (1992). *Chaos and fractals: New frontiers of science.* New York: Springer-Verlag.

Perfetti, C. A., & Zhang, S. (1995). The universal word identification reflex. In D. L. Medin (Ed.), *The psychology of learning and motivation* (Vol. 33, pp. 159–189). San Diego: Academic.

Perfetti, C. A., Zhang, S., & Berent, I. (1992). Reading in English and Chinese: Evidence for a "universal" phonological principle. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 227–248). Amsterdam: North-Holland.

Peter, M., & Turvey, M. T. (1994). Phonological codes are early sources of constraint in visual semantic categorization. *Perception & Psychophysics, 55*, 497–504.

Potter, M. C., Moryadas, A., Abrams, I., & Noel, A. (1993). Word perception and misperception in context. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 3–22.

Potter, M. C., Stiefbold, D., & Moryadas, A. (1998). Word selection in reading sentences: Preceding versus following contexts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*, 68–100.

Putnam, H. (1994). The Dewey lectures: Sense, nonsense, and the senses: An inquiry into the powers of the human mind. *The Journal of Philosophy, 91*, 445–517.

Quine, W. V. O. (1961). Two dogmas of empiricism. In W. V. O. Quine (Ed.), *From a logical point of view* (pp. 20–46). New York: Harper & Row. (Original work published in 1953)

Rapp, B., & Caramazza, A. (1997). From graphemes to abstract letter shapes: Levels of representation in written spelling. *Journal of Experimental Psychology: Human Perception and Performance, 23*, 1130–1152.

Rayner, K. (1993). Eye movements in reading: Recent developments. *Current Directions in Psychological Science, 2*, 81–85.

Rayner, K., & Pollatsek, A. (1989). *The psychology of reading.* Englewood Cliffs, NJ: Prentice Hall.

Rayner, K., Pollatsek, A., & Binder, K. S. (1998). Phonological codes and eye movements in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*, 476–497.

Rayner, K., Sereno, S. C., Lesch, M. F., & Pollatsek, A. (1995). Phonological codes are automatically activated during reading: Evidence from an eye movement priming paradigm. *Psychological Science, 6*, 26–32.

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review, 105*, 125–157.

Rueckl, J. G., & Oden, G. C. (1986). The integration of contextual and featural information during word identification. *Journal of Memory and Language, 25*, 445–460.

Sakuma, N., Sasanuma, S., Tatsumi, I. F., & Masaki, S. (1998). Orthography and phonology in reading Japanese kanji words: Evidence from the semantic decision task with homophones. *Memory & Cognition, 26*, 75–87.

Saltzman, E. L., & Kelso, J. A. S. (1987). Skilled actions: A task dynamic approach. *Psychological Review, 94*, 84–106.

Saltzman, E. L., & Munhall, K. G. (1992). Skill acquisition and development: The roles of state-, parameter-, and graph-dynamics. *Journal of Motor Behavior, 24*, 49–57.

Scarborough, D. L., Cortese, C., & Scarborough, H. S. (1977). Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology: Human Perception and Performance, 3*, 1–17.

Seidenberg, M. S. (1985). The time course of phonological code activation in two writing systems. *Cognition, 19*, 1–30.

Sereno, S. C. (1995). Resolution of lexical ambiguity: Evidence from an eye movement priming paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 582–595.

Sereno, S. C., & Posner, M. I. (1995, November). *Early lexical processing: Evidence from eye movements and ERPs.* Poster session presented at the annual meeting of the Psychonomic Society, Los Angeles, CA.

Shanon, B. (1993). *The representational and the presentational: An essay on cognition and the study of the mind.* New York: Harvester Wheatsheaf.

Shaw, R., & Turvey, M. T. (1981). Coalitions as models for ecosystems: A realist perspective on perceptual organization. In M. Kubovy & J. Pomerantz (Eds.), *Perceptual organization* (pp. 343–415). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Skarda, C. A., & Freeman, W. J. (1987). How brains make chaos in order to make sense of the world. *Behavioral and Brain Sciences, 10*, 161–195.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica, 30*, 276–315.

Stewart, I., & Cohen, J. (1997). *Figments of reality*. New York: Cambridge University Press.

Stone, G. O., & Van Orden, G. C. (1989). Are words represented by nodes? *Memory & Cognition, 17*, 511–524.

Stone, G. O., & Van Orden, G. C. (1993). Strategic processes in printed word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 19*, 744–774.

Stone, G. O., & Van Orden, G. C. (1994). Building a resonance framework for word recognition using design and system principles. *Journal of Experimental Psychology: Human Perception and Performance, 20*, 1248–1268.

Strain, E., Patterson, K., & Seidenberg, M. S. (1995). Semantic effects in single-word naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 1140–1154.

Swinney, D. A. (1979). Lexical access during sentence comprehension: Re(consideration) of context effects. *Journal of Verbal Learning and Verbal Behavior, 18*, 645–659.

Taft, M., & van Graan, F. (1998). Lack of phonological mediation in a semantic categorization task. *Journal of Memory and Language, 38*, 203–224.

Tan, L. H., & Perfetti, C. A. (1997). Visual Chinese character recognition: Does phonological information mediate access to meaning? *Journal of Memory and Language, 37*, 41–57.

Thelen, E. (1995). Motor development. *American Psychologist, 50*, 79–95.

Thelen, E., Fisher, D. M., & Ridley-Johnson, R. (1984). The relationship between physical growth and a newborn reflex. *Infant Behavior and Development, 7*, 479–493.

Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT Press.

Turvey, M. T. (1990a). The challenge of a physical account of action: A personal view. In H. T. A. Whiting, O. G. Meijer, & P. C. W. van Wieringen (Eds.), *The natural-physical approach to movement control* (pp. 57–93). Amsterdam: Vrij Universiteit Press.

Turvey, M. T. (1990b). Coordination. *American Psychologist, 45*, 938–953.

Turvey, M. T., & Carello, C. (1981). Cognition: The view from ecological realism. *Cognition, 10*, 313–321.

Turvey, M. T., & Carello, C. (1995). Some dynamical themes in perception and action. In R. F. Port and T. van Gelder (Eds.), *Mind as motion: Explorations in the dynamics of cognition* (pp. 373–401). Cambridge, MA: MIT Press.

Turvey, M. T., Fitch, H. L., & Tuller, B. (1982). The Bernstein perspective: 1. The problems of degrees of freedom and context-conditioned variability. In J. A. S. Kelso (Ed.), *Understanding human motor control* (pp. 239–252). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Turvey, M. T., Shaw, R. E., Reed, E. S., & Mace, W. M. (1981). Ecological laws of perceiving and acting: In reply to Fodor and Pylyshyn (1981). *Cognition, 9*, 237–304.

Uttal, W. R. (1998). *Toward a new behaviorism: The case against perceptual reductionism*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Van Orden, G. C. (1987). A ROWS is a ROSE: Spelling, sound, and reading. *Memory & Cognition, 15*, 181–198.

Van Orden, G. C. (1991). Phonologic mediation is fundamental to reading. In D. Besner & G. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 77–103). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Van Orden, G. C., Aitchison, C. S., & Podgornik, M. N. (1996). *When a ROWS is not a ROSE: Null effects and the absence of cognitive structures*. Manuscript submitted for publication.

Van Orden, G. C., & Goldinger, S. D. (1994). Interdependence of from and function in cognitive systems explains perception of printed words. *Journal of Experimental Psychology: Human Perception and Performance, 20*, 1269–1291.

Van Orden, G. C., Jansen op de Haar, M. A., & Bosman, A. M. T. (1997). Complex dynamic systems also predict dissociations, but they do not reduce to autonomous components. *Cognitive Neuropsychology, 14,* 131–165.

Van Orden, G. C., Johnston, J. C., & Hale, B. L. (1988). Word identification in reading proceeds from spelling to sound to meaning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 371–385.

Van Orden, G. C., & Paap, K. R. (1997). Functional neuroimages fail to discover pieces of mind in the parts of the brain. *Philosophy of Science, 64,* S85–S94.

Van Orden, G. C., Pennington, B. F., & Stone, G. O. (1990). Word identification in reading and the promise of subsymbolic psycholinguistics. *Psychological Review, 97,* 488–522.

Van Orden, G. C., Pennington, B. F., & Stone, G. O. (1998). *What do double dissociations prove? Modularity yields a degenerating research program.* Manuscript submitted for publication.

Van Orden, G. C., Stone, G. O., Garlington, K. L., Markson, L. R., Pinnt, G. S., Simonfy, C. M., & Brichetto, T. (1992). "Assembled" phonology and reading: A case study in how theoretical perspective shapes empirical investigation. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 249–292). Amsterdam: North-Holland.

Verstaen, A., Humphreys, G., Olson, A., & D'Ydewalle, G. (1995). Are phonemic effects in backward masking evidence for automatic prelexical phonemic activation in visual word recognition? *Journal of Memory and Language, 34,* 335–356.

Watkins, M. J. (1990). Mediationism and the obfuscation of memory. *American Psychologist, 45,* 328–335.

Weber, R. (1970). A linguistic analysis of first-grade reading errors. *Reading Research Quarterly, 5,* 427–451.

Whaley, C. P. (1978). Word–nonword classification time. *Journal of Verbal Learning and Verbal Behavior, 17,* 143–154.

Whittlesea, B. W. A., & Jacoby, L. L. (1990). Interaction of prime repetition with visual degradation: Is priming a retrieval phenomenon? *Journal of Memory and Language, 29,* 546–565.

Wydell, T. N., Patterson, K. E., & Humphreys, G. W. (1993). Phonologically mediated access to meaning for Kanji: Is a *Rows* still a *Rose* in Japanese Kanji? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 491–514.

Xu, B., & Perfetti, C. A. (in press). Nonstrategic subjective threshold effects in phonemic masking. *Memory & Cognition.*

Ziegler, J. C., Montant, M., & Jacobs, A. M. (1997). The feedback consistency effect in lexical decision and naming. *Journal of Memory and Language, 37,* 533–554.

Ziegler, J. C., Van Orden, G. C., & Jacobs, A. M. (1997). Phonology can help or hurt the perception of print. *Journal of Experimental Psychology: Human Perception and Performance, 23,* 845–860.

APPENDIX A

Homophone and Control Items and Their Mean Percentage of Incorrect *Yes* Responses in the Categorization Experiment

	43 msec SOA				*200 msec SOA*			
	Homophone		*Control*		*Homophone*		*Control*	
Category	*Word*	*% Incorrect Yes Responses*	*Word*	*% Incorrect Yes Responses*	*Word*	*% Incorrect Yes Responses*	*Word*	*% Incorrect Yes Responses*
Living thing	BARE	35	BEER	60	BARE	10	BEER	25
Living thing	MADE	25	MILD	25	MADE	0	MILD	5
Living thing	MIGHT	10	MILE	25	MIGHT	0	MILE	5
Living thing	NONE	15	NOON	20	NONE	0	NOON	10
Living thing	SUN	55	SIN	20	SUN	70	SIN	5
Living thing	BE	45	BYE	15	BE	0	BYE	0
Living thing	MAIL	20	MALL	30	MAIL	10	MALL	0
Living thing	NIGHT	10	KNIFE	20	NIGHT	0	KNIFE	15
Living thing	SELL	40	CALL	15	SELL	0	CALL	5
Object	GREAT	15	GRACE	25	GREAT	10	GRACE	0
Object	PALE	25	PAID	20	PALE	25	PAID	0
Object	SALE	25	SAID	20	SALE	35	SAID	0
Object	SHOOT	10	CHEAT	15	SHOOT	10	CHEAT	5
Object	BREAK	15	BROKE	25	BREAK	10	BROKE	20
Object	MEET	30	MELT	15	MEET	0	MELT	15
Object	PLAIN	40	PHASE	35	PLAIN	15	PHASE	30
Object	SENT	25	COUNT	10	SENT	0	COUNT	10
Object	THROWN	20	TYRONE	55	THROWN	5	TYRONE	15

Note. SOA = stimulus-onset asynchrony.

APPENDIX B

Homophone and Control Items and Their Mean Percentage of Incorrect *Yes* Responses in the 50 msec and 200 msec Target-Mask SOA Conditions of the Phrase Evaluation Experiment

50 msec SOA				*200 msec SOA*				
Homophone		*Control*		*Homophone*		*Control*		
Word	*% Incorrect Yes Responses*	*Word*	*% Incorrect Yes Responses*	*Word*	*% Incorrect Yes Responses*	*Word*	*% Incorrect Yes Responses*	*Comparison Phrase*
DIE	70	DUE	30	DIE	20	DUE	10	To color
DEAR	40	DOER	10	DEAR	50	DOER	10	Animal
MEET	30	MOAT	0	MEET	0	MOAT	0	Food
SUN	30	SIN	10	SUN	20	SIN	0	Male relative
BREAK	50	BRAVE	0	BREAK	50	BRAVE	0	Part of a car
HAIR	80	HERE	40	HAIR	30	HERE	30	Rabbit
SEEM	50	SCAM	40	SEEM	40	SCAM	10	Place where cloth is joined
SEE	50	SPA	10	SEE	20	SPA	0	Ocean
FAIR	10	FIRE	0	FAIR	20	FIRE	0	Cost of airline travel
BEAT	50	BELT	20	BEAT	20	BELT	0	Vegetable
WEEK	40	WALK	20	WEEK	20	WALK	0	Not strong
SENT	20	CART	0	SENT	10	CART	0	Penny
PALE	60	PILL	20	PALE	30	PILL	0	Bucket
FEET	60	FELT	20	FEET	20	FELT	10	Accomplishment
NIGHT	30	KNIFE	30	NIGHT	30	KNIFE	10	Medieval soldier
MADE	40	MELD	30	MADE	10	MELD	10	Servant who cleans
BEAR	70	BASE	30	BEAR	40	BASE	10	Naked
NONE	20	NUT	10	NONE	0	NUT	0	Religious female
CELL	30	SEAL	40	CELL	0	SEAL	0	Trade goods for money
MAIN	100	MACE	30	MAIN	40	MACE	20	Part of a horse

Note. SOA = stimulus-onset asynchrony.