The Nonlinear Dynamics of Speech Categorization

Article *in* Journal of Experimental Psychology Human Perception & Performance · March 1994

DOI: 10.1037/0096-1523.20.1.3 · Source: PubMed

Journal of Experimental Psychology:
Human Perception and Performance
1994, Vol. 20, No. 1, 3–16

0096-1523/94/$3.00

The Nonlinear Dynamics of Speech Categorization

Betty Tuller, Pamela Case, Mingzhou Ding, and J. A. Scott Kelso

Little is known about the processes underlying the nonlinear relationship between acoustics and speech perception. In Experiment 1, we explored the effects of systematic variation of a single acoustic parameter (silent gap duration between a natural utterance of *s* and a synthetic vowel *ay*) on judgments of speech category. The resulting shifts in category boundary between *say* and *stay* showed rich dynamics, including hysteresis, contrast, and critical boundary effects. We propose a dynamical model to account for the observed patterns. Experiment 2 evaluated one prediction of the model, that changing the relative stability of the two percepts allows categorical switching. In agreement with the model, an increase in the number of stimulus repetitions maximized the frequency of judgments of category change near the boundary. Thus, a dynamical approach affords the rudiments for a theory of the effects of temporal context on speech categorization.

The perception of speech exhibits an extremely important duality: It is remarkably stable under acoustic variation, but at the same time it can adjust flexibly to changing contexts. The stability of speech perception is exemplified by the phenomenon known as categorical perception, that is, the extreme difficulty humans have in discriminating between different acoustic stimuli that are labeled as the same speech segment (e.g., Liberman, Harris, Hoffman, & Griffith, 1957; see Repp, 1984, for a review). Outside the laboratory as well, a wide range of acoustic productions are categorized identically, such as when the same word is produced by different speakers. Nevertheless, the boundaries between categories are flexible, adjusting with factors such as phonetic context, the acoustic information available, speaking rate, and linguistic experience (see Repp & Liberman, 1987, for a recent synthesis).

These phenomena of speech perception and categorization have motivated a large body of work, published since the 1950s, in which researchers have dealt with several issues in speech perception and proposed mechanisms and models to account for them. Yet the basic issue of how humans sort a continuously changing signal into the proper categories is still incompletely understood, regardless of whether the categories concern speech, objects, emotions, or individuals (see, e.g., a recent compilation by Harnad, 1987). Our motivation for the present work stemmed from the realization that advances in understanding how patterns form in open, nonequilibrium systems may provide a productive conceptual framework within which to formulate ideas and embed results (e.g., Bergé, Pomeau, & Vidal, 1984; Haken, 1983b; Nicolis & Prigogine, 1989). As our point of departure, we advance the notion that changes in speech categorization that occur as the acoustic signal is altered are indicative of a pattern formation process in perception.[1] The concepts of pattern formation, particularly self-organization, and the mathematical tools of nonlinear dynamical systems are playing an increasingly visible role in psychology and biology (for a variety of different contributions in this context, see Haken & Stadler, 1990) and provide the backdrop to our experimental investigation of speech categorization.

We believe that perceptual categorization of speech provides strong hints of what we call *perceptual dynamics* (multistability, loss of stability, switching, etc.). Perceptual dynamics characterizes the time-dependent behavior of the speech system in terms of its (nonlinear) dynamics; that is, equations of motion describing the temporal evolution of the perceptual process, especially the stability and change of perceptual forms. This approach is analogous to the theoretical framework of *coordination dynamics* used to understand goal-directed movements (e.g., Kelso & Schöner, 1987; Schöner & Kelso, 1988b). Coordination dynamics does not refer to biomechanics per se (e.g., masses and stiffnesses of moving segments) but rather, in keeping with the present work, characterizes the spatiotemporal patterns of coordinated actions produced by the nervous system. Such a theme resonates with early Gestalt theory, which stressed the search for natural principles of autonomous order formation in perception and brain function (Köhler, 1940; for a historical treatment, see Stadler & Kruse, 1990). Recent advances in the theory of complex, nonlinear systems (e.g., Haken, 1983a) may provide the tools for exploring the phenomenological aspects of behavior emphasized by Gestaltists (Epstein & Hatfield, in press). A brief description of the main features of such sys-

Betty Tuller, Pamela Case, Mingzhou Ding, and J. A. Scott Kelso, Center for Complex Systems, Florida Atlantic University.

This research was supported by National Institute of Deafness and Communicative Disorders Grant DC00411, National Institute of Mental Health Grant MH42900, and Biomedical Research Support Grant RR07258.

We thank an anonymous reviewer for helpful comments on an earlier draft of this article.

Correspondence concerning this article should be addressed to Betty Tuller, Center for Complex Systems, Florida Atlantic University, P.O. Box 3091, Boca Raton, Florida 33431-0091. Electronic mail may be sent to tuller@walt.ccs.fau.edu.

[1] Our model recognizes that the boundaries of sensation, perception, and decision making are not clearly defined, and we use these terms interchangeably in this article.

tems follows that will provide strategic guidance to the present experiments and the theoretical model developed.

A defining feature of a nonlinear system is that when a control parameter is varied throughout a wide range, often no observable change in behavior results. However, when the control parameter reaches a critical point, behavior may change qualitatively or discontinuously. In an early example of such qualitative change in human behavioral coordination (Kelso, 1984), subjects had no trouble moving their wrists or index fingers rhythmically in an antiphase relationship (homologous muscle groups contracting in an alternating fashion) throughout a range of cycling frequencies. However, when the cycling frequency (the control parameter) increased past a critical frequency, subjects produced an involuntary and abrupt shift to an in-phase coordinative pattern (homologous muscle groups contracting synchronously). Thus, increases in a nonspecific variable, cycling frequency, produced switches from one (initially stable) coordinative pattern to another. The temporal evolution of this self-organized pattern formation process (including instabilities, bifurcations, and so on) is captured by low-dimensional equations of motion for the system's collective modes (Haken, Kelso, & Bunz, 1985).

Our questions are whether the perceptual stability of speech sounds may also be characterized in terms of dynamics and whether these dynamics allow for a deeper understanding of the results of speech identification studies. Typically, listeners judge stimuli as belonging to the same speech category despite large variations in the manipulated acoustic parameter(s). Only when the manipulated parameter reaches a critical value does the percept change, and then it typically does so abruptly and discontinuously. Nevertheless, the similarity between speech identification tasks and the tenets of nonlinear dynamics may be only superficial, because the stimulus set is usually presented to listeners in random order to eliminate short-term sequential effects known to modify category boundaries (e.g., Diehl, Elman, & McCusker, 1978; Diehl, Kluender, & Parker, 1985; Diehl, Lang, & Parker, 1980; Eimas, 1963; Healy & Repp, 1982). This technique locates a statistically defined phoneme boundary (most often, the point corresponding to the 50% crossover of the response function for a two-category set; see Ganong & Zatorre, 1980, for a comparison of different methods for boundary location) at the expense of determining the temporal evolution of the perceptual process, that is, how new percepts form. In the present experiments, rather than using procedures designed to eliminate possible dynamical effects, we studied these effects in their own right, because we believe them to reflect intrinsic properties of the speech perception system. We examined the categorization of each speech token to determine whether the transitions between categories may be understood as a dynamical process, evolving over time.

Our strategy in the present experiments was to use a stimulus continuum for which categorical perception has often been demonstrated but to vary the acoustic parameter sequentially (i.e., as a control parameter) to explore the temporal evolution of phonemic categorization and change. The stimuli chosen comprise two *say–stay* continua (Best, Morrongiello, & Robson, 1981; Hodgson & Miller, 1992) with sequential variation of the duration of the silent gap after the initial fricative. In Experiment 1, this methodology revealed rich contextual effects on the category boundary, including both contrast and hysteresis. We then propose a theoretical model that describes the temporal evolution of perceptual stability and change in terms of factors responsible for both types of context effect. In Experiment 2, we evaluated one prediction of the model—that changing the relative stability of the two percepts promotes categorical switching. Operationally, this means that repetition near the hypothesized instability should increase the likelihood of a spontaneous change in identification. In agreement with the prediction, an increase in the number of stimulus repetitions enhanced the likelihood of category changes near the category boundary. More generally, the experiments demonstrate the fruitfulness of the approach, which helps bridge traditional characterizations of speech perception with conceptual developments in the natural sciences.

Experiment 1

Method

Subjects. The subjects were 9 graduate or undergraduate psychology students who were paid $5 per hour or received credit in their undergraduate psychology course for participating. All subjects reported normal hearing in both ears.

Stimuli. Each stimulus consisted of a natural, 120-ms /s/ excised from a male utterance of *say*, followed by a silent gap of 0–76 ms in 4-ms increments, followed by one of two 290-ms vocalic syllables created on the Haskins parallel-resonance synthesizer. The vocalic syllables were stylized versions of the vocalic portions from natural, male utterances of *say* and *stay* and differed only in the onset frequency of the first formant (F1). F1 onset frequency was either 230 Hz (biased toward perception of *stay* when preceded by /s/) or 430 Hz (biased toward perception of *say* when preceded by /s/). The formant amplitudes and overall amplitude envelopes of the stimuli, as well as the time-varying frequency characteristics of F2 and F3 (and F1 beyond the initial 40 ms), were identical. The stimuli are essentially those used by Best et al. (1981), except that we incremented the duration of the silent gap in smaller steps over a narrower range.

Procedure. There were six experimental conditions which included three presentation orders for each of the two stimulus types (F1 onset = 230 Hz and F1 onset = 430 Hz). Presentation orders included one random and two sequential patterns of change in the duration of the silent gap separating the /s/ from the synthetic syllable. The sequentially ordered patterns involved (a) a silent gap duration that increased in 4-ms increments from 0 ms to 76 ms and then decreased back to 0 ms and (b) a silent gap duration that decreased in 4-ms steps from 76 ms to 0 ms and then increased back to 76 ms.

Use of the two stimulus types with different F1 onsets ensured that listeners could not simply count stimuli in order to produce consistent category boundaries. Including both increasing and decreasing gap durations within any given run in the sequentially ordered conditions allowed for a matched-pairs comparison of where categories changed in a given context.

Experiment 1 contained 10 runs of the randomly ordered condition for each stimulus type and 5 runs of each of the four conditions with sequentially ordered stimulus presentation (two orderings for each of two stimulus types). Presentation was randomized

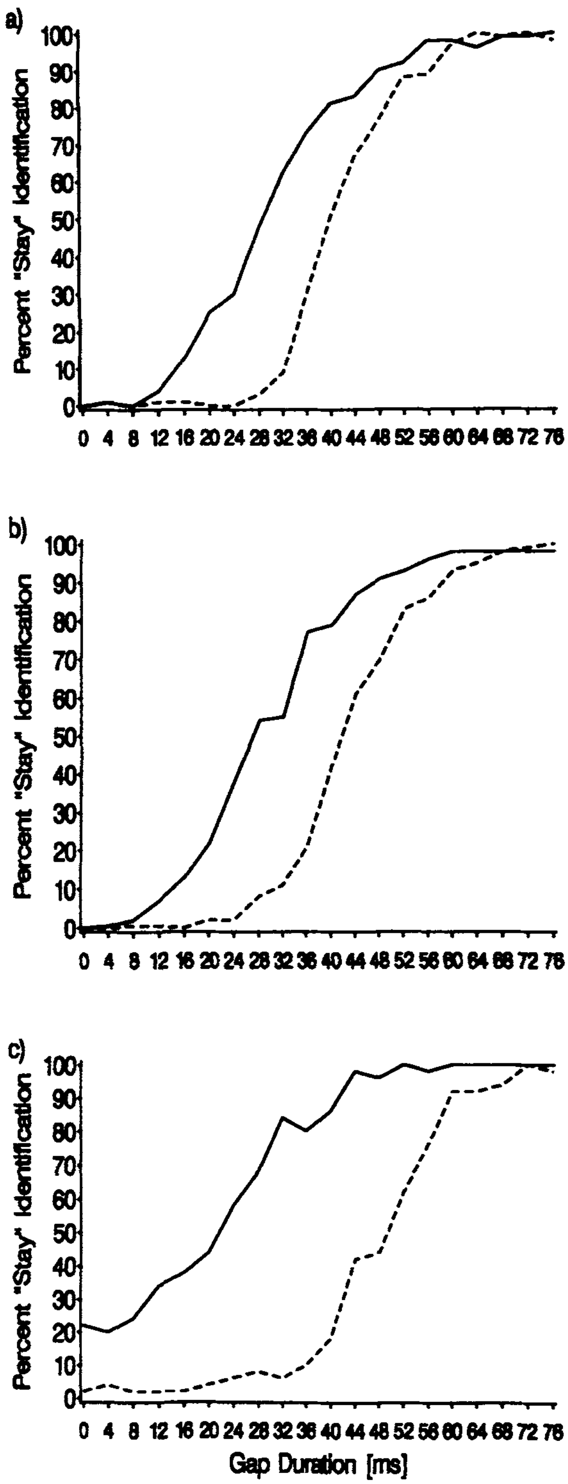

Figure 1. Panels a and b show the percentage of *stay* identifications across subjects in the randomized forced-choice task and the percentage of *stay* identifications for sequentially presented stimuli averaged across direction of gap change in Experiment 1. Panel c shows the percentage of *stay* identifications across subjects in the randomized forced-choice task in Experiment 2. The solid line represents an F1 onset of 230 Hz; the dashed line represents an F1 onset of 430 Hz.

across conditions, with the constraint that sequentially ordered and random runs alternated. Stimuli were separated by a 2-s interstimulus interval (ISI); a 6-s interval was inserted after every 2 runs.

The subject's task was to identify each stimulus as either *stay* or *say* and to indicate the response on an answer sheet during the ISI. Stimuli were presented binaurally through headphones at a comfortable listening level (approximately 70 dB). Subjects were tested individually in a quiet room and took a 5-min break midway through the single 1-hr session.

Results and Discussion

Random presentation. The results of the forced-choice identification task for the randomly ordered stimuli (Figure 1a) replicate classic identification functions (e.g., Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967). They also replicate the trading relations reported by Best et al. (1981), with the category boundary for the stimuli biased toward *stay* occurring at a silent gap that was 12 ms shorter on average than that for stimuli biased toward *say*. Thus, the stimuli and the subjects in the present work seem in no way out of the ordinary.

We examined the results for local sequential effects by computing the conditional probability of judging each stimulus as a member of the same category as the immediately preceding stimulus. When data were averaged across subjects, no systematic contrast or assimilative effects were evident. When data for individual subjects were analyzed, one subject showed a consistently contrastive effect (i.e., the subject tended to classify a stimulus into a category different from that judged for the preceding stimulus); responses from one other subject were consistently assimilative (the subject tended to classify a stimulus into the same category as that judged for the preceding stimulus). Data from the seven other subjects showed varying amounts of contrast and assimilation across the stimulus set. These data differ from those previously reported (e.g., Diehl et al., 1978; Diehl et al., 1985; Diehl et al., 1980; Macmillan, Goldberg, & Braida, 1988), most likely because in our analysis the biasing context was drawn from the entire range of stimuli, not a restricted set.

Sequential presentation. The response patterns produced by sequential variation of the duration of the silent gap between the /s/ and the synthetic syllable were first analyzed by averaging category changes across paired increasing and decreasing series, thus explicitly reducing the sensitivity of the measure to any perceptual dynamics (cf. Ganong & Zatorre, 1980). As shown in Figure 1b, the 50% crossover point determined by this method was close to that obtained with random presentation.

Next, we compared the switch location when the gap duration increased with the switch location when the gap duration decreased. This method can produce only three distinct patterns of results, illustrated in Figure 2 for single runs selected from three different subjects. In the first pattern, the boundary between responding *say* and responding *stay* remains fixed despite changes in the direction of gap change, so that a given gap duration always gives rise to the same response (Figure 2a). This critical boundary occurred only 17% of the time across subjects and conditions.

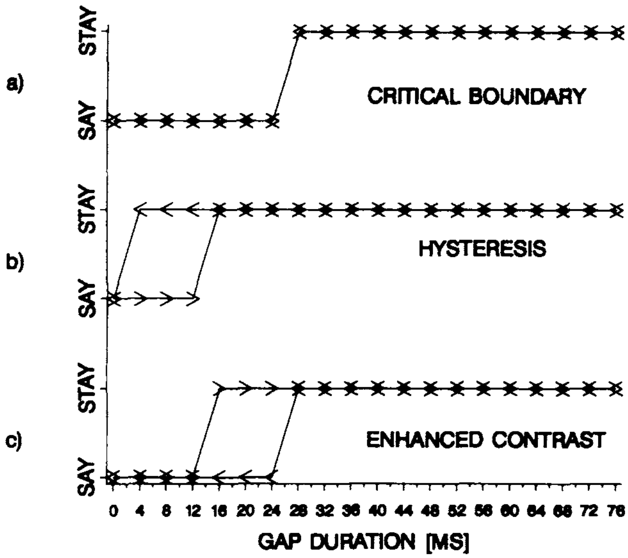

Figure 2. Individual runs showing prototypical patterns of response with sequential stimulus presentation. The greater-than symbol (>) represents responses when gap duration progressively increased; the less-than symbol (<) represents responses when gap duration progressively decreased.

The second possible response pattern is hysteresis: With increasing gap duration, a larger gap is necessary to induce a category change from *say* to *stay* than is necessary for the change from *stay* to *say* with decreasing gap duration (Figure 2b). That is, the initial perceptual state persists even when the current gap duration favors the alternative response when stimuli are randomized. Hysteresis was far more common than a critical boundary in Experiment 1, occurring on 41% of runs across subjects and conditions. The presence of hysteresis—the system's dependence on its recent history—may be regarded as strong evidence of nonlinearity and multistability in speech.

In the third response pattern, switching from *say* to *stay* occurs earlier when gap duration increases than when gap duration decreases (Figure 2c). That is, the subject quickly switches to the alternate percept and does not hold on to the initial categorization. We call this phenomenon *enhanced contrast*, which may be akin to selective adaptation and range-frequency effects in speech perception (e.g., Rosen, 1979), implying that boundary shifts act to enhance the contrast between speech categories (an effect often relegated to response bias; Elman, 1979; Macmillan et al., 1988). Enhanced contrast occurred on 42% of runs across subjects and conditions.

Figure 3 shows the frequency with which each response pattern was observed as a function of condition. Notice first that a critical boundary occurred least frequently of the three response patterns in all conditions. The frequencies with which enhanced contrast and hysteresis occurred did not differ from each other, but both of these response patterns occurred significantly more frequently than did a critical boundary (Wilcoxon matched-pairs signed-ranks test, *ps* $< .05$). Note also that the direction of sequential changes in gap duration influenced the relative frequencies of hysteresis and enhanced contrast. When gap duration of stimuli biased toward *stay* (F1 = 230 Hz) first increased to 76 ms and then decreased back to 0 ms, enhanced contrast was observed more frequently than hysteresis (Wilcoxon matched-pairs signed-ranks test, $p < .05$). When gap duration first decreased, there was no difference between enhanced contrast and hysteresis in frequency of occurrence ($p > .1$).

Individual patterns. Individual variation was seen in the absolute duration of the silent gap at which responses changed category and in the frequency of occurrence of the three response patterns described in the preceding section. Nevertheless, no subject exhibited a critical boundary as the most frequent response pattern. The most frequently a critical boundary was observed was 35% of runs in 1 subject; a critical boundary pattern for the other 8 subjects ranged from 0% to 25% of runs.

Two subjects produced the same response pattern over all runs and conditions, although the gap duration at which category change occurred varied. One of the two always showed hysteresis, despite the fact that the boundary from *say* to *stay* ranged from a 16-ms to a 52-ms gap duration across stimulus types and the boundary from *stay* to *say* ranged from 32 ms to 0 ms. This was the same subject who showed assimilation of the randomly presented stimuli. The other subject always responded with enhanced contrast (the boundary from *say* to *stay* ranged from 12 ms

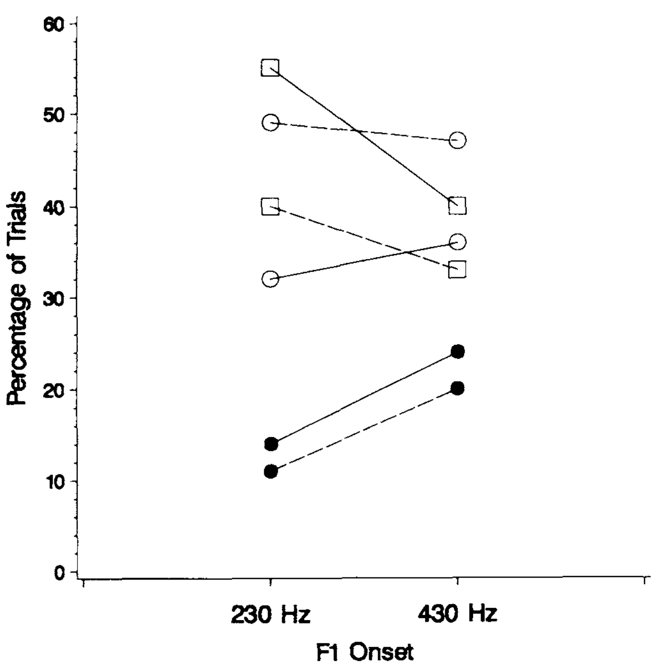

Figure 3. Frequency of response pattern as a function of condition, separately for the two F1 onset frequencies. Solid circles (●) represent critical boundary; open squares (□) represent enhanced contrast; open circles (○) represent hysteresis; solid lines represent gap duration that first increased and then decreased; dashed lines represent gap duration that decreased first.

to 32 ms, and that from *stay* to *say* ranged from 44 ms to 24 ms) and showed contrast effects with random stimulus presentation. The results suggest that the subject who showed consistent hysteresis perceived the continuum as less strongly categorical than did the subject who showed consistent contrast and thus had relatively greater difficulty deciding when a boundary had been crossed. For this explanation to be tenable, however, the subject who showed enhanced contrast should have judged the randomly presented continuum as less continuous than did the subject who showed hysteresis, but this pattern did not occur. Rather, the major effect of sequential presentation was to magnify the local sequential effects observed with random presentation.

Hysteresis was the most common pattern for 3 of the remaining 7 subjects, enhanced contrast was most common for 3 others, and the last subject produced hysteresis and enhanced contrast equally often. These results are in accordance with the lack of consistent local sequential effects in random presentations for these 7 subjects.

To summarize, the results of the forced-choice identification task for randomly ordered stimuli replicate classic identification functions (e.g., Liberman et al., 1967) and show the expected trade-off between spectral (F1 onset) and temporal (gap duration) information. Sequential stimulus presentation allowed a closer examination of contextual effects on judgments of speech category. Two response patterns, enhanced contrast and hysteresis, were observed equally often and far more frequently than a critical boundary.

Models of category boundary change that emphasize range-frequency effects (e.g., Parducci, 1965; Parducci & Wedell, 1986) or consistency (Haubensek, 1992) can accommodate a critical-boundary phenomenon as well as the contrastive effects of sequential presentation on category judgments, but they do not predict hysteresis with stimuli of constant range and frequency. Although effects in speech categorization are usually reported to be contrastive, other sorts of category judgments often exhibit hysteresis (Fender & Julesz, 1967; Haas et al., in press; Hock, Kelso, & Schöner, 1993; Kruse & Stadler, 1990; Tuller, Giangrande, Kelso, & Ding, 1993; Williams, Phillips, & Sekuler, 1986).

One model that addresses both contrast and hysteresis is Helson's (1964) adaptation-level theory, which posits that categorization results from a weighted combination of the stimulus being judged, other stimuli that provide a context, and a residual term that indexes the organism's internal state. (For applications of the model to contrastive effects in speech perception, see Diehl et al., 1980.) Applying this model to the visual perception of brightness, Helson determined that contrast and assimilation result from different combinations of stimulus parameters, with a small zone of interaction in which neither effect occurs. Note, however, that in Experiment 1 the identical change in stimulus parameters yielded three effects: a critical boundary, enhanced contrast, and hysteresis. In the following section, we propose a theoretical model that captures all three patterns of category change within a unified dynamical account.

A Theoretical Model

Determination of an appropriate dynamical model is based on mapping the stable perceptual categories (or forms) for *say* and *stay* onto attractors of the model. For a single percept, a local model containing a fixed point is adequate. If several stable percepts coexist, a nonlinear dynamical model must be found. Task variables are presumed to act as parameters on the perceptual dynamics. In the present experimental paradigm, the percepts studied are parameterized by the silent gap duration between *s* and the synthetic syllable *ay*. An initial perceptual form is established at either extreme of the parameter range. Sequential changes in parameter values eventually result in a switch to the alternative perceptual form. We have shown also that the initial category may persist even when the current value of the parameter favors the alternative response in a randomized stimulus sequence. Such results may be described concisely by the dynamical system

$$V(x) = kx - x^2/2 + x^4/4, \tag{1}$$

where x is the perceptual form and k is the control parameter specifying the direction and degree of tilt for the potential $V(x)$ (see Figure 4). The equation of motion from which the potential function is derived is

$$dx/dt = -dV(x)/dx, \tag{2}$$

in which the long-term behavior of x is captured by stable fixed points (the minima shown in Figure 4).

The parameter k in Equation 1 is generally a monotonically increasing function of the gap duration. For ease of visualization, Figure 4 shows $V(x)$ for several values of k. With $k = -1$, only one stable fixed point exists, corresponding to the percept *say* (see Figure 4a). As k increases from -1 (corresponding to the increase in gap duration), the potential landscape tilts but otherwise remains unchanged in terms of the composition of attractive states (see Figure 4b), and *say* is still the only perceived state. However, when k reaches the critical value of $-k_c$, an additional attractor, corresponding to *stay*, appears via a saddle-node bifurcation (see Kelso, Ding, & Schöner, 1992, for an elementary exposition of dynamical systems theory and its application to biological systems), in which a point attractor (with $x < 0$) and a point repeller (the maximum at $x = 0$) are simultaneously created. The coexistence of both *say* and *stay* continues until $k = k_c$, where the attractor corresponding to *say* ceases to exist via a reversed saddle-node bifurcation, leaving only one stable fixed point in the system corresponding to *stay* (see Figure 4d). Further increases in k only serve to deepen the minimum corresponding to *stay* (see Figure 4e).

An accurate portrait of any real-world problem must take into account the influence of random disturbances, the sources of which, in a perception experiment, may correspond to factors such as fatigue, attention, boredom, and so on. Mathematically, spontaneous switches among attractive states occur as a result of these fluctuations, modeled as random noise. For a given point attractor, the degree of resistance to the influence of random noise is related to its stability, which, in general, is indicated by the depth and width

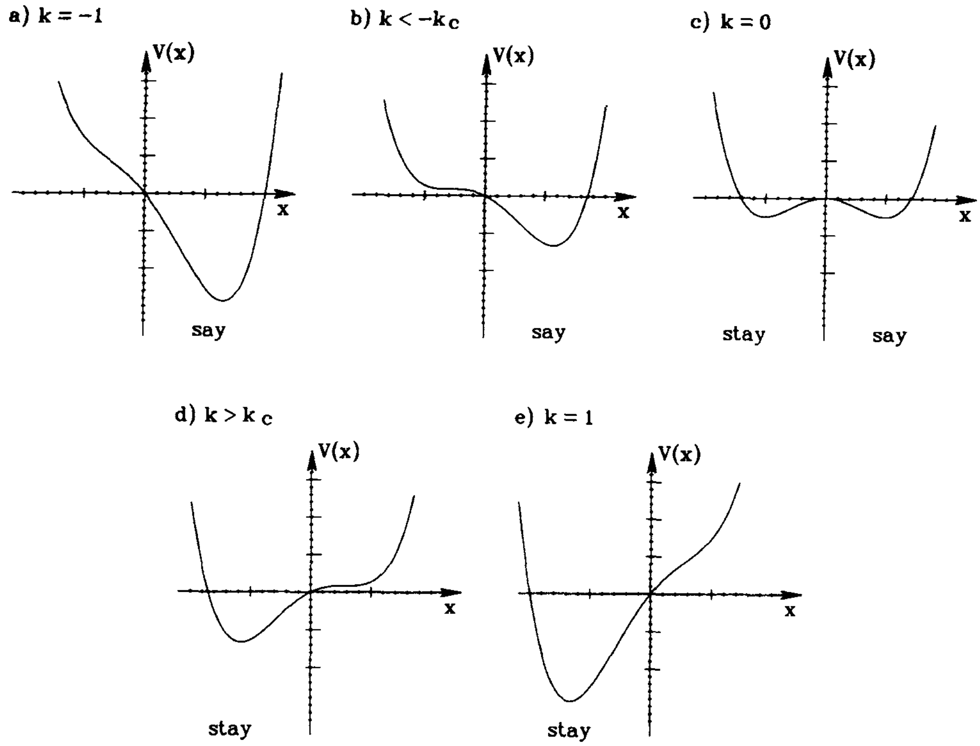

Figure 4. Potential landscape defined by Equation 1 for five values of k. See text for details.

of the potential well (basin of attraction). As k is increased successively in Figure 4, the stability of the attractor corresponding to the initial percept decreases (the potential well becomes shallower and flatter), leading to an increase in the likelihood that the subject will switch to the alternative percept. This implies that perceptual switching is more likely with repeated presentations of a stimulus near the transition point than with repetition of a stimulus far from the transition point. In Experiment 2 we devised measures to test this model prediction.

A few further remarks are in order. First, the correspondence between the theory and experimental observations is qualitative in that the values of coefficients in Equations 1 and 2 are chosen only for the convenience of presentation. Other experimental information (e.g., measures of perceptual fluctuations and local relaxation times of perceptual states) is necessary to establish a precise correspondence. Second, that both the theory and the experimental paradigm are dynamical should be understood in the sense that the ISI between successive presentations of stimuli is small so that the contextual influence of the previous percept is carried into the present percept. More specifically, in terms of the model, this means that the system state clings to its initial attractor, although random fluctuations may push it over the hump. It is worth noting that the paradigm is no longer dynamical if the ISI is very large, thereby eliminating the context effect. This time dependence is also evident in the trace mode of Durlach and Braida's (1969) psychophysical theory of intensity perception (see also Braida et al., 1984; Macmillan, 1987).

To account for the three response patterns observed in Experiment 1—critical boundary, hysteresis, and enhanced contrast—we need to examine in more detail the dependence of k on the gap duration. To do so, we observe that, in general, repetitive presentation of a speech stimulus predictably shifts the location of adjacent phoneme boundaries (for early reviews, see Darwin, 1976; Eimas & Miller, 1978). In Experiment 1, the number of repetitions of each perceptual form before and after a boundary could differ markedly. For example, because most subjects judged the stimulus to change in the first half of the sequence, a run in which gap duration first increased and then decreased had many more contiguous instances categorized as *stay* than categorized as *say*. Taking this factor explicitly into account, we propose the following equation describing the behavior of k as a function of the gap duration:

$$k(\lambda) = k_0 + \lambda + \epsilon/2 + \epsilon\theta(n - n_c)(\lambda - \lambda_f), \qquad (3)$$

where k_0 specifies the initial perceptual configuration at the beginning of a run, λ is a variable linearly proportional to the gap duration, λ_f denotes the value of λ at the other extreme from its initial value, n is the number of perceived stimulus repetitions in a given run, n_c represents a critical number of accumulated repetitions (defined initially as $n_{total/2}$), and $\theta(n - n_c)$ is a step function defined as 0 if $n < n_c$ (i.e., during the first half of each trial) and 1 if $n \geq n_c$ (i.e., during the second half of each trial). In Equation 3, ϵ is a parameter that represents the lumped effects of learning, linguistic experience, and attentional factors. Note that the introduction of criteria stemming from cognitive processes is not without precedent. For example, attention and previous experience play a large role in synergetic modeling of perception of ambiguous visual figures (Ditzinger & Haken, 1989, 1990; Haken, 1990) and contribute to the weighted log mean that determines adaptation level in Helson's (1964) work.

The tilt of the function (Equation 1) thus depends on the initial conditions, the experimental parameter of gap duration, and an additional term incorporating the number of perceived repetitions that is sensitive to cognitive factors. When n, ϵ, or both are sufficiently small, the tilt of the potential depends only on gap duration and the initial configuration.

Figure 5 illustrates the model results corresponding to experimental runs in which the gap duration first increased from 0 ms to 76 ms and then decreased from 76 ms to 0 ms. Similar results hold for situations in which the gap duration first decreased from 76 ms to 0 ms and then increased from 0 ms to 76 ms. Figures 5a and 5b show three regions of different states of the system in the ϵ–λ plane in the first half of each run, when n is small (see Figure 5a), and in the second half of each run, when n is large (see Figure 5b). White regions indicate the set of parameter values for which a stimulus has but a single possible categorization in the represented portion of the run. The white regions at small gap durations are categorized only as *say*, whereas the white regions at large gap durations are categorized only as *stay*. The shaded regions indicate the set of parameter values for which a stimulus may be categorized as either *say* or *stay* (the bistable region) and thus represent the condition from $-k_c$ (the lower border of each shaded region) to k_c (the upper border of each shaded region).

Consider the condition with initial $k_0 = -1$ (*say*) and gap duration, λ, increasing from 0 ms (see Figure 5a). As λ increases, the stimuli are categorized as *say* for any value of ϵ, so long as the (ϵ, λ) coordinate remains below the shaded region. Within the shaded region, the stimuli are categorized as *say* despite the percept's becoming progressively less stable. As λ continues to increase, the percept switches from *say* to *stay* at the upper boundary of the shaded region (the heavy line), after which *say* is not a possible percept. Note that for different values of ϵ, the switch to a new percept occurs at different durations of silent gap.

In the second half of the run, λ is decreasing, $n > n_c$ so that $\theta = 1$, and the resulting division of the ϵ–λ plane looks somewhat different (see Figure 5b). Figure 5b should be read from top to bottom (from large to small gap durations). As λ decreases, the stimuli are categorized as *stay* for any value of ϵ, so long as the (ϵ, λ) coordinate remains above the shaded region. Again assuming the absence of a perturbing force, the subject continues to categorize the stimuli as *stay* within the bistable (shaded) region, despite the percept's becoming progressively less stable. As λ continues to decrease, the *lower* boundary of the shaded region (the heavy curve) marks the switching from *stay* back to *say*.

In Figure 5c, we have simply superimposed the boundaries at which switching occurs in Figures 5a and 5b, because each describes a different segment of a run. The line with negative slope represents the category switch in the first half of the run, and the curve with positive slope represents the switch

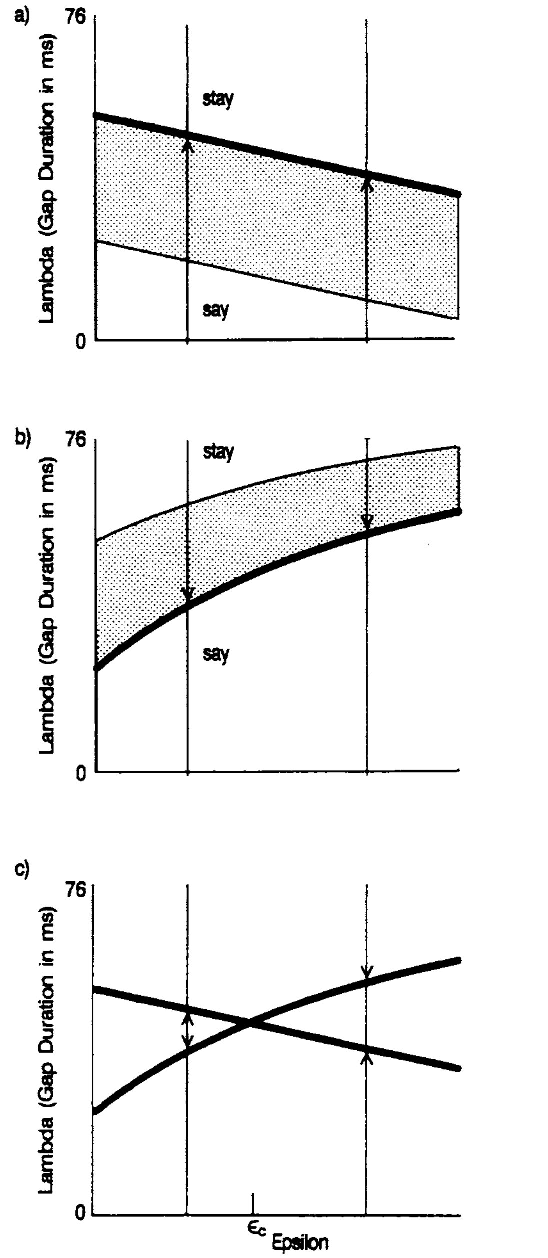

Figure 5. Panel a shows perceptual states in the ϵ–λ plane for increasing λ (small n). Panel b shows perceptual states in the ϵ–λ plane for decreasing λ (large n). Panel c shows superposition of reversed saddle-node bifurcation points in Panels a and b. See text for details.

back to the initial category. The intersection of the line and curve yields a critical value of ϵ, denoted ϵ_c, for which a critical boundary would be observed in the *say*–*stay* continuum (the category change would be reported for the same stimulus, regardless of whether gap duration first sequentially increases or decreases). For $\epsilon > \epsilon_c$ (e.g., the vertical line to the right of ϵ_c in Figure 5c), the system exhibits the enhanced contrast effect. In the region $\epsilon > \epsilon_c$, λ is smaller for the negatively sloped line than for the positively sloped curve. For $\epsilon < \epsilon_c$ (e.g., the vertical line to the left of ϵ_c in Figure 5c), the classical hysteresis phenomenon is obtained: λ is larger for the line than for the curve. The main qualitative features of the observed data are thus captured in the model. Moreover, the parameter plane looks essentially the same when the effect of random fluctuations is explicitly considered. The effect of fluctuations is to enhance the likelihood that switches to the other category will occur as the boundary is approached rather than crossed.

Post Hoc Evaluation of Model Predictions

The observed patterns in the experimental data allowed us to test specific predictions of our model as well as other models. First, when gap duration initially increased and then decreased, subjects reported many more contiguous instances of *stay* than when gap duration first decreased and then increased. Range-frequency models predict that the switch from *stay* to *say* will occur at longer gap durations in the former condition than in the latter condition if one allows that (a) a skewed frequency distribution of category identification acts similarly to a skewed frequency distribution of the individual stimuli (Diehl, 1975) and (b) the effective context for judgment extends over a shorter term than the entire experimental session (Parducci & Wedell, 1986). Similarly, when gap duration first decreased and then increased, the number of contiguous *say* identifications was large; per range-frequency models, the switch from identification of *say* to *stay* should occur on stimuli with smaller gap durations than when gap duration first increases. These range-frequency predictions were met for only 19% of runs across subjects and conditions in Experiment 1.

The present model incorporates a variable, λ, that varies with the range of stimuli and a term, n, that reflects the skewed frequency distribution of category members. However, their effects cannot be interpreted independent of ϵ. Thus, one prediction of the model is that when the silent gap duration initially increases, the switch from perceiving *say* to perceiving *stay* should occur at smaller gap durations when a critical boundary is observed than when hysteresis occurs; switching with enhanced contrast should occur at still smaller gap durations (compare the values of λ in the three regions of the negatively sloped line in Figure 5c). The pattern of relative gap durations should be reversed in the second half of the run, when the duration of the silent gap sequentially decreases; switching from *stay* to *say* should occur at longer gaps when enhanced contrast occurs than when a critical boundary occurs. Critical boundaries should, in turn, occur at longer gap durations than are observed for hysteresis (note these relationships in the positively sloped curve of Figure 5c).

We first compared only the hysteresis and enhanced contrast switch points because of the paucity of responses that exhibited a critical boundary. Fisher's exact probability was calculated across runs for the 7 individuals who showed both hysteresis and enhanced contrast. Results (using a criterion level of $p < .01$) agreed with predictions for 6 of the 7 subjects for both directions of change in gap duration. For the remaining subject, when gap duration increased, switching occurred at smaller gap durations with hysteresis than with enhanced contrast (as predicted), but the switch point for the two response patterns did not differ when gap duration sequentially decreased. When gap durations at which critical boundaries occurred were included in the analyses, there was only a single comparison in which the switch point for the critical boundary was more extreme than that observed for either hysteresis or enhanced contrast.

The proposed effects of ϵ in our model allow a second prediction. As incorporated in the model, ϵ reflects the combined pressures of factors with different effects. Factors such as learning and experience with the stimuli tend to increase ϵ over the course of the experiment and lead to the prediction that the frequency of occurrence of hysteresis relative to enhanced contrast should decrease as the experiment progresses. For the 7 subjects who showed both hysteresis and enhanced contrast, we compared the frequency with which each pattern occurred in the first 10 runs of the experiment with the frequency with which it occurred in the last 10 runs. Hysteresis occurred less often as the experiment progressed for 5 of those subjects.

Another component of ϵ, attention, should, in principle, produce local fluctuations as attention waxes and wanes. Therefore, if ϵ is close to ϵ_c, the local fluctuations are likely to change the observed response pattern (e.g., from hysteresis to enhanced contrast or critical boundary). Such local fluctuations are less likely to change the response pattern when ϵ is far from ϵ_c. One method of assessing distance from ϵ_c is the degree of overlap (the hysteresis region) and "underlap" (in enhanced contrast), which, according to the model, increases with the distance from ϵ_c. Subjects who show only hysteresis or only enhanced contrast are thus predicted to be far from ϵ_c because the local fluctuations do not cause a change in response pattern. These subjects should show a larger hysteresis or contrast region than subjects who exhibit more than one response pattern (who, theoretically, are closer to ϵ_c). Simply ranking the size of the overlap or underlap confirmed this prediction: The subject who showed only hysteresis and the subject who showed only enhanced contrast had the two largest overlap and "underlap" regions, respectively.

To summarize, the model captures changes in speech categorization as a process of self-organization—a spontaneous change in perceptual pattern—and provides a compact description of many qualitative features of the experimental data. In addition, the inclusion of random fluctuations is of fundamental conceptual importance in that it allows switches between attracting states. In physical and certain biological systems, the size of these fluctuations as measured, for example, by the variance or standard deviation of x around the attractive state is a measure of the stability of the state (Haken, 1983b). The more stable the attractor, the smaller the mean deviation from the attractive state for a given strength of stochastic force. These predictions have been worked out in quantitative detail in the case of biological coordination (for reviews, see Jeka & Kelso, 1989; Schöner & Kelso, 1988a). Although with categorical responses we have no direct measure of the variance of x, we can evaluate the changing relative stability of the categories by including trials for which the present control parameter, silent gap duration, changes sequentially but then remains fixed for multiple presentations. In principle, as a category loses stability, the shallower minimum in the potential has a less restraining influence on the fluctuations, and changes in category should occur earlier with repetition than with presentation of a single stimulus. Stimulus repetition also probes the notion that the region $-k_c < k < k_c$ is bistable; spontaneous changes in category (i.e., with no corresponding change in the control parameter) are possible only if more than one stable state exists.

Experiment 2

Method

Subjects. Subjects were 6 graduate or undergraduate psychology students motivated by credit or pay. Three of the subjects had participated in Experiment 1. One of the three had exhibited enhanced contrast in all runs of Experiment 1; the other 2 showed different combinations of enhanced contrast, hysteresis, and critical boundary.

Procedure. The stimuli were the same as in Experiment 1 across the same range of silent gap durations.

There were eight experimental conditions in addition to the more typical random ordering of stimuli. These included two presentation orders (gap duration sequentially increasing and sequentially decreasing) for each of the two stimulus types (F1 onset = 230 Hz and F1 onset = 430 Hz), with the final gap duration presented either once or six times.

Runs in which gap duration sequentially increased all began with a gap of 0 ms and progressed to a gap duration that ranged from 8 ms to 76 ms, for a total of 18 runs with different end points. In addition, the stimulus with the longest gap duration in each run was presented either once or six times. For example, a nonrepeated, increasing run to 24-ms gap duration included stimuli with gaps of 0, 4, 8, 12, 16, 20, and 24 ms. A repeated, increasing run to the same gap duration proceeded as follows: 0, 4, 8, 12, 16, 20, 24, 24, 24, 24, 24, 24 ms.

Runs in which gap duration sequentially decreased all began with a gap of 76 ms and ended with a gap duration that ranged from 68 ms to 0 ms (18 runs with different end points). The stimulus with the shortest gap duration in each run occurred either once or six times. A nonrepeated, decreasing run to a stimulus of 24-ms gap duration included stimuli with gaps of 76, 72, 68, 64, 60, 56, 52, 48, 44, 40, 36, 32, 28, and 24 ms. A repeated, decreasing run to a stimulus of the same gap duration proceeded identically but with six instances of the syllable with the 24-ms gap. In all cases, stimuli were separated by a 1-s ISI.

The experiment contained 144 unique runs (to 18 different end points in each of eight conditions), which were randomized and separated into four blocks of 36 runs each. Ten independent randomizations of the four blocks were constructed and presented to subjects over ten 1-hr sessions (one randomization per session).

Subjects were tested individually in a sound-insulated room. The subject's task was to identify the first stimulus of each run as either *say* or *stay* and to press the return key if a switch from one percept to another was heard. The keypress aborted the remainder of that run and, after a 3-s interval, started presentation of the next run. This procedure was adopted in order to minimize the duration of each session. Stimuli were presented binaurally through headphones at a comfortable listening level (approximately 70 dB). Subjects were given a short break of about 5 min after each block of 36 runs.

At the end of the experiment, subjects were presented with a randomized set of all stimuli to ensure that they did, in fact, demonstrate reliable identification of the most extreme stimuli and reproduced a classical identification function.

Results and Discussion

Random presentation. The results for the forced-choice identification task for the randomly ordered stimuli (Figure 1c) again replicate earlier studies (e.g., Best et al., 1981; Liberman et al., 1967), including Experiment 1. A category boundary difference of 28 ms, on average, was observed between the stimuli biased toward *say* and those biased toward *stay*. In Figure 1c, note first that the curve for stimuli biased toward *stay* is not invariably identified as *say* even at the shortest gap duration. This was due to a single subject who perceived all such stimuli as *stay*. Second, the identification curves in Figure 1c are less smooth than those in Figures 1a and 1b, primarily, we suspect, because fewer subjects participated in Experiment 2 than in Experiment 1. In addition, one subject's data were omitted from further analysis because she exhibited extreme variability of identification of stimuli across the entire range of gap durations used.

Sequential presentation. Our primary concern in Experiment 2 was whether changes in speech identification arise from loss of stability of an initial perceptual category. To this end, we examined whether, at final gap duration, the likelihood of a spontaneous change in identification was maximized by the repetition of the final stimulus. We determined the earliest stimulus on which category change was reported on Tokens 2–6 of the repeated condition. We then determined the percentage of switches occurring on that stimulus in the nonrepeated condition for five runs as follows: If the earliest category change in the repeated condition was reported on stimuli with a silent gap duration of, for example, 12 ms, nonrepeat runs ending in 12, 16, 20, 24, and 28 ms were examined for switches occurring on stimuli up to and including those with 12-ms gaps. This procedure allowed five opportunities for observing a switch in repeated and nonrepeated conditions. Analysis of the repeated condition was broken down further into the earliest stimulus on which category change was reported during Tokens 2–6 of the repeated section compared with the first token of the repeated section (again equated for opportunity of observing a switch).

We calculated *t* tests for related measures separately for the four conditions (two F1 onsets × two directions of gap variation) across the 5 subjects. For all conditions, changes in the identified speech category occurred earlier (i.e., at a shorter gap duration when the duration was increasing and at a longer gap duration when the duration was decreasing) during Tokens 2–6 of the repeated stimulus than during the first instance of that stimulus. Similarly, switches occurred earlier during Tokens 2–6 of the repeated stimulus than when the same stimulus was presented only once. Across all comparisons, $t(4)$ values ranged from 3.47 to 5.37, $ps < .025$, one-tailed. As expected, there were no significant differences in gap duration at which switches occurred between the first token of the last gap duration in the repeated condition and the single token of the corresponding nonrepeated condition ($ps > .1$).

Representative data from two subjects are displayed in Figure 6. These particular subjects were chosen as examples because they had also participated in Experiment 1. Figure 6 plots the proportion of runs (F1 = 430 Hz) on which a category change was reported within the repeated section, on the first token of the repeated section, and on the final, nonrepeated stimulus of the corresponding run as a function of final gap duration. Graphs on the left and right display results when gap duration sequentially increased and decreased, respectively (thus graphs on the right should be read from right to left). For both subjects, repeated tokens induced category changes earlier than did single presentations; that is, in Fig-

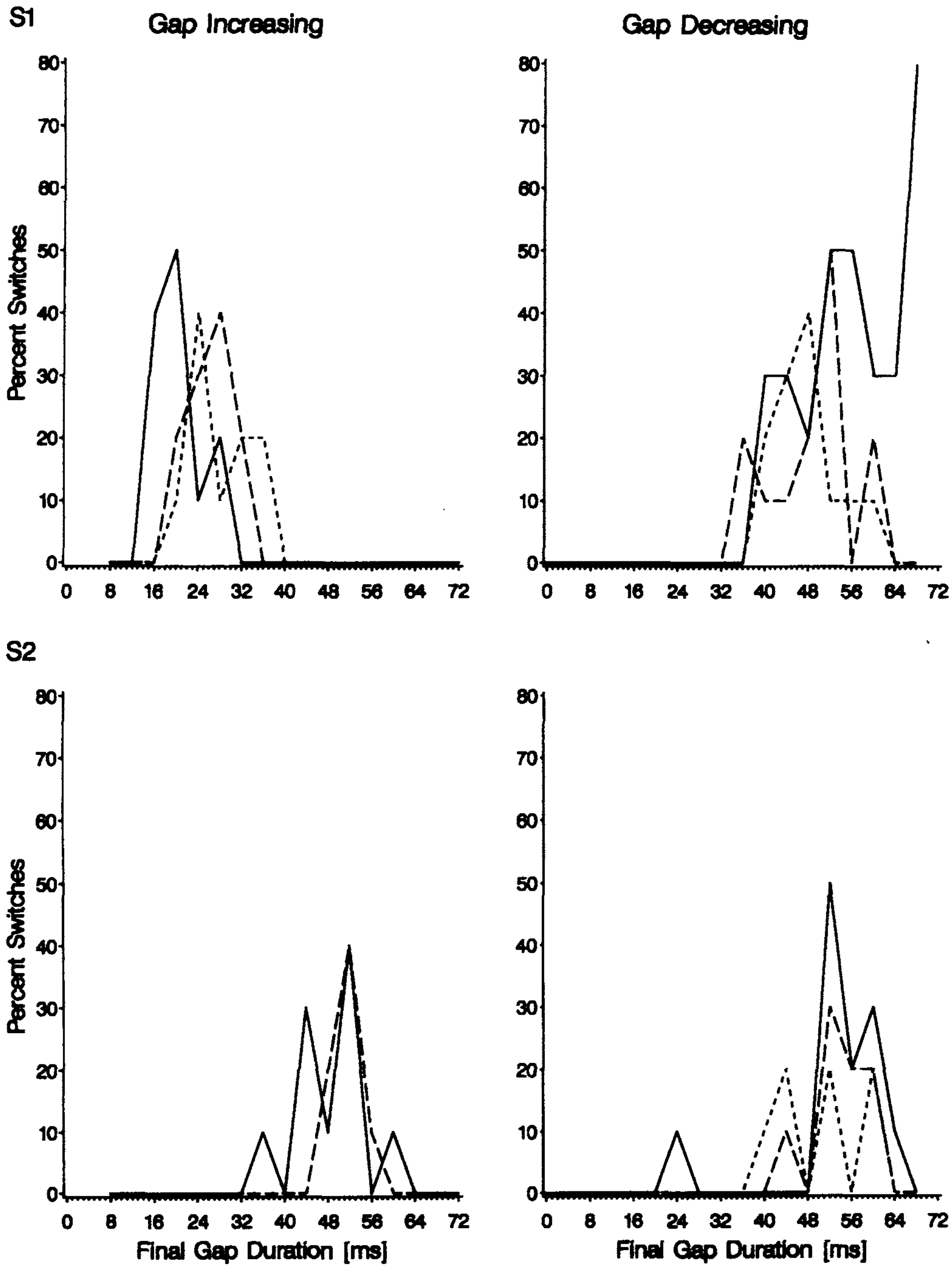

Figure 6. The proportion of runs (F1 onset = 430 Hz) on which a category change was reported within the repeated section (solid lines), on the first token of the repeated section (long dashes), and on the nonrepeated token (short dashes) as a function of final gap duration. Data from two subjects are shown separately. Left panels represent sequentially increasing gap duration; right panels represent sequentially decreasing gap duration (see text).

ure 6 the solid lines lift off the x-axis earlier than either of the dashed lines. Note that the data from this experiment are consistent with the data from Experiment 1. The subject who exhibited enhanced contrast on all runs in Experiment 1 showed category changes at shorter gaps when gap duration sequentially increased than when gap duration sequentially decreased (Figure 6, top left vs. top right). This subject also showed consistently contrastive effects in the randomly ordered conditions of both experiments. The subject who showed a mixture of response patterns and no consistent local sequential effect in Experiment 1 showed a much smaller separation in the distribution of switching points for increasing versus decreasing gap duration in Experiment 2 (Figure 6, bottom left vs. right). [Note that for this subject the data from the two single token conditions are overlaid when gaps increase.]

To summarize, regardless of whether a subject's response patterns showed hysteresis, enhanced contrast, or a critical boundary, an increase in the number of stimulus repetitions maximized the likelihood of category change near the boundary. Although the increased likelihood of category change is in agreement with the theoretical model proposed here, alternative explanations should also be considered. To this end, we explored whether sequential presentation of stimuli and stimulus repetition act primarily on the perceived frequency of the categories, independent of the stability of the categories. Of course, it would be scientifically counterproductive for us to propose a model that did not accommodate the known anchoring effects of stimulus repetition on category boundaries, and such effects are incorporated into the parameter n in Equation 2. Nevertheless, our model separates the effects of repetition alone from the effects of systematically leading the subject toward the category boundary before a stimulus is repeated.

In an effort to tease apart these effects, we examined the ordering of runs when two ascending runs, or two descending runs, were directly abutted. We included runs in which the final stimulus was not repeated and those in which the final stimulus was slated for repetition but the switch in category judgment occurred before that stimulus was reached and the remainder of the run was aborted. Among runs with progressively decreasing gap duration, abutted runs provided two long sequences of stimuli categorized as *stay*, each followed by a single stimulus categorized as *say*. In general, range-frequency and adaptation-level theories predict that with a constant stimulus range (as used in Experiment 2), shifts in category boundary will be in the direction of the more frequently occurring stimulus (see Diehl et al., 1980, for a similar analysis of speech boundary effects). Thus, a boundary shift should be observed in the direction of the *stay* end of the continuum (i.e., a switch at a stimulus with longer silent gap duration) on the second run relative to the first run. In similar fashion, when gap duration progressively increased, abutted runs provided two long sequences of stimuli categorized as *say*, each followed by a single stimulus categorized as *stay*. According to range-frequency and adaptation-level theories, these sequences should produce a shift in category boundary in the direction of the *say* end of the continuum on the second run relative to the first run. In other words, a switch from the perception of *say* to the perception of *stay* should occur at a stimulus of smaller gap duration in the second run than in the first run. These predictions were met only 45% of the time, despite the large difference between the frequencies of the two categories. Thus, the experimental results are unlikely to have been solely a function of frequency of occurrence of category members per se. Rather, the relative stability of the categories also promotes or inhibits the possibility of perceptual change.

In earlier work, Diehl et al. (1980) reported that adaptation-level theory provides a good qualitative fit to adaptation and contrast effects in speech. For a quantitative fit to be obtained, however, an additional parameter is necessary that indexes the inherent stability of a stimulus's category membership. The present model accomplishes the same goal using a somewhat different tack. That is, rather than including a parameter that independently indexes stability, it reflects the relative stability of a percept by the depth and width of the basin of attraction (solutions to Equation 1).

Another related explanation for the present data stems from an analysis of hysteresis effects in perception of ambiguous visual patterns in which a bias parameter and an attentional parameter interact (Ditzinger and Haken, 1989, 1990; Haken, 1990). In the present experiment, the bias parameter maintained the initially perceived word through some range of the bistable region. However, the bias itself is a time-dependent process whose influence fades as the gap duration progressively favors the alternative category. In addition, in the Ditzinger–Haken model attentional parameters become saturated during the recognition process. Within the bistable region, a strong bias for the initial response will maintain the category for a relatively long period of time before the saturation of attention results in a category switch. As gap duration varies, the bias fades and attention to the initial category saturates after a relatively shorter period of time. Thus, as the bias to the initial category fades, stimulus repetition is more likely to result in switching than is a single instance of the same stimulus.

General Discussion

The traditional approach to understanding cognitive and perceptual processes is to emphasize static representational structures that minimize time dependencies and ignore the potentially dynamic nature of such representations. Increasingly, this approach is being challenged by researchers in a wide variety of fields who believe that dynamical systems theory is a more appropriate mathematical framework for the study of cognition than is symbolic computation and have built dynamics directly into their theories and models (Haken & Stadler, 1990). In the last decade, explicit dynamical models have been formulated to express coordination dynamics governing spontaneous pattern formation across the limbs of an individual (e.g., de Guzman & Kelso, 1991; Haken et al., 1985; Jeka & Kelso, 1989; Kelso, 1981, 1984), between individuals (Schmidt, Carello, & Turvey, 1990), and between individuals and their environment (Kelso, DelColle, & Schöner, 1990; Wimmers, Beek, & van Wieringen, 1992). The concepts have been elaborated to characterize intentional changes in behavior (Kelso, Scholz, & Schöner, 1988; Schöner & Kelso, 1988b), learning (e.g., Kelso, 1990; Schöner, Zanone, & Kelso, 1992; Zanone & Kelso, 1992), development (Thelen, Kelso, & Fogel, 1987; Thelen & Ulrich, 1991), visual perception (Fender & Julesz, 1967; Haas et al., in press; Hock et al., 1993; Kruse & Stadler, 1990; Williams et al., 1986), artificial neural networks for pattern recognition and associative learning (Haken, 1991; Haken, Kelso, Fuchs, & Pandya, 1990; Kawamoto & Anderson, 1985), and human brain activity (Fuchs, Kelso, & Haken, 1992; Kelso, Bressler, et al., 1992).

The present experiments indicate that the basic concepts of self-organization, such as control parameters, bifurcations, attractors, multistability, instability, and hysteresis, can also help expose the underlying (nonlinear) dynamics of speech categorization. Some of these concepts have been illustrated in earlier work on speech production in which subjects were required to say /ip/ or /pi/ repetitively, at ever increasing speaking rates (the control parameter), while their glottal and lip movements were monitored (Kelso, Saltzman, & Tuller, 1986; Tuller & Kelso, 1990). In general, the observed interarticulator phase relationships varied little within limited regions of change in the rate parameter. However, at critical values of speaking rate, the relative phase of articulator movements changed qualitatively (see Stetson, 1951, for similar data, although in a different theoretical context). These transition points corresponded exactly to the perceptual shift in syllable affiliation of the consonant (see also Tuller & Kelso, 1991). Thus, the discontinuities in production created by speaking rate, like the discontinuities in perception created by varying an acoustic parameter, are indicative of the dynamical nature of speech processes.

We may interpret several investigations of speech in a similar way, although they have not explicitly followed a dynamical systems approach (Kelso & Tuller, 1984). Most notable among them is Stevens' quantal theory (1972, 1989). Quantal theory demonstrates how the acoustic resonances produced by the vocal tract, together with characteristics of the auditory system, give rise to a set of stable regions within the multidimensional articulatory/auditory space. These regions are considered to be stable because the auditory properties associated with them are relatively insensitive to small articulatory perturbations along some dimensions (for similar considerations in producing the necessary aerodynamic conditions for glottal modes, see Catford, 1977; Kelso & Tuller, 1984). Stevens suggested that the languages of the world have taken advantage of regions with stable articulatory–acoustic correspondences when developing places of articulation for consonants, avoiding the intervening unstable nonlinearities (see Goldstein, 1983, for an intriguing example of articulatory–acoustic correspondences and their effect on historical vowel shifts).

Other, more explicitly dynamical investigations include attempts to form a topology of vowels (Wildgen, 1990) and consonants (Petitot-Cocorda, 1985) in terms of a landscape of attractors and repellers within an articulatory or acoustic space. In a similar vein is the modeling of phonological systems as a self-organized solution of talker- and listener-based constraints (Lindblom, MacNeilage, & Studdert-Kennedy, 1983).

The present investigation of the dynamics of speech perception entails examining individual shifts in category boundaries as a function of continuous variation on some acoustic control parameter. In so doing, we have made no effort to distinguish particular cues or mechanisms for the perception of speech sounds. Instead, we focused on uncovering the dynamical processes involved in categorization and change. The assumption is that the same dynamics apply across particular cues or categories (and, we suspect, provide a promising operational approach to understanding the ontogenetic development of phonetic categories).

As a final note, it is reassuring that despite our singular methodology the present results and the model incorporate the effects of selective adaptation, range, and frequency on the location of category boundaries. The present model accommodates not only contrastive and assimilative effects, but also hysteresis. Thus, the approach may serve to unify a broad set of data and provide a coherent rationale for predicting contextual effects in speech.

References

Bergé, P., Pomeau, Y., & Vidal, C. (1984). *Order within chaos.* Paris: Hermann.

Best, C., Morrongiello, B., & Robson, R. (1981). Perceptual equivalence of acoustic cues in speech and nonspeech perception. *Perception & Psychophysics, 29,* 191–211.

Braida, L. D., Durlach, N. I., Lim, J. S., Berliner, J. E., Rabinowitz, W. M., & Purks, S. R. (1984). Intensity perception: 13. Perceptual anchor model of context coding. *Journal of the Acoustical Society of America, 76,* 722–731.

Catford, J. C. (1977). *Fundamental problems in phonetics.* Bloomington: Indiana University Press.

Darwin, C. (1976). The perception of speech. In E. C. M. Friedman (Ed.), *Handbook of perception* (Vol. 1, pp. 175–226). San Diego, CA: Academic Press.

de Guzman, G. C., & Kelso, J. A. S. (1991). Multifrequency behavioral patterns and the phase attractive circle map. *Biological Cybernetics, 64,* 485–495.

Diehl, R. L. (1975). The effect of selective adaptation on the identification of speech sounds. *Perception & Psychophysics, 17,* 48–52.

Diehl, R. L., Elman, J. L., & McCusker, S. B. (1978). Contrast effects on stop consonant identification. *Journal of Experimental Psychology: Human Perception and Performance, 4,* 599–609.

Diehl, R. L., Kluender, K. R., & Parker, E. M. (1985). Are selective adaptation and contrast effects really distinct? *Journal of Experimental Psychology: Human Perception and Performance, 11,* 209–220.

Diehl, R. L., Lang, M., & Parker, E. M. (1980). A further parallel between selective adaptation and contrast. *Journal of Experimental Psychology: Human Perception and Performance, 6,* 24–44.

Ditzinger, T., & Haken, H. (1989). Oscillations in the perception of ambiguous patterns: A model based on synergetics. *Biological Cybernetics, 61,* 279–287.

Ditzinger, T., & Haken, H. (1990). The impact of fluctuations on the recognition of ambiguous patterns. *Biological Cybernetics, 63,* 453–456.

Durlach, N. I., & Braida, L. D. (1969). Intensity perception: I. Preliminary theory of intensity resolution. *Journal of the Acoustical Society of America, 46,* 372–383.

Eimas, P. D. (1963). The relation between identification and discrimination along speech and non-speech continua. *Language and Speech, 6,* 206–217.

Eimas, P., & Miller, J. (1978). Effects of selective adaptation on the perception of speech and visual patterns: Evidence for feature detectors. In R. Walk & H. Pick (Eds.), *Perception and experience* (pp. 307–345). New York: Plenum Press.

Elman, J. L. (1979). Perceptual origins of the phoneme boundary effect and selective adaptation to speech: A signal detection theory analysis. *Journal of the Acoustical Society of America, 65,* 190–207.

Epstein, W., & Hatfield, G. (in press). Gestalt psychology and the philosophy of mind. *Philosophical Psychology.*

Fender, D., & Julesz, B. (1967). Extension of Parnum's fusional area in binocularly stabilized vision. *Journal of the Optical Society of America, 57,* 819–830.

Fuchs, A., Kelso, J. A. S., & Haken, H. (1992). Phase transitions in the human brain: Spatial mode dynamics. *International Journal of Bifurcation and Chaos, 2,* 917–939.

Ganong, W., & Zatorre, R. (1980). Measuring phoneme boundaries four ways. *Journal of the Acoustical Society of America, 68,* 431–439.

Goldstein, L. (1983). Vowel shifts and articulatory-acoustic relations. In M. P. R. van den Broecke & A. Cohen (Eds.), *10th International Congress of Phonetic Science* (pp. 267–273). Dordrecht, The Netherlands: Foris.

Haas, R., Fuchs, A., Haken, H., Horvath, E., Pandya, A. S., & Kelso, J. A. S. (in press). Recognition of dynamic patterns by a synergetic computer. *Progress in Neural Networks.*

Haken, H. (1983a). *Advanced synergetics: Instability hierarchies of self-organizing systems and devices.* Berlin: Springer-Verlag.

Haken, H. (1983b). *Synergetics, an introduction: Non-equilibrium phase transitions and self-organization in physics, chemistry and biology.* Berlin: Springer-Verlag.

Haken, H. (1990). Synergetics as a tool for the conceptualization and mathematization of cognition and behavior—How far can we go? In H. Haken & M. Stadler (Eds.), *Synergetics of cognition* (pp. 2–31). Berlin: Springer-Verlag.

Haken, H. (1991). *Synergetics, computers and cognition.* Berlin: Springer-Verlag.

Haken, H., Kelso, J. A. S., & Bunz, H. (1985). A theoretical model of phase transitions in human hand movements. *Biological Cybernetics, 51,* 347–356.

Haken, H., Kelso, J., Fuchs, A., & Pandya, A. (1990). Dynamic pattern recognition of coordinated biological motion. *Neural Networks. 3,* 395–401.

Haken, H., & Stadler, M. (Eds.). (1990). *Synergetics of cognition.* Berlin: Springer-Verlag.

Harnad, S. (1987). *Categorical perception: The groundwork of cognition.* Cambridge, England: Cambridge University Press.

Haubensek, G. (1992). The consistency model: A process model for absolute judgments. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 303–309.

Healy, A. F., & Repp, B. H. (1982). Context sensitivity and phonetic mediation in categorical perception. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 68–80.

Helson, H. (1964). *Adaptation level theory: An experimental and systematic approach to behavior.* New York: Harper & Row.

Hock, H., Kelso, J. A. S., & Schöner, G. (1993). Bistability and hysteresis in the organization of apparent motion patterns. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 63–80.

Hodgson, P., & Miller, J. L. (1992). Internal phonetic category structure depends on multiple acoustic properties: Evidence for within-category trading relations (abstract). *Journal of the Acoustical Society of America, 92,* 2464.

Jeka, J. J., & Kelso, J. A. S. (1989). The dynamic pattern approach to coordinated behavior: A tutorial review. In S. A. Wallace (Ed.), *Perspectives on the coordination of movement* (pp. 3–45). Amsterdam: North-Holland.

Kawamoto, A. H., & Anderson, J. A. (1985). A neural network model of multistable perception. *Acta Psychologica (Amsterdam), 59,* 35–65.

Kelso, J. A. S. (1981). On the oscillatory basis of movement. *Bulletin of the Psychonomic Society, 18,* 63.

Kelso, J. A. S. (1984). Phase transitions and critical behavior in human bimanual coordination. *American Journal of Physiology: Regulatory, Integrative and Comparative Physiology, 15,* R1000–R1004.

Kelso, J. A. S. (1990). Phase transitions: Foundations of behavior. In H. Haken (Ed.), *Synergetics of cognition* (pp. 249–268). Berlin: Springer-Verlag.

Kelso, J. A. S., Bressler, S. L., Buchanan, S., deGuzman, G. C., Ding, M., Fuchs, A., & Holroyd, T. (1992). A phase transition in human brain and behavior. *Physics Letters A, 169,* 134–144.

Kelso, J. A. S., DelColle, J. D., & Schöner, G. (1990). Action-perception as a pattern formation process. In M. Jeannerod (Ed.), *Attention and performance* (Vol. 13, pp. 139–169). Hillsdale, NJ: Erlbaum.

Kelso, J. A. S., Ding, M., & Schöner, G. (1992). Dynamic pattern formation: A primer. In A. Baskin & J. Mittenthal (Eds.), *Principles of organization in organisms* (pp. 397–439). Santa Fe, NM: Santa Fe Institute; and Reading, MA: Addison-Wesley.

Kelso, J. A. S., Saltzman, E. L., & Tuller, B. (1986). The dynamical perspective in speech production: Data and theory. *Journal of Phonetics, 14,* 29–60.

Kelso, J. A. S., Scholz, J. P., & Schöner, G. S. (1988). Dynamics governs switching among patterns of coordination in biological movement. *Physics Letters, A134*(1), 8–12.

Kelso, J. A. S., & Schöner, G. S. (1987). Toward a physical (synergetic) theory of biological coordination. *Springer Proceedings in Physics, 19,* 224–237.

Kelso, J. A. S., & Tuller, B. (1984). A dynamical basis for action systems. In M. S. Gazzaniga (Ed.), *Handbook of cognitive neuroscience* (pp. 321–356). New York: Plenum.

Köhler, W. (1940). *Dynamics in psychology.* New York: Liveright.

Kruse, P., & Stadler, M. (1990). Stability and instabilities in cognitive systems: Multistability, suggestion, and psychosomatic interaction. In H. Haken & M. Stadler (Eds.), *Synergetics of cognition* (pp. 201–215). Berlin: Springer-Verlag.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review, 74,* 431–461.

Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology, 54,* 358–368.

Lindblom, B., MacNeilage, P., & Studdert-Kennedy, M. (1983). Self-organization processes and the explanation of phonological universals. In B. Butterworth, B. Comrie, & O. Dahl (Eds.), *Explanations of linguistic universals* (pp. 181–203). The Hague, The Netherlands: Mouton.

Macmillan, N. A. (1987). Beyond the categorical/continuous distinction: A psychophysical approach to processing modes. In S. Harnad (Ed.), *Categorical perception: The groundwork of cognition* (pp. 53–85). Cambridge, England: Cambridge University Press.

Macmillan, N. A., Goldberg, R. F., & Braida, L. D. (1988). Resolution for speech sounds: Basic sensitivity and context memory

on vowel and consonant continua. *Journal of the Acoustical Society of America, 84,* 1262–1280.
Nicolis, G., & Prigogine, I. (1989). *Exploring complexity: An introduction.* San Francisco: Freeman.
Parducci, A. (1965). Category judgment: A range-frequency model. *Psychological Review, 72,* 407–418.
Parducci, A., & Wedell, D. H. (1986). The category effect with rating scales: Number of categories, number of stimuli, and method of presentation. *Journal of Experimental Psychology: Human Perception and Performance, 12,* 496–519.
Petitot-Cocorda, J. (1985). *Les catastrophes de la parole. De Roman Jakobson à René Thom* [Speech catastrophes: From Roman Jakobson to René Thom]. Paris: Maloine.
Repp, B. (1984). Categorical perception: Issues, methods, findings. In N. J. Lass (Ed.), *Speech and language: Advances in basic research and practice* (Vol. 10, pp. 243–335). San Diego, CA: Academic Press.
Repp, B., & Liberman, A. (1987). Phonetic category boundaries are flexible. In S. Harnad (Ed.), *Categorical perception: The groundwork of cognition* (pp. 89–112). Cambridge, England: Cambridge University Press.
Rosen, S. M. (1979). Range and frequency effects in consonant categorization. *Journal of Phonetics, 7,* 393–402.
Schmidt, R. C., Carello, C., & Turvey, M. T. (1990). Phase transitions and critical fluctuations in the visual coordination of rhythmic movements between people. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 227–247.
Schöner, G., & Kelso, J. A. S. (1988a). Dynamic pattern generation in behavioral and neural systems. *Science, 239,* 1513–1520.
Schöner, G., & Kelso, J. A. S. (1988b). Dynamic patterns in biological coordination: Theoretical strategy and new results. In J. A. S. Kelso, A. J. Mandell, & M. F. Shlesinger (Eds.), *Dynamic patterns in complex systems* (pp. 77–102). Singapore: World Scientific.
Schöner, G., Zanone, P. G., & Kelso, J. A. S. (1992). Learning as a change of coordination dynamics: Theory and experiment. *Journal of Motor Behavior, 24,* 29–48.
Stadler, M., & Kruse, P. (1990). The self-organization perspective in cognition research: Historical remarks and new experimental approaches. In H. Haken & M. Stadler (Eds.), *Synergetics of cognition* (pp. 32–51). Berlin: Springer-Verlag.
Stetson, R. H. (1951). *Motor phonetics: A study of speech movements in action* (2nd ed.). Amsterdam: North-Holland. (Reprinted in J. A. S. Kelso & K. G. Munhall [Eds.] [1988]. *R. H. Stetson's motor phonetics: A retrospective edition.* San Diego: College Hill.)
Stevens, K. N. (1972). The quantal nature of speech: Evidence from articulatory-acoustic data. In E. David, Jr., & P. Denes (Eds.), *Human communication: A unified view* (pp. 51–66). New York: McGraw-Hill.
Stevens, K. N. (1989). On the quantal nature of speech. *Journal of Phonetics, 17,* 3–45.
Thelen, E., Kelso, J. A. S., & Fogel, A. (1987). Self-organizing systems and infant motor development. *Developmental Review, 7,* 39–65.
Thelen, E., & Ulrich, B. D. (1991). Hidden skills. *Monographs of the Society for Research in Child Development, 56*(1, Serial No. 223).
Tuller, B., Giangrande, J., Kelso, J., & Ding, M. (1993). *Perceptual dynamics of pitch.* Manuscript submitted for publication.
Tuller, B., & Kelso, J. A. S. (1990). Phase transitions in speech production and their perceptual consequences. In M. Jeannerod (Ed.), *Attention and performance* (Vol. 13, pp. 429–452). Hillsdale, NJ: Erlbaum.
Tuller, B., & Kelso, J. A. S. (1991). The production and perception of syllable structure. *Journal of Speech and Hearing Research, 34,* 501–508.
Wildgen, W. (1990). Basic principles of self-organization in language. In H. Haken & M. Stadler (Eds.), *Synergetics of cognition* (pp. 415–426). Berlin: Springer-Verlag.
Williams, D., Phillips, G., & Sekuler, R. (1986). Hysteresis in the perception of motion direction as evidence for neural cooperativity. *Nature, 324,* 253–255.
Wimmers, R. H., Beek, P. J., & van Wieringen, P. C. (1992). Phase transitions in rhythmic tracking movement: A case of unilateral coupling. *Human Movement Science, 11,* 217–226.
Zanone, P. G., & Kelso, J. A. S. (1992). The evolution of behavioral attractors with learning: Nonequilibrium phase transitions. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 403–421.

Received July 24, 1992
Revision received January 28, 1993
Accepted March 2, 1993 ■