
Acoustic and phonological learning: two different dynamics?

Apprentissage acoustique et apprentissage phonologique : deux dynamiques différentes

Betty Tuller



Édition électronique

URL : <http://journals.openedition.org/msh/8193>

DOI : 10.4000/msh.8193

ISSN : 1950-6821

Éditeur

Centre d'analyse et de mathématique sociales de l'EHESS

Édition imprimée

Date de publication : 1 décembre 2007

Pagination : 127-139

ISSN : 0987-6936

Référence électronique

Betty Tuller, « Acoustic and phonological learning: two different dynamics? », *Mathématiques et sciences humaines* [En ligne], 180 | hiver 2007, mis en ligne le 21 février 2008, consulté le 23 juillet 2020. URL : <http://journals.openedition.org/msh/8193> ; DOI : <https://doi.org/10.4000/msh.8193>

ACOUSTIC AND PHONOLOGICAL LEARNING: TWO DIFFERENT DYNAMICS?¹

Betty TULLER²

RÉSUMÉ – Apprentissage acoustique et apprentissage phonologique : deux dynamiques différentes ?

Qu'est-ce qui détermine le succès ou l'échec lorsque des adultes essaient d'apprendre à percevoir des sons de parole qui ne figurent pas dans leur langue maternelle ? Dans ce travail, l'apprentissage phonologique est exploré dans un cadre théorique explicitement dynamique. Les catégories phonologiques sont conçues comme les attracteurs d'un système dynamique évoluant au fur et à mesure que l'apprentissage se déroule. Le système dynamique lui-même est structuré par le système phonologique maternel, tel que celui-ci s'est mis en place chez l'auditeur individuel. Cette approche devrait contribuer à intégrer la phonétique et la phonologie (descriptions continues et descriptions symboliques et discrètes) à l'intérieur d'un cadre théorique commun, qui devrait lui-même permettre de mieux comprendre comment les individus perçoivent des formes phonologiques nouvelles.

MOTS-CLÉS – Apprentissage phonologique, Catégorie d'apprentissage, Dynamiques de la perception, Systèmes dynamiques,

SUMMARY – *What determines success or failure when adults try to learn to hear speech sounds that are not in their native language? Here we use an explicitly dynamical framework for exploring phonological learning. Phonological categories are conceptualized as attractors in a dynamical system that evolves as learning proceeds. The dynamical system itself is structured by the native phonological system, as it is instantiated in the individual listener. This approach should help integrate phonetics and phonology (continuous and discrete symbolic descriptions) within a common framework, which in turn may elucidate how individuals perceive new phonological forms.*

KEYWORDS – Category Learning, Dynamical systems, Perceptual Dynamics, Phonological learning

1. THEORETICAL APPROACHES TO PHONOLOGICAL LEARNING

Learning a second language as an adult is typically much more difficult than in childhood [Flege, 1995; Ingram, Park, 1997; Iverson *et al.*, 2003; Kuhl, 2000]. This is not surprising given that the neuroplasticity of early development is not evident in adults [Long, 1990; Pallier *et al.*, 2003; Scovel, 2000]. However, not all speech sounds are equal; some nonnative distinctions are easier for adults to learn than others, depending on the similarity or dissimilarity of the nonnative sound to native ones. The

¹ Work Supported by NSF grant NS0414657 and the Office of Naval Research.

² Florida Atlantic University, 777 Glades Road, Boca Raton, FL 33431, tuller@ccs.fau.edu

easiest contrasts to discriminate are those that do not assimilate to any native language category (as when English speaking listeners perceive place and voice contrasts across Zulu oral clicks [Best *et al.*, 1988]). When listeners assimilate two non-native sounds into a single native category, and both sounds are perceived as equally good members of the native category, they are extremely difficult to learn to discriminate (as in the dental-retroflex contrast in Hindi [Tees, Werker, 1984]). This has been formalized by Best as a Single Category contrast in her Perceptual Assimilation Model (PAM) [Best, 1994; Best *et al.*, 2001, 1988] and by Flege [1991], whose Speech Learning Model holds that similar sounds acquire distinctiveness when the adult learner perceives sub-segmental (non-phonemic) features on which sounds differ in the native language (L1) and the second language (L2). Thus ease of learning a distinction between a native sound and a nonnative one is inversely related to their similarity [Flege *et al.*, 2003].

One possible mechanism for this inverse relation is that prototypical sounds in the native language function as “perceptual magnets” or “attractors,” causing acoustically similar tokens to be perceived as members of the same linguistic category [Kuhl, 1991; Kuhl, Iverson, 1995]. The acquisition of new speech sounds occurs on the background of the distribution of prototypes in acoustic/perceptual space. All else being equal, when nonnative sounds are attracted or drawn to the same native prototype or magnet, they will be perceptually similar to each other, and hence less discriminable from each other and from the category prototype. Conversely, when nonnative sounds are attracted to different native prototypes or magnets, they will be perceptually distinct. Importantly, in these theories the definition of the category prototype is based on a formal linguistic description, operationally defined as the average of perceptual ratings across subjects. Defined as such, one cannot truly understand individual differences in learning to distinguish nonnative speech sounds.

Young adults with similar language backgrounds exhibit significant differences in the ability to discriminate difficult nonnative speech sounds, even after identical language training [Case *et al.*, 2003; Jantzen, Tuller, in press; Jenkins, Yeni-Komshian, 1995; Polka, 1991; Pruitt *et al.*, 1998; Tuller *et al.*, in press]. This may be, in part, because adults learning to perceive new speech sounds do so within the context of their individually existing perceptual abilities, which is not identical to a system of formally described linguistic prototypes nor to the average of perceptions of a cohort. Moreover, the question remains whether a category prototype even exists independent of context and conditions since the “best exemplar” of a category changes when a different set of stimuli, or a subset of the original stimuli is used, or when the distinction of interest is presented in a different context, or when attention is directed elsewhere (e.g., [Francis, Nusbaum, 2002; Iverson, Kuhl, 1995; Nosofsky, 1984, 1986]).

The hypothesis here is that an explicitly dynamical framework may be beneficial for exploring and modeling the process of phonological learning. Phonological categories are conceptualized as attractors in a dynamical system that evolves as learning proceeds. The dynamical system itself is initially structured by the native phonological system as it is instantiated in individual listeners. Dynamical approaches take many guises but have been influential in various aspects of speech and language [Browman, Goldstein, 1986, 1991, 2000; Byrd, Saltzman, 2003; Case *et al.*, 1995; Gafos, Benus, 2006; Nguyen *et al.*, 2005; Petitot-Cocorda, 1985; Tuller, 2003; Tuller *et al.*, 1994; Vihman, 1996] as well as other areas of cognition (e.g., [Haken, 1977; Kelso, 1984, 1995; Kugler, Turvey, 1987; McClelland *et al.*, 1986, Schöner, Kelso, 1988]).

2. SPEECH AS A NONLINEAR DYNAMICAL SYSTEM.

A dynamical system is one that evolves over time such that its present state always depends in some rule-governed way on previous states. Differential equations or maps of essential variables offer a mathematical description of how a behavior's essential parameters change as time passes and contextual parameters change (spectral composition, rate of presentation/production etc.). In such a system, phonological categories are equivalent to attractors (stable behaviors of the system) and reliably perceiving *new* phonological categories means the emergence of additional attractors. Attractors in this sense not only warp the parametric space so as to attract nearby tokens (as in native language magnets; [Kuhl, 1991]) but they have several other testable attributes that influence both pattern stability and change. For example, there exist ranges of acoustic parameter variation within which the perceptual form remains relatively stable (i.e. is resistant to change as a function of parameter variation or noise). In other ranges, however, even small variations in the acoustic parameter can cause large (nonlinear) changes in categorization of the input and the changes are hastened in the presence of noise. At these critical values, the existing attractor(s) lose stability and the observed behaviors may change gradually or abruptly as new attractors form. Abrupt, or qualitative, changes are called phase transitions or bifurcations.

Earlier work explored parametric variation of the acoustic composition of the stimuli [Case *et al.*, 1995; Tuller *et al.*, 1994] or of articulatory parameters [Benus *et al.*, 2004; Browman, Goldstein, 1995; Saltzman, Munhall, 1989]. The transition between phonological categories in this conceptualization is an especially important point for uncovering the basic dynamics of speech categorization because the transition points reveal the underlying dynamics [Case *et al.*, 1995; Tuller *et al.*, 1994]. Thus, the transition between categorizations may also be an important entry point for understanding the acquisition of nonnative sounds [Tuller, Jantzen, Jirsa, in press].

3. DYNAMIC PERCEPTUAL LEARNING: AUDITION AND NATIVE LANGUAGE

A dynamical account of learning explicitly recognizes that learning occurs over time and that perceptual change may be continuously dependent on the changing language environment. There are many examples of continuous learning in audition. For example, when adult humans adapt to lenses that compress the visual spatial field by half, adaptive changes in sound localization quickly occur although there is no alteration of the auditory input itself [Zwiers *et al.*, 2003]. Along the same lines, Lackner and Shenker [1985] have shown that when the muscle spindles of the biceps muscle of one arm are activated by a vibratory stimulus, it gives the illusion of forearm extension. A sound source attached to the finger of the vibrated limb also appears to move with the perceived change in finger location. When subjects point to the sound with their other hand, they initially mislocalize the sound source. With visual or tactile information, and a little practice, subjects re-calibrate such that the acoustic localization of the sound source is once again very accurate. Importantly, when the vibration stops, there is an immediate negative aftereffect such that the sound source is again mislocalized but in the opposite direction. It is important to realize that these examples demonstrate auditory perceptual learning in adults that is constantly and continuously modified based on new information. Perceptual adaptation and negative aftereffects suggest the flexibility of the learning process over even very short durations, calling into question the more traditional definition of learning as a persistent change in behavior.

These ideas converge with our view of categorization as a dynamical process (cf. [Lackner, Tuller, in press]). For example, Tuller, Case, Ding, and Kelso [1994] demonstrated dynamical effects in speech perception when acoustic changes yield shifts in perceptual categories. When the duration of a silent gap between an “s” and the following vowel in “say” increases in small steps, a shift between perception of “say” and “stay” occurs. Signature properties of dynamical systems (e.g. hysteresis) are observed depending on the direction of changing gap duration and initial percept. A theoretical model attributed the shift in category to loss of stability in the initial attractor and offered predictions that were later tested and confirmed [Case, Tuller, Ding, Kelso, 1995]. Thus, the enhancement of noise causes perception to switch to the alternative (“say” or “stay”) earlier than when noise is minimized. The switching between categories was modeled as the appearance and disappearance of attractive states in the underlying dynamical system. Changes in perceived category occur when the attractor corresponding to the initial category loses stability.

4. DYNAMIC PERCEPTUAL LEARNING: NONNATIVE SPEECH SOUNDS

Several demonstrations exist of continuous modifiability of nonnative speech patterns, although the explanations have not been explicitly dynamical. For example, Sancier, Fowler [1997] describe perceptually guided changes in production of voiceless stop consonants by a Brazilian-English bilingual speaker as a function of language environment. The authors demonstrate that the adult speaker’s productions tended to drift toward the ambient language characteristics (in a manner reminiscent of pre-linguistic infants; e.g. [Boysson-Bardies, Vihman, 1991; Whalen, Levitt, Wang, 1991]). The speaker’s productions shifted back and forth despite (or due to) multiple trips between the US and Brazil, indicating that learning continuously occurred. Flege [1987] reports similarly continuous changes in production by native speakers of French and English living, respectively, in Chicago and Paris for at least 12 years.

To recast these data in dynamical terms, one would view the potential function for speech as a composite of the attractors for the native and nonnative sounds (e.g., for the native unaspirated voiceless stop in Portuguese and the aspirated voiceless stop in English). When the individual is in an environment in which, for example, the nonnative sound predominates, its influence on the composite becomes stronger over time. Note that this view of the basis of the continuous modifiability of speech is consonant with the idea that the statistical distribution of inputs shapes an individual’s perceptions. If statistical properties of individuals’ linguistic inputs differ in the same native language environment, or the relative distribution changes with language environment, there may be different attraction patterns on new (nonnative) acoustic inputs and both perception and production patterns (e.g. [Iverson, Kuhl, 1996; Iverson *et al.*, 2003; Kuhl, 2000; Kluender *et al.* 1998; Pierrehumbert, 2001, 2002]).

Preliminary work by Case, Tuller, and Kelso [2003] provides a demonstration that an explicitly dynamical approach may be used to predict and understand aspects of learning to categorize new speech sounds. When an individual listener initially perceived a nonnative sound as “different” from a native one, although perhaps still acceptable as an exemplar of the native category, the rate of change of the landscape to include the sound to be learned (the progressive stabilization of the new sound) was relatively fast. In contrast, another listener who initially perceived the nonnative sound as indistinguishable from a native one showed an increasing strength of the attraction of the to-be-learned sound until a qualitative change (a bifurcation, or phase transition) reflected the emergence of a new attractor. The rate of change of the perceptual space to

the new sound was slow and the bifurcation was marked by high variability. Theoretically, the former case represents cooperation between the initial perceptual landscape and the sound to-be-learned, whereas the latter puts the two in competition. The bottom line is that there is more than one way to learn to perceive nonnative sounds, and the strategy used is in large part determined by the perceptual abilities of the individual learner. In what follows, we evaluate the generality of these patterns, especially the availability of different modes of learning, and determine whether they can be understood within a common dynamical framework.

5. A DYNAMICAL ACCOUNT OF INDIVIDUAL MODES OF LEARNING

In Case, Tuller, Ding, Kelso [1995], the categorization of native speech sounds changed when the attractor corresponding to the initial category lost stability. Here we extend this dynamical framework to investigate learning to perceive nonnative speech sounds. Specific questions include whether the initial attractor layout (the structure of an individual's initial perceptual space) has predictive value in determining whether the space reorganizes to include a new, non-native speech sound, how this reorganization occurs over time (is it smooth or does it entail a bifurcation), and whether the reorganization is based on the same information for different learners.

We have explored the patterns of speech perceptual learning in over twenty subjects in several experiments [Case *et al.*, 2003; Tuller, Jantzen, Jirsa, in press; Jantzen, Tuller, 2006]. Here we present data from two monolingual American English (AE) subjects learning to perceive the difference between the voiced alveolar stop consonant /d/, which is phonemic in English, and the voiced dental stop consonant /ɖ/, which is not. The major articulatory distinction between these two sounds is in place of articulation--in /d/ the tongue tip is placed against the alveolar ridge, while in /ɖ/, the tongue tip is against the upper front teeth.

First we determined how individual AE subjects perceived the nonnative and native speech sounds. To this end, we created a synthetic continuum that included 11 consonant-vowel syllables, with the initial consonant ranging from the voiced unaspirated dental /ɖ/ to the voiced unaspirated alveolar /d/, followed by a 215ms vowel /a/. The syllables were each 315ms in duration and differed in the second (F2) and third (F3) formant frequencies. Synthesis details may be found in Tuller, Jantzen, Jirsa [in press]. A parallel unvoiced continuum that ranged from /t/ to /t̪/ was created and used at the end of training with the voiced continuum in order to assess generalization of learning to a featurally close distinction. Five native Malayalam speakers evaluated the synthetic stimuli on the dental end of the voiced and voiceless continua as being good exemplars of the Malayalam dental stop consonants.

The synthetic stimuli were used in a "perceptual mapping," completed before the subject had any exposure to natural dental stimuli. The perceptual mapping included several tasks: a two-alternative forced-choice identification task, judged goodness as an exemplar of the native alveolar, and judged goodness as an exemplar of the nonnative dental. The combination of tasks provides a more complete assessment of perception pre- and post-training than any one task alone.

The day after initial perceptual mapping, subjects began their training sessions, which were performed daily (with the exception of weekends and holidays) for 15 days and lasted approximately 1 hour. Each training session consisted of a free exploration period where the subject was able to listen and compare digitized files of words produced by five native Malayalam speakers. Malayalam contains both the voiced

alveolar stop consonant /d/, as in English, *and* the voiced dental stop consonant /d/ [Jongman, Blumstein, Lahiri, 1985]. The words included variability in target consonant position and word length and also included two-word stimuli in which the 2nd word began with the target consonant. After 10 minutes of free exploration with the words, a single stimulus from the naturally produced training set was presented for identification as either dental or alveolar. When the response was correct, the participant heard, "Correct". If the response was incorrect, the participant heard "Incorrect, listen again" and the stimulus was replayed before the next stimulus was presented. This paradigm of 2AFC with feedback continued for 200 trials.

5.1. THE INDIVIDUAL'S INITIAL PERCEPTUAL ABILITY STRONGLY INFLUENCED LEARNING

Performance on the perceptual mapping completed in the one-hour session just prior to the first training session was compared with a second mapping completed just after the last (15th) training session. All subjects who learned to categorize the nonnative, dental stimuli as reliably different from the native, alveolar stimuli had some initial structuring of perception that was biased in the correct direction. This initial structuring could be apparent from the identification task, the judged goodness task, or both. However, subjects followed one of 3 patterns:

- *Pattern 1 (P1)*. Half the subjects who were able to learn the dental-alveolar distinction showed an initial perceptual pattern exemplified by the subject whose data are shown in Figure 1. In P1, subjects usually identified stimuli on the dental end of the synthetic stimulus continuum as dental and the alveolar-end stimuli as alveolar before any training occurred (Figure 1a). For the subject illustrated, all but three midrange stimuli (stimuli 7, 8, and 9) were identified at levels significantly different from chance ($p < .05$). After training, the identification function is essentially a more extreme version of the pre-training results. Results for the judged goodness tasks are based on a criterion of responses falling outside the 95 % confidence limit centered at the neutral judgment value of 4. Before any training with the dental occurred, the two most extreme stimuli from the dental end of the continuum are perceived as somewhat better dentals than midrange stimuli or those on the alveolar end of the continuum (Figure 1b).

Post-training, all five dental end stimuli are considered good or very good exemplars of the dental, while the five alveolar end stimuli are considered poor or very poor dentals. When judging stimuli as exemplars of the native category, there is a strong initial distinction among the stimuli that increases post-training: three stimuli closest to the dental end of the continuum are initially perceived as poor exemplars of the native category but after training, five dental-end stimuli are now considered poor alveolars. Thus in Pattern 1 there is evidence of *pre-training* ability to categorize the nonnative dental in the identification task and for structuring within both dental and alveolar categories.

- *Pattern 2 (P2)*: The second pattern exhibited by subjects who were able to learn the dental-alveolar distinction is exemplified by the subject whose data are shown in the second row of Figure 1. The pre-training identification pattern is similar to that of P1 in that extreme stimuli are identified at levels significantly better than chance. Also similar to P1, the post-training identification is a more extreme version of the pre-training pattern, with all responses significantly different from chance in the appropriate direction. However, ratings of stimuli show far less structuring in P2 than in P1. In P2, neither judged goodness task differentiated among the stimuli before

training. Thus Pattern 2 is characterized by evidence of pre-training *categorization* of the dental but with no defined internal category structure.

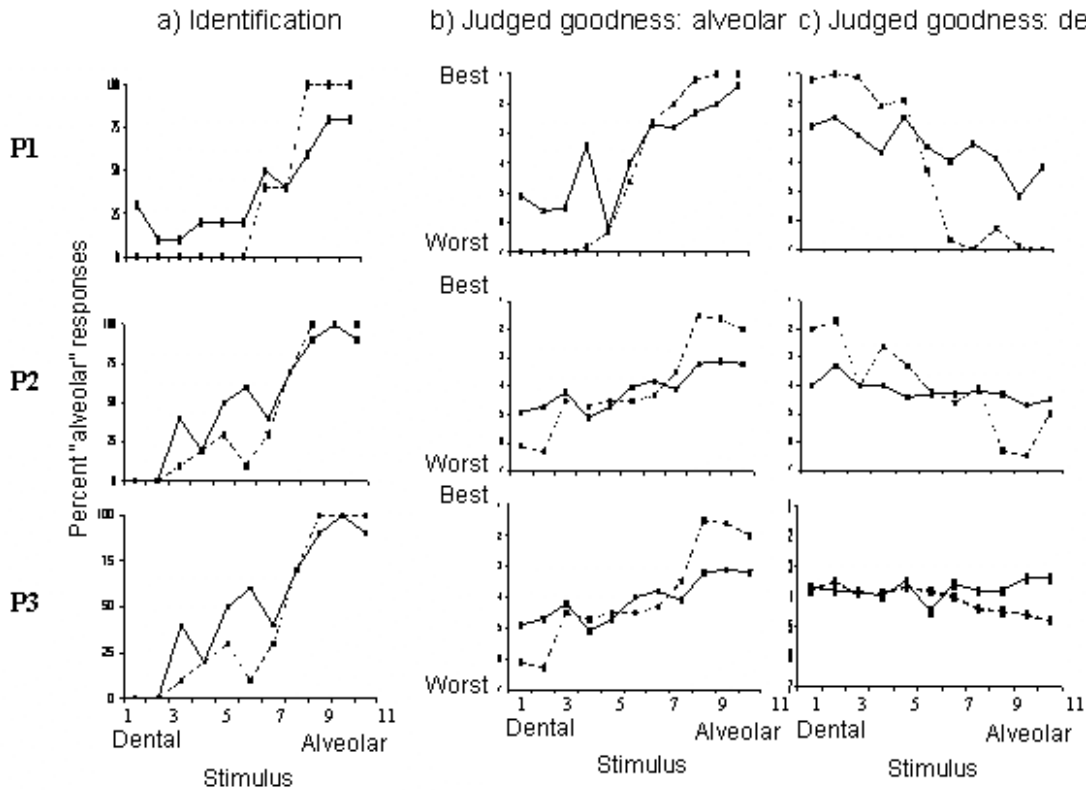


FIGURE 1. a) identification; b) judged goodness as an alveolar (native) stop consonant and c) judged goodness as a dental (nonnative) stop consonant for pattern 1 (p1: top), pattern 2 (p2: middle); and pattern 3 (p3: bottom).

Solid line: pre-training. dashed line: post-training.

- *Pattern 3: Non-learners.* Of course, not all subjects were able to learn to categorize stimuli reliably as dental. Results for a typical “non-learner” are shown in Figure 1, bottom. Note that both pre- and post-training, the subject cannot reliably identify the continuum stimuli as dental or alveolar, all stimuli are judged as “neutral” in the judged goodness tasks (neither good nor bad exemplars of the alveolar or dental stop consonant), and there is no improvement with training.

5.2. INITIAL PERCEPTUAL PATTERNS (P1 AND P2) RESULTED IN DIFFERENT MODES OF LEARNING (ACOUSTIC AND PHONOLOGICAL).

An additional task that required the listener to rate differences between stimuli occurred twice during each training session and responses were used in a multidimensional scaling (MDS) analysis. Day-by-day changes in the MDS solutions and the variability across days provide information regarding the formation of a new phonological category and whether a phase transition in perception occurred as learning progressed over time.

Figure 2 shows the MDS scores for each stimulus as training progressed, for a subject showing initial pattern P1 (left) and one showing initial pattern P2 (right). First

note that for both subjects on Day 1, before the phonemic distinction is learned, stimuli in perceptual space do not follow the physical acoustic ordering of stimuli and no clear grouping of stimuli is apparent. For P1, however, the stimuli partition over training days. Within-category distinctions shrink and are non-linearly related to acoustic differences; across-category distinctions are perceptually larger than would be expected on the basis of the acoustics alone. This pattern suggests true phonological learning.

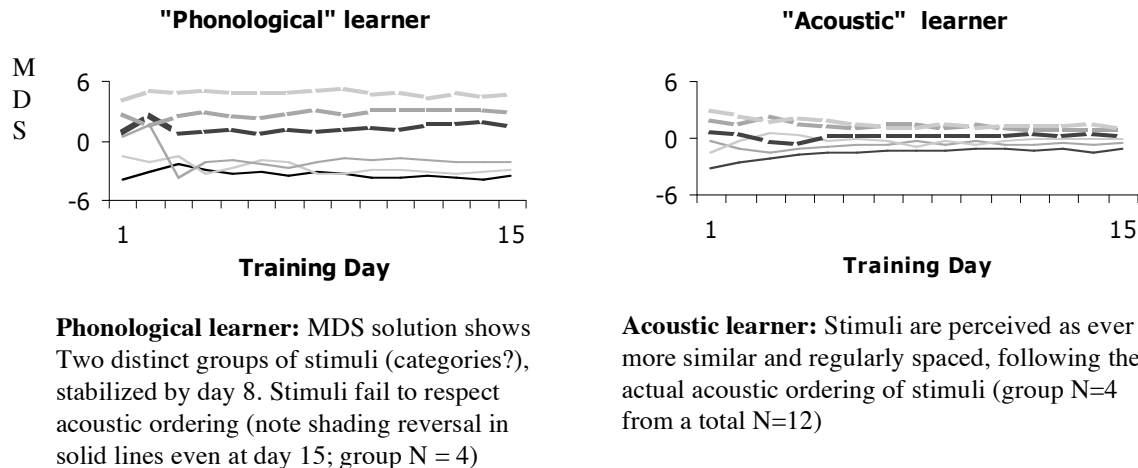


FIGURE 2. Absolute value of the MDS score for each stimulus during training for P1 (left) and P2 (right)

The day-to-day changes in the MDS solution for P2 are very different from those in P1. As training progresses, stimuli order according to acoustic composition (Figure 2, right). Importantly, there is no evidence for a bimodal warping of perceptual space. Instead, across-category distinctions are what would be expected on the basis of the acoustics alone. Although stimuli are systematically labeled as dental or alveolar, the labeling is most likely a criterion threshold determined on the basis of sensitivity to small acoustic differences, a sensitivity that increases with training. Therefore, this second pattern is termed "acoustic learning".

5.3. CHANGES IN PERCEPTUAL SPACE OVER TIME.

The day-to-day variability in perceptual space was calculated by subtracting the Day 1 dimension score from the MDS analysis for each stimulus from the Day 2 score, the Day 2 from the Day 3 score, and so forth. The squared difference scores were then summed and the standard deviation for each interval found. Case *et al.* [2003] found that order of pair elements has a strong influence on the difference ratings; therefore, MDS solutions were obtained separately for stimulus pairs with the acoustically more dental stimulus presented first and for those with the acoustically more alveolar stimulus presented first. The MDS results showed that a more native (alveolar) token provided a stronger attractive force on nonnative tokens than the other way around. The perceptual stability of native tokens had to be de-stabilized in order for the dental and alveolar sounds to separate into distinct categories. This process was evident in an increase in variability of the MDS over the first few days of training, followed by an abrupt drop in variability as the categories separated, *only* when the more native token was presented first in the pair (the thicker line in Figure 3). Variability did not show a local peak when

the more nonnative sound was presented first in the pair (the thinner line in Figure 3). This may be understood as a bifurcation and has been modeled as such (Tuller, Jantzen, & Jirsa, in press). Importantly, the attractive influence of native category tokens was evident for “phonological learners” only (i.e. learners with initial perceptual pattern P1). “Acoustic learners” had smaller day-to-day variability in the MDS solutions and did not show an effect of stimulus order or evidence of a bifurcation.

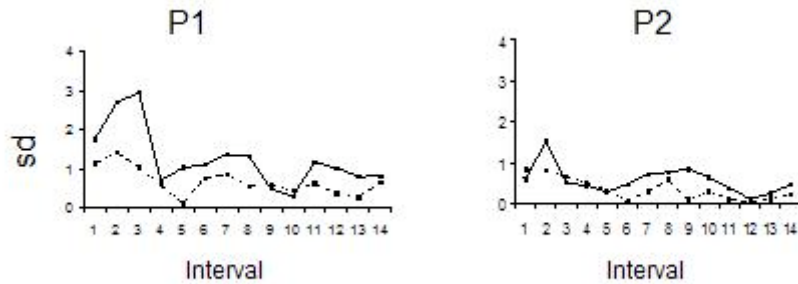


FIGURE 3. Day-to-day variability in the MDS solutions when a more dental token is presented first (solid line) and a more alveolar token first (dashed line) for Pattern 1 (left) and Pattern 2 (right).

6. CONCLUSIONS

The MDS values over time and their variability over time provide converging evidence for two modes of learning, one mode based on increasing sensitivity to acoustic differences and the other based on learning what constitutes a phonological category. In traditional psychological theory, acoustic learners would be described as utilizing an exemplar-based mode of categorization, in part because stimulus confusability decreases over time. In contrast, phonological learners responses form distinct dental and alveolar groups that do not respect the acoustic ordering of the stimuli, thus increasing the confusability of within-group stimuli. Thus rule- or integration-based explanations are more likely to be invoked. In dynamical terms, however, only phonological learners act consistent with the description of self-organization of perceptual space with a transition, or bifurcation, as an attractor is created (see [Nguyen *et al.*, in preparation] for more discussion of abstractionist and exemplar based accounts of phonological representations).

In our work, both phonological learners and acoustic learners did, in fact, learn to perceive a nonnative distinction reliably. This forces the question of how to define categories and understand how they form. A category may be considered as the interaction of two levels of representation, a discrete level and a parametric level [Luce, Galanter, 1963]. In this view, an abstract phonological category (a discrete entity) is defined by a probability density function over parameters (as in the well-known density distribution of vowels in the parametric space defined by F1 and F2 [Hillenbrand *et al.*, 1995; Peterson, Barney, 1952]. Much research on categorization in speech assumes (often implicitly) that the set of such density distributions, whether based on perception, production, or both, defines a phonological system for a given language. Note that although categories are discrete and language specific, they are operationally derived from the average, or the centroid, of each density distribution (i.e., exemplars of the continuous phonetic implementation). Thus a major issue has been how to reconcile the

discrete, symbolic realm of phonology with the continuous, dynamical realm of phonetics (e.g., [Beckman, Kingston, 1990; Hallé *et al.*, 1999; Ladefoged, 1988; Ohala, 1990; Pierrehumbert, 2006]). Theories also typically differ on how new categories form. Do perceiver's use a general rule to incorporate new tokens (rule-based, or information-integration learning [Nosofsky *et al.*, 1994] or do the tokens collectively determine the category (exemplar-based learning [Nosofsky, 1986; Coleman, 2002; Pierrehumbert, 2001])? An exemplar-based model has also been extended to learning to produce new speech sounds, with the average properties of the exemplars in a labeled category constituting the production goal [Pierrehumbert, 2001, 2002]. Note that this potentially links the perception and production of nonnative categories in a nontrivial way [Flege *et al.*, 1999; Hardison, 2005; Sancier, Fowler, 1997].

The work presented here suggests that a nonlinear dynamic approach may help integrate phonetics and phonology (continuous and discrete descriptions) within a common framework which in turn may elucidate how new phonological categories form. The approach also forms the basis for a hybrid view of phonological representations that are both abstract yet influenced by fine phonetic detail of the experienced exemplars [Nguyen *et al.*, 2007]. In short, perceptual learning for nonnative speech sounds is viewed as a process that modifies the existing dynamical system [Case *et al.*, 2003; Kelso, 1995; Lackner, Tuller, in press; Tuller, 2003; Tuller *et al.*, in press; Jantzen, Tuller, in press]. This perspective allows predictions about how learning will proceed depending on how the stimuli are initially perceived by the individual learner and may lead to individualized strategies for enhancing learning. Operationally, the progressive stabilization of an attractor corresponding to a new phonological form and whether or not a bifurcation is observed during learning is dependent on the initial perceptual landscape. Thus there is more than one route to learning, each with its own dynamics.

Acknowledgments. The author would like to thank McNeel Gordon Jantzen for assistance with data collection and analysis. Work supported by NSF grant NS0414657 and the Office of Naval Research.

REFERENCES

- BECKMAN M. E., KINGSTON J., "Introduction", J. Kingston, M. E. Beckman (eds), *Papers in laboratory phonology I: Between the grammar and the physics of speech*, Cambridge (UK), Cambridge University Press, 1990, p. 1-16.
- BENUS S., GAFOS A., GOLDSTEIN L., "Phonetics and phonology of transparent vowels in Hungarian", P. M. Nowak, C. Yoquelet, D. Mortensen (eds), *Proceedings of the 29th annual meeting of the Berkeley Linguistic Society*, Berkeley (CA), Berkeley Linguistic Society, 2004, p. 485-497.
- BEST C. T., "The emergence of native-language phonological influences in infants: A perceptual assimilation model", J. C. Goodman, H. C. Nusbaum (eds), *The development of speech perception: The transition from speech sounds to spoken word*, Cambridge (MA), MIT Press, 1994, p. 167-224.
- BEST C.T., MCROBERTS G., SITHOLE N., "Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants", *Journal of Experimental Psychology: Human Perception and Performance* 14, 1988, p. 345-360.
- BEST C.T., MCROBERTS G., GOODELL E., "Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system", *Journal of the Acoustical Society of America* 109, 2001, p. 775-794.

- BOYSSON-BARDIES B. (de), VIHMAN M.M., "Adaptation to language: Evidence from babbling and first words in four languages", *Language* 67, 1991, p. 297-319.
- BROWMAN C. P., GOLDSTEIN L., "Dynamics and articulatory phonology", R. F. Port, T. van Gelder (eds), *Mind as motion*, Cambridge (MA), MIT Press, 1995, p. 175-193.
- BROWMAN C. P., GOLDSTEIN L., "Competing constraints on intergestural coordination and self-organization of phonological structures", *Les Cahiers de l'ICP, Bulletin de la Communication Parlée* 5, 2000, p. 25-34.
- BROWMAN C.P., GOLDSTEIN L., "Towards an articulatory phonology", *Phonology Yearbook* 3, 1986, p. 219-252.
- BYRD, D., SALTZMAN E., "The elastic phrase: Modeling the dynamics of boundary-adjacent lengthening", *Journal of Phonetics* 31, 2003, p.149-180.
- CASE P., TULLER B., KELSO J. A. S., "The dynamics of learning to hear new speech sounds", *Speech Pathology*, Nov 17, 2003.
http://www.speechpathology.com/articles/arc_disp.asp?article_id=50&catid=560
- CASE P., TULLER B., DING M., KELSO J. A. S., "Evaluation of a dynamical model of speech perception", *Perception and Psychophysics* 57, 1995, p. 977-988.
- COLEMAN J., "Phonetic representations in the mental lexicon", J. Durand, B. Laks (eds.), *Phonetics, phonology and cognition*, Oxford, Oxford University Press, 2002, p. 96-130.
- FLEGE J. E., "Second language speech learning: Theory, findings, and problems", W. Strange (ed.), *Speech perception and linguistic experience: Issues in cross-language research*, Baltimore (MD), York Press., 1995, p. 233-277.
- FLEGE J., MACKAY I., MEADOR D., "Native Italian speakers' production and perception of English vowels", *Journal of the Acoustical Society of America* 106, 1999, p. 2973-2987.
- FLEGE J., SCHIRRU C., MACKAY I., "Interaction between the native and second language phonetic subsystems", *Speech Communication* 40, 2003, p. 467-491.
- FLEGE J.E., "The production of 'new' and 'similar' phones in a foreign language: Evidence for the effect of equivalence classification", *Journal of Phonetics* 15, 1987, p. 47-65.
- FLEGE J.E., "Age of learning affects the authenticity of voice onset time (VOT) in stop consonants produced in a second language", *Journal of the Acoustical Society of America* 89, 1991, p. 395-411.
- FRANCIS A.L., NUSBAUM H.C., "Selective attention and the acquisition of new phonetic categories", *Journal of Experimental Psychology: Human Perception and Performance* 28, 2002, p. 349-366.
- GAFOS A.I., BENUS S., "Dynamics of phonological cognition", *Cognitive Science* 30, 2006, p. 1-39.
- HAKEN H., *Synergetics: An introduction*, Heidelberg (Germany), Springer-Verlag, 1977.
- HALLE P. A., BEST C. T., LEVITT A., "Phonetic vs. phonological influences on French listeners' perception of American English approximants", *Journal of Phonetics* 27, 1999, p. 281-306.
- HARDISON D.M., "Second-language spoken word identification: Effects of perceptual training, visual cues, and phonetic environment", *Applied Psycholinguistics* 26, 2005, p. 579-596.
- HILLENBRAND J., GETTY L.A., CLARK M.J., WHEELER K., "Acoustic characteristics of American English vowels", *Journal of the Acoustical Society of America* 97, 1995, p. 3099-3111.
- INGRAM C. L., PARK S.-G., "Cross-language vowel perception and production by Japanese and Korean learners of English", *Journal of Phonetics* 25, 1997, p. 343-370.
- IVERSON P., KUHL P.K., "Mapping the perceptual magnet effect for speech using signal-detection-theory and multidimensional-scaling", *Journal of the Acoustical Society of America* 97, 1995, p. 553-562.
- IVERSON P., KUHL P. K., "Influences of phonetic identification and category goodness on American listeners' perception of /r/ and /l/", *Journal of the Acoustical Society of America* 99, 1996, p. 1130-1140.
- IVERSON P., KUHL P.K. AKAHANE-YAMADA R., DIESCH E., TOHKURA Y., KETTERMANN A., SIEBERT C., "A perceptual interference account of acquisition difficulties for non-native phonemes", *Cognition* 87, 2003, p. B47-B57.

- JANTZEN M.G., TULLER B., "Learning to perceive nonnative speech sounds: The role of test stimulus variability", *4th joint meeting of the Acoustical Society of America and the Acoustical Society of Japan*, Hawaii, Dec 1, 2006.
- JANTZEN M.G., TULLER B. "Towards an understanding of individual differences in phonological learning", A. Fuchs, V. Jirsa (eds), *Coordination Dynamics 2007*, Springer [in press].
- JENKINS J.J., YENI-KOMSHIAN G. H., "Cross-language speech perception: perspective and promise", W. Strange (ed.), *Speech perception and linguistic experience: Issues in cross-language research*, Baltimore (MD), York Press, 1995, p. 463-479.
- JONGMAN A., BLUMSTEIN S.E., LAHIRI A., "Acoustic properties for dental and alveolar stop consonants - a cross-language study", *Journal of Phonetics* 13, 1985, p. 235-251.
- KELSO J. A. S., *Dynamic patterns: The self-organization of brain and behavior*, Cambridge (MA) MIT Press, 1995.
- KELSO J. A. S., "Phase transitions and critical behavior in human bimanual coordination", *American Journal of Physiology: Regulatory, Integrative and Comparative* 15, 1984, p. R1000-R1004.
- KLUENDER K.R., LOTTO A.J., HOLT L.L., BLOEDEL S.L., "Role of experience for language-specific functional mappings of vowel sounds", *Journal of the Acoustical Society of America* 104, 1998, p. 3568-3582.
- KUGLER P. N., TURVEY M. T., *Information, natural law, and the self-assembly of rhythmic movement*. Hillsdale (NJ), Lawrence Erlbaum Associates, Inc. 1987.
- KUHL P. K., "Language, mind, and brain: Experience alters perception", M. S. Gazzaniga (ed.), *The new cognitive neurosciences* [2nd ed.], Cambridge (MA), MIT Press, 2000, p. 99-115.
- KUHL P.A., IVERSON P., "Mapping the perceptual magnet effect for speech using signal detection theory and multidimensional scaling", *Journal of the Acoustical Society of America* 97, 1995, p. 553-562.
- KUHL P.A., "Human adults and human infants show a 'perceptual magnet effect' for the prototypes of speech categories, monkeys do not", *Perception & Psychophysics* 50, 1991, p. 93-107.
- LACKNER J.R., TULLER B., "Dynamical systems and internal models", A. Fuchs, V. Jirsa (eds), *Coordination Dynamics 2007*, Springer, [in press].
- LACKNER J.R., SHENKER B., "Proprioceptive influences on auditory and visual spatial localization", *The Journal of Neuroscience* 5, 1985, p. 579-583.
- LADEFOGED P., "The many interfaces between phonetics and phonology", *UCLA Working Papers in Phonetics* 70, 1988, p. 13-23.
- LONG M., "Maturation constraints on language development", *Studies in Second Language Acquisition* 12, 1990, p. 251-285.
- LUCE R.D., GALANTER E., "Discrimination", R. D. Luce, R. R. Bush, E. Galanter (eds), *Handbook of mathematical psychology*, New York, John Wiley and Sons, Inc., 1963, p. 191-243.
- MCCLELLAND J. L., RUMELHART D. E., AND THE PDP RESEARCH GROUP, *Parallel distributed processing: Explorations in the microstructure of cognition: Vol. 2. Psychological and biological models*. Cambridge (MA), MIT Press, 1986.
- NGUYEN N., LANCIA L., BERGOUNIOUX M., WAUQUIER-GRAVELINES S., TULLER B., "Role of training and short-term context effects in the identification of /s/ and /st/ in French", Hazan, V., Iverson, P. (eds), *ISCA Workshop on Plasticity in Speech Perception (PSP2005)*, London, 2005, p. A38-39.
- NGUYEN N., WAUQUIER-GRAVELINES S., TULLER B., *The dynamical approach to speech perception: From fine phonetic detail to abstract phonological categories*, I. Chitoran, C. Coupé, E. Marsico and F. Pellegrino (eds.), Mouton [in press].
- NOSOFSKY R.M., "Choice, similarity, and the context theory of classification", *Journal of Experimental Psychology: Learning Memory and Cognition* 10, 1984, p. 104-114.
- NOSOFSKY R.M., "Attention, similarity, and the identification-categorization relationship", *Journal of Experimental Psychology: General* 115, 1986, p. 39-57.

- OHALA J., "There is no interface between phonetics and phonology: A personal view", *Journal of Phonetics* 18, 1990, p. 153-171.
- PALLIER C., DEHAENE S., POLINE J.-B., LEBIHAN D., ARGENTI A.-M., DUPOUX E., MEHLER J., "Brain imaging of language plasticity in adopted adults: can a second language replace the first?", *Cerebral Cortex* 13, 2003, p. 155-162.
- PETERSON G.E., BARNEY H.L., "Control methods used in a study of the vowels", *Journal of the Acoustical Society of America*. 24, 1952, p. 175-184.
- PETITOT-COCORDA J., *Les catastrophes de la parole: de Roman Jakobson à René Thom* [The catastrophes of speech. From Roman Jakobson to René Thom], Paris, Maloine, 1985.
- PIERREHUMBERT J., "The next toolkit", *Journal of Phonetics* 34, 2006, p. 516-530.
- PIERREHUMBERT J., "Word-specific phonetics", C. Gussenhoven, N. Warner (eds.), *Laboratory phonology VII*, Berlin, Mouton de Gruyter, 2002, p. 101-140.
- PIERREHUMBERT J., "Why phonological constraints are so coarse-grained", *Language and Cognitive Processes* 16, 2001, p. 691-698.
- POLKA L., "Cross-language speech perception in adults: Phonemic, phonetic, and acoustic contributions", *Journal of the Acoustical Society of America*, 89, 1991, p. 2961-2977.
- PRUITT J. S., KAWAHARA H., AKAHANE-YAMADA R., KUBO R., "Methods of enhancing speech stimuli for perceptual training: Exaggerated articulation, context truncation, and 'STRAIGHT' re-synthesis", R. Carlson, C. Dunger, B. Granstrom, A. Oster (eds), *Speech technology in language learning*, Stockholm, 1998, p. 105-108.
- SALTZMAN E., MUNHALL K., "A dynamic approach to gestural patterning in speech production", *Ecological Psychology* 1, 1989, p. 333-382.
- SANCIER M., FOWLER C.A., "Gestural drift in a bilingual speaker of Brazilian Portuguese and English", *Journal of Phonetics* 25, 1997, p. 421-437.
- SCHÖNER G., KELSO J. A. S., "A dynamic pattern theory of behavioral change", *Journal of Theoretical Biology* 135, 1988, p. 501-524.
- SCOVEL T. "A critical review of critical period research", *Annual Review of Applied Linguistics* 20, 2000, p. 213-223.
- TEES R. C., WERKER J. F., "Perceptual flexibility: Maintenance of recovery of the ability to discriminate nonnative speech sounds", *Canadian Journal of Psychology* 38, 1984, p. 579-590.
- TULLER B., "Computational models in speech communication", *Journal of Phonetics* 31, 2003, p. 503-507.
- TULLER B., CASE P., DING M., KELSO J. A. S., "The nonlinear dynamics of speech categorization", *Journal of Experimental Psychology: Human Perception and Performance* 20, 1994, p. 1-16.
- TULLER B., JANTZEN M.G., JIRSA V., "A dynamical approach to speech categorization: Two routes to learning", *New Ideas in Psychology. Special Issue on Dynamics and Psychology* [in press].
- VIHMAN M. M., *Phonological development: The origins of language in the child*, Oxford, Basil Blackwell, 1996.
- WHALEN D.H., LEVITT A.G., WANG Q., "Intonational differences between the reduplicative babbling of French- and English-learning infants", *Journal of Child Language* 18, 1991, p. 501-516.
- ZWIERS MO, VAN OPSTAL A. J., PAIGE GD., "Plasticity in human sound localization induced by compressed spatial vision", *Nature Neuroscience* 6, 2003, p. 175-181.