University Press Potsdam

Article published in:

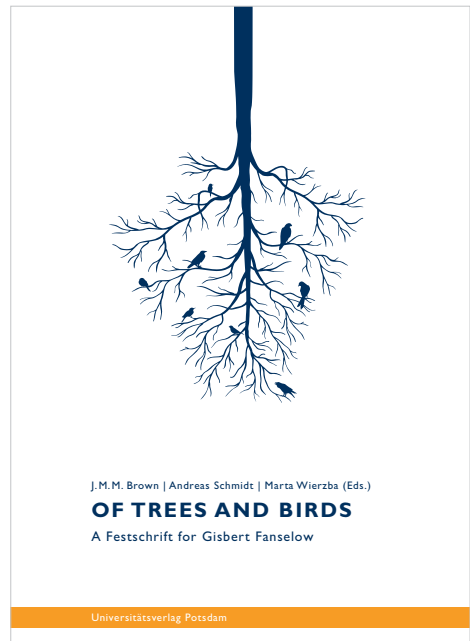J.M.M. Brown | Andreas Schmidt | Marta Wierzba (Eds.)

OF TREES AND BIRDS

A Festschrift for Gisbert Fanselow

Universitätsverlag Potsdam

Multistability in speech and other activities

Adamantios Gafos, Universität Potsdam

1 Introduction

All theories of language and speech agree on one basic working hypothesis: an infinite set of new forms or larger structures comprising chunks of speech such as syllables or whole sentences can be constructed out of a finite set of atoms, for us, consonants and vowels: "von endlichen Mitteln einen unendlichen Gebrauch machen" (von Humboldt 1836: 30) and later Chomsky's (2000: 3) "discrete infinity". Modern linguistic theory and all models of speech production have inherited a specific form of this hypothesis by assuming that the message is essentially a concatenation of speech gestures. We propose to explore an alternative hypothesis, which generalizes Humboldt's "infinite use of finite means" insight. That alternative states that concatenation is just one way to combine units; there are other ways of composing or other functional organizations of what on the surface appear to be the same sequences of units. We refer to this as the non-uniqueness of functional organization hypothesis or simply the non-uniqueness hypothesis.

Consider the task of repeating a syllable, /ta ta ta …/, a rather simple utterance. One way—the standard way for theories of language and models of speech production—of describing this task is: make a /t/ (that is, configure the vocal tract to the state required for /t/), then make an /a/, then a /t/, then an /a/, and so on. The individual instructions in this sequence may be overlapped in time to some extent ('coarticulation') but the instructions are concatenated and the control regime for flesh-

ing out this utterance has as many instructions as there are phonemes in it. An alternative way to describe the organization is: cycle through /t/ and /a/. Now, the description length for the organization underlying the utterance is independent of the utterance's length. Specifically, the organization is not one of concatenative, sequential control. Rather, it is a so-called periodic attractor giving rise to persistent cyclicity in movement (other ways of composing units become possible when we expand the scope of mathematical models to be entertained in this proposal).

What appears on the surface in the aforementioned task is a monotony of reiterant /ta/s. What lies beneath is a so far unexplored thesis: there is no isomorphism between a spoken utterance and the control regime effecting it. The question of existence and typology of (additional to concatenation) control regimes is uncharted territory. New theoretical work along with attendant experiments are required to explore it. That concatenation may not be the only principle of composition may be of value to theoretical linguistics beyond phonology and phonetics. In syntax and semantics, there exist phenomena where composition with the regular interpretation of the individual lexical items does not add up to an appropriate meaning of the whole as in the so-called pluractionals, first introduced in Newman (1980) with verbal constructions where consonant duplication conveys iteration, and reduplicative numerals (Gil 1982, 2013). The nature of the compositional mechanisms implicated in these phenomena is an open problem (e.g., for pluractionals, as in 'she wrote book after book', see Beck & von Stechow 2007, Henderson 2013). Our view brings to the forefront the notion of composition also for phonology and phonetics. It does so by offering a formal foundation on which this notion can be elaborated on in the domain of speech. Finally, enriching the notion of composition in speech leads to a number of consequences. Here we highlight one with significant potential long-term benefits. Tasks involving syllable repetitions are widely used in diagnostic and research paradigms on speech. It is a topic of a major debate whether such tasks pertain to regular speech (with clinical implications of considerable socioeconomic weight; see Ballard et al. 2000 vs. Ziegler 2002, 2003a,b, Staiger et al. 2017). Our approach enables a formal expression of the above debate and opens up ways to resolve it. Specifically, it opens up the way for a characterization of speech

movements as separate from other movements which may be similar to speech but engage distinct organizations from speech.

2 Change in functional organization

We seek to develop a novel theoretical basis for speech movements using dynamical models (a basic tool of science based on ordinary differential equations, Thompson & Stewart 2002) which, crucially for our main thesis, admit distinct regimes. In these models, each regime corresponds to an equivalence class of movement patterns with potentially distinct spatio-temporal properties and two different regimes (two equivalence classes of movements) are mathematically incommensurate; one cannot be reduced to the other. Transitions between distinct regimes are known as bifurcations, formally, qualitative changes in the topology of the phase space—the space defined by the minimum set of dimensions needed to describe a system. An example of such topological space representations is in section 4. Mathematical models with these properties will be explored for the first time in speech.

Here we offer an example of our data. In preliminary work, we used a latest generation 3D electromagnetic articulometer (Carstens, Bovenden) to track sensors attached on the lips, jaw, tongue tip and tongue body at 1250 Hz. We registered speech movements at eight distinct speech rates, from a very slow 30 beats per minute (bpm) to a very fast 570 bpm (0.5 to 9.5 Hz). In each trial, the subject spoke a sequence, as in /ta ta …/ or /ta ka ta ka …/, at the rate indicated by an audible metronome. Fig. 24.1 shows some data. For /t/, the tongue tip raises to make contact with the alveolar ridge in a closing movement (blue-shaded rectangles in Fig. 24.1), followed by an opening movement for the vowel /a/ (brown-shaded rectangles in Fig. 24.1).

Analysis of our so registered data reveals a qualitative change of disappearing asymmetries in movement properties at a rate of approximately 210 bpm (3.5 Hz). This can be seen in different ways. Below 210 bpm, opening and closing movements differ substantially with respect to peak velocity and relative time to peak velocity (the time at which peak velocity occurs as a percentage of the movement duration).
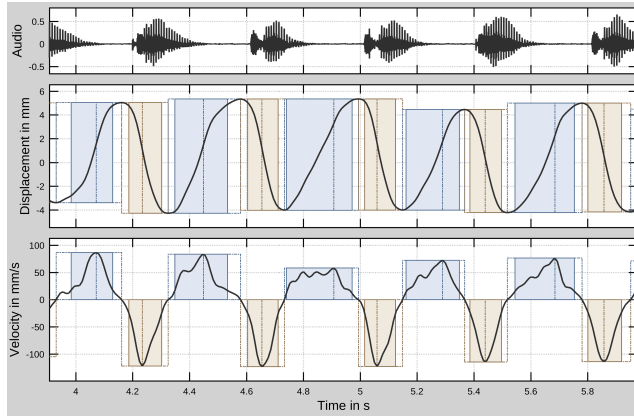
Figure 24.1: Audio (top), trajectory (displacement, on y-axis, as a function of time, on x-axis) of the tongue tip sensor during /ta ta ta …/ at 150 bpm (middle) and its velocity (bottom). Peak velocities for closing movement are lower than for opening movements (compare bottom, blue vs. brown rectangle heights). [Color image available in the online version of this article.]

At or above 210 bpm, a rate which in further experiments we expect to be speaker-specific, openings and closings converge to the same peak velocities; see Fig. 24.2, top. They also converge to the same relative time to peak velocity (0.5, that is, peak velocity is reached halfway within the opening or closing movement) of the undamped harmonic oscillator (Nelson 1983) indicating symmetrical velocity profiles; see Fig. 24.2, bottom. In other words, two movement classes change to one class: movements become cyclic at faster rates. As Schrödinger (1945) remarked, living systems are not statically stable—they are dynamically stable: the spatio-temporal order underlying what is macroscopically a /ta ta ta …/ flexibly changes as rate is increased. Both spatio-temporal orders are stable within their separate ranges of rates, a property known as multistability. The control regime at common speech rates (below 210 bpm) is a concatenation of discrete targeted movements: do a /t/, move to /a/, do another /t/ and so on. At higher rates, the mode of composition changes to a cyclic organization (see later for other modes under
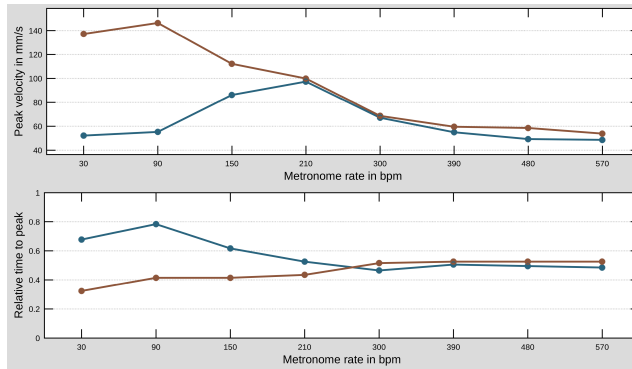
Figure 24.2: As rate increases, asymmetries in opening vs. closing movements in peak velocity (top) and relative time to peak velocity (bottom) change to symmetries: two different classes of movements converge to one class. [Color image available in the online version of this article.]

different mathematical models). Further investigation of kinematic parameters and trajectories will be performed to verify the nature of the distinct organizations. But the essential point is that there is no fixed control regime. There are different qualitative modes.

3 Informational constraints

In low speed quadruped locomotion, the limbs of the same girdle are half a period out of phase (one limb's motion is identical to that of the other limb if shifted by half of its period). But as speed increases, there is a transition from an asymmetric to a symmetric gait. Examples of such shifts in organization abound in nature (Winfree 1980). Even animals which one would not a priori think of as possessing gaits show them. For example, in flying birds, see Fig. 24.3, vortex patterns show a shift from concatenated vortex rings to a single continuous vortex ring as speed increases (the shift is seen at 7 m/s for Cockatiels and Ring Turtledoves; Hedrick et al. 2002).

However, speech is different. Speech gestures are not mere physical instantiations of organs moving. They also carry meaning: they func-
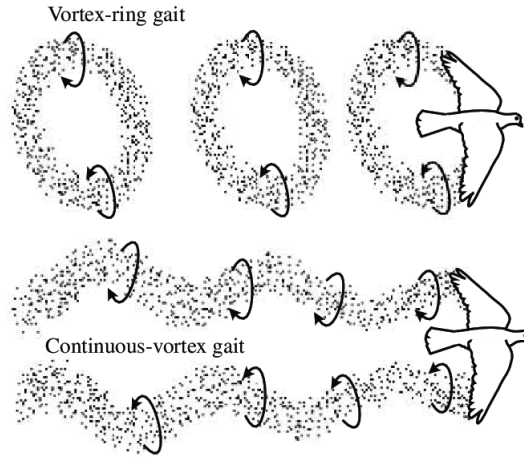
Figure 24.3: When birds fly, the vortices shed by their wings show qualitative shifts as speed of flight increases. Such reorganizations indicate that a multiplicity of qualitative modes and bifurcations among these characterize not just limbs and their kinematics properties (as in quadrupeds) but also parameters that refer to the animal-environment as a system. [Adapted with permission from Figure 1, panel A of Hedrick et al. 2002.]

tion as a message (Pattee 1969, 1972). How this communicative function shapes the movements of speech has remained a major open problem. Consider a task where participants repeat /ta ta .../ at fast rates. In a series of studies, Ziegler and colleagues (Ziegler 2002, 2003a,b, Staiger et al. 2017) have argued that such tasks, widely used in normal and disordered speech assessments, employ "distinct traits" in control, not characteristic of speech. A major debate has thus unfolded on whether such tasks pertain to speech (Ballard et al. 2000 vs. Ziegler 2003a,b). So far, neurological, clinical and experimental observations have served as the main basis of this debate.

Our research offers a theoretical foundation on which aspects of this debate can be expressed and potentially resolved. As shown in the ensuing, in models with more than one dynamical regime, the properties

of movements depend both on the model and on the regime underlying these movements. Thus, characterizing the nature of an observed movement is not possible without a parallel understanding of the control regime effecting it. In other words, Ziegler's thesis is entirely consistent with our main hypothesis. The former asks us to appreciate that organizational modes may differ across speech and non-speech orofacial movement tasks; the latter proposes to express the notion of organization by mathematical models admitting multiple regimes. However, one cannot a priori preclude a more nuanced view wherein speech and non-speech differ but may also share organizations under some parameter settings: see the next section. A major lesson from the biological world is that abstract organization in behavior is function-specific but not effector-specific (Schöner 1986: 258, Kelso 1995). In our domain, the same vocal tract organs may be governed by different dynamics in other functions (e.g., non-speech) and different organs may be governed by similar dynamics in the same function.

We will set up designs where participants utter sequences of different complexity. We will explore both meaningless and meaningful utterances, including modifying their prosody. We limit description here to one manipulation concerning vowels. In the lowest complexity utterances, participants will repeat identical syllables as in /ta ta …/ as fast as possible. In a next, higher complexity condition participants will repeat /ta tu …/ as fast as possible, then switch to /ta ta …/ as fast as possible and continue to alternate in this way. A vowel contrast is present in the latter but not in the former case. We will continue adding complexity in the form of a consonant contrast. Any qualitative differences in the movements along the complexity scale would be evidence that the introduction of complexity has consequences for the dynamics even within the space of repetition tasks which prior work argues may be in a separate class from speech altogether. There are indications that such differences do exist (e.g., see Staiger et al. 2017: 488 on their same vs. alternating consonants condition). We will also employ tasks which vary the segmental content and rhythmic structure of utterances as in Kelso et al. (1985) and Sevald & Dell (1994). We will depart from prior work in that we will register movements (not only acoustics, as Staiger et al. do for their syllable repetitions) and pursue among other methods analyses using phase space representations. An example follows.

4 Speech-like tongue movements

Do movements similar to those of speech but embedded in different biological or cultural settings abide by the same organization(s) as in speech? Work on orofacial movements has so far focused on externally accessible organs such as the jaw or the lips (mastication vs. speech: Max et al. 2003, Moore et al. 1988, Nelson et al. 1984, Ostry & Munhall 1994). For the tongue, arguably the most crucial organ for speech, some work looks into swallowing (Serrurier et al. 2012, Bennett et al. 2007). Oral movements in swallowing are partly voluntary (Green & Wang 2003), but control of tempo is infeasible. Changing tempo is crucial for us. In the perspective of our proposal, different modes in organization are revealed via continuous scaling of parameters (e.g., rate) which causes changes from one mode to another.

Music seems to offer a highly apt case. Wind-instrument players, when training, perform repetitive attacks as in /tata tata .../, called 'single tonguing', or /taka taka .../ with alternating /t/, /k/, called 'double tonguing'. Trainees are explicitly told to form speech-like /t/, /k/. Thus, speech-like movements are implicated (Bertsch & Hoole 2014). Crucially, keeping on tempo is important to trainees in honing their skill.

To assess feasibility, a trumpetist was recorded double tonguing. Movies of the tongue body and tip can be found online.[1] Prior work seeks differences between speech and other orofacial action in individual parameter values (e.g., jaw acceleration in Nelson et al. 1984). Our hypothesis shifts focus to the qualitative topology of the underlying dynamical regimes. Portraits in phase space of tongue body movements in speech vs. double tonguing across different rates (indicated in bpm within each panel) are in Fig. 24.4. Speech-/taka/ (top row), at slower and normal rates (below 300 bpm), shows a double-periodicity. This pattern disappears at or above (>=) 300 bpm. In trumpet-/taka/ (bottom row), phase portraits at slower rates show eight-figure-like shapes. These also disappear >= 300 bpm. It is clear that the phase portraits of speech and speech-like movements are distinct below (what we expect to be a participant-specific rate of about) 300 bpm. Beyond that rate,

---

1. http://www.ling.uni-potsdam.de/phonolab/trumpet/

speech- and trumpet-/taka/ appear to converge to so-called limit cycle dynamics.
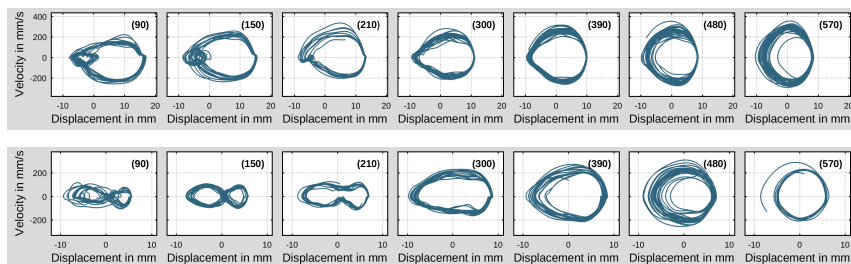

Figure 24.4: Phase portraits of /taka/ in speech (top) vs. in trumpet double tonguing (bottom). Speech and non-speech show different portraits at slower rates, but converge to a common pattern at higher rates.

Fig. 24.4 illustrates the novel viewpoint reached when the main question of this section and the debate in the previous section are expressed in topological terms. In conformity with non-uniqueness, at least three organizations seem to be involved, one for each task, speech-/taka/ and trumpet-/taka/, below a critical rate and at least one more above that rate; 'at least one' as further investigation of trajectories via analyses, as sketched later on, is needed to verify the precise nature of the dynamical regimes.

We return to the issue we were confronted with earlier. All speech involves vocal tract movement but not all vocal tract movement is speech: when is vocal tract movement not speech? A movement can serve as a message, that is, convey information, only in the context of the larger set of contrasts in which it operates. In the speech task, at rates >= 300 bpm, the vowel identity is obscured. It is not possible to discern at these rates whether the V in /tVkV/ was /a/ or /i/. The movements cannot convey contrast. This is crucial for information; Shannon (1948: 380) started his landmark paper on information theory by explaining that messages should be thought of as sequences of choices among a set of elements (for us, phonemes). It is thus not surprising that it is at these rates where the topology of what started as a speech task resembles that of the non-speech task. The movements are of course movements of the

same physical structures as in speech, but they cannot convey contrast. At normal rates, however, speech and non-speech show distinct topologies.

5 From the standard model to multistability

The standard model of the speech gesture is a special case of the linear second order system $\ddot{x} = -\omega^2(x - x_0) - 2\zeta\omega\dot{x}$ with natural frequency ω and damping ζ (Browman & Goldstein 1986, Fowler 1980, Fowler et al. 1980, Saltzman 1986). When $\zeta = 1$, this system is critically damped and approaches the so-called fixed point of the dynamics, the x_0 of the equation, but does not oscillate around it (as required for discrete movements). An appealing feature of the model is that it is intrinsically self-equilibrating: the 'target' of the system, the x_0, is achieved regardless of initial conditions. Hence, in C-V-C, where C stands for any consonant and V for any vowel, going from the C to the V and then to the second C can be achieved regardless of where in the vocal tract the constrictions (for the C) or opening (for the V) are located. This is achieved by switching between the fixed points corresponding to the vocal tract targets of each segment in a C-V-C. The standard model is a single-regime model. See Sorensen & Gafos (2016) for its phase space topology and a review of other such models; see Gafos (1999, 2002) on how the model relates to theories of phonological grammar.

Mathematical models which admit more than a single regime (e.g., Schöner 1986, Jirsa & Kelso 2005) allow for qualitatively different stable spatio-temporal patterns (multistability). Such models are of particular interest due to our non-uniqueness hypothesis. We illustrate how such models offer an appropriate starting point in seeking formal expressions for our main hypothesis. We use the Jirsa & Kelso (2005) model, henceforth JK, but the point applies equally to the Schöner model. JK is described by the second order equation $\ddot{x} = (1 - x^2)\tau\dot{x} - x - b(\dot{x}/\tau - x + x^3/3) + \alpha$ and admits three regimes: for $a \in [0, 2]$, $b \in [1, 2]$, a monostable regime governed by a single fixed point and a bistable regime with two fixed points, both regimes generating discrete movements, and a limit cycle regime which generates rhythmic movements ($a = b = 0$).

For discrete movements, JK makes predictions similar but not identical to those of the standard model. To wit, let us compare predictions about how the three kinematic variables (amplitude, peak velocity and duration) relate to one another, so-called kinematic relation predictions. For the standard model, Fig. 24.5 (top), peak velocity over amplitude is proportional to the inverse of duration and relative time to peak velocity for the typically used $\zeta = 1$ is around 0.20 (which is too low; 0.38 … 0.70 is the range seen in empirical studies as in Perkell et al. 2002). In preliminary work, JK's corresponding predictions are in Fig. 24.5 (bottom): the relation between peak velocity over amplitude and the inverse of duration is maintained but relative times to peak velocity now lift to 0.50. Furthermore, kinematic relations are regime-dependent. In the limit cycle regime (not shown in Fig. 24.5), JK predicts a nonlinear relation between peak velocity and amplitude, but a different one from that of its fixed point regime. Overall, then, kinematic relations depend on the model **and** on its dynamical regime.
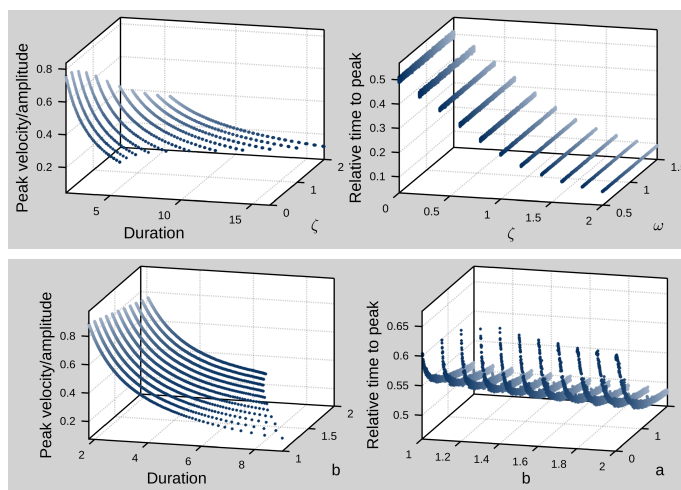

Figure 24.5: Similarities and differences in predicted kinematic relations between peak velocity, amplitude and duration of the standard (top) vs. the Jirsa & Kelso model (bottom), one of the multiple-regime models.

Consider now the ways of composing a simple utterance, a CV, in different models. In the standard model, a CV is a combination of two fixed point regimes. The same applies to JK (but as just shown JK's predicted kinematic relations are different from those of the standard model). In Schöner's (1986) model, a CV receives a very different analysis. Starting from the regime with one fixed point representing the C target, the dynamics is changed via a parameter adjustment in the model to the regime of the limit cycle; following this adjustment, the dynamics is restored to the fixed point regime now of the second V target, resulting in movement to the vowel. Kinematic relation predictions are to be worked out. The implication stated once again is: the regimes involved in effecting some sequence cannot be known a *priori*. Rigorous analyses of the different models is needed and constitutes an entirely new undertaking. In pursuing such analyses, we will use two approaches.

Such analysis can proceed in two ways. First, it is possible to make explicit the predicted kinematic relations, as exemplified in Fig. 24.5, from each candidate model by mathematical analysis, using methods from ordinary differential equations, and simulation. At this stage, no model appears to predict all the (known) kinematic relations at the same time. For example, the standard model describes well the observed kinematic relations between peak velocity and amplitude and the inverse proportionality between amplitude-normalized peak velocity and duration, but fails to offer reasonable relative times to peak velocity (Fig. 24.5, top, for $\zeta = 1$, relative time to peak is about 0.20 which is too low). There exists no prior comparative analysis of the different models.

Another way to assess models is to consider entire trajectories, that is, the time-series of position, velocity and acceleration, $(x, \dot{x}, \ddot{x})$. Let the form of any model be $\ddot{x} = f(\dot{x}, x, a, b)$, where a, b are posited parameters for that candidate model; we use two here as the JK model, one of the multiple-regime models, uses two. Via bootstrapping methods, it is possible to estimate ranges for the a, b parameters. Once such ranges are obtained, a stability analysis can be done. Here, one can ask: does change in these parameters lead to a bifurcation, that is, to a qualitative shift in the topology of the phase space of the candidate model? The presence of bifurcations would offer analytical proof for the non-uniqueness hypothesis.

6 Open questions

Our main hypothesis forces us to rethink a most familiar, in theories of language, concept of concatenation and develop extensions using topological space notions of organization. We have seen evidence that doing so promises to expand our insights on the nature of the principles that underwrite speech units. It also raises a number of specific open questions, each with attendant linguistic phenomena awaiting further understanding. We mention a few examples. Are regimes sequence (in)dependent: does /tapa/ implicate the same regime(s) as /pata/? Are regimes dependent on phonemic content: does /kita/ implicate the same regime(s) as /kata/? Relevant to these two questions are observations about the kinds and frequencies of different phonological patterns. Thus, labial-coronal consonant sequences are more numerous than coronal-labial sequences across languages as well as in acquisition when the child shifts from babbling /ma ma …/ to more wordlike shapes (MacNeilage & Davis 2000). Parallel to this is the fact that studies on fast speech show evidence that /daba/ (coronal-labial) changes to /bada/ (labial-coronal) but not the inverse (Rochet-Capellan & Schwartz 2007). A notion of sequence stability seems involved. The approach promoted is highly pertinent here. A multiple-regime model has been used to express relative stability and changes in organization as a result of movement speed in other non-speech domains (Huys et al. 2008). Consider next prosody, encompassing phenomena "characteristic not so much of the individual segments as of their relations to each other" (Catford 1977: 172). A notion of organization among units is clearly involved. Here, we can ask: does /TAka/, with stress on the first syllable as in German, engage the same organization(s) as /taKA/ with final stress as in French? In other words, are there language-specific regimes expressing foot structure and the within-foot distinction in stressed vs. unstressed syllables. Progress on the issue of how to express prominence or stress in extensions of the standard model is still at its early stages (Saltzman et al. 2008). A phase space topology approach has not been used heretofore. When topological information on phase spaces is combined with different models of the units partaking in these organizations, a new qualitative and quantitative understanding is to be expected. In sum, for each of these questions,

Bibliography

Ballard, Kirrie J., Jay P. Granier & Donald A. Robin. 2000. Understanding the nature of apraxia of speech: Theory, analysis, and treatment. *Aphasiology* 14(10). 969–995.

Beck, Sigrid & Arnim von Stechow. 2007. Pluractional adverbials. *Journal of Semantics* 24(3). 215–254.

Bennett, Janice W., Pascal H. van Lieshout & Catriona M. Steele. 2007. Tongue control for speech and swallowing in healthy younger and older adults. *International Journal of Orofacial Myology* 33. 5–18.

Bertsch, Matthias A. & Phil Hoole. 2014. Tonguing on brass instruments: High speed visualization and benchmarks of fastest tempi. In *Proceedings of the International Symposium on Musical Acoustics (ISMA)*, 408–412. Le Mans, France.

Browman, Catherine P. & Louis M. Goldstein. 1986. Towards an articulatory phonology. *Phonology Yearbook* 3. 219–252.

Catford, John C. 1977. *Fundamental problems in phonetics*. Edinburgh, Scotland: Edinburgh University Press.

Chomsky, Noam. 2000. *New horizons in the study of language and mind*. Cambridge, MA: Cambridge University Press.

Fowler, Carol A. 1980. Coarticulation and theories of extrinsic timing. *Journal of Phonetics* 8(1). 113–133.

Fowler, Carol A., Philip Rubin, Robert E. Remez & Michael T. Turvey. 1980. Implications for speech production of a general theory of action. In Brian Butterworth (ed.), *Language production: Speech and talk*, 373–420. New York City, NY: Academic Press.

Gafos, Adamantios I. 1999. *The articulatory basis of locality in phonology* (Outstanding Dissertations in Linguistics). New York City, NY: Garland.

Gafos, Adamantios I. 2002. A grammar of gestural coordination. *Natural Language & Linguistic Theory* 20(2). 269–337.

Gil, David. 1982. *Distributive numerals*. Los Angeles, CA: University of California, dissertation.

Gil, David. 2013. Distributive numerals. In Martin Haspelmath & Matthew S. Dryer (eds.), *The world atlas of language structures online*.

Leipzig, Germany: Max Planck Institute for Evolutionary Anthropology. http://wals.info/chapter/54 (5 July, 2018).

Green, Jordan R. & Yu-Tsai Wang. 2003. Tongue-surface movement patterns during speech and swallowing. *The Journal of the Acoustical Society of America* 113(5). 2820–2833.

Hedrick, Tyson L., Bret W. Tobalske & Andrew A. Biewener. 2002. Estimates of circulation and gait change based on a three-dimensional kinematic analysis of flight in cockatiels (Nymphicus hollandicus) and ringed turtle-doves (Streptopelia risoria). *Journal of Experimental Biology* 205(10). 1389–1409.

Henderson, Robert. 2013. Quantizing scalar change. In Todd Snider (ed.), *Proceedings of the 23rd Semantics and Linguistic Theory Conference (SALT)*, 473–492. Ithaca, NY: CLC Publications.

von Humboldt, Wilhelm. 1836. *Über die Verschiedenheit des menschlichen Sprachbaues.* Berlin, Germany: Druckerei der Königlichen Akademie der Wissenschaften.

Huys, Raoul, Breanna E. Studenka, Nicole L. Rheaume, Howard N. Zelaznik & Viktor K. Jirsa. 2008. Distinct timing mechanisms produce discrete and continuous movements. *PLoS Computational Biology* 4(4). e1000061.

Jirsa, Viktor K. & J. A. Scott Kelso. 2005. The excitator as a minimal model for the coordination dynamics of discrete and rhythmic movement generation. *Journal of Motor Behaviour* 37(1). 35–51.

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

Kelso, J. A. Scott, Eric Vatikiotis-Bateson, Elliot L. Saltzman & Bruce Kay. 1985. A qualitative dynamic analysis of reiterant speech production: Phase portraits, kinematics and dynamic modeling. *The Journal of the Acoustical Society of America* 77(1). 266–280.

MacNeilage, Peter F. & Barbara L. Davis. 2000. On the origin of internal structure of word forms. *Science* 288(5465). 527–531.

Max, Ludo, Anthony J. Caruso & Vincent L. Gracco. 2003. Kinematic analysis of speech, orofacial nonspeech, and finger movements in stuttering and nonstuttering adults. *Journal of Speech, Language and Hearing Research* 46(1). 215–232.

Moore, Christopher A., Anne Smith & Robert L. Ringel. 1988. Task-specific organization of activity in human jaw muscles. *Journal of Speech, Language and Hearing Research* 31(4). 670–680.

Nelson, Winston L. 1983. Physical principles for economies of skilled movements. *Biological Cybernetics* 46(2). 135–147.

Nelson, Winston L., Joseph S. Perkell & John R. Westbury. 1984. Mandible movements during increasingly rapid articulations of single syllables: preliminary observations. *The Journal of the Acoustical Society of America* 75(3). 945–951.

Newman, Paul. 1980. *The classification of Chadic within Afroasiatic.* Leiden, Netherlands: Universitaire Pers Leiden.

Ostry, David J. & Kevin G. Munhall. 1994. Control of jaw orientation and position in mastication and speech. *Journal of Neurophysiology* 71(4). 1528–1545.

Pattee, Howard H. 1969. How does a molecule become a message. *Developmental Biology Supplement* 3. 1–16.

Pattee, Howard H. 1972. Laws and constraints, symbols and languages. In Conrad H. Waddington (ed.), *Towards a theoretical biology*, 248–258. Edinburgh, Scotland: Edinburgh University Press.

Perkell, Joseph S., Majid Zandipour, Melanie L. Matthies & Harlan Lane. 2002. Economy of effort in different speaking conditions: i: a preliminary study of intersubject differences and modeling issues. *The Journal of the Acoustical Society of America* 112(4). 1627–1641.

Rochet-Capellan, Amélie & Jean-Luc Schwartz. 2007. An articulatory basis for the labial-to-coronal effect: /pata/ seems a more stable articulatory pattern than /tapa/. *The Journal of the Acoustical Society of America* 121(6). 3740–3754.

Saltzman, Elliot L. 1986. Task dynamic coordination of the speech articulators: a preliminary model. *Experimental Brain Research Series* 15. 129–144.

Saltzman, Elliot, Hosung Nam, Jelena Krivokapic & Louis Goldstein. 2008. A task-dynamic toolkit for modeling the effects of prosodic structure on articulation. In Plinio A. Barbosa, Sandra Madureira & César Reis (eds.), *Proceedings of the 4th International Conference on Speech Prosody*, 175–184. Campinas, Brazil.

Schöner, Gregor. 1986. A dynamic theory of coordination of discrete movement. *Biological Cybernetics* 63(4). 257–270.

Schrödinger, Erwin. 1945. *What is life?* London, England: Cambridge University Press.

Serrurier, Antoine, Pierre Badin, Anna Barney, Louis-Jean Boö & Christophe Savariaux. 2012. The tongue in speech and feeding: Comparative articulatory modelling. *Journal of Phonetics* 40(6). 745–763.

Sevald, Christine A. & Gary S Dell. 1994. The sequential cuing effect in speech production. *Cognition* 53(2). 91–127.

Shannon, Claude E. 1948. A mathematical theory of communication (part 1). *Bell Systems Technical Journal* 27(3). 379–423.

Sorensen, Tanner & Adamantios I. Gafos. 2016. The gesture as an autonomous nonlinear dynamical system. *Ecological Psychology* 28(4). 188–215.

Staiger, Anja, Theresa Schölderle, Bettina Brendel, Kai Bötzel & Wolfram Ziegler. 2017. Oral motor abilities are task dependent: a factor analytic approach to performance rate. *Journal of Motor Behavior* 49(5). 482–493.

Thompson, John Michael Tutill & H. Bruce Stewart. 2002. *Nonlinear dynamics and chaos.* 2nd ed. West Sussex, England: Wiley & Sons.

Winfree, Arthur T. 1980. *The geometry of biological time.* New York City, NY: Springer.

Ziegler, Wolfram. 2002. Task-related factors in oral motor control: Speech and oral diadochokinesis in dysarthria and apraxia of speech. *Brain and Language* 80(3). 556–575.

Ziegler, Wolfram. 2003a. Speech motor control is task-specific: Evidence form dysarthria and apraxia of speech. *Aphasiology* 17(1). 3–36.

Ziegler, Wolfram. 2003b. To speak or not to speak: Distinctions between speech and nonspeech motor control. *Aphasiology* 17(2). 99–105.