Articulatory representation and organization

Adamantios Gafos and Louis Goldstein

1. Introduction[1]

As a scientific enterprise, phonology is a theory of speakers' knowledge of the abstract patterns of speech events that characterize and differentiate languages. Work has largely proceeded from the assumption that the observables used to build the theory are transcriptions of speech as a sequence of segmental units. Not surprisingly, therefore, the internal representation of speech that theories of phonology have traditionally postulated is likewise a sequence of segmental units, not much different in kind from the observables themselves. With the advent of practical acoustic analysis in the 1950s, it became possible to consider using acoustic patterns as the observables, rather than relying on transcription. However, the acoustics revealed by spectrographic analysis appeared so unlike the transcriptions of speech, and so apparently incompatible with it (Harris 1953, Hockett 1955), that it was rejected as a primary observable. It was assumed that somehow a listener must re-construct a segmental transcription when listening to speech (through something like categorical perception, e.g., Liberman et al. 1967), so neither the basic observables, nor the nature of the hypothesized representations were much changed.

Two independent developments later together spawned novel work that challenged the standard view of phonological representation and its reliance on transcription for observables. These were autosegmental phonology (Goldsmith 1976) and the beginnings of availability of techniques (x-ray, magnetometer, ultrasound) for obtaining quantitative records of articulator motion during speech. Autosegmental theory hypothesized that the phonological representation was composed of multiple feature sequences whose boundaries were not necessarily aligned in time in a manner that would be implied by a single sequence of segments. While many autosegmental theorists continued to employ segmental transcriptions as the observables, others saw that the temporal structure of these representations was in many cases isomorphic with the (newly) observable structure of articulatory events (e.g., Fujimura 1981, Browman & Goldstein 1986). Thus, it became possible to use articulatory events and their timing as observables informing autosegmental-type representations.

While phonological representations are no longer seen as autosegmental, the working assumption that phonological representations are isomorphic with speech production events has produced a new, deeper, and more general understanding of several phonological phenomena. This is so partly because the isomorphism makes it possible to test specific hypotheses about representations and processes using (quantitative) articulatory observables. This has proven particularly revealing in circumstances in which the nature of the articulatory-acoustic mapping obscures an articulatory event in the acoustic (and transcription) record. The hypothesis of such “hidden” events can afford a novel, simple description of particular phonological processes, and can be tested in laboratory experiments.

2. Units of articulatory organization and units of phonological encoding

One of the major stumbling blocks in being able to systematically relate the observables of speech to phonological representations and processes is the apparent incompatibility in the nature of the entities involved at the two different levels of description. Combinatorial phonological units are discrete, qualitative, and context-invariant, while speech is continuous (in space and time) and highly context-dependent. The same issue of the relation between the qualitative and the continuous is met in (diachronic) sound change (see **Harrington, this volume**).

Advances in understanding of the coordination and control of action, beginning with Bernstein (1967) and Turvey (1977), have provided a principled way of unifying these descriptions. This approach was first applied to speech in the work of Fowler (1980) and Fowler et al. (1980), and was made explicit in the concept of speech *gesture,* as developed in the Articulatory Phonology of Browman & Goldstein (1986; 1989; 1990), and the Task Dynamics model of Saltzman (1986), Saltzman & Munhall (1989) and Kelso, Tuller & Saltzman (1986).

A gesture is a functional unit of action that achieves a specified task (most often in the case of speech, a vocal tract constriction task). Two properties of gestures are key to bridging the conceptual divides between qualitative and continuous, and between context-invariant and context-dependent. The first is the notion of *coordinative structure*. The constriction tasks of gestures are defined abstractly in the space of constriction variable goals. For example, reducing to 0 (or less) the distance between the upper and lower lips (or Lip Aperture) results in producing a lip closure gesture. The many articulatory (e.g., upper lip, lower lip, jaw) and muscular (orbicularis oris, anterior belly of the digastric, risorius, etc.) components that can contribute to this task form a coordinative structure (or functional synergy) within which they cooperate flexibly to achieve the task. This flexibility means that the task will be achieved with quantitatively different movements of the articulatory components, depending on the context in which the gesture is produced. The articulator motions are context-dependent, but the task description guiding them is invariant. The motions are not themselves gestures, but are guided by the current active gesture(s). A similar approach to contextual flexibility is also found in Guenther's (1995) neurally inspired model of speech production.

The second relevant property of gestures is that the continuous motion of a controlled task (its kinematics) is modeled as the output of a dynamical system, i.e., a system of differential equations. The signature property of such systems is that while the state (position and velocity of the object or, here, constriction) is changing continuously, the equations that give rise to the time-varying state are fixed during the lifetime of the gesture and constitute an underlying law governing this surface variability (e.g., Saltzman 1995). Most speech gestures have been hypothesized to be governed by *point-attractor* dynamics: all possible trajectories converge on a single state over time, as specified by the target, or equilibrium position parameter of the system. The hypothesized dynamics will give rise to quite different trajectories depending on the initial condition (as determined by context). Dynamical laws defining gestures thus stand at the same level of abstraction as the invariant context-independent units of representation in a purely symbolic view of phonological units. Crucially, however, it would be misleading to view the kinematics as *implementing* these (dynamical) symbols because no additional formal construct is needed to go from the dynamical law defining gestures to the continuity and context-specificity of their kinematic patterns.

While point-attractor dynamics (and the related Equilibrium Point Hypothesis, Perrier et al. 1996) provide a good first approximation to a lawful description of speech kinematics, other findings suggest that speech gestures do not have targets that are single points but rather ranges of values. The arguments for this approach have been developed most explicitly in the work of Keating (1990), wherein "windows" prescribe ranges of variability within individual articulatory dimensions, and by Guenther (1995), wherein targets are defined as convex regions in a multidimensional space of orosensory parameters (such as tongue body height with respect to jaw, tongue tip pressure receptor and so on). Within the dynamical systems model of speech gestures, for a proposal on defining targets as ranges see Saltzman & Byrd (2000) and for a different proposal on specifying targets using "activation fields" see Gafos & Kirov (2010).

Research on speech errors has also shown that the choice of observables can strongly influence conclusions about the minimal units of phonological encoding in speech production. Levelt, Roelofs & Meyer's (1999) theory of phonological encoding in speech production hypothesizes that these are wholistic, symbolic units, and one of the major sources of evidence presented for this view is the nature of (segmental) speech errors. Analyses of transcriptions of speech error corpora (Fromkin 1971, Shattuck-Hufnagel & Klatt 1979, Shattuck-Hufnagel 1983) have been argued to show that errors result from the insertion of an intended phonological segment in the wrong *slot* within a prosodic *frame* for an utterance. Apart from this misplacement of a unit (or units), an errorful utterance is assumed to be both phonotactically and phonetically well-formed. Fromkin (1971) originally argued for the abstractness of the unit involved in errors by uncovering errors in which a phonological segment is phonetically accommodated to its new position, e.g. *slumber* [p^h]*arty* → *lumber s*[p]*arty*.

The observables that were used to develop these theories of speech errors are segmental transcriptions. However, recent work reveals a very different picture when quantitative measures of speech articulation (Mowrey & McKay 1990, Pouplier 2003, 2007, 2008, Pouplier & Goldstein 2005, Goldstein et al. 2007) and/or acoustics (Frisch & Wright 2002, Goldrick & Blumstein 2006) during error production are examined. The key result is that the measured properties of a segment when it appears as a substitution in an error are usually not identical to those of the same segment when it is produced in a non-errorful utterance. In fact, the substitution combines properties of the intended and substituted segments. In the most extreme case of this (Goldstein, Pouplier et al. 2007), errors appear to involve simultaneous production of the intended constriction gesture and the substituted gesture. For example, when talkers repeat the phrase 'cop top' or 'top cop,' they produce *gestural intrusion* errors, in which the tongue dorsum gesture for /k/ and the tongue tip gesture for /t/ are coproduced (Goldstein et al. 2007). Such intrusions are the most frequent type of error observed, both in their repetition task and in a SLIP task (Pouplier, 2007) in which there is no overt repetition. These errors (and their frequency) call into question the classic arguments for segments as units of phonological encoding. First, they show that many errors are not in fact phonetically well-formed (coproduced tongue tip and tongue dorsum constrictions are not licensed in English). Second, the occurrence of gestural intrusions can be explained as qualitative shifts in systems of coupled oscillators to a dynamically more stable state (Goldstein, Pouplier et al. 2007), which relates the errors to a wider class of nonlinear phenomena. Finally, for multigestural segments such as nasals, either of the component gestures (oral constriction, velum lowering) can intrude without the other intruding (Goldstein, Pouplier et al. 2007). This argues that gestures function as units of encoding.

The fact that gestural intrusion errors are the most frequently observed error does not, of course, rule out the possibility that segments as well as gestures are units of phonological encoding. It merely shows that the classic arguments are inadequate, because they are based on purely transcriptional observables. Richer experimentation will be required to determine if there exist certain types of errors that provided unambiguous support for segments as units of phonological encoding.

3. Articulatory events and phonological processes

Another insight of the gestural approach is the idea that phonological units and processes may be realized as patterns of gestural coordination among the gestures that constitute these units or take place in phonological processes.

We will illustrate this with examples of allophonic variation, assimilation and harmony. First consider the difference between 'clear' and 'dark' allophones of English /l/, as in *lip, late, lie* versus *pill, feel, cool* ([l] versus [l□]). In a-temporal models of phonology, the difference would be expressed by saying that the basic allophone is the clear /l/ and in syllable-final position this /l/ changes to a 'dark' or velarized version by a feature-change rule adding the feature [+back]. Looking closely at this variation with the X-ray microbeam system, Sproat and Fujimura (1993) discovered that English /l/ is composed of two gestures, a tongue tip 'consonantal' gesture and a tongue dorsum 'vocalic' gesture, and that the relative timing of these varies as a function of syllable position and adjoining prosodic boundary. In syllable-initial position, the two gestures show a synchronous pattern of relative timing, with tongue tip and tongue dorsum attaining their goals at the same time. In syllable-final position, the tongue dorsum gesture significantly precedes the tongue tip gesture, with the tongue dorsum attaining its target at the onset of the tongue tip gesture. In syllable-final position, then, the acoustic portion of the syllable corresponding to the vowel is significantly more overlapped with the tongue dorsum gesture. The acoustic consequence of this difference in overlap is what gives rise to the distinction between the 'clear' and 'dark' variants of /l/ (see also Browman & Goldstein 1995). Krakow (1989, 1999), on English nasals, finds a strikingly similar pattern of timing between the component gestures of velic lowering and oral closing, and shows how the timing differences can be used to explain the allophonic variation between oral and nasalized vowels as in *meat* versus *team* (see also Byrd et al. 2009 for a recent replication of these results using real-time MRI). The insight of expressing phonological processes such as allophony as patterns of gestures and their coordination has inspired the development of grammar models based on gestural representations. In a study of the phonological system of Moroccan Colloquial Arabic, Gafos (2002) argues that phonological knowledge can make reference to the temporal dimension of linguistic form. This proposal adopts Optimality Theory (Prince & Smolensky, 1993/2004) by expressing language-particular patterns as the result of optimization under a set of violable constraints, some of which must crucially refer to temporal relations among gestures. Angermeyer (2002), Benus, Smorodinsky and Gafos (2004), Bradley (2002), Davidson (2003, 2006) and Hall (2003) also pursue a model of grammar based on gestural representations and Optimality Theory in analyzing independent phenomena in other languages.

Another area of focus has been assimilation. A sample inventory of experimental studies on local assimilation includes: Bulgarian [t d] palatalization, Wood (1996); Chaga nasal-stop sequences, Browman & Goldstein (1986); English /s/ palatalization, Zsiga (1995); English /s/ to [ʃ] assimilation, Holst & Nolan (1995), Nolan, Holst & Kühnert (1996); English ca<u>mp</u>er, ca<u>mb</u>er, Browman & Goldstein (1986); English casual speech, Browman & Goldstein (1989, 1990); Castillian Spanish nasal place assimilation, Honorof (1999); German CC clusters, Kohler (1990), Kröger (1993); Igbo vowel assimilation, Zsiga (1997); Greek vowel hiatus altenations, Baltazani (2002); Italian CC clusters, Farnetari & Busà (1994); Russian palatalization, Keating (1988); Russian coronal-dorsal sequences, Barry (1991). For a discussion of assimilation and experimental data on lexical access, see **Ernestus_chap5**.

Here we review two examples. Zsiga (1995, 1997) compared [s + j] sequences as in 'confess your', whose acoustic consequences resemble [ʃ], especially under fast speaking rates, to other [ʃ]s as in 'fresh' and 'confession'. In 'fresh' the [ʃ] is part of the mental lexicon entry. In 'confession', the [ʃ] is assumed to be derived by a lexical phonological rule of palatalization changing [s] to [ʃ] when an [s]-final verb combines with the Latinate suffix –*ion* to form its deverbal noun. Using electropalatography, Zsiga found that the tongue-palate contact pattern during the acoustic interval corresponding to the [ʃ] in 'confession' is indistinguishable from that of the [ʃ] in 'fresh'. However, in 'confe<u>ss y</u>our', tongue-palate contact patterns during the underlined portion of the utterance change in a way that reveals the bisegmental make-up of such sequences. Across word boundaries, therefore, an [ʃ]-like acoustic output arises via coarticulation; that is, as the by-product of the temporal overlap between [s] and [j]. Thus, coarticulatory overlap and the result of the presumed phonological rule of palatalization may have similar acoustic consequences, but the two can be teased apart by examining how articulation unfolds in time.

In his work on Castillian Spanish nasal place assimilation, Honorof (1999) finds that the alveolar nasal [n] assimilates completely to the place of the following labial or dorsal obstruent, e.g. in /digan # paxa …/ → [diɣampaxa…] 'say (form. pl) straw', the alveolar /n/ assimilates completely to a labial nasal. This subset of the data is therefore fully consistent with standard phonological treatments of assimilation. According to these, the place specification for the nasal is replaced by a copy of the place specification of the following obstruent (Chomsky & Halle 1968), or in an autosegmental view the domain of the place specification of the obstruent extends via spreading to also encompass the nasal with concomitant delinking of the nasal's specification (McCarthy 1988). However, when the obstruent trigger of nasal place assimilation was the dental [t], Honorof's data showed that the result of the assimilation is not a dental [n]. Rather, [n] and [t] blended variably with a constriction location intermediate between an alveolar and a dental. The blending seen in the /n/ plus alveolar sequences is a notable result that speaks to the issue of underspecification. In particular, the radical underspecification theory of representations has promoted the idea that unmarked segments are not specified for certain features, which marked segments are (Archangeli 1988, Stemberger 1991), and since coronals are considered to be the prototypical unmarked segments, coronals should lack a specification for their place of articulation. Such unmarked segments receive fully specified representations by the action of default rules that fill in the missing values or by assimilation rules that spread the values from nearby segments to the underspecified targets (see Steriade 1995 for a review). The Castillian Spanish blending facts indicate that, if /n/ is considered to be a target of assimilation, then it cannot be said to be underspecified.

So called *long distance* assimilations such as vowel and consonant harmony have also been investigated using laboratory techniques. Since Gay (1977, 1978) it has been known that a non-contiguous sequence of identical vowels such as [u-u] in [kutup] is produced by speakers of English with a discontinuity both in the articulatory and the electromyographic measures of lip rounding (see also Boyce 1988, 1990). For example, in the electromyographic signal there is a trough coincident with the production of the intervening consonant. The cessation of muscle activity during the consonant is consistent with the hypothesis that the linguistic representation underlying the production of lip rounding schedules the rounding of the two identical vowels as two independent events, $[u]^{Round}C[u]^{Round}$ where C is a variable for any permissible intervocalic consonant or consonant cluster. A number of other studies have documented the same trough pattern in the production of non-contiguous, identical vowels in Spanish, French (Perkell 1986) and Swedish (McAllister 1978, Engstrand 1981). In contrast to these cases, Boyce (1988, 1990) found a plateau of continuous activity in Turkish for [uCu] utterances both in muscle excitation patterns (of the orbicularic oris) and in lower lip protrusion kinematics. This pattern of results, the English trough versus the Turkish plateau, seems to reflect the fact that Turkish but not English has vowel (rounding) harmony.

Cohn (1990, 1993) studies a case of nasal harmony in Sundanese, in which nasality spreads rightward from a nasal consonant until it encounters a supralaryngeal consonant, e.g. [ɲãũr] 'say', but [ŋãtur] 'arrange'. But the laryngeals /h ʔ/ can intervene in the domain of nasal spread as if they were skipped by the spreading, e.g. [mĩhãk] 'take sides' and [nũʔũs] 'dry'. Using oral/nasal airflow traces, Cohn presents evidence that these 'transparent' consonants are in fact nasalized. This result is consistent with the standard autosegmental treatment which sees harmony as an extension of the domain of the assimilating property. Gerfen (1999) studies nasal harmony processes in Coatzospan Mixtec using airflow recordings, and Walker (1999) is an acoustic study of nasal harmony in Guarani. For vowel harmony, using a combination of Electromagnetic Articulometry and Ultrasound methods, Benus and colleagues studied transparent vowels in Hungarian vowel harmony (Benus 2005, Benus & Gafos 2007, Benus, Gafos & Goldstein 2004, Gafos & Benus 2003). Gick *et al*. (2006) used Ultrasound to study the transparency of the low vowel [a] in Kinande tongue root harmony. Walker *et al*. (2008) studied transparency in the consonant harmony of Kinyarwanda using Electromagnetic Articulometry. From the perspective of the typological richness and specificity of harmony systems across languages much remains to be done, both in terms of charting the phonetic data in a more rigorous way than with transcriptions and in terms of integrating that data with phonological theory (for a review see Archangeli & Pulleyblank 2007).

We highlight a critical outstanding issue in relating experimental data on harmony to phonological theories. Focusing on an apparently simple case, we can ask what relation can be established between phonological theory and, for example, the continuous activation of lip rounding observed in Turkish [uCu] sequences? Two hypotheses suggest themselves: the continuum is an extended unitary rounding gesture, or the continuum is the aggregate by-product of overlap of separate shorter rounding gestures. According to the former view, in Turkish, rounding would extend over a domain encompassing both vowels in [uCu] and this is what gives rise to the plateau seen in Boyce's study. In the latter view, the plateau is the result of two separate rounding instructions, each with its own temporal domain, and it is the juxtaposition of these two rounding domains which results in a rounding plateau across the entire [uCu] sequence. The choice between the two hypotheses corresponds to a fundamental issue in phonological theory. This is the issue of assimilation and harmony as feature spreading (Goldsmith 1976, 1990, 1985, Clements 1976, 1977, 1985, Kiparsky 1981, Hayes 1986, Sagey1986) versus feature change (Chomsky & Halle 1968). Although the former view is widely assumed, it has never been subjected to systematic investigation across languages and across assimilating parameters. Deciding between these two views is not an easy matter. It is well-known that due to coarticulation the shape of the vocal tract at any time is an aggregate of multiple gestures associated with different segments. Aggregation has been observed for gestures that involve different constriction variables and for gestures that involve the same constriction variables. For different constriction variables, Hardcastle (1985) and Marchal (1988), using electropalatography, find that the gestures of two successive consonants, such as those in /kt/ (one with a tongue dorsum constiction goal, the other with a tongue tip constriction goal), show different degrees of overlap, and that the amount of overlap increases with speaking rate (Hardcastle 1985). Similar results are reported when overlapping gestures are specified for the same constriction variable. Munhall & Löfqvist (1992), for example, study the effects of speaking rate on two successive laryngeal abduction movements in *kiss Ted*, where the two units with laryngeal abduction gestures correspond to *ss* and *T*. The basic finding is that the distance between the two glottal peak openings decreases as rate increases. At slow rates, two opening movements occur and the glottis is closed between these two openings. At fast rates, a single movement is found with similar durations for the abduction and adduction phases (see also Boyce et al. 1990 for similar results that relate the occurrence of a one- or two-movement pattern for the velum with rate of speech). Munhall & Löfqvist (1992) find that the *shape* of the observed trajectories could be reasonably well modeled by adding two underlying gestures at different degrees of overlap. When the gestures do not overlap, summation produces two clear peaks in the shape of the simulated trajectory. As overlap increases, the simulated trajectory resembles in shape the blends or single movement patterns observed in the actual trajectories. However, there were inconsistencies between the amplitude of the simulated trajectories and that of the actual trajectories, especially at intermediate to large amounts of overlap. These inconsistencies derive from the simplifying assumption that the aggregation function can be estimated by simple algebraic summation or linear superposition.

An alternative is to hypothesize that the dynamical parameter values (target and dynamical stiffness, or time constant) of overlapping gestures of a constriction variable are averaged, rather than added (see Saltzman & Munhall 1989). In the case of a partially overlapping sequence of identical gestures, as might represent certain types of geminate consonants, this would mean that the same dynamical regime would be involved in single vs. geminate consonants, with the only difference being the span of time over which the regime is active. This representational difference could account in a simple way for Löfqvist's (2005) recent findings on the kinematic properties of geminates in Japanese and Swedish. In addition to geminates being of course longer, they are produced with greater articulatory displacements and result in tighter constrictions. Yet they lack the increase in peak velocity that is usually associated with an increase in displacement (cf. Beckman et al. 1992, Cho 2006). This combination of greater displacement without a corresponding increase in peak velocity could result from effective undershoot in the case of the single consonant. If we hypothesize that the relation between gestural time constant (or stiffness) and the activation duration of single consonants is such that single consonants do not have sufficient time to reach the target value, then they will exhibit undershoot. The longer activation time of the geminate would allow the target to be reached.

Other results suggest that in some cases the aggregation function must be more complex than either adding or averaging. Nolan et al. (1996) investigated the properties of s-ʃ overlap in

English examples like *this shop*. Using electropalatography and acoustics, they find that for modest degrees of overlap, the results are consistent with the predictions of parameter averaging. However, for extreme degrees of overlap, the palatographic and acoustic characteristics of the maximum constriction are not significantly different from those of [ʃ] by itself; i.e., there appears to be no influence of [s] at all on those characteristics. Yet, the overall constriction duration is longer than that for a single [ʃ], suggesting that the [s] gesture is still somehow contributing to the observed movements.

The key unresolved issue thus can be summarized by asking: how do the planning or execution systems combine multiple inputs for a given constriction variable? This is a critical question for the study of the relation between linguistic representation and articulatory organization and it is a question we can only ask if coordination of gestures is a fundamental part of our model.

4. Syllable organization

Laboratory phonology work over the last 20 years has developed both theoretical models and empirical methods that pursue the consequences of defining syllable structure as patterns or modes of temporal coordination among phonetic primitives. This approach is possible when the primitives are articulatory units that have observable, dynamic temporal properties among which abstract coordination relations can be defined (Browman & Goldstein 1988, 1995, Gafos 2002). Thus, the organization of compositional primitives into syllables, and the structural relations among units within a syllable (onset, rime, nucleus, coda) are implicit in the same representation (coordination topology) as required to adequately model the temporal regularities of speech. The consequences of this view have been investigated for a wide range of phenomena from syllable structure-sensitive allophony (see section 2 above) to universal preferences (markedness) of syllable structures (Nam et al. 2009). For a broader discussion of syllables examined with laboratory phonology approaches see **Coté, this volume**. For higher prosodic structure and rhythm, see **Turk, this volume**.

One specific theory of syllable structure developed in this framework is based on coupled oscillators (Goldstein *et al*., 2006, Nam *et al*., 2009). The theory attempts to account for why CV syllables are preferred to VC in several ways: they are more frequent cross-linguistically (and may be the only universally attested syllable type), they are acquired earlier than VC syllables, and they afford relatively freer internal combination (between onsets and nuclei) than do VC (between nuclei and coda). That theory attempts to relate these generalizations in a principled way to the fact that C and V gestures are triggered relatively synchronously in onsets (Löfqvist & Gracco 1999), but not in codas. In this theory, stable temporal coordination among articulatory units during speech production is achieved by associating each unit with a clock responsible for triggering that articulatory action, and by coupling the clocks for a given syllable's gestures to one another in a plan or *coupling graph* (a specific model of coordination topology). The coupling relations within the graph are hypothesized to leverage the intrinsically available modes of coupling oscillatory motions (Haken, Kelso & Bunz 1985, Turvey 1990), in-phase and anti-phase. Much work summarized in those papers shows that the in-phase mode is more *accessible* and more *stable* than the anti-phase mode. Thus if a syllable is to be composed of a consonant unit and a vowel unit, there are only two ways of coordinating them using these intrinsically available modes: in-phase, in which C and V are triggered synchronously is hypothesized for the onset-nucleus (CV) relation, and anti-phase (sequential triggering) is hypothesized for the nucleus-coda (VC) relation. Given the independently motivated properties of in-phase and anti-phase modes, the differences between CV and VC syllables can be explained.

A promising implication of the coordination topology model of syllable structure is that it opens the possibility of using temporal properties of articulatory events to infer syllabification. Whereas in English strings such as /kru/ 'crew' or /gli/ 'glee' are parsed into a single syllable with a complex two-consonant cluster as its onset, in Moroccan Arabic similar strings are claimed to be parsed into two syllables, e.g. /kra/ 'rent' → [k.ra], /skru/ → [sk.ru] 'they got drunk', /glih/ → [g.lih] 'he grilled' ('.' marks syllabic divisions; Dell & Elmedlaoui 2002). In terms of coordination topology, the consonants composing the onset in English should all share the same (onset) coordination in relation to the vowel, while they are coordinated sequentially with respect to one another (Browman & Goldstein 2000).

However, in Arabic, only the single (simplex) onset consonant bears the onset relation to the vowel. The different topologies should be associated with distinct temporal patternings of articulatory intervals. Pursuing this prediction, articulatory studies of syllable structure have examined the variability of structurally relevant intervals. Two distinct patterns of stability have emerged, each characteristic of a particular qualitative syllabic organization. In languages that admit complex onsets, the most stable interval across CVC, CCVC and CCCVC utterances (where C is any consonant and V is any vowel) is an interval defined by the center of the prevocalic consonantal string and the end of the hypothesized syllable (Browman & Goldstein 1988, Honorof & Browman 1995, Byrd 1995). The stability of this interval is predicted by models in Browman & Goldstein (2000) and Gafos (2002) as the result of optimization in systems of competing C-V and C-C constraints on coordination, and also by the coupled oscillator model (Saltzman et al. 2008) as the result of a loop in the coupling graph in which all onset Cs are coupled in-phase with the V and anti-phase with one another. In contrast, in languages that do not admit complex onsets such as Arabic, the most stable interval across CVC, CCVC and CCCVC utterances is defined by the immediately prevocalic consonant and the end of the hypothesized syllable (Shaw, Gafos, Hoole and Zeroual 2009). See ** Cote, this volume** for a summary from other languages.

Shaw et al. (2009) introduce computational and analytical methods in the study of the relation between syllable structure and experimental data. Given a hypothesized coordination topology, their models generate simulated temporal structure via a probabilistic version of a theory of temporal coordination constraints (Gafos 2002). The simulated data are then compared to the experimental speech movement data for their goodness of fit. Using this method, Shaw & Gafos (2010) show that for a CCV string in a language that does not admit complex onsets like Moroccan Arabic, the simplex onset topology provides a better fit to the experimental data from that language than the complex onset topology. The situation is reversed for English data. Shaw et al. (2009) also show that variability in the experimental data can influence the behavior of stability indices projected from an underlying qualitative syllabic organization. As variability across the lexical sample over which stability measures are assessed is increased moderately, the stability indices corresponding to the qualitative organization of a simplex onset parse remain in the quantitative region characteristic of simplex onsets. But as variability increases further, a tipping point can be seen beyond which the stability pattern turns to a state characteristic of complex onsets. The stability pattern can therefore change, thus exposing the range of validity of earlier heuristic methods (discussed above) that do not employ explicit stochastic modeling. Overall, instead of ignoring variability or treating it as a nuisance, Shaw et al. (2009) develop methods which harness variability as a tool for elucidating the relation between mental organization, in the form of syllable structure, and its complex behavioral instantiations.

5. Conclusion

We have reviewed how research in articulation has informed phonological inquiry across a wide range of domains. In each case, we have presented the key results obtained and noted areas of convergence or lack thereof with phonological theorizing.

A shared notion in the research reviewed is the construct of the speech gesture, a dynamic event with both spatial and temporal properties. This notion and the model in which it is embedded have sustained research in laboratory phonology by keeping in perspective theoretical developments in phonological theory and rigorous experimental data. Because the model is formally fleshed out it can be used to derive explicit predictions. These predictions have been pursued in various studies using a wide range of experimental methods. In turn, the studies pursuing these predictions have produced new data patterns which present opportunities for sharpening the theory, the model or the relation between the two. We have discussed examples of this interleaving of theory and experiment around the issues of, most notably, the nature of speech errors, the formal mechanism of assimilation (spreading versus feature change), the notion of transparency in harmony systems, the relation between phonological plan and surface produced output, and finally syllable structure.

References

Angermeyer, Philipp (2003). 'Copying contiguous gestures: an articulatory account of Bella Coola reduplication', in Elsi Kaiser and Sudha Arunachalam (eds.), *University of*

Pennsylvania Working Papers in Linguistics 9.1., Proceedings of the 26th Annual Penn Linguistics Colloquium. Philadelphia: Penn Linguistics Club, 17–30.

Archangeli, Diana (1988). Aspects of underspecification theory. *Phonology* 5, 183-207

Archangeli, Diana & Douglas Pulleyblank (1994). *Grounded Phonology*. Cambridge: MIT Press.

Archangeli, Diana & Douglas Pulleyblank (2007). 'Harmony', in Paul de Lacy (ed.), *The Cambridge Handbook of Phonology*. Cambridge: Cambridge University Press, 353-378.

Baltazani, Mary (2006). 'Focusing, prosodic phrasing and hiatus resolution in Greek', in L. Goldstein, D. Whalen & C. Best (Eds.), *Varieties of phonological competence (Laboratory Phonology 8*. Berlin, New York: Mouton de Gruyter, 473-494.

Barry, M. (1991). Temporal modeling of gestures in articulatory assimilation, in *Proceedings of the XIIth International Congress of Phonetic Sciences*. Aix-en-Provence: University of Provence, 14-17.

Beckman, M. E., Edwards, J., Fletcher, J. (1992). 'Prosodic structure and tempo in a sonority model of articulatory dynamics', in G. J. Docherty, D. R. Ladd (eds.), *Papers in laboratory phonology II: Gesture, segment, prosody*. Cambridge: Cambridge University Press, 68–86

Benus, S. (2005). *Dynamics and transparency in vowel harmony.* Unpublished Ph.D. Dissertation, New York University.

Benus, S., Gafos, A., & Goldstein, L. (2004). 'Phonetics and phonology of transparent vowels in Hungarian', in P. M. Nowak, C. Yoquelet & D. Mortensen (eds.), *Proceedings of the 29th Annual meeting of the Berkeley Linguistic Society*. Berkeley Linguistic Society, 485-497.

Benus, S., Smorodinsky, I. & Gafos, A. (2004). Gestural coordination and the distribution of English "geminates", in Sudha Arunachalam and Tatjana Scheffler (eds.), *University of Pennsylvania Working Papers in Linguistics 10.1., Proceedings of the 27th Annual Penn Linguistic Colloquium*. Philadelphia: Penn Linguistics Club, 33–46.

Benus, S., & Gafos, A. (2005). 'Qualitative and quantitative aspects of vowel harmony: A dynamics model', in B. G. Bara, L. Barsalou & M. Bucciarelli (eds.), *CogSci2005, XXVII Annual Conference of the Cognitive Science Society (Stresa, Italy, 2005)*. New Jersey: Lawrence Erlbaum, 226-231.

Bernstein, N. (1967). *Coordination and regulation of movement.* New York: Pergamon Press.

Boyce, S. E. (1988). *The influence of phonological structure on articulatory organization in Turkish and in English: howel harmony and coarticulation*. PhD dissertation, Yale University, New Haven, Conn.

Boyce, S. E. (1990). Coarticulatory organization for lip rounding in Turkish and English. *Journal of the Acoustical Society of America* 88, 2584-2595.

Boyce, S. E., R. A. Krakow, F. Bell-Berti, and C. Gelfer. (1990). Converging sources of evidence for dissecting articulatory movements into core gestures. *Journal of Phonetics* 18, 173-188.

Bradley, Travis G. (2002). 'Gestural timing and derived environment effects in Norwegian clusters', in Line Mikkelsen and Christopher Potts (eds.), *WCCFL 21 Proceedings*. Somerville, MA: Cascadilla Press, 43-56.

Browman, C. P. and L. Goldstein (1986). Towards an articulatory phonology. *Phonology* 3, 219-252.

Browman, C. P. and L. Goldstein (1988). Some notes on syllable structure in Articulatory Phonology. *Phonetica* 45, 140-155.

Browman, C. P. and L. Goldstein (1989). Articulatory gestures as phonological units. *Phonology* 6, 201-251.

Browman, C. P. and L. Goldstein (1990). 'Tiers in articulatory phonology, with some implications for casual speech', in J. Kingston and M. Beckman (eds.) *Papers in Laboratory Phonology I: Between the Grammar and the Physics of Speech*. Cambridge: Cambridge University Press, 341-397

Browman, C. P. and L. Goldstein (1995). 'Gestural syllable position effects in American English', in F. Bell-Berti & L. J. Raphael (eds.), *Producing speech: contemporary issues*. Woodbury, N.Y.: American Institute of Physics Press, 19-33.

Browman, C. P. and L. Goldstein (2000). Competing constraints on intergestural coordination and self-organization of phonological structures. *Les Cahiers de l'ICP, Bulletin de la Communication Parlée* 5, 25–34.

Byrd, D., S. Tobin, E. Bresch, S. Narayanan (2009). Timing effects of syllable structure and stress on nasals: a real-time MRI examination. *Journal of Phonetics* 37, 97-110.

Byrd, D. & E. Saltzman (2003). 'The elastic phrase: modeling the dynamics of boundary-adjacent lengthening', *Journal of Phonetics* 31(2), 149–180.

Cho, T. (2006). 'Manifestation of prosodic structure in articulation: evidence from lip kinematics in English', in L. M. Goldstein, D. H. Whalen, & C.T. Best (wds.), *Laboratory Phonology*. Berlin: Mouton de Gruyter, 519-548.

Chomsky, N., and M. Halle (1968). *The Sound Pattern of English*. New York: Harper and Row.

Clements, G. N. (1976). *Vowel harmony in nonlinear generative phonology: an autosegmental model*. [Published in 1980 by Indiana University Linguistics Club, Bloomington, IN.]

Clements, G. N. (1977). 'Neutral vowels in Hungarian vowel harmony: an autosegmental interpretation'. In *Proceedings of the North Eastern Linguistic Society* 7, 49-64.

Clements, G. N. (1981). 'Akan vowel harmony: a nonlinear analysis', in G. N. Clements (ed.), *Harvard Studies in Phonology* 2. Department of Linguistics, Harvard University, 108-177.

Clements, G. N. (1985). The geometry of phonological features. *Phonology Yearbook* 2, 225-252.

Cohn, A. (1990). *Phonetic and phonological rules of nasalization*. UCLA PhD dissertation. Distributed as *UCLA Working Papers in Phonetics*, 76.

Cohn, A. (1993). 'The status of nasalized continuants', in M. Huffman and R. Krakow (eds.), *Nasals, Nasalization, and the Velum*. San Diego: Academic Press, 329-367.

Davidson, L. (2003). *The atoms of phonological representation: Gestures, coordination, and perceptual features in consonant cluster phonotactics*. Ph.D. dissertation, Department of Cognitive Science, Johns Hopkins University.

Davidson, L. (2006). Schwa elision in fast speech: segmental deletion or gestural overlap? *Phonetica* 63, 79–112.

Dell, F. & M. Elmedlaoui (2002). *Syllables in Tashlhiyt Berber and in Moroccan Arabic*. Dordrecht, Boston: Kluwer Academic Publishers.

Engstrand, O. (1981). Acoustic constraints of invariant input representation? An axperimental study of selected articulatory movements and targets. *Reports of the Uppsala University Department of Linguistics* 7, Department of Linguistics, Uppsala, Sweden, 67-94.

Farnetari, E., and M. G. Busà. 1994. Italian clusters in continuous speech. In *Proceedings of the International Conference on Spoken Language Processing*, Yokohama, Volume 1, 359-362.

Fowler, C. A. (1980). Coarticulation and theories of extrinsic timing. *Journal of Phonetics, 8,* 113-133.

Fowler, C. Rubin, P., Remez, R. and Turvey, M. (1980). 'Implications for speech production of a general theory of action', in B. Butterworth (ed.), *Language Production, Volume 1: Speech and Talk*. London: Academic Press, 373–420

Frisch, S., & Wright, R. (2002). The phonetics of phonological speech errors: An acoustic analysis of slips of the tongue. *Journal of Phonetics, 30,* 139-162.

Fromkin, V. A. (1971). The non-anomalous nature of anomalous utterances. *Language 47,* 27-52.

Fujimura, O. (1981). Temporal organization of articulatory movements as a multidimensional phrasal structure. *Phonetica 38,* 66-83.

Gafos, A. (2002). A grammar of gestural coordination. *Natural Language and Linguistic Theory* 20, 269-337.

Gafos, A. (2006). 'Dynamics in grammar: comments on Ladd an Ernestus & Baayen', in L. Goldstein, D. Whalen & C. Best (eds.), *Varieties of phonological competence (Laboratory Phonology 8*. Berlin, New York: Mouton de Gruyter, 51-79.

Gafos, A. and Benus, S. (2003). 'On neutral vowels in Hungarian', in M.-J. Solé, D. Recasens & J. Romero (eds.), Proceedings *of the 15th International Congress of Phonetic Sciences*, Universitat Autònoma de Barcelona, 77-80.

Gafos, A., & Benus, S. (2006). Dynamics of phonological cognition. *Cognitive Science 30*, 905-943.
Gafos, A. and C. Kirov (2010). 'A dynamical model of change in phonological representations: the case of lenition', in François Pellegrino, Egidio Marsico, Ioana Chitoran & Christophe Coupé (eds.), *Approaches to Phonological Complexity, Phonology & Phonetics Series*. Berlin: Mouton de Gruyter, 225-246.
Gay, T. (1977). Articulatory movements in VCV sequences. *Journal of the Acoustical Society of America* 62, 183-191.
Gay, T. (1978). 'Articulatory units: segments or syllables?', in ed. A. Bell and J. B. Hooper (eds.), *Syllables and segments*. Amsterdam: North-Holland Publishing Company, 121-131.
Gerfen, Chip (1999). *Phonology and phonetics in Coatzospan Mixtec*. Dordrecht: Kluwer.
Gick, B., Pulleyblank, D., Campbell, F., & Mutaka, N. (2006). Low vowels and transparency in Kinande vowel harmony. *Phonology 23,* 1-20.
Goldrick, M., & Blumstein, S. (2006). Cascading activation from phonological planning to articulatory processes: Evidence from tongue twisters. *Language and Cognitive Processes 21*, 649-683.
Goldsmith, J. (1976). *Autosegmental phonology*. PhD dissertation, MIT. [Published by New York: Garland Press, 1980.]
Goldsmith, J. (1985). Vowel harmony in Khalka Mongolian, Yaka, Finnish and Hungarian. *Phonology Yearbook* 2, 253-275.
Goldsmith, J. (1990). *Autosegmental and metrical phonology*. Oxford: Blackwell.
Goldstein, L., Byrd, D., and Saltzman, E. (2006). 'The role of vocal tract gestural action units in understanding the evolution of phonology', in M. Arbib (ed.) *From Action to Language: The Mirror Neuron System.* Cambridge: Cambridge University Press. 215-249.
Goldstein, L., Pouplier, M., Chen, L., Saltzman, E., & Byrd, D. (2007). Dynamic action units slip in speech production errors. *Cognition 103*, 386-412.
Guenther, Frank H. (1995). Speech Sound Acquisition, coarticulation, and rate effects in a neural network model of speech production. *Psychological Review* 102, 594–621.
Haken, H., Kelso, J.A.S. & Bunz, H. (1985). A theoretical model of phase transitions in human hand movements. *Biological Cybernetics 51*, 347-356.
Hall, N. E. (2003). *Gestures and segments: vowel intrusion as overlap*. Ph.D. dissertation, University of Massachusetts, Amherst.
Hardcastle, W. J. (1985). Some phonetic and syntactic constraints on lingual coarticulation during /kl/ sequences. *Speech Communication* 4, 247-263.
Harris, C. M. (1953). A study of the building blocks in speech. *Journal of the Acoustical Society of America 25*, 962-969.
Hayes, B. (1986). Assimilation as spreading in Toba Batak. *Linguistic Inquiry* 17, 467-499.

Hockett, C. A. (1955). *A Manual of phonology.* Indiana University Publications in Anthropology and Linguistics 11.

Holst, T., and F. Nolan. (1995). 'The influence of syntactic structure on [s] to [ʃ] assimilation', in B. Connell and A. Arvaniti (eds.), *Phonology and Phonetic Evidence: Papers in Laboratory Phonology IV.* Cambridge: Cambridge University Press, 315-333.

Honorof, Douglas (1999). *Articulatory gestures and Spanish nasal assimilation*. Doctoral dissertation, Yale.

Huffman, M. (1990). *Implementation of nasal: timing and articulatory landmarks.* UCLA PhD dissertation. Distributed as *UCLA Working Papers in Phonetics* 75.

Hulst, H. v. d. and J. v. d. Weijer. (1995). 'Vowel harmony', in J. Goldsmith (ed.), Handbook of Phonological Theory. Cambridge, Mass.: Blackwell Publishers, 495-534.

Keating, Patricia. (1988). Underspecification in phonetics. *Phonology* 5, 275-292.

Keating, Patricia (1990). 'The window model of coarticulation: articulatory evidence', in John Kingston and Mary Beckman (eds.), *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*, 451-470.

Kelso, J.A.S., Saltzman, E. L., & Tuller, B. (1986). The dynamical perspective on speech production: Data and theory. *Journal of Phonetics 14,* 29-59.

Kiparsky, P. (1981). *Vowel harmony*. Ms. Stanford University, Stanford, Calif.

Kiparsky, Paul. (1982). 'Lexical morphology and phonology', in Yang, I.-S. (ed.), *Linguistics in the Morning Calm*. Seoul: Hanshin, 3–91.

Kohler, K. J. (1990). 'Segmental reduction in connected speech in German: phonological facts and phonetic explanation', in W. J. Hardcastle and A. Marchal (eds.), *Production and Speech Modeling*. Dordrecht: Kluwer, 69-92.

Kröger, B. J. (1993). A gestural production model and its implications for reduction in German. *Phonetica* 50, 213-233.

Krakow, R. A. (1989). *The articulatory organization of syllables: a kinematic analysis of labial and velic gestures*. PhD dissertation, Yale University.

Krakow, R. A. (1999). Physiological organization of syllables: a review. *Journal of Phonetics* 27, 23-54.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences 22*, 1-75.

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

Lindblom, B. (1963). Spectrographic study of vowel reduction. J*ournal of the Acoustical Society of America 35* 1773-1781.

Lofqvist, A. (2005). Lip kinematics in long and short stop and fricative consonants. *Journal of the Acoustical Society of America 117*, 858-878.

Löfqvist, A. & Gracco, V. (1999). Interarticulator programming in VCV sequences: lip and tongue movements. *Journal of the Acoustical Society 105*, 1864-1876.

Marchal, A. (1988). Coproduction: evidence from EPG data. *Speech Communication* 7, 287-295.

McAllister, R. (1978). Temporal asymmetry in labial coarticulation. *Working Papers, Institute of Linguistics, Stockholm University* 35, 1-29.

McCarthy, J. J. (1988). Feature geometry and dependency: a review. *Phonetica* 45, 84-108.

Mowrey, R. A., & MacKay, I. R. (1990). Phonological primitives: Electromyographic speech error evidence. *Journal of the Acoustical Society of America 88,* 1299-1312.

Munhall, K., and A. Löfqvist. (1992). Gestural aggregation in speech: laryngeal gestures. *Journal of Phonetics* 20, 111-126.

Nam, H., Goldstein, L., & Saltzman, E. (2009). 'Self-organization of syllable structure: a coupled oscillator model', in F. Pellegrino, E. Marisco, & I. Chitoran (eds.), *Approaches to phonological complexity.* Berlin/New York: Mouton de Gruyter, 299-328.

Nolan, F., T. Holst, and B. Kühnert. (1996). Modelling [s] to [ʃ] assimilation in English. *Journal of Phonetics* 24, 113-137.

Perrier P., Ostry, D. J., Laboissière, R. (1996). The equilibrium point hypothesis and its application to speech motor control. *J Speech Hear Res., 39* 365-78.

Perkell, J. S. (1986). Coarticulation strategies: preliminary implications of a detailed analysis of lower lip protrusion movements. *Speech Communication* 5, 47-68.

Pouplier, M. (2003). *Units of phonological encoding: Empirical evidence*. Unpublished PhD dissertation, Yale University.

Pouplier, M. (2007). Tongue kinematics during utterances elicited with the SLIP technique. *Language and Speech 50*, 311-341.

Pouplier, M. (2008). The role of a coda consonant as error trigger in repetition tasks. *Journal of Phonetics 36*, 114-140.

Pouplier, M., & Goldstein, L. (2005). Asymmetries in the perception of speech production errors. *Journal of Phonetics 33*, 47-75.

Prince, A., & Smolensky, P. (1993/2004). *Optimality theory: Constraint interaction in generative grammar.* Technical Report CU-CS-696-93, Department of Computer Science, University of Colorado at Boulder, and Technical Report TR-2, Rutgers Center for Cognitive Science, Rutgers University, New Brunswick, NJ, April 1993. [Published: Blackwell, 2004].

Sagey, E. (1986). *The representation of features and relations in non-linear phonology.* [Published by New York: Garland Press, 1991.]

Saltzman E. (1986). Task dynamic coordination of the speech articulators: a preliminary model. *Experimental Brain Research Series* 15, 129-144.

Saltzman, E. (1995). 'Dynamics and coordinate systems in skilled sensorimotor activity', in R. F. Port and T. van Gelder (eds.), *Mind as Motion.* Cambridge: MIT Press, 149-173.

Saltzman, E., & Munhall, K. G. (1989). A dynamical approach to gestural patterning in speech production. *Ecological Psychology* 1, 333-382.
Saltzman, E., Nam, H., Krivokapic, J., & Goldstein, L. (2008). A task-dynamic toolkit for modeling the effects of prosodic structure on articulation. In P. A. Barbosa & S. Madureira (Eds.), *Proceedings of the Speech Prosody 2008 Conference,* Campinas, Brazil.
Shattuck-Hufnagel, S. (1983). 'Sublexical units and suprasegmental structure in speech production planning', in P. F. MacNeilage (ed.), *The production of speech*. New York: Springer, 109–136.
Shattuck-Hufnagel, S., & Klatt, D. (1979). The limited use of distinctive features and markedness in speech production: evidence from speech error data. J*ournal of Verbal Learning and Verbal Behavior 18*, 41–55.
Shaw, J., Gafos, A., Hoole, P. & C. Zeroual (2009). Syllabification in Moroccan Arabic: evidence from patterns of temporal stability in articulation. *Phonology* 26, 187-215.
Shaw, J. and Gafos, A. (2010). 'Quantitative evaluation of competing syllable parses', in Jeffrey Heinz, Lynne Cahill and Richard Wicentowski (eds.), *ACL SIGMORPHON, Proceedings of the 11th Meeting of the ACL special interest group in computational phonology and morphology*. Uppsala, July 11-16, Sweden, page nos. tba.
Smith, C. L. (1995). 'Prosodic patterns in the coordination of vowel and consonant gestures', in B. Connell & A. Arvaniti (eds.), *Papers in Laboratory Phonology IV: Phonology and Phonetic Evidence.* Cambridge: Cambridge University Press, 205-222.
Sproat, R. & O. Fujimura. (1993). Allophonic variation in English /l/ and its implications for phonetic implementation. *Journal of Phonetics* 21, 291-311.
Stemberger, J. (1991). Radical underspecification in language production. *Phonology* 8, 73-112.
Steriade, Donca. (1995). 'Underspecification and markedness', in John Goldsmith (ed.), *Handbook of phonological theory*. Cambridge, Mass.: Blackwell, 114–174.
Steriade, Donca. (1999). 'Phonetics in phonology: the case of laryngeal neutralization', Matthew Gordon (ed.), *Papers in phonology 3 (UCLA working papers in linguistics 2).* Los Angeles: Department of Linguistics, University of California, 25–145.
Stevens, K. N. (1972). 'The quantal nature of speech: evidence from articulatory-acoustic data', in D. Denes (ed.), *Human communication: A unified view*. New-York: McGraw-Hill, 51-66.
Stevens, K. N. (1989). On the quantal nature of speech. *Journal of Phonetics 17,* 3-45.
Turvey, M. T. (1977). 'Preliminaries to a theory of action with reference to vision', in R. Shaw & J. Bransford (eds.), *Perceiving, acting and knowing: toward an ecological psychology.* Hillsdale, NJ: Lawrence Erlbaum Associates, 211-265
Turvey, M. T. (1990). Coordination. *American Psychologist 45*, 938-953.
Walker, R. (1999). Guaraní voiceless stops in oral versus nasal contexts: an acoustical study. *Journal of the International Phonetic Association* 29, 63-94.

Walker, R., Byrd, D., Mpiranya, F. (2008). An articulatory view of Kinyarwanda coronal harmony. *Phonology* 25, 499-535.
Wood, A. J. S. (1996). Assimilation or coarticulation: evidence from the temporal coordination of tongue gestures for the palatalization of Bulgarian alveolar stops. *Journal of Phonetics* 24, 139-164.
Zsiga, Elizabeth C. (1995). 'An acoustic and electropalatographic study of lexical and postlexical palatalization in American English', in B. Connell and A. Arvaniti (eds.), *Phonology and Phonetic Evidence: Papers in Laboratory Phonology IV*. Cambridge: Cambridge University Press, 282-302.
Zsiga, Elizabeth C. (1997). Features, gestures, and Igbo vowels: an approach to the phonology/phonetics interface. *Language* 73, 227-274.

[1] Acknowledgments: Research supported by NIH NIDCD DC008780 and NSF 0922437 grants to Haskins Laboratories and ERC 249440.