7 Interarticulatory Coordination
Speech Sounds

PHILIP HOOLE AND
MARIANNE POUPLIER

7.1 Introduction

There is probably no speech utterance that involves only a single organ of the vocal tract. In a sense, full coverage of the topic of interarticulatory coordination therefore requires coverage of speech production in its entirety, and it is important at the beginning of this chapter to outline how we intend to define a task of manageable proportions. Thinking in terms of the three processes defined by Catford (2001), initiation, phonation, and articulation, we will focus on articulation, defined as the process of modifying an airstream. So we will not go into areas such as the coordination of respiratory and phonatory activity, nor the patterns of coordinated movement that are required for non-pulmonic initiation. We will, however, be devoting a section to laryngeal-oral coordination in which we treat the alternation between voicing and voicelessness as articulatory behavior of the larynx.

A basic distinction in articulation is that between temporal and spatial aspects. We will be looking at interarticulatory coordination above all in its temporal aspects. Spatial aspects, namely how multiple articulators collaborate to achieve some functional goal, are dealt with by Perrier and Fuchs in Chapter 11 of this volume under the title of motor equivalence, a key theoretical issue in speech motor control.

Even from the temporal point of view, we will not attempt to cover one big area, namely timing issues with respect to coarticulation. Coarticulation necessarily involves issues of interarticulatory coordination. For example, when investigating the extent of anticipatory labial coarticulation, the activity of the lips in the articulation of some sound must be referenced in time to the activity of some other speech organ, or perhaps to some acoustic goal. Since good reviews of coarticulation are readily available, and full coverage of this topic would involve discussion of a multitude of acoustic studies, we refer the reader to the relevant sources (e.g., Farnetani and Recasens 2010).

The Handbook of Speech Production, First Edition. Edited by Melissa A. Redford.

Turning now to the structure of the chapter itself, the most fundamental division we will be making is between the coordination of multiple articulators in the articulation of single segments (section 7.2) versus coordination of multiple articulators in the articulation of multiple segments (section 7.3). In section 7.2, attention will be paid to influences on timing patterns when, for example, the tongue and jaw or the lips and jaw are coordinated to form a constriction. Attention will also be paid to a particularly clear case in which interarticulatory timing has linguistic relevance, namely laryngeal-oral timing in various forms of the voicing distinction. We thus also need to distinguish between cases where multiple articulators contribute to a single gesture in a single segment and those where the different articulators also involve different gestures (gesture is used here in the sense of a constriction in the vocal tract). We will also review the influence of syllable position on multiple-gesture segments.

In section 7.3, which focuses on the case of the coordination of multiple articulators across multiple segments, we examine the timing of consonant sequences, such as the movements of tongue-dorsum and tongue-tip for /kl/, and also the timing of consonant-vowel sequences. Discussion will again focus on the role of syllable position in determining timing patterns. We will also consider the influence of the segmental properties of the elements in a sound sequence, such as voicing or place order, as well as language-specific preferences for specific coordination patterns.

It is worth mentioning at this point that the traditional term "segment" we have used here is often regarded by proponents of gestural theories (e.g., Task Dynamics, Articulatory Phonology) as the epiphenomenal expression of underlying coordination patterns. Regarding gestures as the basic units ("atoms") of articulatory representation, "segment" is simply a convenient term for multigestural ("molecular") structures with a particularly high degree of cohesiveness. Evidence for this view comes from, for example, Saltzman et al. (1998) and Saltzman, Löfqvist, and Mitra (2000), who analyzed phase-resetting in a perturbation paradigm and found stronger intergestural "glue" within than between segments. Along these lines, Pouplier (2011) has recently noted that a consonant cluster is not fundamentally different from a single consonantal segment given that closure and release can be modeled by separate but coupled dynamic building blocks, as in the split-gesture dynamics approach of Nam, Goldstein, and Saltzman (2009).

7.2 Coordination of multiple articulators for single segments

7.2.1 Multiple articulators for a single gesture

A considerable body of work has looked at the coordination of multiple articulators to form a specific vocal tract constriction, such as lip-closure, from the point of view of the cohesiveness of the component movements. A high level of cohesiveness in coordination patterns provides evidence for the existence of higher-level functional groupings of the articulators. These functional grouping are often referred to as coordinative structures following Turvey (1977, 1990) and Bernstein (1967). Such groupings are considered necessary to understand how speakers manage the potentially very large number of degrees of freedom in the vocal apparatus (Kelso, Saltzman, and Tuller 1986). Moreover, coordinative structures form the substrate for the spatial tradeoffs characteristic of motor equivalence (Perrier and Fuchs, this volume, Chapter 11).

For the particularly well-researched lip-jaw complex considerable evidence has been found for consistent interarticulatory timing patterns. For lip closing movements the typical pattern has been for peak velocities to occur in the temporal order (1) upper lip, (2) lower lip, (3) jaw (e.g., Gracco 1988). The tightness of the coordination is generally more precise for oral closure than for oral opening (Gracco 1988; Kollia, Gracco, and Harris 1995). This variation in tightness of coordination provides one example of the observation that movements into and out of a constriction are quite asymmetric, which in turn has led to the proposal that the closure and opening phase of a gesture are best modeled as separately controlled movements (Browman 1994; Nam 2007; Harrington, Fletcher, and Roberts 1995).

Interarticulatory coordination can also be influenced by the precise identity of the target sounds even when the articulators involved are basically the same. For example, Mooshammer, Hoole, and Geumann (2006) looked in particular at the influence of manner of articulation on tongue-jaw coordination. The results showed a substantial degree of consistency in that the jaw almost always reached its target position for the consonant later than the tongue tip. The more interesting feature of the results came from some subtle but consistent timing differences between the consonants: the sibilants showed a symmetrical tongue-jaw coordination pattern, meaning late location of jaw target onset and early location of jaw target offset within the encompassing tongue target phase. In contrast, the voiceless plosive /t/ showed an asymmetrical pattern with very late attainment of the jaw target, but also a late location of jaw target offset close to tongue target offset. These differences in coordination can be readily understood in terms of the different aerodynamic and acoustic requirements of the sounds involved. The sibilant fricatives require a stable, high position of the jaw centered on the oral constriction phase in order to ensure generation of turbulence. The critical location for voiceless plosives is at the release of the constriction where a high jaw position can help to support a salient burst. Interestingly, for other consonants (e.g., voiced plosive, nasal, lateral), where none of these constraints so obviously apply, Mooshammer et al. found more variability (i.e., not such a clear preference) for either the symmetrical or asymmetrical timing pattern. These results indicate that, in spite of stable timing patterns for many multi-articulator constriction gestures, these patterns are not totally stereotypic. Instead, they can be modulated to fulfill differing aerodynamic and acoustic demands.

Given that speech motor control obviously exhibits a whole range of very precise timing relationships, a question arises as to the most appropriate framework in which to express these relationships.

One answer is simply based on what kinematic measures exhibit the most stable patterns. For example, various studies have suggested that timing relationships expressed in terms of the time-points of peak velocities are more stable than position-based ones, such as attainment of target position (see Kollia et al. 1995, for discussion).

A more controversial answer is that temporal invariance in coordination patterns can be captured using a phase-plane representation. For example, when cyclical movements of the jaw are plotted with velocity as a function of position then a roughly circular pattern in this two-dimensional representation is traced out. Characteristic time-points in the movement of another articulator, such as the lip, can then be expressed as an angle in this plane (see, e.g., Kelso et al. 1986; Nittrouer et al. 1988, for more details on procedures). The original hypothesis, based on other work in motor control and the characterization of dynamic systems (Kelso et al. 1986) was that stability in interarticulatory timing might express itself as a phase-angle that remains constant over manipulations such as stress and rate. However, the balance of later work indicates that systematic changes in phase-angle probably occur (e.g., Nittrouer et al. 1988 and Shaiman, Adams, and Kimelman 1995, for lip-jaw; Nittrouer 1991, for tongue-jaw).

7.2.2 *Multiple articulators for multiple gestures*

In this section we first compare the degree of interarticulator cohesion within versus between gestures, then move on to look at effects of syllable position on multiple-gesture segments, and then at within and cross-language patterns of laryngeal-oral coordination.

7.2.2.1 Inter- vs. intragesture cohesion With regard to the degree of interarticulator cohesion within versus between gestures, we can return to Kollia et al. (1995). This paper takes as its point of departure the many findings on interarticulator coordination for the lip-jaw complex and integrates velar movement into the analyses. The question then asked is: How cohesive is the interarticulator timing for lip-velum or jaw-velum (for example during lip closing and velum closing for /b/) in comparison to the intragestural lip-jaw case discussed in the previous section? Kollia et al. found that for the two closing movements associated with /b/ in the sequence /mabnab/, intergestural (velum-jaw, velum-upper lip) cohesion was clearly weaker than intragestural cohesion when estimated from time-points of peak velocities. However, the interpretation of this result is not entirely straightforward since velar raising is probably influenced not only by the /b/ but also, if only weakly, by the preceding /a/. Thus, intergestural cohesion may have an inherently more complex and variable velocity profile than intragestural cohesion. Support for this interpretation comes from the fact that, when the analysis was based on peak position rather than peak velocity timing, the intragestural advantage was much less evident, albeit the correlations were overall weaker when based on peak positions (peak velar position in this sound sequence is more easily attributed to a single segment than is peak velocity). By contrast, for the oral and velar opening movements in the /bna/ part of the sequence, the peak velocity that referenced velum-jaw coordination suggested tighter cohesion than for jaw-upper lip coordination. A transition such as that from /b/ to /n/ is probably a good example of where tight intergestural coordination is required and possible. Kollia et al. also suggest that compared to peak positions "peak velocity relations may be more illustrative of critical or constrained timing relations" (1323).

A study by Gracco and Löfqvist (1994) followed a somewhat similar approach to that of Kollia et al. (1995), but focused on the intergestural coordination of oral and laryngeal movements. Specifically, they looked at the temporal cohesion of jaw lowering and glottal closing in the movement from /s/ to the vowel in words like "supper," and of lower lip raising and glottal opening from the vowel to /p/. At a gross level of analysis, timing patterns showed evidence of stable organization in both cases. The time-point of peak glottal closing velocity always preceded peak jaw lowering velocity, while the time-point of peak lip raising velocity always preceded peak glottal opening velocity. When detailed correlation analyses were used to assess temporal cohesion between gestures, it emerged that correlations between oral and glottal time-points were noticeably lower for the oral opening movements. At first glance, this suggests a similar intergestural pattern of coordination to the intragestural pattern of coordination already noted above (less cohesion for oral opening movements, section 7.2.1). However, the authors emphasize that the intergestural cohesion level they found was remarkably high compared to results reported in other studies for intragestural lip-jaw coordination in opening movements (intergestural cohesion is typically lower than intragestural cohesion, Saltzman et al. 1998, 2000). Assessment of the conclusion is not entirely straightforward, though, because Gracco and Löfqvist did not explicitly measure lip-jaw cohesion in their study. Nevertheless, the results may indicate that intergestural cohesion can be as high as that involved in forming a single constriction.

7.2.2.2 Influence of syllable position on multi-gesture segments We will first consider the effects of syllable position on multi-gesture segments with respect to the coordination of the oral and velar gestures in nasal consonants. These particular multi-gesture segments have been the object of some clear and influential results with respect to syllable position. They thus provide a good starting point for discussing generalizations in interarticulatory coordination for a much wider range of multi-gesture segments. Liquids have also been a particular focus in studies that investigate the effect of syllable position on interarticulatory coordination. In fact, liquids might be construed as the tip of an iceberg consisting of everything that would fall under the traditional phonetic terms of secondary and double articulation.

In an influential paper, Krakow (1999) reports results from an investigation into the coordination of lips and velum in the nasal /m/ in American English. Critically, she also embedded the empirical work in a wide-ranging review of articulatory evidence for syllabic structure. The basic timing result for lip-velum coordination was that in syllable-initial position the velum and the lower lip moved more or less in synchrony. That is, the end of the velar lowering movement for /m/ roughly coincided with the beginning of the lip-raising movement for the oral closure. In contrast, in syllable-final position, the velar movement was timed much earlier. Specifically, the end of the velar lowering movement roughly corresponded to the *beginning* of the lip closing movement. This basic timing pattern was recently confirmed by Byrd et al. (2009), who also gave an interesting overview of how the coupling relations among the articulators could be modeled.

Given Krakow (1999) and Byrd et al.'s (2009) results, it is quite clear that syllable position can have striking effects on articulatory coordination. The much more difficult issue is the extent to which such findings are generalizable from lip-velum coordination to other patterns of interarticulatory coordination. Both Byrd et al. and Krakow make the link to other multi-gesture segments. Byrd et al. formulated the link as follows:

> Strikingly, there appear general parallels (see also Browman and Goldstein, 1992; Krakow, 1999) in the effect of syllable structure on nasals, glides, and liquids – namely, that in onset position the primary oral constriction gesture target precedes or is synchronous with the secondary (non-primary constriction) gesture target (i.e., the pharyngeal gesture for [r], dorsal for [l], or velum for nasals), whereas in coda, the secondary gesture target occurs far earlier, during the preceding vocalic nucleus.
>
> (2009: 98)

However, one could ask whether these "parallels" are not just a fortuitous coincidence of specific properties of certain sounds in a particular dialect of English. For example, the fact that English has no contrastive vowel nasalization may allow early velar lowering before the post-vocalic nasal, which then just happens to meet a preferred timing pattern for dark /l/. Certainly, it is not obvious to us that it is attractive to regard the velar gesture for /m/ as "secondary" in the same sense as velarization for dark /l/ (or as "wider" as suggested in Browman and Goldstein 1995: 25). In particular, there do indeed seem to be well-documented differences in the timing of nasals across languages. French (Cohn 1993) and the Australian aboriginal language Arrernte (Butcher 1999) both show synchronous coordination of the oral and ~~velic~~ constrictions in both onset and coda. In the case of French, this could be explained by the need to maintain a difference between oral and nasal vowels, as these constitute phonologically contrastive units. Arrernte, on the other hand, has no oral-nasal vowel contrast. Cross-linguistic variety in the coordination patterns for the multiple lingual gestures has also been reported for liquids by Gick et al. (2006), as discussed in more detail below. These cross-language differences underscore a fact which will come up again in section 7.3; namely, that of systematic language-specific differences in articulatory organization.

Although the range of possible patterns for liquids and other sound categories will be discussed in detail immediately below, it is first worth emphasizing that Krakow (1999) did not confine herself to the basic syllable-initial/final contrast in discussing the empirical findings. Instead, she highlighted a whole range of timing issues that the basic finding led to. For example, in considering why Cantonese (Wang 1995) did not show the same pattern as English, Krakow considered whether fast/casual speech could lead to re-syllabification of final consonants:

> In casual (or fast) speech, the pattern characteristic of the final consonant (whether /m/ or /l/) was subject to a shift in the direction of that for the initial consonant. It appears that this shift is more likely to occur when the consonant is word-medial than -marginal. On the other hand, the syllable-initial consonants (/m/ and /l/) remained relatively stable with respect to their inter-articulator timing, regardless of the style or rate of speech.
> (Krakow 1999: 43)

In this way, Krakow makes the important point that not only timing itself, but also the *stability* of timing can be position dependent: there is less stability in the coda (a point that will also be relevant later in coordination of segment sequences). The question of whether syllable affiliation is better seen as gradient or discrete is currently still under discussion (e.g., Scobbie and Pouplier 2010: 252).

Let us now consider a wider range of multi-gesture segments. A good point of departure is the extensively cited paper of Sproat and Fujimura (1993) on the component gestures of American English /l/ (extensive discussion in Krakow 1999, who also reproduces one of the main figures). In many dialects of English, /l/ is considered light in onsets and dark in codas. The assumption is that the latter allophone involves an additional gesture of the dorsum, traditionally referred to as velarization. Note, though, that recent work claims that liquids are generally composed of two gestures (see Proctor 2011: Recasens 2012). In the variety of American English investigated by Sproat and Fujimura, there was clear dorsal activity for the onset /l/s as well. The striking finding, if /l/ is regarded as inherently multi-gestural, is that the gestural timing showed a very clear dependence on position of /l/ in the syllable. Target achievement for the dorsal movement tended to precede target achievement for the apical movement in final position, but to follow it in initial position. If we consider the dorsal component of /l/ articulation a vocalic gesture, then we might regard the positional effect as one in which the vocalic gesture is attracted toward the syllable nucleus while the consonantal apical gesture is associated with the syllable margin.

Browman and Goldstein (1995) looked at the pattern of multi-gesture coordination for /l/ as synchronous in initial position and asynchronous in final position. They pointed to the fact that in Sproat and Fujimura's (1993) data the amount by which the dorsum led the tip in final position was greater in absolute terms than the amount by which it followed the tip in initial position (averaged over subjects about 60ms vs. 30ms).

In addition to illustrating effects of syllable position, Sproat and Fujimura's (1993) data illustrate very elegantly how the timing for syllable final laterals can vary over a whole continuum of values depending on the segmental and prosodic context. This variation probably represents a syllable-related timing feature that goes well beyond laterals and nasals.

We indicated above some general reservations about regarding the velar gesture in a nasal and the dorsal gesture in a liquid as somehow similar in status (e.g., secondary, more open). To this general reservation, one could now add the observation that the dorsal gesture for dark /l/ in English is probably not secondary in the same sense as, for example, palatalization in Russian or Serbo-Croat. The dorsal component of /l/ is very often much more resistant to reduction than the coronal component (cf. Scobbie and Pouplier 2010). It is also important to realize that the gestural constellation referenced in Sproat and Fujimura's (1993) investigation is probably just one of quite a large number of possibilities that could be found for liquids within and across language. Consider, for example, languages with a dorsal retraction gesture only in final position (e.g., two speakers of Quebec French in Gick et al. 2006), or only a clear /l/ with a raised tongue body or potentially no secondary gesture at all (e.g., German, Romanian: Marin and Pouplier 2014; Recasens, Fontdevila, and Pallarès 1995), or a dorsal gesture both initially and finally but timed synchronously with the coronal gesture in both positions (two speakers of Serbo-Croat, Gick et al. 2006). The latter case is particularly interesting because of the presence of palatalization in the Serbo-Croat sound system. The Serbo-Croat /l/ investigated by Gick et al. was the non-palatalized liquid. The presence of a palatalized variant might lead to constraints on interarticulatory coordination. Unfortunately, little about the effects of secondary articulations on intergestural coordation is currently known (but see Zsiga 2003, on a timing comparison of English and Russian).

In point of fact, "classical" secondary articulations like palatalization have indeed themselves hardly been investigated in terms of interarticulatory coordination. A rare exception to this generalization is a study by Kochetov (2006). Kochetov examined the labial and dorsal gestures for Russian /p^j/ in initial and final position (also comparing it with the phonemic sequences /pj/ and /jp/). As in the studies discussed above there was an effect of syllable position. Moreover, the general direction of the effect was the same in the sense that the dorsal gesture was timed earlier relative to the labial gesture in final position compared to initial position. However, there was also one substantial difference: the gestures were more synchronous in final than initial position. Kochetov raised the important issue of gestural recoverability as a constraint on interarticulatory timing. For initial /p^j/ the observed slight delay of the dorsal gesture relative to the labial gesture is advantageous in ensuring firstly that the dorsal gesture is not completely obscured by the more anterior, more constricted labial gesture, and also that it contributes to probably very salient perceptual differences in the burst properties of palatalized vs. non-palatalized /p/. The notion of perceptual recoverability will also be a major issue in our discussion of segment sequences below.

Russian is actually a very interesting language for investigating recoverability constraints because, as Kochetov's (2006) data show, the dorsal component of /p^j/ cannot be delayed to such an extent as to lead to confusion with /pj/. Kochetov suggests that the syllable-position-dependent timing pattern for /p^j/ is more similar to findings of Gick (2003) for the English approximant /w/ than most of the nasal and liquid findings outlined above (the approximant also showed greater asynchronicity in initial than final position). Once again we would be inclined to not push any parallels too far because of the controversial phonological status of coda /w/ in English. However, the case of /w/ does provide a convenient link to a further gap in our understanding. If /w/ is regarded in traditional terms as a labial-velar double articulation, then one might ask whether it has any similarities in timing with the only other doubly articulated consonant of appreciable frequency, namely the corresponding plosive /kp/ (not found in English, of course, but quite common in West African languages). Here the current balance of evidence would seem to be that the labial component is timed slightly later than the dorsal component (see in particular Ladefoged and Maddieson 1996: 338). This is the opposite pattern to that found by Gick for /w/ in initial position. But this comparison must be seen as tentative in the extreme because to our knowledge articulatory timing for /w/ and /kp/ has not yet been compared within the same language.

As for avenues for future research in this area, one sound that is likely to attract increasing attention is American English /r/. This sound is a classic case of a multi-gestural segment, but most likely with three constrictions (labial, palatal, pharyngeal). Timing investigations into the articulation of this sound have been difficult until recently because the pharyngeal component is not observable with traditional tools, such as electromagnetic articulography (EMA). But with the increasing availability and frame-rates of ultrasound and real-time MRI, this situation is clearly changing (see, e.g., Campbell et al. 2010). In spite of the recent availability of techniques that can be used to study American English /r/, we would like to suggest that it might be more fruitful to focus on secondary articulations of the traditional type exemplified by Kochetov's (2006) investigation. The reason for this is that secondary articulations are extremely pervasive and linguistically rich whereas American English /r/ (and perhaps its Mandarin counterpart) is a genuinely exotic sound (even if a large number of speakers happen to have it in their repertoire).

7.2.2.3 Laryngeal-oral coordination for single segments The combination of plosive articulation with the laryngeal abduction-adduction movement illustrates particularly well how intimately interarticulatory timing is bound up with linguistic distinctions. By varying the timing of the glottal gesture from early to late, one can move through the categories voiceless pre-aspirated, voiceless unaspirated, voiceless aspirated, and voiced aspirated. Cross-linguistically, less use seems to be made of the pre-aspirated category than of the other three categories, even though it is a well-known feature of Icelandic (see Hoole and Bombien 2010, for recent data and discussion). Hindi is an example of a language that contrasts voiceless unaspirated, voiceless aspirated, and voiced aspirated (plus normal voiced) stops. Accordingly, several studies on Hindi that have used fiberoptics and transillumination provide convenient illustrative material of the timing relationships between glottal and oral gestures (e.g., Kagaya and Hirose 1975: Figure 1; Benguerel and Bhatia 1980: Figures 1–2; Dixit 1989: Figures 1–4). Of course, the timing relationships are not all that is different across these three categories. For example, the voiceless aspirated category generally has a longer, larger movement than the other two; the voiced aspirated category is also striking for the particularly strong suppression of cricothyroid activity (cf. Dixit and MacNeilage 1980). Nonetheless, appropriate timing is clearly a key feature of the distinction between the voiceless unaspirated, voiceless aspirated, and voiced aspirated categories.

Of course, undoubtedly the single most famous index of interarticulatory coordination of any kind is the Voice Onset Time (VOT) measure introduced in Lisker and Abramson's (1964) seminal work. Because it can be conveniently measured in the acoustic signal, it has found widespread use not only as a measure of language differences, but also to investigate the development of timing skills in first and second language acquisition (e.g., Lowenstein and Nittrouer 2008; Flege et al. 1998) and the disintegration of coordinated behavior in speech pathology (e.g., Hoole, Schröter-Morasch, and Ziegler 1997; Ackermann and Hertrich 1997; Auzou et al. 2000). Although we will not attempt to review the vast literature on VOT, we will look briefly at one specific VOT-related issue where instrumental data of laryngeal kinematics is particularly relevant: this is the common finding that VOT varies fairly systematically with place of articulation in plosives, with VOT for /p/ generally being shorter than /t, k/ (for reviews see, e.g., Docherty 1992; Cho and Ladefoged 1999).

VOT variation as a function of place of articulation is an interesting case study in interarticulatory coordination. One simple explanation for a shorter VOT in /p/ is that the later oral release relative to the glottal abduction-adduction cycle falls out passively from differences in occlusion duration, since it is also well known that /p/ tends to have a longer oral occlusion phase. Evidence has indeed been found in the literature that the duration of the interval from peak glottal opening to release of the plosive is inversely related to the duration of the occlusion (e.g., Jessen 1999; further data and review in Hoole 2006; Hoole and Bombien 2014), but this may not be the only explanation for longer VOT in /t/ than /p/ since there is also evidence that the duration of the glottal gesture is not constant over all places of articulation, and a longer glottal gesture duration may in some cases also contribute to longer VOT in /t/. Nevertheless, speakers and languages may try to get as much mileage as possible out of a fairly constant duration of the glottal abductory-adductory cycle (cf. Weismer 1980; Shipp 1982); for example, there appears to be a cross-language tendency that long aspiration phases are associated with short occlusion phases and vice versa (Hutters 1985; recently Bombien and Hoole 2013), and the occlusion duration of voiceless fricatives is typically longer than that of aspirated voiceless plosives (fricatives are normally neither pre- nor post-aspirated).

We close this section on laryngeal-oral coordination with a brief look at differences across manner of articulation, specifically plosives vs. fricatives. Here the timing of the start of glottal abduction is of more interest than the timing of peak glottal opening since the latter, as just seen, is mainly conditioned by the control of aspiration for plosives, so plosives and fricatives differ radically but trivially in many languages. It turns out that an earlier onset of glottal abduction for fricatives (relative to the formation of the oral occlusion) is an extremely stable feature of interarticulatory coodination. For example, it is a consistent feature of the many early studies on a variety of languages conducted by Löfqvist and colleagues (e.g., Löfqvist and Yoshioka 1984; Löfqvist and McGarr 1987) and is probably the most consistent timing pattern in extensive data for German (Hoole 2006; Hoole and Bombien 2014). The stability of this relationship is probably because resistance to airflow at the glottis must be reduced quickly and early in the occlusion phase for fricatives in order to allow sufficient air-pressure build-up to drive the noise source. This is in contrast to plosives, where precise control of the aerodynamic conditions is probably more critical at the release rather than at the formation of the occlusion.

7.3 Coordination of multiple articulators for multiple segments

In this section, we will focus on the articulatory coordination of multiple gestures over longer extents of time, resulting in the articulation of successive segments. As before, we will subdivide the present section with respect to the coordination of supraglottal articulators (7.3.1) versus laryngeal-oral coordination (7.3.2).

7.3.1 Supraglottal coordination for multiple segments

First, we will address consonant-consonant (CC) and consonant-vowel (CV) timing, which brings us again into the realm of the syllable since the timing of successive segments is to a significant degree determined by syllabic structure. In keeping with the Articulatory Phonology framework adopted in this chapter, our focus in this discussion will be the Coupled Oscillator Model of Syllable Structure (COMS), which is described more completely in Goldstein and Pouplier (2014). We will also consider how timing that is conditioned by syllable position interacts with other factors known to affect CC timing, such as voicing and place of articulation.

In the previous section, we noted that, at least in English, the multiple, segment-internal gestures of liquids and nasals show differential timing patterns in syllable onset and coda position: in onset position the target plateaus of the two gestures are timed to coincide, in coda position the gestures occur more sequentially, and the earlier gesture overlaps substantially with the vowel. The coda timing pattern leads to an audible vowel darkening/rhotacization in case of liquids and vowel nasalization in the case of nasals. Importantly, these effects truly pertain to the syllable, as shown in Krakow (1993). Specifically, she found that the intervening syllable boundary prevents early lowering of the velum during the preceding vowel in the case of *see$me* compared to *seem$E*.

The observations about the differential timing of segment-internal gestures in onset and coda position have been the basis for the Articulatory Phonology model of syllable structure, namely, COMS (Browman and Goldstein 1988, 2000; Goldstein, Byrd, and Saltzman 2006; Nam et al. 2009; Kelso et al. 1986). In this model, syllabic organization is hypothesized to emerge from the specific coupling relationships between consonants and vowels. In onset position, each consonantal gesture is coupled in-phase to the vowel and anti-phase to the other onset consonant(s), if present. In coda position, VC and CC coupling is exclusively anti-phase. Thus, for onsets, the temporal midpoint of the onset cluster as a whole (i.e., the "c-center") maintains an invariant timing relationship to the vowel, irrespective of how many consonants the onset comprises. Coda clusters, on the other hand, form a "chain" rather than a "loop" (for more details, see Goldstein and Pouplier 2014). Modeling studies support the hypothesis that the difference in underlying coupling topology gives rise to the known differences in (C)CV versus VC(C) timing (Nam and Saltzman 2003). However, we have already seen in the case of liquids and nasals that the assumption of strictly anti-phase coupling for codas is probably not tenable cross-linguistically. Accordingly, more recently, the postulate that languages only ever make use of in-phase and anti-phase coupling relationships has been relaxed somewhat such that the notion of an eccentric phase has been used to characterize any phase relationships other than an in-phase relationship (see, e.g., Goldstein and Pouplier 2014).

Whereas much of the work on CV timing goes back to the pioneering studies on coarticulation by Öhman (1966, 1967) and Kozhevnikov and Chistovich (1965; for an overview of theories and data on coarticulation see Farnetani and Recasens 2010), empirical support for the COMS hypothesis that onsets are characterized by in-phase coupling between a consonant and vowel comes from a study on American English by Löfqvist and Gracco (1999), who showed that in simple consonant-vowel syllables with bilabial stops the lip movements for the bilabial closure and the tongue movements for the vowel constriction are initiated approximately synchronously (within 25 ms of one another, across speakers and vowel contexts). Similar data have been reported for Catalan, German, and Italian (Mücke et al. 2012; Niemann et al. 2011).

As for onset clusters, the c-center timing pattern described by COMS has been by and large confirmed for a variety of languages: English (Nam et al. 2009; Marin and Pouplier 2010), French (Kühnert, Hoole, and Mooshammer 2006), Georgian (Goldstein, Chitoran, and Selkirk 2007), Italian (Hermes, Mücke, and Grice 2013), Mandarin (Gao 2008), German (Pouplier 2012), and Romanian (Marin 2013). Yet exceptions to the general rule have also been reported. Romanian and Georgian may serve here to illustrate cases in which other factors seem to partly override syllable position-specific timing effects. We will also briefly discuss Slovak, which has been reported to not follow a c-center timing pattern at all.

With respect to Romanian, there are two exceptional patterns which we will consider here, one concerning onset clusters (Marin 2013) and the other one liquids in coda position (Marin and Pouplier 2014). Firstly for onsets, Marin has found that clusters fall into two groups: s-initial clusters (/sp, sk, sm/) adhere to the c-center pattern, while stop-initial clusters (/kt, kn, ks, ps/) are timed in a fashion more akin to sequential coda-timing, contrary to the predictions of COMS. One factor which separates the two groups is lexical frequency, with the latter group being of low lexical frequency (borrowings from Slavic or Greek). Marin hypothesized that there may be a causal relationship between a low lexical frequency and the lack of a complex c-center coupling topology. Specifically, she hypothesizes that the "default" onset pattern only emerges with sufficient learning and that low lexical frequency items are in fact not produced as complex onsets. The precise coupling topology that best describes the Romanian pattern will have to be addressed on the basis of modeling studies.

The other exceptional pattern in Romanian concerns the timing of liquids in coda position. Marin and Pouplier (2010) had earlier confirmed the predictions of COMS for American English codas with the exception of liquids (/-lk, -lp/; see also Katz 2012). These liquid + plosive clusters failed to exhibit the sequential timing pattern predicted by COMS; the surface timing measurements were more in line with what is expected for onsets (vowel shortening or increasing vowel overlap with increasing syllable margin complexity). Pouplier (2012) came to the conclusion that this might be related to the strong velarization of English /l/ (see above), with the dorsal gesture having a close affinity to the nucleus. This conclusion was motivated in part by data showing that German /l/ is not velarized, in contrast to American English /l/, and indeed German coda /l/ clusters behaved exactly as predicted by COMS in terms of strictly sequential timing (Pouplier 2012). However, Marin and Pouplier (2014) compared the English and German pattern to Romanian, which, like German, does not have a velarized /l/ and should therefore be expected to be similar to German. Indeed Romanian coda /l/ clusters behaved like German, yet Romanian coda /r/ clusters patterned with American English /l/ in terms of a more onset-like timing pattern. The Romanian rhotic is realized as a trill. Trills are produced with a dorsal constriction similar to the one for velarized /l/ (Proctor 2011; Recasens 2013) therefore Marin and Pouplier point out that this "special" timing pattern seen in American English dark /l/ (and /r/, Katz 2012) and Romanian /r/ may simply be a measurement issue since the measures performed in these studies do not take the dorsal gesture into account. They argue that the temporal characteristics of the ~~timing~~ gesture ~~was~~ a confounding factor in previous analyses, in terms of the duration of the ~~velar~~ gesture itself as well as in terms of possible intergestural timing change between dorsal and tongue tip gesture in clusters. Relatedly, Gick et al. (2006) found ~~for~~ the intrasegmental organization of liquids across languages, and suggest that the unexpected diversity in timing may be part of arbitrary, grammatical parameter settings that languages display. This study points to the possibility that language-specific timing differences for the multi-gesture segments may interact with syllable-position specific timing relations.

Due to the greater degrees of freedom inherent in the COMS for codas (less coupling relations are specified) there may also be a greater potential for language-specific effects to emerge in coda position compared to onset position. That said, onset clusters also show considerable variation in timing patterns. For example, Georgian onset clusters present another exception to the basic COMS pattern. This Caucasian language has come to be known for a place order effect, first described by Chitoran, Goldstein, and Byrd (2002) and Chitoran and Goldstein (2006). Chitoran and colleagues found that the relative order of place of articulation of the two consonants of a cluster is a significant determinant of articulatory timing. In

Georgian, onset clusters with place of articulation ordered from front-to-back, such as in /bg/, show more overlap than corresponding back-to-front clusters (here: /gb/). The place-order effect also exists in other languages (French: Kühnert et al. 2006; German: Bombien et al. 2010, Pouplier 2012; Korean and Russian: Kochetov, Pouplier, and Son 2007; Moroccan Arabic: Gafos et al. 2010). Importantly, there is some evidence that global onset coordination for these two cluster types is not the same, with back-to-front onset clusters deviating from the COMS pattern (Goldstein et al. 2008), although a systematic analysis is missing. Yet there is still another layer of complexity. For example, Gafos et al. (2010) argued, on the basis of data from two Moroccan Arabic speakers, that the emergence of a place-order effect may be speaker-specific and may vary with the degree of consonantal overlap characteristic for a given speaker. They assume that timing patterns with inter-speaker variability are not part of the phonological plan, and that only consistent speaker behavior (e.g., overlap in homorganic vs. heteroganic clusters) should be taken as an index of phonologically controlled timing. With respect to the place-order effect, Gafos et al. suggested that is not phonological and emerges only if overlap impacts on perceptibility. For speakers who time their consonants far enough apart, the consonants will be recoverable in either order. However, their argument for a perceptual basis to the place-order effect is not uncontroversial (Kühnert et al. 2006; Chitoran and Goldstein 2006), and given that their study comprises only two speakers, it awaits further corroboration. Along these lines, it should also be noted that the word-initial Moroccan Arabic consonant sequences of the Gafos et al. study probably do not form onset clusters (Shaw et al. 2009), in contrast to the unequivocally tautosyllabic status of these word-initial clusters in Georgian.

But there is more to cluster timing than place order. For example, Bombien and colleagues (Bombien, Mooshammer, and Hoole 2013; Bombien et al. 2010; Hoole et al. 2013) have shown for German data that even within the same place order, there may be systematic differences in overlap. For example, the onset cluster /kn-/ shows less CC overlap compared to /kl-/ even though both are back-to-front clusters. Like Gafos et al. (2010), Bombien and colleagues assumed that perceptual constraints, driven in turn by the aerodynamic requirements of each specific sound in the cluster, can govern cluster timing. Unfortunately, the relationship between onset cluster and vowel was not investigated so we do not know whether the different timing patterns result in different coordination patterns with the vowel analogously to the Georgian place-order effect.

Another factor known to affect CC overlap is voicing. Bombien and Hoole (2013) investigated how CC timing differs in French and German as a function of the voicing specification of C1 (cf. Pouplier 2012 for an interaction of voicing and place order in German). When C1 is voiced in a C1C2 onset cluster, the consonants show more overlap for German. French clusters, however, are consistently similar to the German voiceless pattern, irrespective of the voicing status of C1. This is quite surprising given that voiced stops in French have negative VOT, while in German voiced stops are rather voiceless unaspirated.

The cross-language variation in timing patterns described above strongly suggests that languages differ systematically in how much they allow their consonants to overlap (Kochetov et al. 2007; Cho, this volume, Chapter 22; Zsiga 2003). If this is the case, then one might expect that the timing patterns described by COMS interact with language-specific patterns. Pouplier and Beňuš (2011) made this suggestion in the context of Slovak timing in syllables with syllabic consonants.

Slovak is a language in which word onset clusters that are phonologically analyzed as true tautosyllabic onset clusters nonetheless fail to exhibit the c-center pattern predicted by COMS. This may be related to the fact that Slovak allows for syllabic consonants (e.g., *smrt*, "death"). Pouplier and Beňuš (2011) present data showing that syllables with a consonantal nucleus are generally timed differently from syllables with a vocalic nucleus, which suggests that in-phase coupling between onset and nucleus may only be possible at all if the nucleus is a vowel: Slovak C-C onset nucleus sequences were found to be less overlapped than C-V onset nucleus sequences. It could be that this difference emerges as follows. When onset consonant and vowel sequences are timed synchronously at movement onset, the vowel will extend temporally beyond the consonant, due to the slower movement parameterization of vowels compared to consonants. When the nucleus is a consonant, there is no such asymmetry in movement dynamics between onset and nucleus, therefore in-phase coupling would endanger perceptual recovery.

The Slovak data once more underscore the reality of considerable language-specific effects in intersegmental timing. These effects have led to the proposal that Articulatory Phonology be incorporated into a formal phonological model such as Optimality Theory (Prince and Smolensky 2008), in order to be able to formalize the arbitrary choices that languages make with respect to articulatory timing (for work in this area see Gafos 2002; Gafos and Beňuš 2006; Davidson 2006). This modification would also allow us to formalize in a predictive fashion how perceptual constraints may impact articulatory timing, even though much more experimental work is required to uncover this relationship.

There have been only a few other concrete suggestions to date how to modify the basic assumptions of COMS in order to capture the range of empirically observed timing relations across speakers and languages. With respect to interspeaker variability, Goldstein et al. (2008) suggested that such variability is due to variations in coupling strength and/or articulator weights. For instance, they argued that sequential onset effects could arise in a C1C2V sequence when C1 is less tightly coupled to the vowel compared to C2 in a particular speaker's production. This kind of speaker-to-speaker variation is seen as a natural by-product of a self-organizing system, and is therefore not phonological. With respect to cross-language variability, Browman and Goldstein (2000) suggested that anti-phase V-C coupling in coda may be generally weaker than in-phase C-V coupling in onset, and discuss the incorporation of coupling strength as a means to think about the phonological notion of syllable weight in gestural terms (see also Nam 2007). Arguably, it also provides a basis for the kind of language-specific effects we saw for liquids to emerge in coda position. Also, on the topic of cross-language variability in timing relations, there are the cases of timing differences between cluster types, such as the already mentioned place order effect, which is phonologically relevant in Georgian (see Chitoran et al. 2002). Goldstein et al. (2008) proposed that the back-to-front and front-to-back clusters are differentiated in Georgian by the coupling of the consonantal release gesture. For a front-to-back cluster, only the closure, but not the release gestures of C1 in a C1C2 are coupled in-phase to the vowel, resulting in more overlap compared to back-to-front clusters in which it is the *release* of C1 and closure of C2 that are coupled in-phase to the vowel.

Another route to achieve a greater variety of coupling topologies is discussed by Goldstein et al. (2008) in the context of English /l/. The proposal is that "secondary" gestures (our reservations about this term have already been expressed above) such as the tongue body gesture of English /l/ may not be directly coupled to the vowel. In that case, only the tongue tip gesture would bear a direct relation to the vowel (in-phase coupling), the tongue body gesture would be linked to the tongue tip gesture only. Note that this proposal arguably acknowledges the segment as a unit of speech production, a topic on which there has been little consensus among Articulatory Phonology proponents (see for discussion Browman and Goldstein 1986; Saltzman and Munhall 1989; Byrd et al. 2009; Hoole 2006).

Overall, the detailed ramifications of the various proposals for modifications to the basic COMS model will have to be worked out by modeling studies in combination with more comprehensive datasets, but the list of factors affecting CC and CV timing is long, and we have covered only a few. Moreover, language-specific temporal settings can be expected to interact with basic COMS patterns in a complex fashion.

In concluding this section, we would like to point out that higher-order prosodic boundaries, such as feet or prosodic words, likewise have an effect on articulatory organization. There is, however, relatively little work on articulatory coordination in this area. The interested reader is referred to, among others, Cho (2006), Saltzman et al. (2008), Byrd and Choi (2010), Hoole et al. (2008), and Tilsen (2009). There are also effects of stress on interarticulatory and intergestural organization. Relevant studies on these effects include de Jong, Beckman, and Edwards (1993) and Harrington et al. (1995). Finally, there is recent work on tonal alignment and its effects on the c-center timing pattern of the supralaryngeal articulators. The interested reader is referred to Gao (2008) and Mücke et al. (2012).

7.3.2 Laryngeal-oral coordination for multiple segments

In this section we will look briefly at two areas that in our opinion present interesting challenges to our understanding of how laryngeal-oral coordination relations are most appropriately specified: clusters that have a sonorant as their last element (e.g., /pl/); and sequences of purely voiceless consonants.

For clusters with sonorants, we will concentrate on syllable-initial consonants, comparing the articulation of the cluster with that of the simple voiceless onset, for example, /pl/ with simply /p/. For sounds with a clear devoicing gesture, it is intuitively appealing to specify laryngeal-oral coordination with respect to the timing of peak glottal opening. Thus, for languages like English or German, with clearly aspirated plosives in initial position, Browman and Goldstein (1986) formulate the following two rules: (1) If a fricative gesture is present, coordinate the peak glottal opening with the midpoint of the fricative. (2) Otherwise, coordinate the peak glottal opening with the release of the stop gesture.

We have already seen in section 7.2.2.3 that some fine-tuning of rule (2) may be required, in that the interval from peak glottal opening to oral release may vary with place of articulation. But if we start from the naive assumption that the devoicing gesture "belongs" to the /p/, then whether a vowel or /l/ follows the /p/ should be irrelevant to the timing, making any shift in the timing of peak glottal opening rather surprising.

In fact, there is a reasonable amount of evidence that shifts in timing of peak glottal opening do occur. Hoole (2006) investigated quite a wide range of onsets that contrasted in terms of the presence or absence of an /l/ as last element (not just /p/ vs. /pl/, but also longer sequences such as /ʃp/ vs. /ʃpl/). Fairly consistent timing differences were found, particularly when expressed as the time within the glottal abduction-adduction cycle at which the release of the last non-sonorant consonant in the onset occurred (i.e., the release of the /p/ in /p vs. pl/ as well as in /ʃp vs. ʃpl/). The shift is always in the direction of the oral release being relatively earlier in the glottal cycle for Cl or CCl onsets (compared to the corresponding simple C or CC onsets). Further evidence that such timing shifts are frequently – even if not invariably – present can be found in Jessen (1999), Tsuchida, Cohn, and Kumada (2000), and Hoole and Bombien (2014).

Although it is clear from the data presented above that timing shifts occur, it is much less clear whether there is a single organizational principle underlying these shifts. For example, it is conceivable that timing shifts occur because the duration of the glottal abduction-adduction movement remains essentially constant, while oral occlusion durations shorten in the more complex onsets with final sonorant (cf. section 7.2.2.3). While this pattern can indeed be found, it does not appear to be the only one. For example, Hoole (2006) found that the laryngeal movement was sometimes even lengthened. Faced with a variety of movement patterns, Hoole suggested that the unifying principle might be the acoustic goal of the speaker: typically, VOT is longer in the onsets terminated by a sonorant than in those without. Assuming that this represents planned behavior by the speaker (i.e., assuming that it is not purely attributable to aerodynamic conditions, cf. Docherty 1992), then there are quite a range of laryngeal-oral coordination patterns that would fulfill this goal. Crucially, this means that coordination may then be better expressed in terms of fulfilling the acoustic or communicative demands of the onset as a whole rather than with respect to a specific segment (for a similar conclusion reached from a more phonological perspective see, e.g., Kehrein and Golston 2004). This in turn points toward an issue that is too large to deal with here; namely, how interarticulatory coordination is affected by the status of the segments and syllables within the prosodic hierarchy (but see, e.g., Cho, this volume, Chapter 22).

Turning now to sequences of purely voiceless consonants, Löfqvist (1990) summarized the observable kinematics of laryngeal behavior as follows: "sounds requiring a high rate of airflow, such as fricatives and aspirated stops, are produced with a separate gesture" (296). The interesting issue with regard to understanding gestural coordination in voiceless sequences is to determine whether the number of gestures observable at the articulatory surface corresponds to the underlying gestural input. Munhall and Löfqvist (1992) investigated this question for a single sequence /s#t/, embedded in the phrase "kiss Ted" (i.e., with a word boundary), over a wide range of stress and speech rate conditions. They showed that the observable kinematic behavior ranged from two clearly distinct gestures at the slowest rates to a single-peaked pattern at the fastest. On this basis, they argued that it may be preferable to assume the presence of two underlying gestures in onset clusters like /#st/, where almost invariably only a single-peaked movement is actually observable. In this respect, their account differs from Browman and Goldstein (1986) who assume that English syllable onsets show only a single glottal gesture, where the timing of that gesture is captured in the two rules given at the beginning of this section.

The Munhall and Löfqvist (1992) study conveniently allows us to recall the concept of gestural dominance that is undoubtedly of considerable importance in understanding gestural coordination generally, and as it is applied to laryngeal-oral coordination. As discussed in Saltzman and Munhall (1989), if one assumes that fricatives dominate glottal timing more strongly than plosives do, and that dominance can be enhanced in initial position, then this may allow a single-peaked pattern at the surface to be reconciled with two underlying gestures in the input. The concept of dominance is also relevant for Browman and Goldstein's (1986) account as well, since the high dominance value for the fricative determines the order of the two rules given above. That said, we note that rather similar shifts in the timing of peak glottal opening in the /s/ of /st/ compared to single /s/ can also be found for /sl/ vs. /s/ (Hoole 2006). This makes it difficult to argue that the timing shift in /st/ is caused by an additional, albeit weak, gesture for /t/ perturbing the location of peak glottal opening away from the midpoint of the fricative, since in /sl/ there is no obvious reason to assume any competition between /s/ and /l/ for control of the glottal articulator.

Similarly to the account proposed in the first part of this section for obstruent-sonorant sequences, Hoole (2006) has argued that laryngeal-oral coordination in complex syllable onsets of purely voiceless consonants is also best understood as fulfilling constraints imposed by the aerodynamic demands of the onset as a whole. For example, in a sequence of voiceless fricative + plosive there is a very strong constraint that glottal abduction must be initiated promptly as the fricative constriction is formed; however, there are only very weak constraints as to where glottal adduction is completed. Even if adduction is completed well before the plosive occlusion is released there is no danger that the plosive will inadvertently become voiced because intraoral pressure will prevent vocal fold vibration. The aerodynamic explanation suggests that peak glottal opening may not be the crucial timing parameter and accounts for the quite radical departures from the timing rules proposed by Browman and Goldstein (1986), which are not that difficult to find (see Goldstein 1990, for further discussion).

Summarizing this section on laryngeal-oral coordination it can be said that laryngeal kinematics are attractive to study because the movement patterns are well suited to studying issues such as gestural blending. Even though we now have a wealth of information going back many years, the area remains an interesting one because it also indicates that interarticulatory coordination does not proceed segment by segment; larger structures must be taken into account. This conclusion is reminiscent of the conclusion to section 7.3.1.

7.4 Conclusion

Although we hope to have shown that a good understanding of many aspects of articulatory timing already exists, it should equally be clear that there still remains much to do. Two major unresolved issues are to understand how segmental effects interact with higher-level prosodic effects, and to join this understanding with an understanding of how individual languages develop preferred timing patterns.

REFERENCES

Ackermann, Herman and Ingo Hertrich. 1997. Voice onset time in ataxic dysarthria. *Brain and Language* 56(3): 321–333.

Auzou, Pascal, Canan Ozsancak, Richard J. Morris, Mary Jan, and Francis Eustache. 2000. Voice onset time in aphasia, apraxia of speech and dysarthria: A review. *Clinical Linguistics Phonetics* 14(2): 131–150.

Benguerel, André-Pierre and Tej K. Bhatia. 1980. Hindi stop consonants: An acoustic and fiberoscopic study. *Phonetica* 37: 134–148.

Bernstein, Nikolaj A. 1967. *The Coordination and Regulation of Movements*. Oxford: Pergamon Press.

Bombien, Lasse and Philip Hoole. 2013. Articulatory overlap as a function of voicing in French and German consonant clusters. *Journal of the Acoustical Society of America* 134: 539–550.

Bombien, Lasse, Christine Mooshammer, and Philip Hoole (2013). Articulatory coordination in word-initial clusters of German. *Journal of Phonetics* 41: 546–561.

Bombien, Lasse, Christine Mooshammer, Philip Hoole, and Barbara Kühnert. 2010. Prosodic segmental effects on EPG contact patterns of word-initial German clusters. *Journal of Phonetics* 38(3): 388–403.

Browman, Catherine. 1994. Lip aperture and consonant releases. In P. Keating (ed.), *Papers in Laboratory Phonology III: Phonological Structure and Phonetic Form*, 331–353. Cambridge: Cambridge University Press.

Browman, Catherine and Louis Goldstein. 1986. Towards an Articulatory Phonology. *Phonology Yearbook* 3: 219–252.

Browman, Catherine and Louis Goldstein. 1988. Some notes on syllable structure in Articulatory Phonology. *Phonetica* 45(2–4): 140–155.

Browman, Catherine and Louis Goldstein. 1992. Articulatory phonology: An overview. *Phonetica* 49: 155–180.

Browman, Catherine and Louis Goldstein. 1995. Gestural syllable position effects in American English. In F. Bell-Berti and L.J. Raphael (eds.), *Producing Speech: Contemporary Issues (for Katherine Safford Harris)*, 19–33. Woodbury, NY: AIP Press.

Browman, Catherine and Louis Goldstein. 2000. Competing constraints on intergestural coordination and self-organization of phonological structures. *Bulletin de la Communication Parlée* 5: 25–34.

Butcher, Andrew R. 1999. What speakers of Australian aboriginal languages do with their velums and why: The phonetics of the nasal/oral contrast. In J.J. Ohala, Y. Hasegawa, M. Ohala, D. Granville, and A.C. Bailey (eds.), *Proceedings of the XIVth International Congress of Phonetic Sciences*, 479–482.

Byrd, Dani and Susie Choi. 2010. At the juncture of prosody, phonology and phonetics: The interaction of phrasal and syllable structure in shaping the timing of consonant gestures. *Laboratory Phonology*,10: 31–60.

Byrd, Dani, Stephen Tobin, Erik Bresch, and Shrikanth Narayanan. 2009. Timing effects of syllable structure and stress on nasals: A real-time MRI examination. *Journal of Phonetics* 37: 97–110.

Campbell, Fiona, Bryan Gick, Ian Wilson, and Eric Vatikiotis-Bateson. 2010. Spatial and temporal properties of gestures in North American English /r/. *Language and Speech* 53(1): 49–69.

Catford, John Cunnison. 2001. *A Practical Introduction to Phonetics*. Oxford: Oxford University Press.

Chitoran, Iona and Louis Goldstein. 2006. Testing the phonological status of perceptual recoverability: Articulatory evidence from Georgian. Poster presented at the 10th Conference on Laboratory Phonology, Paris.

Chitoran, Iona, Louis Goldstein, and Dani Byrd. 2002. Gestural overlap and recoverability: Articulatory evidence from Georgian. *Laboratory Phonology* 7: 419–448.

Cho, Taehong. 2006. Manifestation of prosodic structure in articulatory variation: Evidence from lip kinematics in English. *Laboratory Phonology* 8: 519–548.

Cho, Taehong and Peter Ladefoged. 1999. Variation and universals in VOT: Evidence from 18 languages. *Journal of Phonetics* 27: 207–229.

Cohn, Abigail. 1993. The status of nasalized continuants. In M. Huffman and R. Krakow (eds.), *Phonetics and Phonology V: Nasals, Nasalization, and the Velum*, 329–367. Orlando, FL: Academic Press.

Davidson, Lisa. 2006. Phonotactics and articulatory coordination interact in phonology: Evidence from non-native production. *Cognitive Science* 30(5): 837–862.

de Jong, Kenneth, Mary E. Beckman, and Jan Edwards. 1993. The interplay between prosodic structure and coarticulation. *Language and Speech* 36(2–3): 197–212.

Dixit, R. Prakash. 1989. Glottal gestures in Hindi plosives. *Journal of Phonetics* 17: 213–237.

Dixit, R. Prakash and Peter F. MacNeilage. 1980. Cricothyroid activity and control of voicing in Hindi stops and affricates. *Phonetica* 37: 397–406.

Docherty, Gerard J. 1992. *The Timing of Voicing in British English Obstruents*. Berlin: Foris.

Farnetani, Edda and Daniel Recasens. 2010. Coarticulation and connected speech processes. In W.J. Hardcastle, J. Laver, and F.E. Gibbon (eds.), *The Handbook of Phonetic Sciences*, 2nd edn., 316–352. Oxford: Wiley-Blackwell.

Flege, James E., Elaina M. Frieda, Amanda C. Walley, and Lauren A. Randazza. 1998. Lexical factors and segmental accuracy in second-language speech production. *Studies in Second Language Acquisition* 20: 155–188.

Gafos, Adamantios I. 2002. A grammar of gestural coordination. *Natural Language and Linguistic Theory* 20: 269–337.

Gafos, Adamantios I. and Štefan Beňuš. 2006. Dynamics of phonological cognition. *Cognitive Science* 30(5): 905–943.

Gafos, Adamantios I., Philip Hoole, Kevin Roon, and Chakir Zeroual. 2010. Variation in timing and phonological

grammar in Moroccan Arabic clusters. *Laboratory Phonology* 10: 657–698.
Gao, Man. 2008. Tonal alignment in Mandarin Chinese: An Articulatory Phonology account. PhD dissertation, Yale University.
Gick, Bryan. 2003. Articulatory correlates of ambisyllabicity in English glides and liquids. In J. Local, R. Ogden, and R. Temple (eds.), *Papers in Laboratory Phonology 6: Constraints on Phonetic Interpretation*, 222–236. Cambridge: Cambridge University Press.
Gick, Bryan, Fiona Campbell, Sunyoung Oh, and Linda Tamburri-Watt. 2006. Towards universals in the gestural organization of syllables: A cross-linguistic study of liquids. *Journal of Phonetics* 34: 49–72.
Goldstein, Louis. 1990. On articulatory binding: Comments on Kingston's paper. In J. Kingston and M.E. Beckman (eds.), *Papers in Laboratory Phonology 1: Between the Grammar and Physics of Speech*, 445–450. Cambridge: Cambridge University Press.
Goldstein, Louis and Marianne Pouplier. 2014. The temporal organization of speech. In V. Ferreira, M. Goldrick, and M. Miozzo (eds.), *The Oxford Handbook of Language Production*, 210–227. Oxford: Oxford University Press.
Goldstein, Louis, Dani Byrd, and Elliot Saltzman. 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*, 215–249. Cambridge: Cambridge University Press.
Goldstein, Louis, Ioana Chitoran, and Elisabeth Selkirk. 2007. Syllable structure as coupled oscillator modes: Evidence from Georgian vs. Tashlhiyt Berber. *Proceedings of the XVIth International Congress of Phonetic Sciences, Saarbrücken*, 241–244.
Goldstein, Louis, Hosung Nam, Elliot Saltzman, and Ioana Chitoran. 2008. Coupled oscillator planning model of speech timing and syllable structure. *Proceedings of the 8th Phonetic Conference of China and the International Symposium on Phonetic Frontiers, Beijing, April 18–20, 2008*, CD-ROM.
Gracco, Vincent L. 1988. Timing factors in the coordination of speech movements. *The Journal of Neuroscience* 8: 4628–4639.
Gracco, Vincent L. and Anders Löfqvist. 1994. Speech motor coordination and control: Evidence from lip, jaw and laryngeal movements. *The Journal of Neuroscience* 14(11): 6585–6597.
Harrington, Jonathan, Janet Fletcher, and Corinne Roberts. 1995. Coarticulation and the accented/unaccented distinction: Evidence from jaw movement data. *Journal of Phonetics* 23: 305–322.
Hermes, Anne, Doris Mücke, and Martine Grice (2013). Gestural coordination of Italian word initial clusters – the case of impure "s". *Phonology* 30(1): 1–25.
Hoole, Philip. 2006. Experimental studies of laryngeal articulation. Habilitationsschrift, Ludwig-Maximilians Universität München. http://www.phonetik.uni-muenchen.de/~hoole/pdf/habilpgg_chap_all.pdf (accessed November 29, 2014).
Hoole, Philip and Lasse Bombien. 2010. Velar and glottal activity in Icelandic. In S. Fuchs, P. Hoole, C. Mooshammer, and M. Zygis (eds.), *Between the Regular and the Particular in Speech and Language*, 171–204. Frankfurt am Main: Peter Lang.
Hoole, Philip and Lasse Bombien. 2014. Laryngeal-oral coordination in mixed-voicing clusters. *Journal of Phonetics* 44: 8–24.
Hoole, Philip, Heidrun Schröter-Morasch, and Wolfram Ziegler. 1997. Patterns of laryngeal apraxia in two patients with Broca's aphasia. *Clinical Linguistics and Phonetics* 11: 429–442.
Hoole, Philip, Lasse Bombien, Barbara Kühnert, and Christine Mooshammer. 2008. Intrinsic and prosodic effects on articulatory coordination in initial consonant clusters. In *Proceedings of the 8th Phonetic Conference of China and the*

International Symposium on Phonetic Frontiers, Beijing, April 18–20, 2008, CD-ROM.
Hoole, Philip, Marianne Pouplier, Štefan Beňuš, and Lasse Bombien. 2013. Articulatory coordination in obstruent-sonorant clusters and syllabic consonants: Data and modelling. In L. Spreafico and A. Vietti (eds.), *Proceedings of Ratics3*, pp. 81–97. Bolzano: Bolzano University Press.
Hutters, Birgit. 1985. Vocal fold adjustments in aspirated and unaspirated stops in Danish. *Phonetica* 42: 1–24.
Jessen, Michael. 1999. Redundant aspiration in German is primarily controlled by closure duration. In *Proceedings of the XIVth International Congress of Phonetic Sciences*, 993–996.
Kagaya, Ryohei and Hajime Hirose. 1975. Fiberoptic, electromyographic and acoustic analyses of Hindi stop consonants. *Annual Bulletin, Research Institute of Logopedic and Phoniatrics* 9: 27–46. University of Tokyo.
Katz, Jonah. 2012. Compression effects in English. *Journal of Phonetics* 40: 390–402.
Kehrein, Wolfgang and Chris Golston. 2004. A prosodic theory of laryngeal contrasts. *Phonology* 21(3): 325–357.
Kelso, J.A. Scott, Elliot L. Saltzman, and Betty Tuller. 1986. The dynamical perspective on speech production: Data and theory. *Journal of Phonetics* 14: 29–59.
Kochetov, Alexei. 2006. Syllable position effects and gestural organization: Evidence from Russian. *Laboratory Phonology* 8: 565–588.
Kochetov, Alexei, Marianne Pouplier, and Minjung Son. 2007. Cross-language differences in overlap and assimilation patterns in Korean and Russian. *Proceedings of the XVIth International Congress of Phonetic Sciences, Saarbrücken*, 1361–1364.
Kollia, H. Betty, Vincent L. Gracco, and Katherine S. Harris. 1995. Articulatory organization of mandibular, labial and velar movements during speech. *Journal of the Acoustical Society of America* 98: 1313–1324.
Kozhevnikov, Valerii A. and Ludmilla Chistovich. 1965. *Speech, Articulation and Perception*. Washington, DC: Joint Publications Research Service.
Krakow, Rena A. 1993. Nonsegmental influences on velum movement patterns: Syllables, sentences, stress and speaking rate. In M.K. Huffman and R.A. Krakow (eds.), *Nasals, Nasalization and the Velum*, 87–116. San Diego, CA: Academic Press.
Krakow, Rena A. 1999. Physiological organization of syllables: A review. *Journal of Phonetics* 27: 23–54.
Kühnert, Barbara, Philip Hoole, and Christine Mooshammer. 2006. Gestural overlap and c-center in selected French consonant clusters. In H.C. Yehia, D. Demolin, and R. Laboissière (eds.), *Proceedings of the 7th International Seminar on Speech Production, UFMG Belo Horizonte*, 327–334.
Ladefoged, Peter and Ian Maddieson. 1996. *The Sounds of the World's Languages*. Oxford: Blackwell.
Lisker, Leigh and Arthur S. Abramson. 1964. A cross-language study of voicing in initial stops: Acoustical measurements. *Word* 20: 384–422.
Löfqvist, Anders. 1990. Speech as audible gestures. In W.J. Hardcastle and A. Marchal (eds.), *Speech Production and Speech Modelling*, 289–322. Dordrecht: Kluwer Academic.
Löfqvist, Anders and Vincent L. Gracco. 1999. Interarticulator programming in VCV sequences: Lip and tongue movements. *Journal of the Acoustical Society of America* 105(3): 1864–1876.
Löfqvist, Anders and Nancy McGarr. 1987. Laryngeal dynamics in voiceless consonant production. In T. Baer, C. Sasaki, and K.S. Harris (eds.), *Laryngeal Function in Phonation and Respiration*, 391–402. Boston, MA: College-Hill Press.
Löfqvist, Anders and Hirohide Yoshioka. 1984. Intrasegmental timing: Laryngeal-oral

coordination in voiceless consonant production. *Speech Communication* 3: 279–289.
Lowenstein, Joanna H. and Susan Nittrouer. 2008. Patterns of acquisition of native voice onset time in English-learning children. *Journal of the Acoustical Society of America* 124(2): 1180–1191.
Marin, Stefania. 2013. The temporal organization of complex onsets and codas in Romanian: A gestural approach. *Journal of Phonetics* 41(3–4): 211–227.
Marin, Stefania and Marianne Pouplier. 2010. Temporal organization of complex onsets and codas in American English: Testing the predictions of a gestural coupling model. *Motor Control* 14(3): 380–407.
Marin, Stefania and Marianne Pouplier. 2014. Articulatory synergies in the temporal organization of liquid clusters in Romanian. *Journal of Phonetics* 42: 24–26.
Mooshammer, Christine, Philip Hoole, and Anja Geumann. 2006. Inter-articulator cohesion within coronal consonant production. *Journal of the Acoustical Society of America* 120(2): 1028–1039.
Mücke, Doris, Hosung Nam, Anne Hermes, and Louis Goldstein. 2012. Coupling of tone and constriction gestures in pitch accents. In P. Hoole, L. Bombien, M. Pouplier, C. Mooshammer, and B. Kühnert (eds.), *Consonant Clusters and Structural Complexity*. Berlin: Mouton de Gruyter.
Munhall, Kevin and Anders Löfqvist. 1992. Gestural aggregation in speech: Laryngeal gestures. *Journal of Phonetics* 20: 111–126.
Nam, Hosung. 2007. Syllable-level intergestural timing model: Split-gesture dynamics focusing on positional asymmetry and moraic structure. In J. Cole and J.I. Hualde (eds.), *Papers in Laboratory Phonology*, vol. 9, 483–506. Berlin: Mouton de Gruyter.
Nam, Hosung and Elliot Saltzman. 2003. A competitive, coupled oscillator model of syllable structure. In M.-J. Solé, D. Recasens, and J. Romero (eds.), *Proceedings of the XVth International Congress of Phonetic Sciences, Barcelona, Spain*, 2253–2256.
Nam, Hosung, Louis Goldstein, and Elliot Saltzman. 2009. Self-organization of syllable structure: A coupled oscillator model. In F. Pellegrino, E. Marisco, I. Chitoran, and C. Coupé (eds.), *Approaches to Phonological Complexity*, 299–328. Berlin: Mouton de Gruyter.
Niemann, Henrik, Doris Mücke, Hosung Nam, Louis Goldstein, and Martine Grice. 2011. Tones as gestures: The case of Italian and German. *Proceedings of the XVIIth International Congress of Phonetic Sciences,* Hongkong, 1486–1489.
Nittrouer, Susan. 1991. Phase relations of jaw and tongue tip movements in the production of VCV utterances. *Journal of the Acoustical Society of America* 90(4): 1806–1815.
Nittrouer, Susan, Kevin Munhall, J. Scott Kelso, and Betty Tuller. 1988. Patterns of interarticulator phasing and their relation to linguistic structure. *Journal of the Acoustical Society of America* 84: 1653–1661.
Öhman, Sven E. 1966. Coarticulation in VCV utterances: Spectrographic measurements. *Journal of the Acoustical Society of America* 39(1): 151–168.
Öhman, Sven E. 1967. Numerical model of coarticulation. *Journal of the Acoustical Society of America* 41(2): 310–320.
Pouplier, Marianne. 2011. The atoms of phonological representations. In Marc van Oostendorp, Keren Rice, Beth Hume, and Colin Ewen (eds.), *The Blackwell Companion to Phonology*, 107–129. Oxford: Wiley-Blackwell.
Pouplier, Marianne. 2012. The gestural approach to syllable structure: Universal, language- and cluster-specific aspects. In S. Fuchs, M. Weirich, D. Pape, and P. Perrier (eds.), *Speech Planning and Dynamics*, 63–96. Berlin: Peter Lang.

Pouplier, Marianne and Štefan Beňuš. 2011. On the phonetic status of syllabic consonants: Evidence from Slovak. *Journal of Laboratory Phonology* 2(2): 243–273.
Prince, Alan and Paul Smolensky. 2008. *Optimality Theory: Constraint Interaction in Generative Grammar*. Oxford: Wiley-Blackwell.
Proctor, Michael. 2011 Towards a gestural characterization of liquids: Evidence from Spanish and Russian. *Laboratory Phonology* 2(2): 451–485.
Recasens, Daniel. 2012. A cross-language acoustic study of initial and final allophones of /l/. *Speech Communication* 54: 368–383.
Recasens, Daniel. 2013. Coarticulation in Catalan dark /l/ and the avleolar trill: General implications for sound change. *Language and Speech* 56(1): 45–68.
Recasens, Daniel, Jordi Fontdevila, and Maria Dolors Pallarès. 1995. Velarization degree and coarticulatory resistance for /l/ in Catalan and German. *Journal of Phonetics* 23: 37–52.
Saltzman, Elliot and Kevin G. Munhall. 1989. A dynamical approach to gestural patterning in speech production. *Ecological Psychology* 1: 333–382.
Saltzman, Elliot, Anders Löfqvist, and Subhobrata Mitra. 2000. "Glue" and "clocks": Intergestural cohesion and global timing. In M.B. Broe and J.B. Pierrehumbert (eds.), *Papers in Laboratory Phonology V: Acquisition and the Lexicon*, 88–101. Cambridge: Cambridge University Press.
Saltzman, Elliot, Hosung Nam, Jelena Krivokapic, and Louis Goldstein. 2008. A task-dynamic toolkit for modeling the effects of prosodic structure on articulation. In P.A. Barbosa and S. Madureira (eds.), *Proceedings of the Speech Prosody 2008 Conference, Campinas, Brazil*, 175–184.
Saltzman, Elliot, Anders Löfqvist, Bruce Kay, Jeff Kinsella-Shaw, and Philip Rubin. 1998. Dynamics of intergestural timing: A perturbation study of lip-larynx coordination. *Experimental Brain Research* 123: 412–424.
Scobbie, James and Marianne Pouplier. 2010. The role of syllable structure in external sandhi: An EPG study of vocalisation and retraction of word final English /l/. *Journal of Phonetics* 38: 240–259.
Shaiman, Susan, Scott Adams, and Mikael Kimelman. 1995. Timing relationships of the upper lip and jaw across changes in speaking rate. *Journal of Phonetics* 23: 119–128.
Shaw, Jason A., Adamantios I. Gafos, Philip Hoole, and Chakir Zeroual. 2009. Dynamic invariance in the phonetic expression of syllable structure: A case study of Moroccan Arabic consonant clusters. *Phonology* 26: 187–215.
Shipp, Thomas. 1982. Aspects of voice production and motor control. In S. Grillner, B. Lindblom, J. Lubker, and A. Persson (eds.), *Speech Motor Control*, 105–112. Oxford: Pergamon Press.
Sproat, Richard and Osamu Fujimura. 1993. Allophonic variation in English /l/ and its implications for phonetic implementation. *Journal of Phonetics* 21(3): 291–311.
Tilsen, Sam. 2009. Multitimescale dynamical interactions between speech rhythm and gesture. *Cognitive Science* 33: 839–879.
Tsuchida, Ayako, Abigail Cohn, and Masanobu Kumada. 2000. Sonorant devoicing and the phonetic realization of [spread glottis] in English. *Working Papers of the Cornell Phonetics Laboratory* 13: 167–181.
Turvey, Michael T. 1977. Preliminaries to a theory of action with reference to vision. In R. Shaw and J. Bransford (eds.), *Perceiving, Acting and Knowing: Toward an Ecological Psychology*, 211–265. Hillsdale, N J: Lawrence Erlbaum Associates.
Turvey, Michael T. 1990. Coordination. *American Psychologist* 45(8): 938–953.

Wang, Qi Emily. 1995. Are syllables units of speech motor organization? A kinematic analysis of labial and velar gestures in Cantonese. PhD dissertation, University of Connecticut.

Weismer, Gary. 1980. Control of the voicing distinction for intervocalic stops and fricatives: Some data and theoretical considerations. *Journal of Phonetics* 8: 427–438.

Zsiga, Elizabeth C. 2003. Articulatory timing in a second language: Evidence from Russian and English. *Studies in Second Language Acquisition* 25(3): 399–432.