

# Temporal Organization of Complex Onsets and Codas in American English: Testing the Predictions of a Gestural Coupling Model

Stefania Marin and Marianne Pouplier<sup>1</sup>

This study systematically investigates the temporal organization of American English onset and coda consonant clusters on the basis of kinematic data. Results from seven speakers suggest that consonants in complex onsets are organized globally with respect to the following vowel, while consonants in complex codas are organized locally relative to the preceding vowel. These results support the competitive coupling model hypothesized for complex onsets, a model according to which consonant gestures in onsets are each coupled in-phase to the vowel, and antiphase with each other. The results are overall also consistent with the noncompetitive coupling relations assumed for codas, by which only the first consonant in a cluster is coupled antiphase with the vowel, and any subsequent consonants are coupled antiphase to each other. However, our data also show that the segmental composition of the cluster affects the timing relationship in codas, particularly /lC/ coda clusters pattern differently from other clusters and do not adhere to the predicted timing relations. The data contribute to our understanding of the interaction of linguistic structure and motor control of the articulators in speech production.

**Keywords:** motor control in speech production, articulatory timing, syllable organization, gestural coupling, c-center

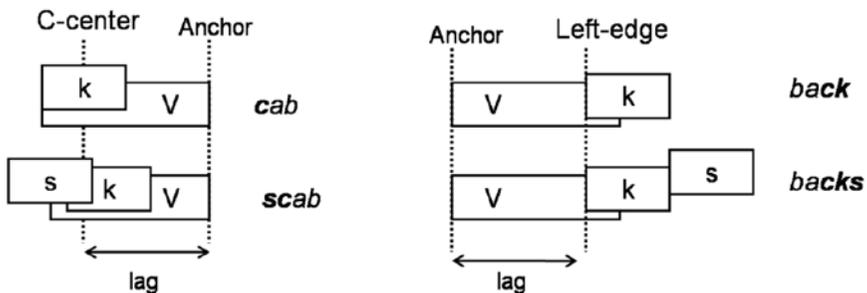
In speech production, the complex coordination of the articulatory movements is governed by linguistic structure. For example, the signaling of old and new information (Fowler & Housum, 1987), but also prosodic structure such as intonational boundaries and syllable affiliation (Beckman, Edwards, & Fletcher, 1992; Byrd & Saltzman, 2003; Fougeron & Keating, 1997; Turk & Shattuck-Hufnagel, 2000) shape the phonetic realization of an utterance, yet the precise ways in which linguistic structure and motor control interact in speech production are not well understood. The syllable is one prominent case in point: Decades of linguistic analysis of language have shown that the syllable is a relevant unit of linguistic organization in the world's languages, and that many grammatical processes can only

---

The authors are with the Institute of Phonetics and Speech Processing, University of Munich, Munich, Germany.

be understood with reference to the syllable (Blevins, 1995). Tracing the phonetic correlates of syllable structure in articulatory and acoustic records of speech has proven difficult (cf. Krakow, 1999 for an overview), yet some headway has been made in work focusing on differences in the temporal coordination of articulatory gestures in onset and coda position (Browman & Goldstein, 1988; Byrd, 1995; Gick, 2003; Giles & Moll, 1975; Goldstein, Nam, Saltzman, & Chitoran, 2008; Hermes, Grice, Mücke, & Niemann, 2008; Honorof & Browman, 1995; de Jong, 2003; Krakow, 1993; Shaw, Gafos, Hoole, & Zeroual, 2009; Sproat & Fujimura, 1993). This work has shown that both consonant-vowel timing and consonant-consonant timing in clusters differ in onset and coda position. Successive onset consonants overlap relatively less with each other and relatively more with the vowel compared with coda consonants. However, the results are, as shall be detailed below, partly contradictory and the timing patterns have only been investigated for a few specific cluster types. It is therefore unclear how general the differences in timing as a function of syllable position are. In this paper, we examine the temporal organization of syllable-initial (onset) and syllable final (coda) consonant clusters in American English over a wider range of onset and coda clusters, to investigate whether and how syllable structure governs articulatory timing.

The empirical observations of consonantal timing differences as a function of syllable position have led to the proposal that onset clusters are timed globally as an ensemble of articulatory gestures with the following vowel, so that the midpoint of the entire cluster, the so-called “c-center”, maintains a stable relationship with the vowel, regardless of onset composition (singleton or complex onset). Coda consonants on the other hand have been hypothesized to be timed locally with the preceding vowel, so that the left-most edge of the cluster is in a stable relationship with the vowel regardless of the number of consonants in the cluster (Browman & Goldstein, 1988, 2000). These predicted patterns are shown schematically in Figure 1 and will be referred to as the “C-center hypothesis”: After lining up both singleton (*cab*, *back*) and cluster (*scab*, *backs*) words with respect to a constant



**Figure 1** — Timing patterns for complex onsets and codas as predicted by the C-center hypothesis: onsets line up along the c-center landmark, codas line up along the left-edge, indicated by the dashed vertical lines. The lag represents the constant articulatory timing predicted by the C-center hypothesis. Boxes are schematic general temporal representations of consonant and vowel gestures (because vowel gestures are slower, hence longer, than consonant gestures, they have been arbitrarily pictured as twice the size of consonants).

reference point in the word (an anchor), simple and complex onsets are hypothesized to align along their c-centers, while simple and complex codas are hypothesized to align along their left-edge.

These distinct timing patterns as a function of syllable position have been hypothesized to arise from specific coupling modes which can be observed in many different domains of motor control and skilled action. Finger tapping and limb coordination research has shown that there are two intrinsic coupling modes that require no learning and can be stably maintained—in-phase and antiphase (Haken, Kelso, & Bunz, 1985; Turvey, 1990). Two actions coupled in-phase start synchronously, and if they have the same natural frequency, remain synchronous throughout; when two actions are coupled antiphase, one action starts as the other reaches its target. In connection with speech, it has been claimed that intergestural coupling generally employs these intrinsic modes, and provides the basis for syllable structure (Browman & Goldstein, 2000; Goldstein, Byrd, & Saltzman, 2006; Nam, 2007; Nam, Goldstein, & Saltzman, 2009; Saltzman, Nam, Goldstein, & Byrd, 2006). Specifically, the onset pattern (with synchronous CV timing) is hypothesized to be the result of in-phase coupling between the consonant and the vowel, while the coda pattern (with sequential VC timing) results from antiphase coupling. In the case of multiple onset gestures, they are assumed to be coupled in-phase (onset-mode) with the vowel, but antiphase with each other (to ensure recoverability) (Figure 2). This results in competitive coupling demands on these gestures which cannot be fully met simultaneously (i.e., onset gestures cannot at the same time be all synchronous with the vowel and sequential to each other). The c-center organization pattern, with one gesture shifting away from the vowel and the other toward the vowel, resulting in a global alignment of the entire cluster relative to its midpoint, represents the “compromise” pattern that best satisfies the competing demands hypothesized to be at play in onsets. On the other hand, multiple coda gestures are assumed to be coordinated in a noncompetitive manner antiphase (coda-mode) to each other, with only one of the consonants being directly coordinated to the vowel (Figure 2). This has also been called left-edge organization. Under this view, therefore, distinct coordination patterns give rise to the stable linguistic structures traditionally referred to as syllables, with onsets being defined by in-phase coupling (with additional competitive coupling relations in the case of complex onsets), and codas by antiphase coupling.

At the articulatory level, an underlying c-center organization, as predicted for onsets, is indicated by a shift in the timing of the rightmost onset consonant toward the vowel as more consonants are added (e.g., *cab* vs. *scab*). This will cause an increasing overlap of the rightmost consonantal gesture and the following vowel,



**Figure 2** — Gestural organization for complex onsets (left) and for complex codas (right). In-phase coordination is indicated by continuous lines, antiphase coordination by dotted lines.

resulting in acoustic shortening of the vowel following a cluster compared with the vowel following a singleton (cf. Figure 1). Local gestural coordination, as hypothesized for codas, affects no such shift in timing, and consonants are expected to be added sequentially to the left edge (e.g., *back* vs. *backs*). Therefore, increasing coda complexity will not change the overlap between the leftmost consonant gesture and the preceding vowel, and hence the vowel preceding the cluster should not be acoustically shortened as a consequence (Browman & Goldstein, 1988). Note that the prediction that a preceding vowel remains durationally unaffected by increasing coda complexity runs counter the compensatory shortening phenomenon reported by Fowler (1983), Munhall, Fowler, Hawkins, & Saltzman (1992), and Shaiman (2001), who observed that vowels shortened with increasing coda complexity. Byrd (1995) attempted to resolve this issue but found that, while vowel duration remained comparable before simple codas and 2-consonant clusters, vowels actually lengthened before 3-consonant clusters, contra the predictions of both the C-center hypothesis and the compensatory shortening hypothesis.

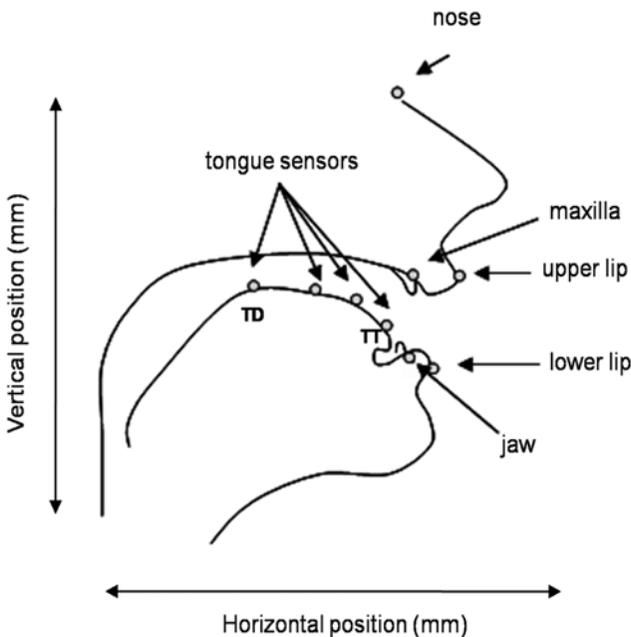
Although the hypothesized c-center organization of onsets and local organization of codas are incorporated in the task-dynamic model of speech planning and production (Browman & Goldstein, 1990, 2000; Goldstein et al., 2006; Nam, 2007; Nam et al., 2009; Saltzman & Byrd, 2000; Saltzman & Munhall, 1989; Saltzman et al., 2006), and is even used as an empirical diagnosis tool for determining syllable affiliation of consonant sequences in a variety of languages (e.g., Goldstein, Chitoran, & Selkirk, 2007 for Georgian and Tashlihyt Berber; Shaw et al., 2009 for Moroccan Arabic; Hermes et al., 2008 for Italian), previous empirical results supporting the hypothesis are relatively limited and partly contradictory. Thus, while some previous studies on English have confirmed a c-center organization for onsets and a left-edge organization for codas (Browman & Goldstein, 1988; Honorof & Browman, 1995), others found a c-center organization for coda consonants as well, at least for some speakers (Byrd, 1995). However, with the exception of the study of Honorof & Browman (1995), none of the previous studies directly compared complex onsets and complex codas: Browman & Goldstein (1988) included simple and complex onsets, but only simple codas, while Byrd (1995) included simple and complex codas but only simple onsets. Since these studies used different sets of clusters (onset /sp-, /pl-, /spl-/ for Browman & Goldstein (1988); onset /sp-, /pl-, /spl-/ and coda /-lp/, /-ps/, /-lps/ for Honorof & Browman (1995); coda /-sk/, /-ks/, /-sks/ for Byrd (1995)), it is not straightforward to resolve the partly contradictory results. For example, Honorof & Browman (1995) reported a local organization for the coda clusters they analyzed (/ -lp/, /-ps/, /-lps/) for all their speakers, while Byrd (1995) reported a c-center for the coda clusters she recorded (/ -sk/, /-ks/, /-sks/) for two of her five speakers, and a local organization for the other speakers. One possible source for the ambiguous results could be the different clusters investigated, suggesting that the hypothesized syllable timing pattern may not be as general as previously thought. With respect to onset clusters, while the previous studies unanimously reported a c-center organization (Browman & Goldstein, 1988; Goldstein et al., 2008; Honorof & Browman, 1995), they analyzed a limited range of cluster types, that is, only clusters combining a lingual and a labial constriction (e.g., /sp-, /pl-/), but not clusters with two contiguous tongue constrictions (e.g., /sk-, /kl-/), so one likewise cannot exclude the possibility that the uncovered timing pattern was cluster-specific.

In this context, we considered that a systematic revisitation of the temporal organization of consonant sequences in English was opportune. The current study used kinematic articulatory data to systematically investigate the temporal organization of several American English consonant sequences in both onset and coda positions. To overcome the limitations of the previous studies, we used a larger set of cluster types, both in onset and coda position, and we also used several analyses to uncover the temporal patterns defining onsets and codas.

## Method

### Data Collection

Articulatory movement data were collected using the electromagnetic articulograph (EMA) system at the Munich Institute of Phonetics (AG500, Carstens Medizinelektronik, cf. Hoole & Zierdt, 2006; Hoole, Zierdt, & Geng, 2007; Zierdt, Hoole, & Tillmann, 1999, for a detailed description and a comparison with other articulography systems). The system records articulatory movement over time by tracking, using an electromagnetic field, the three-dimensional position of sensors glued to various points on the subject's vocal tract. For the recordings, sensors were positioned as shown in Figure 3. Four sensors were placed on the tongue, spaced fairly equidistant: tongue tip (TT; attached approximately 1cm behind the actual tongue tip), anterior tongue body, posterior tongue body, tongue dorsum (TD; in the estimated velar constriction region). Additional sensors were placed on the upper



**Figure 3** — Position of the sensors for the articulatory data collection (sagittal view).

and lower lips, and on the lower teeth to measure jaw movement. Reference sensors were placed on the nose bridge, upper incisor (maxilla), and behind the ears (on the right and left mastoid process); the reference sensors, were used for head movement correction. All sensors except for the ones on the right and left mastoid process were fixed midsagittally. For the current study, only information on the vertical movement time series of the sensors (indicative of constriction degree) was used.

For each experiment, the system was calibrated using the AG500's standard program. Postprocessing was performed using an algorithm developed by Phil Hoole at the Munich Institute of Phonetics. The articulatory data were sampled at a frequency of 200 Hz, and band-pass filtered during postprocessing with a FIR filter (Kaiser window design, 60 dB at 5 Hz for the reference sensors, and at 20 Hz for all other sensors). Audio data were simultaneously collected with a shotgun microphone at a sampling rate of 24 kHz.

We collected data from seven adult native speakers of American English (six female and one male), with no reported speech, hearing or language problems. The speakers were familiarized with the list of utterances before data collection, and they were instructed to speak at a comfortable rate. During the actual data collection, speakers saw the target utterance on a computer screen and were visually cued (by a green box framing the utterance) when to speak. They repeated each utterance twice per trial in four randomized blocks, resulting in a targeted number of eight repetitions per utterance.

## Stimuli

The utterances were designed to allow a comparison of the consonant timing between simple onsets/codas and complex onsets/codas; therefore the stimuli consisted of target words containing onset and coda consonant clusters, and singleton controls (Table 1). We refer to the series SP, SK, SM collectively as the /s/-series, and to PL and KL as the /l/-series.

For each consonantal target set (cluster and corresponding singleton), vowel context was held constant (e.g., *tea scab*, *tea cab*; *backs aim*, *back aim*). To be able to include a variety of clusters, also clusters spanning a morpheme boundary were recorded for codas (parallel to Byrd [1995] and Honorof & Browman

**Table 1 Experimental Items: Relevant Consonants are Shown in Bold Face; Consonants Serving as Anchor Points are underlined. A Full List of the Words in Both the Vowel and Consonant Contexts and the Carrier Phrases is Given in Appendix A.**

	Series	Onsets			Codas		
/s/-series	SP	<b>span</b> <u>k</u>	<b>pan</b> <u>g</u>	<b>sank</b>	<u>gaps</u> / <b>gasp</b>	<b>gap</b>	<u>gas</u>
	SK	<b>scab</b>	<b>cab</b>	sap	<u>backs</u> / <b>bask</b>	<b>back</b>	<u>bass</u>
	SM	<b>smug</b>	<b>mug</b>	suck	<u>gums</u>	<b>gum</b>	<u>Gus</u>
/l/-series	PL	<b>plug</b>	<b>lug</b>	<b>pug</b>	<u>gulp</u>	<b>gull</b>	<u>cup</u>
	KL	<b>club</b>	<b>luff</b>	<b>cup</b>	<u>bulk</u>	<b>ball</b>	<u>buck</u>

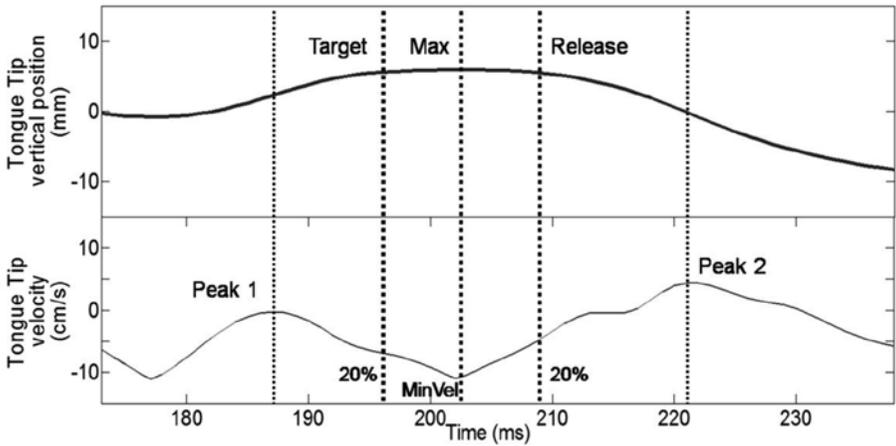
[1995]). Thus coda clusters were either monomorphemic (/sk/, /sp/, /lk/, /lp/), or hetero-morphemic (/ms/, /ks/, /ps/, with the second consonant in the cluster being the plural morpheme). The effect of morpheme structure on articulatory timing was not a focus of this study and therefore not a controlled variable. For series SP and SK, both types were recorded to determine the extent to which presence of a morpheme boundary affected the cluster's temporal organization. All onset clusters were monomorphemic. Besides onset and coda clusters, corresponding cross-word clusters were also collected (*tease cab, back same*), which allowed us to test the effect of syllable affiliation on consonant sequences. For each cluster, the consonant adjacent to the vowel (e.g., /k/ in *scab/backs*) was thus compared with a singleton in a vowel context (e.g., /k/ in *tea cab/back aim*), and also to a singleton in a consonantal context (e.g., /k/ in *tease cab/back same*). Thus a full series for an onset cluster included for example: *tea scab, tea sap, tea cab, tease cab*. An example for a full coda series is *bask aim, bass aim, bass came, backs aim, back aim, back same*.

The target words were embedded in a simple carrier phrase (e.g., *He read "tea scab" from the list*), which varied slightly across cluster types to avoid monotony (cf. Appendix A for a full stimulus list by speaker). For each target set, a constant consonantal anchor (underlined in Table 1) was used with respect to which the relative timing of the consonants of interest was measured.

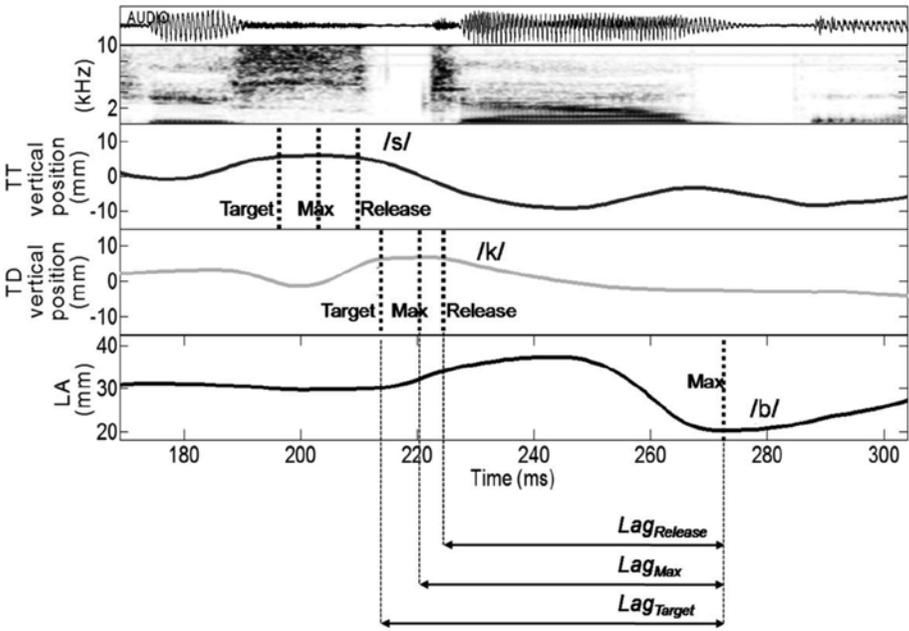
## Data Measurement

The movement time functions obtained from the EMA system were analyzed using software algorithms developed by Mark Tiede at Haskins Laboratories. The consonants produced with lip constriction (/p/, /m/) were defined on the basis of the variable lip aperture, calculated as the Euclidean distance between upper and lower lip sensors. The consonants produced with tongue constrictions were defined on the basis of the vertical movement of the relevant tongue sensor (tongue tip (TT) for /s/ and /l/, tongue dorsum (TD) for /k/). Kinematic events defining articulatory constriction—achievement of constriction, maximum constriction, constriction release, were determined on the basis of changes in the velocity profile of lip aperture/vertical tongue movement (Figure 4). The point of maximum constriction (henceforth Max) was defined as the kinematic event where velocity was at its minimum (MinVel) between two velocity maxima corresponding to the articulator moving toward and away from constriction (henceforth Peak 1 and Peak 2 respectively). Achievement of constriction (henceforth Target) was defined as the point in time at which velocity fell below 20% of the range between preceding velocity peak (Peak 1) and minimum velocity between peaks (Max). Constriction release (henceforth Release) was defined as the point where velocity exceeded 20% of the range between velocity at Max and following velocity peak (Peak 2).

While for each target word eight repetitions were recorded per speaker, not all the repetitions could be used in the analysis due to either mis-readings that went undetected during the recording session, or in some cases due to technical problems with data recording. A total of 48 tokens across subjects and experimental conditions had to be excluded from the analysis for these reasons. In addition, five of the seven speakers (Speakers 1–5) produced coda /l/ with a reduced tongue tip constriction and the measuring algorithm failed. Consequently, all the tokens produced by these



(a)



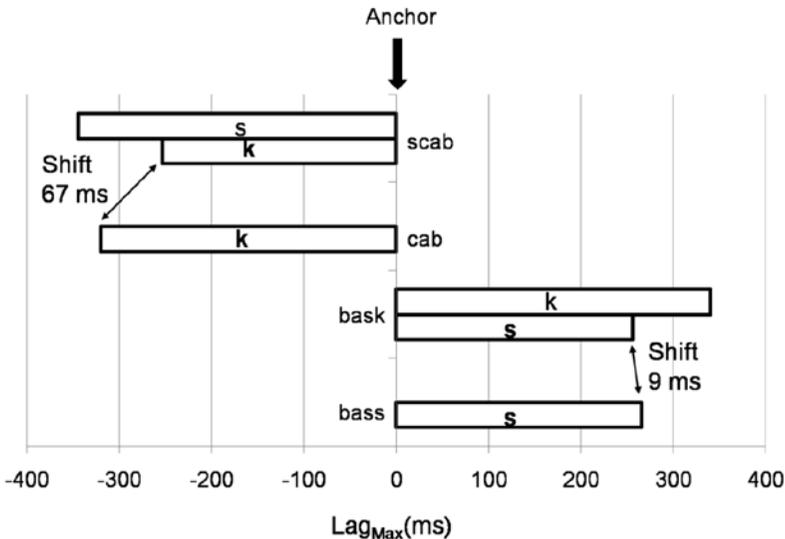
(b)

**Figure 4** — Example articulatory measurements for the word “scab”, as produced by Speaker 5. (a) The panels show vertical position (mm) and vertical velocity (cm/s) profiles of the Tongue Tip sensor during production of /s/ in “scab”. The gestural events Target, Max, Release, as well as the maximum velocity points during movement toward constriction (Peak 1) and away from constriction (Peak 2) are indicated by dotted vertical lines. (b) Illustration of relevant measurements for the entire word: Target, Max and Release for each consonant of experimental interest are indicated with a dotted vertical line for each consonant of interest; the lag from each of these landmarks ( $Lag_{Target}$ ,  $Lag_{Max}$ ,  $Lag_{Release}$ ) to the anchor point (Max of anchor consonant) is indicated by horizontal arrows. The articulatory movement data (bottom three panels) are synchronized with the acoustics, shown in the top two panels.

five speakers in the coda-/l/ condition were excluded from the articulatory analysis, but they were included in the acoustics-only analysis.<sup>2</sup>

### Timing Analysis: Consonant Shift Measurement and Vowel Duration Acoustic Measurement

For each experimental word, relative timing of the consonant gestures to the anchor point was calculated as the lag between the maximum constriction of the relevant consonant and maximum constriction of the anchor consonant ( $Lag_{\max} = \text{Consonant (Max)} - \text{Anchor (Max)}$ , cf. Figure 4b). These lags were averaged across repetitions for each subject, and relative consonant shifts were calculated to compare the average lags of single consonants with the average lags in consonant clusters. For example, in the case of onset /sk-/, shift of /k/ was calculated as the difference between /k/-lag to the anchor in *scab* and /k/-lag to the anchor in *cab*. This is illustrated in Figure 5. The same method was used for all other clusters, and the computations were done so that positive values indicated a shift toward the vowel, while negative values indicated a shift away from the vowel in the case of both onsets and codas. Recall that for the consonants of interest, singletons were recorded both in a vowel context (e.g., *tea cab*), and in a consonantal context (*tease cab*). Accordingly, two shift measures were computed for each cluster: one



**Figure 5** — Example of consonant shift measurement for cluster SK, as produced by Speaker 5 (singleton words are produced in the vowel context). The horizontal bars represent mean lag values ( $Lag_{\max}$ ) from Max of the consonant to the anchor point ( $Anchor_{\max}$ ). Arrows indicate the temporal shift of the consonant in the cluster with respect to the control singleton consonant. This example shows a shift of /k/ toward the vowel in “scab” (67 ms), and almost no shift of /s/ in “bask” (9 ms).

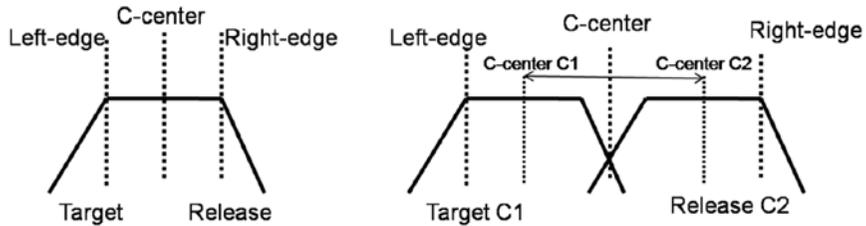
relative to the singleton in the vowel context (i.e., shift of /k/ in *scab* relative to /k/ in *tea cab*), and one relative to the singleton in the consonantal context (i.e., shift of /k/ in *scab* relative to /k/ in *tease cab*). This was done to test for coarticulatory influences on timing which may be independent of syllable position but simply due to differences in segmental context.

By this measure, if a cluster showed a c-center organization, it was expected that the consonant adjacent to the vowel in that cluster would exhibit a shift toward the vowel compared with its timing as a singleton. If a cluster was locally timed (left or right edge), no such shift toward the vowel was expected. Given that the c-center organization is hypothesized to hold for onsets but not for codas, consonants adjacent to the vowel were predicted to shift toward the vowel in onset clusters but not in codas.

Articulatory consonant shift toward the vowel was also predicted to influence the acoustic vowel duration: if the consonant adjacent to the vowel in a cluster (e.g., /k/ in *scab*) shifted toward the vowel relative to its timing in the singleton condition (/k/ in *cab*), the vowel in the cluster condition was expected to be acoustically shorter than the vowel in the singleton condition. To test this prediction, we measured the acoustic duration of the vowel following the consonants in onsets, and the vowel preceding the consonants in coda. The acoustic signal was manually labeled using Praat speech analysis software (Boersma & Weenink, n.d.). For the /s/ series, the vowel was measured from the release of the consonant to the end of periodic vibration and of vowel-specific formant contours. For /l/ onsets, the vowel was measured from the point of gross spectral change from lateral to vowel, to the end of periodic vibration (and of formant structure). Because for most of the tokens in the /l/-coda condition, no discrete acoustic boundary could be found between the vowel and the lateral, the duration of the entire vowel plus lateral sequence was used for this condition. To test whether the vowel plus consonant duration pattern (used for /l/-codas) was qualitatively different from the vowel-only duration pattern (used for the other conditions), we performed, in the /s/-coda condition, a second analysis using vowel plus consonant duration (similar to the analysis used for /l/-codas). We then qualitatively compared the result patterns of the two analyses: the one based on duration of vowel alone, and the one based on vowel plus consonant duration.

### **Stability Analysis: Comparison of Temporal Lag Measurements From Multiple Landmarks**

In terms of stability, distinct temporal events or landmarks within a cluster or singleton were predicted to be more stable in onset vs. coda, that is different landmarks were predicted to bear a constant relationship to the vowel (and hence to the anchor) across singletons and clusters. For onsets, the c-center was hypothesized to be the most stable landmark, while for codas, the left-edge was hypothesized to be most stable. Following previous studies (Browman & Goldstein, 1988; Honorof & Browman, 1995), we defined the temporal landmarks left-edge, right edge and c-center as shown schematically in Figure 6. For a singleton, the left-edge landmark coincided with achievement of target of the consonant (Target), right-edge with Release, and the c-center landmark was defined as the temporal midpoint between these two events. For a cluster, left-edge coincided with the achievement of target of the first consonant, right-edge with release of the second consonant, and c-center



**Figure 6** — Schematic illustration of the temporal landmarks left-edge, right-edge and c-center for singletons and clusters.

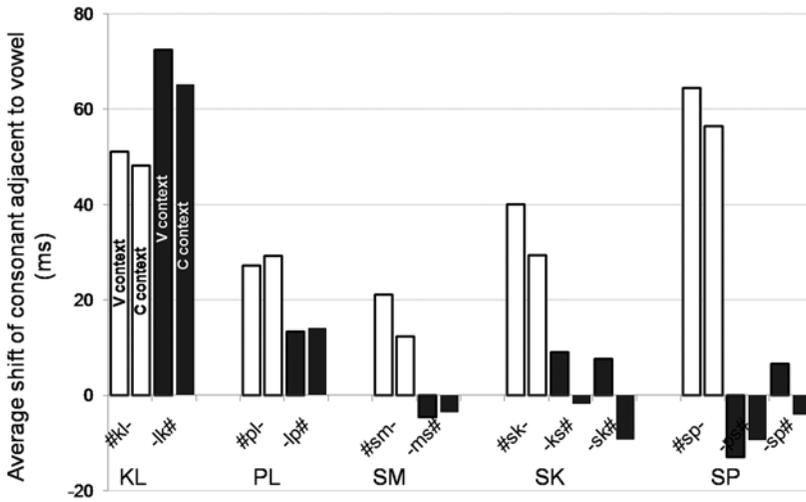
was defined as the arithmetic mean between the c-centers of the two consonants in the cluster (previously calculated in the same manner as for singletons). The temporal lags from each of these landmarks to the anchor point were calculated and used in the stability analysis. We refer to the lag from the left-edge of a singleton or a cluster to the anchor point as the Left-edge Lag; this coincided with the lag from Target of (first) consonant to Anchor (cf. Figure 4b). We refer to the lag from the right-edge to anchor as the Right-edge Lag; this coincided with the lag from Release of (second) consonant to Anchor. Finally, we refer to the lag from the c-center (as defined above) to anchor as the C-center Lag. By this measure, we expected the C-center Lag to be stable (i.e., not significantly different) across singletons and clusters in onsets, and the Left-edge Lag to be stable in codas across singletons and clusters. We expected the other lags (Left- and Right-edge Lag in onset, Right-edge and C-center Lag in coda) to be different across singletons and clusters.

## Results

### Timing Analysis: Consonant Shift

The question addressed using this measure was whether shift patterns of the consonant adjacent to the vowel in a complex cluster were different depending on their syllable affiliation (onset vs. coda), and in a manner consistent with the c-center temporal organization hypothesized for onsets (but not for codas). Specifically, we tested whether the cluster consonant adjacent to the vowel shifted toward the vowel relative to its timing as a singleton. This timing pattern was predicted for onsets but not for codas.

Across-subject means (Figure 7) indicated that for all /s/-clusters, the rightmost consonant in onsets exhibited a greater shift toward the vowel than the leftmost consonant in coda, which exhibited either a relatively smaller shift toward the vowel ([-ks], [-sk], [-sp] in the vowel context), or a small shift away from the vowel ([-ms], [-ps] in both vowel and consonant context, and [-ks], [-sk], [-sp] in the consonant context). For /l/-clusters, shifts of comparable magnitudes were observed for both onsets and codas. The shift pattern for both /s/-onsets and codas, and for /l/-onsets, was as expected, given the theoretical predictions on the organization of complex onsets and codas illustrated in Figures 1 and 2, while the shift pattern for /l/-codas



**Figure 7** — Average shift of consonant adjacent to the vowel in onset and coda clusters (7 subjects for all /s/-clusters and /l/-onsets, 2 subjects for /l/-codas). Positive values indicate a shift toward the vowel; negative values indicate a shift away from the vowel. Shift of consonant in onset cluster is shown by light bars, shift in coda cluster by dark bars. For each cluster, the left-hand bar shows shift calculated with respect to singleton consonant in a vowel (V) context (e.g., tea cab); the right-hand bar shows shift calculated with respect to singleton consonant in a consonant (C) context (e.g., tease cab).

was not anticipated, either on theoretical or empirical grounds: the only study (to our knowledge) that previously looked at the temporal organization of complex /l/ codas (Honorof & Browman, 1995) did not report a different organization pattern for /s/ and /l/ codas, possibly because they pooled /s/ and /l/ clusters together in their analysis.

Statistical analyses for the shift patterns in onsets and codas were limited to the /s/-clusters, because for the /l/-coda condition, articulatory measurements were available from only two subjects (recall that the other subjects produced coda /l/ with a reduced tongue tip constriction that could not be measured). The statistical tests were performed on the means across repetitions for each experimental word for each subject, separately on shifts calculated with respect to singleton consonants in the vowel context (e.g., /k/ in tea cab), and on shifts calculated with respect to singleton consonants in the consonant context (e.g., /k/ in tease cab). Mixed factorial ANOVAs were carried out on the consonant shift values averaged across repetitions, with fixed factors: Syllable Position (onset or coda) and Cluster Type (SM, SK, SP), and random factor: Subjects. For the vowel context, the ANOVA was significant for both main effects (Syllable Position:  $F(1, 6) = 50.894$ ,  $p < .001$ ; Cluster Type:  $F(2, 12) = 5.088$ ,  $p = .025$ ) and the interaction (Syllable Position\*Cluster Type:  $F(2, 12) = 23.315$ ,  $p < .001$ ). For the consonant context, the

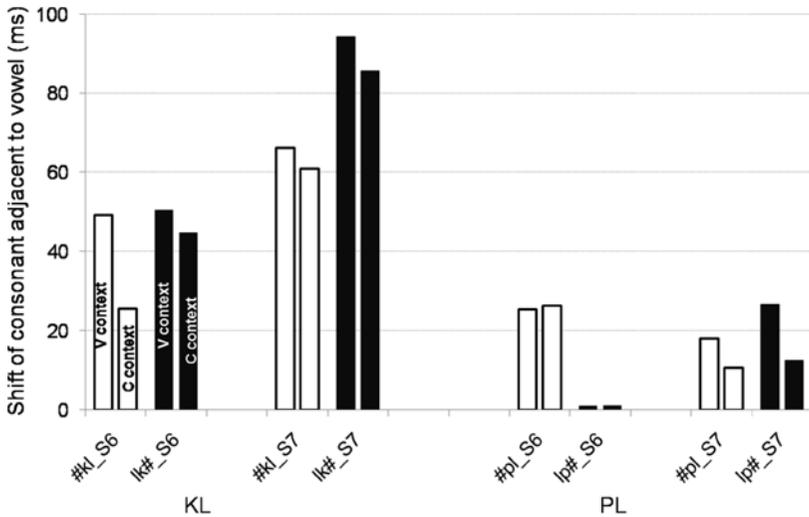
ANOVA was significant for Syllable Position ( $F(1, 6) = 15.090, p = .008$ ), and the interaction between main effects ( $F(2, 12) = 6.285, p = .014$ ), but not for Cluster Type ( $F(2, 12) = 2.619, p = .114$ ).

For the SK and SP coda clusters, paired samples *t* tests were conducted to test whether the monomorphemic clusters (/sC/) differed from the hetero-morphemic ones (/Cs/). Differences between the two coda types (Cs/sC) turned out not significant (Vowel condition: /-ks/ vs. /-sk/:  $t(6) = .158, p = .880$ ; /-ps/ vs. /-sp/:  $t(6) = -1.505, p = .183$ ; Consonant condition: /-ks/ vs. /-sk/:  $t(6) = 1.288, p = .254$ ; /-ps/ vs. /-sp/:  $t(6) = -0.375, p = .721$ ).

The statistical result showed that the shift patterns observed in onsets and codas were indeed significantly different from each other, with greater shifts in onsets than in codas, as expected if onsets were organized globally around a c-center, and codas were organized locally with respect to the cluster's left-edge. Furthermore, this result was independent of whether the shifts were calculated relative to singletons in a vowel or consonant context, suggesting that the difference in shift between onsets and codas was quite robust to contextual variation. Thus, the consonant adjacent to the vowel in an onset cluster was timed differently from a singleton, relative to the anchor point, regardless of whether the singleton was in a context similar to its production in a cluster (the consonant context), or different from it (the vowel context). To exemplify, /k/ in *scab* (preceded by /s/ and followed by a vowel) was timed differently compared with both singleton /k/ in *tea cab* and in *tease cab*. The observed temporal shift of the consonant in the cluster relative to its timing as a singleton could therefore not be due to its coarticulation context in the cluster, but rather to syllable affiliation.

The effect of cluster type was significant only in the vowel context, while the interaction between syllable position and cluster type was significant in both contexts. We suspected, on the basis of inspection of the overall means, that the interaction effect was attributable to the differences between clusters in shift magnitudes in onsets (/sp-/ > /sk-/ > /sm-/). To check this possibility, post hoc comparisons were carried out between the different onset clusters. The paired samples *t* tests showed that the shift of /p-/ in onset /sp-/ was indeed significantly greater than shift of /m-/ in /sm-/ (Vowel condition:  $t(6) = 4.989, p = .002$ ; Consonant condition:  $t(6) = 5.547, p = .001$ ). Shift of /k/ in onset /sk-/, with an intermediate shift magnitude, was at trend level in comparison with both shift of /p-/ in /sp-/ and shift of /m-/ in /sm-/ (Vowel condition: /sk-/ vs. /sp-/:  $t(6) = -2.819, p = .03$ , /sk-/ vs. /sm-/:  $t(6) = 2.963, p = .025$ ; Consonant condition: /sk-/ vs. /sp-/:  $t(6) = -2.501, p = .046$ , /sk-/ vs. /sm-/:  $t(6) = 1.565, p = .169$ ). To determine statistical significance, a Bonferroni-corrected level of 0.017 was used (a level of 0.05 divided by three comparisons). Because the overall differences across cluster types in codas were small (as shown by the means), with the largest of them (between /-ps/ and /-sp/ in the vowel condition) not being significant, no further comparisons were carried out between coda cluster types, to avoid excessive statistical testing on the same data.

With respect to the /l/-clusters, the cross-subject shift averages in Figure 7 showed that these clusters behaved quite differently from /s/-clusters: /l/-coda clusters showed quite large shifts toward the vowel, and in the case of /-lk/, the average shift was actually greater than that observed in the onset. Individual subject patterns (Figure 8) for the subjects for which measurements were available for both onsets and codas (Speakers 6 and 7) showed that the subjects differed with

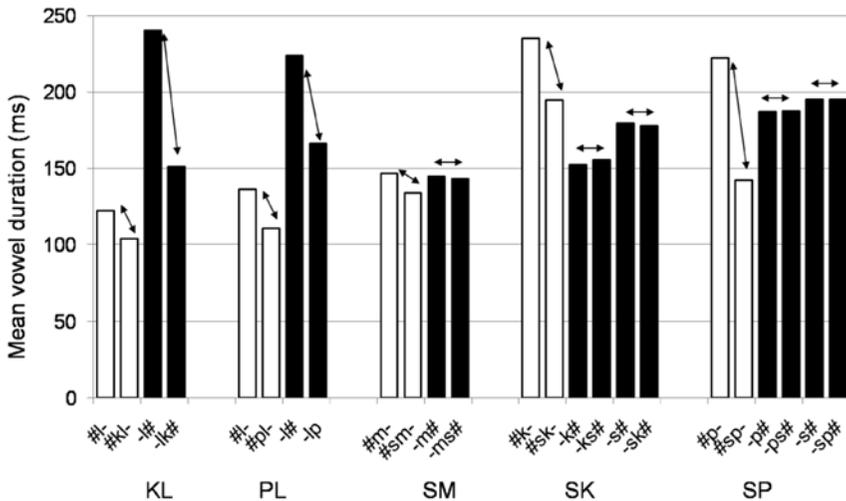


**Figure 8** — Individual shift patterns for /l/-clusters for Speakers 6 and 7. Positive values indicate a shift toward the vowel. Shift of consonant in onset cluster is shown by light bars, shift in coda cluster by dark bars. Leftward bars show shift calculated with respect to singleton consonant in a vowel (V) context; rightward bars show shift calculated with respect to singleton consonant in a consonant (C) context.

respect to this measure. Speaker 7 showed the same temporal pattern regardless of syllable position for both PL and KL series: /l/ in both onset and coda clusters shifted toward the vowel relative to singleton /l/, and the shift magnitudes were comparable (if not greater for the coda than for the onset cluster). Speaker 6 showed a pattern similar to Speaker 7 for KL (with comparable shifts toward the vowel in both syllable positions), but a pattern similar to /s/-clusters for PL (a shift toward the vowel in onset and no shift in coda). While the /l/-onset shift pattern for these subjects was uniform across subjects and clusters, exhibiting a shift toward the vowel, and thus matching the /s/-onset patterns, the /l/-coda pattern was variable across subjects and cluster types.

### Timing Analysis: Vowel Acoustic Duration

One additional way of testing the C-center hypothesis was by looking at the acoustic duration of the vowel preceding codas and following onsets, useful especially in light of the limited available articulatory measurements for /l/ in coda. On this measure, a c-center organization would be indicated by a shortening of the vowel as more consonants were added, while a local coordination should not cause such vowel shortening. The results of this measure are presented graphically in Figure 9. The pattern for the /s/-series was indeed what we expected: in onsets, the vowel was shorter after a cluster than after a singleton, indicating a c-center organization, while in codas, the vowel had roughly the same duration regardless of whether it was followed by a cluster or a singleton, indicating a left-edge organization. The coda



**Figure 9** — Duration patterns for /s/- and /l/-series: average duration of vowels across 7 speakers in cluster words is compared with duration of singletons in the vowel context. Codas are represented by dark bars, and onsets by light bars. The relevant comparisons are flagged by arrows: falling arrows indicate that the vowel in the singleton context is significantly longer than the vowel in the cluster. For the /l/-coda condition, vowel duration includes duration of the lateral.

/s/ pattern matched the results of Byrd (1995) and Honorof & Browman (1995), who observed no difference between the acoustic duration of the vowel preceding singleton and 2-consonant codas, but contradicted the pattern reported by Munhall et al. (1992), and by Shaiman (2001), who observed a consistent (albeit small for some speakers) shortening of the vowel before complex codas compared with the vowels before simple codas (the clusters and singletons collected in these studies matched our /s/-coda condition). For the /l/-series, on the other hand, the vowel was shorter in the context of a cluster than in the singleton context, regardless of syllable position, and this pattern was consistent across speakers in our data (including the speakers for which the articulatory timing data could not be evaluated).

Separate statistical tests were carried out for each syllable position (onset or coda) and for /s/- and /l/-series. Mixed factorial ANOVAs were carried out on the mean vowel duration (means across repetitions for each experimental word for each subject), with fixed factors: Syllable Composition (singleton or cluster), and Cluster Type, and with random factor: Subjects. The comparison included the words illustrated in Figure 9, that is, cluster words and those singleton controls matching the consonant adjacent to the vowel in the cluster (e.g., the control singleton for *scab* was *cab*). Only singletons in the vowel context were included (*tea cab*).<sup>3</sup>

For the /s/-series, the vowels were significantly shorter after complex onsets than after simple onsets, but they were comparable in duration before either simple or complex codas (Onset /s/:  $F(1, 6) = 93.357, p < .001$ ; Coda /s/:  $F(1, 6) = 0.002$ ,

$p = .965$ ). The factor Cluster Type (SM, SK, SP) was significant both in onset and coda position (Onset /s/:  $F(2, 12) = 36.016, p < .001$ ; Coda /s/:  $F(4, 24) = 36.535, p < .001$ ), reflecting the fact that in both syllable positions, the vowel tended to be longer in SP words than in SK words than in SM words, regardless of syllable composition (singleton or cluster). The interaction between the main factors was only significant in onset position (Onset /s/:  $F(2, 12) = 28.914, p < .001$ ; Coda /s/:  $F(4, 24) = 0.358, p = .836$ ), reflecting the fact, observable from an inspection of the means, that the vowel shortened the most after onset /sp-/, and the least after onset /sm-/ (the interaction between factors is reflected in the magnitude of the arrows in Figure 9). However, post hoc paired samples  $t$  tests carried out for each cluster type showed that the shortening pattern itself was independent of cluster type, that is, the vowel was significantly shorter after each cluster relative to the corresponding singleton, in spite of the difference in magnitude between cluster types (Onset SM:  $t(6) = 4.172, p = .006$ ; SK:  $t(6) = 5.678, p = .001$ ; SP:  $t(6) = 8.877, p < .001$ ).

For /l/ clusters, vowels were significantly shorter in the cluster words than in the singleton words both in onsets and in codas (Onset /l/:  $F(1, 6) = 36.953, p = .001$ ; Coda /l/:  $F(1, 6) = 51.908, p < .001$ ). The factor Cluster Type (KL, PL) was significant in onsets but not codas, reflecting the fact that the vowels in PL words in onsets were slightly longer than in KL words (Onset /l/:  $F(1, 6) = 10.499, p = .018$ ; Coda /l/:  $F(1, 6) = 0.006, p = .940$ ). The interaction between main factors was significant in coda position but not in onset, reflecting the fact that in codas the vowels shortened more before /-lk/ than before /-lp/, while the shortening in onsets was comparable for both cluster types (Onset /l/:  $F(1, 6) = 4.707, p = .073$ ; Coda /l/:  $F(1, 6) = 8.796, p = .025$ ). Post hoc paired  $t$  tests for the coda cluster types showed that the shortening pattern was significant for both, in spite of the difference in magnitude (Coda KL:  $t(6) = 13.620, p < 0.001$ , PL:  $t(6) = 3.886, p = .008$ ).

While the acoustic vowel duration pattern observed for the /s/ series and for the /l/ onsets confirmed the predictions of the C-center hypothesis, the duration pattern for the /l/-coda condition was unexpected. Not only was the vowel shorter in the /l/ coda cluster than in the singleton, it was overall considerably shorter relative to the shortening observed for the /l/ onsets. This pattern could not be explained as a function of the vowel involved: coda /s/ clusters were preceded by two different vowels (e.g., *gums* vs. *bask/back*s), and neither shortened consistently before /s/-codas, however the same vowel from *gums* shortened in *bulk* and *gulp*. In addition, two types of coda /s/ clusters were available (monomorphemic /-sk/, /-sp/, and heteromorphemic /-ms/, /-ks/, /-ps/), and the vowel did not shorten before either of these coda types, so it cannot be argued that shortening applied before monomorphemic (/ -lk/, / -lp/) but not before heteromorphemic clusters, since it did not apply before either type in the /s/ coda condition.

One additional confounding factor could have been the acoustic measure used in the coda /l/ condition (vowel plus lateral), distinct from the measure used in the other conditions (duration of the vowel-only). To exclude the possibility that the different pattern observed for coda-/l/ was in fact due to the different measure used, we compared, for the /s/-coda condition, the vowel-only duration with the vowel plus first consonant duration. This comparison showed that the pattern was roughly the same independent of measure used: while, not surprisingly, durations of vowel plus consonant sequences were overall longer, the duration difference as a function

of coda composition (singleton vs. cluster) was not significant regardless of measure used. A mixed factorial ANOVA on /s/-coda, carried out as before with fixed factor: Syllable Composition (singleton or cluster), and random factor: Subjects, showed that the duration of vowel plus consonant was not significantly different between cluster and singleton words ( $F(1, 6) = 2.886, p = .140$ ). This suggested that the pattern observed for the /l/-coda condition, with a significant vowel shortening before clusters compared with singletons, was not due to the measure used (vowel plus lateral), since the same measure in the /s/-coda condition (vowel plus first consonant) was not qualitatively different from the vowel-only measure.

Individual speaker patterns (Table 2) further showed that there was no difference between speakers as a function of whether they produced /l/ in coda with a full tongue tip gesture (Speakers 6 and 7), or with a reduced tongue tip (Speakers 1–5). While there was a large variation in the magnitude of shortening observed in the cluster condition as a function of both speaker and series type, ranging from about 30 ms to over 100 ms, comparable values were observed regardless of type of /l/ produced, full or reduced (e.g., for the PL series, Speaker 1, 2, 6, and 7 had comparable shortening values, and for the KL series, Speaker 2, 4, 7 had comparable shortening values, although some produced a full tongue tip /l/, and other produced a reduced tongue tip /l/; cf. also Scobbie & Pouplier (2010) who showed for British English data that the timing of the two /l/-gestures was constant independently of the spatial magnitude of the tongue tip gesture). Since no confounding factors (such as measure used or type of /l/ produced) could explain the observed pattern, it follows that both the consistent shortening pattern in the cluster condition relative to the singleton condition, and the variability in magnitude of this shortening across subjects (ranging from 30 ms to over 100 ms for the same series), must have to do with properties of coda /l/ that cause/allow this marked shortening of the vowel before an /l/ cluster, reflecting various degrees of articulatory overlap between the lateral and the vowel in the coda cluster.

## Stability Analysis: C-center vs. Left-edge Lag Stability Patterns

With respect to temporal stability patterns, the prediction was that the most stable gestural landmark in a cluster—its left-edge, its middle point (c-center) or its right-edge would bear a constant relationship to the vowel. Further, the C-center

**Table 2 Vowel Shortening Pattern by Speaker for the /L/ Coda Condition. Values Represent the Difference (In Ms) Between Vowel Plus Lateral Duration in the Singleton Condition and Vowel Plus Lateral Duration in the Cluster Condition (Positive Values Indicate a Shorter Duration in the Cluster Condition Relative to the Singleton Condition).**

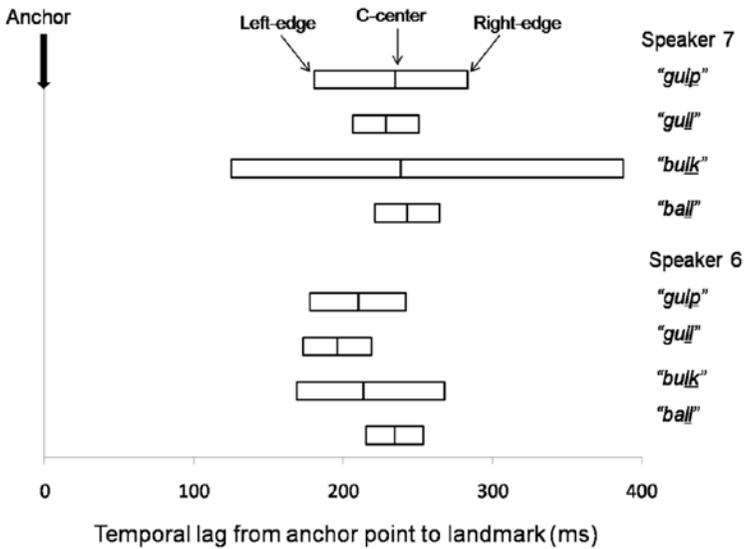
Series	Vowel duration difference (ms) by Speaker						
	1	2	3	4	5	6	7
KL	91	84	106	81	116	64	81
PL	33	37	67	71	137	29	30

hypothesis (as summarized in Figures 1 and 2 in the introduction) predicted that a stable relationship should be observed between the c-center and anchor in onsets and between left-edge and anchor in coda. To test these predictions, mixed factorial design ANOVAs were carried out separately on each of the landmark lags (Left-edge Lag, Right-edge Lag, C-center Lag) with fixed factor: Syllable Composition (singleton, cluster), and random factor: Subject. Separate tests were carried out for each syllable position (onset or coda) and for the /s/- and /l/-series (the statistical tests were carried out only on /l/-onsets since only data from 2 subjects were available for /l/-coda clusters). Because of the complex design, Cluster Type (SM, SK, SP for the /s/-series; PL, KL for the /l/-series) was not included as a factor in this analysis. Since the singleton context (vowel or consonant) did not make a qualitative difference on the syllable position effect in the shift analysis, only the singletons in the vowel context were included in the analysis (to avoid having the same consonant twice for each subject). The statistical tests were carried out on the means across repetitions for each target word for each subject. Statistical results are reported in Table 3.

For the onset condition (both /s/ and /l/ onsets), the C-center Lags were not different between clusters and singletons, but the Left-edge and Right-edge Lags were significantly different. This indicated that the stable temporal landmark across singletons and clusters in onsets was the c-center, that is, the c-center was the landmark with respect to which singletons and complex clusters lined up best.<sup>4</sup> For the /s/-coda condition, the stable landmark was the left-edge: the Left-edge Lag was not different across singletons and clusters, while Right-edge and C-center Lags differed as a factor of coda composition, indicating that the left-edge was the landmark with respect to which singleton and complex /s/-codas lined up best. For /l/-codas, because of insufficient number of speakers for statistical analyses (and hence in absence of a way to assess consistency across speakers), we informally looked at the line-up patterns separately for each speaker (Figure 10), restricting ourselves to qualitative observations. For Speaker 7, the best line-up was along the c-center (as opposed to left-edge), both within cluster type, and across both PL and KL. For Speaker 6, the best line-up across clusters was also the c-center, however /-l/ and /-lp/ alone lined up better with respect to the left-edge than the c-center, as it would be expected for coda clusters, and matching the /l/-shift pattern observed for this speaker for coda PL. Recall that Speaker 6 showed no shift toward the vowel of /l/ in the cluster /-lp/, with respect to singleton /-l/, and absence

**Table 3 Statistical Results for the Temporal Stability Measure. Nonsignificant Differences, Indicating a Stable Relation Between Singletons and Clusters with Respect to Lag Between Landmark and Anchor, are Shown in Bold Face.**

Lag measure	Onset		Coda
	/s/	/l/	/s/
Left-edge	F(1, 6)=22.118, <i>p</i> = .003	F(1, 6)=29.334, <i>p</i> = .002	<b>F(1, 6)=0.147, <i>p</i> = .714</b>
Right-edge	F(1, 6)=19.634, <i>p</i> = .004	F(1, 6)=157.137, <i>p</i> < .001	F(1, 6)=408.691, <i>p</i> < .001
C-center	<b>F(1, 6)=1.980, <i>p</i> = .209</b>	<b>F(1, 6)=2.376, <i>p</i> = .174</b>	F(1, 6)=155.535, <i>p</i> < .001



**Figure 10** — Temporal lags from anchor point to the landmarks left-edge, c-center, and right edge, for 2 speakers for /l/-codas.

of shift was taken to indicate that the second consonant in the cluster was added sequentially, without affecting the timing of the first consonant. Consequently, in the absence of a temporal shift, both singleton /-l/ and /-l/ in the coda cluster were expected to line up with respect to their left-edge, and this was indeed the pattern observed for the coda PL series for this particular speaker.

## Discussion

Our study tested the C-center hypothesis for several different onset and coda cluster types. Overall, for /s/-clusters, results from seven speakers showed a significant difference in shift patterns in onsets and codas: the consonant adjacent to the vowel shifted toward the vowel significantly more in onsets than in codas. Related to this, it was also observed that the vowel following a complex onset was acoustically shorter than the vowel following a singleton, while no acoustic duration difference was observed for the vowel preceding either singleton or complex codas. With respect to stability, the c-center was the most stable temporal landmark in onsets across singletons and clusters, while the left-edge was the most stable landmark in coda position. This pattern was the predicted one under the assumption that a global organization is at play in onsets, and a local organization is at play in codas. These results thus confirmed the competitive coupling structure hypothesized for complex onsets, and the sequential, noncompetitive coupling structure for complex codas (cf. Figure 2). Competing demands on the onset consonantal gestures—by hypothesis coupled in phase with the vowel and sequentially with each other—would result in a tight temporal organization; the observed c-center

effect was the temporal organization emerging from this competition: to satisfy the consonants' simultaneous synchronous coupling demands with the vowel, and the sequential coupling with each other, a stable compromise solution would be reached, by which the consonants in the cluster shifted leftwards and rightwards respectively compared with their singleton timing (cf. also Goldstein et al., 2008). As a result of this temporal shift, onset clusters did not line up at either left- or right-edge with onset singletons, but rather at the midpoint (c-center). On the other hand, gestures in a complex coda have been hypothesized to be coupled serially to the vowel, and this was indeed the temporal pattern observed in /s/-codas, where the consonant closest to the vowel had a tendency to preserve its timing regardless of whether it was part of a complex coda or not; in the absence of any consonantal shift, the best line up between simple and complex codas was at the edge adjacent to the vowel (the left-edge).

For /l/-clusters, the results diverged from the hypothesized pattern. Speakers showed an organization similar to that of a c-center in both complex /l/-onsets (predicted) and complex /l/-codas (unexpected), indicated by a shift toward the vowel of the consonant in the cluster adjacent to the vowel (and subsequent shortening of the relevant vowel), and also by the alignment pattern of clusters and singletons along the c-center rather than along any of the edges. This pattern for coda /l/ clusters has not been reported before—a previous study that analyzed complex /l/ codas (Honorof & Browman, 1995) looked at both /s/- and /l/-clusters pooled together, which probably obscured the different timing pattern for /l/-clusters.

One possible source for the difference observed in /l/-clusters may be the nature of American English /l/—involving both a tongue tip (consonantal) gesture and a tongue body retraction (vocalic) gesture (Browman & Goldstein, 1995; Giles & Moll, 1975; Sproat & Fujimura, 1993). While in onsets, the tongue tip and tongue body were described as being both produced roughly synchronously with the vowel, in codas the tongue body gesture preceded the tongue tip gesture. In addition, in coda position the vocalic tongue body gesture associated with /l/ was described as more extreme than in onset, while the tongue tip gesture was often more reduced relative to the onset.<sup>5</sup> The relative weakening of the more consonantal gesture (the tongue tip), concurrent with a more extreme and earlier vocalic gesture (the tongue body retraction), would result in a more “vowel-like” coda /l/ (be it a cluster or a singleton). In this context, we speculate that the temporal pattern observed for /l/-codas may be due to the perceptual consequences of coda /l/'s overlap with the vowel, combined with requirements for parallel transmission of vowels and consonants (Lieberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967). Because of coda /l/'s more vocalic (or approximant) characteristic, an increasing overlap of the vowel and the postvocalic gestures (which enhances parallel transmission) would not obscure the vowel itself perceptually. Thus, the temporal shift into the vowel (with a consequent shortening of the vowel) may be a possible organization pattern for /l/ coda consonants because perceptibility of both the vowel and the lateral can be maintained even with increased overlap (which would not be the case for obstruent coda clusters). Under this assumption, the difference between /l/-clusters and other clusters would not stem from a different underlying phonological organization of /l/ and /s/-clusters, but rather from recoverability demands, and from the overall more approximant nature of /l/ in codas. Note that the multiple coupling relations for /l/ would not be enough (by themselves) to explain the /l/ coda timing

pattern. If the observed pattern for coda /l/ clusters were the result of the relations between /l/'s multiple gestures alone, the same pattern would be expected for the coda /-ms/ cluster, since the nasal also involves two gestures—an oral gesture and a velum lowering gesture, with the velum preceding the oral gesture in coda position (cf. Krakow, 1993, 1999). However, /-ms/ patterned in our data with the other /s/ coda clusters, in spite of the nasal's composition and timing, rather than with /l/ coda clusters.

The overall greater variability in the timing pattern of /l/ coda clusters than /s/ coda clusters could thus potentially be understood as resulting from coda /l/'s relative perceptual robustness to timing variability: while the underlying coupling relations for all complex codas may be the same, varying degrees of overlap between the lateral and the vowel may be available to speakers at the production level, as long as perceptibility of the vowel is maintained, while no such variation in overlap is available for obstruents without significantly masking the vowel and hence endangering its perception. Evidence for this hypothesis, possibly in the form of computational modeling and perceptual experiments testing the relation between different overlap patterns for /l/ and perception, remains to be obtained in future research.

Also, the /s/ clusters showed an interesting interaction of cluster type and syllable position. While all onsets lined up best with respect to their c-center, as a result of the cluster consonant adjacent to the vowel shifting rightwards, the magnitude of this shift (and the resulting vowel acoustic shortening) depended on the identity of the consonant that shifted. That is, /p-/ in /sp-/ shifted significantly more than /m-/ in /sm-/, and as a tendency, /k-/ in /sk-/ showed an intermediate pattern. A previous study using the same methodology (Goldstein et al., 2008) also observed differences in shift magnitudes as a function of cluster type, with an overall greater shift for /p-/ in /sp-/ than for /l-/ in /pl-/ (we observed the same tendency in our data, cf. Figure 7). They hypothesized that the difference between the /sp-/ and /pl-/ onsets was due to the presence of /l/ (and its multiple gesture composition) in one of the clusters. Specifically, if both the tongue tip and the tongue body constrictions for /l/ were coupled with the vowel, this could result in an overall tighter coupling of /l/ with respect to the vowel (because of the multiple coupling links), and this in turn would result in a more constant timing relation of /l/ with the vowel regardless of onset composition, and hence in less shift of /l/ toward the vowel when part of a complex onset. We hypothesize that the smaller shifts for the /sm-/ onset in our data could be explained in the same way: if the velum lowering and the oral constriction for /m/ were both coupled with the vowel (cf. Byrd, Tobin, Bresch, & Narayanan, 2009 for a discussion), this would result in a tighter coordination of the nasal with the vowel and hence in less shift when the nasal was part of a complex onset. This however would leave other questions open for consideration: why for example /l/ shifted impressionistically more in /kl-/ than in /pl-/ onsets, why the glottal opening gesture in /sk-/ and /sp-/ did not have a similar effect as the velum opening in /sm-/, or why /k/ in /sk-/ tended to shift less than /p/ in /sp-/. It is probable therefore that the differences in shift magnitudes are the result not only of number of coupling relations (resulting in tighter or looser coordination patterns), but also of the exact gestures involved (for example, clusters involving two tongue gestures might result in a tighter temporal organization than clusters

involving tongue and lip gestures, although /kl-/ seems to be an exception to this generalization, see also Bombien et al. [2010] for a discussion on the temporal differences between word-initial /kl/ and /kn/ in German).

With respect to /s/-coda clusters, while our results were consistent across analyses, confirming the temporal organization pattern predicted by the C-center hypothesis (i.e., a local left-edge organization), they were in partial contradiction with previous results reported by Byrd (1995), who observed a c-center organization in coda for two of her speakers. Following Browman & Goldstein (1988), Byrd's study used standard deviation as a measure of variability of the temporal lag from left-edge, right-edge or c-center to the anchor point; the lag with the smallest standard deviation across singletons and clusters was taken to indicate the most stable relationship between consonant landmark and reference point. The rationale for this measurement was that the landmark along which both singletons and clusters lined up would have the least variability (and hence smallest standard deviation) across singletons and clusters. By this measure, Byrd (1995) found that in coda position the Left-edge Lag was smallest for three speakers, while the C-center Lag was smallest for the other two speakers. However the difference in standard deviations for these two measures was often minimal and not evaluated statistically: across subjects, the standard deviation for the C-center Lag was of 47.1 ms, and for the Left-edge Lag of 46.4 ms, with individual subjects showing a difference of no more than 6 ms between these two measures (standard deviation for the Left-edge Lag was between 0.8–5.8 ms shorter than the C-center Lag for 3 speakers, and it was between 1.4 ms–3.1 ms longer for 2 speakers). Both Left-edge Lag and C-center Lag showed considerably smaller standard deviations than the Right-edge Lag (59.3 ms across subjects). Given these minimal differences between two of the measures, it would be difficult to determine the extent to which the results in Byrd's study reflected a significantly different timing pattern between subjects, or rather reflected a limitation of the standard deviation analysis. A standard deviation analysis was probably inevitable in some of the past studies, which relied on the limited articulatory data available at the time, often of single repetitions by a single speaker (cf. Browman & Goldstein, 1988). With a greater number of speakers, more repetitions per speaker, and more stimuli, our study could employ a variety of quantitative methods, which enabled us to address the question of cluster temporal organization with greater statistical confidence, hence allowing us observe both general patterns holding across speakers and cluster types, and patterns particular to some clusters/speakers.

Finally, our coda /s/ results on the acoustic duration of the vowel as a function of syllable composition (singleton vs. cluster) contradicted the compensatory vowel shortening pattern reported by Munhall, Fowler, Hawkins, & Saltzman (1992), and by Shaiman (2001). In the coda /s/ condition, even when we ran our analysis by subject and over individual repetitions (replicating Munhall et al.'s design), the vowel was still not significantly shorter before a complex coda than before a simple coda (cf. Appendix B for mean vowel durations by coda stimulus word by subject). While we have no insights to offer regarding this discrepancy in results, we point out that one difference between the studies, which may have influenced the observed patterns, was that Munhall et al. (1992) and Shaiman (2001) used nonsense monosyllables with the target sequences (some of them possible words,

some not), while Byrd (1995), Honorof & Browman (1995) and our study used real words exclusively.

In conclusion, our results were overall consistent with the hypothesis that gestures in complex onsets and codas exhibit distinct organization patterns: a c-center organization was confirmed for all of our onset clusters, and a sequential organization was observed for /s/-coda clusters. Beyond experimentally confirming some of the theoretically predicted patterns of organization, our data also uncovered an unexpected pattern for complex /l/ codas, a pattern that remains to be further examined in the future. Our results also confirmed the reliability of using the c-center as a diagnostic tool for syllable affiliation: The overall uniform c-center organization observed for all our onset clusters suggests that testing for a c-center effect within word-initial consonant sequences is probably a good empirical way of determining syllable affiliations of these sequences in languages such as Tashlhiyt Berber or Moroccan Arabic (Goldstein et al., 2007; Shaw et al., 2009), where this issue is the object of theoretical dispute.

## Notes

1. The authors are listed alphabetically. They contributed in equal measure.
2. Reduction of the tongue tip gesture during production of coda /l/ in American English is a well-documented phenomenon (Browman & Goldstein, 1995; Giles & Moll, 1975; Sproat & Fujimura, 1993; Wrench & Scobbie, 2003).
3. The result pattern did not change if all singletons were included, nor if the comparison was done with respect to singletons in the consonant context rather than in vowel context. Only one type of singleton was included (the one matching best the context in the cluster) for a balanced number of clusters and singletons.
4. The /s/ and /l/ onsets were tested separately because of the difference between /s/ and /l/ in the coda condition. However, the statistical pattern was the same if all onset series (PL, KL, SM, SK, SP) were pooled together: C-center Lags were not different between clusters and singletons, but the other lags were significantly different ( $F_{\text{c-center}}(1, 6) = 1.269, p = .303$ ;  $F_{\text{left-edge}}(1, 6) = 77.568, p < .001$ ;  $F_{\text{right-edge}}(1, 6) = 100.839, p < .001$ ). This indicated that the observed c-center stability pattern across singletons and clusters was consistent across all onset types.
5. In our data, we could not reliably measure tongue body retraction in the context of low vowels, and we could not include adequate series with nonlow vowel context (given the limited availability of both words with /l/ coda clusters in nonlow vowel context, and matching singletons for these clusters). It is therefore not clear what the exact timing relations are between the tongue body retraction and tongue tip gestures in the cluster /l/ condition.

## Acknowledgments

We thank Phil Hoole, Susanne Waultl, and Alexandra Bigaji for help with data collection and measurements, and Mark Tiede for providing the measuring algorithm. This work was supported by the Deutsche Forschungsgemeinschaft (PO 1269/1-1).

## References

- Beckman, M.E., Edwards, J., & Fletcher, J. (1992). Prosodic structure and tempo in a sonority model of articulatory dynamics. In G. Docherty & D.R. Ladd (Eds.), *Papers in Laboratory Phonology II: Gesture, Segment, Prosody* (pp. 66–86). Cambridge UK, New York: Cambridge University Press.
- Blevins, J. (1995). The syllable in phonological theory. In J. Goldsmith (Ed.), *Handbook of Phonological Theory* (pp. 206–244). London: Blackwell.
- Boersma, P., & Weenink, D. (n.d.). Praat: doing phonetics by computer. Available from <http://www.praat.org/>.
- Bombien, L., Mooshammer, C., Hoole, P. & Kühnert, B. (2010). Prosodic and segmental effects on EPG contact patterns of word-initial German clusters. *Journal of Phonetics*, doi:10.1016/j.wocn.2010.03.003.
- Browman, C.P., & Goldstein, L. (1988). Some notes on syllable structure in Articulatory Phonology. *Phonetica*, 45, 140–155.
- Browman, C.P., & Goldstein, L. (1990). Gestural specification using dynamically-defined articulatory structures. *Journal of Phonetics*, 18, 299–320.
- Browman, C.P., & Goldstein, L. (1995). Gestural syllable position effects in American English. In F. Bell-Berti & L.J. Raphael (Eds.), *Producing Speech: Contemporary Issues* (pp. 19–34). Woodbury, NY: American Institute of Physics.
- Browman, C.P., & Goldstein, L. (2000). Competing constraints on intergestural coordination and self-organization of phonological structures. *Bulletin de la Communication Parlée*, 5, 25–34.
- Byrd, D. (1995). C-centers revisited. *Phonetica*, 52, 285–306.
- Byrd, D., & Saltzman, E. (2003). The elastic phrase: modeling the dynamics of boundary adjacent lengthening. *Journal of Phonetics*, 31, 149–180.
- Byrd, D., Tobin, S., Bresch, E., & Narayanan, S. (2009). Timing effects of syllable structure and stress on nasals: a real-time MRI examination. *Journal of Phonetics*, 37, 97–110.
- Fougeron, C., & Keating, P.A. (1997). Articulatory strengthening at edges of prosodic domains. *The Journal of the Acoustical Society of America*, 101, 3728–3740.
- Fowler, C.A. (1983). Converging sources of evidence on spoken and perceived rhythms of speech: Cyclic production of vowels in sequences of monosyllabic stress feet. *Journal of Experimental Psychology. General*, 112, 386–412.
- Fowler, C.A., & Housum, J. (1987). Talkers' signaling of "New" and "Old" words in speech and listeners' perception and use of the distinction. *Journal of Memory and Language*, 26, 489–504.
- Gick, B. (2003). Articulatory correlates of ambisyllabicity in English glides and liquids. In J. Local, R. Ogden, & R. Temple (Eds.), *Phonetic Interpretation: Papers in Laboratory Phonology VI* (pp. 222–236). Cambridge: Cambridge University Press.
- Giles, S.B., & Moll, K.L. (1975). Cinefluorographic study of selected allophones of English /N/. *Phonetica*, 31, 206–227.
- Goldstein, L., Byrd, D., & Saltzman, E. (2006). The role of vocal tract gestural action units in understanding the evolution of phonology. In M.A. Arbib (Ed.), *Action to Language via the Mirror Neuron System* (pp. 215–249). Cambridge: Cambridge University Press.
- Goldstein, L., Chitoran, I., & Selkirk, E. (2007). Syllable structure as coupled oscillator modes: evidence from Georgian vs. Tashlhiyt Berber. In J. Trouvain & W. J. Barry (Eds.), *Proceedings of the XVIIth International Congress of Phonetic Sciences* (pp. 241–244). Saarbrücken.
- Goldstein, L., Nam, H., Saltzman, E., & Chitoran, I. (2008). Coupled oscillator planning model of speech timing and syllable structure. In *Proceedings of the 8th Phonetic Conference of China and the International Symposium on Phonetic Frontiers, Beijing, April 18-20, 2008* (CD-ROM).

- Haken, H., Kelso, J.A.S., & Bunz, H. (1985). A theoretical model of phase transitions in human hand movements. *Biological Cybernetics*, *51*, 347–356.
- Hermes, A., Grice, M., Mücke, D., & Niemann, H. (2008). Articulatory indicators of syllable affiliation in word initial consonant clusters in Italian. In R. Sock, S. Fuchs & Y. Laprie (Eds.), *Proceedings of the 8th International Seminar on Speech Production*. (pp. 433–436). Strasbourg. Available from <http://issp2008.loria.fr/Proceedings/PDF/issp2008-102.pdf>.
- Honorof, D., & Browman, C.P. (1995). The center or edge? How are consonant clusters organized with respect to the vowel? In K. Elenius & P. Branderud (Eds.), *Proceedings of the XIIIth International Congress of Phonetic Sciences, Stockholm, Sweden* (pp. 552–555). Stockholm: KTH and Stockholm University.
- Hoole, P., & Zierdt, A. (2006). Five-dimensional articulography. *Stem-, Spraak- en Taalpathologie*, *14*(Supplement), 57.
- Hoole, P., Zierdt, A., & Geng, C. (2007). Beyond 2-D in articulatory data acquisition and analysis. In J. Trouvain & W. J. Barry (Eds.), *Proceedings of the XVIth International Congress of Phonetic Sciences* (pp. 265–268). Saarbrücken.
- de Jong, K.J. (2003). Temporal constraints and characterizing syllable structure. In J. Local, R. Ogden, & R. Temple (Eds.), *Papers in Laboratory Phonology VI* (pp. 253–268). Cambridge: Cambridge University Press.
- Krakow, R. (1993). Nongsegmental influences on velum movement patterns: Syllables, sentences, stress and speaking rate. In M.K. Huffman & R.A. Krakow (Eds.), *Nasals, Nasalization, and the Velum* (pp. 87–116). San Diego: Academic Press.
- Krakow, R. (1999). Physiological organization of syllables: a review. *Journal of Phonetics*, *27*, 23–54.
- Lieberman, A.M., Cooper, F.S., Shankweiler, D., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, *74*, 431–461.
- Munhall, K.G., Fowler, C.A., Hawkins, S., & Saltzman, E. (1992). ‘Compensatory shortening’ in monosyllables of spoken English. *Journal of Phonetics*, *20*, 225–240.
- Nam, H. (2007). *A Gestural Coupling Model of Syllable Structure*. PhD Dissertation, Yale University.
- Nam, H., Goldstein, L., & Saltzman, E. (2009). Self-organization of syllable structure: A coupled oscillator model. In F. Pellegrino, E. Marsico, I. Chitoran, & C. Coupé (Eds.), *Approaches to phonological complexity* (pp. 299–328). Berlin: Mouton de Gruyter.
- Saltzman, E., & Byrd, D. (2000). Task dynamic coordination of the speech articulators: A preliminary model. *Human Movement Science*, *19*, 499–526.
- Saltzman, E., & Munhall, K.G. (1989). A dynamical approach to gestural patterning in speech production. *Ecological Psychology*, *1*, 333–382.
- Saltzman, E., Nam, H., Goldstein, L., & Byrd, D. (2006). The distinctions between state, parameter and graph dynamics in sensorimotor control and coordination. In M.L. Latash & F. Lestienne (Eds.), *Motor Control and Learning over the Life Span* (pp. 63–73). Berlin, New York: Springer-Verlag.
- Scobbie, J. M. & Pouplier, M. (2010). The role of syllable structure in external Sandhi: An EPG study of vocalization and retraction in word-final English /l/. *Journal of Phonetics*, *38*, 240–259.
- Shaiman, S. (2001). Kinematics of compensatory vowel shortening: the effect of speaking rate and coda composition on intra- and inter-articulatory timing. *Journal of Phonetics*, *29*, 89–107.
- Shaw, J., Gafos, A., Hoole, P., & Zeroual, C. (2009). Temporal evidence for syllabic structure in Moroccan Arabic: data and model. *Phonology*, *26*, 187–215.
- Sproat, R., & Fujimura, O. (1993). Allophonic variation in American English /l/ and its implications for phonetic interpretation. *Journal of Phonetics*, *21*, 291–311.
- Turk, A., & Shattuck-Hufnagel, S. (2000). Word-boundary-related duration patterns in English. *Journal of Phonetics*, *28*, 397–440.

- Turvey, M. (1990). Coordination. *The American Psychologist*, 45, 938–953.
- Wrench, A.A., & Scobbie, J.M. (2003). Categorising vocalisation of English /l/ using EPG, EMA and ultrasound. In S. Palethorpe & M. Tabain (Eds), *Proceedings of the Sixth International Seminar on Speech Production*, (pp. 314-319). Sydney, Australia.
- Zierdt, A., Hoole, P., & Tillmann, H.G. (1999). Development of a system for three-dimensional fleshpoint measurement of speech movements. In J. J. Ohala, Y. Hasegawa, M. Ohala, D. Granville & A. C. Bailey (Eds.), *Proceedings of the XIVth International Congress of Phonetic Sciences, San Francisco, USA*. (pp. 73-75). San Francisco.

## Appendix A

Utterance lists. Stimulus phrases (target word + flanking word) are between quotation marks (as they were presented visually to the speakers). There were 44 target words for each speaker, with eight repetitions per word, resulting in a total of 2464 target utterances across all seven speakers. The carrier phrases were slightly changed for Subjects 3–7 to save experiment time.

Series	Context	Speakers 1–2	Subjects 3–7
PL	V#pl	I say “tea plug” again.	I say “tea plug” again.
	-p#l-	I say “deep lug” again.	I say “deep lug” again.
	V#p-	I say “tea pug” again.	I say “tea pug” again.
	V#l-	I say “tea lug” again.	I say “tea lug” again.
	-lp#V	I say “gulp it” again.	I say “gulp ax” again.
	-l#p-	I say “gull pit” again.	I say “gull packs” again.
	-p#V	I say “cup it” again.	I say “cup ax” again.
	-l#V	I say “gull it” again.	I say “gull ax” again.
KL	V#kl-	I say “tea club” again.	I say “tea club” again.
	-k#l-	I say “teak luff” again.	I say “teak luff” again.
	V#k-	I say “tea cup” again.	I say “tea cup” again.
	V#l-	I say “tea luff” again.	I say “tea luff” again.
	-lk#V	He read “bulk in” once.	I say “bulk ape” again.
	-l#k	He read “ball kin” once.	I say “ball cape” again.
	-k#V	He read “buck in” once.	I say “buck ape” again.
	-l#V	He read “ball in” once.	I say “ball ape” again.
SP	V#sp-	The phrase “bee spank” is unusual.	He saw “pea spank” on the screen.
	-s#p-	The phrase “bee’s pang” is unusual.	He saw “peace pang” on the screen.
	V#s-	The phrase “be sank” is ungrammatical.	He saw “pea sank” on the screen.

(continued)

## Appendix A (continued)

Series	Context	Speakers 1–2	Subjects 3–7
	V#p-	The phrase “be pang” is ungrammatical.	He saw “pea pang” on the screen.
	-sp#V	He read “gasp it” from the list.	He saw “gasp odd” on the screen.
	-s#p-	He read “gas pit” from the list.	He saw “gas pod” on the screen.
	-ps#V	He read “gaps it” from the list.	He saw “gaps odd” on the screen.
	-p#s-	He read “gap sit” from the list.	He saw “gap sod” on the screen.
	-p#V	He read “gap it” from the list.	He saw “gap odd” on the screen.
	-s#V	He read “gas it” from the list.	He saw “gas odd” on the screen.
SK	V#sk-	He said “pea scab” once.	I say “tea scab” once.
	s#k-	He said “piece cab” once.	I say “tease cab” once.
	V#s-	He said “pea sap” once.	I say “tea sap” once.
	V#k-	He said “pea cab” once.	I say “tea cab” once.
	-sk#V	He saw “bask it” on the screen.	I say “bask aim” once.
	-s#k-	He saw “bass kit” on the screen.	I say “bass came” once.
	-ks#V	He saw “backs it” on the screen.	I say “backs aim” once.
	-k#s-	He saw “back sit” on the screen.	I say “back same” once.
	-k#V	He saw “back it” on the screen.	I say “back aim” once.
	-s#V	He saw “bass it” on the screen.	I say “bass aim” once.
SM	V#sm-	She heard “pea smug” again.	She saw “pea smug” on the screen.
	-s#m-	She heard “peace mug” again.	She saw “peace mug” on the screen.
	V#s-	She heard “pea suck” again.	She saw “pea suck” on the screen.
	V#m-	She heard “pea mug” again.	She saw “pea mug” on the screen.
	-ms#V	The phrase “gums it” is silly.	She saw “gums odd” on the screen.
	-m#s-	The phrase “gum sit” is silly.	She saw “gum sod” on the screen.
	-m#V	The phrase “gum it” is silly.	She saw “gum odd” on the screen.
	-s#V	The phrase “Gus it” is silly.	She saw “Gus odd” on the screen.

## Appendix B

Vowel duration (ms) by word and subject for the coda /s/ series. Shown in bold are the cases where the vowel before a complex cluster was shorter than before a singleton (matching Munhall et al. (1992)'s and Shaiman (2001)'s results). Note that vowel shortening for coda in our data were not consistent either within speaker or within series.

Series	Utterance	Type	Speaker						
			1	2	3	4	5	6	7
SM	<b>gum</b>	VC	124	163	169	130	147	150	128
	<b>gums</b>	VCC	130	155	167	141	131	161	117
	<b>Gus</b>	VC	128	159	169	136	140	140	123
SK	<b>back</b>	VC	151	155	166	151	154	153	135
	<b>backs</b>	VCC	152	160	171	152	145	163	143
	<b>bass</b>	VC	178	181	200	177	166	187	167
	<b>bask</b>	VCC	178	188	189	167	167	195	160
SP	<b>gap</b>	VC	179	231	203	177	156	201	162
	<b>gaps</b>	VCC	186	226	198	165	155	221	162
	<b>gas</b>	VC	186	228	218	179	170	207	175
	<b>gasp</b>	VCC	193	210	199	199	168	220	175