CHANGES IN VOWEL VELOCITY PROFILE WITH VOWEL-CONSONANT OVERLAP

Tanner Sorensen[1] and Adamantios Gafos[1,2*]

[1]University of Potsdam, Potsdam, Germany, and [2]Haskins Laboratories, New Haven, CT, USA
{tsorense,gafos}@uni-potsdam.de

ABSTRACT

We report the effect of a coronal consonant constriction on the movement of a preceding but temporally overlapping /a/. The dependent measure is the shape of the /a/ velocity profile and the independent variable is the temporal overlap of /a/ with the subsequent coronal constriction. We show that the /a/ velocity profile develops a shallower peak and a fatter right tail as the temporal overlap of /a/ with the subsequent coronal increases. These results challenge the idea that anticipatory consonant-to-vowel coarticulation is tightly localized in time and space by showing that anticipatory consonant-to-vowel coarticulation has a longer temporal window and greater spatial extent than originally thought.

Keywords: coarticulation, velocity profile, functional data analysis, electromagnetic articulography.

1. INTRODUCTION

The properties of a consonant constrain tongue position during a preceding vowel. This anticipatory consonant-to-vowel coarticulation is thought to be tightly localized in time and space. In particular, vocal tract shape varies contextually over a short time interval before the consonant closure and over a small spatial region around the consonant constriction location [8, 12]. Nevertheless, electropalatography has revealed that anticipatory consonant-to-vowel coarticulation influences the palate contact of certain vowels as early as their acoustic midpoint (e.g. palate contact for /a/ varies contextually at its midpoint with a subsequent coronal in Italian [5]). This suggests that anticipatory consonant-to-vowel coarticulation has a longer temporal window than originally thought. Since an /a/ produces little palate contact [22], electromagnetic articulography (EMA) is well suited to measure the temporal and spatial extent of anticipatory consonant-to-vowel coarticulation in regions of the vocal tract far from the palate-air boundary. We provide the formal means to characterize this extent and report early anticipatory effects of a coronal consonant on a spatially distal /a/ movement. In particular, temporal overlap of /a/ with a subsequent coronal affects the global velocity profile shape of /a/ measured six centimeters posterior to the apex from start to end of the to-phase of /a/. The to-phase corresponds to the interval between the onset of movement and the achievement of target (i.e. it precedes the release or fro-phase of the movement). This effect indicates that anticipatory consonant-to-vowel coarticulation takes place between even spatially distal sections of the tongue over a relatively long time window preceding consonant closure.

[*]AIG, TS acknowledge support by ERC AdG 249440.

Table 1: Repetitions per target word by subject.

	/kat/	/gat/	/kaʃa/	/galebna/	/galba/	/kamʃa/	/kam/
S1	16	16	8	8	8	8	16
S2	20	20	8	8	8	8	20
S3	16	16	8	8	8	8	16
S4	16	16	9	8	8	7	16

2. DATA SET

The subjects are four speakers of Moroccan Colloquial Arabic: S1, S2, S3, and S4. S2 has training in phonetics. The other three have none. S4 is female. The other three are male.

Target words have a velar stop, then an /a/, then /t,l,ʃ,n/, possibly followed by other segments. The temporal overlap of /a/ with the coronal is made to vary by presenting stimuli with zero (/k**at**/, /g**at**/, /k**aʃ**a/, /g**al**ebna/, /g**al**ba/), one (/k**a**m**ʃ**a/), or two (/k**a**m#h**n**aja/) nonlingual consonants intervening between /a/ and the coronal.

Electromagnetic articulography (EMA) data were collected in the earlier experiment of [21], who report the experimental procedure used for S1. This is the same procedure used in other sessions with S2, S3, and S4. Three receiver coils were placed on the tongue. The most anterior receiver coil T1 is 1 cm posterior to the apex and the most posterior receiver coil T3 is 6 cm posterior to the apex. One more coil T2 is between these. Movement was elicited by presenting the target words of Table 1 in the carrier phrase /ʒibi ___ hnaja/ on a monitor. The number of

repetitions of each stimulus is given in Table 1.

3. METHODS

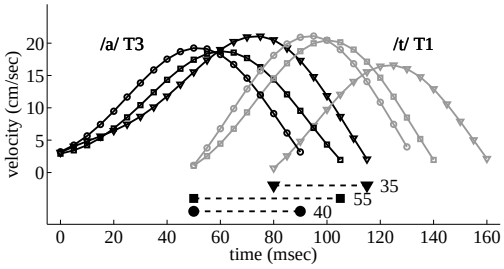
3.1. Segmentation

We identify the /a/ using the T3 coil signal and the subsequent coronal constriction using the T1 coil signal using the following procedure. Coil signals are first filtered by applying a low-pass sixth order Butterworth filter with a cutoff frequency of 20 Hz. The /a/ and the coronal constriction are segmented by delimiting the left and right edges of an interval I over which the movement trace γ approaches the /a/ target and the coronal target, respectively. To find I, we proceed as follows. For each token, we find the principal component $\mathbf{u}$ of γ over an 80 msec interval around the time of peak velocity of the movement. We project γ down onto $\mathbf{u}$ to obtain $\text{proj}_{\mathbf{u}}(\gamma)$. By projecting each γ down onto some $\mathbf{u}$, we simplify a two dimensional time series to one dimension in a nonarbitrary way which captures nearly all variability over the 80 msec frame (minimum eigenvalue ratio across all movement traces is 7.3). The derivative $\frac{\mathrm{d}}{\mathrm{d}t}\text{proj}_{\mathbf{u}}(\gamma)$ of $\text{proj}_{\mathbf{u}}(\gamma)$ is estimated with central differences. The maximum of $|\frac{\mathrm{d}}{\mathrm{d}t}\text{proj}_{\mathbf{u}}(\gamma)|$ is the peak velocity v_p of γ along $\mathbf{u}$. For each token we find the quantity $0.2v_p$. The interval I is delimited on the left and right by the time at which $|\frac{\mathrm{d}}{\mathrm{d}t}\text{proj}_{\mathbf{u}}(\gamma)|$ rises above and falls below, respectively, the threshold $0.2v_p$.

3.2. Temporal overlap

Temporal overlap of /a/ with the subsequent coronal constriction is quantified as the arithmetic difference, in msec, between the right edge of the T3 velocity profile over /a/ and the left edge of the T1 velocity profile over the coronal constriction. Positive overlap means that /a/ and the coronal constriction are unfolding simultaneously on T3 and T1, respectively. Figure 1 illustrates this measure, which is the predictor in the functional model of Section 3.5.

Figure 1: Temporal overlap of three selected /gat/ trials (triangle, square, circle) for S4 plotted with the velocity profiles in absolute value.



3.3. Velocity profile shape

We choose as our dependent measure the tongue body velocity profile during /a/ because, in conformity with the motor control literature, measures of velocity profile shape reflect the evolution of a movement in time and bear readily on dynamical theories of movement [15, 17, 18]. One measure of velocity profile shape is skewness. We fit a seventh degree polynomial function $f(t)$ to each /a/ velocity profile $|\frac{\mathrm{d}}{\mathrm{d}t}\text{proj}_{\mathbf{u}}(\gamma(t))|$, $t \in I$. The skewness S of f is the third standardized moment

$$(1) \quad S = \mathrm{E}\left[\left(\frac{f-\mu}{\sigma}\right)^3\right],$$

where μ is the mean of f, σ is the standard deviation of f, and E is the expectation operator. Positive S means a long right tail and negative S means a long left tail. The variable S is similar to proportional time to peak velocity [2, 3, 4], but captures subtle differences in the global shape of the velocity profile which remain unexpressed by the time to peak velocity measure.

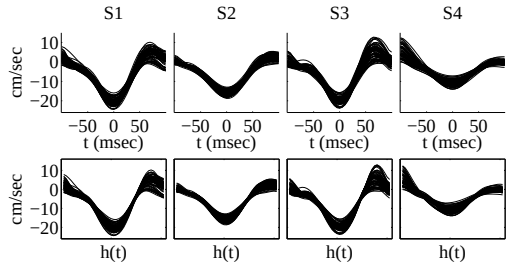
3.4. Time-warping of velocity profiles

The velocity of an EMA receiver coil at time t msec after the estimated onset of movement varies trial by trial. However, the time intervals over which movements unfold also vary in length. Thus, trial by trial variability in velocity at time t is partially due to observing the unfolding movements at different points in their evolution.

Functional data analysis offers tools for dealing with this variability [9, 10, 20, 23]. A key to these tools is a transformation in time applied to each velocity profile via time-warping, also known as registration, functions. Registration transforms the time parameter t of the i^{th} profile to $h_i(t)$ so that qualitative features of the different profiles line up. By transforming the time parameter of each profile, we observe trial-level velocity variability at each transformed time point which is due to genuine velocity differences at time $h_i(t)$.

In our data, we register each /a/ velocity profile over a 200 msec interval centered on the time of peak velocity. A continuous, iterative optimization procedure [19] finds the transformation $h_i(t)$ of time t for the i^{th} /a/ velocity profile to yield $\frac{\mathrm{d}}{\mathrm{d}t}\text{proj}_{\mathbf{u_i}}(\gamma_i(h_i(t)))$, shown in Figure 2 for /a/. Velocity is negative because $\text{proj}_{\mathbf{u_i}}(\gamma_i(h_i(t)))$ decreases as /a/ unfolds. The /a/ begins and ends as the velocity profile falls below and rises back above zero, respectively.

Figure 2: Unregistered (top) and registered (bottom) /a/ velocity profiles over a 200 msec frame around the time of the vowel's peak velocity.



3.5. Functional mixed model

The registered /a/ velocity profile becomes the dependent variable of a functional mixed effects model [14]. The model expresses how the /a/ velocity profile depends on the temporal overlap of /a/ with the subsequent coronal constriction. Specifically, the model allows inference on how the shape of the velocity profile, which is a function of time, depends on temporal overlap, a continuous independent variable. The model specifies a likelihood function $\mathcal{L}(\theta|\mathbf{Y})$, which indicates how the matrix $\mathbf{Y}$, whose i^{th} row is the velocity profile $\mathbf{y}_i = \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{proj}_{\mathbf{u}}(\gamma_i(h_i(t)))$, is generated by the model parameters θ. The likelihood function $\mathcal{L}$ specifies that $\mathbf{y}_{ij}$, which is produced by subject j, is

(2)
$$\mathbf{y}_{ij} = \beta_0(t)x_{0i}+\beta_1(t)x_{1i}+\ldots+\beta_4(t)x_{4i}+u_j(t)+e_i(t).$$

The predictor x_{0i} is 1 for all i; $\beta_0(t)$ is the intercept. The predictor x_{1i} is the temporal overlap in msec of /a/ with the subsequent coronal constriction; $\beta_1(t)$ is the fixed effect for temporal overlap. The variable x_{2i} is the sum coded categorical predictor for whether /k/ ($x_{2i} = +0.5$) or /g/ ($x_{2i} = -0.5$) precedes /a/ in γ_i. The variables x_{3i} and x_{4i} code a three level treatment contrast between /l/, /t/, and /ʃ/ ($x_3 = 1$ if /t/ follows /a/, else $x_3 = 0$; $x_4 = 1$ if /ʃ/ follows /a/; else $x_4 = 0$). The by subject random intercept $u_j(t)$ and the residual error $e_i(t)$ are realizations from Gaussian processes. These predictors code factors which give the model the expressive power to characterize the manipulated variables.

A Markov Chain Monte Carlo (MCMC) procedure samples from the posterior probability distribution $p(\theta|\mathbf{Y})$ of the model parameters θ, given the observed velocity profiles $\mathbf{Y}$ and a prior $p(\theta)$. See [7], Chapter 11 for an overview of MCMC. Our model is the implementation of [13]. The posterior sample size is 10,000.

4. RESULTS

Velocity profile shapes, our main observables, reflect the evolution of a movement in time. Figure 3 plots /a/ velocity profile skewness S against the temporal overlap of /a/ with the subsequent coronal constriction. The figure includes /kat/, /gat/, /kaʃa/, /galebna/, /galba/, and /kamʃa/. Temporal overlap and S are weakly positively correlated, $r(264) = 0.25$, $p = 5.3 \times 10^{-5}$. S does not in turn correlate with the amplitude of /a/ movement along the principal component of T3, $r(264) = -0.006$, $p = 0.92$. See Figure 4. Therefore, temporal overlap specifically affects some measures like S but does not affect others like amplitude.

Figure 5 is the sampling distribution of /a/ velocity profile skewness for /kam/. The coronal constriction for /kam/ trials is in the subsequent word /hnaja/ and does not overlap at all with the /a/ constriction. We do not further analyze /kam/.

We now turn to the posterior distribution of the fixed effects functions $\beta_0(t), \ldots, \beta_4(t)$ of Equation 2. The functional model generates each velocity profile as the sum of a fixed part and a random part. The fixed part of each velocity profile is a linear combination of the functions $\beta_0(t), \ldots, \beta_4(t)$. We interpret the mean, 5 percentile, and 95 percentile of the marginal posterior distribution of these functions. The 5 and 95 percentiles of the marginal posterior distribution indicate for each time t an interval over which ninety percent of the posterior probability density lies. The marginal posterior mean is an estimator of $\beta_0(t), \ldots, \beta_4(t)$.

Figure 6 shows summary statistics of the posterior distribution of the fixed effects intercept $\beta_0(t)$ and the fixed effect $\beta_1(t)$ for temporal overlap. The intercept $\beta_0(t)$ captures roughly the shape of the velocity profile. The effect $\beta_1(t)$ for temporal overlap is near zero before the time of peak velocity. At the time of peak velocity, $\beta_1(t)$ is positive. This means that peak velocity diminishes as overlap increases because the linear combination of $\beta_0(t)$ and $\beta_1(t)$ is closer to zero around the time of peak velocity than $\beta_0(t)$ alone (cf. the velocity minima of Figure 8). After the time of peak velocity, in the right tail of the velocity profile, $\beta_1(t)$ is negative. This means that the right tail becomes fatter as temporal overlap increases because the linear combination of $\beta_0(t)$ and $\beta_1(t)$ has a fatter right tail than $\beta_0(t)$ alone (cf. the right tails of Figure 8). Thus, we infer that an increase in overlap of /a/ with a subsequent coronal has a slow-down effect around the time of peak velocity and a speed-up effect afterward.

Figure 7 shows summary statistics of the poste-

Figure 3: /a/ velocity profile skewness against temporal overlap of /a/ with the subsequent coronal (plus jitter).

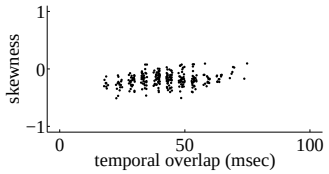


Figure 4: /a/ amplitude against temporal overlap of /a/ with the subsequent coronal (plus jitter).

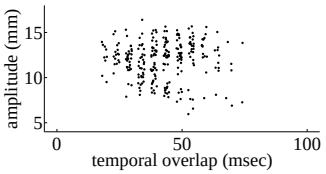


Figure 5: Sampling distribution of /a/ velocity profile skewness for /kam/.

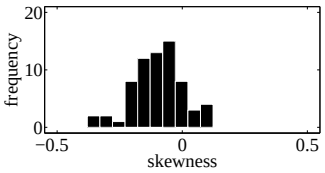


Figure 6: Posterior mean (solid) and 5 and 95 percentiles (dashed) of the fixed intercept $\beta_0(t)$ (left) and fixed effect $\beta_1(t)$ (right).

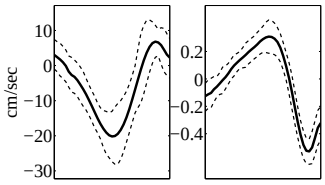


Figure 7: Posterior mean (solid) and 5 and 95 percentiles (dashed) of fixed effects $\beta_2(t)$ (left), $\beta_3(t)$ (middle), and $\beta_4(t)$ (right).

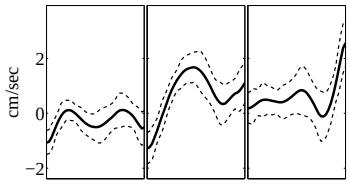
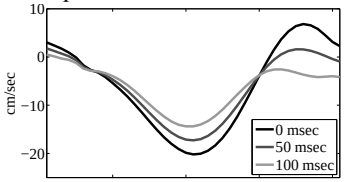


Figure 8: Functional model predictions for /a/ velocity profile shape. Temporal overlap with a subsequent coronal varies.



rior distribution of the fixed effects functions $\beta_2(t)$, $\beta_3(t)$, and $\beta_4(t)$. The fixed effect $\beta_2(t)$ for the voicing of the preceding velar is near zero except on the left tail and around the time of peak velocity, where $\beta_2(t)$ is negative. This means that the magnitude of /a/ velocity is greater after /k/ ($x_2 = +0.5$) than after /g/ ($x_2 = -0.5$) on the left tail and around the time of peak velocity. The functions $\beta_3(t)$ and $\beta_4(t)$ code a fixed effect for the identity of the subsequent coronal (baseline: /l/; $\beta_3(t)$ treatment: /t/; $\beta_4(t)$ treatment: /ʃ/). The function $\beta_3(t)$ is negative on the left tail and positive around the time of peak velocity. This means that the magnitude of /a/ velocity is greater over the left tail and smaller over the peak velocity before /t/ than before the baseline /l/. The function $\beta_4(t)$ is near zero until the far right tail, where $\beta_4(t)$ spikes. This means that the magnitude of /a/ velocity is smaller over the far right tail before /ʃ/ than before the baseline /l/.

In sum, we detect in our data influences of a subsequent coronal consonant on the vowel velocity profile as a function of vowel-consonant overlap.

5. DISCUSSION

The results indicate that a subsequent coronal affects the global velocity profile shape of /a/ measured six centimeters posterior to the apex from start to end of the to-phase of /a/. More specifically, the /a/ velocity profile develops a shallower peak and fatter tail as overlap with this coronal increases. Nevertheless, overlap with the subsequent coronal does not influence /a/ movement amplitude (cf. Figure 4). This is consistent with vowel movement amplitude remaining relatively stable as overlap varies. In contrast, the velocity profile undergoes global change in shape as overlap varies. Figure 8 shows how the functional model predicts that /a/ velocity profile shape changes in response to increasing overlap with the subsequent coronal. As overlap increases, a deep peak becomes shallow and a slender right tail grows fat.

Global changes in velocity profile shape are likewise the response to experimental manipulations in discrete aiming [11] and reciprocal tapping [1] tasks, as well as in saccade eliciting tasks [16]. In the manual domain the velocity profile develops a fatter tail in response to increasing task difficulty (in the sense of Fitts [6]). In the saccadic domain the velocity profile develops a fatter tail in response to a decrease in peak velocity (as the result of fatigue or the drug diazepam). Whether the phenomenon reported here bears more than superficial similarity to these phenomena from other domains of human movement warrants further investigation.

Our results open new empirical ground for modeling of speech movements, namely, to predict that as temporal overlap between a vowel and a subsequent consonant increases, the vowel velocity profile drops in peak velocity and gets a fatter right tail, even as the increasing overlap induces no systematic variability in vowel movement amplitude. Modeling the temporal extent and influence of anticipatory consonant-to-vowel coarticulation will lead to new, quantitatively verifiable hypotheses about coarticulation and the linkages among spatially distal sections of the tongue.

6. REFERENCES

[1] Bootsma, R. J., Fernandez, L., Mottet, D. 2004. Behind Fitts' law: kinematic patterns in goal-directed movements. *International Journal of Human-Computer Studies* 61, 811–821.

[2] Byrd, D., Saltzman, E. 1998. Intragestural dynamics of multiple prosodic boundaries. *Journal of Phonetics* 26, 173–199.

[3] Byrd, D., Saltzman, E. 2003. The elastic phrase: modeling the dynamics of boundary-adjacent lengthening. *Journal of Phonetics* 31, 149–180.

[4] Cho, T. 2006. Manifestation of prosodic structure in articulation: Evidence from lip kinematics in English. In: Goldstein, L., Whalen, D. H., Best, C. T., (eds), *Laboratory Phonology 8*. Berlin/New York: Mouton de Gruyter 519–548.

[5] Farnetani, E., Recasens, D. 1993. Anticipatory consonant-to-vowel coarticulation in the production of VCV sequences in Italian. *Language and Speech* 36(2,3), 279–302.

[6] Fitts, P. 1954. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology* 47, 381–391.

[7] Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., Rubin, D. B. 2014. *Bayesian Data Analysis*. Chapman and Hall/CRC 3 edition.

[8] Kiritani, S., Itoh, K., Hirose, H., Sawashima, M. 1977. Coordination of the consonant and vowel articulations — X-ray microbeam study on Japanese and English. *Annual Bulletin of the Research Institute of Logopedics and Phoniatry* (11), 11–21.

[9] Lucero, J. C., Löfqvist, A. 2005. Measures of articulatory variability in VCV sequences. *Acoustics Research Letters Online* 6, 80–84.

[10] Lucero, J. C., Munhall, V. L., Ramsay, J. O. 1997. On the registration of time and the patterning of speech movement signals. *Journal of Speech, Language, and Hearing Research* 40, 1111–1117.

[11] MacKenzie, C. L., Marteniuk, R. G., Dugas, C., Liske, D., Eickmeier, B. 1987. Three dimensional movement trajectories in Fitts' task: Implications for motor control. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology* 39, 629–647.

[12] MacNeilage, P. F., DeClerk, J. L. 1969. On the motor control of coarticulation in CVC monosyllables. *Journal of the Acoustical Society of America* 45(5), 1217–1233.

[13] Morris, J. S., Carroll, R. J. 2004. Wavelet-based functional mixed models. Technical Report UTMDABTR-006-04 Department of Biostatistics and Applied Mathematics, MD Anderson Cancer Center, University of Texas.

[14] Morris, J. S., Carroll, R. J. 2006. Wavelet-based functional mixed models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 68, 179–199.

[15] Munhall, K. G., Ostry, D., Parush, A. 1985. Characteristics of velocity profiles of speech movements. *Journal of Experimental Psychology: Human Perception and Performance* 11, 457–474.

[16] van Opstal, A. J., van Gisbergen, J. A. 1987. Skewness of saccadic velocity profiles: a unifying parameter for normal and slow saccades. *Vision Research* 27, 731–745.

[17] Ostry, D., Munhall, K. G. 1985. Control of rate and duration of speech movements. *Journal of the Acoustical Society of America* 77, 640–648.

[18] Ostry, D. J., Cooke, J. D., Munhall, K. G. 1987. Velocity curves of human arm and speech movements. *Experimental Brain Research* 68, 37–46.

[19] Ramsay, J., Silverman, B. W. 2005. *Functional Data Analysis*. New York: Springer-Verlag 2 edition.

[20] Ramsay, J. O., Munhall, K., Gracco, V. L., Ostry, D. J. 1996. Functional data analyses of lip motion. *Journal of the Acoustical Society of America* 99, 3718–3727.

[21] Shaw, J., Gafos, A., Hoole, P., Zeroual, C. 2011. Dynamic invariance in the phonetic expression of syllable structure. *Phonology* 28, 455–490.

[22] Stone, M., Lundberg, A. 1996. Three-dimensional tongue surface shapes of English consonants and vowels. *Journal of the Acoustical Society of America* 99, 3728–3737.

[23] West, P. 1999. The extent of coarticulation: an acoustic and articulatory study. Ohala, J. J., Hasegawa, Y., Ohala, M., Granville, D., Bailey, A. C., (eds), *Proceedings of the 14th International Congress of Phonetic Sciences* volume 3 San Francisco: University of California. Linguistics Department, University of California 1901–1904.