JSLHR

Research Note

A Modeling Study of the Effects of Vocal Tract Movement Duration and Magnitude on the F2 Trajectory in CV Words

Kimberly D. Neely,[a] Kate Bunton,[a] and Brad H. Story[a]

Purpose: This study used a computational vocal tract model to investigate the relationship of diphthong duration and vocal tract movement magnitude to measures of the F2 trajectory in CV words.
Method: Three words (*bough*, *boy*, and *buy*) were simulated on the basis of an adult female vocal tract model, in which the model parameters were estimated from audio recordings of a female talker. Model parameters were then modified to generate 35 simulations of each word corresponding to 7 different durations and 5 movement magnitude settings. In addition, these simulations were repeated with vocal tract lengths representative of an adult male and an approximately 6-year-old child.
Results: On the basis of univariate analysis, measures of frequency predicted changes in magnitude, and temporal measures predicted changes in speaking rate consistent with the hypothesis. The combined effects of duration and magnitude showed that F2 was more sensitive to changes in magnitude at shorter word durations compared with longer word durations. This finding held across words and vocal tract length.
Conclusions: Results suggest that there is an interaction between duration and magnitude that affects the slope of the F2 trajectory. The next step is to relate kinematics to F2 trajectory output using real speakers.

The trajectory of the second formant (F2) has been studied as an index of gesture overlap in studies of speaking rate and contrastive stress as well as studies of disordered speech production (cf. Liss & Weismer, 1994; Tjaden & Weismer, 1998; Weismer & Berry, 2003; Weismer, Tjaden, & Kent, 1995). F2 trajectories have shown a reasonable correlation with presumed articulatory movement during transitions; however, determining specifically how rate and movement magnitude individually contribute to F2 trajectory measures has been inconclusive (Liss & Weismer, 1994; Tjaden, 1999; Tjaden & Weismer, 1998; Weismer & Berry, 2003).

Coarticulation is thought to be reflected in measures of the F2 transition and thus can be a useful measure for studying speech production in populations with speech disorders. Weismer, Kent, Hodge, and Martin (1988) devised a definition to index the beginning and end of F2 transitions by marking F2 transition onsets (points in time at which a 20-ms segment occurs in conjunction with a 20-Hz minimal change) and transition offsets (points in time at which a 20-ms segment does not co-occur with a change of 20 Hz or more), allowing investigators to obtain a measure of slope. Steeper F2 slope and F2 onset values that are close to the steady-state target vowel frequency are considered indicators of articulatory overlap. For example, in vowel F2 trajectories that increase in frequency following a consonant release, higher F2 onset values and steeper F2 slope values are believed to suggest more overlap between neighboring consonant–vowel gestures.

Several studies have investigated the effect of speaking rate changes on F2 onset and F2 slope produced by typical adult speakers; for example, Tjaden and Weismer (1998) and Weismer and Berry (2003). Both studies found higher F2 onset frequencies and steeper F2 slope measures for the faster speaking rate condition, suggesting more consonant–vowel movement overlap than in slower speaking conditions (Tjaden & Weismer, 1998; Weismer & Berry, 2003). However, high individual variance of F2 measures across participants in both studies makes interpretation of the specific contributions of the articulatory timing to changes in F2 measures difficult to determine. It is also likely that changes in speaking rate are accompanied by both changes in timing and magnitude of movement.

---

[a] University of Arizona, Tucson

Correspondence to Brad H. Story: bstory@email.arizona.edu

Editor: Jody Kreiman
Associate Editor: Julie Liss

Received November 25, 2014
Revision received September 11, 2015
Accepted February 19, 2016
DOI: 10.1044/2016_JSLHR-S-14-0331

Disclosure: *The authors have declared that no competing interests existed at the time of publication.*

In an effort to determine the effects of the magnitude of articulatory movement on the F2 trajectory, stress manipulation tasks were used with both typical speakers and speakers with neuromotor disorders (Liss & Weismer, 1994; Tjaden, 1999). These studies used contrastive stress tasks to elicit magnitude changes of articulatory movement under tightly defined conditions for both typical speakers (Tjaden, 1999) and those with apraxia of speech and ataxic dysarthria (Liss & Weismer, 1994). Results suggested that the speakers with ataxic dysarthria relied on segmental durations to produce stress, whereas speakers with apraxia altered movement magnitude as well as movement speed as evidenced by larger F2 frequency transitions and longer segment durations (Liss & Weismer, 1994). Tjaden (1999) reported a lower F2 onset frequency during stressed consonant–vowel sequences for words with an upward vowel F2 trajectory (e.g., /ɔɪ/, /aɪ/). This implied a reliance on greater articulatory movement and reduced gestural overlap during stressed syllable production. There was large variance across speakers in both studies, and thus, the exact contribution of articulatory movement and duration changes could not be determined.

X-ray microbeam and electromagnetic articulography data have been used to determine the relationship of F2 measures to vocal tract kinematics (Rong, Loucks, Kim, & Hasegawa-Johnson, 2012; Weismer, Yunusova, & Westbury, 2003). All kinematic data on both typical speakers (Weismer, Yunusova, & Westbury, 2003) and speakers with disorders (Rong et al., 2012) showed reasonable correlations between F2 slope measures and measures of articulatory movement and timing within each subject. However, high variability between subjects was reported in all three studies.

On the basis of published work, it is reasonable to consider F2 slope as an acoustic index of articulatory gesture overlap and that timing and magnitude of movement have an effect. What is not known is how these variables contribute independently to F2 trajectory changes. Changes in articulatory movement and timing are not entirely separate parameters during speech production (Weismer, 1990), making investigations of the specific contributions to F2 trajectory changes challenging. Thus, a model of speech production that allows for direct and independent manipulation of parameters related to articulatory movement and timing is appropriate for examining the individual effects of each variable.

The purpose of the present study was to use a computational model of speech production to determine how changes to timing and magnitude affect the F2 trajectory. We hypothesized that F2 frequency measures would reflect changes in the magnitude of vocal tract movement and temporal measures would reflect duration changes. Measures of F2 slope with both frequency and temporal information should predict changes in both dimensions. Although preliminary, the goal of this work was to lead to more accurate inferences regarding real speakers' coarticulatory strategies on the basis of measures of the F2 trajectories.

Method

This study was based on simulating time-varying configurations of the vocal tract that produce the words *bough*, *boy*, and *buy*. These words were chosen so that they included a voiced bilabial stop followed by a diphthong, hence providing test cases in which F2 traverses a relatively broad frequency range.

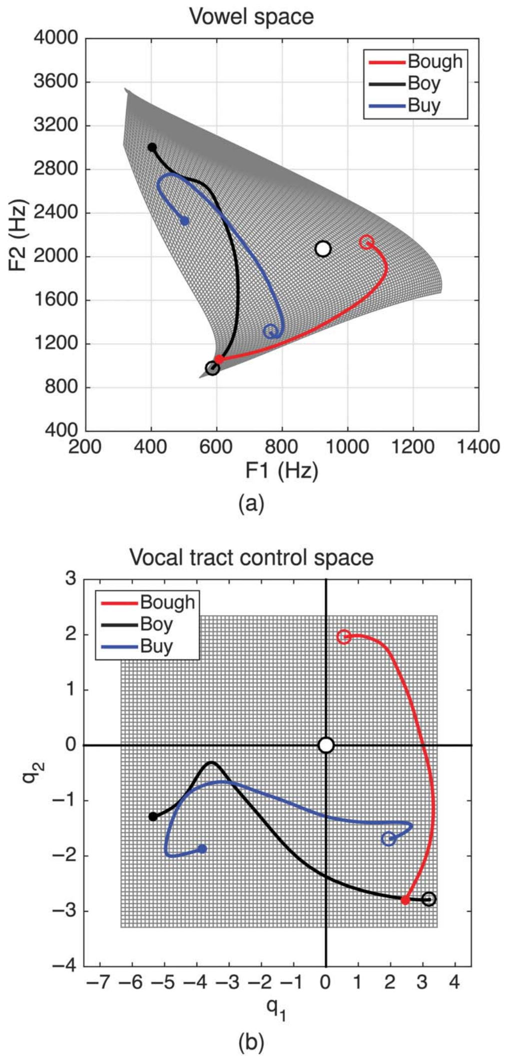
The first step was to generate a baseline simulation that emulated the formant characteristics of natural speech versions of each word. Audio recordings were obtained from a 30-year-old female who was a native, monolingual English speaker of Southwest dialect. She passed a hearing screening and had no history of neurological impairment. Words were read from a list and repeated three times. One instance of each word was chosen on the basis of the clarity of the formants revealed in a wide-band spectrogram. F1 and F2 were then tracked over the time course of the chosen words using a standard linear predictive coding method. Resulting (F1, F2) trajectories of the diphthong portions of *bough*, *boy*, and *buy*, which were spoken as /aʊ/, /ɔɪ/, and /aɪ/, are shown in Figure 1a as the black, red, and blue lines, respectively.

The next step was to map the (F1, F2) trajectories of each word to the time-dependent parameters of the vocal tract model described by Story (2005a):

$$\begin{aligned} V(x,t) &= (\pi/4)[\Omega(x) + q_1(t)\varphi_1(x) + q_2(t)\varphi_2(x)]^2 \\ x &= \Delta[1, N_{vt}] \end{aligned} \qquad (1)$$

where the vocal tract configuration at any given time instant, t, is defined by the variation in cross-sectional area as a function of x, the distance along the curved axis extending from glottis to lips. The variable $\Omega(x)$ is the mean diameter function, and $\varphi_1(x)$ and $\varphi_2(x)$ are two shaping functions, called modes, all of which are dependent on x. For the present study, the $\Omega(x)$, $\varphi_1(x)$, and $\varphi_2(x)$ were based on those reported in Story (2005b) for the female participant denoted as SF1. The scaling coefficients, $q_1(t)$ and $q_2(t)$, are functions of time and control the temporal variation of vocal tract shape from vowel to vowel. The area function at any instant of time is composed of $N_{vt} = 44$ discrete "tubelet" sections (cf. Story, 2005a), each with a length of $\Delta = 0.32$ cm, giving a total tract length of 14.1 cm. The temporal discretization for the model consists of a sampling interval equal to 6.866 ms or 146 Hz. It has been previously shown (cf., Story, 2005b; Story & Titze, 1998) that a given pair of (q_1, q_2) values in Equation 1, defining the vocal tract shape, corresponds to a specific set of (F1, F2) values in the vowel space plot, thus producing a one-to-one mapping. The gray mesh in Figure 1a is a set of 6,400 pairs of (F1, F2) values resulting from the 6,400 (q_1, q_2) pairs shown as a similar mesh in Figure 1b; the white dot within each mesh indicates the formant and coefficient pairs that correspond to the neutral area function condition (i.e., when $q_1 = q_2 = 0$). This mapping was then used to transform the formant pairs along the vowel space trajectories for each of the three words into time-varying coefficient values, $q_1(t)$ and $q_2(t)$, that,

Figure 1. Mapping of (F1, F2) vowel space to vocal tract control space. (a) Baseline (F1, F2) trajectories for *bough*, *boy*, and *buy* superimposed on a formant mesh on the basis of the female vocal tract model (Equation 1). (b) The (q_1, q_2) coefficient trajectories for the baseline cases resulting from mapping the formant trajectories in (a) to the underlying coefficient mesh. For all trajectories, the open circle denotes the beginning of the diphthong.



(a)

(b)

when used in Equation 1, will produce the changes in vocal tract configuration necessary to simulate the diphthong portion of each word.

Localized constrictions, characteristic of obstruent consonants, can be imposed on any given vocal tract shape with an additional scaling function $C(x)$ (Story, 2005a). This function is equal to 1.0 along the entire vocal tract length except near a desired constriction location l_c, where it drops to zero to produce an occlusion. A consonant magnitude function $m_c(t)$ makes the scaling function time dependent, $C(x, t)$, by allowing for gradual onset and offset. When $m_c(t) = 0$, the constriction function has no effect, but when equal to 1.0, the constriction is fully realized within the area function. The composite area function $A(x, t)$ is the product of each of the N_{vt} elements along the x dimension of $V(x, t)$ and $C(x, t)$ so that at any time t,

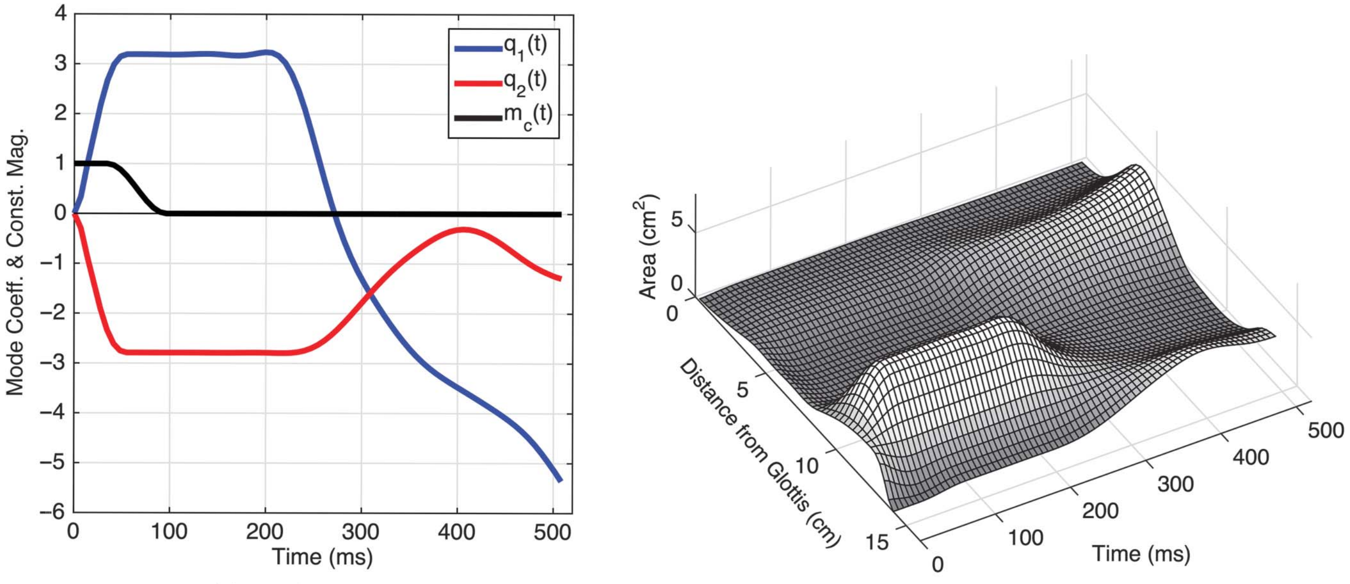
$$A(x,t) = V(x,t)C(x,t). \tag{2}$$

The steps required to produce a simulation are demonstrated in Figure 2 for the word *boy*. Plotted in the upper left panel (Figure 2a) are the $q_1(t)$ and $q_2(t)$ coefficients that, from 100 to 510 ms, were taken directly from the one-to-one mapping for *boy* shown previously in Figure 1b. The initial 100 ms for both q_1 and q_2 were generated by a smooth interpolation from a value of zero to the initial value of the diphthong trajectory. The consonant magnitude function $m_c(t)$ is also plotted and shows that the lips remain in contact for 50 ms and then are released. Using these functions in Equations 1 and 2 produces the time-varying area function $A(x, t)$ shown in Figure 1b; the initial bilabial occlusion can be seen in the lower left portion of the plot, which releases at about 50 ms; the area function then moves through the /ɔɪ/ diphthong.

The frequency response of the area function at each consecutive time sample was calculated with a transmission line algorithm that included energy losses due to yielding walls, viscosity, heat conduction, and acoustic radiation at the lips (Sondhi & Schroeter, 1987; Story, Titze, & Hoffman, 2001). The first two formants F1 and F2 of each frequency response were determined by a parabolic peak-picking technique. To maintain continuity with the format of the plots in Figures 2a and 2b, the time-varying frequency response $H(f, t)$ is shown in Figure 2d, whereas the F1 and F2 tracks are given in Figure 2c. For this particular simulation (and also for the baseline simulations of *bough* and *buy*, which are not shown), the logic may appear circular; that is, formant contours were measured from audio recordings and then mapped to the model coefficients that were used in Equation 1 to generate a time-dependent area function, from which time-dependent formant frequencies were finally obtained. The goal, however, is that the baseline simulations provide a means of generating an initial set of vocal tract parameters that can be systematically modified to relate a representation of movement duration and magnitude to the temporal characteristics of F2.

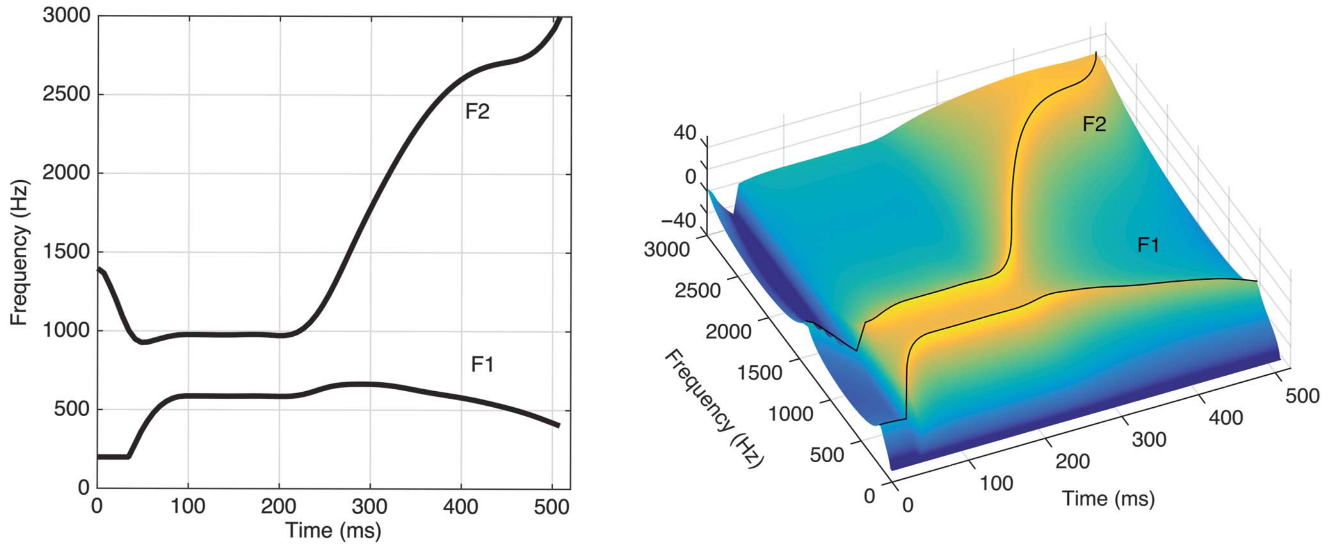
Following from previous literature, additional simulations were generated for each word at effective speech rates that were both faster and slower than the baseline case. In particular, the durations of $q_1(t)$, $q_2(t)$, and $m_c(t)$ were set to be 0.5, 0.625, 0.75, 0.875, 1.0, 1.125, and 1.25 of the baseline duration of 510 ms. This resulted in word durations of 260, 320, 380, 450, 510, 570, and 640 ms, which are similar to the durations produced by talkers in a graded speech rate task reported by Weismer and Berry (2003). Additional conditions were simulated for each word in which the magnitude of the vocal tract shape deformation for the diphthong portion was modified by multiplying the baseline $q_1(t)$ and $q_2(t)$ coefficient trajectories by factors of 1.125, 1.0, 0.875, 0.75, and 0.5. The reason that only one of the factors increases the magnitude is that any

Figure 2. Simulation of the word *boy* on the basis of the female vocal tract model. (a) Time-varying coefficients $q_1(t)$ (blue) and $q_2(t)$ (red) along with the consonant magnitude function $m_c(t)$ (black). (b) Time-varying area function A(x, t) generated with Equations 1 and 2. (c) Time-variation of the formants F1 and F2 produced by the area function in (b). (d) Time-varying frequency response from which F1 and F2 were determined.



(a) $q_1(t)$, $q_2(t)$, & $m_c(t)$

(b) $A(x, t)$



(c) $F1(t)$ & $F2(t)$

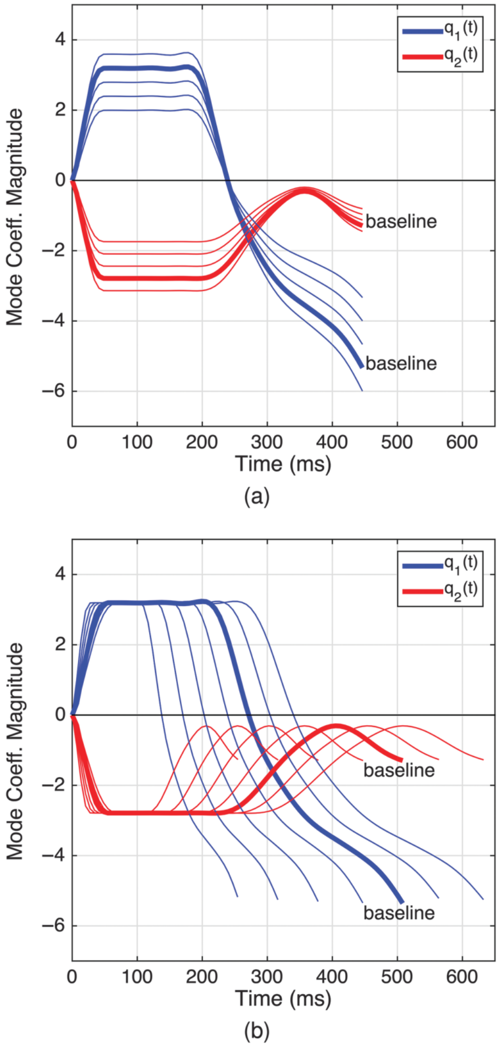
(d) $H(f, t)$

magnitude beyond this value would generate an occlusion in the vocal tract. In addition, the smallest factor was limited to 0.5 to avoid overly centralized vocal tract configurations that would produce little variation in F2. All combinations of the duration and magnitude settings were used so that each word was simulated 35 times.

An example set of time-varying coefficient values for the word *boy* is plotted in Figure 3. The upper panel displays the baseline, reduced, and amplified cases, each with a duration of 450 ms. The blue lines show the time-varying amplitudes of the first mode coefficient $q_1(t)$, whereas the red lines represent the second coefficient $q_2(t)$; the baseline versions of the first and second mode amplitudes are indicated by the thick blue or red lines, respectively. Plotted in the lower panel of Figure 3 are mode amplitudes for the baseline magnitude setting but with varying durations. Again, the baseline cases are indicated with thick blue and red lines.

All 35 versions of each of the three words were simulated two additional times when the overall vocal tract

Figure 3. Example set of time-varying mode amplitude (coefficients) values for the /ɔɪ/ diphthong in *boy*. (a) Baseline, reduced, and amplified cases, each with a duration of 450 ms. The blue lines show the time-varying amplitudes of the first mode, whereas the red lines represent the second mode; the baseline versions of the first and second mode amplitudes are indicated by the thick blue or red lines, respectively. (b) Mode amplitudes for the baseline magnitude setting but with varying durations. The cases with baseline timing are indicated with thick blue and red lines.



length was set to 18.5 cm and 11 cm to reflect the overall size of a large adult male talker and an approximately 6-year-old child, respectively. In these additional cases, the mean vocal tract, mode shapes, and variations of the mode amplitudes in both magnitude and duration were exactly the same as in the female version; the only difference was total vocal tract length. There were a total of 315 simulated words.

Formant Analysis

Analyses were performed directly on the diphthong portion of the time-varying F2 contour produced for each word by the model (e.g., Figure 2c). Each F2 contour was resampled prior to analysis to have a sampling interval of 5 ms to be consistent with previous studies (Tjaden, 1999; Tjaden & Weismer, 1998; Weismer & Berry, 2003). Examples of the F2 measurements are shown in Figure 4 for the diphthong portions of *bough*, *boy*, and *buy* produced with the female vocal tract when the duration and magnitude conditions were 0.875 and 1.0, respectively. Temporal variations of F2 for each word are shown in the top row, and their time derivatives are in the bottom row. The F2 onset was measured as the frequency of the first time sample following the consonant release (Tjaden, 1999), as indicated by the leftmost black dot on each plot. The F2 transition frequency was defined as the F2 value within the diphthong that occurs at a time instant at which the F2 derivative has an absolute value of at least 1 Hz/ms (denoted by the blue dot in each plot); this value corresponds to the target F2 defined by Weismer and Berry (2003) for the /ɔɪ/ diphthong (note that they state the criterion as 20 Hz/20 ms). The F2 slope was measured directly from the F2 derivative as the local maxima or minima following the F2 transition and is shown as the red dot in each plot.

Temporal measures were also calculated from the F2 and F2 derivative plots. The first measure, called Δ_{T1}, is the time interval between the F2 onset and the F2 transition (Δ_{T1} is the same as the Δ_{Trans} measure reported in Tjaden, 1999). A second measure, referred to as Δ_{T2}, is the time interval between the F2 transition point and a later time instant at which the F2 derivative returns to an absolute value of 1 Hz/ms (Tjaden, 1999; Weismer et al., 1995). In cases in which the F2 derivative did not attain this value, Δ_{T2} was measured as the interval between the F2 transition point to the end of the trajectory. In either case, these points are shown as the rightmost black dots in each plot of Figure 4.

Measurement of F2 onset, F2 transition, F2 slope, Δ_{T1}, and Δ_{T2} was automated with a script written in MATLAB (Mathworks, Natick, MA). The result was a table of values for the 35 combinations of duration and magnitude for each word and vocal tract condition (i.e., female, male, and child).

Statistical Analysis

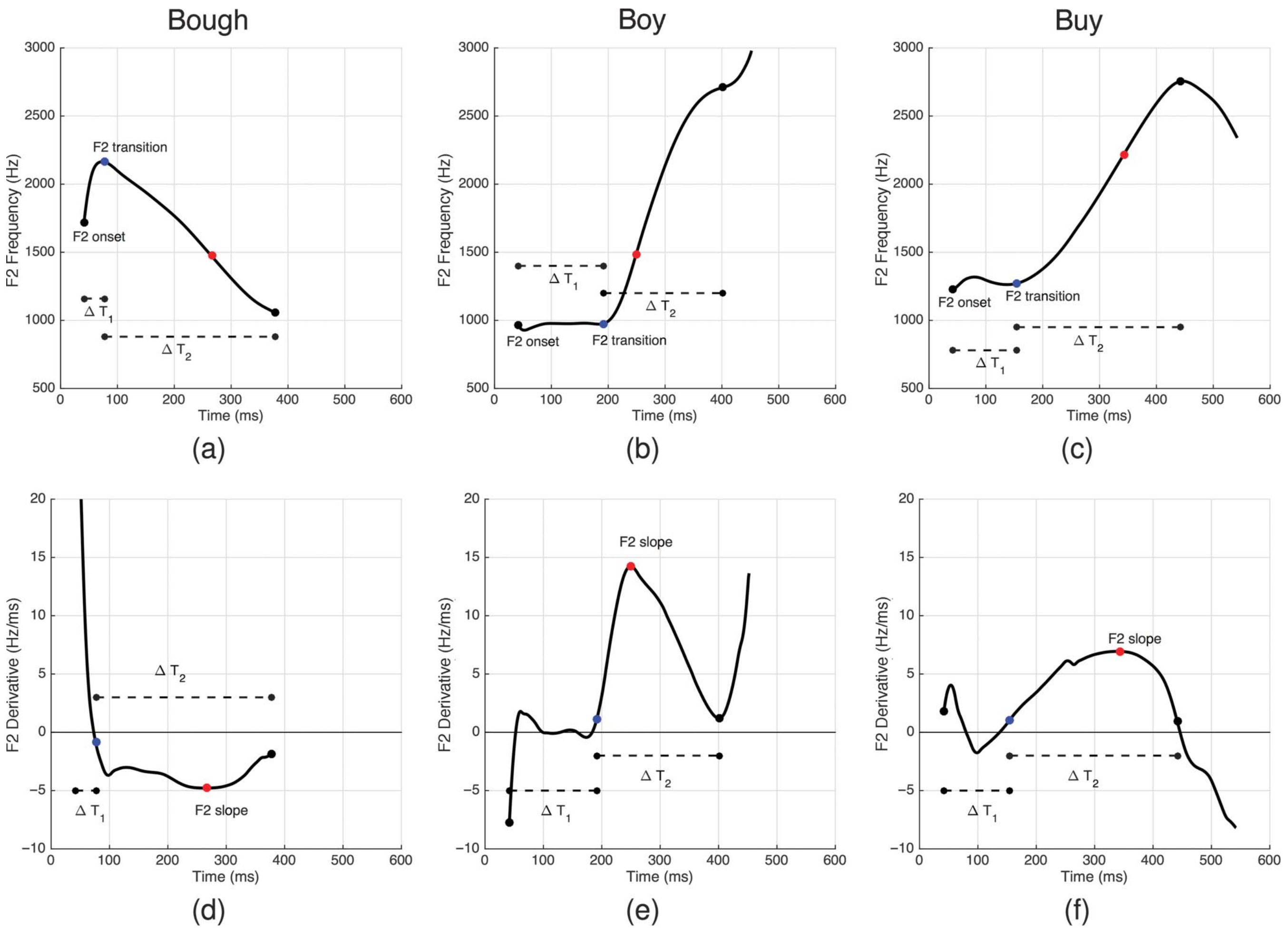
The set of 35 simulations generated for each diphthong with three different vocal tract lengths was treated as a separate experimental unit. The relationship between the dependent variables of timing (n = 7) and magnitude (n = 4) and the independent acoustic formant measures (n = 5; F2 onset, F2 transition, F2 slope, Δ_{T1}, and Δ_{T2}) were quantified with univariate linear regression analyses performed in MATLAB. All statistical analyses used a conservative alpha on the basis of the Bonferroni procedure.

Results

Duration

The effect of duration varied by diphthong but was consistent across vocal tract length. For the word *bough*,

Figure 4. Demonstration of acoustic measures. The top row (a–c) shows the time course of F2 for the diphthong portion of each word, whereas the bottom row (d–f) shows the F2 time derivatives. The measurements described in the text are marked on the appropriate plots.



Δ_{T2} was the most significant predictor ($r^2 \approx 99\%$; $p < .001$), followed by F2 slope ($r^2 \approx 70\%$; $p < .001$). The most significant predictors of variance for the words *boy* and *buy* were Δ_{T1} ($r^2 \approx 99\%$; $p < .001$) and Δ_{T2} ($r^2 > 88\%$; $p < .001$) across all three vocal tract lengths. F2 slope was the next most significant predictor variable for these two words ($r^2 \approx 66\%$; $p < .001$ for *boy* and $r^2 \approx 56\%$; $p < .001$ for *buy*).

Magnitude

For *bough*, Δ_{T1} was the most significant predictor variable and accounted for more than 74% of the variance for the male and female vocal tract lengths ($p < .001$) and 70% ($p < .001$) for the child vocal tract. F2 onset ($r^2 \approx 97\%$; $p < .001$) was also significant for *bough*, followed by F2 transition ($r^2 \approx 51\%$; $p < .001$) across all three vocal tract lengths. For the words *boy* and *buy*, F2 transition and F2 onset were significant predictor variables across vocal tract length. F2 transition accounted for approximately 99% of the variance ($p < .001$), and F2 onset accounted for more than 96% of the variance ($p < .001$). For *buy*, F2 slope was also significant and accounted for roughly 36% of the variance ($p < .001$).

Discussion

The results of the present study, consistent with our hypotheses, showed differences in statistically significant predictor variables across target words (i.e., diphthongs) and vocal tract length. In general, measures made on the basis of frequency predicted changes in magnitude of movement (e.g., F2 onset, F2 transition), and temporally based measures were more sensitive to changes in utterance duration (e.g., Δ_{T1} and Δ_{T2}).

Duration

Seven duration conditions were included in the present study (including the baseline). Relative to the baseline utterance, four durations were made shorter to simulate a faster speaking rate and two were made longer to simulate a slower speaking rate. As expected, measures of F2 varied systematically across diphthong and vocal tract length as utterance duration changed. The two temporally based measures included in the analysis, Δ_{T1} and Δ_{T2}, were expected to be most sensitive to changes in duration. Results showed that both Δ_{T1} and Δ_{T2} were highly predictive of changes in duration for *buy* and *boy* for all vocal tract lengths, but only Δ_{T2} was a significant predictor variable for *bough*. This is likely due to the rapid rise of F2 from the onset frequency to the transition frequency in *bough* regardless of duration, thus making Δ_{T1} relatively insensitive to duration changes.

Magnitude

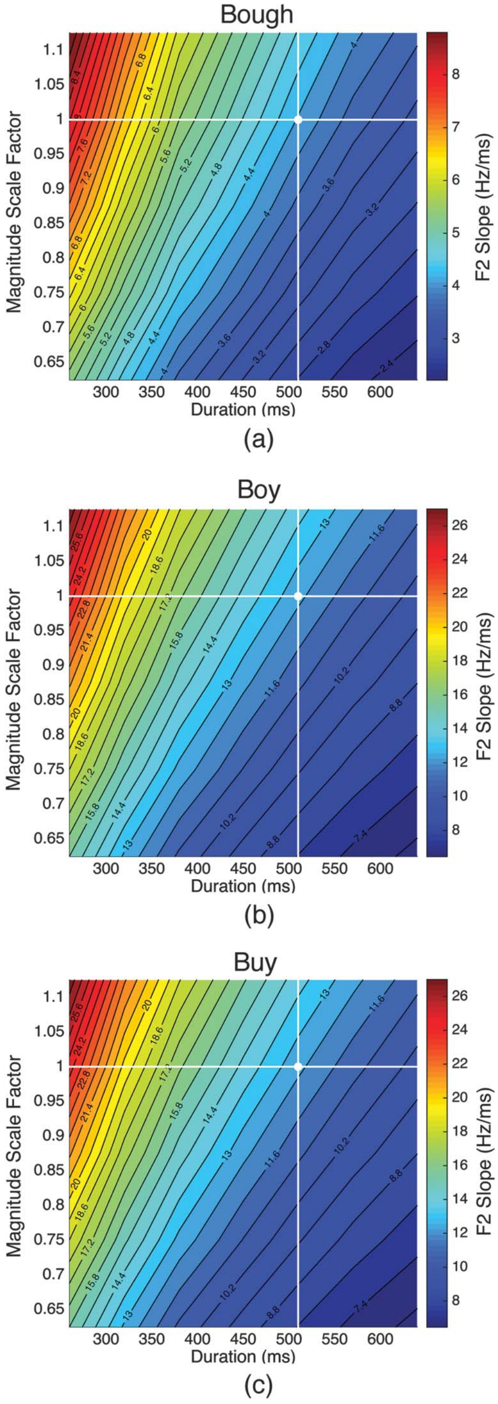
There were five magnitude conditions simulated (including the baseline), three representing a decrease in magnitude and one exceeding the baseline utterance. It was hypothesized that frequency-based measures (F2 onset, F2 transition) would be most sensitive to changes in magnitude across diphthong and vocal tract length. For *boy* and *buy*, the F2 transition frequency was the strongest predictor of change in magnitude for all vocal tract lengths and was the second strongest predictor for *bough*. Weismer and Berry (2003) also found that the F2 transition frequency (called F2tar in their article) explained more variation than F2 onset across speakers, but it did not consistently predict changes in speaking rate. Findings of the present study coincide with their results: The F2 transition frequency was not a predictor of duration (i.e., speaking rate) but was a significant predictor of magnitude, a quantity that Weismer and Berry could not directly measure. Perhaps the variation in response to speaking rate found by Weismer and Berry reflected varying combinations of movement magnitude and articulatory timing produced by real speakers in their experimental rate manipulation task.

Duration and Magnitude

A measure of F2 slope was included in the present study, as it was hypothesized to be sensitive to changes in both duration and magnitude across the diphthong transition, but the degree to which it predicts changes in both domains was unknown. Results of the univariate regression analyses showed that F2 slope was a stronger predictor for duration changes compared with changes in magnitude in all word conditions. This variable was only significant for both duration and magnitude in the word *buy*.

To explore this further, contour plots showing the combined effects of duration and magnitude for the three words were generated for the female vocal tract and are shown in Figure 5. In each plot, the horizontal axis is duration of the word, the vertical axis is magnitude, and the F2 slope is the variable that is coded with the color map. The color bar at the right side of each plot indicates the range of slopes represented, and actual F2 slope values are superimposed on the contour lines. From all three plots, it is

Figure 5. Contour plots showing F2 slope as a function of duration and gesture magnitude for the /ɑʊ/, /ɔɪ/, and /ɑɪ/ portions of *bough*, *boy*, and *buy* made on the basis of the female vocal tract model. Duration is on the *x* axis, and magnitude is on the *y* axis. The color scale shows the range of F2 slope values; red represents the steepest slope; blue represents the most shallow. The F2 slope for the baseline utterance in each case is shown as a white dot. The black lines in each plot represent constant slopes for which values are given by the color map and the key on the right side of each plot. The points along the black lines within the plot represent duration and magnitude pairs that produce the same F2 slope (i.e., isoslope lines).



clear that changes in magnitude have a larger effect on F2 slope when word duration is shorter. Changes in magnitude produce relatively small changes in F2 slope when word duration is longer. For example, words 300 ms in length with varying degrees of magnitude had F2 slopes that covered nearly the entire F2 slope range. At 600 ms, however, F2 slope increased only slightly with increased magnitude. It can be noted that the strongest effect of magnitude was seen for durations shorter than the baseline utterance (white dot). Thus, a talker would apparently have greater ability to affect the degree of F2 slope through control of magnitude at higher, rather than slower, speaking rates. It can also been observed that, although the F2 slope contours for each diphthong appear to be similar, the range of F2 slope values for *boy* (Figure 5a) is nearly twice that of *buy* (Figure 5b), and in a similar manner, the range for *buy* is about twice that of *bough*. The reason is that the "distance" traveled along the frequency axis by F2 is greater in *boy* than the other two diphthongs, and so for a given word duration, the rate of change must also be greater. This result does, however, draw attention to the fact that a given F2 slope may be regarded as a high value in one context but not another; for example, an F2 slope of 8 Hz/ms is at the upper extreme for *bough* but would be considered a low value in *boy*.

Future Directions

This was a preliminary study done on the basis of a kinematic model of the vocal tract shape. As such, studies relating kinematic data and formant measures on the basis of real speakers in a variety of speech tasks are needed to validate findings and further our understanding of how F2 trajectories reflect changes in speech production. Future studies could focus on rate versus magnitude characteristics to determine whether speakers vary their speech similarly depending on the instruction.

Acknowledgments

This work was supported by National Institutes of Health Grant R01-DC011275 (awarded to Brad H. Story) and National Science Foundation Grant BCS-1145011 (awarded to Brad H. Story and Kate Bunton). This was part of the first author's master's thesis.

References

Liss, J., & Weismer, G. (1994). Selected acoustic characteristics of contrastive stress production in control geriatric, apraxic, and ataxic dysarthric speakers. *Clinical Linguistics & Phonetics, 8,* 45–66.

Rong, P., Loucks, T., Kim, H., & Hasegawa-Johnson, M. (2012). Relationship between kinematics, F2 slope, and speech intelligibility in dysarthria due to cerebral palsy. *Clinical Linguistics & Phonetics, 26,* 806–822.

Sondhi, M. M., & Schroeter, J. (1987). A hybrid time-frequency domain articulatory speech synthesizer. *IEEE Transactions on Acoustics, Speech and Signal Processing, 35,* 955–967. doi:10.1109/TASSP.1987.1165240

Story, B. H. (2005a). A parametric model of the vocal tract area function for vowel and consonant simulation. *The Journal of the Acoustical Society of America, 117,* 3231–3254.

Story, B. H. (2005b). Synergistic modes of vocal tract articulation for American English vowels. *The Journal of the Acoustical Society of America, 118,* 3834–3859.

Story, B. H., & Titze, I. R. (1998). Parameterization of vocal tract area functions by empirical orthogonal modes. *Journal of Phonetics, 26,* 223–260.

Story, B. H., Titze, I. R., & Hoffman, E. A. (2001). The relationship of vocal tract shape to three voice qualities. *The Journal of the Acoustical Society of America, 109,* 1651–1667. doi:10.1121/1.1352085

Tjaden, K. (1999). Stress-induced variation in F2 trajectories as evidence for coproduction in CV syllables. *Language and Speech, 42,* 413–432.

Tjaden, K., & Weismer, G. (1998). Speaking-rate-induced variability in F2 trajectories. *Journal of Speech, Language, and Hearing Research, 41,* 976–989.

Weismer, G. (1990). *Assessment of speech and voice production: Research and clinical applications*. Bethesda, MD: The National Institute on Deafness and Other Communication Disorders.

Weismer, G., & Berry, J. (2003). Effects of speaking rate on second formant trajectories of selected vocalic nuclei. *The Journal of the Acoustical Society of America, 113,* 3362–3378.

Weismer, G., Kent, R. D., Hodge, M., & Martin, R. (1988). The acoustic signature for intelligibility test words. *The Journal of the Acoustical Society of America, 84,* 1281–1291.

Weismer, G., Tjaden, K., & Kent, R. D. (1995). Can articulatory behavior in motor speech disorders be accounted for by theories of normal speech production? *Journal of Phonetics, 23,* 149–164.

Weismer, G., Yunusova, Y., & Westbury, J. R. (2003). Interarticulator coordination in dysarthria: An X-ray microbeam study. *Journal of Speech, Language, and Hearing Research, 46,* 1247–1261.