Consonant Dynamics: Rate- and Vowel-Dependence

Claude E. Mauk
University of Texas at Austin
E-mail: claudeed@mail.utexas.edu

ABSTRACT

Many studies describe how vowel formant patterns change as a function of consonant context, stress, speaking tempo and speaking style. Test phrases juxtaposed consonants with high F2 targets with vowels with low or high F2 targets. Strong shifts of F2 targets were intended to encourage variation in F2 onsets. F2 onset values were found to be rate and context dependent, showing that like vowels, consonants behave in a dynamic way, undergoing changes as a function of timing, speaking style and vowel context. A quantitative treatment of the data indicates that F2 onsets can be accurately predicted from information on the F2 midpoint of the preceding vowel, the F2 midpoint of the following vowel, the closure duration and the F2 onset target.

1. INTRODUCTION

Phonetic undershoot is the failure of a phonetic gesture to achieve an articulatory or acoustic target associated with it. The Hypo- & Hyper-articulation (H&H) model of speech production [1, 2] views speech as a motor activity that is subject to the constraint of economy of effort. That is, if there is no motivation to the contrary, speech will tend toward a reduced form, i.e. hypo-articulation. Such constraints may also affect the signed languages of the deaf [3].

Researchers have examined undershoot by comparing fast and normal speech [4, 5] and comparing normal and clear speech [6, 7], often focusing on variation in vowel quality. Moon [7], for example, has found that vowels in regular speech are hypo-articulated relative to vowels in clear speech. Similarly, vowels in fast speech tend to be hypo-articulated relative to their regular speech counterparts (e.g. [4]). This finding agrees with studies of the physical articulation of vowels in fast speech. Electromyographic (EMG) studies [8, 9] have found in that muscle activity during vowel articulation is lower when those vowels are articulated in a fast manner. Not only do vowels display undershoot in certain situations, that undershoot is accompanied by reduced muscular activity.

In contrast, EMG studies of consonants [8, 9] found that fast speech did not involve a reduction of the muscle activities responsible for achieving a consonant's place of articulation. Rather, muscle activity was relatively increased in fast speech over regular speech. This finding may indicate a reorganization of muscle activity on the part of the speaker so that an articulatory target can be achieved. In other words, a greater effort was made to achieve the consonants' places of articulation because those locations had to be achieved in a shorter amount of time.

In a comparison of spontaneous and citation speech, Sussman [10] found that alveolar [d] and palatalized velar [g] show locus equations with slightly higher slopes in spontaneous speech when compared with locus equations for the same places of articulation in citation speech. While this finding indicates that there is increased coarticulation, between consonants and vowels in spontaneous speech, the difference between the locus equations was very small, suggesting that the actual place of articulation of the consonants is relatively unaffected.

Tjaden & Weismer [11] investigated the relationship between F2 onset and speaking rate by studying a range of speaking rates from slow to fast. They found that F2 onsets vary as a consequence of speaking rate. However, they observed that the direction of variation often conflicted with theoretical predictions.

Similar to Tjaden & Weismer, the goal of this report is to provide data on how English stop consonants vary when vowel context, and tempo are manipulated. Test phrases in this study juxtapose consonants with high F2 targets with vowels with low or high F2 targets. Strong shifts of F2 targets were intended to encourage systematic variation in F2 onsets.

2. METHODOLOGY

Two male, FH & MB, and two female, AT & CW, native speakers of American English served as subjects. Four CV combinations were chosen to maximize the acoustic distance between a relatively high consonant F2 target value and a relatively low vowel F2 target ([da], [do], [ga], [gæ]). Each of these C-V combinations was placed into two phrase contexts: a low F2 environment (preceded by [o]) and a high F2 environment (preceded by [i]).

Subjects were asked to say these phrases in four manners: a) normal, relaxed speech, b) speech faster than normal but not so fast that the subject or experimenter felt the speech lost intelligibility, c) speech that was even faster but without any special regard for clarity, and d) speech that was "as fast as possible". These elicitation conditions were intended to produce a range of speaking rates. Subjects were not given any specific feedback to ensure that the conditions be cohesive or discrete. Phrases were randomized with 10 repetitions for each phrase in each

 ISBN 1-876346-48-5

speaking condition.

The data were recorded on a Fostex D-5 Digital Master Recorder (DAT) and then digitized using MacQuirer 5 at a sampling rate of 22,000 Hz. Spectrograms were produced using a 172 Hz filter including formant tracking with a 5 ms step size. For each phrase, durations were measured for stop closure, preceding vowel and following vowel. F2 measurements were taken for the preceding vowel steady state and offset and following vowel steady state and offset. If no steady state was present during a vowel, F2 measurements were taken at the temporal midpoint of that vowel. A combination of visual inspection of the formants and the formant tracking was used to determine measurement values.

3. RESULTS

In phrases where the CV was preceded by a vowel with a low F2 target, i.e. [o], values for the F2 onset of V2 were generally found to decline as speech rate increased. Figure 1 shows the average dynamics of F2 in the sequence [oda] for each elicitation category for one subject. Average F2 values for five points in the phrase are graphed from left to right. The F2 onset value is highest for the normal rate and lowest for the fastest speech rate. The difference between these two values is approximately 350 Hz. Additionally, the F2 midpoint values for V1 and V2 generally rise as rate increases. For V2 midpoint, the difference between the normal and fastest is around 220 Hz. As speaking rate increases, the F2 dynamics of this phrase flatten with vowels and consonant accommodating to each other.

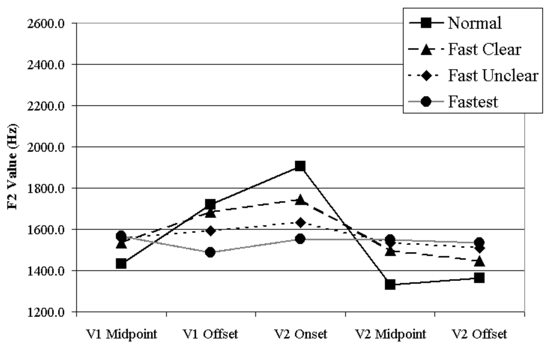

Figure 1: F2 dynamics of [oda] for subject CW

The three faster elicitation conditions did not result in three discrete sets of data, i.e. there tended to be overlap between these conditions with respect to the duration of the spoken phrases. It was considered more useful to analyze the data in terms of consonant closure durations rather than the elicitation conditions. Figure 2 compares F2 onset values with consonant closure durations. The relationship between these two values is positive (r=0.84), that is as durations fall (speaking rate increases), F2 onset values fall.

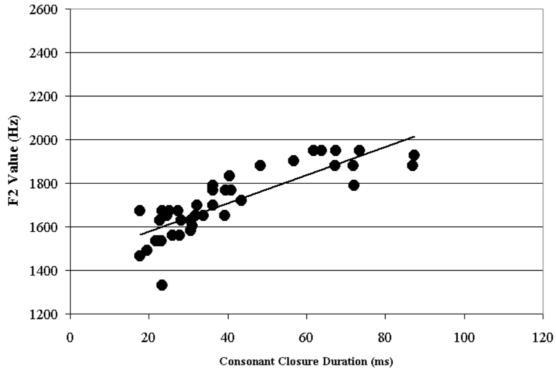

Figure 2: F2 stop closure duration related to V2 F2 onset values for [oda] for subject CW

This pattern was found for [oda], [odo], and [ogæ] for all four subjects. Only one subject, MB, showed this pattern for [oga]. For other subjects there was no compelling relationship between closure duration and F2 onset in [oga]. Table 1 summarizes the relationships between closure duration and F2 onset for each [oCV] sequence for each subject.

S	VCV	r^2	Slope	*As rate increases:*
AT	oga	0.02	0.46	
AT	ogæ	0.31	2.57	lower F2 Onset
AT	oda	0.72	2.51	lower F2 Onset
AT	odo	0.74	2.61	lower F2 Onset
CW	oga	0.00	0.24	
CW	ogæ	0.59	6.31	lower F2 Onset
CW	oda	0.71	6.52	lower F2 Onset
CW	odo	0.64	4.57	lower F2 Onset
FH	oga	0.00	0.05	
FH	ogæ	0.27	3.31	lower F2 Onset
FH	oda	0.51	3.32	lower F2 Onset
FH	odo	0.39	3.28	lower F2 Onset
MB	oga	0.59	7.66	lower F2 Onset
MB	ogæ	0.06	2.11	lower F2 Onset
MB	oda	0.42	3.85	lower F2 Onset
MB	odo	0.38	3.16	lower F2 Onset

Table 1: Relationship between closure duration and F2 onset value for CV sequences preceded by [o].

Out of all of the VCV sets in this study, [oga] would be expected to show the lowest F2 onset, since there is no motivation for the speaker to employ a palatal articulation of [g]. As a result, in the sequence [oga] the contextual pressure on the F2 onset from the neighboring vowels may not have been sufficient to cause consistent variation in the F2 onset value. Instead of variation in the F2 onset, the preceding vowel accommodated to the consonant place of articulation for subjects AT, CW and FH showing a consistent undershoot of the midpoint and offset values of V1 in [oga]. In the case of MB, however, the V1 midpoint and offset values in [oga] were stable across the elicitation conditions. The preceding vowel did not accommodate to the consonant, rather the consonant accommodated to the vowel.

When the CV was preceded by a vowel with a high F2 target, i.e. [i], rate effects on F2 onsets were not as consistent as they were for CVs preceded by [o]. In some cases, effects were similar to those seen with a preceding [o]. Figure 3 shows F2 dynamics for [ida] again from subject CW. The F2 onset value for V2 drops by around 140 Hz from the normal category to the fastest category. Notice that the downward trend here is much smaller than for the [oda] sequence from the same subject (see Figure 1). However, the upward trend of the V2 midpoint is stronger, covering a range of around 320Hz. Figure 4 shows the corresponding scatter plot for F2 onset and consonant closure duration, showing a positive correlation between the two variables (r=0.60).

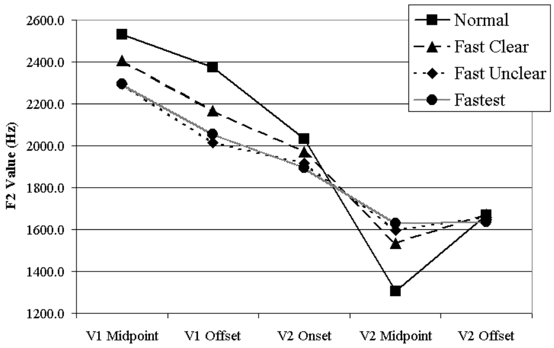

Figure 3: F2 dynamics of [ida] for subject CW

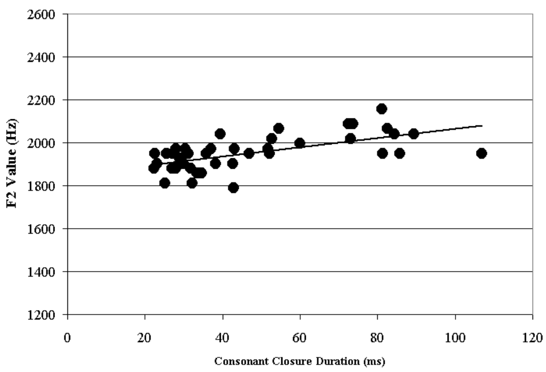

Figure 4: Stop closure duration related to V2 F2 onset values for [ida] for subject CW

In some cases with a preceding high F2 target, F2 onset values rose as rate increased. Figure 5 shows the dynamics of [iga] from subject CW. In this case, the normal category showed the lowest average F2 onset value, while the fast unclear category showed the highest, with a difference of about 150 Hz. The corresponding scatter plot in Figure 6 shows a negative correlation between the consonant closure duration and the F2 onset value (r=-0.54).

Table 2 summarizes the correlation of closure duration and F2 onset values for the cases in which the preceding vowel had a high F2 target. Half of the cases had either extremely

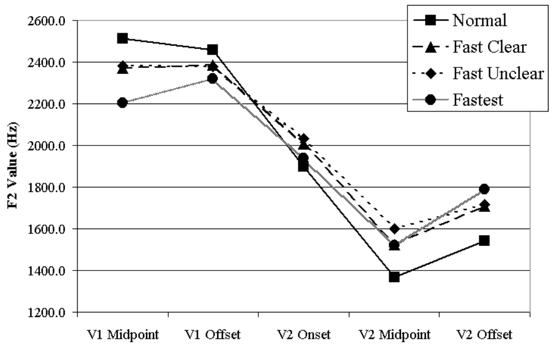

Figure 5: F2 dynamics of [iga] for subject CW

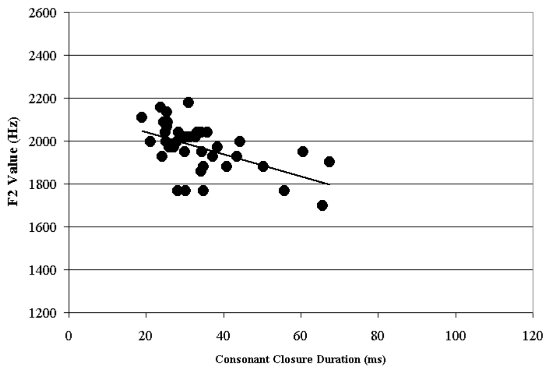

Figure 6: Stop closure duration related to V2 F2 onset values for [iga] for subject CW

S	VCV	r^2	Slope	*As rate increases:*
AT	iga	0.02	-1.14	
AT	igæ	0.15	4.22	lower F2 Onset
AT	ida	0.00	0.00	
AT	ido	0.08	0.74	
CW	iga	0.29	-5.17	higher F2 Onset
CW	igæ	0.00	-0.01	
CW	ida	0.36	2.11	lower F2 Onset
CW	ido	0.23	1.46	lower F2 Onset
FH	iga	0.22	-2.20	higher F2 Onset
FH	igæ	0.07	-1.20	higher F2 Onset
FH	ida	0.19	1.74	lower F2 Onset
FH	ido	0.00	0.19	
MB	iga	0.02	0.82	
MB	igæ	0.18	-1.95	higher F2 Onset
MB	ida	0.02	-0.32	
MB	ido	0.00	-0.01	

Table 2: Relationship between closure duration and F2 onset value for CV sequences preceded by [o].

low r^2 values, or extremely shallow slopes or both, indicating a weak correlation between the two variables. Four out of 16 [iCV]s show a positive correlation between closure duration and F2 onset (c.f. Figures 3 & 4). Another four showed a negative correlation. There is no obvious pattern relating a particular VCV to the type of correlation it shows.

ISBN 1-876346-48-5 © 2003 UAB

Overall it is clear that VCVs where V1 has a low F2 demonstrate stronger and more consistent effect than VCVs where V1 has a high F2 value. In [oCV] sequences in this experiment, both vowels have similar target F2 values. F2 onsets in these cases are influenced by the two vowels in similar ways. In the [iCV] sequences, the F2 target values for the vowels are in conflict, pulling the F2 onset in opposite directions. The F2 onset value may shift toward the F2 value of the preceding vowel or the following vowel or may hover between the two.

4. MODELING F2 ONSETS

F2 onset values are not based entirely upon the duration of the stop closure. A more complex model is needed to better understand the dynamics of this aspect of consonant articulation. F2 onsets were analyzed using multiple linear regressions with F2 midpoint of the preceding vowel, F2 midpoint of the following vowel, duration of stop closure, and F2 onset target as predictor variables.

This model thus takes into account values of F2 before the consonant closure (V1 midpoint F2), values of F2 achieved following the closure (V2 midpoint F2), an indicator of speaking rate (duration of stop closure) and the speaker's ideal F2 onset value (F2 onset target). The F2 onset target was derived from actual F2 onset values for the normal elicitation category. Finally, an interaction term is included in a stepwise manner to allow the model to reflect that the correlation of closure durations with F2 onset values differs substantially depending on the identity of the preceding vowel. Multiple linear regression models are summarized in Table 3.

Factors	*AT*	*CW*	*FH*	*MB*
V2 F2 Onset Target	.621	.365	.790	.756
Closure Duration	.564	.964	.605	1.045
V1 F2 Midpoint	.644	.872	1.056	1.001
V2 F2 Midpoint	.263	.359	.073	ns
Closure Dur X V1 F2	-.422	-.822	-.868	-1.137
Final r^2	*.905*	*.866*	*.881*	*.783*

Table 3: Multiple linear regression models including standardized coefficients and r^2 values.

In general, the main variables were found to be positively related to predicted F2 onset values, i.e. as the value of one of the main variables increases, so does the predicted F2 onset. However, a term representing the interaction of closure duration and preceding vowel midpoint was found to be significant, supporting the idea that the relationship between closure duration and F2 onset are not consistent across all values of V1 midpoint F2.

5. CONCLUSIONS

This investigation shows that like vowels, consonants behave in a dynamic way, undergoing changes as a function of timing, speaking style and vowel context. A quantitative treatment of the data indicates that F2 onsets can be predicted with good accuracy from information on the F2 midpoint of the preceding vowel, the F2 midpoint of the following vowel, the closure duration and the F2 onset target.

REFERENCES

[1] B. Lindblom, "Economy of speech gestures," in *The Production of Speech*, P. MacNeilage, Ed., pp. 215-245. New York: Springer-Verlag, 1998.

[2] B. Lindblom, "Explaining phonetic variation: a sketch of the H&H theory," in *Speech Production and Speech Modeling*, W.J. Hardcastle, and A. Marchal, Eds. Boston: Kluwer Academic Publishers, 1990.

[3] C. Mauk, *Undershoot in two modalities: An examination of fast utterances in English and American Sign Language,* Doctoral dissertation, University of Texas, Austin, 2003.

[4] D.P. Keuhn and K.L. Moll, "A cineradiographic study of VC and CV articulatory velocities," *Journal of Phonetics*, vol. 4, pp. 303-320, 1976.

[5] T. Gay, "Effect of speaking rate on vowel formant movements," *Journal of the Acoustical Society of America*, vol. 63, pp. 223-230, 1978.

[6] M.A. Picheny, N.I. Durlach and L.D. Braida, "Speaking clearly for the hard of hearing II: Acoustic characteristics of clear and conversational speech," *Journal of Speech and Hearing Research*, vol. 29, pp. 434-446, 1986.

[7] S-J. Moon, *An Acoustic and Perceptual Study of Undershoot in Clear and Citation-Form Speech,* Doctoral dissertation, University of Texas, Austin, 1991.

[8] T. Gay and T. Ushijima, "Effect of speaking rate on stop consonant-vowel articulation," *Proceedings of the Speech Communication Seminar, Stockholm, August 1-3*, 1974.

[9] T. Gay, T. Ushijima, H. Hirose, and F.S. Cooper, "Effect of speaking rate on labial consonant-vowel articulation," *Journal of Phonetics*, vol. 2, pp. 47-63, 1974.

[10] H.M. Sussman, E. Dalston, and S. Gumbert, "The effect of speaking style on a locus equation characteristic of stop place of articulation," *Phonetica*, vol. 55, pp. 204-225, 1998.

[11] K. Tjaden and G. Weismer, "Speaking-rate-induced variability in F2 trajectories," *Journal of Speech, Language, and Hearing Research*, vol. 41, pp. 976-989, 1998.

Acknowledgements

This research was supported by a National Institutes of Health grant (R01 DC02014).