ACOUSTIC EFFECTS OF LOUD SPEECH AND INTERRELATIONSHIPS AMONG MEASURES

Laura L. Koenig[a,b] and Susanne Fuchs[c]

[a]Long Island University, Brooklyn NY USA; [b]Haskins Laboratories, New Haven CT USA; [c]Zentrum für Allgemeine Sprachwissenschaft, Berlin Germany
koenig@haskins.yale.edu; fuchs@zas.gwz-berlin.de

ABSTRACT

This paper investigates how louder speech affects measures of formants, duration, intensity, and fundamental frequency in female speakers of German. Past studies of loud speech have not systematically assessed a large number of vowels, or how formant values compare to other acoustic measures. Loudness variation was elicited naturalistically in reading and question-answer tasks with target words designed to sample around the vowel space. Results indicate that the high tense vowels /i/ and /u/ show little formant variation in loud speech; other vowels vary mainly in the form of higher F1 in loud speech. Correlational analyses also showed vowel effects: Overall, the data demonstrated consistent relationships between F1 and intensity, duration, and f0 but effects were weaker and sometimes insignificant for /y u i/. The results indicate that conclusions about effects of loudness need to be made on a vowel-specific basis.

Keywords: loudness, formants, duration, intensity, f0

1. INTRODUCTION

Much previous work has evaluated how louder speech affects measures of intensity and/or fundamental frequency (f0) in typical speakers [e.g., 2, 4, 5, 7]. A few studies of loud speech also find changes in supraglottal articulation: Increased jaw [13] or lingual [11] displacement and higher F1's for /ɑ/ [3, 15]. Most past studies have focused on a single low vowel rather than sampling broadly across the vowel space, and studies that did include high as well as low vowels did not systematically evaluate vowel differences [3, 11]. Finally, past work has not fully considered interrelationships among duration, intensity, f0, and formant values. Such interrelationships can speak both to aspects of speech physiology and to possible speaker strategies for increasing loudness. For example, increased mouth opening, which should correlate with F1, may be a supraglottal strategy for increasing intensity. Researchers have also observed that loud speech may involve longer vowel durations [6], which in turn may permit speakers to reach more extreme articulatory positions [8], possibly for F2 as well as F1. Finally, increased subglottal pressure is known to affect both intensity and fundamental frequency [14]. Some authors have suggested that speakers may actively raise F1 under conditions of higher f0 [10]; thus, increased F1 values in loud speech could effectively be a side effect of source changes.

To evaluate these various possibilities, this paper assesses how loudness affects formants in vowels differing in height and tense/lax quality, and how such differences relate to intensity, duration, and f0.

2. METHODS

2.1. Speakers and elicitation methods

Data were collected from 11 female speakers of Standard (Northern) German. All were in good physical health and had normal body mass index. Data were collected in comfortable and loud conditions. In the comfortable condition, the experimenter/interlocutor stood a few feet from the participant; in the loud condition she went into an adjacent glass-walled room, closed the door, and spoke loudly through the glass. Speakers were not given specific instructions as to how much louder to speak in this second condition, but were told that increased volume might be needed for the experimenter to hear them.

Speakers engaged in two speech tasks: Reading a short set of utterances designed to elicit the point vowels in a connected speech context with minimal coarticulatory influences, and answering questions designed to put a target word in focus. The first task used the following sentences:

- Sie fuhren letzte Woche zur IAA nach Frankfurt [*They drove last week to the IAA [i a ɑ] in Frankfurt*]
- Sie fahren nächste Woche zur LUU nach Hannover [*They're going next week to the LUU [ɛl u u] in Hannover*]
- Wir wollen am Wochenende zur BII nach Hamburg [*We want to go on the weekend to the BII [be i i] in Hamburg*]

In the other task, participants responded to an experimenter question using target words that sampled across the vowel space and used a bilabial-alveolar consonantal context. For example: “Magst du X?” *[Do you like X?]* “X mag ich, aber nicht Y.” *[X I like, but not Y]*. Two words were included for each of the vowels [i ɪ ɑ: ɑ u ʊ y ʏ]: Mieten, Pita, Mitte, Pizza, Mate, Paten, Paddeln, Pasta, Pudel, Pute, Butter, Pudding, Büsten, Büsum, Mützen, München.

The final dataset for the reading task consisted of 788 productions, approximately 72 per speaker. The final dataset for the question-answer task included 1707 productions, an average of ca. 155 per speaker.

2.2. Signals and measures

The full experimental protocol included simultaneous recordings of acoustics, electroglottography, intraoral pressure, and respiratory kinematics (respitrace). The current work analyses the acoustic data only. Data were recorded to a multi-channel acquisition system, sampled at 22,050 Hz. Speakers were seated during recording with a standing microphone placed approximately 30 cm from the mouth.

Onsets and offsets of the target vowels were manually labeled in Praat [1]. Automatic routines were used to obtain f0, intensity, and formant (F1, F2) values at the temporal midpoint of the vowel. Formant values were corrected by hand, adjusting the LPC order, if the automatically-obtained values were inappropriate for the vowel in question.

2.3. Statistics

Statistical analyses on loudness conditions were carried out in R [12]. Correlational analyses were conducted within vowels using Matlab [9].

For the reading data, which included only the corner vowels /i ɑ u/, we used repeated-measures ANOVAs with F1 or F2 as dependent measures, Condition (loud vs. normal) and Vowel (/i ɑ u/) as fixed effects, and Repetition per Speaker as an error term.

The question-answer data yielded a more complex structure; thus these data were analyzed with the more flexible methods of Linear Mixed Models (lme4 package in R). The dependent variables were again F1 or F2, and fixed effects were Condition (loud vs. normal) and Vowel (/i ɪ ɑ: ɑ u ʊ y ʏ/), using /ɑ/ as the reference level. Speaker and Word were included as random intercepts, and Word by Condition and Speaker by Condition were entered as random slopes. The results section reports only significant effects of interest for the current questions. In the following, t-values>|±2| will be considered significant.

3. RESULTS

3.1. Effects of loudness on vowel formants

Figures 1–2 show scatterplots of the vowel spaces from the question-answer task for tense and lax vowels, respectively. In general, formant variation as a function of loudness is evident in F1 but minimal in F2. The data also suggest that loudness effects differ as a function of both vowel height and tense/lax status.

Figure 1: F1-F2 plot of tense vowels, all speakers, in comfortable (grey) and loud (black) conditions for the question-answer task.

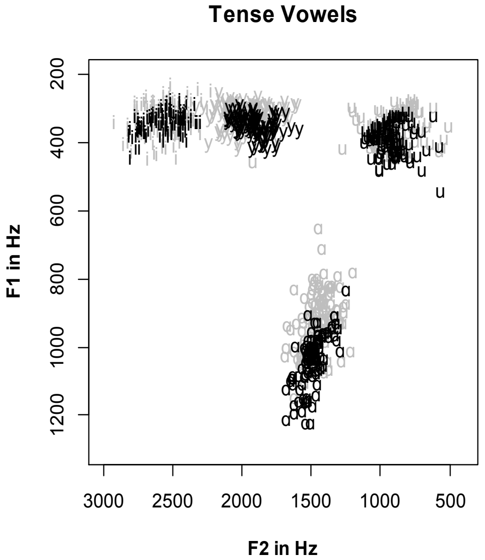
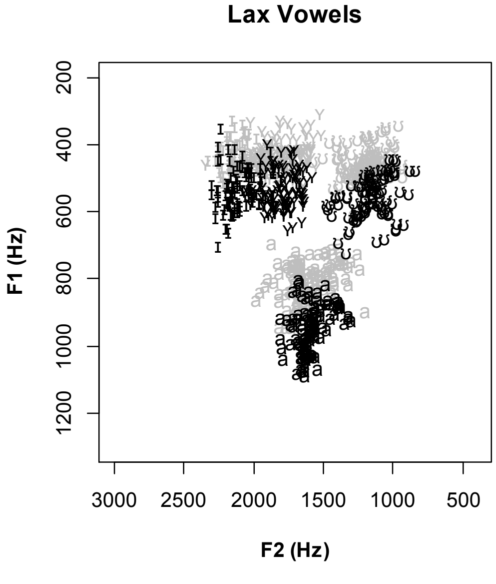


Figure 2: F1-F2 plot of lax vowels, all speakers, in comfortable (grey) and loud (black) conditions for the question-answer task.



The reading task showed significant effects of loudness condition on F1 for each of the three vowels /i ɑ u/ (p<.001). For F2, the condition effect was not significant, but vowel was, as was the vowel x condition interaction. Post-hoc results showed significant (p<.001) effects for /u/ and /ɑ/. In both cases F2 values were slightly higher in the loud condition.

For the question-answer data, statistical results for F1 were significant for all vowels except /i/ (Table 1). The results also show higher t-values for the lax vowels, indicating greater confidence in the coefficient as a predictor. F2, in contrast, only showed a rather weak condition effect in two cases: For /ɑː/ (t=-2.16), where F2 was higher in loud speech, and for /y/ (t=2.99), where F2 was lower in loud speech.

Table 1: Statistical results (t-values) for F1, question-answer task.

Tense vowels	t-value	t-value	Lax vowels
/ɑː/	-10.26	-11.99	/ɑ/
/i/	-0.75	-7.64	/ɪ/
/u/	-2.13	-8.90	/ʊ/
/y/	-2.49	-8.92	/ʏ/

3.2. Correlations among measures

Given that loudness variation was primarily associated with changes in F1, correlational analyses with intensity, duration, and f0 were performed for the first formant only.

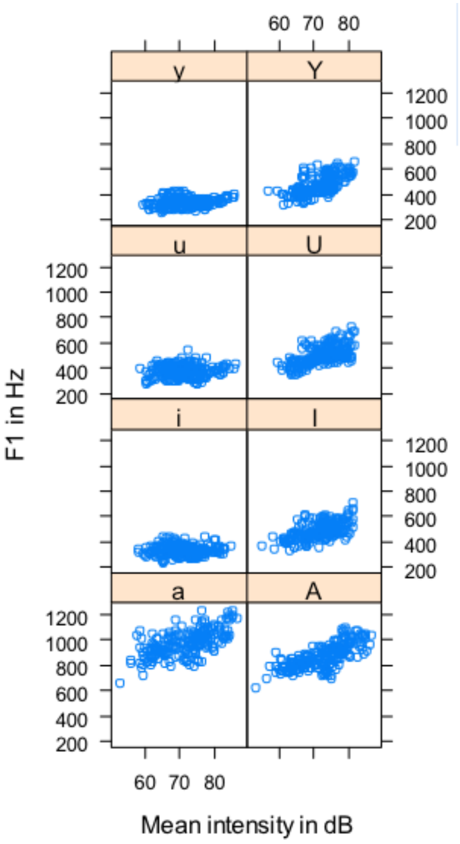
3.2.1. F1 and intensity

Increased F1 in loud speech suggests more open articulatory configurations, and could reflect a supraglottal adjustment for enhancing loudness. The current data show a relationship between F1 and intensity, but primarily for low and lax vowels. Correlation results for the question-answer task are shown in Table 2, and Figure 3 shows F1 values as a function of intensity for each vowel in the question-answer task, with all speakers combined. The relationship was insignificant for /u/ and /i/, and whereas the correlation for /y/ was significant, the ρ-value was considerably lower than for the corresponding lax vowel. Data from the reading task were comparable: A significant correlation was found for /ɑ/ but not for /i/ or /u/.

Table 2: Statistical results (p- and ρ-values) for correlations between intensity and F1, question-answer task.

Tense vowels	p-value	ρ-value	p-value	ρ-value	Lax vowels
/ɑː/	<.001	0.628	<.001	0.730	/ɑ/
/i/	0.7535	0.022	<.001	0.573	/ɪ/
/u/	0.1506	0.100	<.001	0.622	/ʊ/
/y/	0.0027	0.205	<.001	0.613	/ʏ/

Figure 3: Scatterplots of intensity and F1 and intensity for all speakers. The labels Y U I A in the right column correspond to IPA [ʏ ʊ ɪ ɑ].



3.2.2. F1 and duration

Longer durations have been associated with more extreme articulatory positions, which could also facilitate louder speech. Table 3 shows correlation results for the question-answer task, and scatter plots are shown in Figure 4. Correlations were significant for all vowels, but the ρ-values were lower in all cases for the tense vowels compared to their lax counterparts. In the reading data, all correlations were again significant, although that for /ɑ/ was borderline (p=.049).

Table 3: Statistical results (p- and ρ-values) for correlations between duration and F1, question-answer task.

Tense vowels	p-value	ρ-value	p-value	ρ-value	Lax vowels
/ɑː/	<.001	0.449	<.001	0.722	/ɑ/
/i/	<.001	0.303	<.001	0.738	/ɪ/
/u/	<.001	0.327	<.001	0.701	/ʊ/
/y/	<.001	0.254	<.001	0.610	/ʏ/

Figure 4: Scatterplots of duration and F1 for all speakers. The labels Y U I A in the right column correspond to IPA [ʏ ʊ ɪ ɑ].

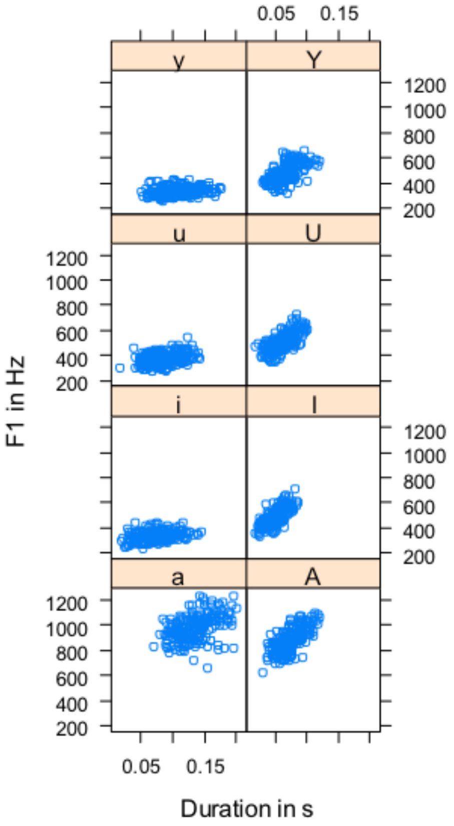


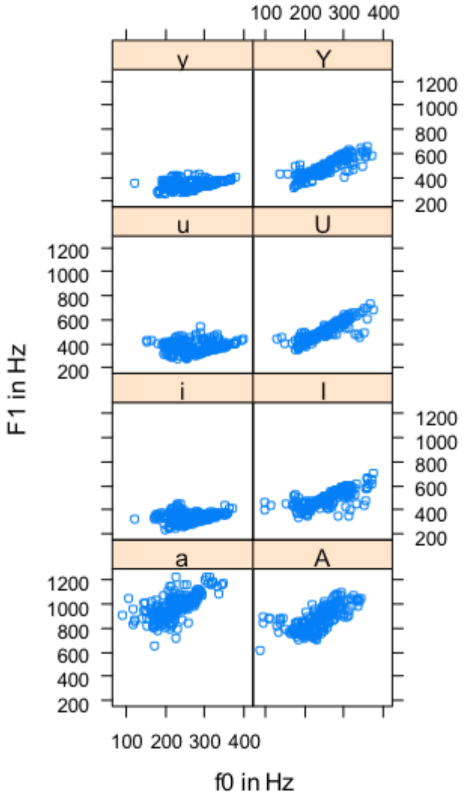
Table 4: Statistical results (p- and ρ-values) for correlations between f0 and F1, question-answer task.

Tense vowels	p-value	ρ-value	p-value	ρ-value	Lax vowels
/ɑː/	<.001	0.636	<.001	0.652	/ɑ/
/i/	0.0015	0.215	<.001	0.710	/ɪ/
/u/	0.0919	0.117	<.001	0.855	/ʊ/
/y/	<.001	0.351	<.001	0.843	/ʏ/

3.2.3. F1 and f0

Some authors have proposed that speakers increase F1 as the f0 increases. Results of correlational analyses are presented in Table 4, and Figure 5 shows the scatterplots for all vowels in the question-answer task. The data are similar to those for intensity: Results for /u/ did not reach significance, and although significant relationships were seen for /y/ and /i/ and the ρ-values were rather low.

Figure 5: Scatterplots of f0 and F1 for all speakers. The labels Y U I A in the right column correspond to IPA [ʏ ʊ ɪ ɑ].



4. DISCUSSION

Despite considerable attention to the characteristics of loud speech in typical speakers, and interest in loud speech as a possible therapy for some speech disorders (e.g., [16]), vowel-specific effects remain relatively unexplored. The current data partly support past reports of more open articulatory positions in loud speech [3, 11, 13, 15], but contrary to [13], effects are not observed for all vowels. Instead, the current data consistently indicate that loudness effects are most prominent for vowels that are low and lax compared to those that are high and tense. In future work we will compare these acoustic results with data on respiration, vocal-fold contact, and intraoral pressure to explore the physiological underpinnings of these vowel differences.

5. ACKNOWLEDGMENTS

Thanks to our participants, and to Jörg Dreyer for technical support.

6. REFERENCES

[1] Boersma, P., Weenink, D. 2011. Praat: Doing phonetics by computer. [Computer program]. Version 5.3. http://www.praat.org.

[2] Finnegan, E., Luschei, E. S., Hoffman, H. T. 2000. Modulations in respiratory and laryngeal activity associated with changes in vocal intensity during speech. *J. Speech Lang. Hear. Res.* 43, 934–950.

[3] Geumann, A. 2001. Vocal intensity: Acoustic and articulatory correlates. In *Fourth International Speech Motor Control Conference*. Netherlands: Nijmegen.

[4] Gramming, P., Sundberg, J., Ternström, S., Leanderson, R., Perkins, W. H. 1988. Relationship between changes in voice pitch and loudness. *J. Voice* 2, 118–126.

[5] Holmberg, E. B., Hillman, R. E., Perkell, J. S. 1988. Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal, and loud voice. *J. Acoust. Soc. Am.* 84, 511–529.

[6] Huber, J. E., Stathopoulos, E. T., Cirione, G. M., Ash, T. A., Johnson, K. 1999. Formants of children, women, and men: The effects of vocal intensity variation. *J. Acoust. Soc. Am.* 106, 1532–1542.

[7] Ladefoged, P., McKinney, N. 1963. Loudness, sound pressure, and subglottal pressure in speech. *J. Acoust. Soc. Am*. 35, 454–460.

[8] Lindblom, B. 1963. Spectrographic study of vowel reduction. *J. Acoust. Soc. Am.* 35, 1773–1781.

[9] Matlab [Computer program]. Version 2014b. The Mathworks.

[10] Maurer, D., Cook, N., Landis, T., d'Heureuse, C. 2009. Are measured differences between the formants of men, women and children due to f0 differences? *J. Int. Phon. Assoc.* 21, 66–79.

[11] Mefferd, A. S., Green, J. R. 2010. Articulatory-to-acoustic relations in response to speaking rate and loudness manipulations. *J. Speech Lang. Hear. Res.* 53, 1206–1219.

[12] R Core Team. 2012. *R: A language and environment for statistical computing*. http://www-R-project.org.

[13] Schulman, R. 1989. Articulatory dynamics of loud and normal speech. *J. Acoust. Soc. Am.* 85, 295–312.

[14] Titze, I. 1994. *Principles of Voice Production.* Englewood Cliffs: Prentice Hall.

[15] Traunmüller, H., Eriksson, A. 2000. Acoustic effects of variation in vocal effort by men, women, and children. *J. Acoust. Soc. Am.* 107, 3438–3451.

[16] Sapir, S., Spielman, J. L., Ramig, L. O., Story, B. H., Fox, C. 2007. Effects of intensive voice treatment (the Lee Silverman voice treatment [LSVT]) on vowel articulation in dysarthric individuals with idiopathic Parkinson disease: Acoustic and perceptual findings. *J. Speech, Lang. Hear. Res.* 50, 899–912.