signal processing laboratory

ieeta instituto de engenharia electrónica e telemática de aveiro

Durational and Voicing Characteristics of Fricatives

Luis M. T. Jesus

Escola Superior de Saúde da Universidade de Aveiro, and

Instituto de Engenharia Electrónica e Telemática de Aveiro

Universidade de Aveiro, 3810 - 193 Aveiro, Portugal

e-mail: lmtj@ieeta.pt

www: http://www.ieeta.pt/∼lmtj

Introdution

In this poster, we present a novel methodology of temporal analysis and results including the durations of the fricatives, and of the VF and FV transitions and a study of devoicing, together with discussion of some possible causes of this phenomenon. The correlation between devoicing and duration was also investigated. The data presented in this poster could be used to improve the naturalness of synthetic speech.

There is some perceptual evidence (Cole and Cooper 1975; Scully 1979; Hogan and Rozsypal 1980) that the vowel to fricative duration ratio is used as a primary cue for voicing, so we've also extracted and analysed new acoustic data from the corpora described by Jesus (2001), namely the ratio of fricative duration to the preceding and following vowel duration.

There have been various studies in the past reporting the duration of fricatives, and of the VF and FV transitions, and investigating the influence of word - and sentence - position, stress and vowel context on fricative durations (Klatt 1971; Klatt 1974; Klatt 1975; LaRiviere et al. 1975; Umeda 1977; Manrique and Massone 1981; Behrens and Blumstein 1988; Crystal and House 1988; Docherty 1992; Shadle and Mair 1996; Crystal and House 1997).

Table 1: Duration of fricatives in American English (Klatt 1975) and Argentine Spanish (Manrique and Massone 1981).

	Duration (ms)	
	American English	Argentine Spanish
[f] stressed	110	147
[f] unstressed	105	192
[v] stressed	75	-
[v] unstressed	65	-
[θ] stressed	100	-
[θ] unstressed	95	-
[ð] stressed	60	98
[ð] unstressed	60	104
[s] stressed	120	148
[s] unstressed	120	187
[z] stressed	60	-
[z] unstressed	60	-
[ʃ] stressed	110	170
[ʃ] unstressed	110	210
[ʒ] stressed	-	98
[ʒ] unstressed	-	149

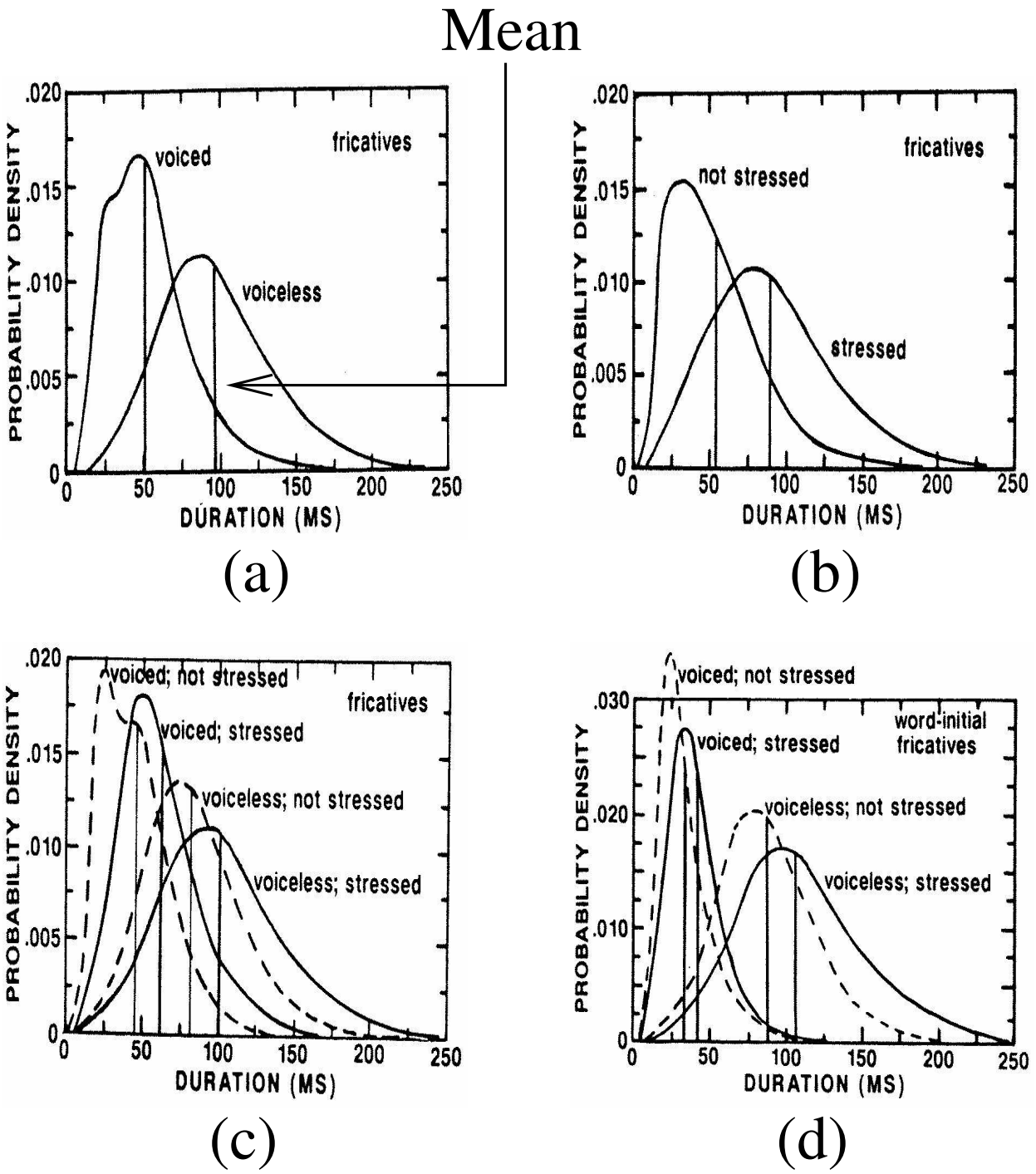


Figure 1: Probability density distribution curves of the durations of: (a) voiced /v, ð, z/ and voiceless /f, θ, s/ fricatives; (b) fricatives occurring in stressed and unstressed syllables; (c) voiced and voiceless fricatives occurring in stressed and unstressed syllables; (d) word-initial voiced and voiceless fricatives occurring in stressed and unstressed syllables. From Crystal and House (1988).

The difference in duration between unvoiced and voiced fricatives, and the frequent devoicing of voiced fricatives, have also been reported in many studies (Docherty 1992; Shadle and Mair 1996; Pirello et al. 1997; Smith 1997; Jongman et al. 2000).

Table 2: Inventory of all cases of devoicing. Values given are in the form x/y, where x = number of devoiced, partially devoiced or voiced examples, and y = total number of examples. After Docherty (1992).

	Word - Initial	Word - Final	All Pos.	
/v/	14/68 (21%)	12/65 (19%)	26/133 (20%)	Devoiced
	16/68 (24%)	20/65 (31%)	36/133 (27%)	Partially Devoiced
	38/68 (56%)	33/65 (51%)	71/133 (53%)	Voiced
/ð/	4/21 (19%)	5/18 (28%)	9/39 (23%)	Devoiced
	-	7/18 (39%)	7/39 (18%)	Partially Devoiced
	17/21 (81%)	6/18 (33%)	23/39 (59%)	Voiced
/z/	3/41 (7%)	17/113 (15%)	20/154 (13%)	Devoiced
	8/41 (20%)	68/113 (60%)	76/154 (49%)	Partially Devoiced
	30/41 (73%)	28/113 (25%)	58/154 (38%)	Voiced
All Fric.	21/130 (16%)	34/196 (17%)	55/326 (17%)	Devoiced
	24/130 (19%)	95/196 (49%)	119/326 (37%)	Partially Devoiced
	85/130 (65%)	67/196 (34%)	152/326 (47%)	Voiced

Corpora

Corpus of Portuguese fricatives /f, v, s, z, ʃ, ʒ/:

- **Corpus 1a** – sustained fricatives preceded by vowels /ɨ, ɐ, u/;
- **Corpus 1b** – fricatives sustained at different effort levels;
- **Corpus 2** – /pV_1FV_2/ repeated 12 times on one breath, V_i=/ɨ, ɐ, u/, and following Portuguese phonological rules;
- **Corpus 3** – 154 real words produced in a frame sentence. Examples of all Portuguese fricatives, with all non-nasalized vowels.
- **Corpus 4** – 12 sentences including 60 words from Corpus 3.

Subjects: 4 adult Portuguese native speakers; two males (**LMTJ** and **CFGA**) and two females (**ACC** and **ISSS**).

Two European Portuguese and British English bilingual siblings, one male (**PS**) and one female (**RS**), were also recorded saying the same Portuguese corpora used for the monolingual speakers.

Recordings: sound - treated booth (B & K 4165 1/2 inch microphone 1 m from the subject, B & K 2636 measurement amplifier), Sony TCD-D7 DAT recorder (16 bits, sampling frequency 48 kHz).

Segmentation and Annotation

The data on the DAT tape were digitally transferred to `.wav` computer files, which contain the acoustic speech signal in the right channel and the laryngograph signal on the left channel.

The time waveforms of all the corpus words were manually analysed to detect the start of the vowel - fricative transition, the start of the fricative, the end of the fricative, and the start of the fricative - vowel transition.

During the vowel - fricative transition, there is a decrease in amplitude, voicing ceases (for unvoiced fricatives) and frication noise starts, as shown in Figure 2.

During the fricative - vowel transition, there is an increase in amplitude, voicing starts (for unvoiced fricatives) and frication noise ceases (Docherty 1992, pp. 118 - 119).

These events do not occur simultaneously or always in the same order, making the segmentation a somewhat subjective process. However, it is important to segment consistently, because the results of the analysis methods depend on where the boundaries are placed (Docherty 1992, pp. 103 - 110).

The amplitude and voicing changes appear in both acoustic and Lx signals, which aids the segmentation process. For example, as can be seen in Figure 2, the FV transition also includes some frication noise because we've established that an unvoiced fricative would only correspond to a steady-state noise segment.

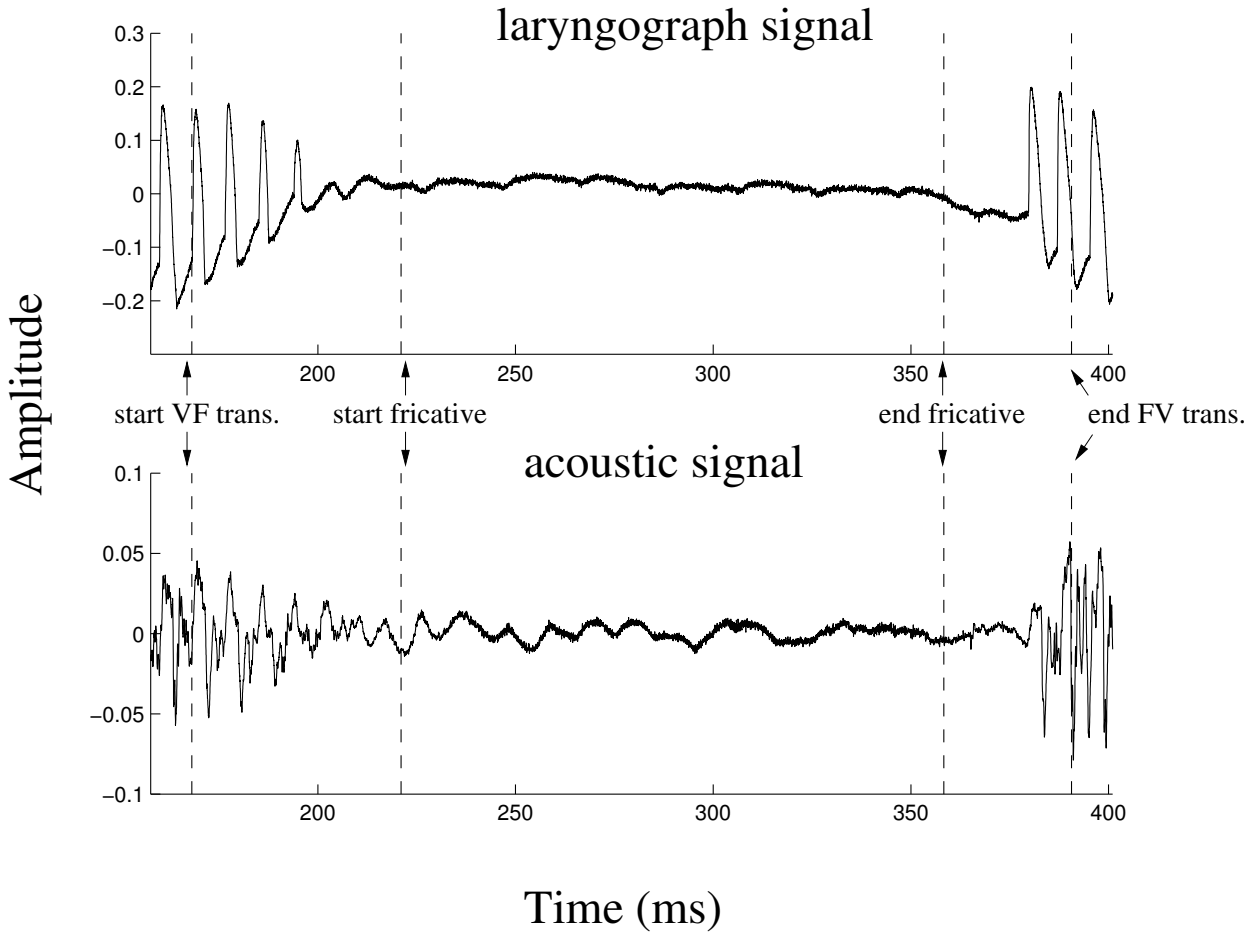


Figure 2: Laryngograph signal and acoustic signal of fricative /f/ in *café* /kɐˈfɛ/, showing the start of the vowel-fricative transition, the start of the fricative, the end of the fricative, and the end of the fricative-vowel transition. Corpus 3 (Speaker LMTJ).

The laryngograph signal was also used in the decision process to determine the VF and FV boundaries (see Figure 2).

For unreduced vowels there was always significant voicing, and for the duration of most fricatives the laryngograph signal changed drastically. Therefore, the amplitude of the laryngograph signal was an important cue in determining the boundaries between the different phones.

When it was not clear from the acoustic signal where the fricative started and ended (especially for voiced fricatives), the laryngograph signal was used as an additional cue, because its amplitude diminishes during the VF transition and increases during the FV transition.

The sentence corpora of speakers PS and RS has been annotated to include more information about the vowels preceding and following the fricatives: start of the vowel preceding the fricative and end of the vowel following the fricative.

Devoicing Criteria

When voiced fricatives devoice, it is with a whisper phonation (Abercrombie 1967, p. 137), distinguishing them from their voiceless counterparts which are realised with a glottal abduction gesture.

Smith (1997) also suggested that the glottis is in a state intermediate between voicing and voicelessness, like the state of the glottis that is used in whisper, with the glottis open but the folds very close together.

The signal shown in Figure 3 is one such case. It corresponds to a segment that starts at the onset of the VF transition, and ends at offset of the FV transition.

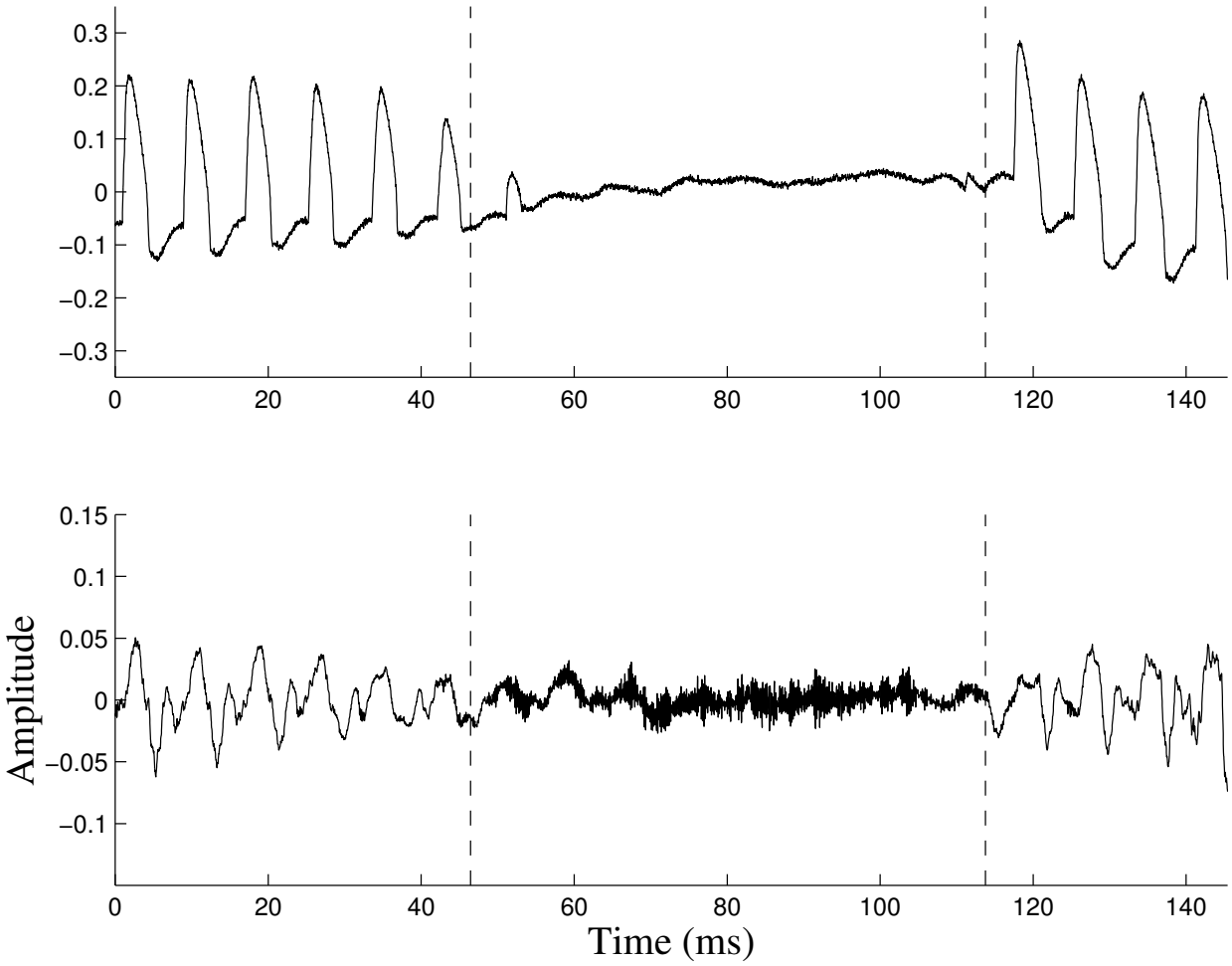


Figure 3: Laryngograph signal and acoustic signal of fricative /z/ in *diga zarpar* /ˈdigɐ zɐrˈpar/. The dashed lines mark the start and end of the fricative. Corpus 3 (Speaker LMTJ).

Both the acoustic signal and the laryngograph signal were used to determine if a fricative was devoiced (Jesus and Shadle 2002).

A fricative was called *devoiced* when less than one-third of the frication interval showed periodic structure in the acoustic or laryngograph signals.

The term *partially devoiced* was used when more than one-third but less than half of the frication interval contained steady acoustic and laryngograph signal cycles.

A fricative was called *voiced* when more than half of the frication interval showed steady acoustic and laryngograph signal cycles, even if the amplitude was much lower than in the vowel (Docherty 1992, p. 13).

If the laryngograph signal was clearly periodic, the interval was classified as voiced; if the laryngograph signal was zero or distorted, the signal was classified as voiced only if the acoustic signal was unambiguously periodic.

Duration Results – Corpus 3

The mean duration of the unvoiced fricatives is always greater than the mean duration of the voiced fricatives, as shown in Figures 4 to 7.

There is no significant difference by place of articulation.

The mean duration of the fricative is greater than the mean duration of the VF and FV transitions, and the mean duration of the VF transition is greater than the mean duration of the FV transition for speakers LMTJ, ACC and ISSS, as shown in Figures 4, 6 and 7.

For Speaker LMTJ's /s/ in word-initial position, as the following vowel's place of articulation moves further back, the duration of the fricative diminishes. This was only observed for this fricative produced by Speaker LMTJ.

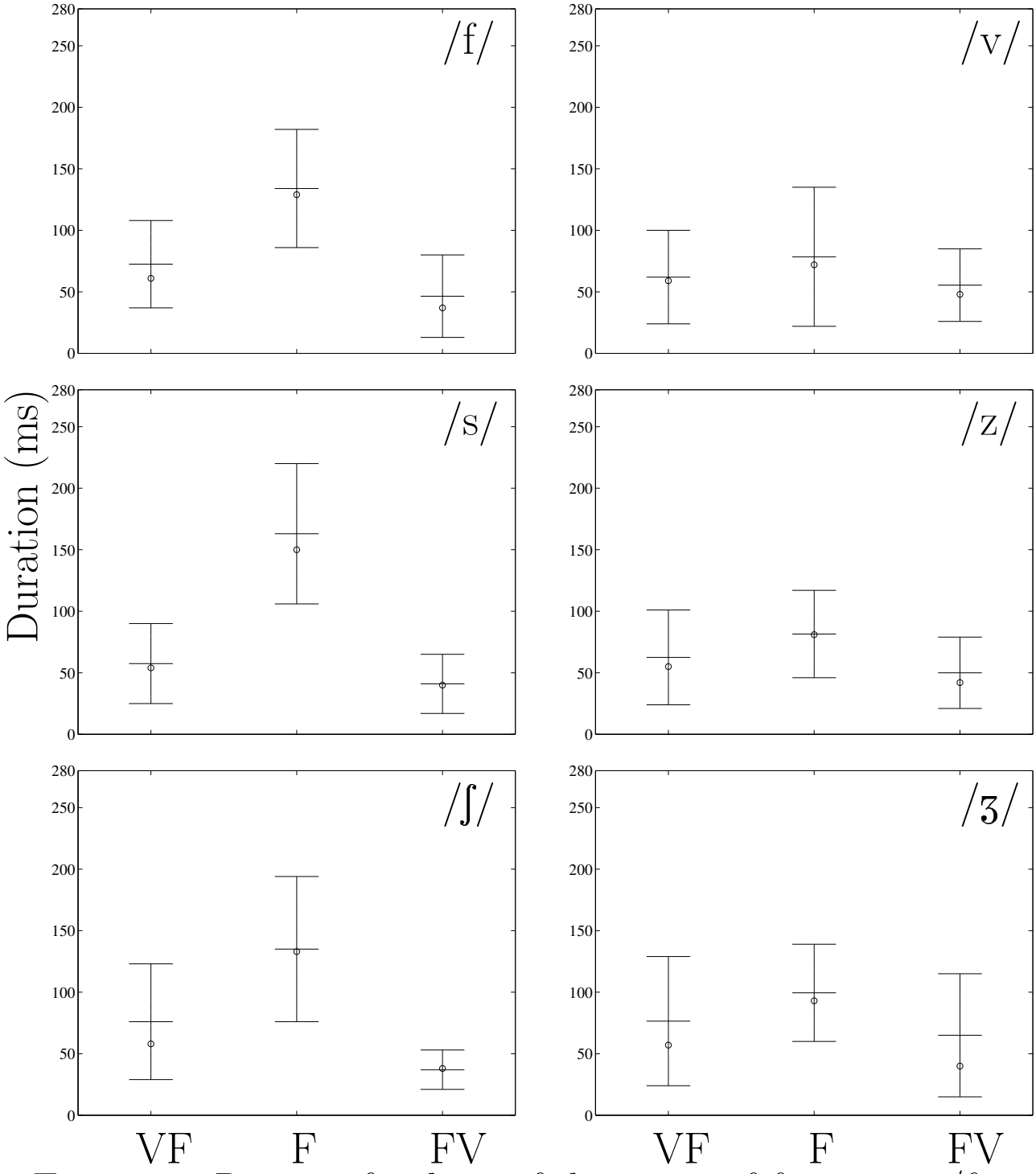


Figure 4: Range of values of duration of fricatives /f, v, s, z, ʃ, ʒ/, and of VF and FV transitions. o is the mean. Corpus 3 (Speaker LMTJ).

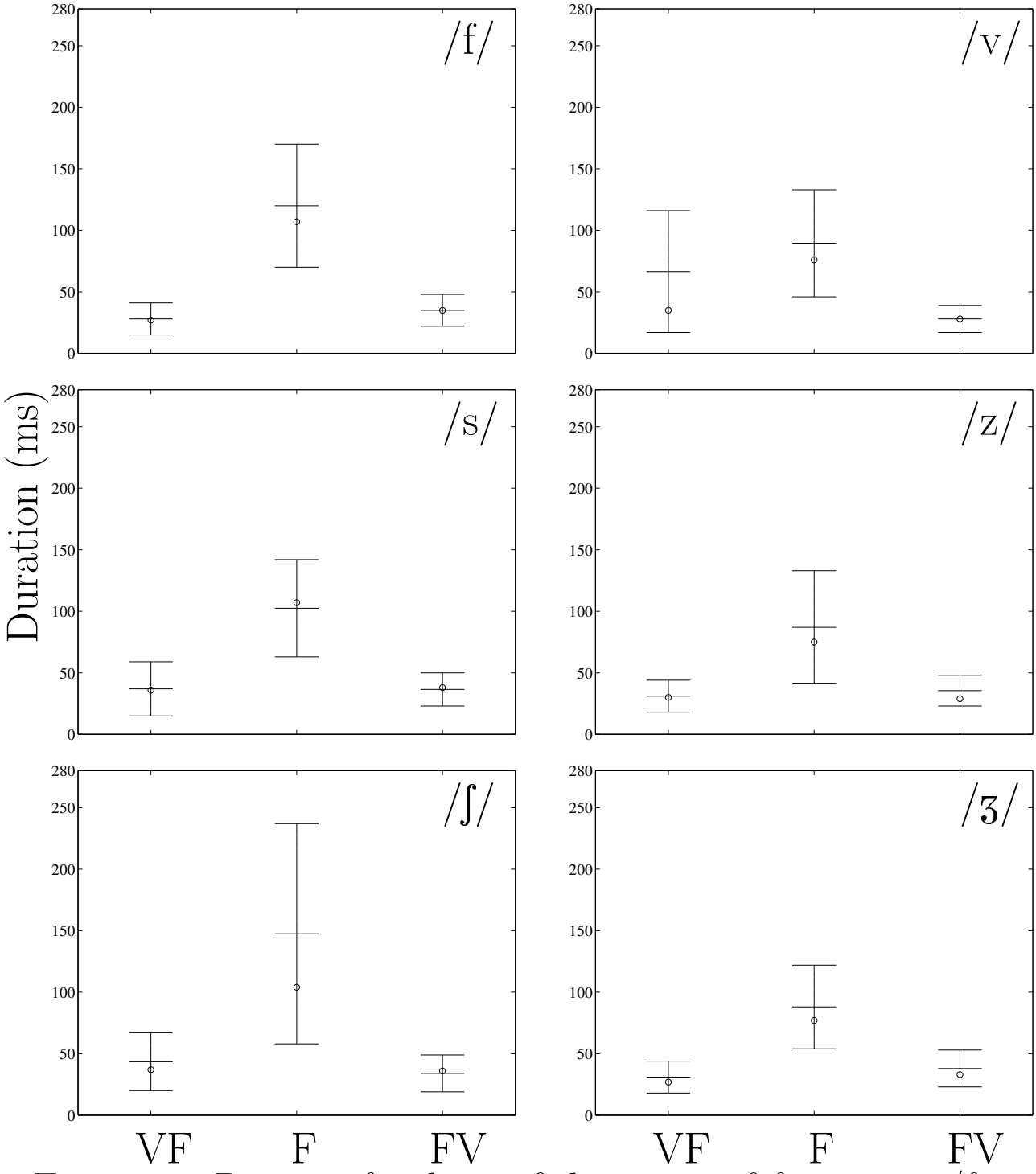


Figure 5: Range of values of duration of fricatives /f, v, s, z, ʃ, ʒ/, and of VF and FV transitions. o is the mean. Corpus 3 (Speaker CFGA).

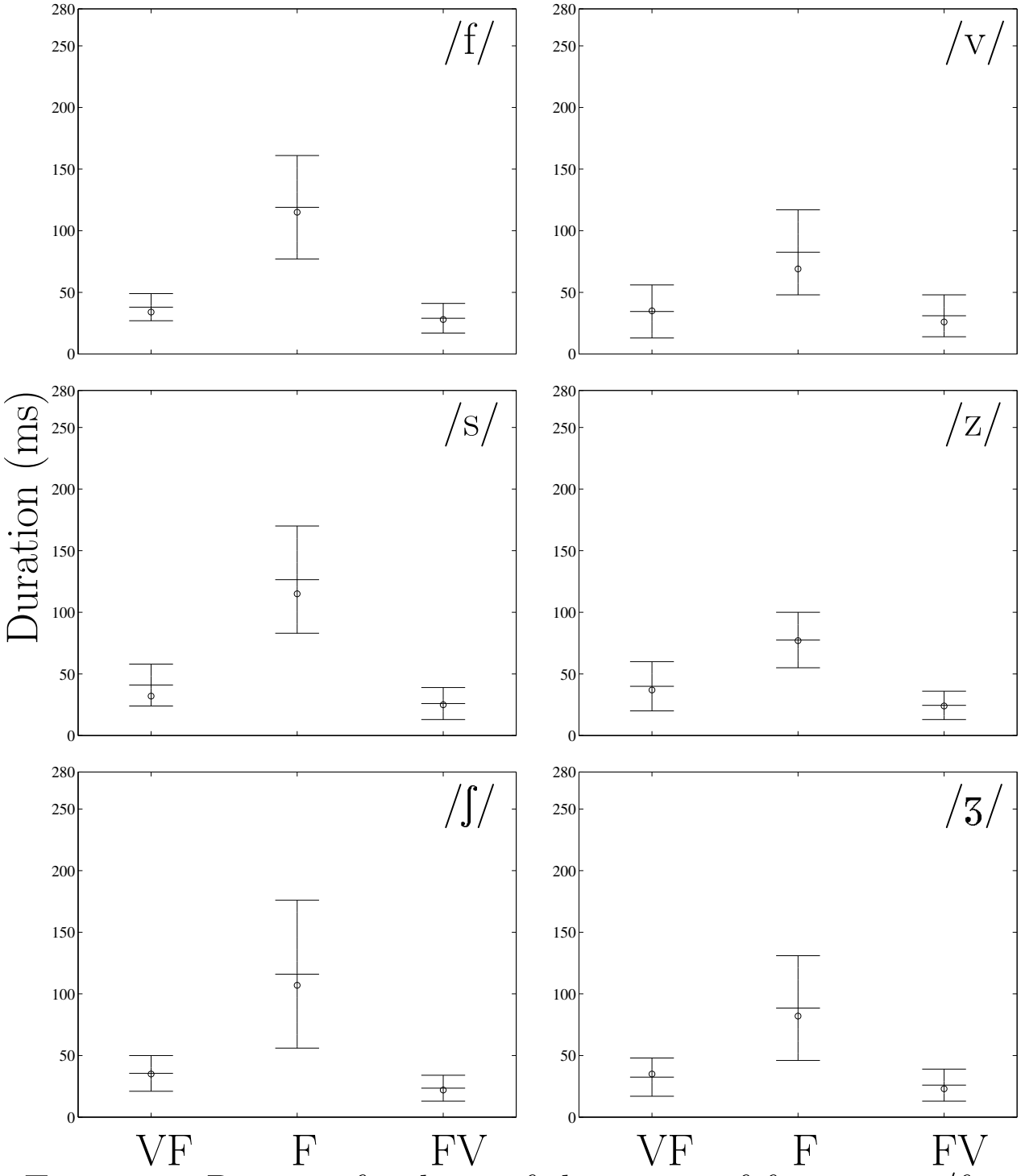


Figure 6: Range of values of duration of fricatives /f, v, s, z, ʃ, ʒ/, and of VF and FV transitions. o is the mean. Corpus 3 (Speaker ACC).

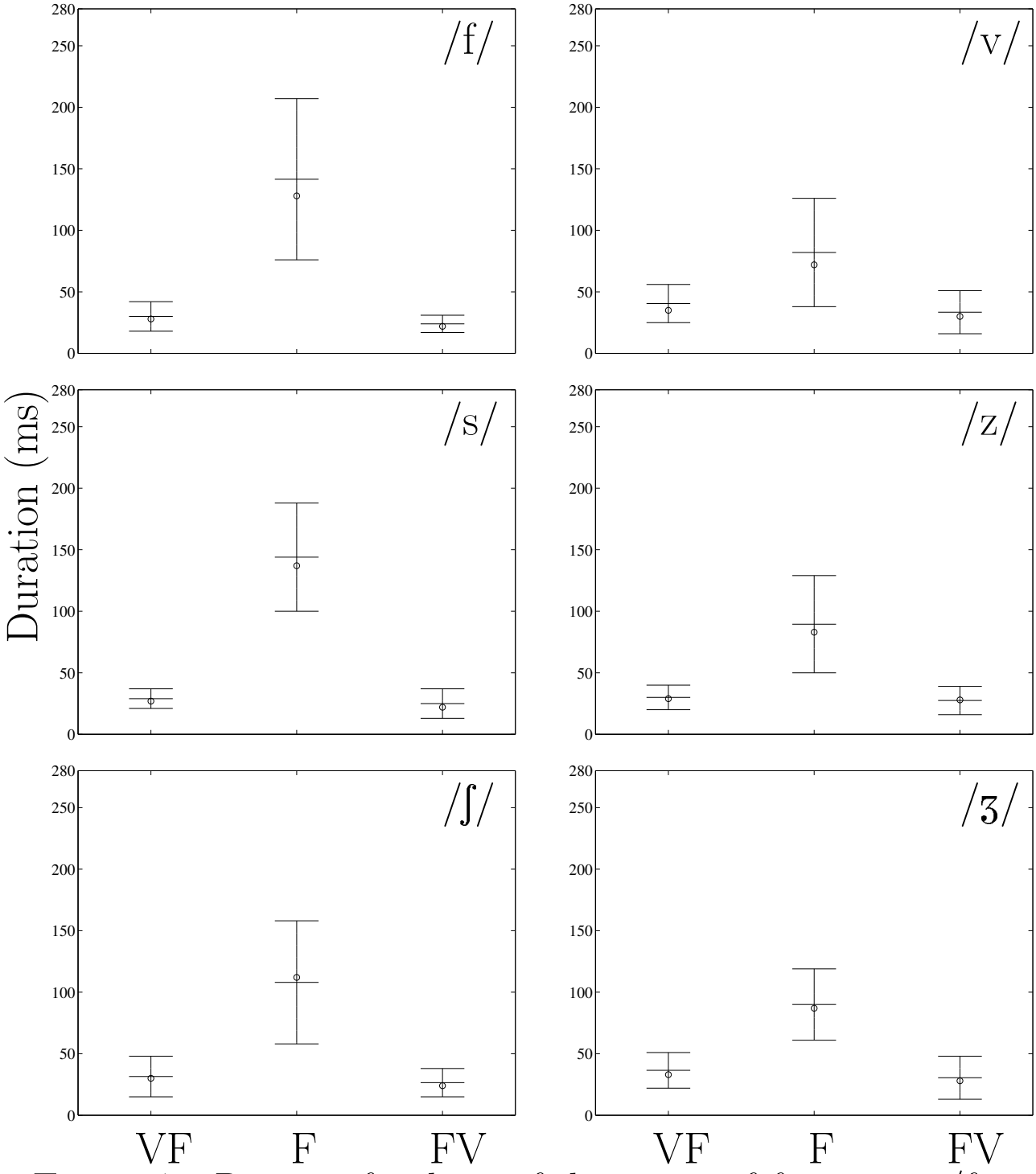


Figure 7: Range of values of duration of fricatives /f, v, s, z, ʃ, ʒ/, and of VF and FV transitions. o is the mean. Corpus 3 (Speaker ISSS).

Duration Results – Corpus 4

As previously observed for Corpus 3 fricatives, the duration of Corpus 4 unvoiced fricatives is always greater than the duration of their voiced counterparts, as shown in Figures 8 to 11, which agrees with results for the English language (Hogan and Rozsypal 1980; Crystal and House 1988; Stevens et al. 1992; Pirello et al. 1997).

The mean duration of the fricatives is greater than the duration of the VF and FV transitions, and comparing the mean duration of the VF and FV transitions, no consistent pattern can be found for any of the speakers (compare with results for Corpus 3), as shown in Figures 8 to 11.

Corpus 4 alveolar fricatives /s, z/ are on average longer than labiodentals /f, v/, a fact that had been previously reported by Pirello et al. (1997) for English, but which was not observed in Corpus 3 fricatives.

Word-final fricatives at the end of the sentences read by Speaker LMTJ have much longer duration than other examples in Corpus 4.

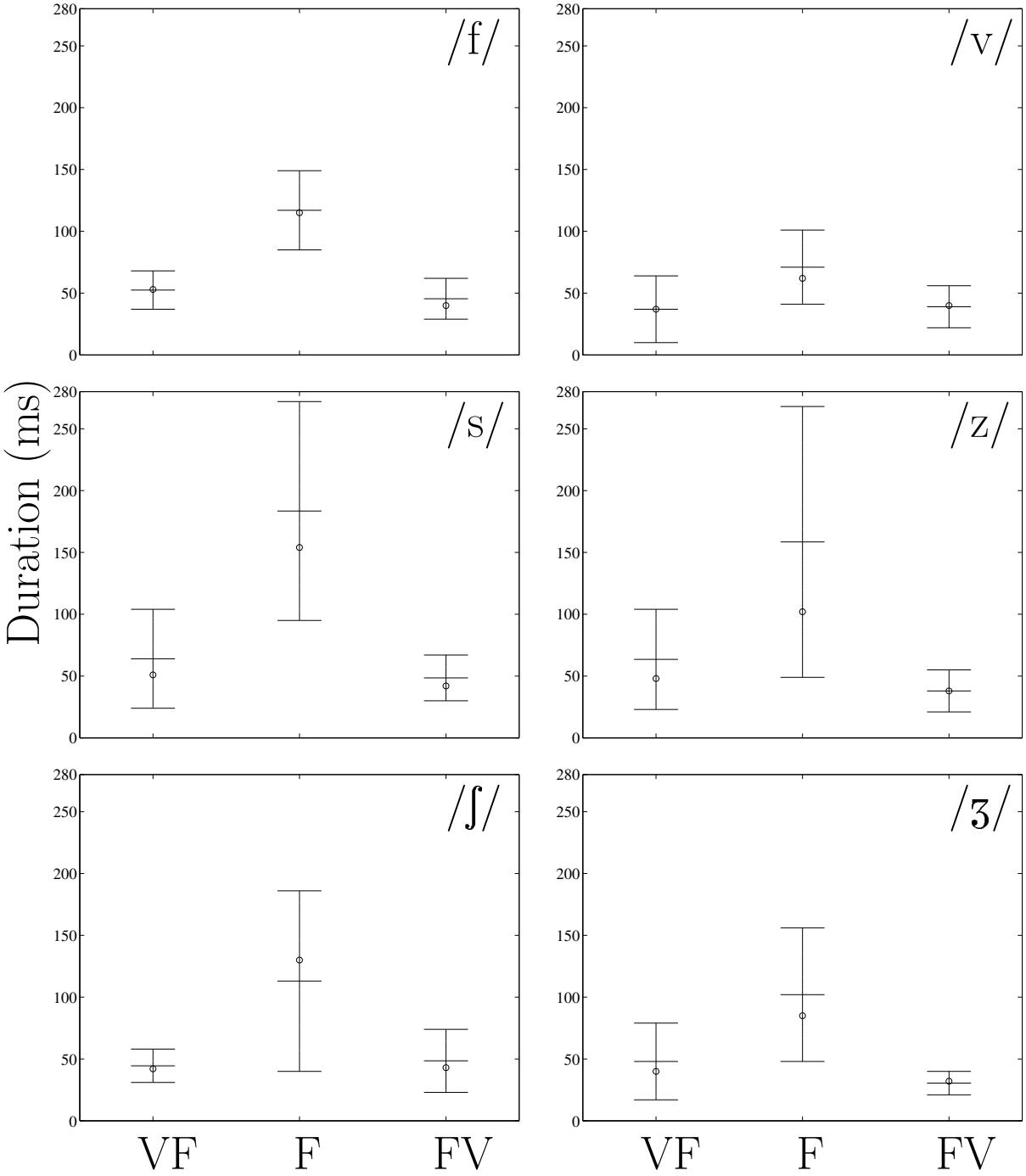


Figure 8: Range of values of duration of fricatives /f, v, s, z, ʃ, ʒ/, and of VF and FV transitions. o is the mean. Corpus 4 (Speaker LMTJ).

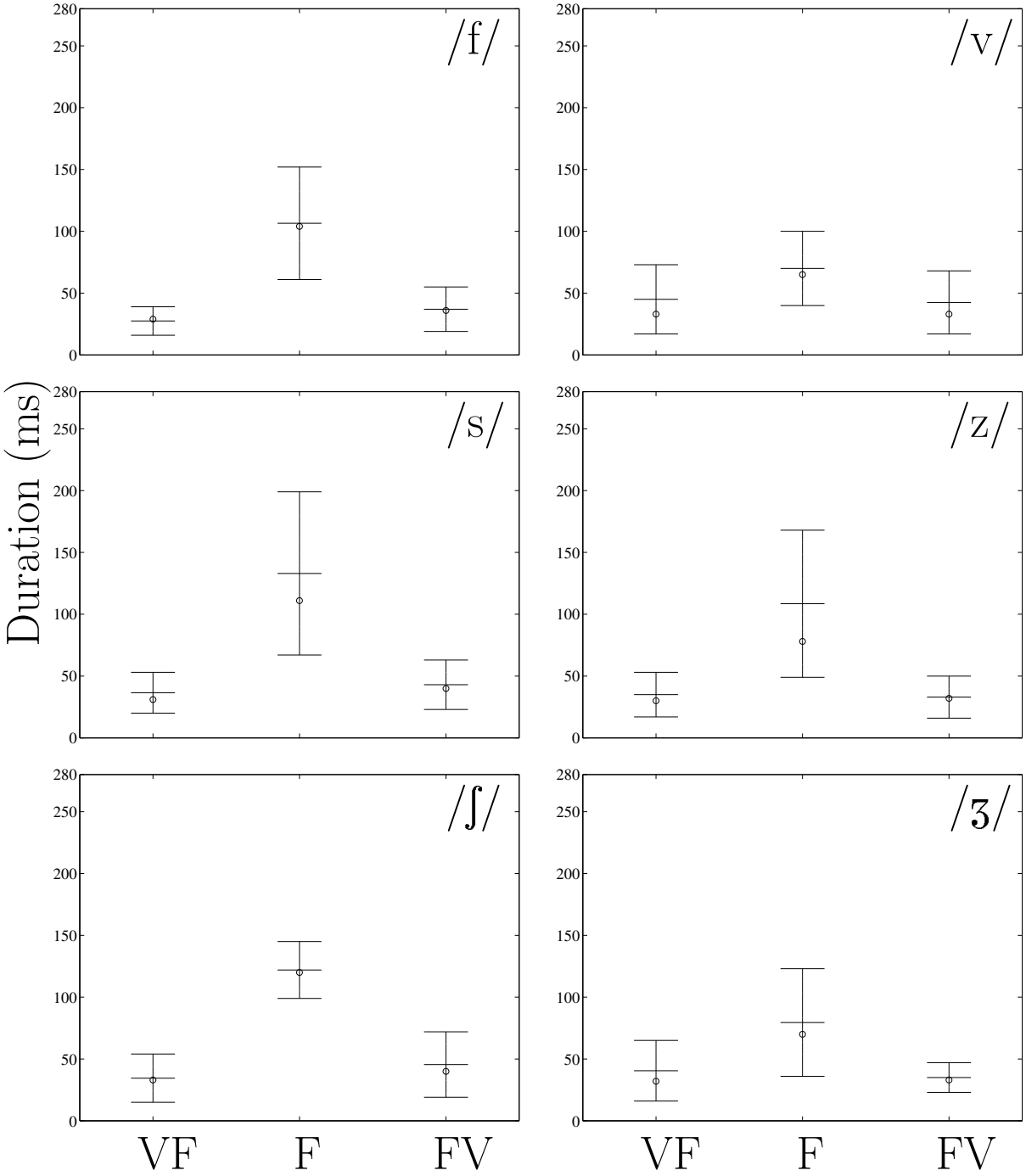


Figure 9: Range of values of duration of fricatives /f, v, s, z, ʃ, ʒ/, and of VF and FV transitions. o is the mean. Corpus 4 (Speaker CFGA).

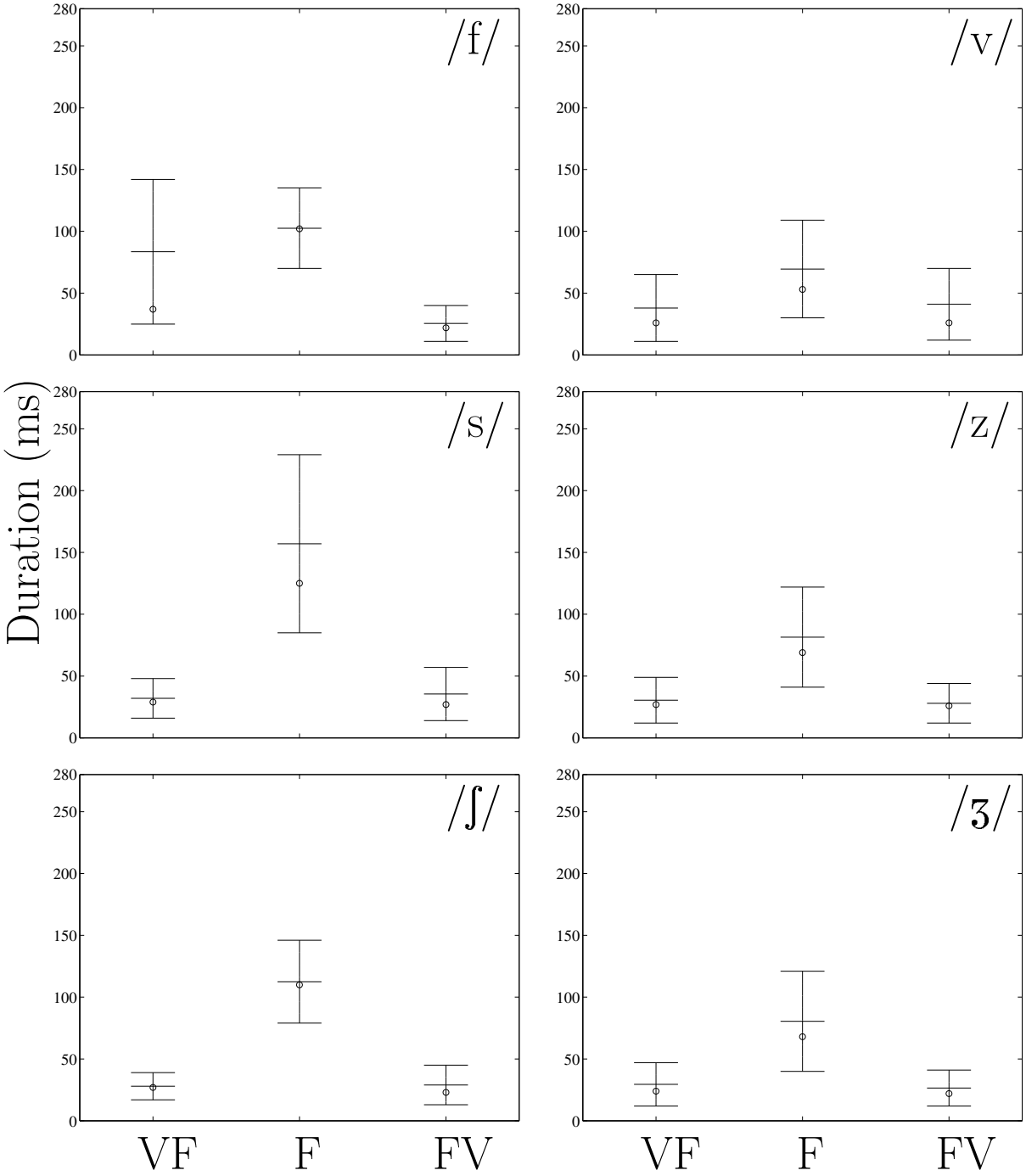


Figure 10: Range of values of duration of fricatives /f, v, s, z, ʃ, ʒ/, and of VF and FV transitions. o is the mean. Corpus 4 (Speaker ACC).

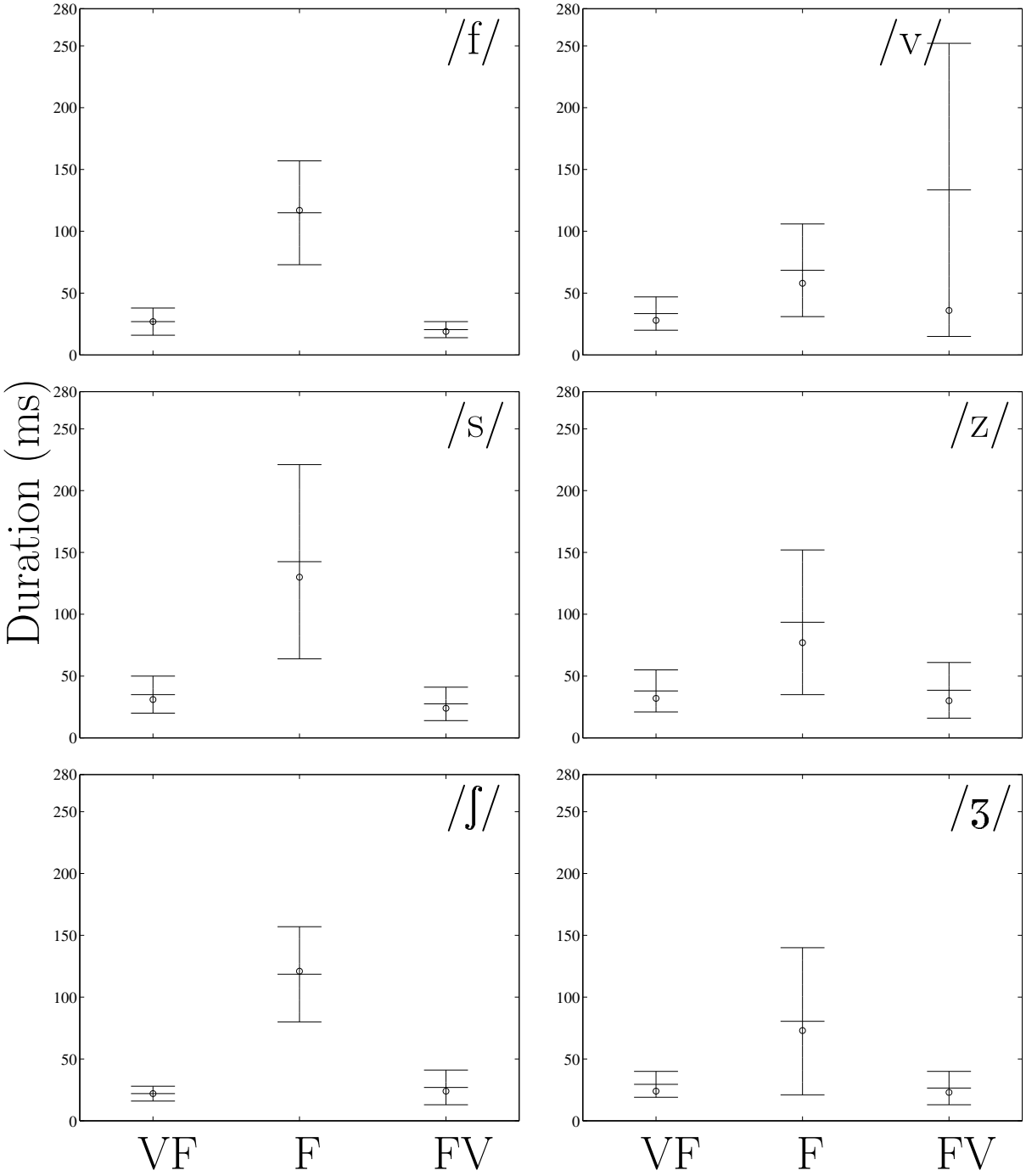


Figure 11: Range of values of duration of fricatives /f, v, s, z, ʃ, ʒ/, and of VF and FV transitions. o is the mean. Corpus 4 (Speaker ISSS).

Devoicing Results – Corpus 3

A complete inventory of devoiced, partially devoiced and voiced examples in Corpus 3 is presented for all four speakers (LMTJ, CFGA, ACC and ISSS) in Tables 3 to 6. Results showed that:

- 55% (70 out of 127) of the examples of fricative /v/ were totally devoiced;
- 74% (79 out of 107) of the examples of fricative /z/ were totally devoiced;
- 86% (92 out of 107) of the examples of fricative /ʒ/ were totally devoiced.

Most word-final fricative examples (93% – 55 out of 59) were totally devoiced, and the percentage of devoicing increased as the place of articulation moved posteriorly.

Table 3: Inventory of all cases of devoicing (using the manual criterion). Values given are in the form x/y, where x = number of devoiced, partially devoiced or voiced examples, and y = total number of examples. Corpus 3 (Speaker LMTJ).

	Word-Initial	Word-Medial	Word-Final	All Pos.	
/v/	6/14 (42.9%)	4/14 (28.6%)	8/9 (88.9%)	18/37 (48.7%)	Devoiced
	2/14 (14.3%)	2/14 (14.3%)	0	4/37 (10.8%)	Partially Devoiced
	6/14 (42.9%)	8/14 (57.1%)	1/9 (11.1%)	15/37 (40.5%)	Voiced
/z/	5/10 (50%)	12/17 (70.6%)	3/3 (100%)	20/30 (66.7%)	Devoiced
	2/10 (20%)	2/17 (11.8%)	0	4/30 (13.3%)	Partially Devoiced
	3/10 (30%)	3/17 (17.7%)	0	6/30 (20%)	Voiced
/ʒ/	7/10 (70%)	13/15 (86.7%)	4/5 (80%)	24/30 (80%)	Devoiced
	1/10 (10%)	0	0	1/30 (3.3%)	Partially Devoiced
	2/10 (20%)	2/15 (13.3%)	1/5 (20%)	5/30 (16.7%)	Voiced
All Fric.	18/34 (52.9%)	29/46 (63%)	15/17 (88.2%)	62/97 (63.9%)	Devoiced
	5/34 (14.7%)	4/46 (8.7%)	0	9/97 (9.3%)	Partially Devoiced
	11/34 (32.4%)	13/46 (28.3%)	2/17 (11.8%)	26/97 (26.8%)	Voiced

Table 4: Inventory of all cases of devoicing (using the manual criterion). Values given are in the form x/y, where x = number of devoiced, partially devoiced or voiced examples, and y = total number of examples. Corpus 3 (Speaker CFGA).

	Word-Initial	Word-Medial	Word-Final	All Pos.	
/v/	7/11 (63.6%)	8/12 (66.7%)	7/7 (100%)	22/30 (73.3%)	Devoiced
	0	2/12 (16.7%)	0	2/30 (6.7%)	Partially Devoiced
	4/11 (36.4%)	2/12 (16.7%)	0	6/30 (20%)	Voiced
/z/	8/10 (80%)	12/14 (85.7%)	2/3 (66.7%)	22/27 (81.5%)	Devoiced
	1/10 (10%)	2/14 (14.3%)	0	3/27 (11.1%)	Partially Devoiced
	1/10 (10%)	0	1/3 (33.3%)	2/27 (7.4%)	Voiced
/ʒ/	8/10 (80%)	12/13 (92.3%)	4/4 (100%)	24/27 (88.9%)	Devoiced
	1/10 (10%)	0	0	1/27 (3.7%)	Partially Devoiced
	1/10 (10%)	1/13 (7.7%)	0	2/27 (7.4%)	Voiced
All Fric.	23/31 (74.2%)	32/39 (82.1%)	13/14 (92.9%)	68/84 (81%)	Devoiced
	2/31 (6.5%)	4/39 (10.3%)	0	6/84 (7.1%)	Partially Devoiced
	6/31 (19.4%)	3/39 (7.7%)	1/14 (7.1%)	10/84 (11.9%)	Voiced

Table 5: Inventory of all cases of devoicing (using the manual criterion). Values given are in the form x/y, where x = number of devoiced, partially devoiced or voiced examples, and y = total number of examples. Corpus 3 (Speaker ACC).

	Word-Initial	Word-Medial	Word-Final	All Pos.	
/v/	4/11 (36.4%)	6/12 (50%)	7/7 (100%)	17/30 (56.7%)	Devoiced
	2/11 (18.2%)	2/12 (16.7%)	0	4/30 (13.3%)	Partially Devoiced
	5/11 (45.5%)	4/12 (33.3%)	0	9/30 (30%)	Voiced
/z/	4/10 (40%)	8/12 (66.7%)	3/3 (100%)	15/25 (60%)	Devoiced
	4/10 (40%)	3/12 (25%)	0	7/25 (28%)	Partially Devoiced
	2/10 (20%)	1/12 (8.3%)	0	3/25 (12%)	Voiced
/ʒ/	9/10 (90%)	11/11 (100%)	3/4 (75%)	23/25 (92%)	Devoiced
	1/10 (10%)	0	1/4 (25%)	2/25 (8%)	Partially Devoiced
	0	0	0	0	Voiced
All Fric.	17/31 (54.8%)	25/35 (71.4%)	13/14 (92.9%)	55/80 (68.8%)	Devoiced
	7/31 (22.6%)	5/35 (14.3%)	1/14 (7.1%)	13/80 (16.3%)	Partially Devoiced
	7/31 (22.6%)	5/35 (14.3%)	0	12/80 (15%)	Voiced

Table 6: Inventory of all cases of devoicing (using the manual criterion). Values given are in the form x/y, where x = number of devoiced, partially devoiced or voiced examples, and y = total number of examples. Corpus 3 (Speaker ISSS).

	Word-Initial	Word-Medial	Word-Final	All Pos.	
/v/	1/11 (9.1%)	5/12 (41.7%)	7/7 (100%)	13/30 (43.3%)	Devoiced
	4/11 (36.4%)	0	0	4/30 (13.3%)	Partially Devoiced
	6/11 (54.6%)	7/12 (58.3%)	0	13/30 (43.3%)	Voiced
/z/	8/10 (80%)	11/12 (91.7%)	3/3 (100%)	22/25 (88%)	Devoiced
	0	0	0	0	Partially Devoiced
	2/10 (20%)	1/12 (8.3%)	0	3/25 (12%)	Voiced
/ʒ/	8/10 (80%)	9/11 (81.8%)	4/4 (100%)	21/25 (84%)	Devoiced
	2/10 (20%)	1/11 (9.1%)	0	3/25 (12%)	Partially Devoiced
	0	1/11 (9.1%)	0	1/25 (4%)	Voiced
All Fric.	17/31 (54.8%)	25/35 (71.4%)	14/14 (100%)	56/80 (70%)	Devoiced
	6/31 (19.4%)	1/35 (2.9%)	0	7/80 (8.8%)	Partially Devoiced
	8/31 (25.8%)	9/35 (25.7%)	0	17/80 (21.3%)	Voiced

Devoicing Results – Corpus 4

The results shown in Tables 7 to 10 indicate that in Corpus 4:

- 44% (77 out of 177) of the examples of fricative /v/ were totally devoiced;
- 78% (86 out of 110) of the examples of fricative /z/ were totally devoiced;
- 71% (89 out of 126) of the examples of fricative /ʒ/ were totally devoiced.

The Corpus 4 fricatives devoiced mostly word-finally, but less often than in Corpus 3, as shown in Figure 12:

- word-initial – 97/157 = 62%
  (in Corpus 3, 75/127 = 59%);
- word-medial – 111/195 = 57%
  (in Corpus 3, 111/120 = 93%);
- word-final – 44/61 = 72%
  (in Corpus 3, 55/59 = 93%).

Devoicing Results – Corpus 4

Some of the fricatives that have been classified as word-final are followed by voiced phonemes. Some of the words that follow these fricatives even start with a vowel.

This might account for the lower word-final average percentage of devoicing in Corpus 4 when compared with Corpus 3.

Indeed, some voiceless fricatives become voiced in Corpus 4, likely as a result of cross-word coarticulation: eleven tokens of word-final /ʃ/ were produced as [ʒ] by speakers LMTJ and ACC when followed by a word starting with a voiced phoneme ([d] or [m]).

There is a slightly lower number of devoiced examples of /ʒ/ than /z/, which contradicts the very clear results of Corpus 3 fricatives (in which the percentage of devoiced examples decreases as the place of articulation moves anteriorly).

One possible explanation could be that /ʒ/ is produced in a more anterior place in continuous speech than in isolated word production.

This hypothesis can only be confirmed with additional articulatory data, which is planned as future work.

Table 7: Inventory of all cases of devoicing (using the manual criterion). Values given are in the form x/y, where x = number of devoiced, partially devoiced or voiced examples, and y = total number of examples. Corpus 4 (Speaker LMTJ).

	Word-Initial	Word-Medial	Word-Final	All Pos.	
/v/	5/14 (35.7%)	7/18 (38.9%)	1/2 (50%)	13/34 (38.2%)	Devoiced
	1/14 (7.1%)	3/18 (16.7%)	0	4/34 (11.8%)	Partially Devoiced
	8/14 (57.1%)	8/18 (44.4%)	1/2 (50%)	17/34 (50%)	Voiced
/z/	5/8 (62.5%)	7/10 (70%)	4/4 (100%)	16/22 (72.7%)	Devoiced
	1/8 (12.5%)	2/10 (20%)	0	3/22 (13.6%)	Partially Devoiced
	2/8 (25%)	1/10 (10%)	0	3/22 (13.6%)	Voiced
/ʒ/	7/8 (87.5%)	5/9 (55.6%)	3/5 (60.0%)	15/22 (68.2%)	Devoiced
	1/8 (12.5%)	3/9 (33.3%)	0	4/22 (18.2%)	Partially Devoiced
	0	1/9 (11.1%)	2/5 (40%)	3/22 (13.6%)	Voiced
All Fric.	17/30 (56.7%)	19/37 (51.4%)	8/11 (72.7%)	44/78 (56.4%)	Devoiced
	3/30 (10%)	8/37 (21.6%)	0	11/78 (14.1%)	Partially Devoiced
	10/30 (33.3%)	10/37 (27%)	3/11 (27.3%)	23/78 (29.5%)	Voiced

Table 8: Inventory of all cases of devoicing (using the manual criterion). Values given are in the form x/y, where x = number of devoiced, partially devoiced or voiced examples, and y = total number of examples. Corpus 4 (Speaker CFGA).

	Word-Initial	Word-Medial	Word-Final	All Pos.	
/v/	11/24 (45.8%)	9/24 (37.5%)	2/6 (33.3%)	22/54 (40.7%)	Devoiced
	4/24 (16.7%)	6/24 (25%)	3/6 (50%)	13/54 (24.1%)	Partially Devoiced
	9/24 (37.5%)	9/24 (37.5%)	1/6 (16.7%)	19/54 (35.2%)	Voiced
/z/	10/12 (83.3%)	12/16 (75%)	5/5 (100%)	27/33 (81.8%)	Devoiced
	1/12 (8.3%)	2/16 (12.5%)	0	3/33 (9.1%)	Partially Devoiced
	1/12 (8.3%)	2/16 (12.5%)	0	3/33 (9.1%)	Voiced
/ʒ/	9/12 (75%)	10/15 (66.7%)	11/12 (91.7%)	30/39 (76.9%)	Devoiced
	1/12 (8.3%)	2/15 (13.3%)	1/12 (8.3%)	4/39 (10.3%)	Partially Devoiced
	2/12 (16.7%)	3/15 (20%)	0	5/39 (12.8%)	Voiced
All Fric.	30/48 (62.5%)	31/55 (56.4%)	18/23 (78.3%)	79/126 (62.7%)	Devoiced
	6/48 (12.5%)	10/55 (18.2%)	4/23 (17.4%)	20/126 (15.9%)	Partially Devoiced
	12/48 (25%)	14/55 (25.5%)	1/23 (4.4%)	27/126 (21.4%)	Voiced

Table 9: Inventory of all cases of devoicing (using the manual criterion). Values given are in the form x/y, where x = number of devoiced, partially devoiced or voiced examples, and y = total number of examples. Corpus 4 (Speaker ACC).

	Word-Initial	Word-Medial	Word-Final	All Pos.	
/v/	14/23 (60.9%)	12/26 (46.2%)	3/4 (75%)	29/53 (54.7%)	Devoiced
	1/23 (4.4%)	2/26 (7.7%)	0	3/53 (5.7%)	Partially Devoiced
	8/23 (34.8%)	12/26 (46.2%)	1/4 (25%)	21/53 (39.6%)	Voiced
/z/	10/12 (83.3%)	16/19 (84.2%)	2/2 (100%)	28/33 (84.9%)	Devoiced
	2/12 (16.7%)	0	0	2/33 (6.1%)	Partially Devoiced
	0	3/19 (15.8%)	0	3/33 (9.1%)	Voiced
/ʒ/	9/12 (75%)	13/18 (72.2%)	3/9 (33.3%)	25/39 (64.1%)	Devoiced
	3/12 (25%)	3/18 (16.7%)	3/9 (33.3%)	9/39 (23.1%)	Partially Devoiced
	0	2/18 (11.1%)	3/9 (33.3%)	5/39 (12.8%)	Voiced
All Fric.	33/47 (70.2%)	41/63 (65.1%)	8/15 (53.3%)	82/125 (65.6%)	Devoiced
	6/47 (12.8%)	15/63 (7.9%)	3/15 (20%)	14/125 (11.2%)	Partially Devoiced
	8/47 (17%)	17/63 (27%)	4/15 (26.7%)	29/125 (23.2%)	Voiced

Table 10: Inventory of all cases of devoicing (using the manual criterion). Values given are in the form x/y, where x = number of devoiced, partially devoiced or voiced examples, and y = total number of examples. Corpus 4 (Speaker ISSS).

	Word-Initial	Word-Medial	Word-Final	All Pos.	
/v/	6/16 (37.5%)	5/16 (31.3%)	2/4 (50%)	13/36 (36.1%)	Devoiced
	2/16 (12.5%)	1/16 (6.3%)	0	3/36 (8.3%)	Partially Devoiced
	8/16 (50%)	10/16 (62.5%)	2/4 (50%)	20/36 (55.6%)	Voiced
/z/	6/8 (75%)	7/12 (58.3%)	2/2 (100%)	15/22 (68.2%)	Devoiced
	1/8 (12.5%)	3/12 (25%)	0	4/22 (18.2%)	Partially Devoiced
	1/8 (12.5%)	2/12 (16.7%)	0	3/22 (13.6%)	Voiced
/ʒ/	5/8 (62.5%)	8/12 (66.7%)	6/6 (100%)	19/26 (73.1%)	Devoiced
	3/8 (37.5%)	3/12 (25%)	0	6/26 (23.1%)	Partially Devoiced
	0	1/12 (8.3%)	0	1/26 (3.9%)	Voiced
All Fric.	17/32 (53.1%)	20/40 (50%)	10/12 (83.3%)	47/84 (56%)	Devoiced
	6/32 (18.8%)	7/40 (17.5%)	0	13/84 (15.5%)	Partially Devoiced
	9/32 (28.1%)	13/40 (32.5%)	2/12 (16.7%)	24/84 (28.6%)	Voiced

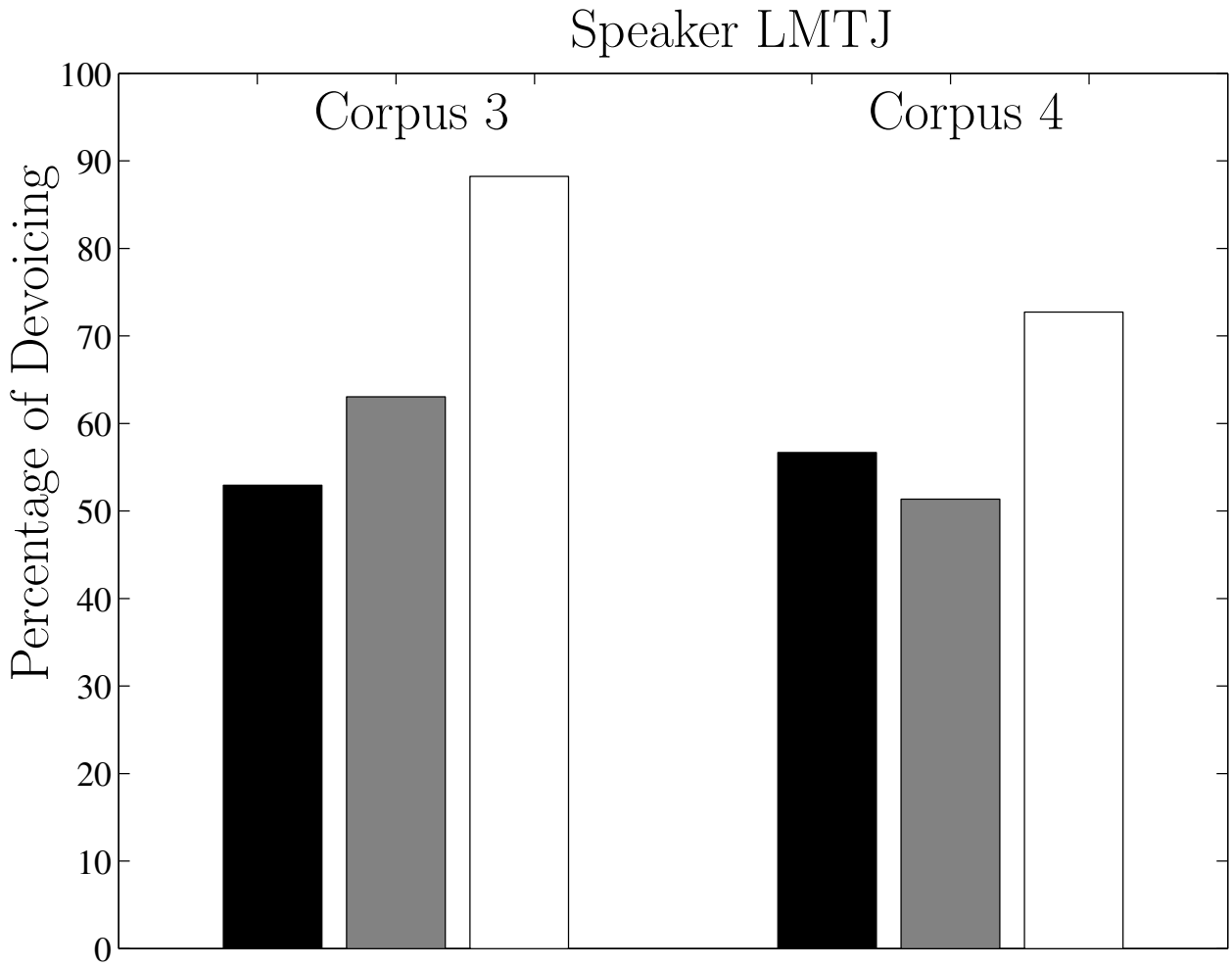
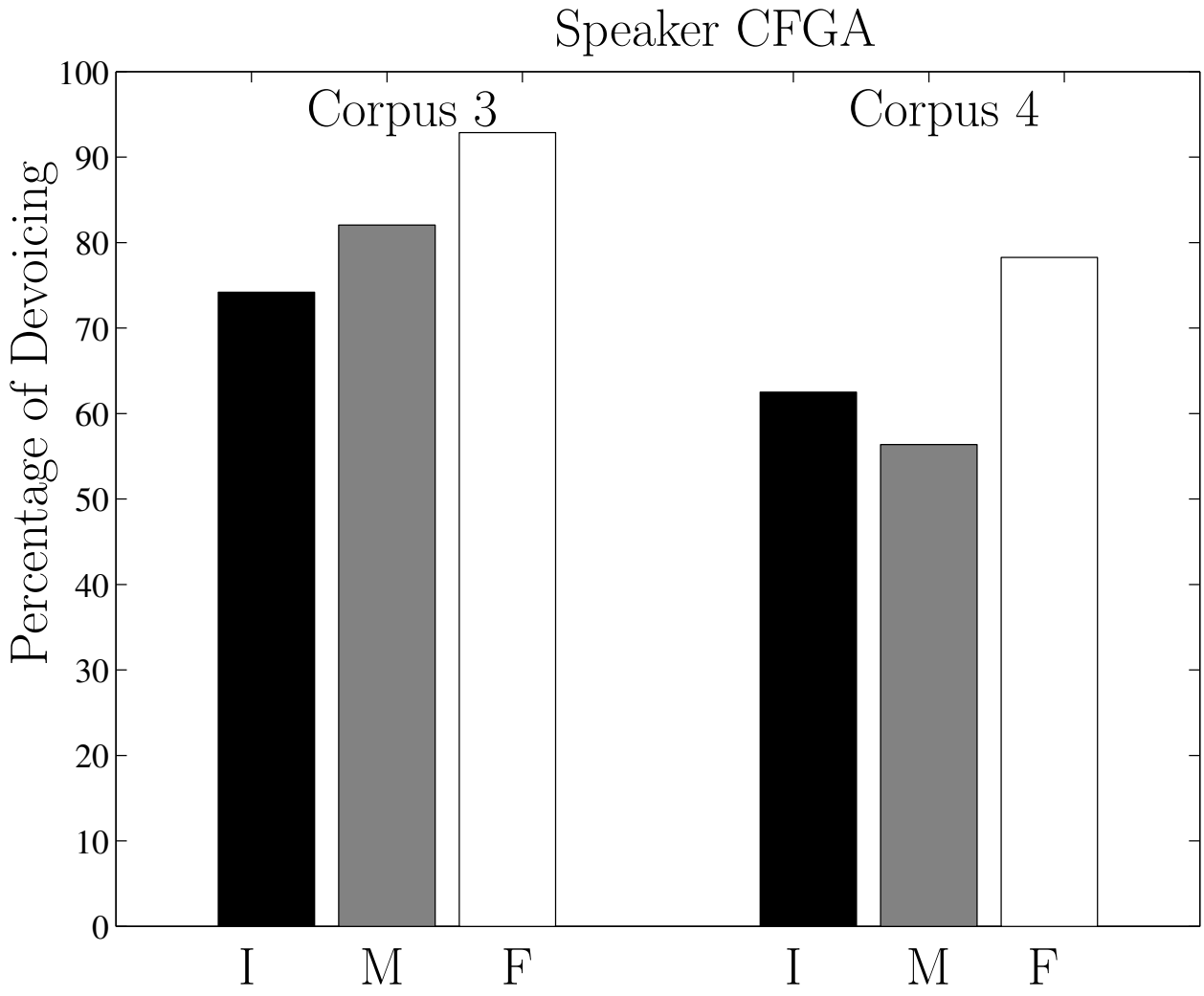
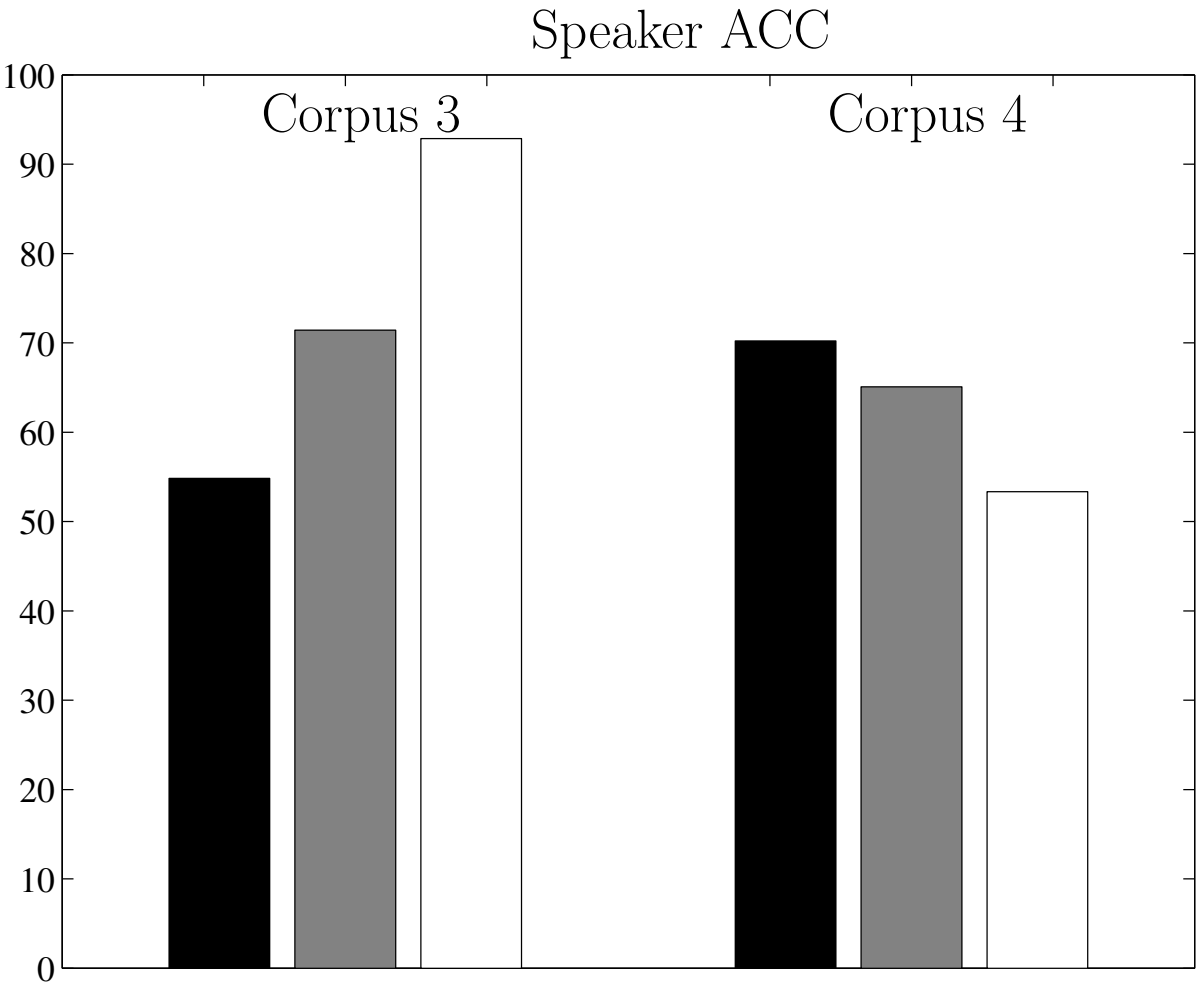
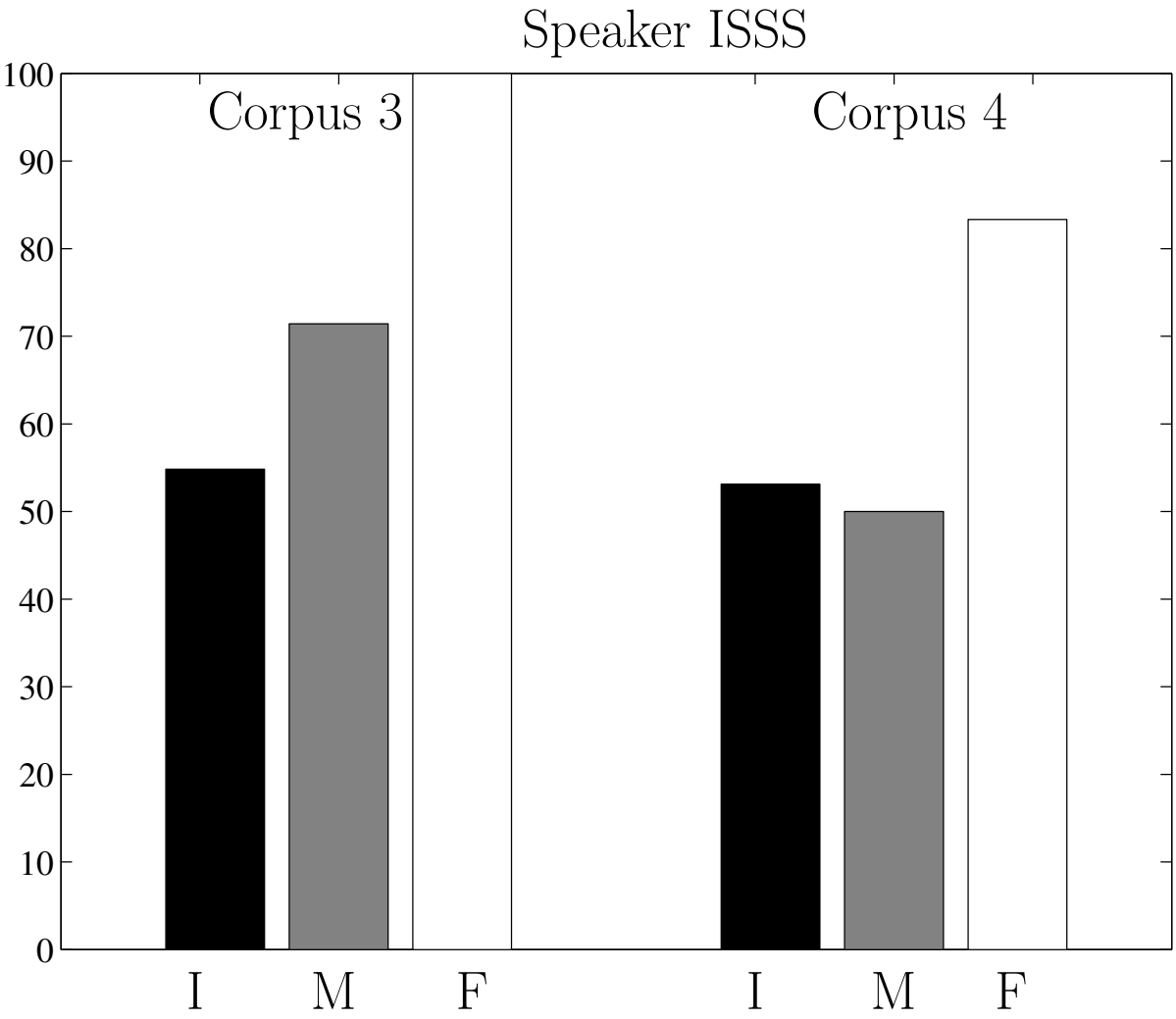


Figure 12: Percentage of total devoicing by position in word, word-initial (I), word-medial (M) and word-final (F), fricatives /v, z, ʒ/ combined.


Speaker CFGA
Corpus 3
Corpus 4
Percentage of Devoicing
100
90
80
70
60
50
40
30
20
10
0
I
M
F
I
M
F


Speaker ACC
Corpus 3
Corpus 4
100
90
80
70
60
50
40
30
20
10
0

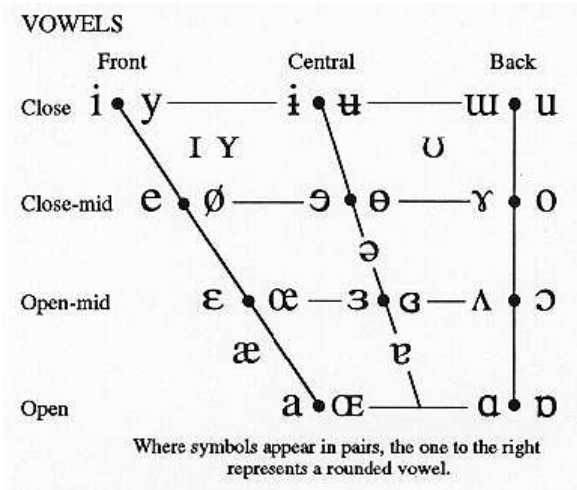

Speaker ISSS
Corpus 3
Corpus 4
100
90
80
70
60
50
40
30
20
10
0
I
M
F
I
M
F

Influence of Vowel Context on Devoicing

In a preliminary study, the influence of vowel context on devoicing was analysed for a limited number of words from Corpus 3 (Speakers LMTJ and ACC) that follow the pattern /V_1FV_2/.

Results are shown in Figures 13 and 14, where V_1 and V_2 are vowels which belong to one of the groups:

- group 1 – /i, ɨ, e/;
- group 2 – /ɛ, ɐ, a/;
- group 3 – /ɔ, o, u/.



From IPA (1999)

Results show that there is no effect of vowel context on devoicing of Portuguese fricatives, just as previously observed by Smith (1997) and Pirello et al. (1997) for the English language. Therefore, this study has not been extended to speakers CFGA and ISSS.

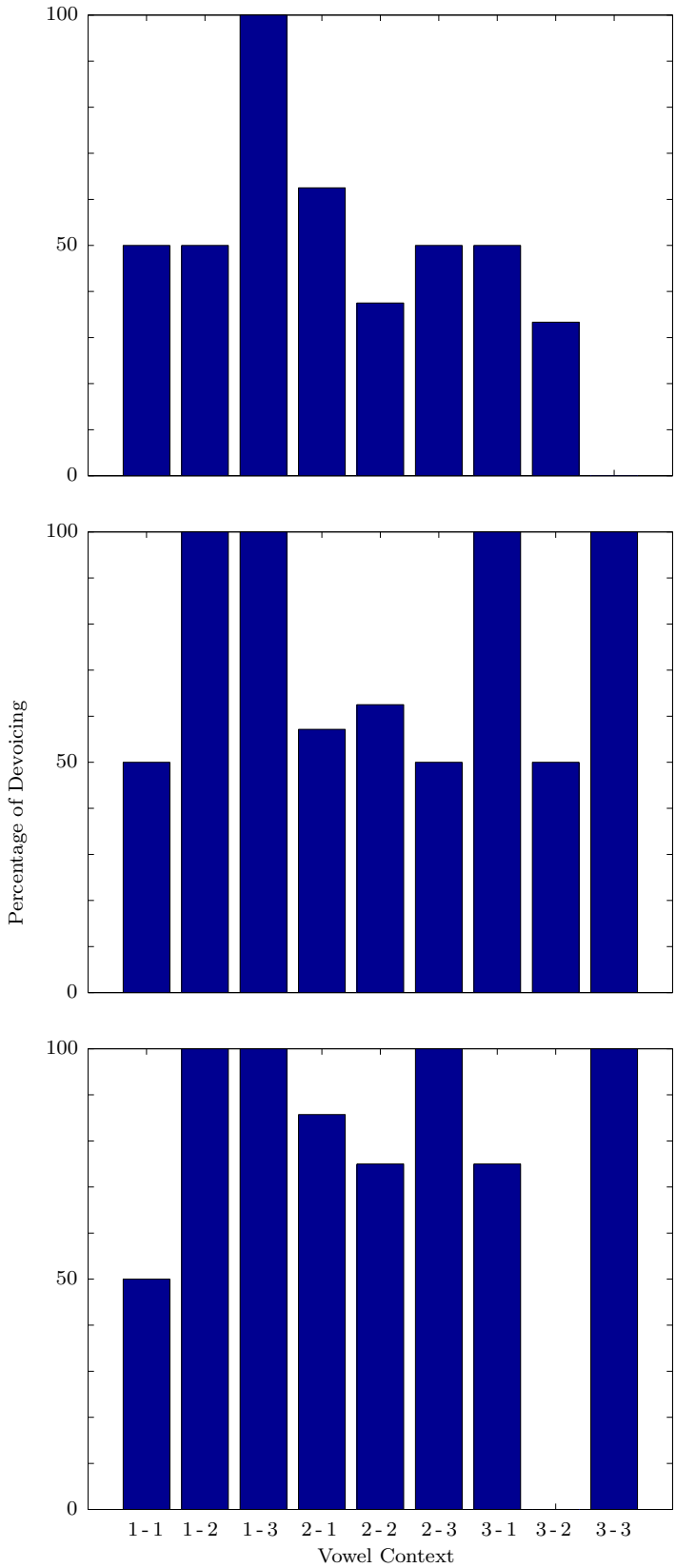


Figure 13: Percentage of totally devoiced fricatives in different vowel contexts. Top: fricative /v/; Middle: fricative /z/; Bottom: fricative /ʒ/. There were no examples of fricative /v/ in 3-3 vowel context. Corpus 3 (Speaker LMTJ).

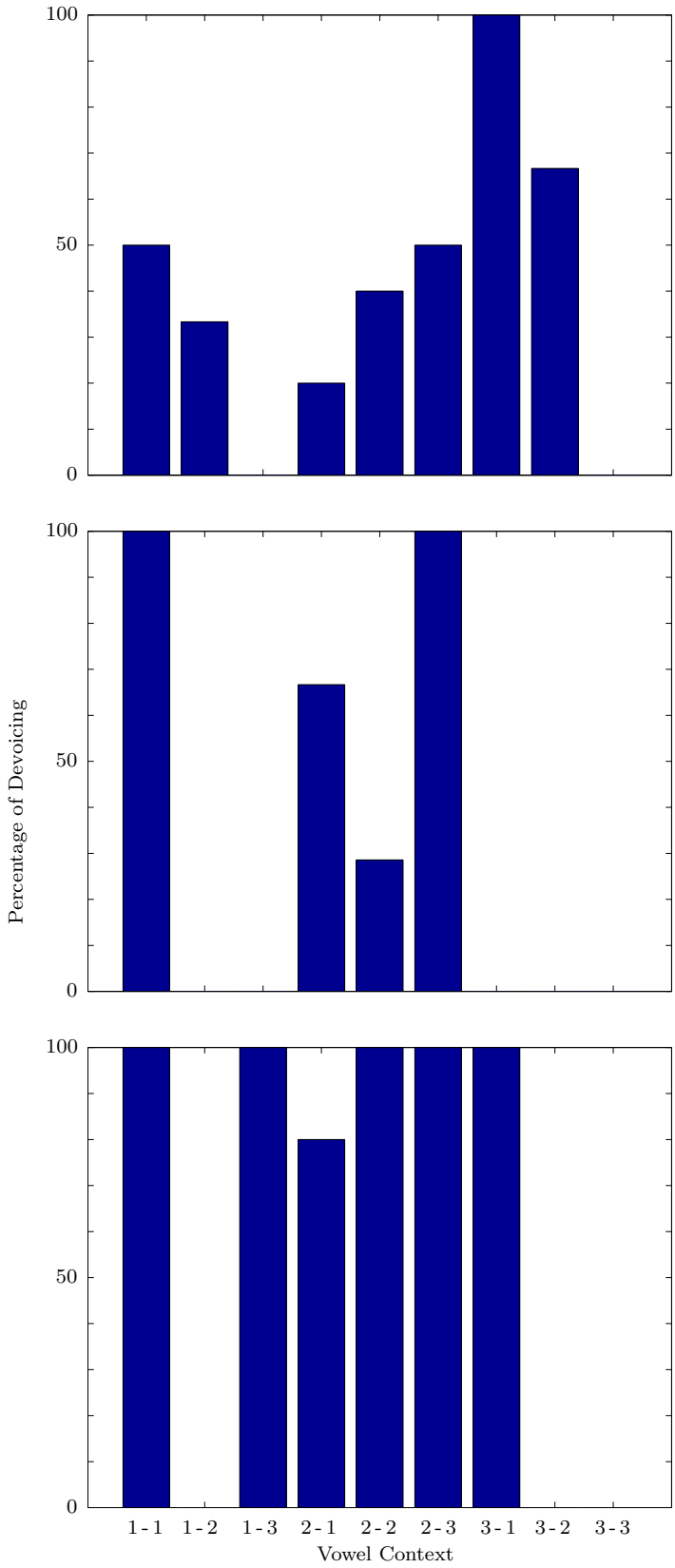


Figure 14: Percentage of totally devoiced fricatives in different vowel contexts. Top: fricative /v/; Middle: fricative /z/; Bottom: fricative /ʒ/. There were no examples of fricative /v/ in 1-3 and 3-3 vowel contexts, of fricative /z/ in 1-3, 3-1 and 3-3 vowel contexts, and of fricative /ʒ/ in 1-2, 3-2 and 3-3 vowel contexts. Corpus 3 (Speaker ACC).

Duration and Devoicing Correlations

Evidence of correlation between duration and devoicing has been reported by Smith (1997) in a study of four American English speakers. The mean duration of /s/ was 101 ms and for /z/ the mean durations grouped into voicing categories were: 81 ms (devoiced), 61 ms (partially devoiced) and 64 ms (voiced).

However, in a previous study by Crystal and House (1988), results show no clear correlation between devoicing and duration, as can be seen from the overlap of the probability density distribution curves of duration shown in Figure 1.

In the present study there is also no consistent pattern between the percentage of devoiced tokens and the average duration of the nine tokens used to ensemble average the spectra of fricatives from Corpus 2, as shown in Figure 15.

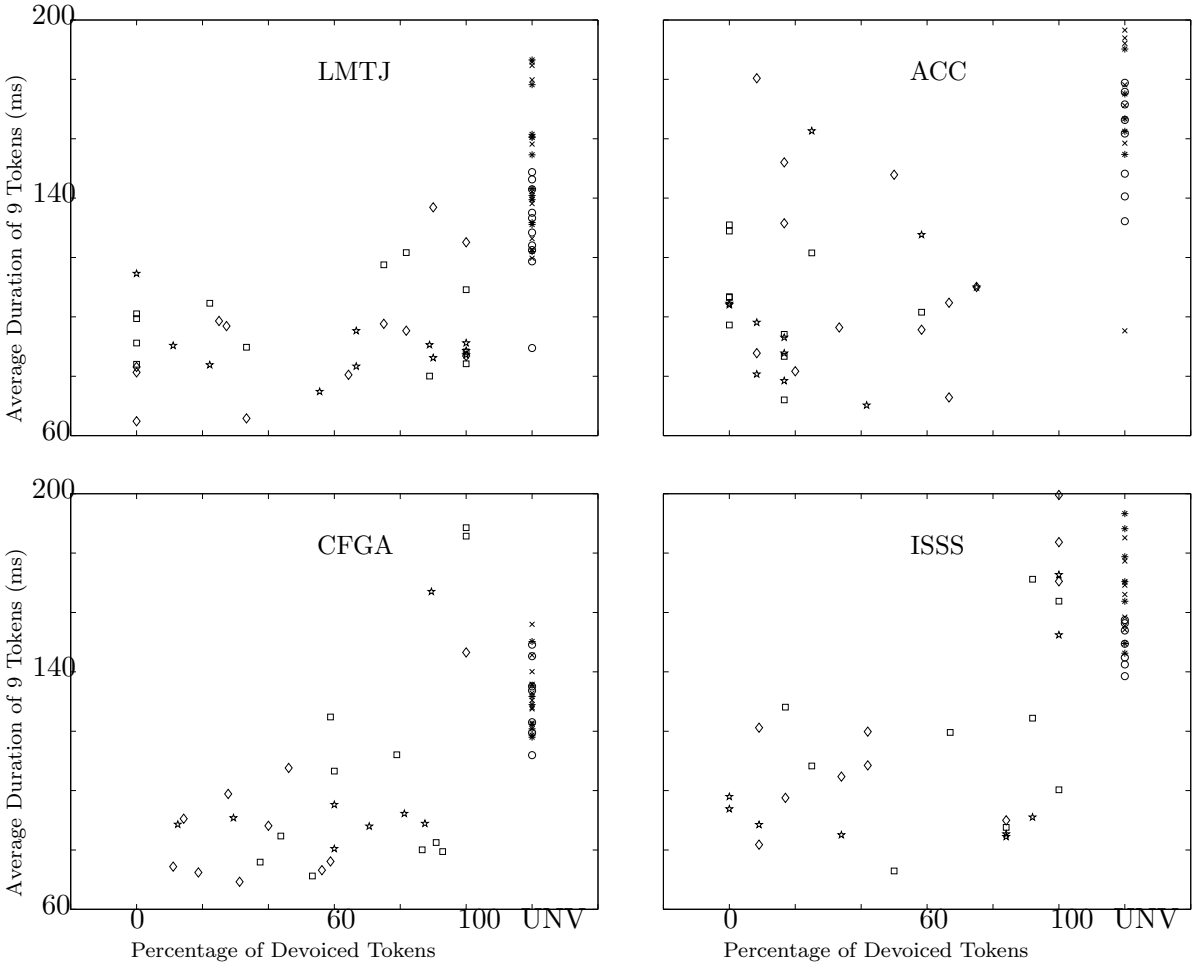
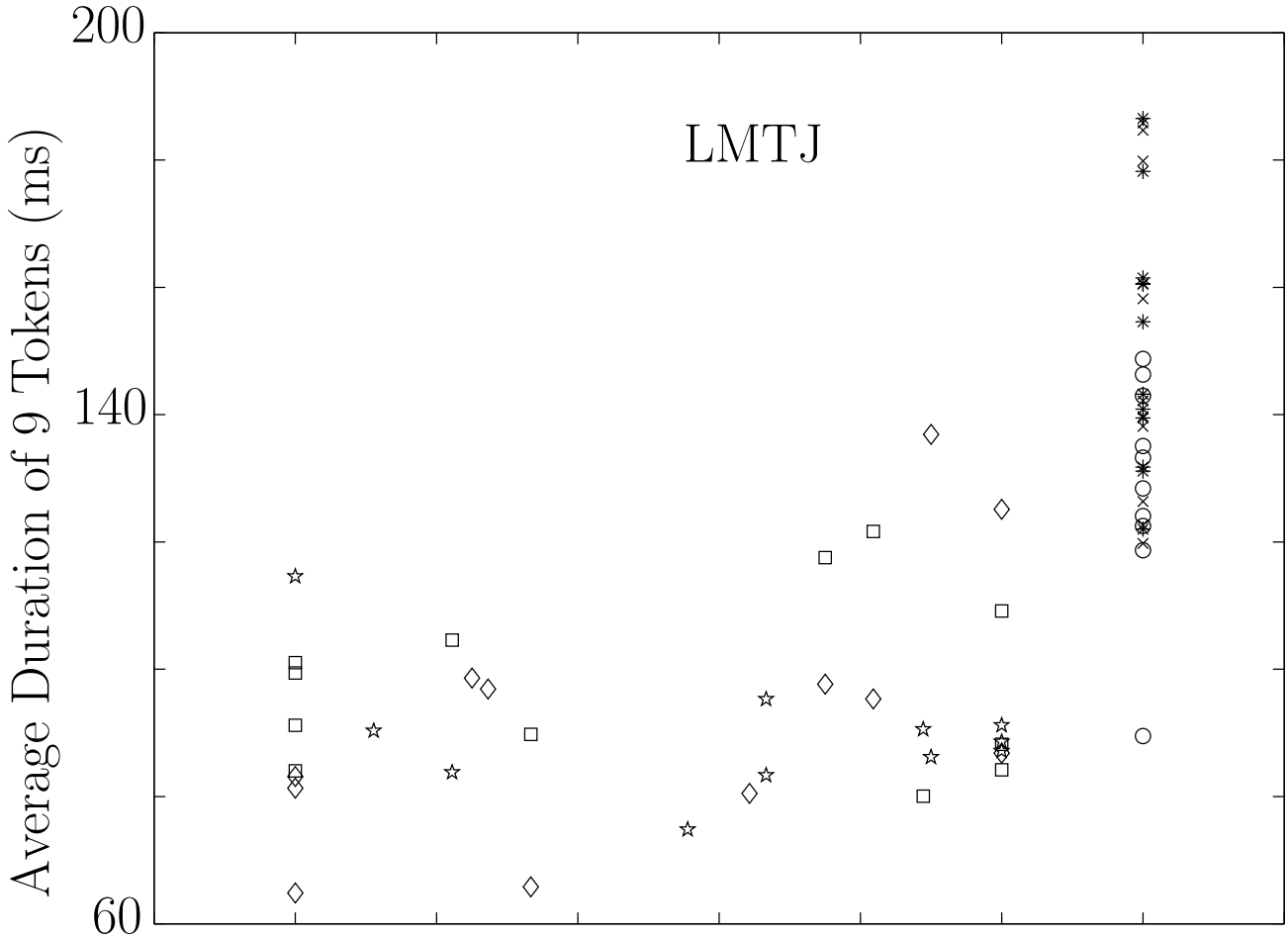
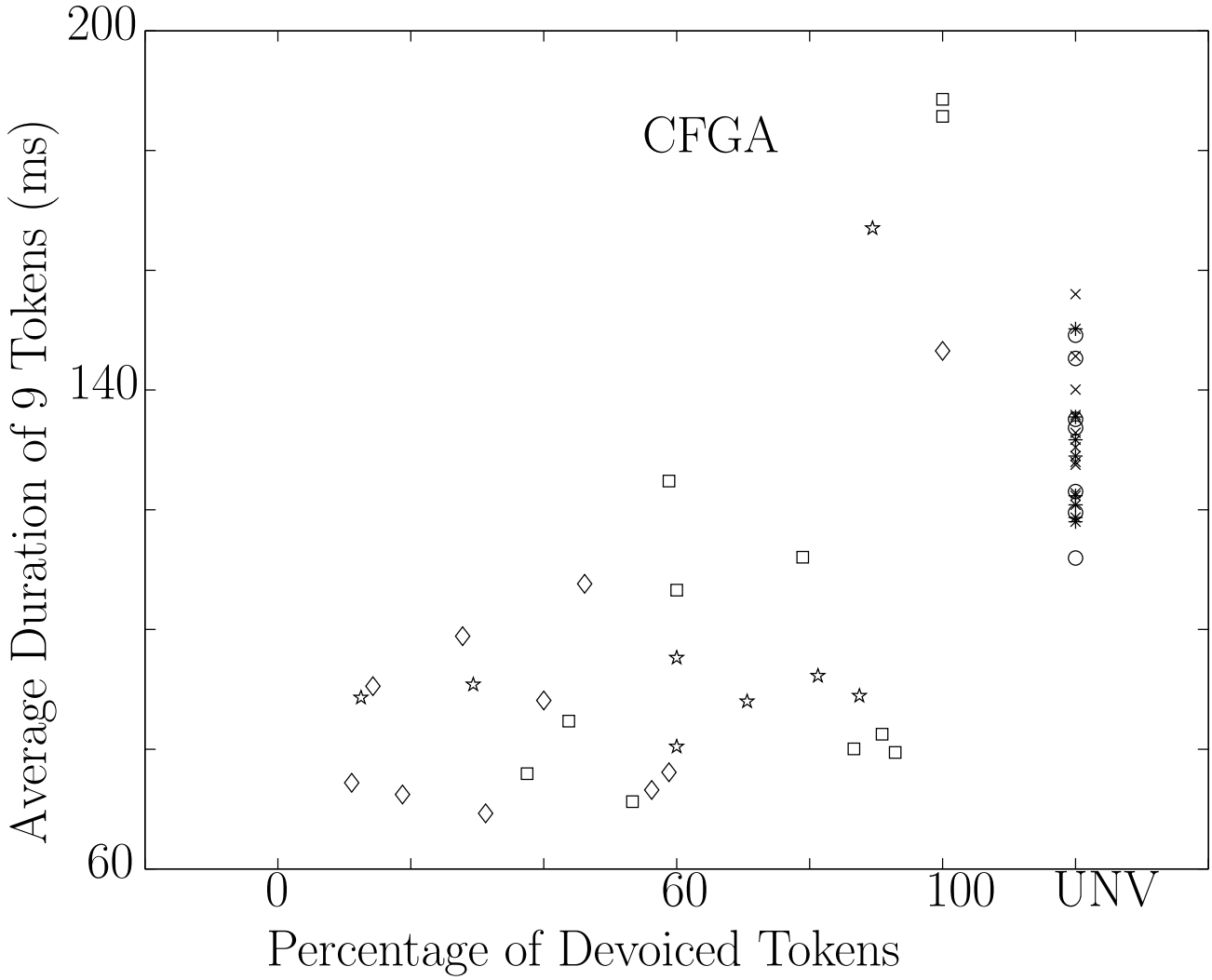
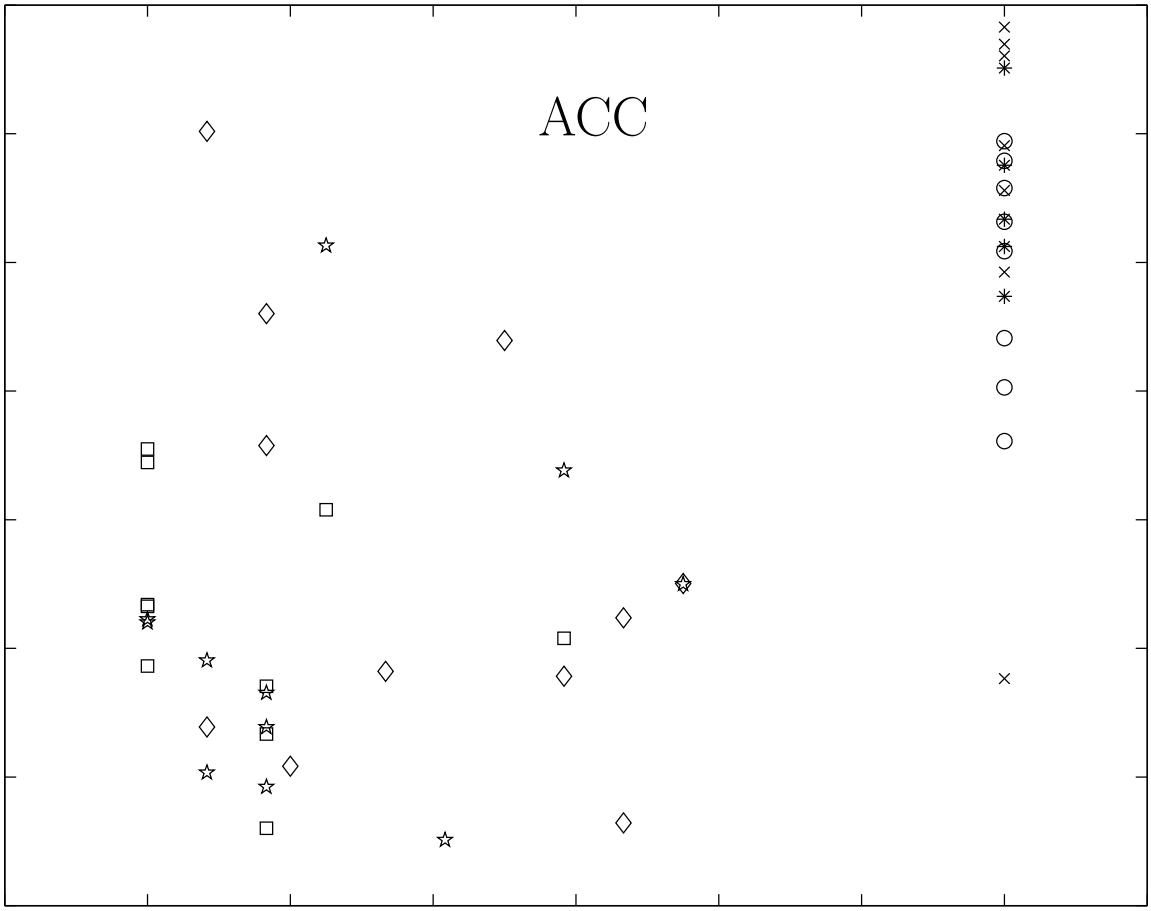
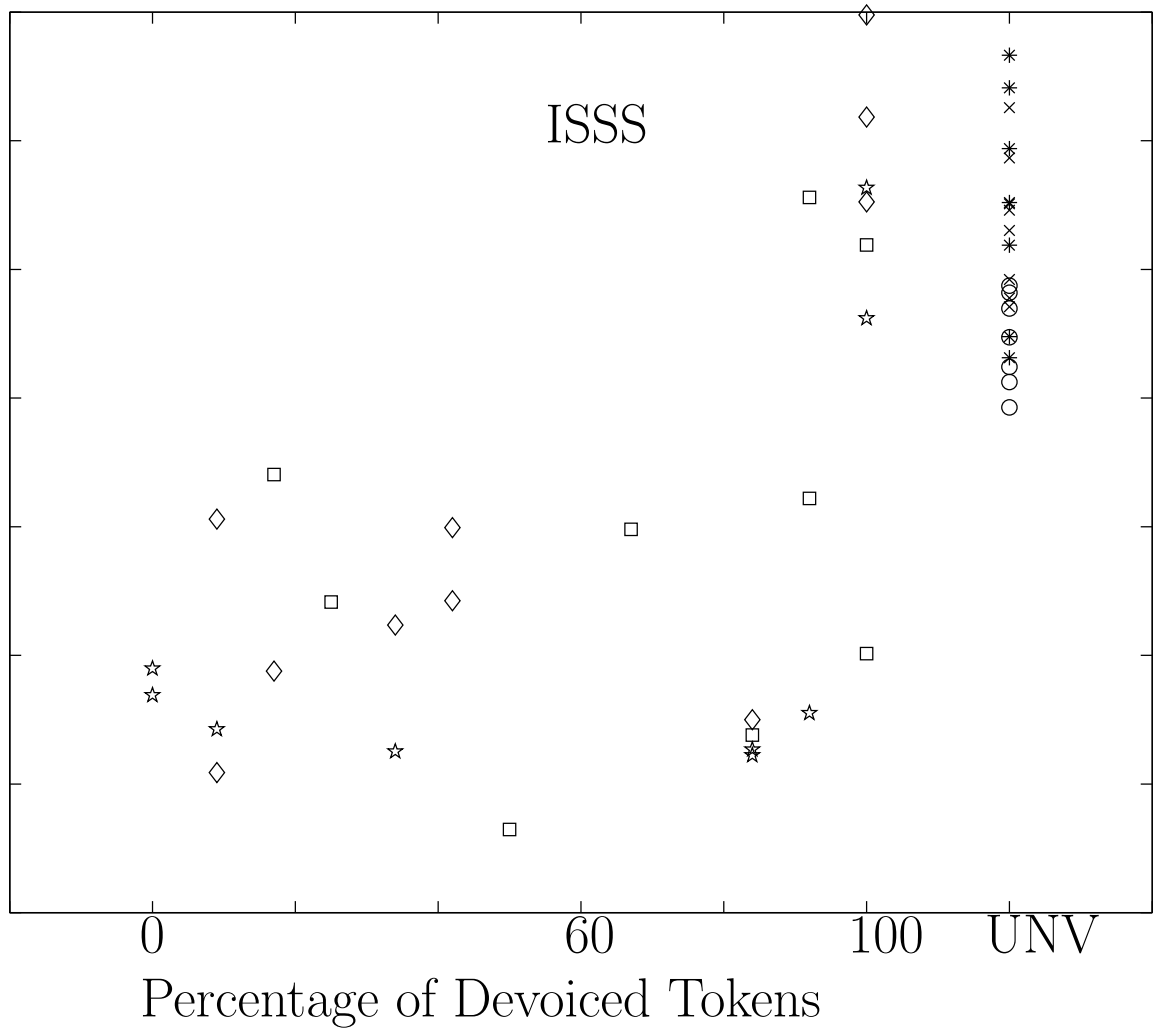


Figure 15: Relationship between the percentage of devoiced tokens and duration in Corpus 2. The graphs also include the duration of unvoiced (UNV) fricatives. ○ – /f/, ⋆ – /v/, ∗ – /s/, ◇ – /z/, × – /ʃ/ and □ – /ʒ/.


LMTJ
Average Duration of 9 Tokens (ms)
200
140
60


CFGA
200
140
60
0
60
100
UNV
Average Duration of 9 Tokens (ms)
Percentage of Devoiced Tokens


ACC


ISSS
0
60
100
UNV
Percentage of Devoiced Tokens

Duration and Devoicing Correlations
–
Corpus 3

Figures 16 to 19 present the average durations of the fricatives /v, z, ʒ/ from Corpus 3 and relate them to devoicing. Only totally devoiced examples were considered.

When a fricative devoices its FV transition, duration diminishes and the duration of fricatives increases (for a few examples the duration remains the same). The VF transition duration is fairly stable.

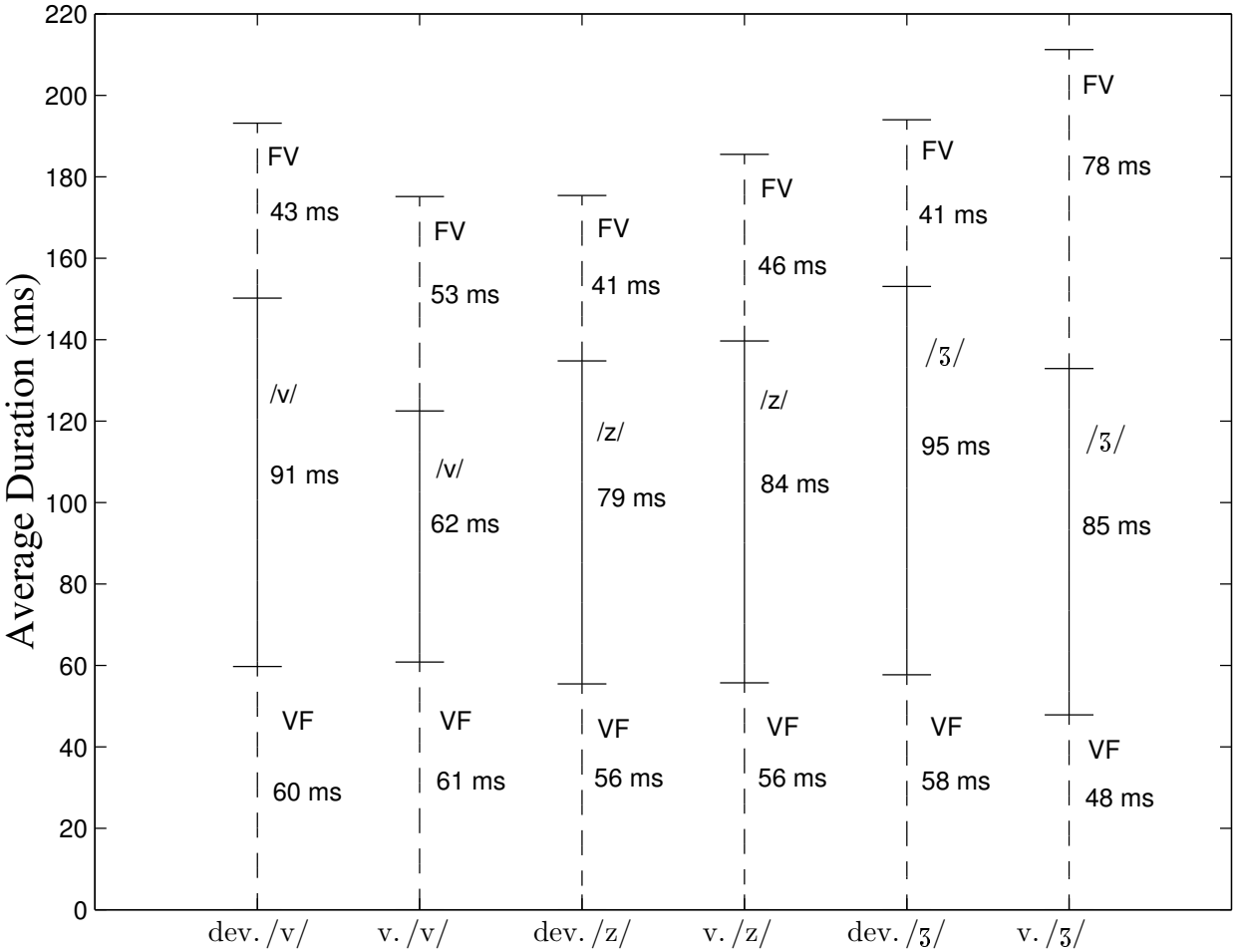


Figure 16: Average duration of voiced and devoiced examples of fricatives /v, z, ʒ/, and of VF and FV transitions in Corpus 3 (Speaker LMTJ).

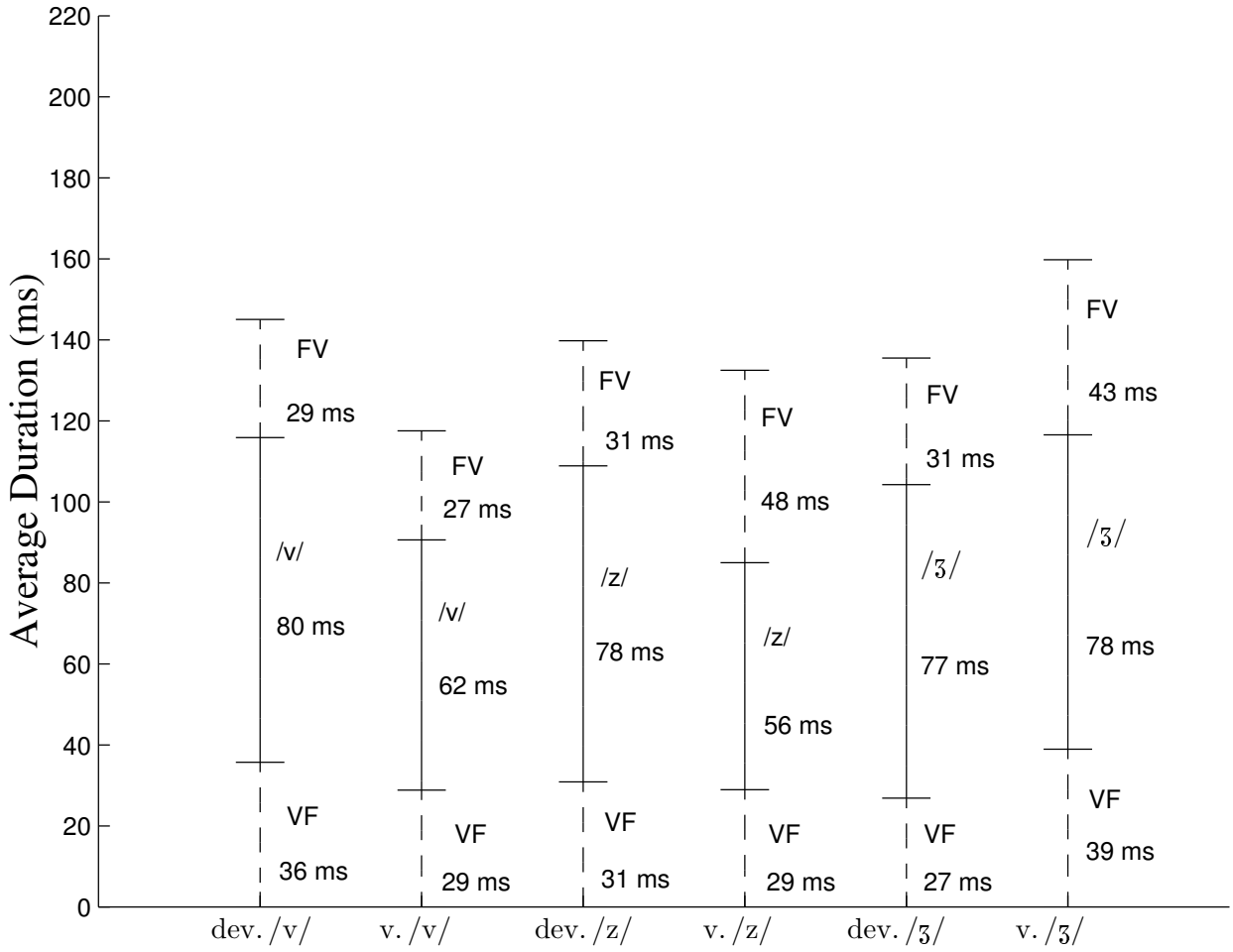


Figure 17: Average duration of voiced and devoiced examples of fricatives /v, z, ʒ/, and of VF and FV transitions in Corpus 3 (Speaker CFGA).

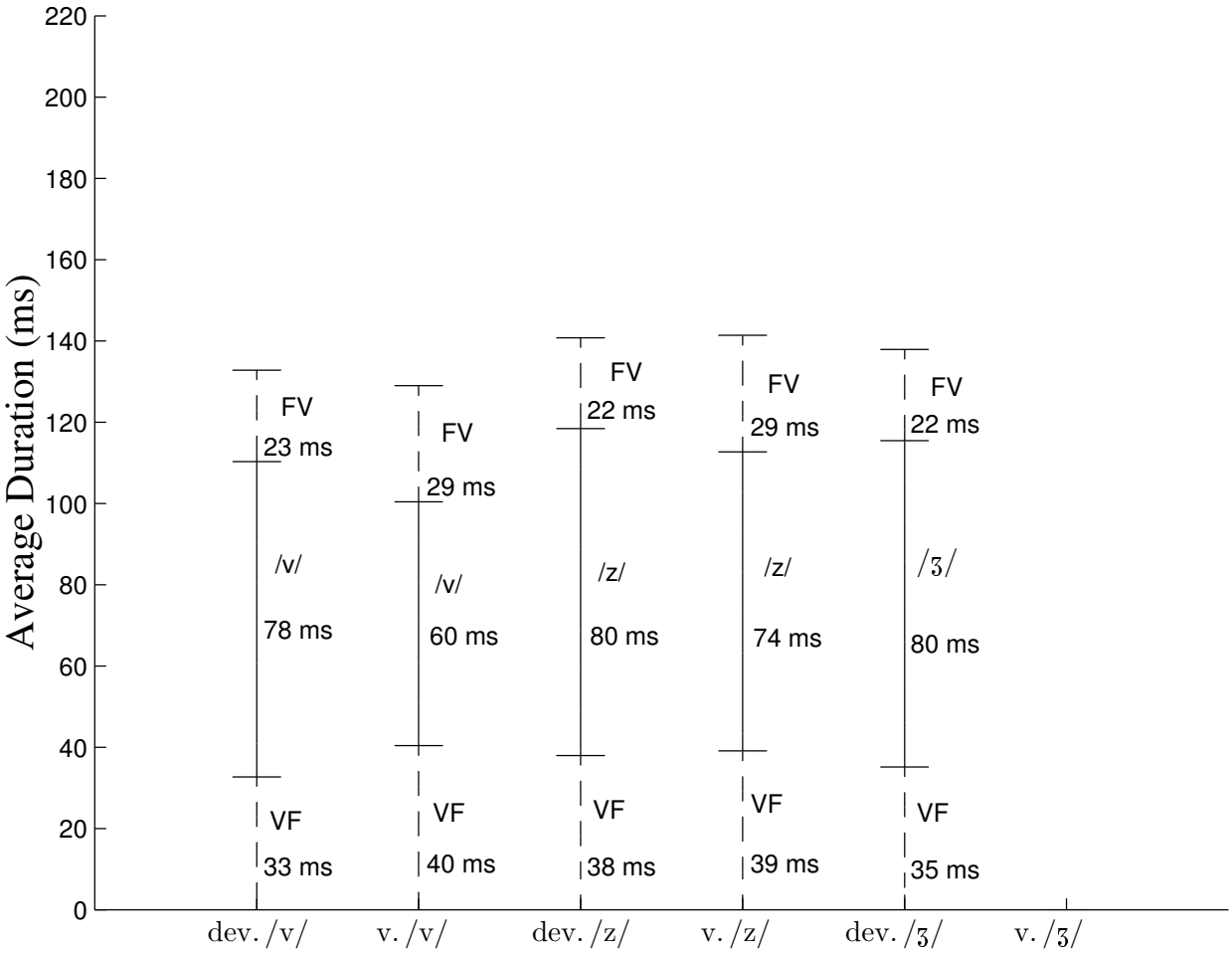


Figure 18: Average duration of voiced and devoiced examples of fricatives /v, z, ʒ/, and of VF and FV transitions in Corpus 3 (Speaker ACC).

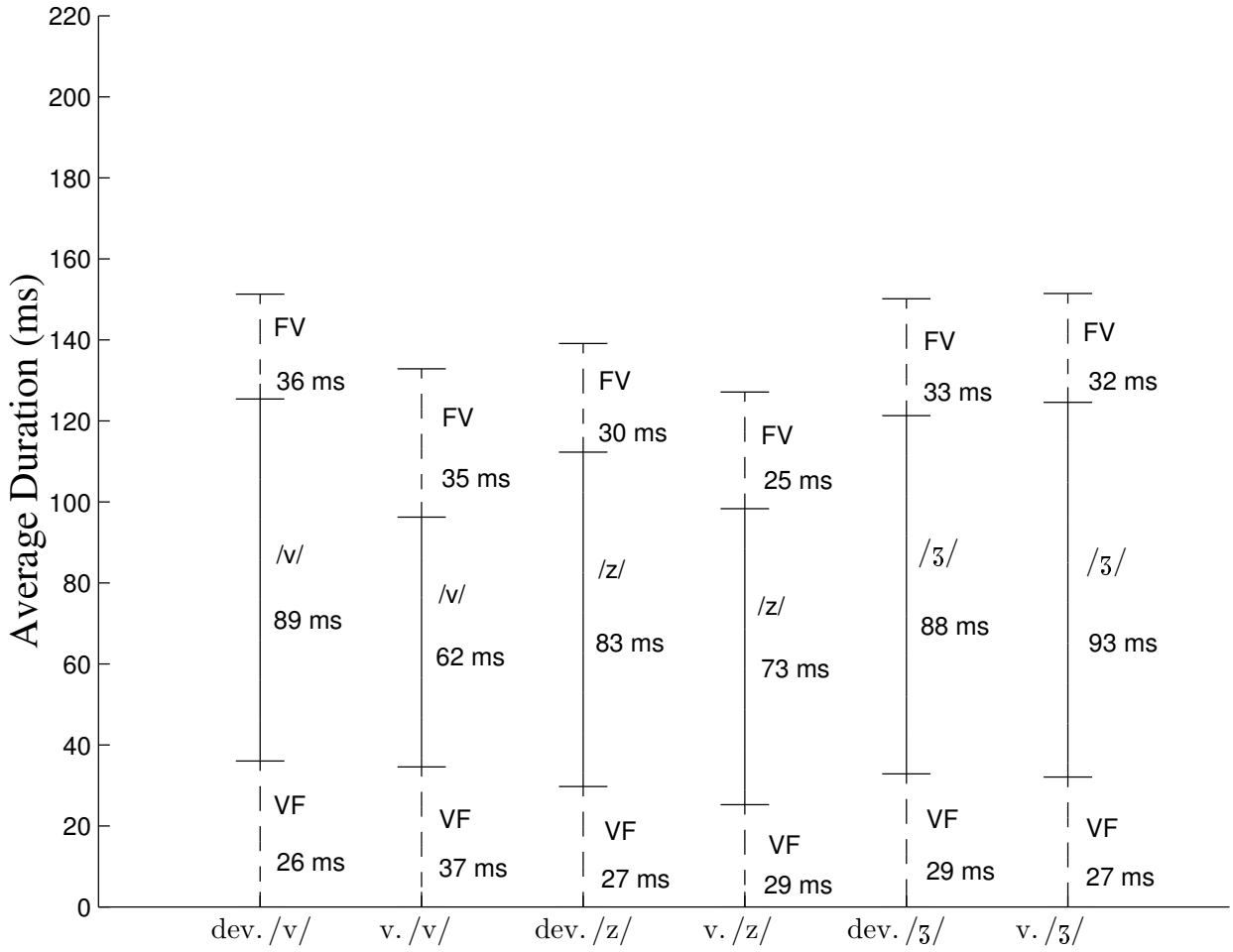


Figure 19: Average duration of voiced and devoiced examples of fricatives /v, z, ʒ/, and of VF and FV transitions in Corpus 3 (Speaker ISSS).

Duration and Devoicing Correlations
–
Corpus 4

For speakers LMTJ and ACC, when Corpus 4 fricatives devoice their duration increases, and the VF and FV transitions are longer (though for a few examples the duration remains the same), as shown in Figures 20 and 22.

This contradicts the result obtained for Corpus 3 where the FV transition duration diminishes when the fricative devoices.

There are some examples (e.g., /ʒ/ produced by Speaker CFGA) where the duration of devoiced fricatives is smaller, but this is because almost all fricatives are devoiced, and so the duration of the few voiced examples is clearly atypical.

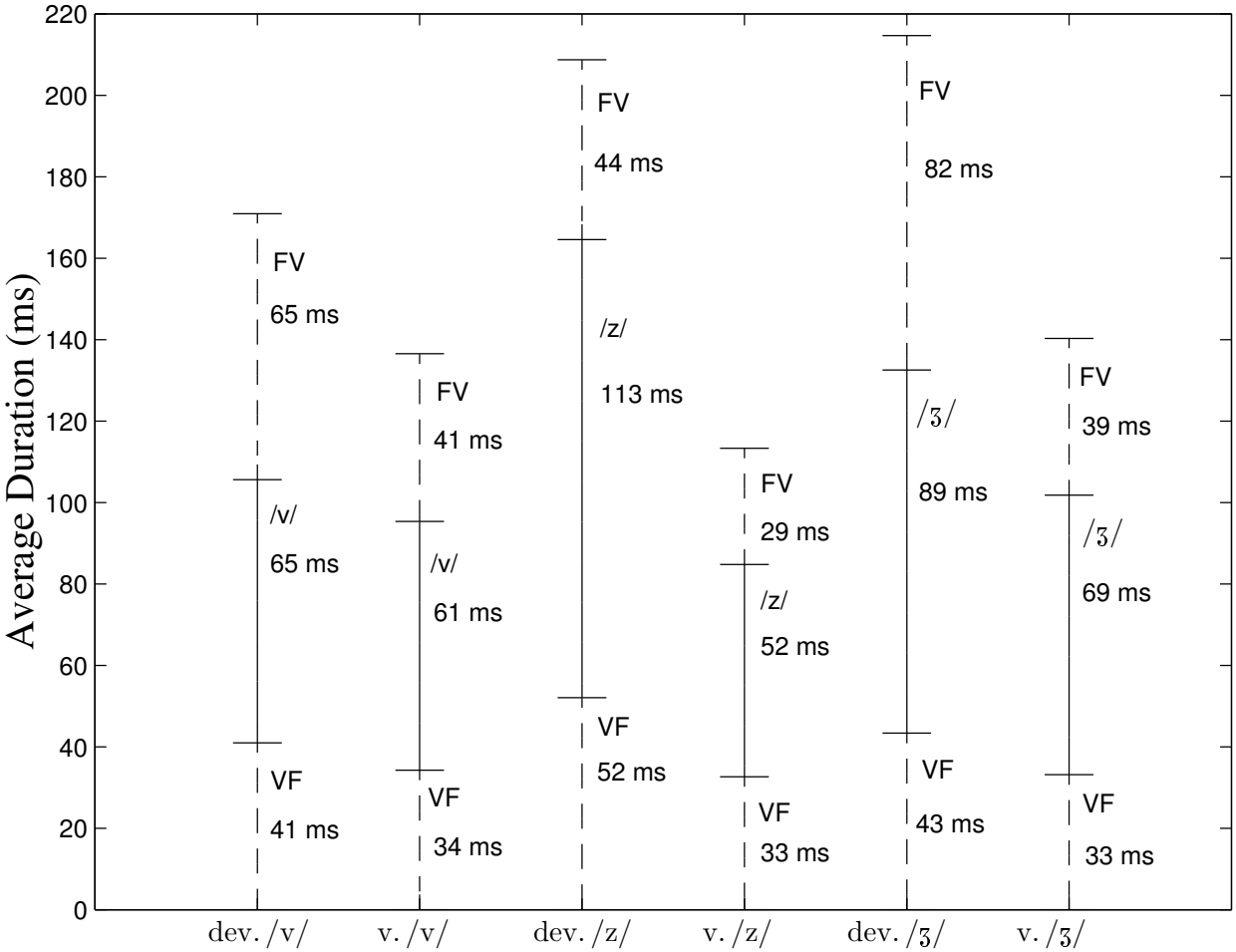


Figure 20: Average duration of voiced and devoiced examples of fricatives /v, z, ʒ/, and of VF and FV transitions in Corpus 4 (Speaker LMTJ).

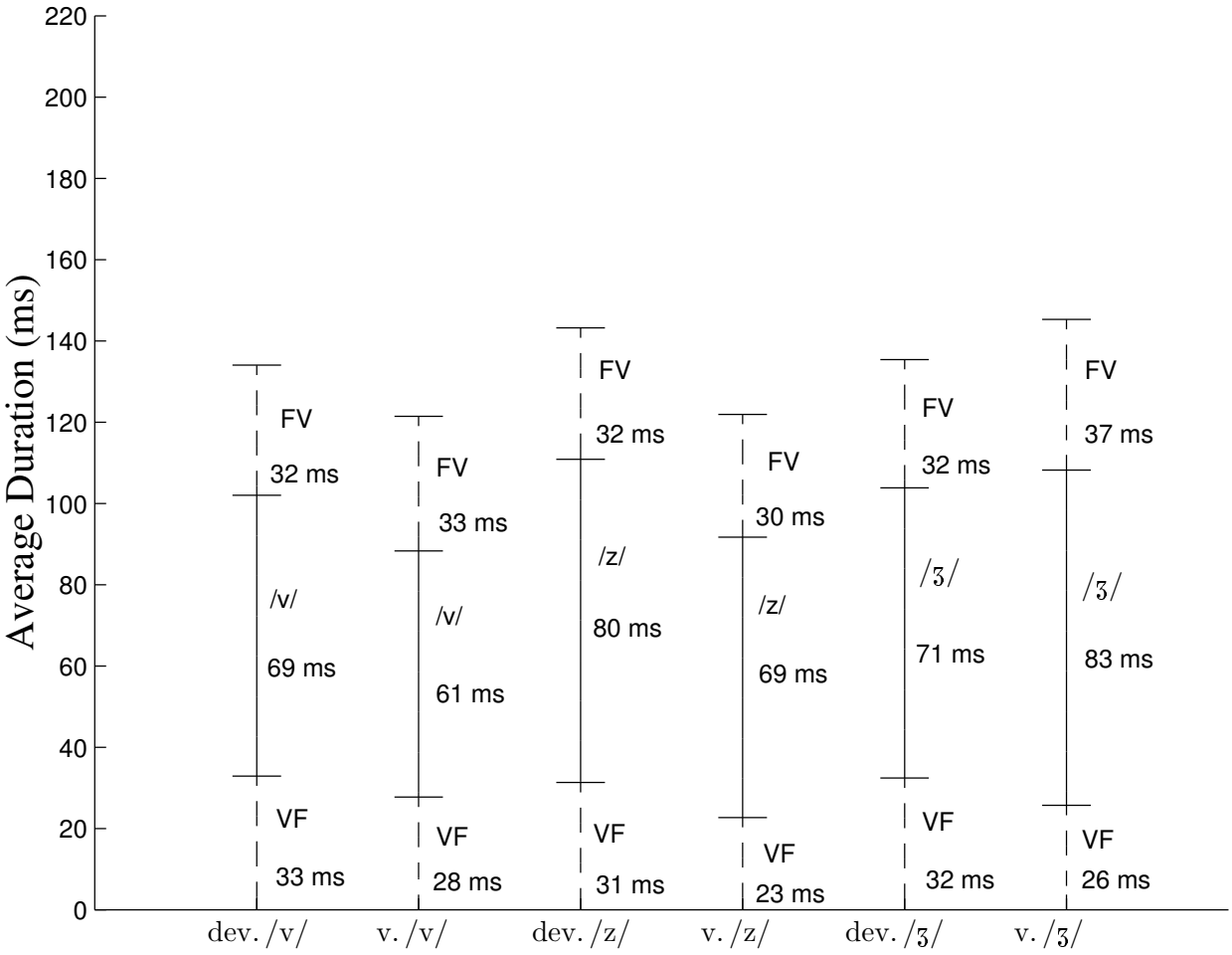


Figure 21: Average duration of voiced and devoiced examples of fricatives /v, z, ʒ/, and of VF and FV transitions in Corpus 4 (Speaker CFGA).

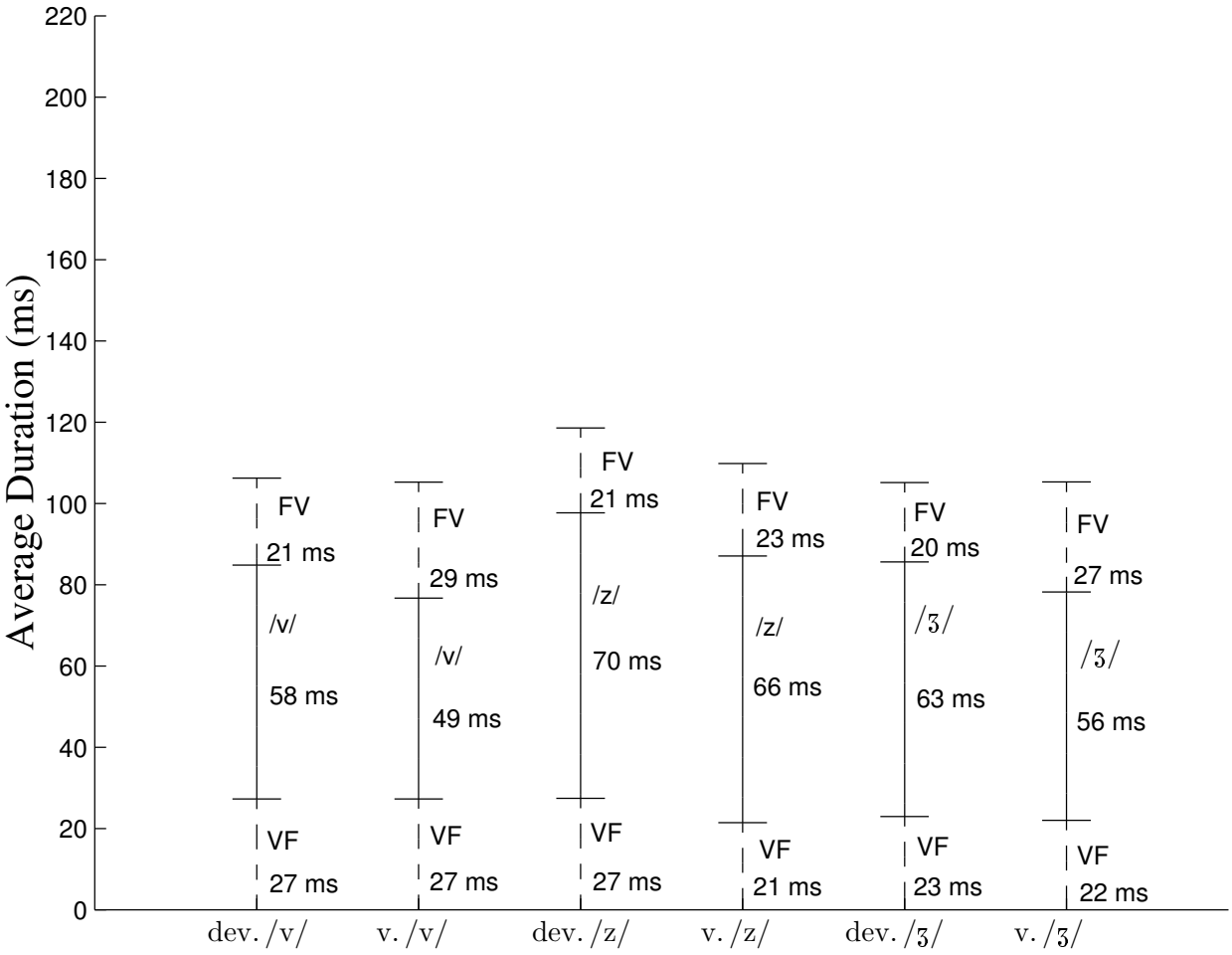


Figure 22: Average duration of voiced and devoiced examples of fricatives /v, z, ʒ/, and of VF and FV transitions in Corpus 4 (Speaker ACC).

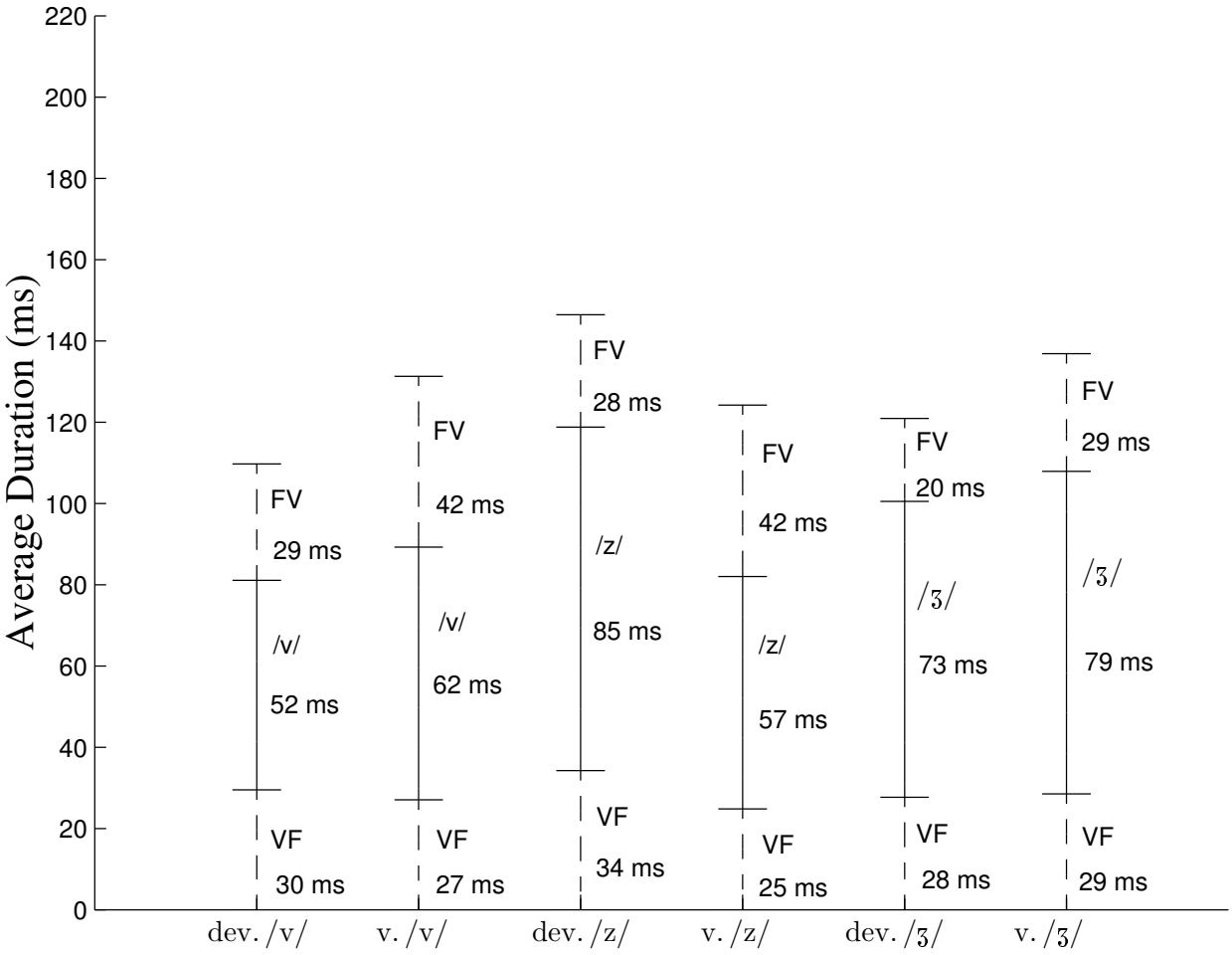


Figure 23: Average duration of voiced and devoiced examples of fricatives /v, z, ʒ/, and of VF and FV transitions in Corpus 4 (Speaker ISSS).

Vowel to Fricative Duration Ratio

The medians of the vowel to fricative duration ratios of unvoiced fricatives is greater than those of voiced fricatives, and devoicing lowers the vowel to fricative duration ratios, as shown in Figure 24.

This new acoustic evidence suggests that Portuguese "maybe changing in the direction of using" fricative to vowel duration ratio "as the primary cue for the voicing contrast" (Klatt 1976, p. 1219).

Segmental durations of fricatives, vowels, and VF and FV transitions, as shown in Figure 25, also provide information about the relative timing of the oral closing / opening and glottal opening / closing gestures (Klatt 1976, p. 1214), and could be used to model the durational behaviour of Portuguese in a speech synthesis by rule system (Klatt 1976; Umeda 1976).

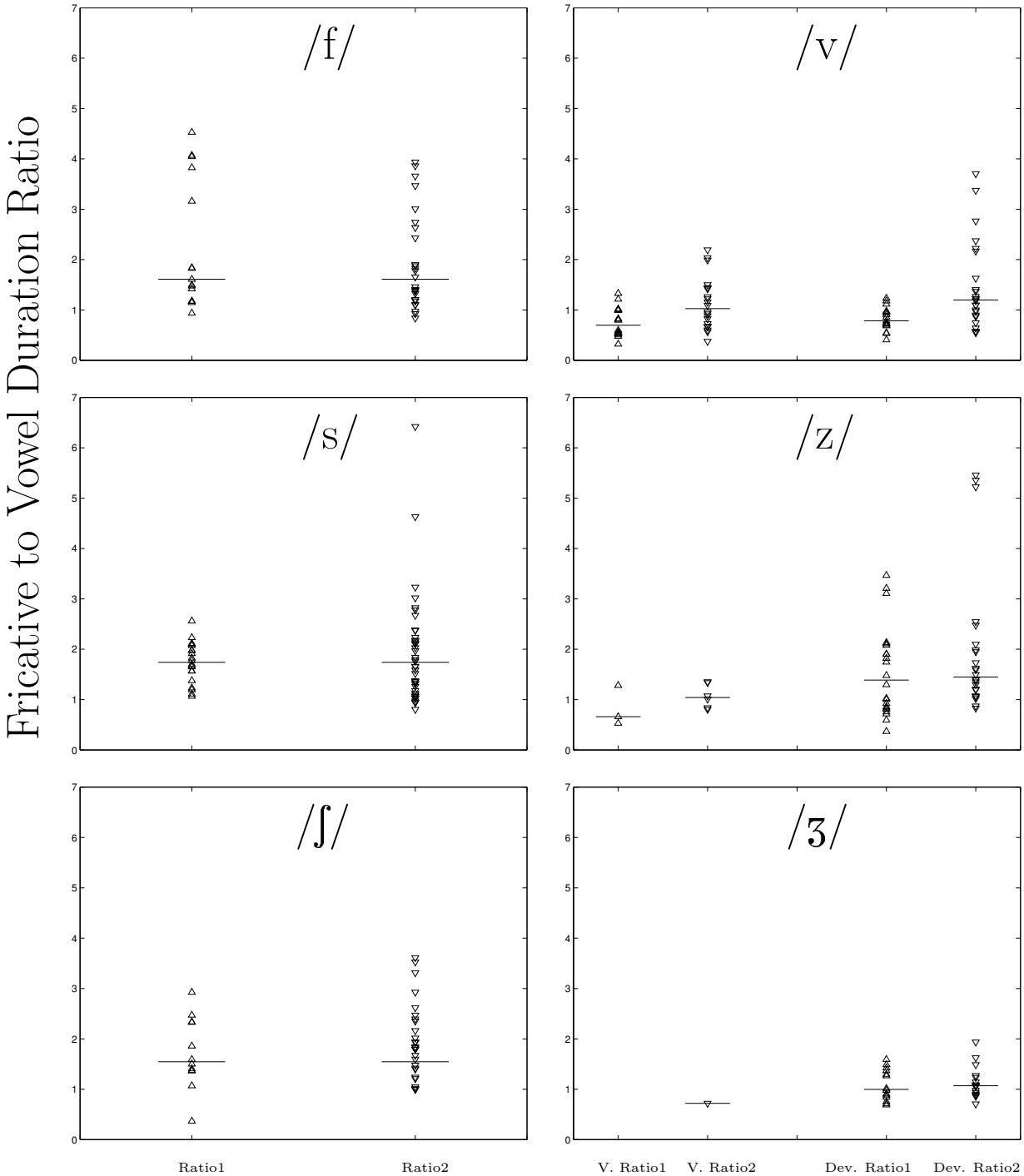


Figure 24: Fricative to vowel duration ratio: ratio of unvoiced fricative duration to the preceding vowel duration (Ratio1); ratio of unvoiced fricative duration to the following vowel duration (Ratio2); ratio of voiced fricative duration to the preceding vowel duration (V. Ratio1); ratio of voiced fricative duration to the following vowel duration (V. Ratio2); ratio of devoiced fricative duration to the preceding vowel duration (Dev. Ratio1); ratio of devoiced fricative duration to the following vowel duration (Dev. Ratio2). The horizontal line is the median. Speaker PS (Corpus 4).

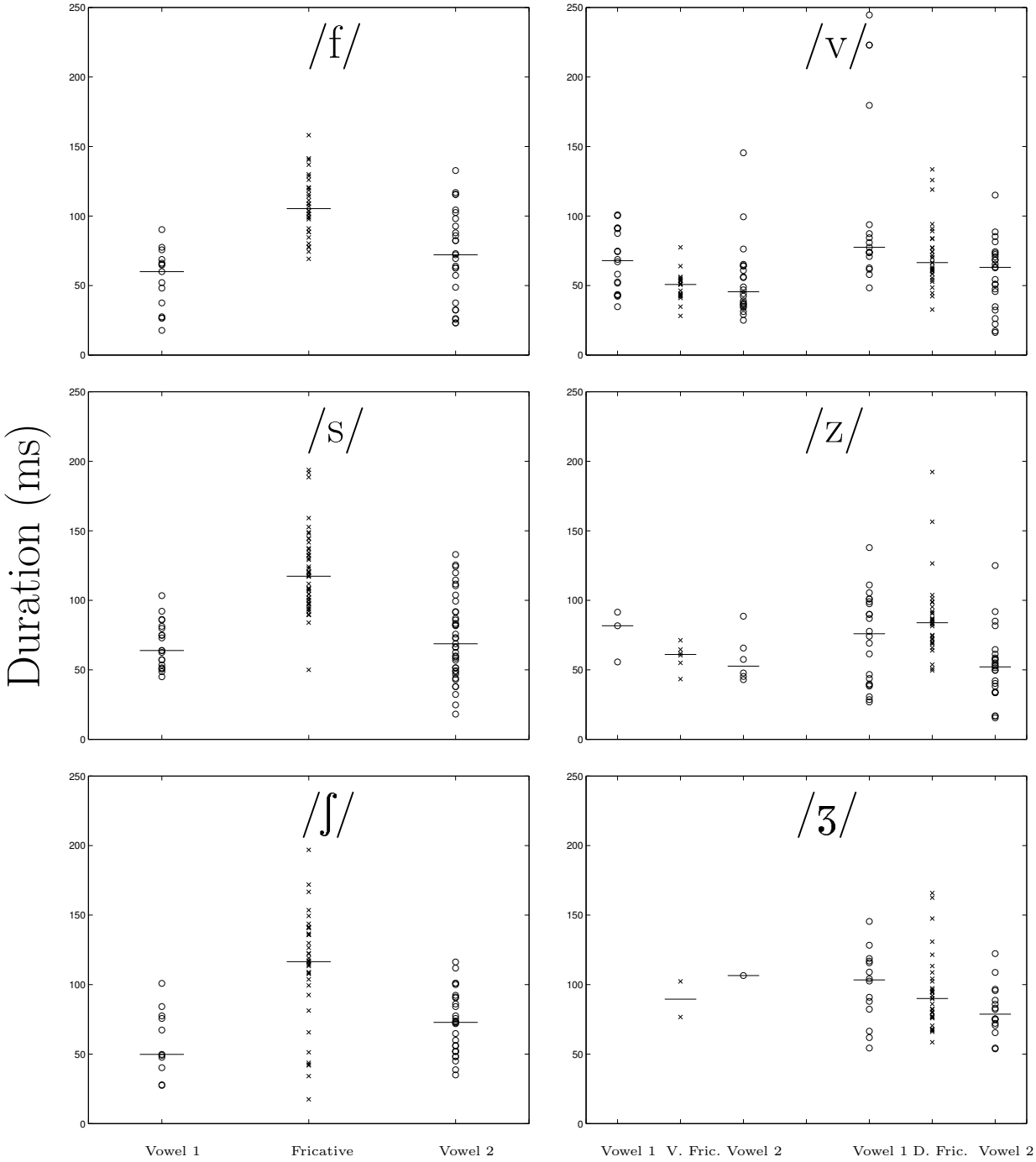


Figure 25: Fricative and vowel durations: vowel preceding fricative (Vowel 1); vowel following fricative (Vowel 2); voiced fricative (V. Fric.); devoiced fricative (D. Fric.). The horizontal line is the median. Speaker PS (Corpus 4).

Acknowledgements

This work was partially supported by Fundação para a Ciência e a Tecnologia, Portugal.

References

Abercrombie, D. (1967). *Elements of General Phonetics*. Edinburgh: Edinburgh University Press.

Behrens, S. J. and S. E. Blumstein (1988). Acoustic characteristics of English voiceless fricatives: A descriptive analysis. *Journal of Phonetics 16*, 295–298.

Cole, R. A. and W. E. Cooper (1975). Perception of voicing in English affricates and fricatives. *Journal of the Acoustical Society of America 58*(6), 1280–1287.

Crystal, T. H. and A. S. House (1988). A note on the durations of fricatives in American English. *Journal of the Acoustical Society of America 84*(5), 1932–1935.

Crystal, T. H. and A. S. House (1997). A note on the durations of American English consonants. In S. Kiritani, H. Hirose, and H. Fujisaki (Eds.), *Speech Production and Language: In Honor of Osamu Fujimura*, Chapter 4, pp. 195–213. Berlin: Mouton de Gruyter.

Docherty, G. J. (1992). *The Timing of Voicing in British English Obstruents*. Berlin: Foris Publications.

Hogan, J. T. and A. J. Rozsypal (1980). Evaluation of vowel duration as a cue for the voicing distinction in the following word - final consonant. *Journal of the Acoustical Society of America 67*(5), 1764–1771.

IPA (1999). *Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet*. Cambridge: Cambridge University Press.

Jesus, L. M. T. (2001). *Acoustic Phonetics of European Portuguese Fricative Consonants*. Ph.D Thesis, Department of Electronics and Computer Science, University of Southampton, Southampton, UK.

Jesus, L. M. T. and C. H. Shadle (2002). A parametric study of the spectral characteristics of European Portuguese fricatives. *Journal of Phonetics, in press*.

Jongman, A., R. Wayland, and S. Wong (2000). Acoustic characteristics of English fricatives. *Journal of the Acoustical Society of America 108*(3), 1252–1263.

Klatt, D. H. (1971). On predicting the duration of the phonetic segment [s] in English. Quarterly Progress Report 103, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, USA.

Klatt, D. H. (1974). The duration of [s] in English words. *Journal of Speech and Hearing Research 17*(1), 51–63.

Klatt, D. H. (1975). Vowel lengthening is syntactically determined in a connected discourse. *Journal of Phonetics 3*(3), 129–140.

Klatt, D. H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America 59*(5), 1208–1221.

LaRiviere, C., H. Winitz, and E. Herriman (1975). The distribution of perceptual cues in English prevocalic fricatives. *Journal of the Acoustical Society of America 18*(4), 613–622.

Manrique, A. M. B. and M. I. Massone (1981). Acoustic analysis and perception of Spanish fricative consonants. *Journal of the Acoustical Society of America 69*(4), 1145–1153.

Pirello, K., S. E. Blumstein, and K. Kurowski (1997). The characteristics of voicing in syllable-initial fricatives in American English. *Journal of the Acoustical Society of America 101*(6), 3754–3765.

Scully, C. (1979). Model prediction and real speech: Fricative dynamics. In B. Lindblom and S. E. G. Öhman (Eds.), *Frontiers of Speech Communication Research*, pp. 35–48. London: Academic Press.

Shadle, C. H. and S. J. Mair (1996). Quantifying spectral characteristics of fricatives. In *Proceedings of the International Conference on Spoken Language Processing (ICSLP 96)*, Philadelphia, USA, pp. 1517–1520.

Smith, C. L. (1997). The devoicing of /z/ in American English: Effects of local and prosodic context. *Journal of Phonetics 25*(4), 471–500.

Stevens, K. N., S. E. Blumstein, L. B. Glicksman, M. Burton, and K. Kurowski (1992). Acoustic and perceptual characteristics of voicing in fricatives and fricative clusters. *Journal of the Acoustical Society of America 91*(5), 2979–3000.

Umeda, N. (1976). Linguistic rules for text-to-speech synthesis. *Proceedings of the IEEE 64*(4), 443–451. Special Issue on Man-Machine Communication by Voice.

Umeda, N. (1977). Consonant duration in American English. *Journal of the Acoustical Society of America 61*(3), 846–858.