CHAPTER 8

CONCLUSION

8.1 INTRODUCTION TO CHAPTER 8

Recall from Ch. 1 that this dissertation had 2 main experimental goals:

1) to determine the effects SR and other variables in the study design had on the frequency of devoiced vowels; and
2) to determine the effects the variables in the study design had on the voicing duration observed in those vowels that were voiced.

Also, since no literature could be found discussing the type of program used to elicit the data, it was deemed necessary to examine the data carefully to make sure that it could be subjected to the checks above.

Before the results of these statistical checks were discussed, however, the supporting goals were presented in Chapters 2 through 4.

8.2 HIGH VOWEL DEVOICING

Chapter 2 discussed the rule of High Vowel Devoicing. First a literature review was presented, beginning with discussion of the various factors surrounding the application of the rule. These included SR, pitch accent placement, dialectal pressures, the mora the target vowel is in , and the consonantal environment of the vowel.

Next, previous characterizations of the rule were presented. In the earliest characterizations (e.g. McCawley 1968) devoicing was viewed as a strictly phonological change of the vowel's [+voice] to [-voice]. However, the recent work (Kondo 1997; Tsuchida 1997) has not upheld this. Based on glottal spreading gestures and the electric potentials of the muscles causing glottal spreading for the one speaker tested, those works maintained that 2 loss of voicing processes are at work in Japanese: 1) phonological devoicing by spreading the feature [spread glottis] from the preceding stop, where the sharing of the feature by the stop and the vowel results in the temporal realignment of the spreading gesture to the midpoint of their productions (Kingston 1990; Iverson & Salmons 1995); and 2) a phonetic process of gestural overlap whereby the glottal spreading gesture of a preceding and following fricative overlap the voicing gesture of the vowel, resulting in gradient amounts of voicing on the vowel. In other words, in Tsuchida's model phonological devoicing occurs between plosives, but phonetic loss of voicing occurs between voiceless fricatives.

It was noted that the data from the current study supported gradient amounts of voicing between the allophonic affricates occurring in the tokens used. It seems reasonable to assume that the fricative portion of affricates affect vowel voicing in the same way that voiceless fricatives do. However, it was also noted that in this study devoicing between velar plosives shows the same gradiency that devoicing between fricatives and affricates do, and further that there appears to be only one glottal spread after productions of affricates as well (as evidenced by examination of the frication in spectrograms of affricate productions). Tsuchida (1997) is therefore supported by this study—i.e. there are both phonological devoicing and phonetic loss of voicing at work in Japanese—with the modification that both processes occur in all environments for some participants but possibly not for others. That is to say, for some speakers, both processes occur in all consonantal environments.

Secondly, a new characterization of the mode of production of many devoiced vowels was discussed. Specifically, it was maintained that many productions are actually fricativized vowels (Ladefoged 1996), usually but not always accompanied by devoicing. This is quite similar to the 'obstruentization' in Uyghur, a Turkic language, discussed in Hahn (1991). Examples of both the frication observable in spectrograms and the number of zero crossings were presented as evidence. The fact that fricativization also occurs in productions of heavily voiced vowels indicates manipulation of the oral closure independent of vowel voicing.

It was also noted that the two types of devoicing specified above lead to the characterization of two parameters for vowel devoicing: 1) manipulation of the glottal opening to control voicing via glottal spread; and 2) manipulation of the oral closure to control voicing via reduced air flow. These two parameters certainly interact, as was seen in Figures 2.11 to 2.13. Further work is needed to determine to what degree they are dependent.

In view of this discussion, an attempt toward clarifying the terminology surrounding HVD was made, as well as a step toward appropriate geometric representations of the various stages of vowel reduction observed in devoicing data. The various stages posited were *fully voiced vowels,* characterized by voicing lasting the duration of the oral cavity configuration with accompanying well-defined formant structure; *partially voiced vowels,* characterized by voicing lasting only a portion of the time formant structure does; *fricativized vowels* which may or may not be devoiced, characterized by heavy amounts of frication attributed to a increased oral closure; *devoiced vowels;* characterized by heavy frication which serves as the vehicle for the formation of the formant frequencies of the original vowel; and *deleted vowels,* characterized by both a loss of voicing and formant activity. It was noted that the timing slot of the deleted vowels may or may not be preserved, with an example where the timing slot was not preserved being offered (i.e. the frication of the preceding obstruent was much shorter in duration than in other productions). *Whispered vowels* were characterized as unrelated to normally occurring vowel devoicing.

Phonological devoicing consists of a spread of the feature [+spread glottis] from the preceding obstruent (based upon place of articulation: [k͜x$^{\check{\text{i}}}$ši], cf. *[k͜š$^{\check{\text{i}}}$ši], as well as the shift of the frication from the preceding obstruent toward the vowel that is observable in spectrograms). Fricativization was characterized as a spread of the feature [+consonantal] from the preceding obstruent. However, it was noted that this characterization of fricativization is not satisfactory due to the gradient nature of frication observable both on the vowel site and the following stop's closure. For the time being, the surface representations of the types of vowels delimited above can be tentatively classified as follows for the 3 features [consonantal], [voice], and [spread glottis]:[1]

---

[1]Minus specifications for [spread glottis] and [voice] are not needed if these features are taken to be monovalent, a seemingly well-justified move (see Kenstowicz 1994 §9.7 for discussion).

fully voiced vowel: [-consonantal, +voice]

partially devoiced vowels: [-consonantal, +voice]
devoicing determined by overlap of [+spread glottis] and/or [+consonantal]

fully devoiced vowels: [+consonantal, +spread glottis]

whispered vowels: [-consonantal, +spread glottis]

fricativized vowels: [+consonantal] {[+voice] [+spread glottis]}
devoicing determined by overlap of [+spread glottis]

deleted vowels: possible deletion of timing slot;
deletion of all/some features

Both partially devoiced and totally devoiced vowels are thought to have the same underlying specifications as fully voiced vowels; i.e. [-consonantal], [+voice]. The degree of devoicing will be determined by the extent of the overlap of the voicing gesture by the glottal spreading gesture ([+spread glottis]), manipulation of the degree of oral closure ([+consonantal]), or a combination of the two. Fully devoiced vowels are characterized above as [+consonantal, +spread glottis] due to the fact that the articulation of the vowel productions in this study is that of a fricative (see §2.3.2 for discussion).

The degree of deletion of a vowel will be determined by whether or not the timing slot itself has been preserved, and if it is, which features remain attached to the timing slot (i.e. residual oral configuration features preserving the original vowel's identity—see §2.5.5). The amount of voicing accompanying a fricativized vowel will be determined by whether there has been total or partial assimilation of the feature [+spread glottis].

Exploration into the proper characterization of the interaction between fricativization and devoicing is left for further work involving articulation studies.

8.3 INTERACTION BETWEEN PITCH ACCENT ASSIGNMENT AND DEVOICING

Chapter 3 discussed the interaction between pitch accent placement and vowel devoicing in Japanese. As noted in Chapter 2, it has been maintained in many works that pitch accent blocks the devoicing of an accented vowel (Bloch 1950; McCawley

1977; Haraguchi 1977). However, many other works have demonstrated that devoicing of accented vowels does occur (e.g. Sugito 1966; Kondo 1993; Kitahara 1997, 1998). Devoicing of accented vowels in some cases was also supported by this study. Example productions of both cases were offered in §3.3.

One of the strategies for avoiding devoicing an accented vowel is to shift the accent location away from the vowel to be devoiced. However, although observed, this does not appear to have been a common strategy used by the participants of this study. More commonly it appears that an accented vowel was devoiced. Further examination of the data is necessary to determine what other, if any, strategies the participants of this study utilized for avoiding pitch accent assignment/devoicing conflict.

8.4 EXPERIMENTAL SETUP

Ch. 4 summarized the experimental setup of the study, including a description of the HyperCard™ stack used to elicit data. The program was used to display sentences on the computer screen for predetermined amounts of time, with participants pacing their production of the sentence to match the display time. The variables affecting the application of HVD were also discussed. Variables for which control was attempted were noted, followed by discussion of the variables that were manipulated in the study. The techniques used for measuring token durations and vowel voicing durations were also presented; both token and voicing durations were made from filtered, expanded waveform displays assisted by inspection of wide-band spectrograms.

8.5 VARIABLES AFFECTING MEAN TOKEN DURATIONS

Ch. 5 presented checks on the data to ensure that the program was successful in eliciting 3 distinct data sets (*slow, normal* and *fast)* allowing further statistical analysis. This was done by checking that the variables affecting the MTDs did not compromise the generation of those sets.

It was seen that while the requirement for *normalcy* was satisfied, the requirement of *homogeneity of variance* (i.e. the nature of the data distribution about its mean value) was not the same for the 3 data sets. However, after examination of the data, and in consideration of discussion on homogeneity of variance in Rosenthal & Rosnow (1993), it was decided that it was appropriate to subject the data to ANOVA and ANCOVA analysis to check for effects of the various factors on the application of HVD.

By far the largest factor affecting MTDs was SR, the display times of the stimuli sentences. The effect of this variable was enormous; as the sentence display time increased, the participants slowed their productions in a step-wise fashion. This resulted in token productions of significantly different durations for each of the 3 data sets. Although the inclusion of both genders, different tokens, and multiple repetitions (and their interactions) did have significant effects on MTDs, their inclusion did not adversely affect the generation of three distinct data sets corresponding to the three targeted SRs.

It is worth noting that in general males produced shorter tokens than the females, supporting a result noted in Yuen & Hubbard (1997). Although the effect of gender in this study was not as significant as in that study, this is likely due to the nature of the two elicitation paradigms—in Yuen & Hubbard (1997), a self-judged SR during a sentence-reading task was utilized; in this study, a computer program exerted control over the SR.

All in all, it was judged that the program did elicit three distinct data sets without being compromised by the variables that affected the MTDs. Further research will therefore utilize the same basic elicitation paradigm.

8.6 FACTORS AFFECTING FREQUENCY OF DEVOICING

Somewhat surprisingly, SR was not the major factor influencing the frequency of vowel devoicing. Instead, the largest influence on devoicing rates was which mora (the first or the second) the vowel was in. This supports one of the findings of the study of Kuriyagawa & Sawashima (1989), and also supports the contention that HVD can no longer be considered a strictly fast speech rule—moraic position, not SR, was the largest factor affecting it. Kuriyagawa & Sawashima (1989: 92) verified that the 1st mora durations in that study's tokens were of shorter duration than the

2nd mora durations; the fact that word-initial consonants and vowels are shorter than medial consonants and vowels (Sagisaka & Tohkura 1984) was proposed as the reason more 1st mora vowels were devoiced. Particularly as SR increased, the shorter-duration 1st mora vowels were even more susceptible to devoicing. Since mora durations have not yet been taken for the data in this study, it remains to be seen whether this explanation is also plausible for the current data. Comprehensive pitch accent placement judgments will also need to be made to rule out accent placement effects, although Homma (1973) found that pitch accent placement does not have a significant effect on vowel duration.

SR did have a significant effect. This was seen in the way that SR interacted with the mora the token was in. Vowels in both mora of the tokens became more likely to devoice as SR increased, but the effect of SR was not the same for both mora. Vowels in the 1st mora of the tokens became much more likely to devoice as SR increased; vowels in the 2nd mora of the tokens also became more likely to devoice to a lesser degree.

The token itself also had a significant effect on whether or not one of its vowels would devoice; i.e. the vowels in some tokens were more likely to devoice than the vowels in other tokens. However, the exact source of this effect is impossible to pinpoint in the current study due to the unbalanced consonantal environments of the tokens utilized.

Finally, the gender of a participant also had a significant effect. In general, males devoiced more vowels than the female participants did. However, this effect depended on which repetition block a vowel was produced in. In general, the males produced more devoiced vowels in the 1st repetition block, while females produced more devoiced vowels in the 2nd block. If there is a source for this observed variation, it will need to be uncovered by further research. It may have been due to the individual effects that were also noted, although the individual participant differences had much less an effect on the frequency of devoicing than gender did.

8.7 FACTORS AFFECTING VOWEL VOICING DURATIONS

The results of the checks for effects of the variables on vowel voicing durations (voiced vowels only) were almost the same as the results of the checks for effects of the variables on the number of devoiced vowels discussed above. The only significant difference was that individual effects due to participant differences were also dependent on which repetition block a vowel was produced in. In other words, some participants produced longer voicing durations in the 1st repetition block than in the 2nd, while others produced longer voicing durations in the 2nd. Still others produced almost the same average amount of vowel voicing in each repetition block. This effect is attributed to simple individual variation.

8.8 INDICATIONS FOR FURTHER RESEARCH

In regard to experimental design, it was shown that although the program did facilitate the production of 3 distinct data sets corresponding to 3 distinct SRs, the length of time available for the participants to produce a sentence at the slow SR led to variation in the overall sentence durations. This was reflected in the token durations as well. Further work will involve finding a way to lessen the observed variance due to the display times of the stimuli sentences.

In addition, it was noted several times in discussion of the effects of the variable TOKEN that the source of the variation could not be judged due to the incomplete set of consonantal environments in the tokens used. Therefore more work needs to be done using a phonetically balanced set of tokens in order to highlight effects of the consonantal environment.

As for the process of devoicing itself, there are many details that still need to be cleared up. In particular, the nature of the phonetic overlap associated with either a shift or a lengthening of the glottal spreading gesture is needed—i.e. are there any factors influencing a participant's choice in using one but not the other? Also, it needs to be further verified that the shift of the spreading gesture, as evidenced by the location of frication in the spectrogram, is not free to be shifted to any temporal location, but is indeed fixed at the midpoint of the preceding obstruent and the vowel

(Kingston 1990; Iverson & Salmons 1995). Both of these points will be addressed by further examination of the current data set.

Related to the process of devoicing, overlap of the specification for [+spread glottis] from the preceding consonant is held to be responsible for SR-dependent devoicing. In addition, as noted in §2.3.2, fricativization necessarily involves only the preceding consonant. Finally, devoicing of vowels when the following consonant is a voiced sonorant has been noted in the literature (Maekawa 1990) and is readily observable in the current data set as well (e.g. [t͡su̥noru] 'to mount, increase'). The question is why this 'illicit' devoicing is limited to before sonorants in the current data set. The interaction between vowel devoicing and the assumed redundant voicing of sonorants warrants further attention.

In addition, the effect of vowel length on vowel devoicing has been noted in the past—short high vowels in devoicing environments are easily devoiced, but devoicing of long high vowels in devoicing environments have to the best of the current author's knowledge never been reported. For SR-dependent devoicing, it is assumed that cases of reduced devoicing of long vowels would be observed (i.e. the first 40 ms of a 120 ms long vowel would show no voicing activity). However, the lack of evidence of totally devoiced long vowels suggests that phonological (i.e. SR-independent) devoicing of the long vowels is not possible.

In light of the reduced effect of the following segment on the application of devoicing noted immediately above, one might expect devoicing of long vowels as well particularly if HVD were acting on feature specifications alone—the doubly-linked specifications of features should appear to the rule to be only one specification of the feature [voice]. Any reduced effect of the following segment due to redundant specification for [voice] found in consonantal sonorants should be seen in vocalic sonorants as well. What exactly is blocking the devoicing of long vowels, be it the double linking of the features itself or something else, warrants further research.

Fricativization of vowels, both voiced and devoiced, was also observed in the current data. This was attributed to a close oral constriction, equal to that of a fricative when the vowel is devoiced. The characterization of the observed frication was given as a spread of the feature [+consonantal] after Kaisse (1992). However, it was noted that this characterization was not satisfactory due to the highly gradient nature of the amount of frication observed. Further work involving articulographic measurements is called for to determine the exact range of oral closure possible, its influence on the observed amount of devoicing, and the factors influencing the degree of closure.

Lastly, it was noted that the vowel reduction processes being subsumed under devoicing appears to involve two separate parameters: 1) degree of glottal spread; and 2) degree of oral closure. (The 3rd parameter, that of oral cavity shape, is not directly involved in the voicing of the vowel.) As was seen in Ch. 2, these occur independently, although in general a close oral closure (i.e. fricativization of the vowel or vowel site) appears to be related to the degree of voicing of the vowel.

Further studies coordinating glottal spreading and degree of oral closure are indicated, ideally combining concurrent glottographic and articulographic examination of the same productions.