Phonological Status of Korean /w/: Based on the Perception Test

Kang, Hyunsook[1])

ABSTRACT

The sound /w/ has been traditionally regarded as an independent segment in Korean regardless of the phonological contexts in which it occurs. There have been, however, some questions regarding whether it is an independent phoneme in /CwV/ context (cf. Kang 2006). The present pilot study examined how Korean /w/ is realized in /S*wV/ context by performing some perception tests. Our assumption was that if Korean /w/ is a part of the preceding complex consonant like /Cʷ/, it should be more or less uniformly articulated and perceived as such. If /w/ is an independent segment, it will be realized with speaker variability.

Experiments I and II examined the identification rates as "labialized" of the spliced original stimuli of /S*-V/ and /Sʷ*-ʷV/, and the cross-spliced stimuli /Sʷ*-V/ and /S*-ʷV/. The results showed that round qualities of /w/ are perceived at significantly different temporal point with speaker and context variability. We therefore conclude that /w/ in /S*wV/ context is an independent segment, not a part of the preceding segment. Full-scale examination of the production test in the future should be performed to verify the conclusion we suggested in this paper.

Keywords: sound /w/, Korean, independent segment, complex segment, perception test

1. Introduction

Glides in general show formant structures similar to those of the vowels but are different from them since they have no steady state of formant structures as Padgett (2007) suggested.

The specific acoustic characteristics of glide in each particular language, however, need to be investigated in terms of not only its own acoustic characteristics within its own temporal period, but also the variations it may have caused to its nearby segments.

Some glides, for example, show their typical acoustic characteristics within their own temporal periods only: English /w/ is known to have low target frequency of F2 around 751Hz in the syllable-initial position across all vowel contexts (cf. Sussman, 1994; Lisker, 1957; O'Connor et al, 1957). It shows

little coarticulation with the upcoming vowel.

Other glides show variations depending on the contexts. Kang (2006) shows that Korean /w/ has different variants in different contexts: as a single onset, it showed the rising F2 transition to the (higher) F2 of the upcoming vowel whereas as a member of the complex onset in /stop-w-vowel/ sequence it showed the flat transition from the release of the stop to the F2 of the following vowel (cf. Yun 2005). Since the latter variant does not seem to have its own definable temporal period with its typical distinctive acoustic features, one comes to wonder whether it should be categorized as an independent phoneme /w/ or rather as a part of the preceding segment like /Cʷ/ or the following segment.

Oh (2010) also argues that Korean /w/ can be acoustically realized in various forms after examining /stop-w-V/ sequences: /w/, /ɥ/, a part of complex /Cʷ/ or null depending on contexts and speakers. In particular, Oh assumed that the consonant before /w/ in /CwV/ sequence is labialized if it shows F2 difference but no VOT difference compared to the corresponding plain /CV/ sequence. It is not clear, however,

why VOT differences affect the preceding consonants' being labialized or non-labialized; secondly, if it is an indicator of the preceding consonants' being labialized, there is also question of whether the whole duration of the stop (closure duration and VOT if it occurs) might be a better indicator. As for the coarticulation of the preceding segment by the following /w/ in Korean, Kang (2010b) also showed that /w/ in fact affects the preceding consonants' spectral shapes in /swi/ sequences.

Though there have been some studies regarding the surface realizations of Korean /w/, there have not been many attempts to consider what acoustic cues Koreans use in perceiving Korean /w/. Are the rounding qualities overlapped on preceding C and/or at the beginning of the following vowel important in perception? Or do the rounding qualities overlapped on preceding C or at the beginning of the following vowel play different roles in perception depending on speakers and contexts? Or do both these cues occur in all the speakers?

By examining Korean listeners' perception of sound /w/, we may answer first, what acoustic cues of /w/ are distinctive and/or redundant in various contexts in perception and whether the distinctive and redundant cues change their roles with speaker and context variability. Secondly, we also expect to answer whether the acoustic form of /w/ in contexts like /stop-w-e/ and /stop-w-i/ is the general form of /w/ in /CwV/ context. Kang (2006) showed that if the preceding consonant is stop, in particular a tense stop, the following /we/ and /wi/ sequence showed more or less uniform acoustic realization like /ɥ/ and /œ/. This raises the question of whether /CwV/ sequences in these contexts should be represented as /CʷV/ or /Cɥ(œ)/. We would like to see whether fricative /S*/ in /S*wV/ contexts may answer this question. Note that in this paper we will represent tense alveolar fricative as /S*/ and lax alveolar fricative as /s/.

In order to answer these questions, we chose the test sequence in which the tense alveolar fricative /S*/ occurs as the segment before /-wV/ sequence since it has its own temporal interval that may show the coarticulatory effects with its nearby segment (cf. Beckman & Shoji, 1984; Heinz & Stevens, 1961; Jongman et al, 2000, Kang, 2010a, b).

This paper consists as follows: In section 2, we run Experiments I and II to examine the perception of the stimuli

of spliced original stimuli /S*-V/ and /Sʷ*-ʷV/, and the cross-spliced stimuli of /Sʷ*-V/ and /S*-ʷV/. Insection 3, we will discuss the roles some acoustic cues of /w/ play in Korean and the phonological status of Korean /w/ based on the results of these Experiments.

2. Experiments

In this section, we ran the pilot test on the perception of /S*wV/ sequences. Experiment 1 examined the perception of the spliced original stimuli of /S*V/ and /S*wV/. This experiment was performed to examine whether all the sequences are correctly produced. Experiment II examined the perception of the cross-spliced stimuli of /S*V/ and /S*wV/. This experiment was performed to examine at what temporal point listeners perceive the rounding qualities of /w/ in the signal. The results of the Experiments are expected to show whether /w/ in /S*wV/ sequence is well produced and if so, at what temporal interval in the acoustic signal "rounding" cues of /w/ are distributed in each speaker.

2.1 Methods
2.1.1 Speakers
Five male speakers from Seoul (age 20-25) read in the normal speed a list of sentences shown in Table 1. The list in Table 1 contains sequences with tense alveolar fricative /S*/ and lax alveolar fricative /s/. The list was written in Korean.

Table 1. List of sentences
1. ikɨldʒanɨn syetayeyo.
2. ikɨldʒanɨn switayeyo.
3. ikɨldʒanɨn sutayeyo.
4. ikɨldʒanɨn sətayeyo.'
5. ikɨldʒanɨn swatayeyo.
6. ikɨldʒanɨn S*yetayeyo.
7. ikɨldʒanɨn swətayeyo.
8. ikɨldʒanɨn S*atayeyo.
9. ikɨldʒanɨn sitayeyo.
10. ikɨldʒanɨn S*witayeyo.
11. ikɨldʒanɨn S*utayeyo.
12. ikɨldʒanɨn syatayeyo.

13. ikɨlʤanɨn S*yotayeyo.

14. ikɨlʤanɨn S*etayeyo.

15. ikɨlʤanɨn S*ətayeyo.'

16. ikɨlʤanɨn S*watayeyo.

17. ikɨlʤanɨn satayeyo.

18. ikɨlʤanɨn syutayeyo.

19. ikɨlʤanɨn sotayeyo.

20. ikɨlʤanɨn syətayeyo.

21. ikɨlʤanɨn S*yatayeyo.

22. ikɨlʤanɨn syotayeyo.

23. ikɨlʤanɨn S*wyetayeyo.

24. ikɨlʤanɨn setayeyo.

25. ikɨlʤanɨn S*otayeyo.

26. ikɨlʤanɨn S*itayeyo.

27. ikɨlʤanɨn swyetayeyo.

28. ikɨlʤanɨn S*wətayeyo.

29. ikɨlʤanɨn S*yutayeyo.

30. ikɨlʤanɨn S*etayeyo.

Five speakers read the list three times. No speaker reported any hearing or speech problem. A small amount of money was paid to the speakers. The recording is done in a sound attenuated booth using a dynamic microphone SHURE KSN44 and a digital recorder Tascam HD-P2 at the Hanyang Phonetics Lab. The speech was digitized on a PC running Praat software (Boersma & Weenink 2002).

2.1.2 Material for Perception Test

Among the recorded material, the words /S*Vta/ and /S*wVta/ were excised from the carrier phrases. In particular, we chose /S*e/, /S*a/, and /S*ə/ from the /S*V/ sequences and /S*we/, /S*wa/, and /S*wə/ from /S*wV/ sequences for the Experiments. The reasons for choosing these sequences are first, that there are native Korean words beginning with these sequences, secondly that the vocalic portion after /S*/ in these sequences are long enough to be closely examined in the acoustic signal unlike /S*i/ and /S*wi/, thirdly, that tense alveolar fricative does not contain aspiration so that it was easy to make cross-spliced tokens, and finally, that the vowels /e/, /a/, /ə/ may represent front, back and central vowels that do not influence the acoustic characteristics of the preceding segment

/S*/ (cf. Kang, 2010a, b).

These six words are then segmented into two components. The first component is the frication portion of the waveform. The frication component, then, may include the fricative /S*/ and whatever features overlapped on the frication portion. The second component included the rest of the word: from the onset of the regular waveform which includes the first vowel V to the offset of the following /-ta/ sequence as well as glide /w/ if it is realized at the start of the regular waveform of the first vowel. Since two components occur in each type of a sequence, the following components are available for the stimulus of the tests: /S*-/ and /-V(ta)/ from /S*Vta/, and /S^w*-/ and /-^{w}V(ta)/ from /S*wVta/. Since qualities of /w/ may be realized on the preceding fricative portion and/or at the start of the following vocalic portion, the two components from /S*wVta/ sequence were both marked with the superscript /w/. From this point on, we will simply refer to these four components as plain /S*/ and /V/ and labialized /S^w*-/ and /-^{w}V/ without the following /-ta/.

Specifically, /S*V/ and /S*wV/ sequences were segmented as in (1).

(1) Creation of the components

a. /S*V_i/ ---> /S*$_i$ – V_i/   (i = 1, 2, 3)

b. /S*wV_j/ ---> /S^w*$_j$ – ^wV_j/ (j = 1, 2, 3)

From (1), we selected the first recorded /S*$_1$/ and /S^w*$_1$/ as the representative frication components. To each representative frication component, the remaining three vocalic parts of the sequences were added. For the /plain-plain/ stimulus, for example, the vocalic components of the plain phrase, /-V_1/, /-V_2/, and /-V_3/ are added to the representative frication noise /S*$_1$-/, creating three tokens as in (2a). Likewise, to the representative labialized frication component /S^w*$_1$-/, the vocalic components of the labialized phrase, /-$^wV_1$/, /-$^wV_2$/, and /-$^wV_3$/ are added as in (2b) for the /labialized-labialized/ stimulus. We also created cross-spliced tokens of /S*$_1$-/ with /-$^wV_1$/, /-$^wV_2$/, and /-$^wV_3$/, and /S^w*$_1$-/ with /-V_1/, /-V_2/, and /-V_3/ as shown in (2c, d).

(2) Stimuli

a. Plain-Plain: /S*$_1$-V_1/, /S*$_1$-V_2/, /S*$_1$-V_3/

b. Labialized-Labialized: /S^w*$_1$-^{w}V$_1$/, /S^w*$_1$-^{w}V$_2$/, /S^w*$_1$-^{w}V$_3$/

c. Labialized-Plain: /S^w*$_1$-V$_1$/, /S^w*$_1$-V$_2$/, /S^w*$_1$-V$_3$/

d. Plain-Labialized: /S*$_1$-^{w}V$_1$/, /S*$_1$-^{w}V$_2$/, /S*$_1$-^{w}V$_3$/

As a result, three new tokens were created for each type of the stimulus.

2.1.3 Participants

Fifteen native speakers of Korean participated in both Experiments, all undergraduate students of Hanyang University. All of them are from near Seoul area and speak Seoul dialect. None of the listeners had any known hearing or speaking disorders. All subjects were individually tested. Experiment II was conducted a week after Experiment I.

2.1.4 Procedures

Two Experiments were conducted. Experiment I consisted of the first two types, the /plain-plain/ and the /labialized-labialized/, and Experiment II consisted of the last two Types, the /labialized-plain/ and the /plain-labialized/. Participants took the Experiment I first, and they came back for the Experiment II a week after.

In each Experiment, the stimuli were presented twice in the random order. Since each token was repeated twice and there were 5 speakers and 15 participants, 450 responses (3 tokens* 2 repetition * 5 speakers * 15 participants) were acquired for each type in (2). Listeners were instructed to choose one response out of /S*Vta/ and /S*wVta/. All instructions were given in Korean.

Each Experiment consisted of two parts. The first part was a warm-up test in which participants were asked to listen to two stimuli from each speaker (total 10 stimuli) and to select one of two choices that is the best fit to the sound they hear (/S*V/ vs. /S*wV/). This warm-up test was set up to make participants to get familiarized with the test procedure and also with the speakers' voice. The main test consists of 3 parts in which participants were forced to choose one from each pair /S*e/ vs. /S*we/, /S*a/ vs. /S*wa/, and /S*wə/ vs. /S*ə/. Each subsection of the main test consisted of 60 tokens. The sound of the token was repeated twice.

2.2. Experiment I

In Experiment I, the perception test consists of the stimuli of /plain-plain/ (/S*-V/) and /labialized-labialized/ (/S^w*-^{w}V/) sequences.

2.2.1 Results of /labialized-labialized/ (/S^w*-^{w}e/) vs. /plain-plain/ (/S*-e/)

Mean identification rates are plotted in Figure 1. The value of 1 indicates that all responses to a given stimulus were "labialized" /S*weta/; the value 0 indicates that all responses to a given stimulus were "plain" /S*eta/.
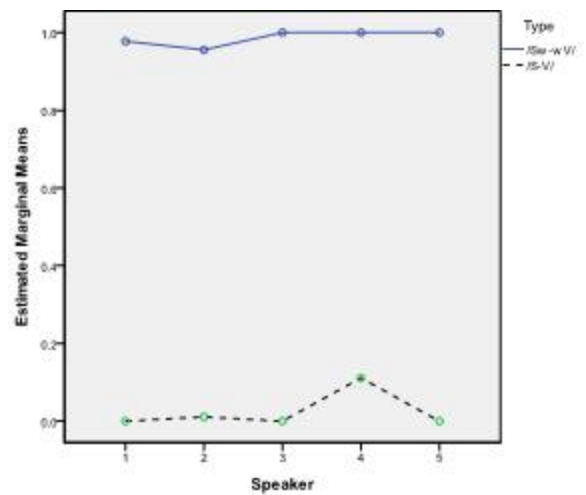

Figure 1. Results of /S^w*-^{w}e/ vs. /S*-e/

With respect to identification rate, there were main effects of Type [F(1, 44) =8119.60, p < .001] and Speaker [F(4, 176) = 7.54, p < .001]. Stimuli composed of /labialized-labialized/ (/S^w*-^{w}e/) were identified by listeners almost always as "labialized" (around 99%) and stimuli of /plain-plain/ (/S*-e/) were identified by listeners almost always as "plain" (around 98%). As for the Speakers, post-hoc tests revealed that identification rates of a stimulus as "labialized" for Speaker 4 were significantly different from all the other Speakers [p<.008-.04]. There was a significant Type*Speaker interaction [F(4, 176) = 5.99, p < .001]. To break the interaction, simple effects were further tested. Simple effects analysis showed that /plain-plain/ Type when produced by Speaker 4 was significantly more perceived as "labialized" (11%) [p < .009-.018] than those by the other Speakers (0%, 1%, 0%, and 0 %

for Speakers 1, 2, 3 and 5 respectively) but there were no other interactions between Type and Speakers.

2.2.2 Discussion

Listeners perceived Type 1 stimuli /Sʷ*-ʷe/ and Type 2 stimuli /S*-e/ as those intended regardless of the speakers. However, the /plain-plain/ (/S*-e/) type stimuli by Speaker 4 were significantly more perceived (11%) as "labialized" than those by the other Speakers. We speculate that Speaker 4 may have some idiosyncratic characteristics to produce /plain-plain/ type stimuli with slight "rounding" qualities which may be misperceived as /labialized-labialized/ type.

2.2.3 Results of /labialized-labialized/ (/Sʷ*-ʷa/) vs. /plain-plain/ (/S*-a/)

Mean identification rates of /Sʷ*-ʷa/ and /S*-a/ are plotted in Figure 2. The value of 1 indicates that all responses to a given stimulus were "labialized" /S*wata/ sequence; the value 0 indicates that all responses to a given stimulus were "plain" /S*ata/ sequence.
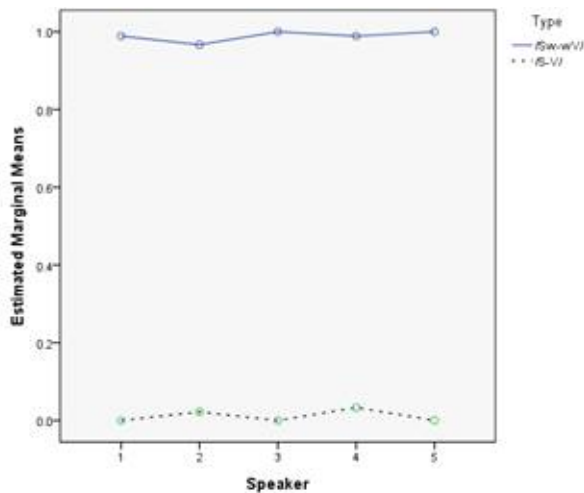

Figure 2. Results of /Sʷ*-ʷa/ vs. /S*-a/

With respect to identification rates, there was a main effect of Type [1, 44] =16072.45, $p < .001$] but no significant effect of Speaker [F(4, 176) = 0.75, $p > .05$]. A post-hoc test revealed that identification rates for two Types were significantly different from each other [$p < .001$]. Stimuli of /labialized-labialized/ components were identified by listeners

almost always as "labialized" (around 99%.); Stimuli with /plain-plain/ were identified by listeners almost always as "plain" (around 99%.). There was no significant Type*Speaker interaction [F(4, 176) = 2.00, p> .05].

2.2.4 Discussion

As is expected, the stimuli /Sʷ*-ʷa/ and /S*-a/ were perceived as they were intended. Unlike /S*-e/ sequence, there was no significant difference between the stimuli of Speaker 4 and those of the other Speakers in /plain-plain/ sequence. It is possible that Speaker 4 have produced less rounding qualities in /S*a/ sequence than in /S*e/ sequence or that the idiosyncratic "rounding" qualities by Speaker 4 might not be strong enough in this context to be misperceived as /S*wa/.

2.2.5 Results of /labialized-labialized/ (/Sʷ*-ʷə/) vs. /plain-plain/ (/S*-ə/)

Mean identification rates are plotted in Figure 3. The value of 1 indicates that all responses to a given stimulus were /S*wəta/ sequence; the value 0 indicates that all responses to a given stimulus were /S*əta/.
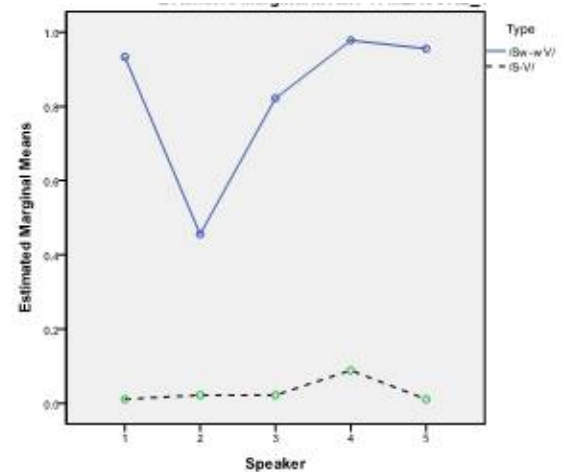

Figure 3. Results of /Sʷ*-ʷə/ vs. /S*-ə/

With respect to identification rates, there were main effects of Type [F(1, 44) =1775.44, $p < .001$] and Speaker [F(3.35, 147.36) = 29.44, $p < .001$]. A post-hoc test revealed that identification rates for two Types were significantly different from each other [$p < .001$]. As for the Speakers, post-hoc tests revealed that identification rates for speaker 2 of a stimulus as

"labialized" were significantly different from all the other Speakers [$p < .001$]. Stimuli by speaker 3 and speaker 4 also showed significantly different identification rates of the stimulus as "labialized" from each other [$p < .01$].

There was a significant Type*Speaker interaction [$F(3.40, 149.68) = 26.16$, $p < .001$]. Simple effects analysis showed that the /labialized-labialized/ (/Sʷ*-ʷə/) type led to lower labialized identification rates, 46%, of a stimulus as "labialized" produced by speaker 2 than by the other speakers [$p < .001$]: 93%, 82%, 98%, and 96 % for speakers 1, 3, 4 and 5 respectively. For the same type of stimuli, the identification rates of the stimulus as "labialized" were marginally different between speaker 3 and speaker 4 [$p < .047$]. There was marginal difference between speaker 4 and the other Speakers [$p < .035$-.038]. Other than that, no significant difference for the identification rates of /plain-plain/ stimulus as "labialized" among the other speakers.

2.2.6 Discussion

As is expected, the stimuli /Sʷ*-ʷə/ and /S*-ə/ were mostly perceived as they were intended. However, the identification rates of a /labialized-labialized/ stimulus as "labialized" produced by speaker 2 were quite low, and somewhat low for Speaker 3 compared to other Speakers. We suggest that listeners perceived less "rounding" qualities in the stimuli by these speakers because they were less articulated in these stimuli. It is also possible that "rounding" qualities are in the process of being "deleted" presently.

2.3 Experiment II

Experiment 1 showed that participants well perceived the spliced original stimuli of /S*-V/ and /Sʷ*-ʷV/ as they were intended except /S*wə/ sequence. In Experiment II, we will examine where in the waveform of /S*wV/ listeners perceive labial characteristics. If labial /w/ was a part of a preceding consonant, labial characteristics should be more or less uniformly realized in all the /S*wV/ sequences with little speaker variability and it will be well proved with the perception test results.

For this purpose, we cross-spliced the frication component of one type of stimuli with the vocalic component of the

corresponding type as was shown in (2c) and (2d). This experiment will lead us to observe whether there are speaker-dependent variances in the realization of the rounding qualities.

2.3.1 Results of /labialized-plain/ (/Sʷ*-e/) vs. /plain-labialized/ (/S*-ʷe/)

Mean identification rates are plotted in Figure 4. Again, the value of 1 indicates that all responses to a given stimulus were /S*weta/; the value 0 indicates that all responses to a given stimulus were /Seta/.
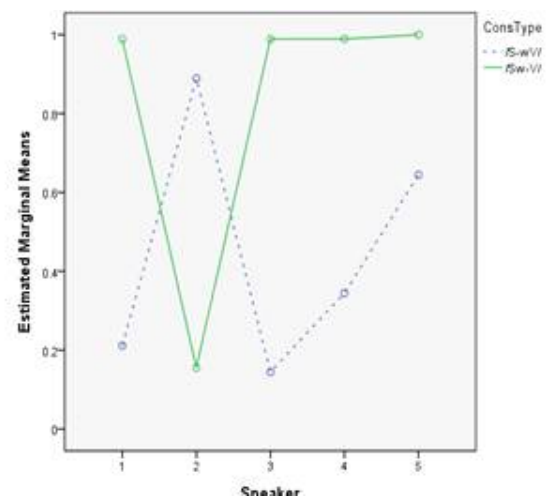

Figure 4. Results of /Sʷ*-e/ vs. /S*-ʷe/

An inspection of Figure 4 reveals that the lines of Type are not parallel, but rather intersect, a cross-over interaction. That is, the /Sʷ*-e/ stimuli led to higher identification rates of "labialized" than the /S*-ʷe/ stimuli produced by all the speakers except speaker 2. As expected, there was significant Type*Speaker interaction [$F(4, 176) = 148.04$, $p < .001$].

Specifically, simple main effects showed that for the /Sʷ*-e/ type, the identification rates as "labialized" produced by speaker 2 were significantly lower (16%) than those by all the other speakers, 99%, 99%, 99% and 100% for Speakers 1, 3, 4 and 5 respectively [$p < .001$], while the /S*-ʷe/ stimuli by speaker 2 showed significantly higher identification rates (89%) of a stimulus as "labialized" than those produced by the other speakers, i.e. 21%, 14%, 34% and 64% for Speakers 1, 3, 4 and 5 respectively [$p < .001$-0.02]. The /S*-ʷe/ stimulus by Speaker 5 also showed significant difference in the identification

rates of a stimulus as "labialized" from those by the other speakers [p < .001-0.02]. The main effect of Type is thus no significant since it depends on speakers we compare.

2.3.2 Discussion

The results showed that listeners perceived significantly more "labialized" sequences from the stimulus if the first component, the frication part, came from the original "labialized" sequence and significantly less "labialized" sequences if the second component, the vocalic part, came from the original "labialized" sequence except when they are produced by speaker 2.

Consequently, we may speculate that listeners perceived more "labialized" sequences from the cross-spliced stimulus of /S^w*-e/ by Speakers 1, 3, 4 and 5 because they distributed more "rounding" qualities on the frication part than on the vocalic part in the original /S*wV/ sequence. In the case of speaker 2, we may conjecture that he overlapped the "rounding" qualities more at the beginning of the vocalic part than on the consonantal part, unlike the other speakers and that this influenced the perception of the listeners. As for speaker 5, we suggest that he might have distributed considerable "rounding" qualities on the vocalic part, if not more, as well as on the frication part, in the original /S*we/ stimulus unlike speakers 1, 3 and 4.

2.3.3 Results of /labialized-plain/ (/S^w*-a/) vs. /plain-labialized/ (/S*-^{w}a/))

Mean identification rates are plotted in Figure 5. As in the previous case, the value of 1 indicates that all responses to a given stimulus were "labialized" /S*wa/; the value 0 indicates that all responses to a given stimulus were "plain" /S*a/.
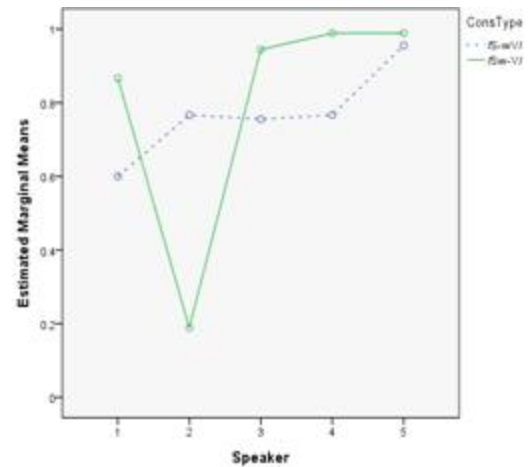

Figure 5. Results of /S^w*-a/ vs./S-^{w}a/

An inspection of Figure 5 reveals that the lines of Type again intersect. As expected, there was a significant Type * Speaker interaction [$F(3.19, 140.30) = 40.12$, p < .001]. The graph showed that the /S^w*-a/ stimuli by speaker 2 led to lower identification rates of "labialized" than by those of the other speakers. The graph also revealed that the effect of Type looked minimal for Speaker 5; the identification rates of the /S^w*-a/ as "labialized' would be slightly higher than those of the /S*-^{w}a/ stimuli (98% and 96%, respectively).

Simple main effects analysis showed that speaker 2 showed significant difference from all the other speakers [p < .001] in the /S^w*-a/ Type; identification rates of the /S^w*-a/ stimulus as "labialized" were 19% when produced by speaker 2 whereas they are 87%, 94%, 99% and 98% when produced by speakers 1, 3, 4 and 5 respectively. It also showed that the /plain-labialized/ type (/S*-^{w}a/) by speaker 5 had significantly higher identification rates of a stimulus as "labialized" [p < .001- .05] than the other speakers. The main effect of Type is non-significant since it depends on which two Speakers we compare as we can see in the graph.

2.3.4 Discussion

The results showed that listeners perceived "labialized" sequences in high percentage from the /S^w*-a/ type when produced by speakers 1, 3, 4 and 5. This supports our suggestion that speakers 1, 3, 4 and 5 might have distributed considerable "rounding" qualities on the frication part, rather than on the vocalic part. Again there was some difference

between speaker 5 on the one hand and speakers 1, 3 and 4 on the other hand. As was the case with /Sʷ*-e/, we suggest that speaker 5 coarticulated much of "rounding" qualities on the vocalic part as well as on consonantal part unlike speakers 1, 3, and 4 who mainly distributed "rounding" qualities on the frication part.

Again, the identification rates of a stimulus of the /Sʷ*-a/ type as "labialized" by speaker 2 was significantly lower for compared to those by the other speakers and we speculate that it was because speaker 2 coarticulated little "rounding" qualities on the frication part.

The identification rates of a stimulus as "labialized" were high (77%) for the /S*-ʷa/ type by speaker 2 as was expected because speaker 2 seem to distribute the "rounding" qualities more at the beginning of the vocalic part than at the frication part. Interestingly, high identification rates of a stimulus as "labialized" for the /S-ʷa/ sequences occurred when they were produced by the other speakers (60%, 76%, 77%, and 96% for Speaker 1, 3, 4 and 5 respectively). We speculate that F2 transition from /w/ to the following /a/ may show rapid rising which takes considerable time to affect the initial part of the vocalic part. Thus, vocalic part of /-ʷa/ may contain enough "rounding" cues for the listeners to perceive the stimulus as "labialized".

2.3.5 Results of /labialized-plain/ (/Sʷ*-ə/) vs. /plain-labialized/ (/S*-ʷə/))

Mean identification rates are plotted in Figure 6. The value of 1 indicates that all responses to a given stimulus were "labialized" /Swə/; the value 0 indicates that all responses to a given stimulus were "plain" /Sə/.

There was a significant Type*Speaker interaction [F(3.26, 143.16) = 31.37, p < .001]. An inspection of Figure 6 reveals that Type variable may not be significant: the identification rates of the /Sʷ*-ə/ stimuli as "labialized" would be slightly higher than those of the /S*-ʷə/ stimuli when produced by speaker 2 and speaker 4 but the difference would be small. In fact, simple effects test showed that they are not significant [p > .71 and p > .77, respectively].
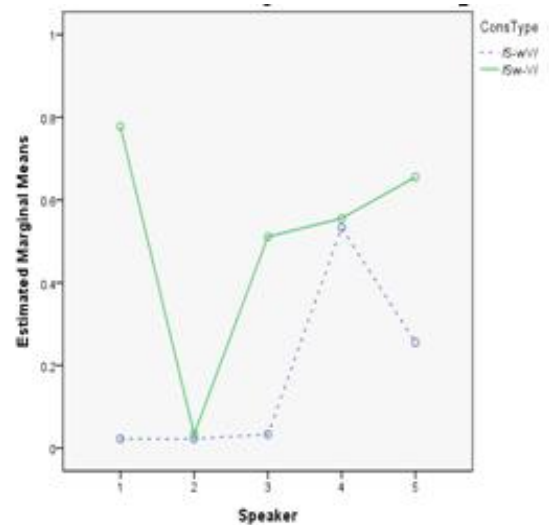

Figure 6. Results of /Sʷ*-ə/ vs. /S*-ʷə/

As for the /S*-ʷə/ type, higher identification rates of a stimulus as "labialized" (53%) occur for speaker 4 compared to all the other speakers, i.e. 2%, 2%, 3%, and 26% for speakers 1, 2, 3 and 5 respectively [p < 0.001-0.01]. There was significant difference in identification rates of a stimulus as "labialized" by speaker 5 from those by the other speakers [p <.001-.01] for the /S*-ʷə/ type. This showed that listeners perceived more "labialized" sequence from the /S*-ʷə/ stimulus if produced by speakers 4 and 5.

As expected, the /Sʷ*-ə/ type by speaker 2 showed significantly lower identification rates of "labialized" (3%) than that by all the other speakers, 78%, 51%, 56%, and 66% for speakers 1, 3, 4 and 5 respectively [p < 0.001], and speaker 1 showed significant difference from speaker 2, 3 and 4 [p < .01-.05] but not from speaker 5 [p > .62].

2.3.6 Discussion

As expected, the results showed significantly low identification rates of "labialized" when listeners heard /Sʷ*-ə/ produced by speaker 2 compared to those by the other speakers. Interestingly, however, the identification rates of the stimulus /S*-ʷə/ by speaker 2 as "labialized" is quite low contrary to the expectation. Since the /-ʷə/ comes from the original stimulus, /Sʷ*-ʷə/, and that labialized vocalic part /-ʷə/ was expected to contain most "rounding" qualities for speaker 2, we expected that listeners would perceive much rounding characteristics from the vocalic part /-ʷə/. This was not the case, however. We

suggest that this is possibly due to less rounding qualities in the original /Sʷ*-ʷə/ sequence than in the other the /Sʷ*-ʷV/ sequences produced by speaker 2. In fact, Experiment I showed that the identification rates of /Sʷ*-ʷə/ sequence as "labialized" by speaker 2 are mere 46% whereas those of /Sʷ*-ʷe/ and /Sʷ*-ʷa/ by speaker 2 are 96% and 97% respectively. We suggest that /w/ is in the process of being deleted for speaker 2 in /S*wə/ context..

As for speaker 5, we have already shown that listeners well perceived the stimulus as "labialized" when either component of /Sʷ*-ʷV/ is included in the cross-spliced stimulus.

It is interesting to note that the identification rates of /S*-ʷə/ as "labialized" produced by speaker 4 were higher than those produced by speaker 5. For the highest identification rates of /S*-ʷə/ as "labialized" produced by speaker 4, we suggest that it is due to the idiosyncratic characteristics in producing /SV/ sequences with some rounding qualities on the plain /S/ by speaker 4.

2.4 General Discussion

In Experiment I, it was shown that sound /w/ was well perceived in /S*we/ and /S*wa/ sequences. However, in /S*wə/ sequences, listeners had more trouble in perceiving /w/, in particular in the sequences produced by speaker 2. This shows that /w/ may be in the process of being deleted currently depending on the following vowel and the speaker.

In Experiment II, we have shown that /Sʷ*-V/ Type resulted in higher identification rates of "labialized" than the /S*-ʷV/ Type when produced by speakers 1, 3 and 4. That is, listeners perceived strong distinctive cues of /w/ mostly from the consonantal part and only weak acoustic cues from the vocalic part in the speech of these speakers. In the future production study, we expect to find most of the "rounding" qualities were on the consonantal part in /S*wV/ sequences in these speakers.

On the other hand, /Sʷ*-V/ / Type resulted in significantly lower identification rates of "labialized" than the /S*-ʷV/ Type when produced by speakers 2. We again presume that listeners perceived main cues of /w/ in speaker 2 from the vocalic part and only some weak redundant cues of /w/ from the consonantal part. In the production study, we expect to find that most of the "rounding" qualities were distributed at the vocalic part in /S*wV/ sequences in speaker 2.

We again presume that listeners perceived main cues of /w/ in speaker 2 from the vocalic part and only some weak redundant cues of /w/ from the consonantal part. In the production study, we expect to find that most of the "rounding" qualities were distributed at the vocalic part in /S*wV/ sequences in speaker 2.

Unlike the other speakers, both /Sʷ*-V/ and /S*-ʷV/ Types by speaker 5 induced high identification rates of the stimulus as "labialized". We expect to find in the production test that speaker 5 have distributed rounding qualities both on the friction part and the vocalic part.

The perception test showed that /w/ in /S*wV/ sequence is realized as various acoustic characteristics at various temporal points as long as it occurs between /S*/ and the following vowel /V/, and depending on speakers, main distinctive cues may occur on the consonantal part, on the vocalic part or on both parts. Even though the speech produced by only five speakers were tested in this perception test, it seemed that at least in /S*wV/ context, listeners perceive distinctive cues mostly from the consonantal part rather than from the vocalic part. Still, vocalic parts are sometimes utilized to specify /w/ if speakers prefer to use that portion.

The vowel qualities also seemed to play a role in different responses. For example, if the vowel sound is /ə/, unexpected results occurred: significantly low identification rates of the stimulus /S*-ʷə/ as "labialized" when produced by speaker 2. We suggest a possible on-going process of /w/ deletion for speaker 2 in /S*-ʷə/ and possibly to speaker 3, and other speakers.

3. Conclusion

Glide /w/ in Korean is known to show different acoustic realizations in different contexts: One variant of /w/ showed the formant transition from the lower F2 to the higher F2 of the following vowel (Kang 2006). The other variant of /w/ in /stop-w-V(i, e)/ showed only flat transition of F2 from the release of the stop to the following vowel. This suggested the possibility that the sound /w/ in this context should rather be classified as a part of the preceding consonant like Cʷ or the

following vowel. Infact, Chinese /w/ / was argued to show more or less uniform realization of labial characteristics in /CwV/ sequence and thus was argued that it should be regarded as a part of a complex segment /Cʷ/ (cf. Duanmu, 2000).

The perception test in this paper showed that main cues of /w/ in /S*wV/ are perceived at different temporal points depending on speakers. In /S*-w-V/ contexts, as the preceding consonant /S*/ has its own temporal period, /w/ tends to mainly affect the acoustic characteristics of the preceding consonant /S*/ and listeners perceived the existence of /w/ mostly from these cues in /S*/. When /S*/ is little coarticulated with /w/, listeners perceived the existence of /w/ at the upcoming vocalic part as in the case of speaker 2 and thus we expect to find acoustic characteristics of /w/ at the vocalic part in such sequences. However, we can also tentatively draw the conclusion that main cues of /w/ in /S*wV/ context are mostly distributed on the consonantal part as three out of five speakers are perceived to place most of /w/ acoustic cues on /S*/. Another speaker, speaker 5, also put most of acoustic cues of /w/ on the consonantal part as well though listeners were also able to induce the existence of /w/ from the vocalic parts in /S*wV/ sequence produced by him.

In /stop-w-V/ context which Kang (2006) examined, the preceding consonant stop has only closure duration and thus does not show any acoustic effects it may have due to the following segment. Therefore, main cues of /w/ should be sufficiently realized on the release burst (and in the VOT period if it exists) and the transition of F2 from the release of the stop to the F2 of the upcoming vowel. It will be interesting to see what acoustic cue differences exist at the beginning of the vocalic parts between the vowel in /S*-w-V/ and in /stop-w-V(i, e)/. We expect that the vocalic part in /S*wV/ sequences, however they are realized, will show more or less redundant cues in most speakers whereas that in /stop-w-V(i, e)/ may show more or less distinctive cues.

In sum, the present pilot study experimentally showed that labial characteristics of sound /w/ in Korean cannot be regarded as a part of complex segment in /S*wV/ context: It does not show uniform realization either on the preceding segment or on the following segment. Rather, this study showed that Korean /w/ is an independent segment and as an independent segment,

/w/ can be acoustically realized as various forms in various contexts: The flat transition of F2 shown in /stop-w-V/ context should be regarded as the complex form of /w/ with the upcoming vowel to effectively encode the labial characteristics in the limited temporal period of following vowel only. With the preceding consonant like /S*/, /w/ could be differently realized as longer temporal period is available at nearby segments for the realization of labial characteristics. In the future production study, we will compare the F2 formant transition of the vocalic parts in /S*wV/ sequences produced by five speakers we dealt with in this paper and show whether the differences we expected in this paper are realized. Also we will compare the F2 transition in /S*wV/ sequence with that in /stop-w-V/ sequences.

We also would like to point out that more extensive study should be performed to see whether there is any trend in realizing /w/ at certain specific location of /S*wV/ sequence and whether this trend reflects any age difference.

References

Beckman, M. & Shoji, A. (1984). Spectral and perceptual evidence for CV Coarticulation in devoiced /si/ and. /syu/ in Japanese, *Phonetica*, Vol. 41, 61-71.

Boersma, P. & Weenink, D. (2002). Version 5.1.07 Praat: doing phonetics by computer (Version) [computer program].

Duanmu, S. (2000). The phonology of standard Chinese. New York: Oxford University Press.

Heinz, J. M. & Stevens, K. N. (1961). On the properties of voiceless fricative consonants, *Journal of the Acoustical Society of America*, Vol. 33, 589-596.

Jongman, A., Wayland, R. & Wong S, (2000). Acoustical characteristics of English fricatives, *Journal of the Acoustical Society of America*, Vol. 108, 1252-1263.

Lisker, L. (1957). Minimal cues for separating /w, r, l, y/ in intervocalic position, *Word*, 13, 256-267.

O'Connor, J. D., Gerstman, L. J., Liberman, A. M., Delattre, P. C. & Cooper, F.A. (1957). Acoustic cues for the perception of initial /w, j, r, l/ in English. *Word,* 13, 24-43.

Oh, M. (2010) W-variants in Korean, *Speech Sciences*, 9, 65-73.

Kang, H. (2006). An acoustic study of the perceptual significance

of F2 transition of /w/ in English and Korean, *Speech Sciences*, Vol. 13, 7-21.

Kang, H. (2010a). Coarticulation effects of Korean sibilant /s/ before a high front segment, *Studies in phonetics, phonology and morphology*, Vol. 16, 3-18.

Kang, H. (2010b). Korean sibilant /s/ before a high front and a round segment, *Journal of the Korean society of speech sciences*, Vol. 2, 59-65.

Padgett, J. (2007). Glide, vowels, and features, *Lingua*, 118, 1937-1955.

Sussman, H. M. (1994). The phonological reality of locus equations across manner class distinctions: preliminary observations, *Phonetica*, Vol. 51, 119-131.

Yun, Y. (2005). Phonetic analysis of the Korean glide [w] and [y], *Korean Journal of Linguistics*, 30, 395-421.

• **Kang, Hyunsook**
  Address: Department of English Language and Culture
           Hanyang University, ERICA Campus
           1592, Sa-3 Dong, Sangnok-gu, Ansan-si
           Kyunggi-do
  Affiliation: Hanyang University
  Telephone: +82-31-400-5348
  E-mail: hskang@hanyang.ac.kr