Merger-in-Progress of Tonal Classes in Masan/Changwon Korean*

Akira Utsugi
(RIKEN)

Utsugi, Akira. (2009). Merger-in-Progress of Tonal Classes in Masan/Changwon Korean. ***Language Research*** **45.1, 23-42.**

In Masan/Changwon Korean, a dialect in the South Gyeongsang (Kyungsang) Province, two lexical tonal classes—"Final" and "Medial-Double" (M-Double)—have similar phonetic realizations. In data collection, three sets of materials were recorded from younger speakers: (i) words with particles, (ii) compounds and (iii) sentences with contrastive focus, each set containing either Final or M-Double words. The results revealed that the two classes were indeed under a merger-in-progress in younger speakers. In addition, between-speaker and between-condition (material set) differences were found on how classes merge. The between-condition difference is a challenge for theories on lexical tonal representations.

Keywords: prosody, tone, intonation, merger-in-progress, Gyeongsang (Kyungsang) Korean

1. Introduction

Merger, one of the phenomena in sound change, involves two or more phonological categories that diachronically lose their distinction. However, it is not a sudden change; rather there is a stage of merger-in-progress before two categories are completely merged. Merger-in-progress involves various complicated phenomena that can lead to reconsideration of phonetic and phonological theories. For example, studies by Warren and his colleague on New Zealand English diphthongs have been a useful source for reconsidering the traditional generative phonological viewpoint on the mental lexicon (e.g., Warren & Hay 2006, Hay et al. 2006). The present paper is concerned with merger-in-

* This study is a revised and extended version of a poster presented at the Laboratory Phonology 11 Conference (LabPhon 11), Victoria University of Wellington, Wellington, New Zealand, 1 July 2008, and a paper in Tsukuba Working Papers in Linguistics, Special Issue Dedicated to Professor Hakutarō Jōo, 2009. I would like to thank Bob Ladd, Alice Turk, Bert Remijsen and two anonymous reviewers of LabPhon 11 for their invaluable comments. I would like to specially thank consultants of the research and Yasuo Akizuki who helped me recruit the consultants. This work was supported by JSPS Postdoctoral Fellowships for Research Abroad (1/2007 - 1/2009).

progress of tonal classes and its influence on phrase-level (postlexical) variation in Masan/Changwon Korean. The subsequent section overviews lexical and postlexical phenomena and their change in this dialect. Sections 3 and 4 report on a phonetic study showing the merger-in-progress.

2. Overview of Masan/Changwon Korean Prosody

It is widely known that Korean dialects spoken in the North and South Gyeongsang (or Kyungsang) Province employ pitch for lexical purposes unlike most of the other dialects (see Fukui 2003 for a review). The Masan/Changwon dialect is one of the South Gyeongsang dialects spoken in two neighbouring cities—Masan and Changwon.[1] Previous litreature has examined the prosodic system of this dialect. Since there seem to be generational differences, we shall first overview the system in the speech of older speakers (Section 2.1) and then in the speech of younger speakers (Section 2.2).

2.1. Prosody in the Speech of Older Generation

The prosodic system of the older generation of Masan/Changwon Korean speakers has been reported by C-G Gim (1970) and YS Kang (2005). Although it is not explicitly stated in his study, C-G Gim's study seems to be based on his intuitions as a native speaker of this dialect (He was born in Changwon, 1939.). YS Kang's study is based on data from speakers whose ages range from the 60's to the 80's, and who are from the region around Masan and Changwon, including Haman, Uireong, Jindong, and Jinhae.[2] Their results agree on major points. The approximate results are shown in Table 1 with some modifications by the author, which are explained later. The modifications have been made for convenience and do not reflect my theoretical viewpoint.[3] Labels shown in the first column are named after the author for convenience.

---

[1] In previous literature, "the North Gyeongsang (Kyungsang) dialect" and "the South Gyeongsang dialect" have often been used as dialect names based on province names. However, several studies have reported within-province differences in tonal phenomena. For example, the Jinju dialect in the western part of South Gyeongsang Province is known to have different tonal phenomena from dialects in the Eastern part of South Gyeongsang Province (Ooe 1977). Thus, we rather use city-based dialect names such as the Masan/Changwon dialect. Although we focus on the two cities in this paper, the same tonal system may be found in some neighbouring cities/towns.

[2] Jindong town was incorporated into Masan city in 1995.

[3] In this study, fully specified tonal sequences such as HLLL and LHHHL are used to describe pitch, only for convenience. I do not consider it to be the case that tones are fully specified on every syllable in underlying and surface representations. Rather, I support the phonetic underspecification approach as in Pierrehumbert and Beckman's (1988) model of Tokyo Japanese.

Table 1. Masan/Changwon Korean Prosodic System in Citation Form, Based on Previous Studies of Older Speakers

	Number of Syllables			
	1	2	3	4
Non-Final (-4)				HLLL
Non-Final (-3)			HLL	LHLL
Non-Final (-2)		HL	LHL	*LHHL*
Final	H	*LH*	*LHH*	LHHH
I(nitial)-Double	H	HH	HHL	HHLL
M(edial)-Double	L	*LH*	*LHH*	*(LHHL)*

Two modifications have been made, both of which are related to disagreements between the two reports. First, C-G Gim uses three-level tones (M in addition to L and H). For example, trisyllabic Non-Final (-2) is described as MHM rather than LHL and trisyllabic M-Double as LMM rather than LHH. The four syllable M-Double is described as LMMM, while he elsewhere states that there is a slight fall from the third to the fourth syllable (C-G Gim 1970: 138f., 2002a: 43, 2002b: 266f.). Second, four-syllable M-Double is not found in YS Kang's report.[4]

A significant difference underlies the descriptive differences. C-G Gim heard (or perhaps sensed, based on his intuition) the difference in pitch register between Non-Final/Final and M-Double; peaks and initial valleys in Non-Final/Final are higher than those in M-Double. YS Kang did not identify four-syllable M-Double since she could not employ such register differences in her description, which are the only cues between four-syllable Non-Final (-2) and M-Double.[5]

As shown in italics in Table 1, three pairs of Non-Final/Final and M-Double words are potentially ambiguous: disyllabic Final vs. M-Double, trisyllabic Final vs. M-Double and four-syllable Non-Final (-2) vs. M-Double. Although the four-syllable pair is distinguished only in pitch register, the other pairs are clearly distinguished when post-positional particles are followed (C-G Gim 1970, YS Kang 2005).[6] Thus, YS Kang identifies disyllabic and trisyl-

[4] It should also be noted that H-ending words actually end with a falling pitch in the utterance-final position (YS Kang 2005: 81). This falling pitch is interpreted as a sequence of lexical H and utterance-final L% boundary tone.

[5] With a lack of raw data, we cannot determine the reality of the pitch register. We speculate about three possibilities. First, there is a dialectal difference between C-G Gim and YS Kang's speakers. Second, the register difference is real, which YS Kang failed to hear. Third, Non-Final/Final and M-Double words are different in underlying representations, but neutralised either completely or incompletely in some conditions, yielding no or very small differences in phonetic realisation. Further studies are needed to determine the actual reason.

[6] Monosyllabic Final and I-Double are also distinguished when followed by particles. See C-G

labic M-Doubles.

Tonal patterns of Final words plus particles depend on the type of particle. When a Final word is followed by a particle such as *–i/ka* (nominative), pitch falls from the final syllable of the noun to the first syllable of the particle, as in the second column of (1).[7] When a Final word is followed by a particle such as *–pota* ('rather than') and *–kkaci* ('up to'), pitch falls from the first to the second syllable of the particle, as in the third column of (1). (Data in (1) are taken from YS Kang 2005.) The first type of particle is considered an unaccented particle, while the second type is considered an accented particle whose accent is on the first syllable.

(1)	a.	*palam*	*palam-i*	*palam-pota*
		wind	wind-NOM	wind-than
		LH	LHL	LHHL
	b.	*taynamu*	*taynamu-ka*	*taynamu-pota*
		bamboo	bamboo-NOM	bamboo-than
		LHH	LHHL	LHHHL

On the other hand, tonal patterns of M-Double words are constant irrespective of the type of particle. It is always LHH when the prosodic word (i.e., the combination of a noun and a particle in this case) is trisyllabic, and is LHHL when the prosodic word is four-syllable, as in (2).

(2)	a.	*salam*	*salam-i*	*salam-pota*
		person	person-NOM	person-than
		LH	LHH	LHHL
	b.	*holangi*	*holangi-ka*	*holangi-pota*
		tiger	tiger-NOM	tiger-than
		LHH	LHHL	LHHLL

Thus, the following differences are found between Final and M-Double:

(3)	a.	Disyllabic word		
		Final:	*palam-i*	LHL
		M-Double:	*salam-i*	LHH
	b.	Trisyllabic word		
		Final:	*taynamu-pota*	LHHHL
		M-Double:	*holangi-pota*	LHHLL

Gim (1970) and YS Kang (2005) for details.

7 A particle *–i* is attached after a noun ending with a closed syllable, while *–ka* is attached after the one ending with an open syllable.

In addition to the case of noun–particle combination, differences between the disyllabic Final and M-Double are found in compounds and focused sentences.

Compound rules of Masan/Changwon and the neighbouring dialects have been studied by many researchers (e.g., Hayata 1974). Focusing on patterns that are directly related to the following discussion, the rules are summarised as follows: When the first constituent is Final and the second constituent is Final or Non-Final, the initial L of the second constituent changes to H.[8] On the other hand, when the first constituent is Final and the second constituent is I-Double or M-Double, pitch falls from the final syllable of the first constituent to the first syllable of the second constituent. When the first constituent is M-Double, the whole word has an M-Double pattern (i.e., L, LH, LHH, etc.), irrespective of the tonal class of the second constituent.[9]

The prosody of focused sentences has been studied by C-G Gim (1970, 2002b). According to C-G Gim (2002b: chapter 9), when a word is focused, either 'reduction' (*yakhwa*) or 'combination' (*kyelhap*) takes place. The former means a compression of pitch range in post-focal words. The latter indicates a phenomenon in which a focused word and post-focal words form a tonally unified prosodic unit.[10] Even though the pattern of pitch in the latter case is not described explicitly, it seems that rules similar to compounds are applied within the unified prosodic unit.

The differences between Final and M-Double described thus far are well accounted for when we assume that the Final and Non-Final have a lexical pitch accent, H*+L, while M-Double and I-Double have word melodies, Lω+H+H+L and Hω+H+L respectively, following Utsugi (2007, in preparation). Although Utsugi's (2007, in preparation) view was originally proposed for a tonal system in younger speakers, it actually fits the older generation better. A lexical pitch accent is a tonal element associated with a certain tone-bearing unit (TBU). On the other hand, a word melody is a tone or a sequence of tones, which primarily links a prosodic-word node, yielding a constant pitch contour irrespective of the number of syllables.[11] This view explains why M-

---

[8] It is considered to be the case that the initial L in Non-Final and Final is a phrasal tone rather than a word-level tone. This viewpoint accounts for the lack of L in the second element of compounds. This also accounts for the pitch patterns in post-focal words described in the subsequent paragraph.

[9] Previous studies do not have many examples of compounds consisting of two disyllabic constituents, which are the primary concern in this study. There is a need to collect more data from other speakers.

[10] The two phenomena seem to be the same as "compressed" and "wrapped" patterns in Utsugi's (2007, in preparation) study of the speech of the younger generation of Masan/Changwon Korean. The latter is reminiscent of "[+ Multiword AP]" in Igarashi's (forthcoming) typological study as well.

[11] See Utsugi (2007, in preparation) for details. Related discussions are found in Gussenhoven (2004: 30ff) and Igarashi (forthcoming). The concept of 'word melody' is similar to "N-type accent" in Uwano (1984, 1989, and 1999) and "word tone" in Hayata (1999a, 1999b) and

Doubles always have LHHL(…) pitch patterns irrespective of the number of syllables and the nature of the following constituents. It also explains the register difference between Final and M-Double in C-G Gim (1970); it is probable that different tonal entities, H*+L and Lω+H+H+L, have different phonetic realizations, similar to the phonetic difference between accentual H and phrasal H in Tokyo Japanese (Pierrehumbert & Beckman 1988).

2.2. Prosody in the Speech of Younger Generation

A prosodic system as well as several lexical and post-lexical tonal phenomena in the younger generation was studied by Utsugi (2007, in preparation), based on the data from two speakers born in the 1980s. Several differences were found compared to the system in the older generation which was reported in the previous studies. First, the difference of pitch register between Non-Final/Final and M-Double was not found.[12] Second, in four-syllable words, the originally Non-Final (-2) words and originally M-Double words seemed to be completely merged. There were no differences found between the two types of words in both citation form and sentence level production.[13] Third, in trisyllabic words, a difference such as (3b) was not found between Final and M-Double, suggesting that merger is completed in trisyllabic Final vs. M-Double as well.

In disyllabic words, a difference such as (3a) was found in younger speakers in Utsugi (2007, in preparation). However, my recent investigation involving more speakers and more materials has revealed that disyllabic words are indeed under a merger-in-progress. This study reports such recent results. In this study, I report not only citation form results but also phenomena in noun–particle combinations, compounds, and focused sentences.

3. Methods

In this Section I report on an experiment that shows the merger-in-progress of Final and M-Double disyllabic words among younger speakers of Masan/Changwon Korean. Section 4 presents the results.

Donohue (1997).

[12] A similar report is found in Kenstowicz et al.'s (2008) study of the Busan dialect, one of the South Gyeongsang dialects.

[13] Since Masan/Changwon Korean tones have regular correspondence with tones in other modern Korean dialects as well as Middle Korean, original tonal classes are extrapolated from those tones. In compounds, compound tone rules can also be a clue for the extrapolation.

3.1. Consultants

The consultants were seven native speakers of Masan/Changwon Korean. They were all students of Chang-Shin College, Masan, South Korea. Demographic information about these consultants is shown in Table 2.

Table 2. Information on Consultants

Speaker No.	Sex	Year of birth	City where born and grew up
1	female	1987	Masan
2	female	1988	Changwon
3	female	1988	Masan
4	male	1987	Masan
5	male	1982	Masan
6	female	1987	Masan
7	female	1988	Changwon

3.2. Materials

Three sets of materials were used in the recordings: Set 1 consisted of words followed by particles, Set 2 consisted of compounds and Set 3 consisted of focused sentences. All the sets included the same words: four originally Final disyllabic words and four originally M-Double disyllabic words. These words are shown in Table 3. Two more words were used in the recordings but were not analysed since those words were often pronounced in unexpected tonal patterns such as Non-Final or I-Double.

Table 3. Words Used in Material Sets

Original tonal class	Word	IPA	Meaning
Final	*kaul*	[kaɯl]	autumn
	sakwa	[sagʷa][a]	apple
	kwutwu	[kudu]	shoes
	namu	[namu]	tree
M-Double	*tampay*	[tambe]	cigarette
	taychwu	[tetɕʰu]	jujube
	ankyeng	[angʲʌŋ]	glasses
	hopak	[hobak˺]	courgette

[a] [gʷa] is often pronounced as [ga] in Masan/Changwon Korean as well as most of the other Gyeongsang dialects.

Set 1 consisted of noun–particle combinations. Nouns shown in Table 3 were followed by a nominative particle *–i/ga*.

Set 2 consisted of compounds whose first element was one of the words shown in Table 3 and whose second element was either disyllabic Final or M-Double. Table 4 illustrates the compounds used in Set 2.

Table 4. Compounds Used in Set 2

Original tonal class of the first constituent	Word	IPA	Meaning
Final	*kaulpalam*	[kauɯlp*aɾam][a]	autumn wind
	sakwanamu	[sagʷanamu][b]	apple tree
	kwutwukakey	[kuduk*age]	shoe shop
	namukaci	[namuk*adʑi]	branch
M-Double	*tampaykakey*	[tambek*age]	cigarette shop
	taychwunamu	[tetɕʰunamu]	jujube tree
	ankyengtali	[angʲʌŋt*aɾi]	temples (of glasses)
	hopaknamul	[hobaŋnamul]	courgette side dish

[a] [*] stands for a tense obstruent.

[b] [gʷa] is often pronounced as [ga] in Masan/Changwon Korean as well as most of the other Gyeong-sang dialects.

Set 3 consisted of sentences that included words shown in Table 3. Each sentence consisted of an object, a verb and a quotative *–ko hayssta* ('(I) said that …'), with focus on the object, as in (4). (Words in uppercase letters indicate focused words.)

(4) *SAKWA* *meknunta-ko* *hayssta.*
apple eat-QUOT said
'I said I would eat an apple.'

The focused utterances were elicited by playing question sentences recorded by another native speaker. For example, to elicit utterance (4), a question sentence (5) was played.

(5) *mwe* *meknunta-ko* *hayss-no?*
what eat-QUOT said-Q
'What did you say you would eat?'

The material sentences are shown in Table 5.

Table 5. Sentences in Set 3

Original tonal class of the first constituent	Sentence	
Final	*KAUL salanghantako hayssta.*	'I said I love autumn.'
	SAKWA meknuntako hayssta.	'I said I would eat an apple.'
	KWUTWU patnuntako hayssta.	'I said I would receive shoes.'
	NAMU naluntako hayssta.	'I said I would carry a tree.'
M-Double	*TAMPAY patnuntako hayssta.*	'I said I would receive a cigarette.'
	TAYCHWU meknuntako hayssta.	'I said I would eat a jujube.'
	ANKYENG pakkuntako hayssta.	'I said I would change glasses.'
	HOPAK meknuntako hayssta.	'I said I would eat a courgette.'

It was expected that the two tonal classes would show different tonal patterns in all three sets if the two classes were not merged. Table 6 illustrates the expected patterns based on the rules summarised in Section 2.1. Note that the originally H-ending words are expected to end with a falling pitch since the utterance-final L% is predicted (see footnote 4).

Table 6. Expected Patterns When the Two Classes Are Not Merged

	Originally Final	Originally M-Double
Set 1 (Particles)	LHL	LHF (or LHH) [a]
Set 2 (Compounds)	LH+HF (or LH+HH) or LH+LL	LH+HL
Set 3 (Focus)	LH#H(…)HL…	LH#HL…

[a] F stands for falling pitch.

3.3. Recording Procedures

The recordings were conducted in a vacant room in Chang-Shin College. In each of the three sets, material words (or sentences in Set 3) were shown on a computer screen one by one in random order. Each set was repeated three times. The recordings were made using a WAV/MP3 recorder (Roland EDIROL R–09) and an electret condenser microphone (Sony ECM-MS957). Sounds were recorded as WAV files (44100Hz, 16bit) directly by the recorder and then downsampled by half using Audacity (open source software).

3.4. Analysing Procedures

An analysis was first conducted using the author's auditory judgement. A tonal pattern of each token was coded in this process. The results are summarised in Section 4.1.

As a non-native speaker of Masan/Changwon Korean, my coding could have been subjective and inaccurate. Thus, an acoustic analysis was also conducted. In the analysis, the detection of shoulder turning points was conducted using two-piece linear regression in the following manner. First, F0 contours were extracted from WAV files by an autocorrelation algorithm on Praat (software developed by Paul Boersma and David Weenink). F0 values computed during, 10 ms before and 10 ms after pronunciation of obstruents were removed to avoid microprosodic perturbations. Then turning point detection was conducted using ExcelVBA. The algorithm was essentially the same as the one used in Utsugi and H Jang (2008).[14] See Appendix for details. Input for the algorithm was the pitch maximum in the second syllable to the minimum in the final syllable. In Set 3, the final word *hayssta* was ignored.

4. Results

4.1. Results Based on Auditory Judgements

As stated in Section 3.2, it was expected that the originally Final and M-Double words would show different pitch patterns. Hereafter, we call the patterns expected for originally Final words and those for originally M-Double words as "Final patterns" and "M-Double patterns" respectively.

Figure 1 shows the results based on the auditory judgements.[15] If the two classes were not merged, it would be expected that all the originally Final words (left bars in each speaker) would show the Final patterns (black) and all the originally M-Double words (right bars) would show the M-Double patterns (grey). If the two classes were completely merged, it would be expected that all the words (both of the left and right bars) would show either the Final (black) or M-Double (grey) patterns. As can be seen, none of these two expectations was supported, suggesting that the two classes are neither distinguished nor merged, but under a merger-in-progress.[16]

[14] Note that the algorithm in Utsugi and H Jang (2008) is made to detect elbow turning points. In this study, the algorithm was modified to detect shoulder turning points.

[15] As can be seen in the figure, patterns that are neither Final nor M-Double were also found. These were labeled as "Others" in the figure. A major case of "Others" was a pattern that has a shoulder at the boundary between an object and a verb in Set 3. This is probably due to the pitch range compression in post-focus words. The same pattern was reported in Utsugi (2007, in preparation) with the name "reduced peak".

[16] In words beginning with /h/ or /s/, high pitch was sometimes heard although low pitch was expected. This is reminiscent of Accentual-Phrase-initial high tone in phrases beginning with aspirated and forced obstruents, /s/ and /h/ in Seoul and South Jeolla Korean (S-A Jun 1993, 1996). Similar effects were reported for North and South Gyeongsang dialects by Kenstowicz and C Park (2006). It is assumed that consonantal effects on prosody are getting stronger in younger Gyeongsang speakers, similar to that in speakers of other dialects. In the present analy-

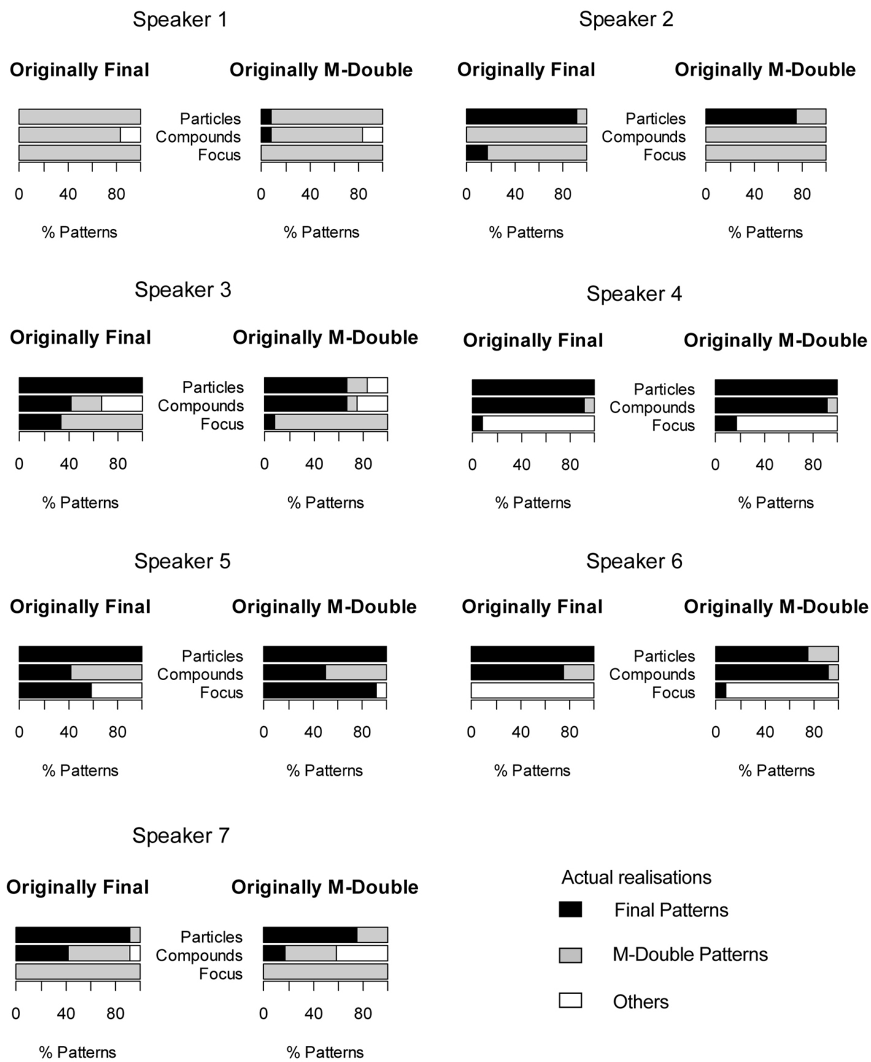

Figure 1. Results based on auditory judgements.

Two interesting emerge from looking closely at Figure 1. First, there is a clear between-subject difference. Speaker 1 favours the M-Double patterns while Speakers 4 and 6 favour the Final patterns. Second, clear between-set differences are found in some speakers. For example, in Speaker 2, the Final patterns are dominant in Set 1 (Particles) while the M-Double patterns are dominant in Set 2 (Compounds) and Set 3 (Focus).

---

sis, high-pitch-beginning pronunciations after /s/ and /h/ are regarded as phonetic variants of low-pitch-beginning pronunciations.

4.2. Results Based on Acoustic Analysis

As noted in Section 3.4, turning point detection was conducted to analyse data in an objective manner. In this analysis, some tokens were not included. First, tokens labelled as "Others" in the auditory judgements were not included. Second, between two Final patterns, LHHF (LHHH with L%) and LHLL, the latter pattern was not included in Set 2 in order to simplify our results.[17] Third, tokens whose F0 contour did not appear clearly due to glottalisation or devoicing were not included.

In addition to these exclusions, there were several tokens in which the turning point detection was not successful. This occurred in two cases. The first case was a case of less than six F0 points. In this case, two-piece linear regression cannot be applied. The second case was a case in which the algorithm returned an error. As in the Appendix, the algorithm returns an error when a contour is not convex.

Figures 2-8 show the results of turning point detection in each speaker. In Set 1 (Particles), the Final patterns (LHL) predict earlier turning points than the M-Double patterns (LHF) do. On the other hand, in Set 2 (Compounds) and Set 3 (Focus), the Final patterns predict later turning points than the M-Double patterns do. Since the *x*-axis is reversed in only the top panels of the figures (panels for Set 1), the more right-aligned points suggest the more Final-like in each panel.

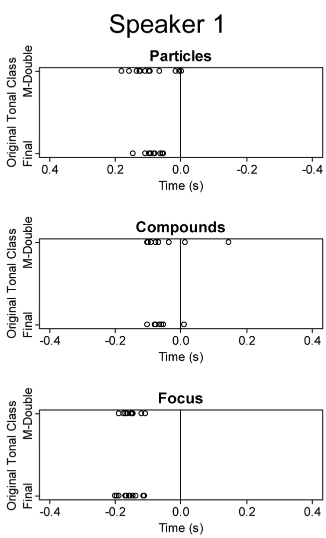

Figure 2. Results based on the detection of shoulder turning points (Speaker 1).

[17] Only a small number of tokens were pronounced with the LHLL pitch pattern: 1 token in Speaker 1 and 7 tokens in Speaker 7.

In Figure 2 (Speaker 1), most of the turning points were aligned on the left. This result is consistent with the result in the auditory judgements; most of the tokens were judged as the M-Double patterns.

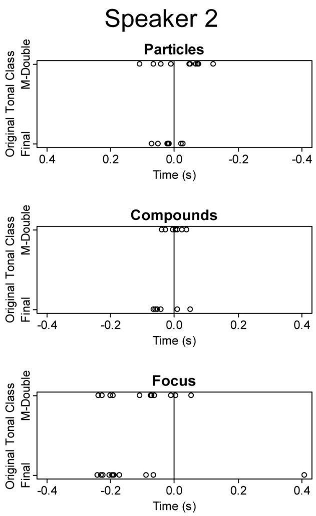

Figure 3. Results based on the detection of shoulder turning points (Speaker 2).

In Figure 3 (Speaker 2), the turning points in the top panel (Set 1, Particles) and the second panel (Set 2, Compounds) were aligned around the centre, while most of the points in the bottom panel (Set 3, Focus) were aligned on the left. Although the turning-point results agreed with the judgement results in Set 3, there were inconsistencies between the two results in Set 1 and Set 2. The turning point should have appeared more to the right in Set 1 and more to the left in Set 2.

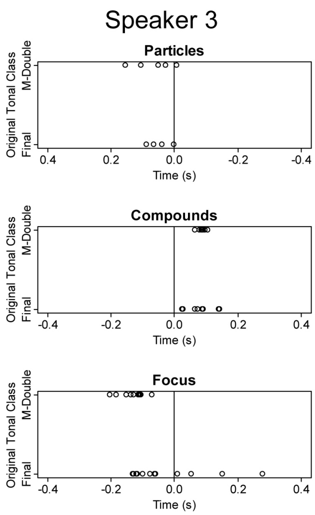

Figure 4. Results based on the detection of shoulder turning points (Speaker 3).

In Figure 4 (Speaker 3), most of the turning points in the top (Set 1, Particles) and bottom (Set 3, Focus) panels were aligned on the left, while most of the points in the middle panel (Set 2, Compounds) were aligned on the right. The result of Set 1 is inconsistent with that of the auditory judgements; most of the tokens in Set 1 were judged as the Final patterns.

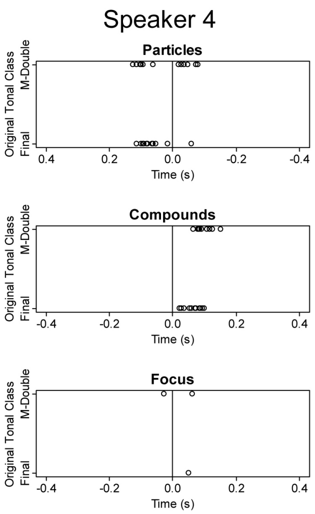

Figure 5. Results based on the detection of shoulder turning points (Speaker 4).

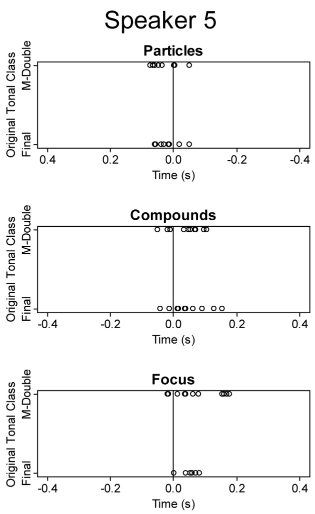

Figure 6. Results based on the detection of shoulder turning points (Speaker 5).

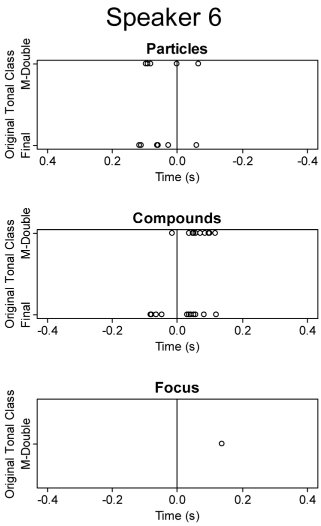

Figure 7. Results based on the detection of shoulder turning points (Speaker 6).

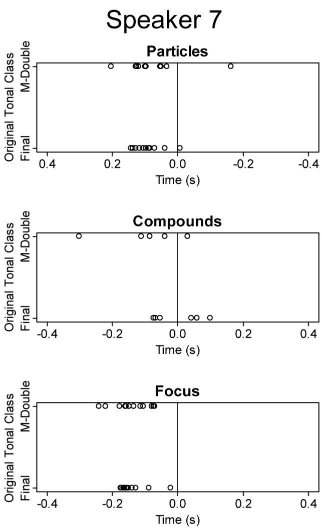

Figure 8. Results based on the detection of shoulder turning points (Speaker 7).

The inconsistencies between the two results are found in Figures 5-8 as well, especially in Set 1 and Set 2. This inconsistency could be due to glottalisation at the utterance-final syllable, which distorted F0 contours, and a small phonetic difference between HL and HF. The turning point detection was successful for Set 3 probably because the target phrases were not located at the utterance-final position and the difference between the Final and M-Double patterns were phonetically large.

Although we have methodological problems in the turning point detection, we can still see the interesting findings stated in Section 4.1. The between-subject differences can be seen clearly in Figures 2-8. The between-set differences are also found (e.g., between Set 2 and Set 3 in Figure 4). These effects were supported by ANOVA.[18] The three-way ANOVA (random factor: [speaker], fixed factor: [original tonal class] and [material set]) revealed that there was a significant main effect of the speaker on the alignment of turning points ($F(6, 10.649) = 4.430$, $p = 0.017$) and a significant main effect on the material set ($F(2, 13.849) = 5.275$, $p = 0.020$). In addition, there was a significant interaction effect between the speaker and material-set factors ($F(12, 12.410) = 5.167$, $p = 0.004$). The main effect of the original tonal class and the other interaction effects were not significant.

To summarise, the results based on the turning point detection were broadly consistent with those based on the auditory judgements. However, some inconsistencies were identified. Methodological improvements are necessary to resolve this issue.

5. Discussion

The results suggest that the two patterns having diachronically different origins co-exist in the dialect, and even in the idiolect of speaker in some cases. The preference of the patterns differs with speakers and material sets. It is unclear whether the patterns bear different intonational meanings and/or sociolinguistic indices or are only selected at random by the younger generation.

The between-set variation suggests that the word-level and phrase-level variations go in different directions, raising an interesting question concerning phonological theories. As referred in Section 2.1, Utsugi (2007, in preparation) proposed an intonational phonological model in which the Final pattern has the H*+L lexical pitch accent, whereas the M-Double pattern is associated with Lω+H+H+L lexical word-edge tones from the left edge of the phonological word, in order to account for different postlexical behaviours between the two. The question is how we reconcile such a model with the between-set variation observed in this study. One possibility is that the two classes are indeed completely merged, having either H*+L or Lω+H+H+L in the underlying representation, and the variation is due to the phonological grammar. However, such a grammar must be bizarre. We cannot account for the fact that the alignment of the shoulder turning point is earlier in some cases and later in the others.

Another possibility would be to assume that each word has two underlying

[18] Before conducting ANOVA, −1 was multiplied to values in Set 1 to compare Set-1 results with results in other sets. As stated in the third paragraph of Section 4.2, the relationship between the expected alignments according to the tonal class were opposite between Set 1 and Sets 2 and 3.

representations, H*+L and Lω+H+H+L, at the same time, one of which is realized at the surface according to certain conditions. Such an approach is a departure from the traditional generative phonology, but can be reconciled with the exemplar-based model (e.g., Pierrehumbert 2001, Bybee 2003). This question remains open at this point.

6. Conclusion

This study reported a merger-in-progress of the two tonal classes, Final and M-Double, in Masan/Changwon Korean. The results of the study revealed that patterns which originated from the Final and M-Double co-exist in this dialect. In addition, between-speaker and between-condition (material set) variations were found.

As this is a preliminary study, further studies need to be conducted. We need to obtain data from older speakers to draw the entire picture of the change. In addition, we need methodological improvements, examinations of factors affecting the variation, such as intonational and sociolinguistic factors, and theoretical modelling, which may lead to reconsideration of current phonological theories.

References

Bybee, Joan. (2003). *Phonology and Language Use*. Cambridge University Press.

D'Imperio, Mariapaola. (2000). *The Role of Perception in Defining Tonal Targets and Their Alignment*. Ph.D. dissertation. Ohio State University.

Donohue, Mark. (1997). Tone systems in New Guinea. *Linguistic Typology* 1, 347-386.

Fukui, Rei. (2003). Pitch accent systems in Korean. In Shigeki Kaji, ed., *Cross-linguistic Studies of Tonal Phenomena: Historical Development, Phonetics of Tones, and Descriptive Studies*, 275-286. Research Institute for Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.

Gim, Cha-Gyun. (1970). Kyengnam pangen-uy sengco yenkwu. *Hangul* 145, 109-149.

Gim, Cha-Gyun. (2002a). *Kwuk.e Pangen Sengcolon*. Yeklak.

Gim, Cha-Gyun. (2002b). *Yenghonam Pangen Wun.yul Pikyo*. Yeklak.

Gussenhoven, Carlos. (2004). *The Phonology of Tone and Intonation*. Cambridge University Press.

Hay, Jennifer, Paul Warren and Katie Drager. (2006). Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics* 34, 458-484.

Hayata, Teruhiro. (1974). Accent in Korean: Synchronic and diachronic studies. *Gengo Kenkyu (Journal of the Linguistic Society of Japan)* 66, 73-116.

Hayata, Teruhiro. (1999a). *Ontyoo no taiporojii*. Taisyuukan Syoten.

Hayata, Teruhiro. (1999b). Accent and tone: Towards a general theory of prosody. In Shigeki Kaji, ed., *Cross-linguistic Studies on Tonal Phenomena: Tonogenesis, Typology,* *and Related Topics*, 221-234. Research Institute for Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.

Igarashi, Yosuke. (forthcoming). Typology of intonational phrasing in Japanese dialects. In Sun-Ah Jun, ed., *Prosodic Typology,* V. 2. Oxford University Press.

Jun, Sun-Ah. (1993). *The Phonetics and Phonology of Korean Prosody*. Ph.D. dissertation. Ohio State University.

Jun, Sun-Ah. (1996). Influence of microprosody on macroprosody: A case of phrase initial strengthening. *UCLA Working Papers in Phonetics* 92, 97-116.

Kang, Young Suk. (2005). Kankokugo no toobu keesyoonandoo hoogen no akusento taikee. *Tookyoodaigaku Gengogaku Ronsyuu (Tokyo University Linguistic Papers)* 24, 79-97.

Kenstowicz, Michael, Hyesun Cho and Jieun Kim. (2008). A note on contrasts, mergers, and acquisitions in Kyungsang accent. *Toronto Working Papers in Linguistic* 28, 107-122.

Kenstowicz, Michael and Chiyoun Park. (2006). Laryngeal features and tone in Kyungsang Korean: A phonetic study. *Studies in Phonetics, Phonology and Morphology* 12.2, 247-264.

Ooe, Takao. (1977). Chinju hoogen no akusentogata ni tuite. *Gengo Kenkyu (Journal of the Linguistic Society of Japan)* 71, 1-20.

Pierrehumbert, Janet. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In Joan Bybee and Paul Hopper, eds., *Frequency and the Emergence of Linguistic Structure*, 137-157. John Benjamins.

Pierrehumbert, Janet and Mary E. Beckman. (1988). *Japanese Tone Structure*. MIT Press.

Utsugi, Akira. (2007). The interplay between lexical and postlexical tonal phenomena and the prosodic structure in Masan/Changwon Korean. A paper presented at the ICPhS 2007 Satellite Meeting: Workshop on "Intonational Phonology: Understudied or Fieldwork Languages," Saarbrücken, Germany.

Utsugi, Akira. (in preparation). Lexical pitch accent, word melody and intonation in Masan/Changwon Korean.

Utsugi, Akira and Hyejin Jang. (2008). Lexical pitch accent and tonal targets in Daegu Korean. Ms, University of Edinburgh.

Uwano, Zendo. (1984). N-kee akusento-no ippan tokusee-ni tuite. In Hirayama Teruo Hakusi Koki Kinenkai, ed., *Gendai hoogengaku-no kadai*, V. 2, 167-209. Meezi Syoin..

Uwano, Zendo. (1989). Nihongo-no akusento. In Miyoko Sugito, ed., *Kooza nihongo-to nihongo kyooiku*, V. 2, 178-205. Meezi Syoin.

Uwano, Zendo. (1999). Classification of Japanese accent systems. In Shigeki Kaji, ed., *Cross-linguistic Studies of Tonal Phenomena: Tonogenesis, Typology, and Related Topics*, 151-178. Research Institute for Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.

Warren, Paul and Jennifer Hay. (2006). Using sound change to explore the mental lexicon. In Claire Fletcher-Flinn and Gus Haberman, eds., *Cognition and Language: Perspectives from New Zealand*. Bowen Hills, Australian Academic Press.

Appendix: Algorithm for Turning Point Detection

The algorithm used in this study is almost the same as one used in Utsugi and H Jang (2008), which is different from the common two-piece linear regression (e.g., Pierrehumbert and Beckman 1988, D'Imperio 2000) in two ways. First, a point outside the horizontal (time) range of the input, which is not what we assume, is avoided in the present algorithm. Second, the present algorithm only detects either of a 'shoulder' or 'elbow' turning point (it is the 'shoulder' in the present study). The part written in bold font is the one newly added by the author. Except for that part, the algorithm is the same as the common two-piece linear regression. (Note that x indicates time and y indicates f0 when this algorithm is applied to an f0 contour.)

[Algorithm]

- Input: a set of points: $\{(x_i, y_i) \mid 1 \le i \le N\}$. ($i$: the point number in the order arranged according to x. N: the total number of points.)
- For each i in which $2 < i < N - 2$,
  - Assume that a turning point is located between (x_i, y_i) and (x_{i+1}, y_{i+1});
  - Fit linear regression lines for two parts, one from (x_1, y_1) to (x_i, y_i) (hereafter, "left"), and the other from (x_{i+1}, y_{i+1}) to (x_N, y_N) (hereafter, "right"), and calculate the following values for each part;
    - Q: the least sum of the squared residuals calculated by the least squares method;
    - a: the intercept;
    - b: the slope;
  - Calculate a coordinate of the intersection of the two linear regression lines (X_i, Y_i);
- **If $X_i < x_0$ or $X_i > x_N$, or if $b_{left} < b_{right}$, then Q_i is not available;** else, $Q_i = Q_{left} + Q_{right}$.
- Select the i whose Q_i is the least, among all of the i's.
- Return (X_i, Y_i) for the i selected above as the turning point. If there is no available Q_i, return an error.

As can be seen, this algorithm can return errors when the contour markedly deviates from a convex contour.

Akira Utsugi
Laboratory for Language Development
Brain Science Institute, RIKEN
2-1 Hirosawa, Wako-shi
Saitama, 351-0198, Japan
E-mail: autsugi@brain.riken.jp

Received: March 30, 2009
Revised version received: June 15, 2009
Accepted: June 19, 2009