Parallel Voice Onset Time shift in Chinese Korean

A case for linguistic drift

Wenhua Jin and David J. Silva
Kennesaw State University / Salem State University

This study reveals the existence of a Voice Onset Time shift in the Korean spoken by native speakers residing in northeast China, a shift parallel to those reported in other Korean varieties in Korea, the USA, and Canada. The VOT pattern observed in the Chinese Korean community is argued to represent a change that cannot be simply explained in terms of diffusion via recent dialect contact, or as a feature directly inherited from the source language when it was transplanted into China over a century ago. We suggest that behind the parallel VOT shifts is the power of “drift” that drives the different Korean varieties along similar journeys of language evolution. This study presents an intriguing case where internal changes driven by “drift” may actually be initiated and further supported by language/dialect contact.

Keywords: Voice onset time (VOT), drift, parallel shift, Chinese Korean, language change

1. Introduction

Korean is spoken all over the world by more than 77 million people (Lewis, Simons, & Fennig, 2016). In addition to the nearly 70 million speakers of the language residing on the Korean Peninsula, the language is spoken by an additional 7 million Koreans living in diaspora: 36.7% in China, 29.82% in the United States, 12.73% in Japan, 2.94% in Canada, 2.52% in Russia, and the remaining 15.29% scattered around the world (Current Status of Overseas Compatriots, 2013). Just as the spread of English beyond Britain has led to the rise of distinct English varieties worldwide, the spread of Korean to locations beyond the Korean Peninsula has given rise to different varieties of the language (Silva, 2011). Given the different geographic and sociolinguistic contexts, studies on different Korean varieties shed
new light not only on the understanding of Korean language evolution, but also on language variation and change in general.

Korean is typologically marked for its three-way-contrastive stop system, one that contrasts lax (or plain) segments from aspirated and tense (or reinforced) counterparts (Table 1). This system of consonants has drawn attention in recent years, especially in terms of reported changes in observed mean values associated with the voice onset time (VOT) of the stops.

Table 1. Korean stop consonants

<table>
<thead>
<tr>
<th></th>
<th>labial</th>
<th>alveolar</th>
<th>velar</th>
</tr>
</thead>
<tbody>
<tr>
<td>lax</td>
<td>p</td>
<td>t</td>
<td>k</td>
</tr>
<tr>
<td>aspirated</td>
<td>pʰ</td>
<td>tʰ</td>
<td>kʰ</td>
</tr>
<tr>
<td>tense</td>
<td>pp</td>
<td>tt</td>
<td>kk</td>
</tr>
</tbody>
</table>

In Seoul Korean, three different VOT patterns have been reported: Traditional, intermediate, and innovative (Silva, 2006, 2011). For the “traditional pattern”, one observes the existence of an overlap in the relatively short VOT between tense and lax Korean stops, though they both can be distinguished from aspirated stops (Han & Weitzman, 1965; Kim, 1965; Lisker & Abramson, 1964). For the “intermediate pattern”, there is a clear three-way VOT distinction among the three stops types, with the lax stops being more aspirated than those in the traditional pattern and standing between the other two categories (Cho, 1996; Han, 1996; Silva, 1992). As for the “innovative pattern”, which predominates in the speech of younger speakers of contemporary Korean, one finds a neutralization of the VOT differences between lax and aspirated stops, suggesting a diachronic shift in VOT which makes VOT no longer a primary cue to distinguish these two phonemic categories. Instead, fundamental frequency (F0) at the onset of the immediately following vowel, which is generally agreed to be the lowest for lax, higher for tense, and highest for aspirated stops, has become the primary cue (Kang, 2014; Kang and Guion, 2006; Kim, Beddor, & Horrocks, 2002; Kim & Stoel-Gammon, 2009; Park & Iverson, 2008; Silva, 2006). Interestingly, this new innovative pattern is also observed among younger Korean speakers in an English-speaking country, Canada (Kang & Nagy, 2016), which is geographically remote from the Korean homeland. Is the same VOT shift also going on in other Korean varieties? If so, what is the nature and mechanism for this change in Korean? With an aim to answer these questions, we turn to the Korean variety spoken in northeast China (Chinese Korean) by members of communities which have robustly maintained the Korean language and culture outside the Korean peninsula (Jin, 2003).
2. Processes leading to language similarity

Before we move on to the current research, it is necessary to review some of the important mechanisms that may lead to similarity between different geographically separated varieties of a single language. Relevant here to our research are three major processes, including transmission, diffusion, and drift.

Labov (2007) defines transmission as the “unbroken sequence of native-language acquisition by children” through the family tree model. The family tree shows how a single language develops dialects that in time become distinct languages through the accumulation of changes and how these daughter languages can split up into daughters of their own. As daughter languages of a parent language, they carry on certain features of the parent language that define them as a language family. Along this line, we assume that different varieties of a single language may also share certain features that are transmitted from their source language.

An alternative process in the explanation of language changes is diffusion, a process whereby the contact and the transfer of features between speech communities often lead to linguistic similarity. Major models proposed for the diffusion process include the wave model (Bailey, 1973; Labov, 2007), the gravity model (Trudgill, 1974), and the social network model (Milroy & Milroy, 1985; Milroy, 2005). In the wave model, changes are typically started by a group of speakers in a particular locale at a given point in time, spreading from that locale outward in successive stages as waves in a pond do when a stone is thrown into it. The gravity model suggests that the population density is also important in that changes often skip a smaller sparsely populated area even if it is closer to the origin and, instead, spread first to larger densely populated urban areas. The social network model looks not just at the population density, but at the nature of the relationship and contact, that is, the social network, that people of a certain area have established as the explanation for language change and diffusion. A general principle of social network analysis on language change is that networks constituted chiefly of strong ties support localized linguistic norms, while those constituted of weak ties are more susceptible to language change.

The third process we consider relevant to our study is drift, the inherent tendency of language to change constantly (Sapir, 1921). Sapir believed that “language moves down time in a current of its own making. It has a drift … if each language continued as a firm, self-contained unity, it would still be constantly moving away from any assignable norm, developing new features unceasingly …” (Sapir, 1921, p. 150). In addition, he further revealed the phenomenon of “parallelisms” in drift, where “the momentum of the more fundamental, the pre-dialectic, drift is often such that languages long disconnected will pass through the same or strikingly similar phases” (Sapir, 1921, p. 172). If we say that Sapir’s “parallelisms in
”drift” explains the similar changes in daughter languages of a protolanguage, such as the shared noun plural in English and German, then Trudgill (2000, 2004) further extended Sapir’s notion of “drift” to the level of dialects to account for parallel independent innovations in related varieties of one language. Trudgill (2004) argued that some similarities between different geographically separated varieties of English may be due “not to their having derived from similar dialect mixtures, nor to characteristics inherited directly from some parent variety, nor to any diffusion or direct contact between them, but to ‘drift’” (Trudgill, 2004, p. 132). He categorized two types of drift in his analysis on the Southern Hemisphere Englishes: 1) “Linguistic changes that are already in progress in the common source continue after separation”. 2) “No linguistic change was underway in the common source, but varieties derived from it inherit shared tendencies or propensities which can subsequently lead to the development of similar but new changes”. This current study will pursue Trudgill’s line of interpretation of drift.

Alongside the interpretation of drift, we will also consider the relationship between drift and language/dialect contact. Sapir’s (1933) pronouncement, as cited by Malkiel (1981, p. 547), clearly indicated three symptoms of a change attributable to drift: its inherent character as distinct from change due to language contact; its slow and gradual process; and its fundamentally unconscious character. Drift then is unambiguously distinguishable from diffusion. But what is the relationship between them? Is drift completely incompatible with diffusion through language/dialect contact? While we have seen two reports in the literature whereby drift does not necessarily exclude language contact (Keiser, 2009; Sneddon, 1993), would this relationship be generally applicable to other cases? We hope this research will contribute to the understanding of these issues.

3. Methodology

3.1 The community

This study focuses on the Korean Chinese community in Shenyang, the capital city of Liaoning province, the third largest Korean Chinese population in China. According to China’s sixth national census (2010), Shenyang is home to about 94,000 Korean Chinese, comprising 39.2% of the total Korean Chinese population in Liaoning province. This community is well known for its successful maintenance of Korean ethnicity in every aspect of its social infrastructure. For instance, it operates a comprehensive ethnic Korean educational system (from kindergarten to higher education), circulates an ethnic Korean newspaper, runs ethnic business centers, and practices religious belief in ethnic Korean churches. With the
Parallel Voice Onset Time shift in Chinese Korean

implementation of China’s “opening up” policy in the 1980’s and the establishment of diplomatic relations between China and South Korea in 1992, the formerly closed social network of the Shenyang Korean community has gradually evolved into a more open one through the labor mobility and the increasingly frequent commercial and cultural exchanges between the two countries.

The vast majority of the residents in Shenyang’s Korean Chinese community are descendants of Koreans who migrated from the Korean peninsula between the mid-19th century and 1945 (Gao, 1989; Piao, 1995). The first large-scale immigration from the Korean peninsula to China started in 1881, after the Qing government abolished the policy that prohibited cultivation of the vast land in northeast China. Once started, the process continued until the end of the Second World War in 1945 (Jin, 1996). This great migration of Koreans into China has been analyzed into three periods, each with a different main reason. The first period (free migration) ran from the 1840s until 1910; the main reasons for these movements include natural disasters in Korea and the abolishment of the no-cultivation policy of the Qing government. The second period (exile migration) was from 1910 to 1931; it was a response to Japan’s colonization of Korea. The third period (mandatory migration) ran from 1931 to 1945, when the Japanese government implemented a mandatory systematic move of Koreans to China (Che, 1998; Zhao, 2004; Zheng, 2000). Table 2 shows the number of Korean people who migrated from the Korean peninsula to northeast China during this 100-year period.

<table>
<thead>
<tr>
<th>Period Name</th>
<th>Time</th>
<th>Number of Korean migrants to China</th>
</tr>
</thead>
<tbody>
<tr>
<td>Free migration</td>
<td>1840s–1880</td>
<td>20,000</td>
</tr>
<tr>
<td></td>
<td>1881–1894</td>
<td>45,000</td>
</tr>
<tr>
<td></td>
<td>1895–1909</td>
<td>145,000</td>
</tr>
<tr>
<td>Exile migration</td>
<td>1910–1931</td>
<td>296,359</td>
</tr>
<tr>
<td>Mandatory migration</td>
<td>1932–1945</td>
<td>1,069,018</td>
</tr>
<tr>
<td>Approximate Total</td>
<td></td>
<td>1,575,377</td>
</tr>
</tbody>
</table>

(Data source: Jin, 1996; Xu, 1985)

With Korea regaining independence after Japan’s defeat in 1945, about 600,000 Koreans who had immigrated to China returned to Korea. By the time that the People’s Republic of China was established in 1949, the Korean population in the northeast area in China was 1,110,657 (Jin, 1996). This population constituted the Chinese Korean communities that were almost completely isolated from Korea for over 40 years (1949 to 1992), due to political reasons (Yang & Wang, 2006).
Since 1992, the year in which China and (South) Korea established diplomatic relations, the two countries have advanced their political, economic, diplomatic, and cultural relations with unprecedented speed and scope. Just ten years later, by the end of 2002, trade between the two countries had increased approximately eightfold, reaching about 44 billion dollars; by the end of 2004, this trade figure had increased twelve times over. China became Korea’s largest trading partner and Korea was China’s third largest. Active official and non-official contacts between the two countries prevailed in all areas, and every year the number of personnel exchanges exceeded 2.5 million (Yang & Wang, 2006). This complex history of cross-border migration, political separation, and socio-economic reconnection between China and Korea over the past century and a half provides a fertile field for the study of language change and variation.

3.2 The data

The data collection for this study was conducted in the summer of 2006 in the urban area of Shenyang city. Thirty five native Korean speakers, aged between 18–74 (Table 3), were first recruited via the first author’s social network and then via the networks of the interviewed speakers. To avoid any potential variation due to dialect differences, only speakers of the Phyengan variety of Chinese Korean (the variety spoken by the first author) were asked to participate. All the subjects were born in China and self-reported as bilingual in Korean and Chinese.

Table 3. Gender and age distribution of the subjects

<table>
<thead>
<tr>
<th></th>
<th>1930s</th>
<th>1940s</th>
<th>1950s</th>
<th>1960s</th>
<th>1970s</th>
<th>1980s</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>15</td>
</tr>
<tr>
<td>Female</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>8</td>
<td>20</td>
</tr>
<tr>
<td>Total</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>3</td>
<td>5</td>
<td>11</td>
<td>35</td>
</tr>
</tbody>
</table>

Note. Due to the limitations in the fieldwork, the number of female subjects born in 1980s is a bit larger than that of males, making the female participants a bit younger on average.

The interviews were recorded with a Dell laptop using Cool Edit 2000. An ATR20 Unidirectional Dynamic Microphone attached to the laptop was held at about 3 cm under the chin of the subject by a research assistant. Each recording session took about 5 minutes; participants were also asked to fill out a demographic questionnaire. The material for the speech task included nine isolated three-syllable words, where each of the nine target stop consonants were embedded in the word-initial position followed immediately by the vowel /a/ (Table 4). These words were printed in Korean script on a piece of paper and subjects were asked to read out the words at a normal speed. A total of 315 tokens were collected for this analysis.
3.3 Acoustic measurement

Acoustic measurement and transcription in this study were conducted with Praat software (Boersma & Weenink, 2007). VOT was measured from the first zero crossing of the stop release as indicated by the spike on the wave form after the closure period of the stop to the point where the periodic striation and the first formant (F1) of the immediately following vowel start in the spectrogram, which was set in wide band with View Ranges between 0–3000Hz and Window Length of 0.005s (Figure 1).

Figure 1. Illustration of VOT measurement for [p] of [pa]

F0 onset and offset of the vowel immediately following the target stop were also identified and measured with Praat. F0 onset was measured at the ending point of VOT as defined above, and F0 offset was measured at the point where the second formant of the vowel began to dissipate, that is, where the clear dark periodic striations of the second formant started to fade away in the spectrogram with the accompanying substantial reduction of the amplitude of the periodic spikes in the waveform. F0 measurements in Hz were normalized into semitones (St) using the formula log(Hz/100) × 12 with 100Hz as a reference pitch to allow for comparison
across age and gender (Oh, 2011; Whalen & Levitt, 1995). Statistical analyses of
the acoustic measurements were conducted with SPSS 13.0. Because place of
articulation of the consonants are not of interest for this current study, they are not
addressed here; rather, the analysis focuses on phonation type.

4. Results

Plotting the mean VOT values for each consonant type as a function of each sub-
ject’s year of birth suggests the existence of an apparent VOT shift over the past
60 years in the Chinese Korean speech community for at least two stop categories:
aspirated and lax (Figure 2). For aspirated stops, mean VOT values produced by
older speakers tend to be considerably higher than those produced by younger
speakers; in addition, the range of mean VOT values for younger speakers is also
smaller than that exhibited by their older counterparts. In contrast, mean VOT
values for lax stops produced by older speakers appear to be somewhat lower
than those produced by younger speakers, with a similar reduction in the range,
while mean VOT values of tense stops are consistent for all speakers. Applying a
simple best-fit line \( y = mx + b \) to the data for each stop category further illu-
minates both the apparent VOT shifts and the different patterns by which phonemic
distinctions are realized by means of differentiated VOTs. For example, one can
readily discern that for older speakers, the mean VOT value for the three phona-
tion types are distinct, with aspirated stops exhibiting higher VOT values than lax
stops, which in turn show higher VOT values than tense stops, a pattern well doc-
umented in the literature published during the 1990s (e.g., Cho, 1996; Han, 1996;
Silva, 1992). For younger speakers, however, the difference in mean VOT values
for aspirated and lax stops (i.e., mean\(_{VOT}_{asp} - mean\(_{VOT}_{lax}\)) is notably smaller:
indeed, for subjects born during the last quarter of the 20th century, this difference
is reduced to near zero. For these younger speakers, it appears that mean VOT no
longer serves as a reliable means of differentiating between underlying aspirated
stops and their lax counterparts, a pattern initially documented in Silva, Choi, and
Kim (2004) and subsequently reported in other work (e.g., Kang, 2014; Kang &
Nagy, 2016; Park & Iverson, 2008; Silva, 2006). We now turn to a statistical analysis
of these apparent trends.

In pursuing a more detailed analysis of the Chinese Korean VOT presented
in Figure 2, we follow the lead of Silva, Choi, and Kim (2004) and Silva (2006)
by proposing 1970 as a boundary point demarcating two different groups within
the sample of the speech community under investigation: the older group with
subjects born before 1970 and the younger group with subjects born in and after
1970. Independent of the linguistic data, there is also an external social factor that
makes 1970 an important generational divide: people born in 1970 started to experience in their teenage years China’s reform and opening towards the outside world, which led to China’s economic boom and greater mobility within Chinese society. Analysis of variance (ANOVA) of the VOT values collected for this study, taking as independent variables the subject’s age group (older vs. younger) and phonation type of the target consonant (tense, lax, aspirated), reveals robust differences in behavior. For the older group, as shown in Figure 3, VOT values associated with each of the three phonation types are statistically significantly different at the $p < 0.05$ level [$F(2, 168) = 275.7, p < 0.0001$], suggesting the existence of three distinct phonetic categories, which uniquely correspond to the three phonation types. For the younger group, however, ANOVA results reveal that although mean VOT values of tense stops are significantly lower than those of both lax and aspirated stops ($p < 0.0001$), mean VOT value of lax stops was not significantly different from that of aspirated stops ($p = 0.54$), thereby yielding two distinct phonetic clusters: one of tense stops and the other consisting of lax and aspirated stops, suggesting a merger in VOT values between the latter two phonation categories. This finding is consistent with the VOT shift reported in Korean varieties spoken in Korea, the United States, and Canada (Kang, 2014; Kang & Nagy, 2016; Silva, 2006; Silva, Choi, & Kim, 2004; among others).

Setting aside the VOT data associated with the tense stops – which appear to be highly consistent in apparent time – let us consider in more detail the changing relationship in VOT for lax and aspirated stops. In an effort to quantify the observation that the mean VOT values associated with these two categories were approximating one another (or merging), we calculated the difference in mean VOT values between lax and aspirated stops, that is, “delta VOT”: $\Delta$VOT=Mean

![Figure 2. Mean VOT values for different phonation type (ms)
VOT\textsubscript{asp} – Mean VOT\textsubscript{lax} (Silva, 2006, p. 293). As we can observe in Figure 4, as subjects’ age moves towards the younger end of the age axis, ΔVOT becomes continuously lower until it stabilizes among the youngest speakers in the data. More specifically, the difference in mean VOT values between aspirated and lax stops produced by older speakers is as high as about 70ms, which continuously decreases to around 30ms for middle aged speakers, and drops to the lowest negative values in the younger subjects’ utterances. Will this trend continue with even younger people than the youngest subjects in this study? The answer is no. As indicated by the best fit quadratic line ($R^2 = 0.62$), values for ΔVOT among the youngest subjects in the study stabilize around 0ms. We did notice, however, two speakers around the age of 35 (circled ones on the chart, born in 1970 and 1972 respectively), did not confirm to the behavioral pattern of other speakers of their age, exhibiting themselves as “outliers” to the general tendency. If we disregard the two speakers, the $R^2$ of the best quadratic line could be raised substantially to 0.79 (a difference of 0.17). The demographic information collected of the subjects revealed that these two subjects were Korean language teachers in a Korean ethnic college. As language teachers, they tended to be more conservative by adhering to the traditional pattern rather than adopting the new emerging pattern exhibited by speakers born in and after 1970.

The small values for ΔVOT for the younger speakers, which attest to a VOT merger between the lax and aspirated categories, corroborate prior studies on the matter. As the literature attests, mean ΔVOT lower than 20ms is reflective of such a VOT merger. In Silva (2006), for example, the speech of the younger speakers who exhibited a lax-aspirated VOT merger presented mean ΔVOT values of under 20ms. In Wright’s (2007) report of VOT merger, the mean ΔVOTs were equal to or less than 19ms. Perhaps most crucially, a perception study by Kim, Beddor, and Horrocks (2002) finds that across the tokens of lax [t], as VOT increased by 18ms or 19ms, listeners’ perception of aspirated [\textipa{t\textsuperscript{h}}] increased by about 60%. Such a sudden increase strongly suggested that VOT values of just under 20ms sit at a critical perceptual threshold between the two phonation categories,
further confirming the phonemic significance of the mean VOT values as reported in the current study.

Given this apparent shift from a three-way to a two-way contrast, we conclude that for those in the younger group of Chinese Korean speakers analyzed here, while relatively short VOT values continue to correlate reliably with the production of word- or phrase-initial tense stops, VOT no longer serves to differentiate lax stops from aspirated stops in the word- or phrase-initial context.

In further exploring the relationship between phonation categories and acoustic correlates in Chinese Korean, we consider the role played by the fundamental frequency (F0) of the vowel immediately following the target stop. Here we find that Chinese Korean presents a familiar pattern: As shown in Figure 5, for both groups of speakers (older and younger) the onset and offset values (semitones) of F0 for the vowel immediately following the target segments are lowest after a lax stop, at an intermediate value after a tense stop, and highest after an aspirated stop. These findings are consistent with those offered in multiple other similar studies of Seoul Korean (Ahn, 1999; Cho, 1996; Cho, Jun, & Ladefoged, 2002; Han, 1996; Han & Weitzman, 1970; Kim, Beddor, & Horrocks, 2002; Silva, 2006). One-Way ANOVA reveals some noteworthy patterns. For example, for all speakers, there is a statistically significant difference between mean F0 (onset was significant at the $p < 0.05$ level [$F(2, 310) = 23.1, p < 0.0001$] and offset was significant at the $p < 0.05$ level [$F(2, 310) = 10.9, p = 0.003$]) for vowels following a lax stop as compared to vowels following an aspirated or tense stop. When comparing F0 patterns for vowels following tense and aspirated stops, however, the patterns become more complex. For younger speakers, the mean value of F0 onset associated with the
tense and aspirated categories are statistically significantly different \( (p = 0.037) \), but no such difference obtains for F0 offset \( (p = 0.33) \). For all the three categories, the generalized F0 contour is falling: F0 onsets are consistently higher than offsets as shown in Figure 5, thereby yielding the percept of falling pitch – a percept impressionistically confirmed by the researchers (but one that merits more detailed investigation of the sort initially done by Kim, Beddor, & Horrocks, 2002). For older speakers, F0 mean values associated with preceding tense and aspirated stops present no statistically significant differences at the \( p < 0.05 \) level in both onset \( (p = 0.73) \) and offset \( (p = 0.72) \) contexts, suggesting that F0 is not a reliable correlate for differentiating between the two phonation types. In addition, the generalized F0 contour of the vowel following a lax stop is rising (a percept once again impressionistically confirmed by the researchers).

Out of these more complex patterns arises an important, consistent generalization: for speakers in both age groups, the F0 of a vowel following a lax stop is relatively lower than that of a vowel following either a tense or aspirated stop. Such generalizations about relative differences in F0 are by no means new; indeed, they can be readily found in the Korean linguistics literature as far back as the 1960s (Kim, 1965). When discussed, however, F0 patterns were always considered in a context whereby VOT differences were argued (or perhaps assumed) to represent the dominant acoustic correlate of phonation type. As differences in mean VOT between aspirated and lax stops are no longer consistently evidenced for younger speakers of Korean, however, the relative prominence of F0 appears to have risen. Thus, amidst the apparent shift in VOT associated with lax and aspirated stops, the underlying contrast between these two categories is not neutralized: rather, it

![Figure 5. F0 onset and offset for the vowel immediately following a stop of different phonation type for younger and older groups](image-url)
Parallel Voice Onset Time shift in Chinese Korean

is manifested by younger speakers in terms of differentiated F0 of the immediately following vowel—a tonal distinction.

In this so-called tonogenesis (Kim, 2000; Silva, 2006; Wright, 2007), tones manifested in F0 started out as redundant phonetic features of consonantal contrast and then developed into a robust distinction coexisting with the original consonantal contrast, and finally became the primary contrastive feature, as the original consonantal features were lost (Haudricourt, 1972; Hombert, 1978; Hombert, Ohala, & Ewan, 1979; Kang, 2014; Maran, 1973; Matisoff, 1973; Thurgood, 2002; Silva, 2006).

5. Discussion

Thus, we find in the Chinese Korean variety the same VOT shift and a change in the role of F0 as in the Korean varieties spoken in Korea, the United States, and Canada (Kang, 2014; Kang & Nagy, 2016; Silva, 2006). Questions then naturally follow: is this change in Chinese Korean an internal one or is it a diffusion from other varieties through recent dialect contact? What is the mechanism behind this observed parallel VOT shift in Chinese Korean?

We suggest that this VOT shift in the Chinese Korean community is an internal change, not a recent borrowing from other Korean varieties. We now return our attention to the ΔVOT data in Figure 4. As discussed above, the observed pattern of ΔVOT constitutes a curve with a negative slope of decreasing magnitude such that it appears to plateau around zero for the youngest speakers in the sample. In apparent time, VOT differences between aspirated and lax stops are reduced to a point of being neutralized. Having established an account for the right side of the curve in Figure 4, what are we to make of the left side? Given the data gathered for this study, it would seem that values for ΔVOT on the leftmost side of the distribution (representing subjects in their seventies, born in the 1930s) lie in a relatively broad range between 40ms and 70ms, but then immediately begin to drop as one moves rightward along the X-axis. Extrapolating leftward, one might be led to conclude that for Korean speakers born prior to 1930 would present values for ΔVOT increasingly in excess of 70ms. However, human physical capacity does not support this extrapolation, as the air in the vocal tract and the lungs that is compressed by the action of the respiratory muscles is limited in volume. Also, calculations of ΔVOT made on the basis of reexamining any available per-subject data found in the earliest foundational Korean VOT literature of the 1960s and early 70s reveal that the aggregated mean VOT difference between aspirated and lax stops is about 72ms (Silva, Choi, & Kim, 2004). In other words, ΔVOT appears to stabilize somewhere around the highest ΔVOT value in the current analysis. Thus, with a little bit of imagination, we can see the quadratic line of Figure 4 is actually
a component of a more comprehensive s-shaped or z-shaped curve, which is often believed to be indicative of a language change in apparent time (Labov, 1994; Guy, 2003). It begins with a clear three-way distinction in VOT produced by the older speakers, and then the decrease in ∆VOT is advanced further in the direction indicated by the inherited age vectors, with the younger speakers exhibiting the most robust pattern. The gradual incrementation of the VOT shift over the span of several decades as presented in Figure 4 suggests that the changes in VOT are not recent borrowings from Seoul Korean. Rather, it is an internal change led by younger speakers in the Chinese Korean community.

This finding is supported by several other observations. The geopolitics of the late 20th century support such a position: given that the Chinese Korean community was geographically and socio-politically separated from Korea for several decades, there were little means for frequent and significant interactions of the sort to establish a density of communication sufficient to effect language change via contact. It is true that with the recent changes in the foreign diplomatic relationship between the People’s Republic of China (PRC) and Korea and the policy changes within the PRC, the network structure of the Chinese Korean community has changed “from closed to open, from static to dynamic” (Kim, 2003, p. 116), creating more channels through which linguistic innovation and influence may flow from one group to another. Indeed, Jin’s (2012) study on this same group of subjects did find that the diphthong variant of the vowel /y/ is diffused from Korea into the Chinese Korean community through the early adopters of the middle-aged higher class females in the community who have more frequent access to Seoul Korean speakers than others. Jin reports that this group of subjects are often more status-conscious and tend to use more prestigious forms (which, they believe, is Seoul Korean) than men, often because their insecure and subordinate position makes it necessary for them to secure social status linguistically (e.g., Coates, 1998; Eckert, 1989; Holmes, 1992; Labov, 1982, 2001; Trudgill, 1972). They also exhibit higher frequency use of prestigious forms than females of other socioeconomic classes.

Thus, one might be tempted to believe that the VOT shift has been diffused into the Chinese Korean community through the recent dialect contact with Korea, which is not necessarily the case. We have seen in the Chinese Korean literature cases that clearly verify that even though one linguistic change has diffused from Korea into the Chinese Korean community (i.e., the case of /y/), it does not necessarily mean that other changes would follow the same route. For example, the front mid rounded vowel /ø/ has diphthongized into /we/ and the front mid unrounded vowel /ε/ has merged with /e/ in the standard Korean spoken in Korea (Ahn, 1987; Chae, 1995; Hong, 1991; Kang, 1997; Silva, 2008). In Chinese Korean, /ø/ was reported to have undergone the same diphthongization into /we/ (Shen
& Cui, 1985) when the Chinese Korean community was still isolated from Korea (the foreign diplomatic relationship between PRC and Korea was established in 1992) with possible channels of diffusion blocked. As for the two front unrounded vowels, a clear distinction between them, that is, no merger, is observed in the study of Chinese Korean (Silva & Jin, 2008) even though the channels between the two nations and the speech communities are open. This speaks to two things about Chinese Korean: an independent change in Chinese Korean similar to that of Seoul Korean is possible; a change in Seoul Korean might not be diffused into Chinese Korean even though the channels are available.

In addition, if it is the case that VOT shift in the Chinese community is a result of recent diffusion from Korea, it would be expected that the change would most likely be led by the same subgroup: middle-aged females from a higher socioeconomic class who led the /y/ change. Indeed, the literature suggests that subjects who lead in one change also lead in another change in the community. For example, Trudgill, Gordon, Lewis, and MacLagan (2000) reported a prime example of a speaker at the forefront of language change, Mrs. Catherine Dudley, whose speech is the most advanced case of New Zealand English with quite marked diphthong shifting; not surprisingly, she is also at the forefront of a second (and linguistically unrelated) change: the loss of postvocalic /r/. However, as we see from the results of this current study, the VOT shift in the Chinese Korean community is led not by middle-aged speakers, but by younger speakers. In the larger project of this research, we incorporated factors of age, gender, socioeconomic classes into various correlation analyses of the mean VOT data and found that only age was correlated ($r = 0.77, p < 0.0001$) with the apparent VOT shift (Jin, 2008). The Point biserial correlation test on gender and $\Delta$VOT showed that the correlation between the two variables was not significant ($r = 0.22, p = 0.21$). In addition, the Independent-Samples $t$ Test on the $\Delta$VOT values of male and female subjects confirmed that gender effect on $\Delta$VOT was not significant ($t = -1.28, p = 0.21$). We thus conclude that while mid-aged higher class females in the Chinese Korean community are early adopters of another variable (diphthongization of /y/), they do not lead in the case of VOT shift. It is possible, however, that middle-aged higher class female Chinese Korean speakers find changes in /y/ salient, while younger Chinese Korean speakers could find changes in VOT salient.

Then, are the younger speakers the early adopters of the VOT shift from Seoul Korean? The answer is no. Most of these younger subjects were college students from the Korean Ethnic Normal University in Shenyang. Following a highly regulated routine, they lived a life basically confined to the campus, taking classes most of the day, eating at the school cafeteria, studying each evening in a classroom, dorm or library, and then returning every night to a specified dorm in which they were required to turn off the lights at a set time. The ethnographic information
collected during the fieldwork confirmed that compared to others in the community (and other subjects in this study), these university students had less access and exposure to the outside society. Thus, the younger speakers in our sample do not have as much contact with Seoul Korean speakers as other subject groups and are arguably the least possible candidates to serve as early adopters of changes from Seoul Korean. Nevertheless, we observe that the apparent VOT shift is, in fact, led by these younger speakers. In addition, as we analyzed above, the gradual incrementation of the VOT shift over the decades as presented in Figure 4 suggests that the changes in VOT are not recent borrowings from Seoul Korean. Based on all the above facts and analysis, we thus conclude that the VOT shift in the Chinese Korean community is an internal change and not a recent diffusion from Seoul Korean through dialect contact.

However, is this Chinese Korean VOT internal change truly independent? We assert that it is not. In claiming so, we turn to a reexamination of published VOT data for Seoul Korean. In Kang’s (2014) study of VOT shift in Seoul Korean, it is reported that for female subjects born in 1930s, the mean VOT difference between aspirated and lax stops is lower than 10ms while for male subjects born in 1930s, the mean VOT difference is around 50ms (Kang, 2014, p. 82, Figure 3). These differences are much lower than the aggregated mean VOT difference of 72ms found in the foundational Korean VOT literature of the 1960s and early 70s (see Silva, Choi, & Kim, 2004). This means that when the subjects born in 1930s were in their teenage years, that is, during the 1940s, at least some Seoul Korean speakers were beginning to neutralize the VOT distinction that had long differentiated lax versus aspirated stops. Returning to the Chinese Korean data in Figure 4, we find the mean VOT difference for certain subjects born in the 1930s to be as small as 35ms, which is much shorter than 70ms (as recorded for some of the oldest speakers). These data suggest that when these speakers were in their teenage years, that is, in 1940s, the process of VOT neutralization was underway in China, too. Yet one must also recall that the large-scale Korean immigration to China that began after 1881 did not end until 1945; moreover, the almost complete isolation of China from Korea did not start until after 1949. On the basis of the converging evidence, we conclude that the process of VOT neutralization that was already underway during the 1940s in Seoul Korean was imperceptibly imported to the Chinese Korean community through migration, thus the change was initiated by dialect contact. In this sense, we say that the internal VOT change in the Chinese Korean community is not completely independent, as its “seed” was from Seoul Korean.

Once the impetus for the neutralization process was carried over to China, however, a long period of almost complete isolation between China and Korea followed from 1949 to 1992. During these about 40 years of isolation, there were no official or non-official channels for dialect contact between the Chinese Korean
community and Korea (Yang & Wang, 2006). In addition, the political, socio-cultural differences, and the extreme geographical distance between China, USA, and Canada ruled out the possibility of contact between the Chinese Korean community and the Dallas Korean community/Canadian Korean community that did have connections with Korea. In this sense, the Chinese Korean community of the late 20th century may be considered a de facto “speech island”, an instance of the geographic, social, and historical discontinuities that are prerequisites for parallel changes.

“Speech island” has been defined in different ways in the literature. Hutterer (1982), for example, defines speech island as “internally structured settlements of a linguistic minority on a limited geographical area in the midst of a linguistically different majority” (cited in Rosenberg, 2005, p. 221). Mattheier (1994) defines it as:

a speech community resulting from the delay or prevention of linguistic and cultural assimilation and that – as a speech minority cut off from its main territory – is surrounded and /or ‘roofed’ by a linguistically and ethnically different majority society, distinguishing or separating itself from the contact society via a socio-psychological attitude motivated by contrastive self-identity

(quoted in Keiser, 2009, p. 4)

In either interpretation, the Chinese Korean community presents itself as a valid case of a speech island. It is one of the settlements of the Korean minority on a limited geographical area in northeastern corner of China in the midst of the linguistically different Chinese majority. It is cut off from its main territory in Korea and “roofed” by the Chinese majority society. Korean Chinese have cherished a profound pride in their ethnic identity and maintained a strong sense of Korean ethnicity by using their own language, attending ethnic schools, observing their ethnic customs and traditions, etc. They are among the most advanced ethnic minorities in China in many areas including education, culture, and economy. They have established the Yanbian Autonomous Prefecture, Changbai Autonomous County, and hundreds of autonomous villages (xiang) (Cui, 2004). They have established a complete ethnic educational system from kindergarten to graduate school, and their illiteracy rate is the lowest and college attendance rate is the highest among the 56 ethnic groups in China (Choi, 2001). During the field trip for this study, all the Korean Chinese the first author encountered believed that because they are ethnic Korean, they should learn and use Korean, and none of their family members have married anyone from other ethnic groups. These practices have been a Korean tradition in China since ancient times, as a means of preserving the Korean language and culture, and the purity and continuity of this ethnic group. Given the strong ethnic Korean identity they uphold amidst the mainstream
Chinese society and their geographical, social, and historical separation from their homeland territory on the Korean peninsula, the Chinese Korean community constitutes a speech island of Korean in the surrounding of Chinese majority which makes it possible for the parallel independent VOT shift to develop.

We believe that behind the parallel VOT shift in the Chinese Korean community is the power of “drift” that drives Chinese Korean to evolve along a similar line of evolution with Seoul Korean. It meets the three symptoms suggested in Sapir’s (1933) later pronouncement of drift: it is an inherent change as distinct from changes due to contact with other Korean communities (i.e., not a matter of diffusion), it is a slow and gradual process encompassing several generations, and it is fundamentally unconscious. We find here an exact case of what Trudgill (2004) categorized as the “type 1” drift: “linguistic changes that are already in progress in the common source may be continued even after separation”. As analyzed above, the VOT neutralization was already in progress in the common source Korean language in the 1940s, and it was continued in the Chinese Korean community after over four decades of separation. Note here that the neutralization was not an already completed process before the separation that resulted in a ready feature that could be directly inherited by the Chinese Korean speakers from the source Seoul Korean variety through unbroken transmission. Rather the process was just underway and was imported to China in its initial stage, and then was further pushed along the strikingly similar phases as the Seoul Korean by the power of drift after the separation. We believe that drift is a process, a mechanism that drives related languages or varieties of a source language to go through similar phases of changes, not a metaphor for outcomes or a post hoc descriptive label of normal changes (Joseph, 1992; Keiser, 2009). Our argument is also supported by studies on several other languages such as Tibeto-Burman languages, northern and southern Wakashan, Hausa, Sulawesi languages, etc. which provided multiple evidences of drift as mechanism in diachronic changes (Fortescue, 2006; LaPolla, 1994; Newman, 1992; Sneddon, 1993).

However, as in a couple of cases reported in the literature (Fox, 2008; Sneddon, 1993) where a documented drift does not necessarily occur in all the varieties of a source language or there might even be a reversing of the drift, we see a similar counter case in the Korean spoken in Yanbian, the largest Korean concentration area in China. Zheng and Li (2005) reports that the VOT merger among younger speakers in Yanbian area is not between lax and aspirated stops, but between tense and lax stops, which is consistent with a traditional pattern reported in Kim (1965). A tentative explanation proposed here about the Yanbian case is that, being the largest Korean Prefecture in China, Yanbian has better preservations of the Korean language and culture than anywhere else in China, and the resulted more closed social network would make it less vulnerable to language changes. We hypothesize
that the Korean variety spoken there still preserves a very traditional VOT pattern where the overlap is between tense and lax stops (see Kim, 1965). If this is the case, then the VOT pattern observed in the young speakers in Yanbian would simply be an indication of language maintenance rather than language change. Why the observed Korean VOT drift appears to be geographically restricted in China to certain area is a question that may be answered with systematic studies on other Korean concentration areas in China and also on the Korean spoken in locations other than Seoul area in Korea.

The case of drift in Chinese Korean is unique in the sense that after decades of isolation of the speech community (1949–1992), there was the beginning of close contact with the source Seoul Korean. While drift is distinct from diffusion in its nature, it appears not necessarily totally incompatible with language/dialect contact. It may be possible that drift may even be further facilitated by language/dialect contact when changes from both processes are heading in the same direction. The increased, persistent contact between Seoul Korean speakers and Chinese Korean speakers through labor export, business, cultural exchange, etc. will continue to create channels for features of Seoul Korean to be diffused into Chinese Korean (and vice versa). When Seoul Korean speakers manifest the results of a VOT shift similar to that experienced by their Chinese Korean counterparts, it may reinforce and facilitate the VOT shift already progressing in Chinese Korean, providing thereby a support to the drift. Also, the fact that both Chinese and English have a two-way contrast for the stop consonants in their respective phonemic systems may also contribute to the VOT shift occurring in China and Canada. Table 5 presents the comparison of the mean VOT values produced by the subjects in this study with those of native VOT norms of Mandarin Chinese (Wu & Lin, 1989) and English (Lisker & Abramson, 1964).

Table 5. Comparison of mean VOT values of Mandarin, English, and Chinese Korean (CK)(ms)

<table>
<thead>
<tr>
<th></th>
<th>Short-Lag</th>
<th></th>
<th>Long-Lag</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mandarin</td>
<td>English</td>
<td>CK lax</td>
<td>CK aspirated</td>
</tr>
<tr>
<td></td>
<td>unaspirated</td>
<td>voiced</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>11</td>
<td>9</td>
<td>17</td>
<td>112</td>
</tr>
<tr>
<td>Mandarin voiced</td>
<td>112</td>
<td>69</td>
<td>65</td>
<td>87</td>
</tr>
</tbody>
</table>

In Table 5, we can see that the mean VOT value of Chinese Korean tense stops is close to the short-lag category of Mandarin and English, while those of the lax and aspirated stops match well with the long-lag category of the other two languages. We have seen reports in the bilingualism literature where L2-to-L1 influence occur between similar-sounding phones, manifesting phonetic shift of L1 in the direction of the L2 (Flege, 1987; Sancier & Fowler, 1997; Yao & Chang, in press). Given
the VOT similarity of CK stops to those of Mandarin and English, it is possible that the bilingual speakers in the Chinese Korean community and the Canadian Korean community may shift towards their ambient languages, that is, from a three-way VOT contrast to a two-way contrast. Thus, in practical linguistics situations, with increasing density of communication, drift and language/dialect contact may not be mutually exclusive. In our case then, both the process of internal change driven by drift and the process of external change induced by the language/dialect contact played a role in accounting for the VOT change in Chinese Korean.

A word is in order here about the VOT shifts in the Korean spoken in Dallas area in US (Silva, 2006) and Canadian Korean (Kang & Nagy, 2016). The data for Silva (2006) were elicited from 36 adult native speakers of Korean residing in the area of Dallas-Fort Worth, Texas, born between 1943 and 1982. However, all the subjects, as reported in the demographic questionnaire, were born, raised and educated in the capital region of Korea (i.e., Seoul city or Gyeonggi province). They spoke the standard Seoul variety and entered the United States after the age of 18. Thus, while they can be considered as representatives of Seoul Korean, we cannot claim them as representatives of American Korean in the way we claim of Chinese Korean. This would mean that the VOT shift reported in Dallas area is part of the VOT shift reported in the Korean spoken in Korea, not an independent one occurring in the United States. Kang and Nagy’s (2016) report of the VOT shift in the Korean spoken in Toronto, Canada, however, is from heritage Korean who were born in Toronto or came to Toronto before the age of seven, hence the legitimacy of “Canadian Korean”. Whether the VOT shift in the Canadian Korean is independent of the one in Korea, however, is currently not clear to us. Kang and Nagy (2016) mentioned the existence of continued close contact between the Canadian Korean community and Korea which might make diffusion possible, but they also admitted the possibility of VOT shift in the Canadian Korean community being an independent change. More detailed ethnographic and social network study about the Canadian Korean community will help illuminate the nature of the change.

6. Conclusion

This study reveals the existence of a VOT shift and a change in the role of F0 in the Chinese Korean variety which are similar to those reported in the Korean varieties spoken in Korea, the United States, and Canada. Subjects born in and after 1970 in the Chinese Korean community manifest a substantial overlap in mean VOT value and an enhanced F0 contrast between lax and aspirated stops categories. Thus, amidst the VOT shift, the underlying contrast between lax and aspirated stops is
maintained by younger speakers in terms of differentiated F0 of the immediately following vowel.

We believe that the observed VOT neutralization in the Chinese Korean community is not a recent diffusion from Seoul Korean or other Korean varieties spoken in USA and Canada through dialect contact, nor is it inherited as a ready feature directly from the source language through transmission. The “seed” of the VOT change is originally from Korea and then is further driven by the power of drift that leads the separated varieties of Korean to pass through strikingly similar phases. Drift is distinct in nature from diffusion through language/dialect contact, but diffusion nevertheless may be compatible with and support drift when changes from both processes head in the same direction. While isolation entailed in a “speech island” is a prerequisite for drift to occur (Keiser, 2009), the ideal clear-cut isolation in a speech island is not possible in reality, nor is it necessary in its absolute sense for drift to occur.

As Fortescue (2006) says, “the last word has not been spoken on the complex relationship between contact and drift”, but we do hope this current study has contributed to our understanding of that relationship. Future studies will need to systematically incorporate the Korean spoken in other Korean concentration areas in China and other countries such as Japan, Russia, Spain, etc. to better understand the mechanism, direction, and course of Korean language change as well as language development in general.

Acknowledgements

We would like to thank the editors and the two anonymous reviewers for their constructive comments and suggestions on the first draft of the paper. An earlier pilot study of the VOT shift in Chinese Korean was presented at the 2007 Annual Meeting of the Berkeley Linguistics Society (BLS) and included in the Proceedings of the BLS-33.

References


Cui, Qingzhi (2004). *Thinking under the global background: China's ethnic policy and ethnic Korean history, present and future.* Doctoral dissertation, Central Minzu University, Beijing, China.


Silva, David James (2011). Out of one, many: The emergence of world Korean(s). In James Hye-Suk Yoon, Yongsoon Kang, Jae-Young Han, & Chung-Kon Shi (Eds.), *Inquiries into Korean linguistics*, volume IV (pp. 9–37). Seoul: Taehaksa.


**Abstract (Chinese)**

中国朝鲜语中的VOT平行变化：一则语言学漂流案例

本项研究揭示了居住在中国东北部的朝鲜语本族人的语言中存在着VOT变化，而这一变化与韩国、美国、加拿大的韩语中所报道的变化是平行的。中国朝鲜语社区中观察到的这一VOT变化，不能简单地从近期的方言接触而传播的角度予以解释，也不是一个多世纪以前移植到中国时从源语言直接传承下来的一个特性。我们认为是“漂流”的力量支配这些平行的VOT变化，驱使不同的韩语方言沿着相似的语言轨迹演化。本项研究展示了一个有趣的案例，由“漂流”驱动的内部变化实际上可以由语言/方言接触引发并得到其进一步支持。

**Address for correspondence**

Wenhua Jin  
Department of Foreign Languages  
Kennesaw State University  
1000 Chastain Road, MD 1804  
Kennesaw, Georgia 30144  
wjin@kennesaw.edu