A sociophonetic approach to variation in Japanese pitch realizations

Region, age, gender and stylistic parameters

Shoji Takano and Ichiro Ota
Hokusei Gakuen University / Kagoshima University

This paper explores the flattening of Japanese sentential pitch as a possible nationwide change in progress among younger generations. Comparative data collected by identical protocols are examined in terms of speakers’ age, gender, and native dialects (in the city of Sapporo and in a rural town in Hokkaido, and in the city of Kagoshima in Kyushu). The paper also stresses the significance of including different registers in prosodic analysis and addressing potential problems with standard practices in which read-aloud materials comprise the primary resource. Based on naturalistic speech production data (i.e., spontaneous speech from a picture story description), our results reveal that: (1) regardless of the accentual discrepancies in their native dialects, younger generations characteristically speak in phonetically flattened realizations of pitch accompanying consistent, steeper declination, and (2) an age-linked differentiation also exists in prosodic phrasing, which is closely linked to the flattening of sentential pitch.

Keywords: sociophonetics, age-related variation, style, register, prosody, Japanese

1. Background

Based primarily on his observations of Tokyo Japanese, Sibata (1978) first suggested that pitch in the speech of younger generations of Japanese had increasingly been flattened. Years later, Sibata (1995, pp. 178–187) reconsidered this phenomenon in more detail, arguing that among younger speakers, the flattening (heiban-ka) of lexical accents “is explosively diffusing as an ingroup code, convey-
ing the images of being novel, youthful, and cosmopolitan” (pp. 182–183). Sibata also suggested that the flattening of sentence intonation could be widely observed in the speaking styles of the younger generation of contemporary Japanese society, which may be due to their “general preference for non-prominence” (p. 185). Sibata further conjectures that the flattening of “an overall pitch of speaking” initiated the flattening of lexical accents (pp. 185–186), which could be interpreted to mean that the flattening of sentential pitch has helped the flattening of lexical accents in order to act as a reinforcer. All of these claims, however, are based on Sibata’s native intuition or limited informal observations. A large-scale scrutiny of their validity is required.

Another intriguing facet of Sibata’s (1978, 1995) thesis is whether the flattening of sentential pitch is simply a localized phenomenon specific to Tokyo or a nationwide linguistic change in progress in the speech of younger generations throughout the country. Sibata restricts his observations to the speech of younger speakers of Tokyo Japanese, while suggesting the need to examine other local dialects for further confirmation of this potential change.

A review of past studies of dialect contact in Japan provides us with an indirect, but strong incentive to explore the flattening of sentential pitch contours as a change in progress in Japanese society today. Since the early 20th century, increasing mobility through continual societal urbanization has continuously and extensively exposed speakers of local dialects to different dialects, with Tokyo Japanese being the socially dominant variety (Sanada, 2000). For over half of the past century, studies of dialect contact in Japan have unanimously identified the loss or ambiguation of lexical accentuation through its flattening as a change in progress (e.g., Haga, 1961; Inoue, 1981; National Institute for Japanese Language and Linguistics [NINJAL hereinafter], 1965; Nomoto, 1960; Sibata, 1959). Based on a real-time paradigm, more recent studies (NINJAL, 1994; NINJAL, 1997 – based on 1986–88 fieldwork) have also demonstrated that while segmental sounds of local dialects tend to adopt Tokyo features more easily, prosodic aspects such as lexical accents continue to entail a great deal of instability, including the innovation of untraceable features as well as the loss of original patterns (NINJAL, 1997). Though lexical accents of subordinate varieties (i.e., regional dialects) became obsolete under the dominance of socially prestigious Tokyo Japanese, they did not fully merge with it. As a result, some in-between obscure patterns of lexical accents remain (cf. Heffernan, 2006).

1. Japanese is a pitch-accent language. Each word is phonologically specified as accented (yuukaku) or unaccented (mu-kaku), forming an accentual phrase. There is, however, a great deal of dialectal variation in the specification of lexical accents.
In view of these findings, our study aims to address the important, empirical question of how this could be extended to a larger prosodic domain of successive accentual phrases (AP hereinafter) in everyday language use as part of the nationwide change in progress.\(^2\) Our study is exploratory and challenging in three respects. Firstly, no prior study of Japanese speech sounds has taken into account the flattening of pitch beyond the word level.\(^3\) Secondly, our study sheds light on stylistic dimensions of variability as one of the critical issues in studies of prosody. While the majority of previous studies of variation and change in the prosodic aspects of Japanese are concerned exclusively with lexical accents in de-contextualized read words or sentences in isolation,\(^4\) we deal with prosodic variability in a more naturalistic speech register including read-aloud speech. Finally, our study is the first attempt to analyze this type of phenomenon from a sociophonetic perspective, focusing particularly on such social parameters as region, age, gender and style (cf. Cole & Thomas, 2005).

1.1 Pitch flattening in read-aloud register

Our prior investigation of read-aloud sentences in isolation (Takano & Ota, 2005a, 2005b, 2005c, 2006; Ota & Takano, 2008a, 2008b, 2014) have found that Japanese pitch contours are differentiated by age cohorts. These contours, particularly in the speech of younger generations, seem to have become similar

\(^2\) Following Beckman and Pierrehumbert (1986, p. 261), we define an Accentual Phrase as a prosodic unit marked by the presence of two delimitative tones (i.e., the phrasal H and boundary L tones). For example, the very first utterance in Figure 3 consists of two consecutive accentual phrases (APs): /senshuu no/ ‘of last week’ and /a’ru hi/ ‘one day’. The former AP can be identified by a phrasal H− tone as well as two boundary L tones: %L at its beginning and %L at its end. The latter AP, on the other hand, consists of its lexically specified pitch accent (A) linked to the first mora, which makes it impossible for a phrasal H− to activate, and the two boundary tones at its beginning (%L) and at its end (L%).

\(^3\) Other types of studies of the flattening of lexical pitch accents, for example, include Kubozono and Ota (1998), which focuses on correlations between the flattening of lexical accents and the types of lexicon (e.g., loanwords), and Mogami (1994), which describes a general tendency to flatten lexical pitch accents based primarily on usage of loanwords in Japanese media.

\(^4\) Although we are unable to refer to them in detail here due to limited space, recent rigorous studies of prosody with a strong emphasis on the significance of spontaneous data have been conducted. Those of lexical accents, for example, include Koiso, Yoneyama, Maki, and Fon (2003), Ito and Speer (2006), and Maekawa and Igarashi (2006) while those of intonational variability include Ayers (1994), Venditti and Swerts (1996), and Sugitoo (1997). Studies of prosody are also extended to other branches of linguistic investigation such as Conversation Analysis (e.g., Schegloff, 1998; Tanaka, 2004).
phonetically on a nationwide scale despite large discrepancies in the accentual systems of different dialects and the lack of speakers’ direct mutual contact.

Based on read-aloud sentences elicited from two age groups (younger – 14 to 23 years; older – 51 to 77 years) of speakers living in Hokkaido (the city of Sapporo and the rural town of Shin-Hidaka) and in the city of Kagoshima (Kyushu) located at opposite ends of Japan, we demonstrated the presence of age-related differentiation as in Figure 1. The dialects spoken in Hokkaido and Kagoshima differ typologically from each other in accentual systems (Hirayama, 1960) and can also be assumed to lack direct mutual contact.

Figure 1 demonstrates that the two younger groups (Shin-Hidaka Young and Kagoshima Young) show almost identical patterns of pitch movement, as do the two older groups (Sapporo Old and Shin-Hidaka Old), which creates two homogeneous age groups. The younger group shows relatively flattened realizations of pitch as compared with the older group, which shows greater degrees of rise and fall in pitch contours. Of particular significance is the fact that the younger participants from Hokkaido and Kagoshima speak in almost identical pitch contours despite the infrequency of their mutual contact and large differences in their accentual systems. Note that the pattern of younger speakers from Sapporo (Sapporo Young) is rather similar to that of the older groups, but it is still consistent with smaller indices of rise and fall than those of the older groups. We attributed this particular finding to gender-linked differentiation. Regardless of their region/age cohorts, women consistently manipulated greater degrees of shifts in pitch than men. This gender-linked difference has also been found repeatedly in our study of more spontaneous style of speech, and is discussed in detail in Section 2. These observations may suggest that Sapporo younger speakers’ deviations from the other

5. Hokkaido dialects, are generally divided into two subgroups namely, inland and coastal (Ono, 1981, 1993). The city of Sapporo (inland) is the cultural center of Hokkaido comprising a population of over 1.85 million people. Shin-Hidaka (coastal) is a rural town with a population of about 20,000. Hokkaido dialects in general have undergone a massive degree of standardization (Inoue, 1981; NINJAL, 1965; Ono, 1981) and are now regarded as a variety closer to Tokyo Japanese (i.e., standard Japanese). In contrast, Kagoshima dialect exhibits a striking auditory contrast to Tokyo Japanese because of the sharp discrepancy of high tone and low tone alignment specified in lexical accent. For example, while ame ‘rain’ is realized as A.me (HL) in Tokyo Japanese, a.ME (LH) is the tone of Kagoshima Japanese (Kubozono, 2012, p. 1398). This tonal (or accentual) feature is still well preserved after having lost almost all of other dialectal features in segmental phonology and grammar in the trend of language change toward the standard variety.

6. Our data did not include the older generation of Kagoshima-dialect speakers. In the course of our fieldwork, we decided not to consider older Kagoshima speakers for our region/age comparisons because even in the read-aloud task their sentential pitch contours were far too distinct to be comparable with the other groups.
A sociophonetic approach to variation in Japanese pitch realizations

The two younger groups are due to the skewed composition of gender in the data (i.e., 15 females and 1 male).

The age-linked differentiation in Figure 1 has also been confirmed statistically. The two Hokkaido age groups (Sapporo Young, Shin-Hidaka Young vs. Sapporo Old, Shin-Hidaka Old) differ to a statistically significant extent ($p < .05$) with respect to all the points of measurement of pitch shifts. Among the three younger groups, no statistically significant differences were found between the Shin-Hidaka and Kagoshima groups in any of the indices. The Sapporo younger group, however, significantly differed from both Shin-Hidaka and Kagoshima in the shift from 1AP-top to 1AP-bottom and the shift from 1AP-bottom to 2AP-bottom, and

Figure 1. Pitch shifts (based on z-scores) according to region and age in read-aloud sentences in isolation (based on Ota & Takano, 2014)*

1AP-top = the peak of the 1st accentual phrase, 1AP-bottom = the terminal point of the 1st accentual phrase, 2AP-top = the peak of the 2nd accentual phrase, 2AP-bottom = the terminal point of the 2nd accentual phrase, 3AP-top = the peak of the 3rd accentual phrase

* See Ota and Takano (2014, pp. 190–194) for the details of analytical procedures and results including the measurement of pitch contours, statistical information, the number of subjects involved as well as the number of tokens analyzed.
from Shin-Hidaka alone in the shift from 1AP-bottom to 2AP-top, which may be due to the skewedness in gender composition as we suggested above.

Read-aloud speech also shows that declination of pitch (i.e., the shift from 1AP-bottom to 2AP-bottom in Figure 1) differs to a statistically significant extent between the two age groups in Hokkaido, that is, the younger speakers show relatively steeper declination compared to the older speakers (see Table 5 in Ota & Takano, 2014, p. 192 for the indices for all of the measuring points mentioned so far).

Similar studies of other regional dialects such as Fukuoka dialect (northern Kyushu) and Tokyo Japanese in addition to Hokkaido and Kagoshima dialects as in Ota, Takano, Nikaido, Utsugi, and Asahi (2012) have also validated the age-linked differentiation in pitch contours. As compared with older counterparts,7 greater degrees of what we call prosodic subordination8 was consistently found in the speech of younger speakers, regardless of the speech production tasks involved (such as isolated sentence reading, performing roles in scripted conversations etc.). These results confirmed increasingly flattened realizations of pitch as nationwide phenomena prevalent among the younger generations.

The above findings also find confirmation in our earlier study of native speakers’ perception of the flattened pitch realizations (see Ota & Takano, 2007). The task included two pairs of read aloud sentences – one in its original flattened pitch and the other, a synthesized version, with greater degrees of pitch shifts both produced by the same speakers. A total of 156 college students (Sapporo 77, Kagoshima 79) who participated in the study judged the speakers with the synthesized version involving more dynamic pitch contours to be older than the identical speakers with their original flattened contours. The results suggest that the flattened pitch was associated by the listeners with “youthfulness” of the speaker, confirming Sibata’s insights into the social meaning of flattened pitch.

1.2 Issues of style and register in studies of prosodic variation

Generally, descriptions of prosody are stylistically limited in past research on prosody. A majority of studies have exclusively focused on a single register – a sentence typically elicited out of its context of use, as the sole resource for analysis. Ladd (1996, pp. 197–199) refers to this kind of research orientation as a “potential problem” in his chapter on prosodic prominence. A speaker’s manipulation of sentence intonation may vary tremendously based on his/her individual interpretations of

---

7. These older groups included only the speakers from Fukuoka and Sapporo.

8. The degree of prosodic subordination is based on the depth of a pitch valley between the first and second APs. The two show an inverse relationship where the shallowness of the pitch valley reflects greater degree of prosodic subordination implying greater flattening of pitch contours.
A sociophonetic approach to variation in Japanese pitch realizations

the unspecified background context of the sentence. Moreover, neutrality (and thus the baseline for all types of comparisons) in prosodic patterns is also quite a dubious notion because of the difficulty of identifying a single neutral intonation pattern for every sentence. Such an analysis of decontextualized speech represents a mainstream practice in both theoretical and sociolinguistic research (e.g., the NINJAL’s series of dialect studies) on Japanese prosody, in that carefully selected words and phrases or artificially constructed sentences out of context are read by native speakers of the language or regional dialects in question.

Several researchers have critically pointed out that prosodic descriptions of prior research may not exactly represent the reality of what native speakers actually do prosodically in multiple layers of styles and registers available in their everyday language use. For example, Umeda (1982), who analyzes pitch declination across registers, portrays systematic variability linked to different registers (i.e., reading of unrelated isolated sentences, reading of an essay, and natural conversation). In contrast to a general catathesis rule proposed in the phonological model of prosody (Beckman & Pierrehumbert, 1986; Pierrehumbert, 1980), Umeda (1982) argues that pitch declination in English is not a universal phenomenon but is rather strongly contextual. A steady declination can be observed (but not consistently) only in the reading of isolated sentences, whereas the other registers hardly display declination patterns of pitch.

Cohen, Collier, and ‘T Hart’s (1982) review of pitch declination studies find a similar claim. The authors report that the phenomenon involves stylistic differentiation – declination is relatively steeper in statements than in continual discourse. Moreover, the rates of declination, which they found to generally increase with shorter utterances, tend to become less consistent in spontaneous utterances as compared with read-aloud materials.

In the Japanese context, Sugitoo (1983), who analyzes regional variation in lexical accents from a passage-reading task in a laboratory setting, points out that the researcher’s overt instructions to speakers engaged in the task seem to heavily affect how words are accented. An instruction such as, “read as if you were a television announcer” significantly contributes to the speakers’ production of standard Japanese accents as compared with the reading of an identical passage in their native dialects.

Koori’s (2007) more recent study of spontaneous conversational data in Tokyo Japanese suggests that the weakening of lexical accents due to phonological downstep is specific to read-aloud speech (cf. Kubozono, 1989) and argues that accentual weakening should be reinterpreted as a marked phenomenon (rather than unmarked as previously claimed) in natural, spontaneous Japanese speech production. Furthermore, commenting on prior lexical accent research with regard to the widespread standardization process in Japan, Sanada, Shibuya, Jinnouchi, and Sugito (1999, p. 101) caution that all of the research outcomes are based
exclusively on a single register of isolated word reading, and that the question of whether the speakers actually speak the same way in their everyday lives is thus still unanswered. They further emphasize that investigation on how younger generations speak in more informal registers is vital.

These observations suggest that, in general, native-speaker participants are readily capable of manipulating systematically varied patterns of prosody across multiple registers and styles (Takano, 2008; Yaeger-Dror, 1996, 1997, 2002a, 2002b). It then follows that a framework of analysis oriented exclusively toward a single register or style could fall into quite limited, and possibly misleading, generalizations. We contend that research on prosody, which is highly sensitive to the context of use, should accommodate an appropriate perspective on multifaceted properties of prosody in its analytical framework.

2. The study

In the present study, we analyze pitch variability in a more naturalistic mode of speech. We asked the participants to describe a story spontaneously based on a series of six pictures (i.e., a picture storytelling task, with all participants viewing the same pictures) in order to approximate prosodic variability in naturally occurring speech production as closely as possible. Our results were then compared to our outcome (Ota & Takano, 2014) from the much less spontaneous register, namely (decontextualized) read-aloud sentences as reviewed in Section 1.

2.1 Data for analysis

Speech production data were collected from the same three sites as in our previous studies of read-aloud sentences (see Figure 1). Table 1 shows the number and demographic characteristics of the participants and the five age/region groupings.

Each subject engaged in the task of spontaneously constructing a story depicted by a series of six cartoons. An expected story begins with a given phrase, “One day during the last week,” and describes six sequential scenes using the past tense in which a woman goes to a department store, buys a tie, stops by on other floors to buy shoes and clothes, and ends up going home holding a number of shopping bags.

2.2 Analytical procedures

We followed the same coding procedures as in Ota and Takano (2014, p. 192) for read-aloud sentences in isolation (see Figure 2), except that the participants were
A sociophonetic approach to variation in Japanese pitch realizations

given no fixed text to read aloud this time and were expected to produce short narratives describing the story spontaneously. We measured five dimensions of pitch shifts (Figure 2): (1) the shift from the first AP top to its bottom (Shift 1), (2) the shift from the first AP bottom to the second AP top (Shift 2), (3) the shift from the second AP top to its bottom (Shift 3), (4) the shift from the second AP bottom to the third AP top (Shift 4), and (5) the shift from the first AP bottom to the second AP bottom as a declination (Shift 5).

To identify and segment a single intonation unit in a successive flow of speech, we adopted the Japanese ToBI system for prosodic transcriptions (Venditti, 1995, 2005). We determined the beginning of a new intonational unit based on Break Index 3 for the marker of the intonation phrase (IP hereinafter) which is the largest disjuncture typically accompanied by pitch reset, the boundary pitch movement, pause, and final syllable lengthening, according to the X-JToBI coding scheme (Igarashi, Kikuchi, & Maekawa, 2006). Using PitchWorks, we coded a variety of information such as the f0 value of the pitch peak (A)9 or that of the AP-initial high phrase tone (H−),10 the value of the AP-final lowest f0, pitch range, speaker ID, and so on, in the information tiers.

Figure 3, an example of our coded materials, contains two IPs (marked by 3 in BI Tier): the first is composed of an unaccented AP /senshuu no/ ‘of last week’ and an accented AP /a’ru hi/ ‘one day’, and the second is composed of a sequence of three APs: an accented /depa’ato e/ ‘to a department store,’ an unaccented /kai-mono/ ‘shopping,’ and an accented AP /ikima’shita/ ‘went.’

### Table 1. Participants

<table>
<thead>
<tr>
<th>Regions</th>
<th>Age group</th>
<th>Number</th>
<th>Age range</th>
<th>Gender</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shin-Hidaka (Hokkaido)</td>
<td>Older</td>
<td>10</td>
<td>57 to 77</td>
<td>5 males, 5 females</td>
</tr>
<tr>
<td></td>
<td>Younger</td>
<td>10</td>
<td>17 to 29</td>
<td>5 males, 5 females</td>
</tr>
<tr>
<td>Sapporo (Hokkaido)</td>
<td>Older</td>
<td>9</td>
<td>55 to 73</td>
<td>3 males, 6 females</td>
</tr>
<tr>
<td></td>
<td>Younger</td>
<td>8</td>
<td>20 to 23</td>
<td>5 males, 3 females</td>
</tr>
<tr>
<td>Kagoshima (Kyushu)</td>
<td>Younger</td>
<td>10</td>
<td>19 to 20</td>
<td>5 males, 5 females</td>
</tr>
</tbody>
</table>

Note. Our data did not include the older generation of Kagoshima-dialect speakers. In the course of our fieldwork, we decided not to consider older Kagoshima speakers for our region/age comparisons because even in the read-aloud task their sentential pitch contours were far too distinct to be comparable with the other groups.

---

10. In X-JToBI, the phrasal tone H− indicates the rising movement of pitch at the beginning of an AP, and it is usually (but not always) associated with the second mora from the beginning of the AP.
For comparative purposes, we then grouped the segmented IP tokens in terms of combinatory patterns of APs in sequence (Table 2) and conducted a comparative analysis of the tokens that belong to the same combinatory pattern (e.g., all of the IPs with the sequence “accented AP + accented AP” across the subjects, regions, and age groups). IPs composed of a single AP were discarded.

Our analysis was focused on the variable movement of pitch during the trajectory of the IP-initial. Two APs of all the tokens were collected due to the scarcity of tokens that could cover all of the combinatory patterns. A total of 521 tokens varying in terms of length (i.e., number of APs in sequence), were divided into four types of AP combinations at the IP-initial position: 161 sequences of “accented AP + accented AP” (30.9%), 108 sequences of “accented AP + unaccented AP” (20.7%), 205 sequences of “unaccented AP + accented AP” (39.4%), and 47 sequences of “unaccented AP + unaccented AP” (9.0%). For statistical analyses, we also took into account the length of tokens, that is, the number of APs that constituted a single IP by three levels: 2 APs (253 tokens), 3 APs (165 tokens), and 4 or more APs (103 tokens).
2.3 Systematic pitch variability in story-telling register

Parallel to Figure 1 on pitch shifts in read-aloud sentences in isolation, Figures 5, 6, 7 and Figure 8 in this section are schematic representations of pitch shifts (based on $z$-scores) in the IP-initial sequence of two APs for each of the four combinatory patterns.
produced in the story-telling task, respectively. Specific values of pitch shifts (i.e., Shifts 1, 2, 3 and 5),\textsuperscript{11} which entered statistical analyses are available in Appendix.\textsuperscript{12}

\textbf{2.3.1 Variability involving accented APs}

Overall comparisons of the average values between the two age groups indicate that the age-linked pattern of pitch movement we have found with read-aloud sentences as in Figure 1 is maintained almost perfectly in story telling task except for Shift 3 in “unaccented + accented” APs (average younger −1.43, average older −1.31 as in columns VII and VIII in Appendix).

Figure 4, which compares register-specific pitch trajectories between the two age groups, shows that the older group in story-telling task consistently demonstrates a more dynamic movement of rise (1AP-top, 2AP-top) and falls (1AP-bottom, 2AP-bottom) though the differences from the younger group are very small.

More detailed accounts of each combinatory pattern as in Figures 5, 6 and 7, however, lead us to rather complicated outcomes. The combinatory pattern of “accented + accented” APs (Figure 5), which was the focus of our investigation of read-aloud sentences in isolation earlier (Ota & Takano, 2014), does not yield statistically significant degrees of differences between the age groups in this spontaneous style of speech. In fact, none of the pitch shifts involving accented APs (Figures 5–7) demonstrate age-linked differentiation to a statistically significant extent, except for Shift 3 (2nd AP top to its bottom) in the “unaccented + accented” combination with respect to Hokkaido subjects as in Figure 7.

A robust finding with respect to accented APs is gender-linked differentiation: women in general hold a more dynamic rise and fall in pitch along with a steeper declination than men regardless of their age and region. This gender-specific pattern consistently applies to all of the measuring points in the combinatory patterns involving accented APs (as well as a few of the patterns involving unaccented APs

\textsuperscript{11} Note that Shift 4 (Shift 4 in Figure 2, a pitch shift from the second AP bottom to the third AP peak) is not analyzed in this register because we are concerned only with the sequence of two APs at the IP-initial position.

\textsuperscript{12} Column IX in Appendix indicates (with asterisks) the independent variables that are found to be statistically significant at $p < .05$. We used ANOVA (region x age x gender x IP length, $p < 0.05$) for two age groups from Hokkaido, and ANOVA (region x gender x IP length, $p < 0.05$) for three younger groups from Sapporo, Shin-Hidaka, and Kagoshima. In the same column, the values given with parenthesized (Hokkaido) and (Younger), etc. are concerned with each of the age/region groups, respectively. Age-linked factors that are found statistically significant are underlined. Note that Shift 4 (Shift 4 in Figure 2, a pitch shift from the second AP bottom to the third AP peak) is not analyzed in this register because we are concerned only with the sequence of two APs at the IP-initial position.
A sociophonetic approach to variation in Japanese pitch realizations

in Figures 9b, 9c and 9d) in accord with what was found with read-aloud sentences in isolation (Ota & Takano, 2014).13

Women’s dynamic pitch movement (as compared with men) has also been confirmed in previous studies of Japanese language use (e.g., Hiramoto & Wong, 2005; Ohara, 1992; Okamoto, 1995). Although these studies vary in terms of speech styles ranging from spontaneous role plays in experimental settings to natural conversations, they unanimously demonstrate that women’s uses of pitch involve more frequent occurrences of sharp pitch peaks in a characteristically wider pitch range as compared with the male counterparts. In addition to such findings, our results further demonstrate that it is accented APs that are likely to differentiate between women’s and men’s speech in Japanese prosody, contributing to the dynamic activation of pitch, particularly in regards to the former.

Other factors that are found statistically significant with respect to accented APs include the length of IPs and the region. As for the former, our results indicate

13. Shift 2 (1st AP bottom to 2nd AP top) in the “unaccented + accented AP” sequence (Figure 9c) does not yield statistically significant differences between the gender groups although the gender pattern is maintained: men 0.72, women 0.80 (Hokkaido); men 0.62, women 0.75 (Younger).
Figure 5. Pitch shifts in Accented AP + Accented AP in spontaneous speech
* Indicates the level of statistical significance at $p < .05$

Figure 6. Pitch shifts in Accented AP + Unaccented AP in spontaneous speech
* Indicates the level of statistical significance at $p < .05$. 
A sociophonetic approach to variation in Japanese pitch realizations

Figure 7. Pitch shifts in Unaccented AP + Accented AP in spontaneous speech
* Indicates the level of statistical significance at \( p < .05 \).

Figure 8. Pitch shifts in Unaccented AP + Unaccented AP in spontaneous speech
* Indicates the level of statistical significance at \( p < .05 \).
Figure 9a. Gender differences in Accented AP + Accented AP in spontaneous speech
* Indicates the level of statistical significance at $p < .05$.

Figure 9b. Gender differences in Accented AP + Unaccented AP in spontaneous speech
* Indicates the level of statistical significance at $p < .05$. 
Figure 9c. Gender differences in Unaccented AP + Accented AP in spontaneous speech
* Indicates the level of statistical significance at $p < .05$.

Figure 9d. Gender differences in Unaccented AP + Unaccented AP in spontaneous speech
* Indicates the level of statistical significance at $p < .05$. 
that the longer an IP is, the more dynamic the rises and falls in pitch are likely to become (Shifts 2 [Hokkaido, \(F(2, 122) = 5.573, p < .05\)] [Younger, \(F(2, 92) = 7.606, p < .05\)] and Shift 3 [Younger, \(F(2, 92) = 7.804, p < .01\)] in Figure 5; Shift 2 [Hokkaido, \(F(2, 160) = 3.650, p < .05\)] in Figure 7). This particular finding accords with the generalization by Cohen et al. (1982) that the rates of declination (i.e., the gradual weakening of pitch dynamics) tend to be larger in shorter utterances. Koori’s (2007) claim that the weakening of lexical accents in Tokyo Japanese is specific to the read-aloud register and is not generalizable to spontaneous speech (e.g., conversations) is also partly supported here.

The regional variability involving accented APs (Shifts 3 and 5 in Figure 5, Shift 3 in Figure 7) seems to reflect the dialectal distinctions among speakers in Sapporo [inland Hokkaido], Shin-Hidaka [coastal Hokkaido] and Kagoshima [Kyushu]). Furthermore, Shift 3 in Figure 7 contains a number of inexplicable results: the three independent variables (age, gender, region) interact to a statistically significant extent with respect to Hokkaido subjects (see column IX in Appendix). Further examination (using Bonferroni’s multiple comparisons on SPSS) indicates that (1) age makes a difference only among female subjects from Sapporo (older −1.48, younger −2.13), (2) gender makes a difference among younger subjects from Sapporo (men −1.41, women −2.13) and Shin-Hidaka (men −1.0, women −1.67), and (3) younger female speakers from Sapporo (−2.13) and Shin-Hidaka (−1.67) behave differently. The complexity of these results needs to be further qualified with more extensive data and a more elaborate examination which is not possible within the space of this paper.

2.3.2 Variability involving unaccented APs

Unlike IPs consisting of accented APs discussed so far, a relatively consistent age-linked differentiation in Japanese pitch activation can be found with respect to unaccented APs, which in fact have been understudied in prior research on connected speech in general. Shift 1 (Old −1.02, Young −0.68 [Hokkaido]: \(F[1, 163] = 20.662, p < .001\)) and Shift 5 (declination; Old −0.55, Young −0.80 [Hokkaido]: \(F[1, 160] = 6.284, p < .05\)) in Figure 7 (the “Unaccented + Accented” APs sequence) and Shift 1 (Old −1.20, Young −0.70 [Hokkaido]: \(F(1, 18) = 7.829, p < .05\)) in Figure 8 (the “Unaccented + Unaccented” AP sequence) exhibit statistically significant differences between the two age groups. The remaining shifts (Shifts 2, 3, 5) in the “unaccented + unaccented” AP sequence (Figure 8), though not statistically significant, are also consistent with the aforementioned age-linked differentiation: younger subjects’ flattened realizations of pitch along with steeper declination (Shift 2 – Young 0.51, Old 0.82; Shift 3 – Young −0.71, Old −0.96; Shift 5 – Young −0.21, Old −0.14 in columns VII and VIII in Appendix).
Also notice that Shift 1 in the “unaccented + accented” AP sequence (Figure 7) involves the interaction of the three independent variables (i.e., age, gender, region) with respect to Hokkaido subjects (see column IX in Appendix). Further examination (using Bonferroni’s multiple comparisons on SPSS) indicates that: (1) the speaker’s age makes a difference in all region/gender groups excepting Sapporo female subjects (i.e., there is no statistically significant difference between older [-0.83] and younger [-0.87] female subjects from Sapporo); (2) gender makes a difference only in older subjects from Shin-Hidaka (men −0.85, women −1.35); and (3) older female speakers from Sapporo (−0.83) and Shin-Hidaka (−1.35) behave differently.

Based on these results, it is very likely that as far as more spontaneous modes of speech are concerned, it is the pitch contours specific to unaccented APs (rather than accented APs) that tend to differentiate between the two age groups. The speech of older age groups is characteristic of the dynamic rise and fall of pitch particularly in unaccented APs (at least at IP-initial positions), whereas younger age groups involve relatively flattened realizations of pitch.

This particular finding is quite valuable and we would not have discovered it if we had only been concerned with pre-structured materials such as read-aloud sentences out of context. Our finding also suggests that stylistic considerations can serve as a crucial component of prosodic analysis. A unidimensional approach to prosodic variability (e.g., read speech materials) is likely to provide generalizations that do not necessarily reflect more naturalistic registers, not to mention the reality of everyday language use (cf. Koori, 2007, and also see Ayers, 1994 for American English).

2.4 Variability in the dephrasing of the accentual phrase

Another important discovery from our analysis of a spontaneous mode of speech is concerned with age-linked differences in AP dephrasing. We define AP dephrasing as a phonetic phenomenon in which a lexical accent unit, which presumably formulates an independent AP in read-aloud speech, does not manifest itself as an AP but rather merges with at least one of its neighboring APs.

---

14. This is equivalent to the prosodic unit identified by the Break Index 2 in X-JToBI (Igarashi et al., 2006).
Figure 10. Example of AP dephrasing by a 22-year-old female Sapporo subject in the picture storytelling task: \(/mata/ /huku’/ ga/ /hangaku de/ (/) uraretei’ru no o/ /mi’te/ ‘Also, (She) noticed that the clothes were being sold at half price.’

Figure 10 illustrates an example of AP dephrasing produced by a 22-year-old female subject from Sapporo. An inherently unaccented AP /hangaku de/ ‘at half price’, which would usually be realized as an independent AP in read-aloud materials (ToBI Break Index 2), intonationally does not constitute an independent AP (indicated by 2-) and merges with the immediately following accented AP /uraretei’ru no o/ ‘were being sold’ as a single AP.

Figure 11 describes the occurrence rates of AP dephrasing for the region/age groups, which are based on the relative frequency to the total number of IPs produced (i.e., the number of IPs involving at least one dephrasing occurrence relative to the total number of IPs produced). The raw frequencies of dephrased APs are Shin-Hidaka Older (6), Sapporo Older (9), Shin-Hidaka Younger (39), Sapporo Younger (25), and Kagoshima Younger (35).

---

15. We tried to be as conservative as possible in judging the occurrence of dephrasing. For example, we excluded the compound nouns (e.g., /yoohuku/ /u’ribal/ ‘clothing department’), which were variably realized as either two separate APs or a single AP, depending upon the speakers.
Although a great deal of individual variability can be observed especially within the younger groups (i.e., Shin-Hidaka vs. Sapporo and Kagoshima) in Figure 11, it is evident that the speech of younger subjects in general is more likely to involve AP dephrasing than the speech of their older counterparts. Statistical analysis demonstrates a significant degree of differences in the number of occurrences of AP dephrasing between the two age groups from Hokkaido ($F[1, 29] = 19.169$, $p < .001$), and also indicates that no internal discrepancies exist within each of the age groups.

Here again, the present phenomena involve regional variation to a statistically significant extent. Generally, the speakers in urban areas, regardless of their age, seem to be leading this potential change in progress. Sapporo subjects in general (19.9%) hold higher rates of AP dephrasing than Shin-Hidaka subjects (11.6%) ($F[1, 29] = 4.682$, $p < .05$). Kagoshima younger subjects (25.7%) appear to be almost as advanced as Sapporo younger subjects (28.4%).

Further investigation of greater AP dephrasing among younger speakers suggests a close link between dephrasing in various AP combinatory types (see Table 3) and the prevalent pitch flattening of unaccented APs. Table 3 lists seven combinatory patterns of AP dephrasing observed in the current database.

Types (1), (4), (6), and (7) are all resulted from the non-activation of lexical accents (H*+L) in inherently accented APs. In Type (1), for example, the lexical accent of an inherently accented AP /kutsu ‘mo/ ‘shoes also’ was not realized and that particular AP has become part of the immediately preceding accented AP
Table 3. AP combinatory types of AP dephrasing in the picture story-telling task

<table>
<thead>
<tr>
<th>Types</th>
<th>Example tokens</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) $H^<em>+L \leftarrow H^</em>+L$</td>
<td>josee wa /huku’ ya(//) kutsu mo/ kaimono o shi</td>
<td>The woman did shopping for clothes and shoes as well,’</td>
</tr>
<tr>
<td></td>
<td>woman top clothes or shoes also shopping obj do</td>
<td></td>
</tr>
<tr>
<td></td>
<td>‘The woman did shopping for clothes and shoes as well,’</td>
<td></td>
</tr>
<tr>
<td>(2) $H^*+L \leftarrow H^-$</td>
<td>/a’ru(/) josee ga/ depa’ato e kaimono ni ikima’shita</td>
<td>‘A woman went to a department store for shopping.’</td>
</tr>
<tr>
<td></td>
<td>a woman department store to shopping to went</td>
<td></td>
</tr>
<tr>
<td>(3) $H^- \rightarrow H^*+L$</td>
<td>/hujinyoo no(/) kutsu’ ga / yasuuri- sareteiru no’ o mitsukema’shita.</td>
<td>The woman noticed that women’s shoes were being sold at a bargain.’</td>
</tr>
<tr>
<td></td>
<td>for women gen shoes sub at a bargain- being-sold that obj notice-PAST</td>
<td></td>
</tr>
<tr>
<td>(4) $H^- \leftarrow H^*+L$</td>
<td>/senshuu no(/) aru hi / depa’arto e eikima’shita.</td>
<td>‘One day last week (a woman) went to a department store.’</td>
</tr>
<tr>
<td></td>
<td>last week gen one day department store to went</td>
<td></td>
</tr>
<tr>
<td>(5) $H^- : H^-$</td>
<td>sono a’to ni /yoohuku no(//) hangaku no/ se’eru ga mi’eta node,</td>
<td>‘After that, (the woman) saw a half-price sale for clothing, so…’</td>
</tr>
<tr>
<td></td>
<td>that after clothing gen half-price gen sale sub saw conj</td>
<td></td>
</tr>
</tbody>
</table>
Table 3. (continued)

<table>
<thead>
<tr>
<th>Types</th>
<th>Example tokens</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>(6) $H^<em>+L / H^</em>+L \rightarrow H^-$</td>
<td>takusa'n no /nimotsu o (/) motte/ depa'ato o a'to ni shima'shita.</td>
<td>‘(The woman) left the department store carrying a lot of baggage.’</td>
</tr>
<tr>
<td></td>
<td>a lot of GEN baggage OBJ carry-GER department store OBJ behind did</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(7) $H^*+L \rightarrow H^-$</td>
<td>sono /kaeri no(/) tochuu de /josee wa ...</td>
<td>‘(After) that, on her way home the woman ….’</td>
</tr>
<tr>
<td></td>
<td>that go home GEN on one’s way woman TOP</td>
<td></td>
</tr>
</tbody>
</table>

*Note. $H^*+L =$ accented AP; $H^-= unaccented$ AP; Arrows ($\rightarrow, \leftarrow$) = directions of AP merger; / xxx ( / ) xxx (in bold italics) = dephrased APs*
Shoji Takano and Ichiro Ota

/huku’ ya kutsu mo/ ‘clothes or shoes also.’ In a rather drastic case of Type (6), neither of the lexical accents of the inherently accented APs /ni’motsu o/ ‘baggage OBJ,’ /mo’tte/ ‘carry-GER’ was realized and those APs resulted in formulating a single unaccented AP /nimotsu o motte/ ‘carrying baggage’ as a whole.

Figures 12 and 13 describe the distribution of combinatory types for the younger and older age groups, respectively. While it seems difficult to make direct comparisons of younger and older patterns considering that the latter group has too few tokens to quantify, we could still identify relatively shared patterns of distribution among the three younger groups from different regions in Japan. The majority of dephrasing in the speech of the younger subjects tends to occur in the AP sequences involving unaccented APs – namely, Types 3 (H− → H*+L) and 5 (H− : H−). In Type 3, the first unaccented APs (e.g., /hujinyoo no/ ‘for women GEN’) do not constitute an independent AP and tend to merge with the immediately following accented APs (e.g., /hujinyoo no kutsu’ ga/ ‘women’s shoes SUB’). In Type 5, two adjacent unaccented APs behave like a single AP without any intonational break between them (e.g., /yoohuku no hangaku no/ ‘(sale) of half-price clothes’). In contrast, it seems more difficult to identify any shared patterns in the dephrasing of the two older subject groups. Interestingly, however, Sapporo older subjects are in the lead with respect to both these innovative types (i.e., 3 and 5) of AP dephrasing as in Figure 13.

Variable uses of AP dephrasing by younger subjects also involve regionality. The speakers in urban areas (Sapporo and Kagoshima) in general appear to be leading in the innovative types of AP dephrasing (i.e., 3 and 5 in Figure 12). Type (2), on the other hand, is much more advanced among Shin-Hidaka subjects alone, regardless of their age (Figures 12–13). We leave this dimension of the findings for future investigation.

The findings thus far seem to have much to do with our discovery of a consistent age-linked differentiation in pitch variability in the more spontaneous register referenced in Section 2.3 – that is, younger speakers’ less dynamic (i.e., flattened) movement of pitch contours and its particularly strong association with unaccented APs. In other words, it is very likely that younger speakers, in general, inactivate unaccented APs (or make them less prominent), and such dephrased APs are more likely to be integrated with adjacent APs without any intonational demarcations. Moreover, this innovation could be considered to be a change in progress in Japanese prosody, given the fact that the speakers in urban areas, regardless of their age, tend to be in the lead. A follow-up study based on more systematic sample in the future might throw more light on the progression of this potential change in progress.

Finally, Figures 12 and 13 also demonstrate that it is very unlikely that accented APs are the trigger for AP dephrasing in a spontaneous mode of speech,
A sociophonetic approach to variation in Japanese pitch realizations

contrary to a general catathesis rule in the phonological model (Beckman & Pierrehumbert, 1986; Pierrehumbert, 1980). The rule predicts that the weakening of pitch accentuation (both $H^*+L$ and $H^-$) tends to take place in an environment where accented APs precede other APs (cf. Pierrehumbert & Beckman, 1988). It is then expected that AP dephrasing should be observed more frequently in this environment (Types 1 and 2) than otherwise.

Our spontaneous speech data, however, show that this is hardly the case. In fact, Figures 12 and 13 demonstrate that across the two age cohorts, the AP sequences in which $H^*+L$ precedes (Types 1 and 2) seem to be a relatively minor

---

**Figure 12.** AP combinatory types of dephrasing produced by the three younger subject groups in the picture storytelling task

**Figure 13.** AP combinatory types of dephrasing produced by the two older subject groups in the picture storytelling task
prosodic environment that favors AP dephrasing. This finding supports Koori’s (2007) claim that the weakening of lexical accents due to phonological downstep should be reinterpreted as register-specific (e.g., read-aloud speech) and not necessarily generalizable to more spontaneous modes of register such as conversations (cf. Umeda, 1982). Based on the current results, we also contend that the outcome of analyses of a single style (or register) should not be blindly applied to other styles (or registers). Systematic accounts of multiple registers could be a valuable component of prosodic analysis.

3. Discussion and conclusion

While the robustness of age-linked variability is manifested differentially across distinct registers, relatively flattened realizations of sentential pitch accompanying steeper declination are largely identified in the speech of our younger subjects. The speech of older subjects, on the other hand, is characteristic of assured realizations of accentual phrases regardless of their accentual types (i.e., either accented or unaccented APs), leading to the dynamic movement of sentential pitch.

We initially suggested the possibility that sentential pitch flattening can be considered a nationwide change in progress based on past studies of loss (or ambiguity) of Japanese lexical accents as a recent general trend among younger generations and Sibata’s specific comments on the flattening of overall pitch contours in their speech. The results of the present study (and our matched-guise experiments of social meanings of flattened pitch contours in Ota & Takano, 2007) seem to confirm the younger generations’ general preference for the flattening of sentential pitch contours. Our younger subject groups, who come from both geographically and linguistically distinct communities (i.e., Hokkaido and Kagoshima), tend to share remarkably similar patterns of pitch variability.

The geographical origin of this particular prosodic change and the sociolinguistic implications, however, remain unspecified. Variable patterns of AP dephrasing involving unaccented APs in particular, which we propose as a potential linguistic change in progress, seem to involve regional discrepancies in terms of the progression of change, in that city speakers are likely to be more advanced than rural speakers.

It is rather obvious from a literature review that the observed regionality in the spread of the change may be closely linked to a long-standing nationwide change in progress – the obsolescence of local dialects in favor of the Tokyo variety norms (kyootsuugo-ka). A series of large-scale (both apparent time and real time) studies of Hokkaido dialects, for example, unanimously attribute the flattening of lexical accents among younger participants (under the age of 30) to the kyootsuugo-ka
process (e.g., Inoue, 1981; NINJAL, 1965, 1997). In an intriguing contrast, the change can also be seen in other regions as unifying local dialects, as a form of linguistic maintenance of local pride or localness (Tokugawa & Sanada, 2001) through the formation of new dialects (shin-hoogen) (Inoue, 1985). NINJAL’s (1965) report on Hokkaido dialect speakers across three generations, for example, identified this bidirectional trend of change in the speech of high school students: one toward the common variety (kyootsuu-go) that is spoken in the Tokyo metropolitan area and the other toward new local features that cannot be traced back either to local dialects or to the common variety. The latter type of linguistic change seems to be linked to a preference for non-prominence (mu-akusento-ka): the flattening or loss of accentual differentiation in many of two-mora words investigated (NINJAL, 1965, p. 40). Furthermore, the more recent NINJAL (1997) progress report also confirms that the kyootsuugo-ka process coexists with the emergence of a new dialect in lexical accentuation.

Whichever interpretation is more appropriate as a sociolinguistic generalization, larger-scale apparent time studies are vital and should involve finer age cohorts from broader varieties of both urban and rural dialect regions to obtain a reliable conclusion. It is also certain that any follow-up study should include systematic accounts of parallel data from speakers of Tokyo Japanese as the possible innovators to understand the processes of this potential linguistic change. Furthermore, not all of these observations should be interpreted as direct evidence for a change in progress that eventually leads to phonological changes in the language since the possibility of age grading (i.e., as our younger speakers age, they will not speak in flattened pitch contours) cannot be totally ruled out under the framework of the present study.

The present study also reveals that the observed pitch variability involves another important social correlate beyond the speaker’s age and regionality. Our analysis of accented APs in particular demonstrates that Japanese women are more likely to exploit a dynamic movement of pitch than men in naturalistic language use in accord with previous studies of gender differentiation in Japanese prosody. Gender-linked variability has thus far been relatively understudied in laboratory-based studies of prosody in general. We hope that the current results will contribute to a better understanding of how variably but systematically speech prosody is exploited in everyday language use.

Finally, our sociophonetic approach to multiple registers has produced several significant consequences for describing the reality of variable pitch realizations and the changes that may be occurring in Japanese language use. The realizations of sentential pitch are found to vary in systematic response to different registers. As also demonstrated in several prior studies (e.g., Cohen et al., 1982; Umeda, 1982), generalizations based on a single register or style (typically the reading of isolated
sentences) cannot be blindly applied to other registers and styles (cf. Koori, 2007; Takano, 2008; Yaeger-Dror, 1996, 1997, 2002a, 2002b). In the current results, the readings of isolated sentences yielded the most noticeable differentiation between the age groups (Ota & Takano, 2014), but in the other more spontaneous mode of speech (i.e., story-telling register), the age-linked patterns were still consistently maintained but became less salient. All of these gaps between the registers, however, are precisely what has been predicted in prior studies that share an emphasis on the necessity of multiple registers and styles for legitimate research on prosody.

The present study has also demonstrated that the systematic account of natural speech production, one of the empirical premises in variationist linguistics (Labov, 1972, 1984, 1994, 2001) is equally promising for describing linguistic variation and change in suprasegmental phenomena such as prosody. Our investigation of a naturalistic speech production task has succeeded in revealing several new dimensions in Japanese sentential pitch that would have gone undiscovered if research had remained concerned with artificial production tasks as the sole resource for analytical data.

It has been found, for example, that the flattening of sentential pitch specific to younger age groups is interrelated with relatively frequent dephrasings of APs in their more spontaneous, naturalistic mode of speech. We pointed out that the weakening and the resultant flattening of pitch involve qualitatively different mechanisms, contrary to the predictions of the phonological models based primarily on read-aloud speech. The weakening of pitch contours does not seem to be derived from the prosodic environment that is theoretically assumed to favor it by the phonological rule of downstep (i.e., the environment preceded by accented AP), but instead involves unaccented APs as a way of making inherently non-prominent features even further less prominent.

Contrary to our initial expectations, which motivated our exclusive focus on behavior of accented APs, it seems very unlikely that younger generations who manipulate excessive degrees of sentential pitch flattening will be defaulting lexical pitch accents of accented APs in spontaneous language use. Instead, it may be a likely change for us to watch for with regard to the future of Japanese prosodic changes that (at least in everyday language use) lexical pitch accents of accented APs will be maintained with a smaller magnitude of pitch movement as compared with prior generations, but, at the same time, dephrased unaccented APs will be accommodated increasingly to adjacent APs. Accounts of the causality of this linguistic constraint, however, certainly require serious investigation.
Acknowledgements

The present study is supported by a Japanese Grant-in-Aid for Scientific Research (C) (1) (No. 16520284), the NINJAL Collaborative Research Project (organized by Yoshiyuki Asahi), and a Japanese Grant-in-Aid for Scientific Research (B) (No. 25284082, represented by Kenjiro Matsuda). We are also grateful to Dr. Akira Utsugi at Nagoya University and anonymous APLV referees for their helpful suggestions and insightful comments.

References


Inoue, Fumio (1981). Hokkaido nai-no hoogen sa [Regional differences within Hokkaido dialect]. In *Igarashi Saburoo Sensei koki kinen shukuga ronbun shuu* [Papers dedicated to the celebration of Professor Saburoo Igarashi’s 70th birthday] (pp. 72–83). Sapporo: Hokkaido Hoogen Kenkyuu Kai [Hokkaido dialect circle].


Koiso, Hanae, Yoneyama, Seiko, Maki, Yoichi, & Fon, Janice (2003). *Nihongo Hanashi Kotoba Koopasu-o mochiita danwa koozoo-to inritsu-tono kankei-ni kansuru ichi koosatsu*
An analysis of prosody and discourse structure in the corpus of spontaneous Japanese. *Jinkoo Chinoo Gakkai Kenkyuu kai Shiryo* [Sources for the Japanese Society for Artificial Intelligence Conference], 37, 139–144.


Ono, Yoneichi (1981). *Gengo kenkyuu-to Hokkaido hoogen* [The study of language and Hokkaido dialects]. In *Igarashi Saburo Sensei koki kinen shukuga ronbun-shuu* [Papers...
A sociophonetic approach to variation in Japanese pitch realizations 35

dedicated to the celebration of Professor Saburou Igarashi’s 70th birthday] (pp. 194–209). Sapporo: Hokkaido Hoogen Kenkyuu Kai [Hokkaido dialect circle].


# Appendix

Pitch movement (based on z-scores) in IP-initial sequences of 2 APs: Age/region group averages in the picture story-telling task (also see Figure 2)

<table>
<thead>
<tr>
<th>AP Combinatory Patterns</th>
<th>Column I</th>
<th>Column II</th>
<th>Column III</th>
<th>Column IV</th>
<th>Column V</th>
<th>Column VI</th>
<th>Column VII</th>
<th>Column VIII</th>
<th>Column IX</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shin-Hidaka Younger (SHY)</td>
<td>Group Average, No. of Subjects &amp; Gender</td>
<td>Group Average, No. of Subjects &amp; Gender</td>
<td>Group Average, No. of Subjects &amp; Gender</td>
<td>Group Average, No. of Subjects &amp; Gender</td>
<td>Group Average, No. of Subjects &amp; Gender</td>
<td>No. of subjects &amp; Gender</td>
<td>No. of subjects &amp; Gender</td>
<td>*Types of Statistical Significance (p&lt;.05)</td>
<td></td>
</tr>
<tr>
<td>Total No. of Tokens</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shin-Hidaka Older</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sapporo Younger</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sapporo Older</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kagoshima Younger</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Younger Groups Average</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Older Groups Average</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Shift 1
- Accented + Accented (+161)
  - ANP: −1.26 (m5, f5), −1.28 (m5, f5), −1.41 (m5, f3), −1.63 (m3, f5), −1.29 (m4, f5), −1.33 (m14, f13), −1.44 (m8, f10)
  - *Types of Statistical Significance (p<.05): m −1.16/f −1.60 (Hokkaido); m −1.10/f −1.54 (Younger)

### Shift 2
- ANP: 0.63 (m5, f5), 0.79 (m5, f5), 0.81 (m5, f3), 0.65 (m3, f5), 0.75 (m4, f5), 0.62 (m14, f13), 0.72 (m8, f10)
  - *Types of Statistical Significance (p<.05): m 0.56/f 0.86 (Hokkaido); 2 APs 0.42/3 APs 0.81/4 or more APs 0.90 (Hokkaido); m 0.55/f 0.91 (Younger); 2 APs 0.43/3 APs 0.73/4 or more APs 1.02 (Younger)

### Shift 3
- ANP: −0.86 (m5, f5), −1.11 (m5, f5), −1.46 (m5, f3), −0.86 (m3, f5), −1.0 (m4, f5), −0.94 (m14, f13), −1.0 (m8, f10)
  - *Types of Statistical Significance (p<.05): m −0.82/f −1.23 (Hokkaido); m −0.89/f −1.29 (Younger); Sapporo −1.46/ S-H −0.86/Kagoshima −1.0 (Younger); 2 APs −0.81/3 Aps −1.1/4 or more APs-1.34 (Younger)

### Shift 5
- ANP: −0.23 (m5, f5), −0.32 (m5, f5), −0.64 (m5, f3), −0.21 (m3, f5), −0.25 (m4, f5), −0.32 (m14, f13), −0.28 (m8, f10)
  - *Types of Statistical Significance (p<.05): Sapporo −0.64/S-H −0.23/Kagoshima −0.25 (Younger)
### Appendix (continued)

<table>
<thead>
<tr>
<th>Shift</th>
<th>Column I</th>
<th>Column II</th>
<th>Column III</th>
<th>Column IV</th>
<th>Column V</th>
<th>Column VI</th>
<th>Column VII</th>
<th>Column VIII</th>
<th>Column IX</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-1.61</td>
<td>-1.41</td>
<td>-1.34</td>
<td>-1.83</td>
<td>-1.28</td>
<td>-1.47</td>
<td>-1.65</td>
<td>*m −1.31/f −1.78 (Hokkaido); *m −1.22/f −1.59(Younger)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(m5, f5)</td>
<td>(m5, f5)</td>
<td>(m5, f5)</td>
<td>(m4, f6)</td>
<td>(m4, f5)</td>
<td>(m14, f15)</td>
<td>(m9, f11)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.50</td>
<td>0.76</td>
<td>0.66</td>
<td>0.71</td>
<td>0.54</td>
<td>0.56</td>
<td>0.74</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(m5, f5)</td>
<td>(m5, f5)</td>
<td>(m5, f5)</td>
<td>(m4, f6)</td>
<td>(m4, f5)</td>
<td>(m14, f15)</td>
<td>(m9, f11)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>-0.59</td>
<td>-0.82</td>
<td>-0.95</td>
<td>-0.68</td>
<td>-0.57</td>
<td>-0.66</td>
<td>-0.74</td>
<td>*m −0.57/f −0.94 (Hokkaido); *m −0.52/f −0.88 (Younger); *Sapporo −0.95/ S-H −0.59/Kagoshima −0.57 (Younger)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(m5, f5)</td>
<td>(m5, f5)</td>
<td>(m5, f5)</td>
<td>(m4, f6)</td>
<td>(m4, f5)</td>
<td>(m14, f15)</td>
<td>(m9, f11)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-0.09</td>
<td>-0.06</td>
<td>-0.29</td>
<td>0.04</td>
<td>-0.04</td>
<td>-0.10</td>
<td>-0.01</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(m5, f5)</td>
<td>(m5, f5)</td>
<td>(m5, f5)</td>
<td>(m4, f6)</td>
<td>(m4, f5)</td>
<td>(m14, f15)</td>
<td>(m9, f11)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.64</td>
<td>-1.10</td>
<td>-0.72</td>
<td>-0.98</td>
<td>-0.62</td>
<td>-0.67</td>
<td>-0.87</td>
<td>*Younger −0.68/Older −1.02 (Hokkaido); *m −1.31/f −1.78(Hokkaido); *Age x Gender x Region interact (Hokkaido)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(m5, f5)</td>
<td>(m5, f5)</td>
<td>(m5, f5)</td>
<td>(m4, f6)</td>
<td>(m5, f5)</td>
<td>(m15, f15)</td>
<td>(m10, f11)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.65</td>
<td>0.68</td>
<td>0.84</td>
<td>0.87</td>
<td>0.56</td>
<td>0.67</td>
<td>0.74</td>
<td>*No. of APs 2 = 0.59, 3 = 0.88, 4 = 0.80 (Hokkaido)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(m5, f5)</td>
<td>(m5, f5)</td>
<td>(m5, f5)</td>
<td>(m4, f6)</td>
<td>(m5, f5)</td>
<td>(m15, f15)</td>
<td>(m10, f11)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>-1.33</td>
<td>-1.25</td>
<td>-1.77</td>
<td>-1.34</td>
<td>-1.18</td>
<td>-1.43</td>
<td>-1.31</td>
<td>*Younger −1.55/Older −1.31 (Hokkaido); *m −1.20/f −1.64 (Hokkaido); *m −1.14/f −1.71 (Younger); *Sapporo −1.57/S-H −1.30 (Hokkaido); Sapporo −1.77, S-H −1.33, Kagoshima −1.18 (Younger); Age x Gender x Region interact (Hokkaido)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(m5, f5)</td>
<td>(m5, f5)</td>
<td>(m5, f5)</td>
<td>(m4, f6)</td>
<td>(m5, f5)</td>
<td>(m15, f15)</td>
<td>(m10, f11)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-0.69</td>
<td>-0.58</td>
<td>-0.93</td>
<td>-0.47</td>
<td>-0.62</td>
<td>-0.74</td>
<td>-0.55</td>
<td>*Younger −0.80/Older −0.55 (Hokkaido); *m −0.48/f −0.85 (Hokkaido); *m −0.52/f −0.97 (Younger); *No. of APs 2 = −0.81, 3 = −0.48, 4 or more = −0.73 (Hokkaido)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(m5, f5)</td>
<td>(m5, f5)</td>
<td>(m5, f5)</td>
<td>(m4, f6)</td>
<td>(m5, f5)</td>
<td>(m15, f15)</td>
<td>(m10, f11)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Appendix (continued)

<table>
<thead>
<tr>
<th>Unaccented + Shift 1</th>
<th>Column I</th>
<th>Column II</th>
<th>Column III</th>
<th>Column IV</th>
<th>Column V</th>
<th>Column VI</th>
<th>Column VII</th>
<th>Column VIII</th>
<th>Column IX</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unaccented (47)</td>
<td>-0.65</td>
<td>-1.44</td>
<td>0.73</td>
<td>-1.22</td>
<td>-0.76</td>
<td>-0.74</td>
<td>-1.20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shift 2</td>
<td>0.37</td>
<td>0.97</td>
<td>0.62</td>
<td>0.91</td>
<td>0.43</td>
<td>0.51</td>
<td>0.82</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>Shift 3</td>
<td>-0.73</td>
<td>-1.10</td>
<td>-0.63</td>
<td>-0.76</td>
<td>-0.78</td>
<td>-0.71</td>
<td>-0.96</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>Shift 5</td>
<td>-0.35</td>
<td>-0.14</td>
<td>-0.01</td>
<td>0.15</td>
<td>-0.36</td>
<td>-0.21</td>
<td>-0.14</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Younger −0.70/Older −1.20 (Hokkaido)  

*m 0.06/f −0.28 (Hokkaido); *Sapporo −0.08/S-H −0.23 (Hokkaido); *No. Of APs 2 = −0.49, 3 = −0.37, 4 or more = −0.10 (Younger)
Abstract (Japanese)

本稿は、特に若い世代を中心に全国的規模で進行中の言語変化と想定される音声ピッチの平坦化現象を検証する。北海道の二地域（札幌市・新ひだか町）、および九州南部地方鹿児島市など計3地域において、同一調査手順に則った実地調査を行い、収集された発話データを話者の年齢・性別・母方言などの観点から比較検討した。また、従来の韻律研究では文章の読み上げ音声が分析の主な対象であった実を背景に、異なるレジスターを分析の射程に含める重要性も経験的に立証している。本調査では、6コマのイラストで描かれた物語を見た後、その物語の内容を口頭で即興的に復元するタスクから得られた自然発生的談話に近い音声データとして分析した。その結果、(1) 話者の母方言の差異とは無関係に、若い世代の話者のピッチは、(老年層話者のそれと比較して) 一貫して急激な下降 (declination) を伴う平坦化の傾向を示し、(2) 韻律句生成 (prosodic phrasing) においても、ピッチの平坦化に関連した年齢差が明らかになった。

Address for correspondence

Shoji Takano
Hokusei Gakuen University
Sapporo
Japan
stakano@hokusei.ac.jp