L2 intonation by Swedish learners of Japanese

Analysis of pitch accent and prosodic phrasing errors

Natsumi Goto

Supervisor: Dr Shinichiro Ishihara

Centre for Language and Literature, Lund University
MA in Language and Linguistics, Japanese
SPVR01 Language and Linguistics: Degree Project – Master's (Two Years) Thesis, 30 credits
August 2020

Abstract

This thesis examines the L2 intonation of Swedish learners of Japanese with a focus on pitch accent and prosodic phrasing errors. Previous studies have attempted to predict L2 learners' difficulties and errors, such as Contrastive Analysis Hypothesis (CAH) by Lado (1957) and Markedness Differential Hypothesis (MDH) by Eckman (1977; 1985). L2 Japanese pronunciation has been analysed at the word or phrase level by existing studies (Sakamoto 2010; Ogawasara 2015). Tsurutani (2007; 2011) studies the intonation of English learners of Japanese at the phrasal and sentence level intonation while lacking the perspective of syntax and prosody.

This thesis attempts to apply CAH and MDH to the case of Swedish learners of Japanese and to fill the gap in the findings of the existing L2 Japanese pronunciation studies. This study adds a perspective of the interaction between syntax and prosody and analyses the errors made by Swedish learners of Japanese at the phrasal and sentence level intonation.

Four Tokyo Japanese native speakers and fourteen Swedish learners of Japanese read the stimuli aloud while their speech was recorded for the analysis. How their speech depicts the syntactic structure, and prosodic features of the stimuli is discussed.

It was found that four Tokyo Japanese native speakers produced diverse contours, but the varieties can be explained under the theories of syntax-prosody mapping and prosodic well-formedness. In contrast, the diversity among Swedish learners of Japanese cannot be explained within the theories. Most of the learners' errors seems to be due to negative transfer from their L1 as well as other various factors. The causes behind their errors seem to be complex and need further research.

Keywords: L2 Japanese intonation, Swedish learners of Japanese, L2 errors at the phrasal and sentence levels, pitch accent.

Acknowledgements

Doing a Master degree in Sweden for two years has been really challenging and intense for me. During this challenge, there have been a lot of people helping me go through.

まずスウェーデンで修士課程を二年間やるという私の大きな挑戦に理解を示して支えてくれた私の家族に、一番の感謝を伝えたいと思います。留学中にそばにいれなかったことでたくさんの迷惑をかけたし、おばあちゃんの最期も一緒に見送ることができなくて本当にごめんなさい。多くの機会を与えてもらったので、自分なりに精一杯いろいろなことを学ぶことができた二年間でした。心からありがとう。

それから私の研究に快く参加してくれた日本人 4 人、日本語を勉強しているスウェーデン人の学生 14 人の皆さん、協力してくださって本当にありがとうございました。今年の 2 月くらいからはコロナウイルスの影響で参加してくれた日本人交換留学生も早期帰国、スウェーデン人学生にとっても楽しみにしていた日本への交換留学の中止など、言語を勉強している人にとって、つらく厳しい時期になったと思います。どうかこのような状況の中でも健康に気を付けて言語学習を続け、いつかまた日本または別のどこかの国で英語や日本語を使って、新しい道に進んでいってほしいと心から願っています。また、研究の過程で快く質問に答えてくださったり、授業観察をさせていただいたルンド大学とストックホルム大学日本語科の先生方、本当にありがとうございました。私の研究が日本語教育にすこしでも役に立てば幸いです。

そして石原先生。言語学の基礎があるとは言えない、かつ頭の回転がよく理解も早いともいえない私に、二年間根気強く指導をしてくださりありがとうございました。論文執筆を始めてからは、時間が十分に取れず満足とは言えない草稿を出したりなど、多くの迷惑をかけてしまったと思います。ここまで、本当に丁寧にご指導してくださりありがとうございました。

Thank you, also for Åsa, my study coordinator, who has been helping me since I asked questions about this course even before I was admitted to this Master degree. I really appreciate that you have made sure that everything goes well with my studies and other things like residence permit. Tack så mycket.

I appreciate for Centre of language and linguistics, and the course coordinator Jordan who organised the programme.

Table of contents

List of Figures

List of Tables

Romaji convention

Japanese vocabulary is transcribed based on Hepburn system of romanisation. Japanese words are provided in Italic, and the translation in English is marked with apostrophes. Mora with pitch accent is marked with an apostrophe at the end of mora transcribed in romaji.

Glossing

Glossing for this paper is in accordance with Leipzig Glossing Rules (Max Planck Institute for Evolutionary Anthropology, Department of Linguistics, n.d.) Abbreviations used in this paper are provided below.

Abbreviations

ACC Accusative

AP Accentual phrase

BPM Boundary pitch movement

CAH Contrastive analysis hypothesis

GEN Genitive

IP intonation phrase

LOC Locative

MDH Markedness differential hypothesis

NOM nominative

PAST Past tense

Chapter 1 Introduction

Mastering L2 pronunciation can be challenging for language learners, and it is quite common to hear a foreign accent or non-native elements in second language learners' speech (Major 2008; Tsurutani 2007; Mennen 2015; Zampini 2008). At the segmental or word level, there are many previous studies of analysing L2 learners' production (Major 2008). Also, it is considered to be relatively easy to instruct L2 learners about their segmental or word-level errors (Tsurutani 2011). When it comes to L2 learners' errors in phrasal or sentence intonation, however, there is a severe lack of research. Due to the lack of previous studies, it would be challenging for L2 teachers to instruct learners. Another reason as to why the instructions in L2 intonation has not been introduced well in a classroom setting might be because of the complex features of intonation.

Intonation can be defined as:

> the tonal structure of speech expressed by the melody produced by our larynx. It has a phonetic aspect, the fundamental frequency (F0), and a grammatical (phonological) aspect (Fery 2016, p.6)

There are very few previously conducted studies examining L2 learners' errors in intonation (Tsurutani 2007; 2011). Most of the studies focus on L2 learners' segmental or word level production. In the case of L2 Japanese learners, most of the existing literature works with L2 Japanese learners' segmental or word-level errors (Sakamoto 2010; Ogasawara 2015). Tsurutani (2007; 2011) conducted studies, where she analysed the errors of English learners of Japanese at phrasal and sentence level. Her studies have found that some of the learners' errors seem to be explained as a negative transfer from their L1. In her studies, the production of L2 Japanese learners is analysed in terms of the degree of deviation from Japanese native speakers' model speech. In her studies, the Japanese native participants produced similar contours, that enable her to assume that native speakers would produce similar contours for the same set of stimuli. However, the various shape of contours can be commonly observed even among native speakers of a language because intonation consists of the combination of requirements from prosody and syntax. Therefore, this study aims to explore the errors produced by Swedish learners of Japanese from the perspective of the syntax-prosody mapping and prosodic well-formedness. This study is also motivated by the situation surrounding L2 learners. That is, when L2 learners' speech contains elements of a foreign accent or non-nativeness in intonation, that could affect how the intelligibility of L2 learners is judged by native speakers of their target language (Anderson-Hsieh, Johnson and Koehler 1992; Sato 1995; Mouri, Hirose and Minematsu 2003). Furthermore, L2 learners might not be recognised as proper members of the target language

community due to their foreign accent in their speech (Baker 2011; Mesthrie, Swann, Deumert and Leap 2009). To guide this study, the following two research questions are proposed:

Research question 1: How is the interaction between the syntax-prosody mapping and prosodic well-formedness realised in the utterances produced by Swedish learners of Japanese in comparison to Japanese native speakers?

Research question 2: Can their errors be explained as a transfer from their L1, Swedish?

To answer these research questions, this study explores speech produced by Swedish learners of Japanese. Four stimuli sentences in Japanese were carefully formulated, which have two conditions: accentedness of the first word and syntactic structure. The stimuli were presented to Tokyo Japanese native speakers, and Swedish learners of Japanese, and they were asked to read aloud for recording. The data analysis of their speech takes a qualitative approach to describe the detailed characteristics of the speech produced by each participant.

This study would be able to contribute with a new perspective on the studies related to the error analysis of intonation by L2 Japanese learners. That is, this study analyses L2 learners' errors with a focus on the relation between syntax and prosody. The findings of this research could provide, not only further insight to relevant research, but also help to improve L2 Japanese language education.

This thesis is structured as follows. In Chapter 2, the literature of existing studies is reviewed before being proposed the research questions and aims of this study. Chapter 3 presents the aims and research questions of this study based on the discussion of previous studies as well as describes the method of this study. Chapter 4 contains the results of the studies, and Chapter 5 discusses the result. Chapter 6 provides a summary and a conclusion of this study.

Chapter 2 Theoretical background

This chapter provides the theoretical background that motivates the two research questions proposed in section 3.1. Firstly, the general framework of intonational phonology and some background about intonation of Tokyo Japanese are discussed in section 2.1. Section 2.2 explains intonation of Swedish which is the mother tongue of this study's participants. Section 2.3 introduces the general background about second language phonology. Previous studies on L2 Japanese prosodic acquisition are discussed in section 2.4.

2.1 Intonation of Tokyo Japanese

This section mainly focuses on Tokyo Japanese, since the variety is the most widely used in Japanese language education. For this study, the native Japanese participants were recruited from those who come from the areas where Tokyo Japanese is mainly used. The analysis also concerns Tokyo Japanese. First, the Autosegmental Metrical model of intonation phonology (AM model) and the ToBI system are introduced. Then we move onto the discussion of intonation of Tokyo Japanese.

In order to describe the prosodic structure and tone assignment, the Autosegmental-Metrical model of intonational phonology (AM model) has been developed (Liberman 1975; Bruce 1977; Pierrehumbert 1980). Under the AM model, languages are described in the transcription system termed as the ToBI system (Tone and Break indices). The ToBI system was originally invented for English and has now been adapted for different languages, including Japanese. For Japanese, there are two versions of the ToBI system: the original version by Venditti (2005) and the extended version X-J_ToBI framework (Venditti, Maekawa, and Beckman 2008). This study adopts the latter. The AM framework assumes that intonation consists of a phonological organisation (Jun 2005, p.1). The organisation is described as a sequence of distinctive tones such as High (H tone) and Low tone (L tone). Distinctive tones are aligned with a specific syllable and mark a prominence of the syllable in a word as well as the prosodic phrasing of an utterance. This phonological representation of tone maps onto the phonetic realisation following language-specific phonetical realisation rules and phonological representation (Jun 2005, p.2). The AM framework distinguishes between local events and transitions (Ladd 2008). The local events are tones or combinations of tones that are associated with a prominent syllable. Between each local event, tone assignment is not specified and termed as transitions. Therefore, the ToBI system transcribes distinctive tones (To for tone) that are

associated with each word and the degree of transition and strength (BI for Break Index). The local event, namely, pitch accent, has two aspects; pitch accent builds up an intonational contour as a sentence, and shows what parts of a sentence are more prominent than the others, and marks prosodic phrasing in a sentence (Ladd 2008). This aspect is described as an intonational pitch accent. Another aspect of pitch accent is the lexically specified pitch accent. For instance, Japanese and Swedish words have lexically embedded pitch accent.

In the X-JToBI transcription system, two levels of prosodic phrasing are assumed; the lower level is the *Accentual Phrase (AP)*, and the higher level is the *Intonation Phrase (IP)* (Igarashi 2015). The X-JToBI assumes that the organisation of the two levels follows the Strict Layer Hypothesis (Selkirk, 1986). It means that "any domain at a given level of the hierarchy consists exclusively of domains at the next lower levels of the hierarchy" (Igarashi 2015, p.527). As Table 1 exemplifies, different terminologies for Japanese prosodic phrasing have been proposed (Kawakami 1957; Uwano 1989, 1999; McCawley 1968, Poser 1984; Kubozono 1993)

Table 1: The summary of Japanese prosodic phrasing (taken from Ishihara 2015, p.570)

<table>
<tr><th colspan="2">AM theory / ToBI</th><th colspan="2">Minor/Major Phrase</th><th>Syntax–Prosody Mapping</th></tr>
<tr><td>Pierrehumbert & Beckman</td><td>J_ToBI, X-JToBI</td><td>McCawley, Poser, Kubozono</td><td>Kawahara & Shinya</td><td>Ito & Mester, Selkirk, this chapter</td></tr>
<tr><td rowspan="2">Utterance</td><td rowspan="3">Intonation Phrase</td><td rowspan="2">(not discussed)</td><td>Utterance</td><td>ι_{max}</td></tr>
<tr><td>Intonational Phrase</td><td>PClause or Intonational Phrase (ι)</td></tr>
<tr><td>Intermediate Phrase</td><td>Major Phrase</td><td>Major Phrase</td><td>Phonological Phrase (φ)</td></tr>
<tr><td>Accentual Phrase</td><td>Accentual Phrase</td><td>Minor Phrase</td><td>Minor Phrase</td><td>φ_{min}</td></tr>
</table>

Ishihara (2015, p.570) categorizes the various terminology into three groups; researchers who work under the AM framework such as Pierrehumbert and Beckman (1988) and the ToBI frameworks (Venditti 2005; Maekawa et al. 2002), researchers who adapt McCawley's proposed terminology (1968), and the last group is Ito and Mester (2007, 2012, 2013) and Selkirk (2009, 2011) who adopt syntax-prosody mapping hypothesis. Their views do not adapt to the Strict layer hypothesis but instead assume prosodic recursions of the same level prosodic hierarchy. In the X-JToBI system, the Strict Layer Hypothesis, and the two levels of prosodic phrasing are adapted. Therefore, the discussion in this paper shall follow these as well. In the X-JoBI, the *Accentual Phrase (AP)* is defined as

1) having a delimitative rise to high around the second mora and a subsequent gradual fall to low at the end of the phrase

2) having at most one lexical pitch accent. While a typical AP consists of one lexical word plus any following particles or postpositions (Igarashi 2015, p. 529).

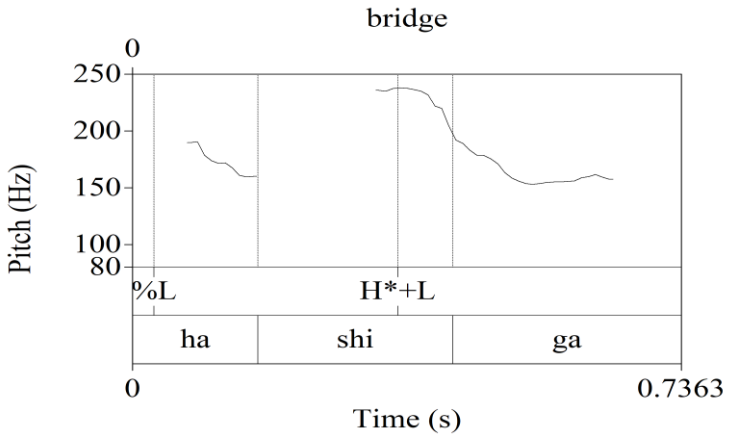


Figure 1: Bridge contour (created by the author)

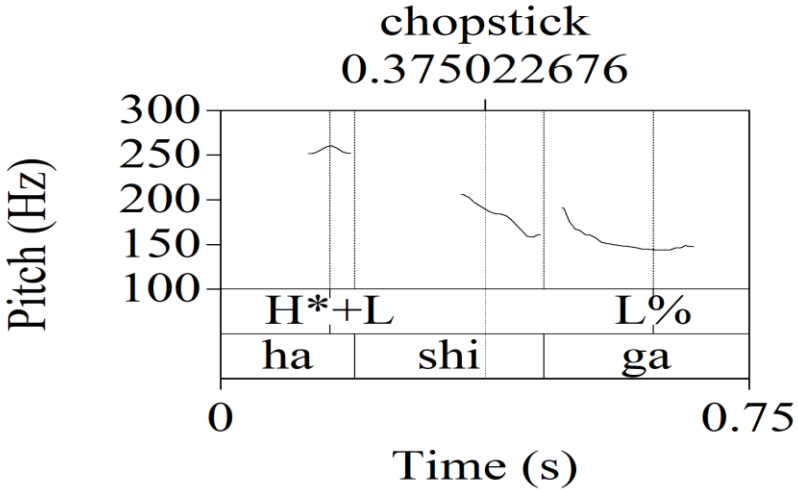


Figure 2: Chopstick contour (created by the author)

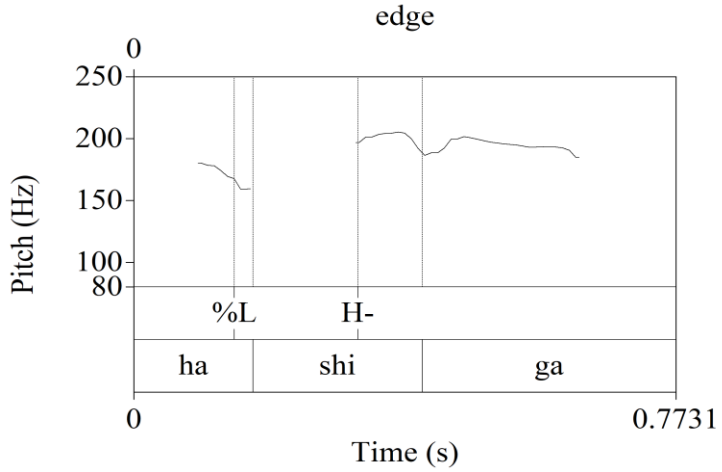


Figure 3: Edge contour (created by the author)

The X-JoBI is based on the AM theory (Ladd 2008; Gussenhoven 2004; Jun 2014) that describes a contour as a series of distinctive tones. As to a typical contour of AP in Tokyo Japanese that contains only unaccented words, the contour is represented as %L H- L%. L% is an initial low boundary tone. H- is called phrasal high that forms the initial rise of AP. L% means a final low boundary tone. The pattern of an unaccented AP can be seen in the example of *Hashi* 'edge' in Figure 3. *Hashi'* 'bridge' has a pitch accent on the second mora *shi* and has a typical contour as AP that contains an accented word. The contour begins with the initial low tone %L, a phrasal high H- and a pitch accent H*+L follows. The asterisk in H* stands for 'the tone associated with the mora that is governed by the accented syllable' (Igarashi 2015, p.529). As in *ha'shi* 'chopstick' in Figure 2, whose first mora *ha* is accented, it is common that an initial low boundary tone is hardly observed. Not only AP with an accented initial mora but also an AP initial syllable whose initial syllable is heavy such as a long vowel and a diphthong, %L is likely to be imperceptible as well (Igarashi 2015). AP is the domain of pitch accent realisation, and one AP has at most one pitch accent (Ito and Mester 2012, 2013).

In Figures 1, 2, and 3, three segmentally homonymous words with three different tonal specifications (*hashi*) are provided. The three words are accompanied by the case marker *ga* to exhibit clearer contour differences. In Tokyo Japanese, lexical contrast is made in two ways; the absence and presence of pitch accent and if there is one, its location (Kawahara 2015). A pitch accent in Japanese is characterised by an abrupt fall in F0, that is transcribed as H*+L in the X-JToBI system. *Ha'shi* 'chopstick' carries a pitch accent (H*+L) on the first mora. *Hashi*' 'bridge' has a pitch accent on the second mora '*-shi.'* The third example, *hashi* 'edge', is an unaccented word, which does not have a pitch accent. The contour is realised with a low boundary tone (%L) and an adjacent phrasal H tone.

The intonation phrase (IP) is above AP in the hierarchy of prosodic phrasing. Within an IP, the range of pitch is determined. When it transits to the next IP, a new range of pitch is set by a speaker, which is termed as pitch reset (Igarashi 2015, pp.541). Within an IP, after an accented AP, the pitch range of the succeeding AP is compressed compared to the previous AP. This phenomenon is called a downstep.

Figure 4: The contour of downstep (generated with Hirose Laboratory's website, n.d.)

(1) Ya'kuza no one'esan ga ima'shi-ta
Yakuza GEN sister NOM exist-PAST
'There was a yakuza sister.'

As in Figure 4, the first word, *ya'kuza* is an accented word and the adjacent word, *one'esan* 'sister' is also an accented word. The pitch range for *one'esan* is compressed in comparison to *ya'kuza*. In other words, the effect of downstep can be seen at *one'esan*. It seems that most of the researchers agree that IP (called a Major phrase by Poser, 1984; called an intermediate phrase by Pierrehumbert and Beckman 1988; Major phrase in Kubozono 1988, 1993) is the domain of downstep.

Boundary pitch movements can mainly be observed at the phrasal final or sentence-final (Igarashi 2015). Igarashi (2015, p.545) presents four main types of BPM in the inventory of the X-JToBI system.

(1) H% (Simple rise)

(2) LH% (Scooped rise)

(3) HL% (Rise-fall)

(4) HLH% (Rise-fall-rise)

There are two levels of prosodic phrasing; accentual phrase (AP) and intonation phrase (IP) in Japanese. The boundary of phrasing can be affected by various linguistic factors such as word accent, syntactic structure, focus and discourse factors (Venditti 2005). As to the relation between syntax and prosody, the field of syntax-phonology interface deals with how these components of

language interact with each other (Ishihara 2015). Ishihara (2015) mentions the three main aspects that have been discussed in the syntax-phonology interface: syntax-prosody mapping, prosodic well-formedness, and information structure-prosody mapping. Considering the relevance to this study, the first two theories are discussed in detail below.

Syntax-prosody mapping is 'the correspondence between the syntactic structure and the prosodic structure' (Ishihara 2015, p. 577). In the theory of syntax-prosody mapping, how much or what syntactic information is mapped onto sentence prosody is discussed. For example, a syntactic structure of the same sentence in Figure 4 is illustrated in Figure 5.

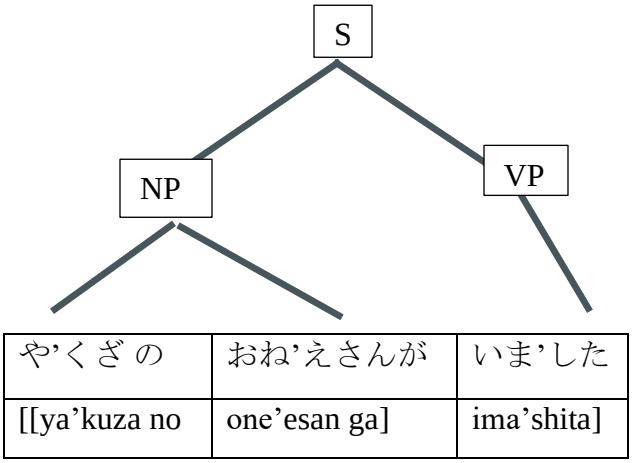


Figure 5: Sentence example with a left-branching tree

The end-based theory assumes a loose relation between syntactic and prosodic structure. It predicts that syntactic structure is mapped onto the prosodic structure, but not the other way around. Another critical point of the end-based theory is that it does not concern the right edge of XP, but only the left edge of XP is mapped on prosody in the case of Japanese prosody. In Figure 5, a noun phrase, *ya'kuza no one'esan ga* 'sister of a yakuza' forms a left-branching constituent since the first phrase, an NP *ya'kuza no* modifies another NP *one'esan* as a specifier. *Ima'shita* forms a verb phrase (VP). At the beginning of the sentence, there are two left edges of XP boundaries. According to Selkirk and Tateishi (1991), the left edge of a syntactic XP is mapped onto prosody as the left edge of an IP in Japanese. Therefore, it predicts that there is no boundary of IP between *one'esan ga* and *imashita*. Since the IP boundary is not predicted, the effect of downstep can be observed at *one'esan* and *ima'shita*. The example sentence is in one IP, but each accented word forms its own AP, which generates three APs. In this particular case, the verb *imashita* is one word that does not contain branching. Due to no XPs branching, the boundary that is predicted for the last VP is ignored in Syntax-prosody mapping.

Not only syntax-prosody mapping, but prosodic well-formedness also comes into play in terms of prosodic phrasing. Prosodic well-formedness is one of the factors that influences how prosodic structure is formed (Ishihara 2015). Kubozono (1993, p.220) illustrates the cases where one of the prosodic well-formedness, namely rhythmic effects, influence prosody. Rhythmic effect is observed in left-branch phrasing that consist of four accented words. It is predicted to see the effect of downstep as shown in Figure 6, but instead the pitch contours can be like in Figure 7. The phenomenon that the F0 peak of the third phrase is depicted as high as the second phrase or even higher is termed as rhythmic effect. Figure 8 is one of the examples in Kubozono (1993). The branching structure is drawn by the author.

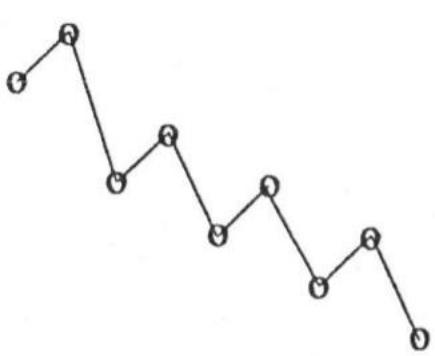


Figure 6: Predicted downstep contour taken from (Kubozono 1993, p. 218)

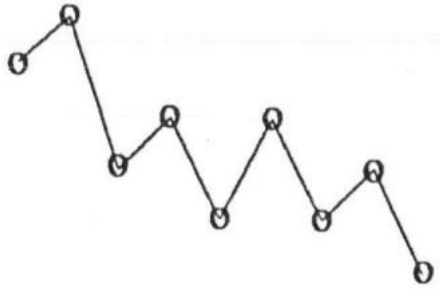
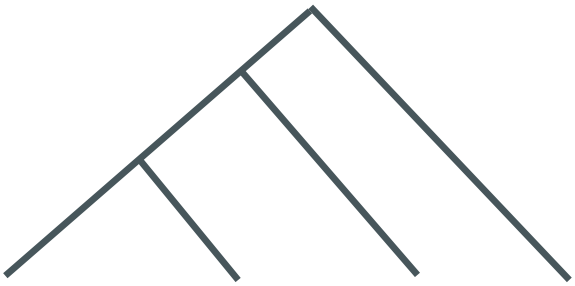


Figure 7: Downstep contour with rhythmic effect taken from (Kubozono 1993, p. 218)



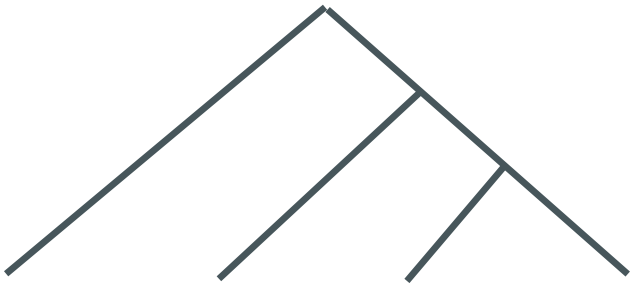
example	[[[ma'riko ga	no'n-da]	wa'in no]	nio'i]
end based theory prediction	{IP1			}
Phrasing with rhythmic effect	{IP1	}{ IP2		}

Figure 8: Rhythmic effect example phrase from (Kubozono 1993, p. 220)

(2) Ma'riko ga no'n-da wa'in no nio'i
Mariko NOM drink-PAST wine GEN smell
'The smell of wine that Mariko drank.'

An example phrase in Figure 8, *ma'riko ga no'nda* forms a relative clause. This noun phrase modifies *wa'in no* and forms a noun phrase as well. *Ni'oi* also is modified by the noun phrase that consists of the previous three words. Therefore, there are three left edges of each noun phrase at the beginning of the example phrase. Based on the end-based theory, no IP boundary is predicted between each word. At the beginning, there would be three left edges of XPs. Since IP is the domain of downstep, the effect of downstep is predicted to be observed at *no'n-da*, *wa'in no* and *nio'i*. A predicted contour is shown in Figure 6 and the effect of downstep appears throughout the phrase. Despite of the prediction, the data in Kubozono's study (1993) shows that the F0 peak of the third word is realised as high as or higher than the preceding word. [1]This phenomenon can be explained by the analysis of prosodic phrasing as shown 'phrasing with rhythmic effect' in Figure 8. Unexpected F0 rise of the third word can be caused by the insertion of a new IP boundary due to rhythmic effect. This analysis is consistent with the study conducted by Shinya, Selkirk and Kawahara (2004) where they use different example sentences with left branching structure of four accented words.

[1] Kubozono (1993) analyses that the rise in the third phrase is due to rhythmical boost and the prosodic phrasing for this example would be one IP as predicted.



example	[ya'kuza ga	[one'esan ni	[nomi'mono o	urima'shita]]]
prediction	{IP1 }{IP2 }{IP3 }			
Phrasing with rhythmic effect	{IP1 }{IP2 }			

Figure 9: Example sentence with the sequence of four accented words

(3) Ya'kuza ga one'esan ni nomi'mono o urima'shi-ta
Yakuza NOM sister LOC drink ACC sell-PAST
'Yakuza sold a drink to sister.'

Another case where the rhythmic effect influences prosodic phrasing is exemplified in Figure 9. The example sentence consists of four accented words. *Ya'kuza ga* forms a noun phrase as a subject of the sentence. The following *one'esan ni nomi'mono o urima'shita* forms a verb phrase that contains the indirect *one'esan n*i and the direct object *nomi'mono o* of the verb. Based on the end-based theory, there would be three boundaries of IP as there are three left edges of syntactic XPs. Due to the rhythmic effect, the prosodic phrasing marked as 'phrasing with rhythmic effect' in Figure 9 can be observed as well. In this case, the predicted IP boundary between *ya'kuza ga* and *one'esan ni* is not realised, and instead, the sentence has two IPs that have two accented words, respectively.

In this 2.1 section, the prosody of Tokyo Japanese was discussed. The main features can be summarised as follows. Tokyo Japanese employs pitch accents in two ways; locations and absence/presence of pitch accent. In Tokyo Japanese, pitch accent is characterised with an abrupt fall in F0, which is transcribed as H*+L in the X-JToBI system. Syntactic and rhythmic factors can influence prosody as can be seen in the phenomena like downstep. In the next section 2.2, the prosody of Swedish, the mother tongue of this study's participants is discussed.

2.2 Intonation of Stockholm Swedish

Since language learners' errors are often influenced by their L1, this study discusses the basic features of Swedish with a focus on its usage of pitch. I acknowledge that Swedish has regional varieties, and there are many studies on the area. This study analyses Swedish learners of Japanese who study Japanese at Lund and Stockholm University. However, considering the purpose of my study, this study mainly concerns Central Swedish, which is widely used in Stockholm. When Swedish is mentioned in this paper, it refers to Central Swedish.

Swedish employs both stress and pitch accent systems. Pitch accent and stress is distinctive from each other (Bruce and Hermans 1999). There are two types of pitch accent: *accent 1* and *accent 2*. In addition to *accent 1* and *2*, there are two levels of prominence named as *big accent* and *small accent* (*focal accent* for *big accent* and *word accent* for *small accent* in Riad, 2013). All words in Swedish have either *accent 1* or *accent 2*, which is one of the points that differs from Japanese since Japanese has unaccented words. Pitch accent is only associated with stressed syllables in both *accent 1* and *accent 2*. Monosyllabic words and words with final stress are always assigned *accent 1*. Table 2 illustrates pitch accent patterns in Central Swedish (Riad 2013). Only accent 1 in Swedish is reported to appear as a foreign accent (Bruce and Hermans 1999). Depending on Swedish regional varieties, the shape of contours for accent 1 is different. Figure 10 is a part of the figure extracted from Gårding and Lindblad (1973, p.48). It visualizes the difference in the accent 1 contour between Skånska and Rikssvenska. The contour in Skånska is the one named as 1A and 2A is Rikssvenska in Figure 10.

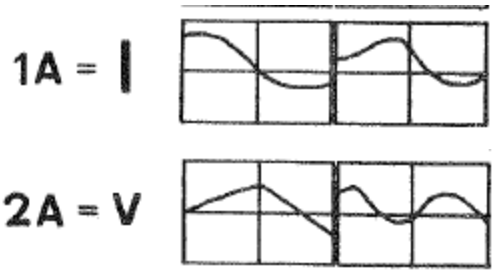


Figure 10: Accent 1 contours in Skånska and Rikssvenska (taken from Gårding and Lindblad 1973, p 48)

In Figure 10, the contours in *accent 1* in Skånska (1A) and Rikssvenska (2A) are on the left side. In Skånska, the timing of pitch peak in *accent 1* is earlier than that of Rikssvenska. Therefore, it is possible that the different location of pitch peak in *accent 1* are observed as a negative transfer of *accent 1* into Japanese with a different location of pitch peaks depending on which regional varieties the Swedish learners of Japanese speak.

Table 2: Swedish word accent (remade based on Riad 2013, p.8)

	Accent 1	Accent 2	Accent 2 for compounds
Big accent	L*H	**H***LH	**H***L*H
Small accent	HL*	**H***L	**H***L

Big accent is assigned based on different reasons, for instance, information structure of sentences such as givenness or focus and word position in utterances. Unique tones for *accent 2* are marked in bold letters in Table 2. At big accent level, *accent 1* words are assigned a contour of L*H where the first L tone is associated with a stressed syllable. For *accent 2*, the contour gets a new H tone that is associated with a stressed syllable. In compound words, the first two tones, H and L tone are associated with stressed syllables since more stressed syllables are available. At the small accent level, both *accent 1* and *accent 2* have HL contour. However, L tone is associated with a stressed syllable in *accent 1*, and H tone is associated with a stressed syllable in *accent 2* due to the difference in timing among them. Before discussing tone at the phrasal level, Swedish prosodic hierarchy, which plays a vital role in tonal representation, is discussed. The Swedish prosodic hierarchy is composed of the following levels: two levels of prosodic words, i.e., minimal and maximal prosodic words, prosodic phrase, and intonation phrase (Myrberg and Riad 2015).

Table 3: Swedish prosodic hierarchy (made by the author based on Myrberg 2010, p. 33)

Prosodic categories	features
Intonation Phrase	It corresponds to a syntactic phrase The domain of initiality accent and right edge boundary tone
Prosodic Phrase	Mostly corresponds to a syntactic phrase The domain of big accent, big accent is a marker of a PP
Maxmimal Prosodic word	One pitch accent per maximal prosodic word
Minimal Prosodic word	Domain of syllabification One stress per minimal prosodic word

As shown in Table 3, each category roughly corresponds to syntactic words for prosodic words, syntactic phrases for prosodic phrases, and clauses for intonation phrase. Minimal prosodic words have one stress and the domain of syllabification. Maximal prosodic words have a one pitch accent.

In addition to big and small accent, there is a tone that marks prominence in Central Swedish. Myrberg (2010, 2013) proposes an initiality accent. Initiality accents have a function as a left-edge prominence marker (Riad 2013). An initiality accent is associated with a stressed syllable and appears both in *accent 1* and *accent 2* words just like other big accents. Initiality accent and big accent behave in a similar way phonologically, while the distribution of initiality accents does not rely on information structure but purely on prosody.

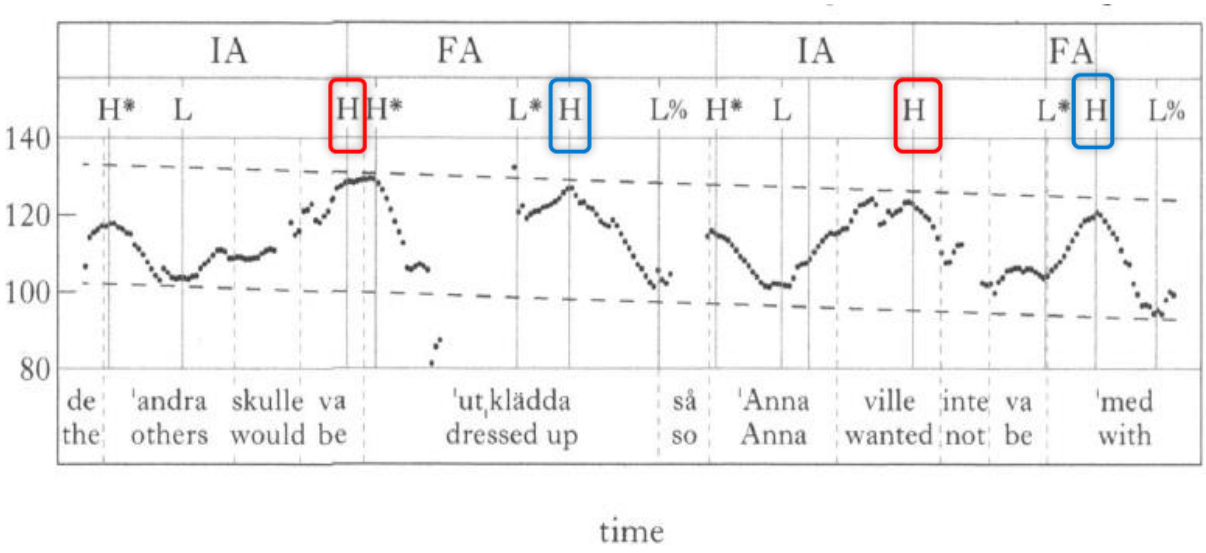


Figure 11: Swedish sentence prosody (taken from Myrberg 2013, p. 99)

{{(De andra skulle va utklädda)$_{PP1}$}$_{IP1}$ så { (Anna ville inte va med.)$_{PP2}$}$_{IP2}$ }$_{IP3}$

(1) De andra skulle va utklädda så Anna ville inte va med

(2) The others would be dressed up so Anna wanted not be with

Figure 11 is taken from Myrberg (2013, p.99) and illustrates Swedish sentence prosody. The sentence forms two prosodic phrases that are marked as PP1 and PP2. The boundaries of the prosodic phrases are marked with ( ). Two coordinated clauses form two intonation phrases which are marked as IP1 and IP2 with { }. These two intonation phrases are contained in a larger intonation phrase as marked with { } $_{IP3}$. IA means initiality accent, and it can be seen as H tone at *va* and *ville*, which are marked with red squares in Figure 11. FA stands for focal accent marked

with blue squares. Focal accent (big accent) is observed at two points; *utkädda* (H*L**H**, bold letter is focal accent) and *med* (L***H**). As many researchers agree that Swedish has tones at the intonational phrase level (Bruce 1977; Gussenhoven 2004), there are boundary tones L% at the end of each intonation phrase. Similarities and differences in intonation between Swedish and Japanese are summarised in Table 4 on page 16.

2.3 Second language phonology

This section reviews the previous studies on theories of negative and positive transfer, and the two influential theories to predict learners' difficulty in second language acquisition: Contrastive Analysis by Lado (1957) and Markedness Differential Hypothesis (Eckman 1977; 1985).

It is common to observe a foreign accent or non-native like pronunciation among adult foreign language learners. Their errors represent the underlying system termed as *Interlanguage* introduced by Selinker (1972). *Interlanguage* consists of the successes and failures of second language acquisition as well as negative and positive transfers from L1 and language universals. Many non-native aspects of *Interlanguage* are caused by negative transfer from learners' native language (Major 2001). For example, Japanese learners of English tend to insert vowels in English words with consonant clusters. Another type of transfer is called a positive transfer or a free ride. Positive transfer occurs when the target languages and the learners' native language share the same linguistic features. For instance, Swedish learners of English do not need to learn a sentence order SVO because Swedish employs the same sentence order as English. Another case of positive transfer is simply when learners successfully acquire linguistic features of target languages. Transfers can occur at any linguistic level, for example, at a phonetic level as in the first example of Japanese learners of English, and at the syntactic level as in the second example of Swedish learners of English. In addition to errors caused by negative transfer, some errors are produced by learners with different native languages. Not only among learners with various native languages, children who are learning the same language as their native language also tend to make the same mistakes. These errors are resulted due to the universals of language acquisition (Major 2001).

Previous studies on transfer have attempted to identify the necessary conditions to trigger transfers (Major 2008). Contrastive Analysis Hypothesis (CAH) proposed by Lado (1957) accounts for what causes learners' difficulties based on the similarities and differences between target languages and learners' native languages. According to CAH, similarities between target

languages and learners' native language predict the lack of difficulties in acquisition. On the other hand, differences between target languages and learners' native languages predict the difficulties in the acquisition. In this study, CAH will be applied to the intonation of Swedish learners of Japanese.

Before applying CAH to the case of Swedish learners of Japanese, Table 4 summarises the features of Swedish and Japanese with a focus on pitch accent.

Table 4: The features of Swedish and Japanese intonation

	Swedish	Japanese
lexical pitch accent	○	○
lexical pitch accent patterns	Both pitch and stress are employed Accent 1, Big accent: L*H Accent 1, Small accent: HL* Accent 2, Big accent: H*L**H** Accent2, Small accent: H*L Accent 2 (compound), Big: H*L***H** Accent 2 (compound), small: H*L Two levels of prominence named a big accent and small accent All words have pitch accent	No stress system (%LH-) H*+L for words with the pitch accent, if a first mora has a pitch accent, %L tone is hardly perceivable. If the second mora of a word has a pitch accent, a phrasal high (H-) is hardly perceivable. %L H- for unaccented words The location of H*+L depends on each lexical word The existence of unaccented words
Prosodic marking of phrase boundaries	Initiality accent (H tone) marks the beginning of intonation phrase Bold and underlined H tone above (focal accent termed in Myrberg 2013) appear on the last word of prosodic phrase it is in	%L and L% mark the beginning and end of Accentual Phrase and Intonation Phrase, respectively the location of %L and L% phrasal tones is fixed no matter the conditions of pitch accent of lexical items in the accentual phrase

CAH predicts that acquiring lexical pitch accent, in general, would be easier for Swedish learners of Japanese than learners whose L1 does not employ lexical pitch accent. Despite this, closer observation, as in Table 4 shows that Swedish and Japanese use pitch accent and stress in different ways, which predicts the difficulties in acquiring Japanese for Swedish learners of Japanese.

Since Swedish uses stress to mark prominence on words, and Japanese does not, it would be predicted that Swedish learners of Japanese might assign stress to Japanese words as a negative transfer from their L1. In Swedish, a penultimate syllable in a word obtains stress as a default (Frid 2001, Riad 2013, p.133). Therefore, CAH predicts that Swedish learners of Japanese might assign stress on a penultimate syllable of Japanese words, and associate a rise in F0 where it is not supposed to have. All Swedish lexical words have a lexical pitch accent. In contrast, Japanese has words that have no lexical pitch accent. This difference predicts the difficulty of acquiring Japanese unaccented words for Swedish learners of Japanese based on CAH. In Swedish, since accent 2 is marked linguistically, only *accent 1* is predicted to appear as Swedish foreign accent (Bruce and Hermans 1999). Contours in *accent 1* are different in terms of the location of pitch peak between Skånska and Rikssvenska which are regional varieties in the participants' mother tongue. Therefore, a negative transfer of *accent 1* in Japanese might be observed differently in terms of which Swedish regional varieties the Swedish participants speak.

As regards the prosodic marking of phrase boundaries, Swedish uses initiality accent and focal accent discussed in Myrberg (2013). Initiality accent marks intonation phrase, and focal accent appears on the last word of prosodic phrase. According to Myrberg (2013), as a contour, initiality accent would be H*LH on an *accent 2* word and L*H on an *accent 1* word. On the other hand, Japanese mainly use L tones to mark the end and beginning of the accentual phrase and intonation phrase. This difference in how the two languages use prosody to mark phrase boundary predicts the difficulty of acquiring Japanese for the learners. As a possible negative transfer, Swedish learners might assign H tone to mark the end of phrase boundaries in Japanese as it is in their native language.

In addition to CAH (Lado 1957), Markedness Differential Hypothesis (MDH) proposed by Eckman (1977; 1985) has contributed to predicting the difficulty that second language learners might face. Eckman argues that CAH fails to explain learners' errors, which occur when learners' L1 and target language do not have linguistic differences. To overcome the CAH's deficiency, Eckman proposed MDH that incorporates the notion of typological markedness in addition to cross-linguistic differences and similarities. Eckman defines markedness as following:

> A phenomenon A in some language is more marked than B if the presence of A in a language implies the presence of B; but the presence of B does not imply the presence of A. (Eckman 1977 p.320)

Based on the definition of markedness, Eckman proposes the following MDH (Eckman 1985, p.291).

> The areas of difficulty that a second language learner will have can be predicted on the basis of a comparison of the NL (native language) and TL (target language) such that:
>
> Those areas of the TL that are different from the NL and are relatively more marked than in the NL will be difficult;
>
> The degree of difficulty associated with those aspects of the TL that are different and more marked than in the NL corresponds to the relative degree of markedness associated with those aspects;
>
> Those areas of the TL that are different than the NL but are not relatively more marked than in the NL will not be difficult.

In the case of Swedish and Japanese prosody, there seems to be no previous studies discussing markedness at the level of sentence prosody and prosodic phrasing. At the word level, Kubozono (1997) discusses the markedness of Japanese pitch accent assignment. In Japanese, when a word has (n) mora, the number of possible patterns in pitch accent assignment is (n+1). For instance, when a word consists of 4 morae, the possible pitch accent patterns would be 5 patterns: pitch accent on the first, second, third, fourth, fifth, and no pitch accent. According to him, words having a pitch accent on its antepenultimate mora is the statistically most frequent accent pattern among accented words. In other words, this pattern is the least marked pattern in terms of pitch accent assignment. Words like *ya'kuza* and *nomi'mono* 'drink' would be examples of unmarked pitch accent patterns. Since Swedish pitch accent system concerns the syllable as a unit, not the mora as in Japanese, it is hard to simply compare and discuss whether the unmarked pitch assignment on the antepenultimate mora in Japanese is different from Swedish. At least it could be argued that the unmarked pitch assignment might be easier to master for Swedish learners of Japanese in comparison to the other pitch accent patterns in Japanese.

In summary, L2 speech produced by second language learners consists of success, failure of language learning, positive and negative transfer from their L1 as well as language universals. CAH by Lado (1957) and MDH by Eckman (1977; 1985) attempt to make predictions of L2 learners' difficulty. CAH predicts based on the differences and similarities between their L1 and target language. In MDH, markedness is also taken into account in addition to cross-linguistic features. In the case of Swedish learners of Japanese, the following difficulties and negative transfer are predicted based on CAH and MDH.

Swedish learners of Japanese are predicted to make errors or have difficulties such as:

- Assigning stress on the penultimate syllable of Japanese words, and associating a rise in F0

-Difficulties of acquiring unaccented words in Japanese

-Assigning H tone to mark phrase boundaries in Japanese when L tone is supposed to be assigned

2.4 Previous studies in L2 Japanese prosodic acquisition

This section is a literature review on L2 Japanese prosody acquisition, and it discusses the findings of the related studies and to identify what has not been studied.

The previous studies dealing with L2 Japanese prosodic acquisition have been conducted from various viewpoints such as word-level production (Sakamoto 2010; Ogasawara 2015), learners' ability to perceive Japanese pitch accent (Wiener and Goss 2019; Sakamoto 2010; Lanz 2003) and sentence level production (Tsurutani 2011). In Sakamoto (2010) study, both the production and perception of English learners of Japanese are examined. According to her findings, English learners of Japanese seem to have the ability to identify the assignment of pitch accent at the word level but show difficulty in assigning pitch accent to proper locations. Her study also reveals that producing unaccented words in Japanese appears to be not hard for learners. Ogasawara (2015) investigates the word level production by speakers of Mandarin Chinese learning Japanese. For Mandarin Chinese learners of Japanese, producing Japanese words with a falling contour that is common in their L1 is easier than rising-falling contour words. In other words, Ogasawara's study shows that L2 words with a similar contour to learner's L1 words cause a lack of difficulty while different contour between learner's L1 and L2 creates difficulty in production. As to perception, Wiener and Goss (2019) examine the learners with different native languages sensitivity toward Japanese pitch accent. The learners are Mandarin Chinese, which employs tones with high informative functions, and English, which has less informative cues. Based on their findings, it can be said that L2 learners with a native language that uses tone for informative cues seem to have an advantage over learners with a native language that does not use tone for the cues in terms of the perception of Japanese pitch accent. While most of the studies working with L2 Japanese learners focus on either speaker of Mandarin Chinese or English, Lanz (2003) analyses the performance in the perception of Japanese pitch accent among speakers of pitch accent languages (Punjabi, Serbo-Croatian, and Swedish) and stress-accent languages (English, Russian, and Samoan). The result shows no statistical difference in their performance regardless of the participant's native language.

The study conducted by Tsurutani (2011) is the most relevant to this study because she examines the production of L2 Japanese learners both at word and sentence level. She assumes

that learner's errors in acquiring Japanese pitch accent are the combination of errors at word and sentence level. In Tsurutani's study, 19 Australian English learners of Japanese who studied Japanese for 200 hours in class were asked to read prepared example sentences aloud for recording. Their recording was extracted from a computer programme where students can do self-assessment in terms of pronunciation. The participants were allowed to try recordings until they were satisfied with their production. Their final productions were considered in the analysis. The learners' speech is graded on a scale of 1to 100 scales by experienced Japanese language teachers for the overall performance and was compared to a modelled speech of ten native speakers. Five Japanese sentences for the recording were prepared. 19 out of 27 words in the sentences have a flat pitch contour. A flat contour is hard to pronounce for English learners of Japanese based on Tsurutani's previous study (Tsurutani 2008). Figure 11 shows the sentence stimuli.

1	*Shachoo-no kekkonshiki-ni okyakusan-ga sennin kita*. (24) LHH-H LLHHLL-L / LHHHH-H /HLLL HL President's wedding- to/ guests / one thousand came One thousand guests came to the president's wedding reception.
2	*Obasan-no yoosu-o byooin-de kiitekudasai*. (23) LHHH-H HHH-H / HHHH-H / HHHHHHL Aunt's condition / the hospital-at / please ask. Please check how my aunt is at the hospital.
3	*Watashi-no kookoo-de isshoni shashin-o torimashoo*. (23) LHH-H HHHH-H / LLHH LHH-H LHHHL My high school-at /together photos let's take. Let's take a photo together at my high school.
4	*Shuumatsu-kara futari-no hito-to shigoto-o suruyoteedesu*.(24) HHHH-HH / LHH-L LH-L / LHH-H LH HHH HL Weekend-from / two people-with / work (I) plan to do. From this weekend I am planning to work with two people.
5	*Yuubinkyoku-de hyakuen-no kitte-o gomai katta*.(23) HHHLLL-L/ LHHH-H LHH-H / LHH LLH The post office-at / 100 yen stamps / five(pieces) bought. I bought five 100-yen stamps at the post office.

Note: H= high pitch mora, L= low pitched mora
"/" indicates accentual phrase boundary. ( ) = length of sentence by the number of mora

Figure 12: Stimuli for Tsurutani study taken from Tsurutani (2011, p.8)

For the gradings by the experienced Japanese language teachers, she notes that most of the participants who received high grades did not produce the same contour to a native speaker's model speech. In her account, it is because other factors besides intonation of the learners' speech are considered when the language teachers judged their production. In my opinion, her account would be one of the reasons to explain that the learners'speech with high grades do not correspond to the native speaker's speech, but not the only reason behind it. In her study, the fact that native speakers produce various contours for the same sentence is not taken into account. Indeed, the participant's production is compared to the model speech by native speakers without discussing the varieties among native speakers. In other words, how much learners' production corresponds to that of the native speakers' speech might not determine native-like speech since there are varieties of contour among native speakers.

She describes the patterns of the learners' mistakes at the level of the prosodic word, accentual phrase (AP), and intonational phrase (IP). According to her results, at the prosodic word level, one of the frequent incorrect errors was assigning a pitch accent on unaccented words or putting a pitch accent on the wrong location in words. The learners tend to put HL contour patterns into words that are not supposed to have a contour pattern. In English, all of the lexical words have stress that leads to a rise in F0, and the HL pattern is a common contour in English (Tsurutani 2007, 2008). In addition to that, the LH contour of unaccented words is cross-linguistically marked. Therefore, she argues that the learners' error of assigning HL contour to non-HL contour words can be explained as a negative transfer and linguistic markedness. In addition to the incorrect assignment of HL contours, she also reports that the learners have a strong tendency to have high pitch and to lengthen at the end of prosodic words. Figure 12 is extracted from the Tsurutani study (2011, p.10). It shows one of the learners in her study where the speaker produces the pitch rise at the end of two prosodic words (indicated with ↓ in the figure) and lengthening. P in Figure 13 stands for a pause, and below the letter, the length of lengthening in milliseconds is presented. The parts of lengthening are underlined.

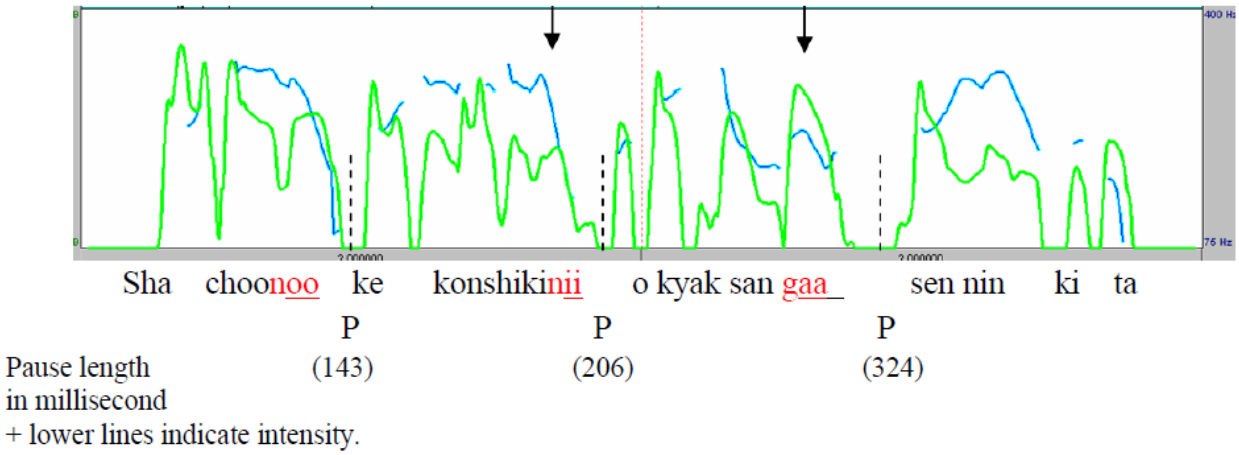


Figure 13: An example of a pitch rise and lengthening at the end of prosodic words (Tsurutani 2011, p. 11)

She accounted that learners' errors of lengthening and a rise in pitch at the prosodic words are due to a negative transfer of their mother tongue. In English, syllable-timed final lengthening is more commonly observed than in Japanese. Lengthening also signals continuation in an utterance. Another possible explanation is that in syllable-timed languages, including English, the location of a stressed syllable can be altered depending on rhythmic factors and sentence prosody. In syllable-timed languages, the interval between each stressed syllable is constant but can differ for each sentence. If the learners attempt to assign stress to Japanese sentences as a negative transfer from their mother tongue, which is a syllable-timed language, it eventually causes a rise in pitch with the same interval. Sentence contour in Japanese primarily consists of lexical pitch accent of words whose locations are fixed and lexically determined. At the prosodic word level, L% and L% tones mark the beginning and end of AP. In sum, if the learners applied the strategy of their syllable-timed languages, it would cause the errors of assigning a rise at the end of prosodic words at where they are supposed to assign low tones. Besides her account that lengthening and a rise in pitch is due to a negative transfer, this can be explained that the learners assigned BPMs. Assigning BPMs at the end of prosodic words is not necessarily defined as an error because Japanese native speakers also assign BPMs at the end of prosodic words (Hara 1993; Kibe, et al. 2013; Hirachi 2008).

At the intonational phrase level, there are some participants in her study who lack downstep where downstep is predicted to be observed. She provides one example of the lack of downstep due to incorrect insertion of a pause that hinders the effect of downstep. Another observation at this level is that some learners produce monotonous speech. According to her, learners' attempts to produce native-like speech based on their impression results in this

monotonous speech. The reason behind this is because Japanese has unaccented words that are characterised by a plateau shaped contour, while English words always have stress.

Before moving onto the discussion of the aims and research question for this study, here I summarise the findings of the previous studies on L2 Japanese prosodic acquisition. In terms of L2 Japanese prosody, studies were conducted to examine learners' perception of pitch accent (Wiener and Goss 2019; Sakamoto 2010; Lanz 2003) and word level production (Sakamoto 2010; Ogasawara 2015). There is one study investigating the sentence level of learners' production (Tsurutani 2011). In her study, the deviation from native speaker's speech in the speech of Japanese learners are defined as errors. However, some of the 'errors' in her study could be explained in another way. The assumption that all of the deviation from native speakers can be defined as errors seems to narrow the perspective in the analysis of Japanese learners' speech.

Based on the findings of the previous studies, this study aims to add two new aspects of L2 Japanese prosody. Firstly, this study acknowledges that there are varieties even among Japanese native speakers who speak Tokyo Japanese. For that reason, this study does not compare learners' speech to Japanese native speakers' 'model' speech based on the degree of deviation from native speakers. Instead, this study tries to explore how to define learners' errors and describes how the sentence prosody produced by Swedish learners of Japanese and native speakers illustrates syntax and prosodic factors in their speech. Secondly, this study aims to add new insights into Swedish learners of Japanese since there is no previous study working with Swedish learners of Japanese.

Chapter 3 Methodology

This chapter explains the methodology that was employed to answer the research questions. Firstly, it provides information about the participants. Then, it moves onto the stimulus sentences, which the participants were asked to read aloud for recording. Thirdly, the detailed procedure of the recording will be described. Lastly, the analysis methods will be explained.

3.1. Aims and research questions

Both Swedish and Japanese use pitch to express lexical meaning and prosodic phrasing. Because of that, it is easy to get an assumption that for Swedish learners of Japanese mastering pitch accent must be easier than learners with languages that do not employ pitch accent. Yet foreign accent is easily detected in Swedish learners of Japanese. Most of the previous studies conducted with the learner's errors in Japanese pitch at the word level (Sakamoto 2010; Ogasawara 2015). Furthermore, there is no previous study that has studied Swedish learners of Japanese. This master thesis aims at examining the errors produced by Swedish learners of Japanese in terms of both word level and sentence level prosody. Since Swedish and Japanese share some phonological features, studying the errors produced by Swedish learners of Japanese would be able to provide interesting insights into the related areas of study. In Tsurutani's study (2011), which is the most relevant to this study, the learners' speech is analysed and compared with native speakers on the assumption that native speakers produce similar contours. Since this study uses sentence stimuli that could get different contours as a result of interaction between the syntax-prosody mapping and prosodic well-formedness, there are varieties among the Japanese native speakers. Therefore, this study attempts to study what kind of similarities and differences in how the interaction is depicted in the speech of Swedish learners and Japanese native speakers. To guide the study, two research questions are formulated.

Research question 1: How is the interaction between syntax-prosody mapping and prosodic well-formedness realised in the utterances produced by Swedish learners of Japanese in comparison to Japanese native speakers?

Research question 2: Can their errors be explained as a transfer from their L1, Swedish?

In order to answer the first research question, four Japanese stimulus sentences that may be possibly pronounced with various contours due to the interaction between syntax and prosody were prepared. Both Swedish learners of Japanese and Japanese native speakers were asked to read the stimuli aloud for recording. There were fourteen Swedish students and four Japanese native speakers participating. In addition to the recording, Swedish students were asked to fill out a questionnaire on their language learning experience and language usage. In addition to the questionnaire, Swedish learners were asked about what kind of regional dialects of Swedish they think that they are speaking. A more detailed explanation of the recording and the questionnaire is provided in the following sections.

3.2. Participants

In total, four Japanese native speakers (three females and one male) and fourteen Swedish learners of Japanese (five females and nine males) participated in the experiment. Two of the Japanese participants were exchange students from Japanese universities. One of them was a Master student at Lund University. Another is a postgraduate student from a Japanese university. All of the Swedish students were undergraduate students majoring in Japanese at either Lund University or Stockholm University. The participants were given a cinema ticket as a compensation to join the experiment. The cinema tickets were provided by the programme of MA in Language and Linguistics. Since the stimulus sentences for this experiment contain Kanji and Hiragana characters, all of the Swedish participants were chosen from students who have studied Japanese for two years or longer at university. The detailed characteristics of the participants will be provided in Table 5 and 6 below. Participant ID is assigned to each participant, which indicates speaker's number, gender (f or m), and mother tongue (J or S). For instance, 01f_J means Japanese female speaker 01, and 03m_S means Swedish male speaker 03.

Table 5: Japanese native participants

Participant ID	Gender	Age	places they have lived (prefectures in Japan)
01f_J	female	19	Saitama
02f_J	female	20	Osaka three years in primary school, Thailand four or five years Tokyo about three years
03m_J	male	23	Saitama
04f_J	female	24	Nagano, Ishikawa

Table 6: Swedish learners of Japanese

Participant ID	Gender	How long learn J	Studying abroad experience	University	Swedish dialect
01m_S	male	five years	two months	Lund	Skånska
02m_S	male	three years six months	two months	Lund	Rikssvenska for family communication For other people, Skånska
03m_S	male	seven years three years in high school, 2013-2015, 2017 to now	12 months	Lund	Rikssvenska, Stockholm Swedish
04f_S	female	two years six months	two months	Lund	Skånska
05m_S	male	three years	four months	Stockholm	Rikssvenska
06f_S	female	five years	three weeks	Stockholm	Rikssvenska
07m_S	male	one year six months	-	Stockholm	Rikssvenska Study Swedish as a second language
08m_S	male	one year six months	-	Stockholm	Rikssvenska, Stockholm Swedish
09f_S	female	one year six months	-	Stockholm	Rikssvenska
10m_S	male	one year three months	one year three months	Stockholm	Rikssvenska
11f_S	female	five years three months	three months	Lund	Rikssvenska

12m_S	male	one year six months	-	Lund	Norrland Svenska
13f_S	female	three years	two months	Lund	Rikssvenska and a bit of Skånska
14m_S	male	four years	-	Stockholm	Rikssvenska

*How long they have learned Japanese is the total years or months they have learned.

I asked Japanese language teachers at both Lund and Stockholm university about what instructions the Swedish participants have received in terms of Japanese pitch accent and prosody. Stockholm University has a language laboratory (språkstudion), where the students can practice Japanese pronunciation with computer aids. With the facility, Stockholm students had classes to practice pronunciation including moraic nasal /N/ (撥音), consonant gemination(促音), length contrast in vowels, downstep and minimal pairs. The students also learned these concepts in a classroom setting. The fact that the pitch height of the first mora and second mora in words are different was also introduced to Stockholm University students. They have not received instructions on prosodic phrasing. At Lund University, pitch accent and prosody were introduced with the concepts such as haku, consonant gemination, and minimal pairs. There are not many opportunities to learn pitch accent at the word level, but they had trained pronunciation at the phrasal or sentence level with a focus on phrasing and how to use prosody in order to express emotional states or focus.

3.3 Stimulus sentences

Four stimulus sentences in Japanese were created for recording. All of the words were chosen from Genki textbook 1 (Banno, E. et al. 2011). The textbook is used in both Stockholm and Lund University first-year Japanese classrooms. This ensures that the vocabulary in the stimuli was not above the participants' proficiency level. In addition to that, words that contain voiceless sounds were avoided as much as possible in order to obtain clear pitch contours from the recording.

There are two factors controlled in the stimuli: pitch accent of the first word in the sentence, and the syntactic structure of the sentence. The factor of pitch accent of the first word is marked with [+Acc] for the accented first word *ya'kuza* and [-Acc] for the unaccented word *tomodachi*. The factor of syntactic structure is named with A [*ga*] and B [*no*]. A [*ga*] means that a stimulus sentence contains a syntactic structure, 'SUBJECT *ga* INDIRECT OBJECT *ni*

nomi'mono o urima'shita.' B [*no*] means that a stimulus sentence has 'SUBJECT (modifier *no* modified) *ga nomi'mono o urima'shita.'* To aid readers, the case particle attached to the first word of the sentence is accompanied by the condition.

Each stimulus sentence will be described in details and the prosodic phrasing of each stimulus sentences predicted based on the syntax-prosody mapping will be explained in Table 7 to 11. Expected AP boundary is marked with ( ) and boundary for IP are shown with { }.

All of the words in the A1 [+Acc, ga] sentence have a pitch accent. Therefore, each phrase 'noun + particle (*ga, ni* or *o*)' is predicted to form an AP on its own. Before *ya'kuza ga, one'esan ni* and *nomi'mono o*, IP boundaries are predicted based on the end-based theory by Selkrik (1986,1996) where the left edge of XPs are projected to prosody. The last two words would be in the same IP since they form a verb phrase together. Hence, the effect of downstep is predicted only on the last word *urima'shita* which is underlined in Table 7.

Table 7: A1 stimulus sentence

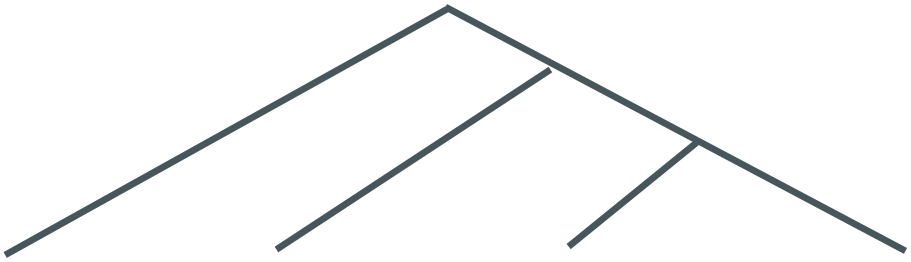
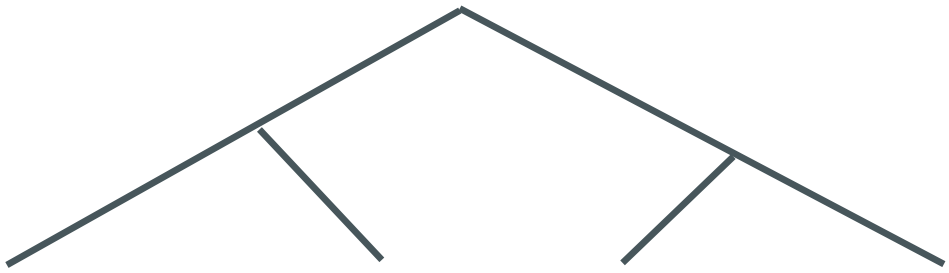
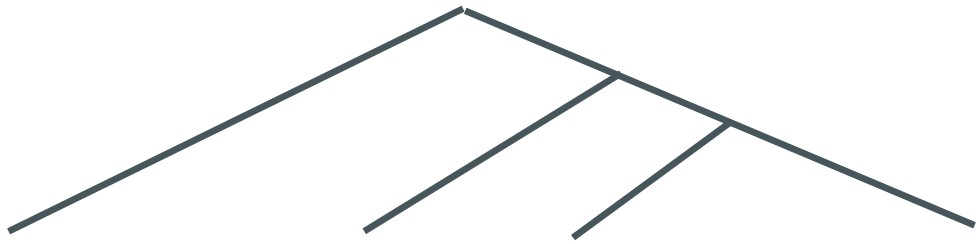
A1 [+Acc, *ga*]	やくざ が お姉さん に 飲み物 を 売りました
	[Ya'kuza ga [one'esan ni [nomi'mono o urima'shi-ta]]]
	Yakuza NOM sister LOC drink ACC sell-PAST
	'A yakuza sold sister a drink.'
 {$_{IP1}$ ($_{AP1}$ ya'kuza ga)} {$_{IP2}$ ($_{AP2}$ one'esan ni)} {$_{IP3}$ ($_{AP3}$ nomi'mono o) ($_{AP4}$ urima'shita)}	

Table 8: A2 stimulus sentence

A2 is [+Acc, *no*]	やくざ の お姉さん が 飲み物 を 売りました
	[Ya'kuza no one'esan ga] [nomi'mono o urima'shi-ta]]
	Yakuza GEN sister NOM drink ACC sell-PAST
	'A yakuza sister sold a drink.'
 {$_{IP1}$($_{AP1}$ya'kuza no) ($_{AP2}$ one'esan ga)} {$_{IP2}$($_{AP3}$ nomi'mono o) ($_{AP4}$ urima'shita)}	

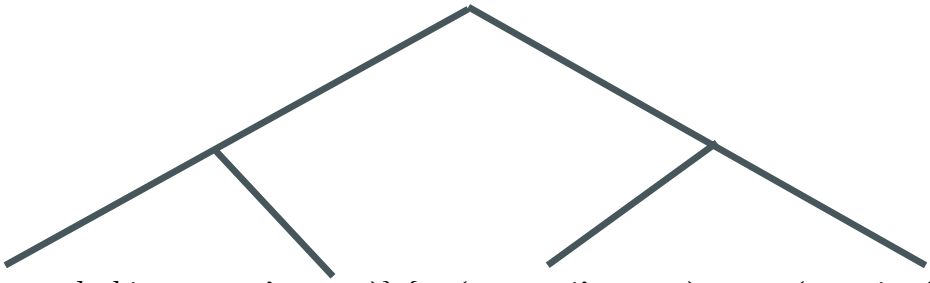
In the A2[+Acc, *no*] stimulus sentence in Table 8, all of the nouns also have a pitch accent. Based on the end-based theory, there would be IP boundaries before *ya'kuza no* and *nomi'mono o* as there are left edges of XPs. The first two words form a noun phrase where *ya'kuza no* modifies the head of the noun phrase, namely *one'esan ga.* The effect of downstep is predicted (underlined in Table 8) on *onee'san ga* since the first two words belong to the same IP. The latter two words consist of a verb phrase where *nomi'mono o* is an object of *urima'shita* 'sold.' The effect of downstep is also predicted on the verb *urima'shita*.

Table 9: B1 stimulus sentence

B1 [-Acc, *ga*]	ともだち が お姉さん に 飲み物 を 売りました
	[Tomodachi ga [one'esan ni [nomi'mono o urima'shi-ta]]]
	Friend NOM sister LOC drink ACC sell-PAST
	'A friend sold a sister a drink.'
 {$_{IP1}$($_{AP1}$tomodachi ga)} {$_{IP2}$($_{AP2}$one'esan ni)} {$_{IP3}$($_{AP2}$nomi'mono o) ($_{AP3}$urima'shita)}	

In Table 9, B1 stimulus sentence has an unaccented word, *tomodachi* 'friend' as the sentence subject. Between the unaccented word and the next word, *one'esan ni*, an IP boundary is predicted as there is a left edge of a noun phrase. The last two words form a verb phrase. Therefore, there would be an IP boundary between *one'esan ni* and *nomi'mono o.*

Table 10: B2 stimulus sentence

B2 [-Acc, *no*]	ともだち の お姉さん が 飲み物 を 売りました
	[Tomodachi no one'esan ga] [nomi'mono o urima'shi-ta]]
	Friend GEN sister NOM drink ACC sell-PAST
	'A sister who is (my) friend sold drink.'
 {$_{IP1}$($_{AP1}$ tomodachi no one'esan ga)} {$_{IP2}$ ($_{AP2}$ nomi'mono o) ($_{AP3}$ urima'shita)}	

In the B2 stimulus sentence, the unaccented word *tomodachi* forms a noun phrase with the head of the noun phrase, *one'esan ga*. The noun phrase is a subject of the sentence. There is a

left edge of a verb phrase that consists of a verb, *urima'shita*, and an object, *nomi'mono o*. There would be an IP boundary between the noun phrase and the verb phrase. Only on the last word *urima'shita* is predicted to get the effect of downstep, and the last two words belong to one IP.

3.4 Procedure

For Japanese native speakers, the details of the experiment were provided in Japanese. For Swedish learners of Japanese, it was given in English. The explanation was conveyed with PowerPoint slides. Firstly, the participants were explained research purpose, which was to analyse the speech of Swedish learners of Japanese. The speech of Japanese native speakers was collected for the comparison with Swedish learners. The instructions also told the participants that all of the collected data would be used only for research purposes only and treated anonymously. The participants were informed that they had the right to withdraw from the study at any time they wanted without providing any reason. After that, the participants could ask questions on the research and were asked to read through the informed consent for joining this study and to sign if they agreed to join. Informed consent for Swedish participants is provided in Appendix A and for Japanese participants is provided in Appendix B. After the participants signed to the informed consent, they were provided with the list of all four stimulus sentences at once. The participants were asked to read them aloud one time each while their speech was recorded to test if there were any problems in recording. This data would not be used for the data analysis.

For recording, a sound recorder Zoom H5 with a head-mounted microphone Shure WH20XLR was used. The recording session was composed of two parts. In each part, the participants were given one stimulus sentence on a PowerPoint slide and asked to read it aloud five times. This was repeated for the four stimulus sentences. In the second part, the four stimulus sentences were shown in a different order from the first part. The participants were then asked to read each sentence five times aloud just like the first part of the recording. Between the recording, for the group of Japanese native speakers, a short questionnaire to ask their name, age, gender, and place of birth was conducted. In addition to this, they were asked whether they have moved within or outside of Japan. The purpose of this was to see whether they had lived in the areas that have dialects with different pitch patterns than Tokyo Japanese. For the Swedish participants, questionnaires were provided to know their name, how long they have been learning Japanese, their mother tongue, their travelling, study abroad experience to Japan, their language usage in daily life, and whether they had a language qualification in Japanese. The questionnaire for the Swedish participants was created by the author based on the questionnaire used for Sakamoto's study (2010). For the Swedish participants, follow-up questions were asked about

their Swedish dialect after the recording. They were asked what Swedish dialect variety they thought that they are speaking. The questionnaires for the Swedish participants were provided in the Appendix B. The recording generates four stimulus sentences for ten times each (five times in the first and five times in the second part of recording) from eighteen participants (four Japanese native speakers and fourteen Swedish learners of Japanese).

3.5 Data Analysis

Sound files of each participant's speech obtained from the recording were annotated with the Praat programme using the Textgrid function. In the annotation, particles and words were annotated separately by inserting boundaries in Textgrid files. Boundaries of each word were evaluated by the author who is a native speaker of Tokyo Japanese.

After the annotation, F0 values were measured using another Praat script. The average of F0 at each measurement point in Table 11 is plotted in a graph per speaker, as exemplified in Figure 14. The vertical axis shows Hertz. Each word has three measurement points (21 points in total for each sentence); the first minimum (appears as 'min1' in Figure 14), maximum and second minimum (appears as 'min2' in Figure 14). maximum and second minimum (appears as 'min2' in Figure 14). These measurement points are marked as positions ('pos01' through 'pos21' in Figure 14). The correspondences of each word and position are presented in Table 11.

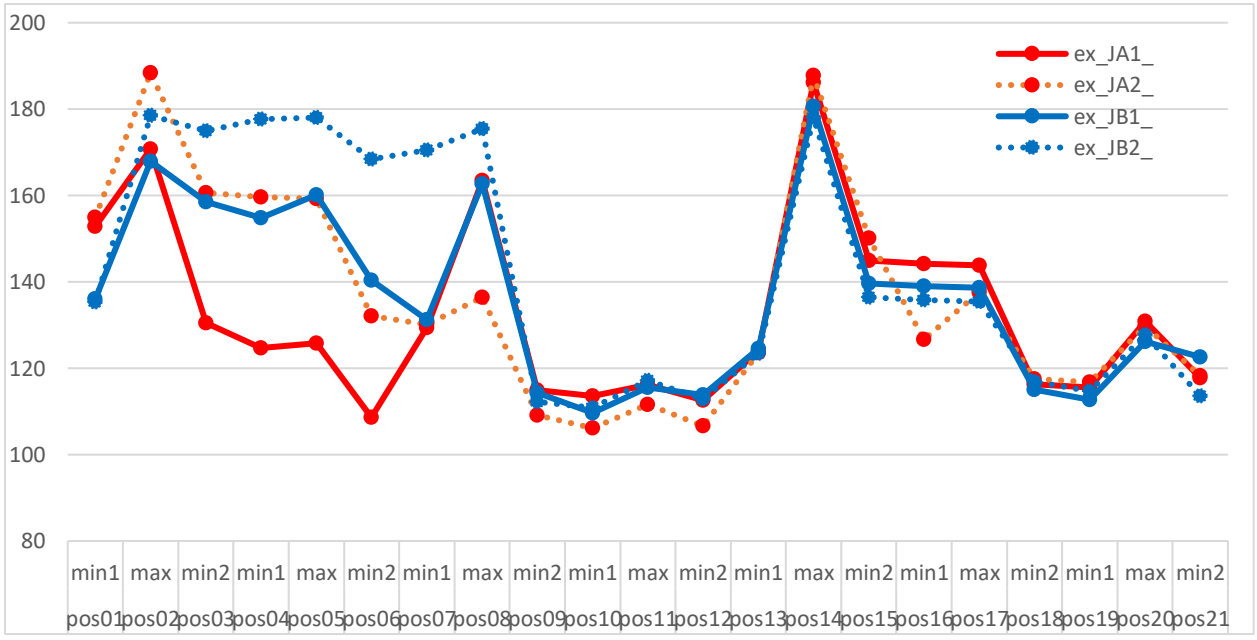


Figure 14: Example of line plots with mean F0 values from 21 measurement points

The solid lines are for the condition of [*ga*], and dotted lines are for the condition of [*no*]. Red lines are for the condition of [+Acc], and blue lines are for the condition of [-Acc].

Table 11: Words corresponding to each position

pos	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
A1	*Ya'kuza*			*ga*			*one'esan*			*ni*			*nomi'mono*			*o*			*urima'shita*		
B1	*Tomodachi*			*ga*			*one'esan*			*ni*			*nomi'mono*			*o*			*urima'shita*		
A2	*Ya'kuza*			*no*			*one'esan*			*ga*			*nomi'mono*			*o*			*urima'shita*		
B2	*Tomodachi*			*no*			*one'esan*			*ga*			*nomi'mono*			*o*			*urima'shita*		

In addition to looking at the average F0 peak at each of the 21 points, actual speech contours were analysed and will be exemplified in the next section on the results as well, since the average of F0 can be skewed if the data have outliers.

The two research questions are the following;

Research question 1: How is the interaction between syntax-prosody mapping and prosodic-well-formedness realised in the utterances produced by Swedish learners of Japanese in comparison to Japanese native speakers?

Research question 2: Can their errors be explained as a transfer from their L1, Swedish?

Based on these research questions, the results are analysed with the following focuses;

-Whether the participants realise pitch accents for each lexical word

-Whether the participants' utterances show downstep where it is expected to occur

-What kind of prosodic phrasing can be assigned by each participant

-Whether other prosodic phenomena, for instance, boundary pitch movement can be observed

Whether speakers produced lexical pitch accent, downstep, or what kind of prosodic phrasing they produced is determined based on the observation of both actual F0 contour, the average of F0 at each of the 21 points and perception by the researcher who is a native Japanese

speaker. Especially, as to downstep, when there is noticeable F0 compression on *one'esan* in A2 [+Acc, *no*] in comparison to *ya'kuza* in A2 [+Acc, *no*], it is described as downstep is produced as predicted. F0 compression is judged based on the difference in F0 height at the highest point within an AP. The predicted downstep to be observed on *urima'shita* in all four stimuli is determined based on the comparison between the highest F0 within each AP *nomi'mono o* and *urima'shita* in each stimulus.

Chapter 4 Results

This chapter presents the results of the experiment. First, a detailed observation of the speech produced by the four Japanese native participants is presented with some examples from their actual speech and the graphs that are generated based on the average F0-values at 21 measurement points. Then, the overall tendency in the speech produced by Swedish learners of Japanese will be discussed, followed by examples.

As a brief overview of the observation, four Japanese participants produced four different contours for the same stimuli. The speech by speaker 03m_J depicts predicted contours based on the syntax-prosody mapping. Speaker 01f_J shows a strong tendency to put BPMs on AP boundaries. Speaker 02f_J divides the four accented words into two IPs containing two words each, due to the rhythmic effect. Lastly, speaker 04f_J produced relatively monotonous speech compared to the other Japanese native speakers. As for the speech produced by Swedish learners of Japanese, their speech is different from Japanese native speakers in terms of the overall contour. How Swedish learners' speech differs from the native speaker also vary among the learners. Frequent errors in lexical pitch accent and the absence of downstep are observed.

4.1 Tokyo Japanese native speakers

The prosodic phrasing produced by speaker 03m_J is overall predicted by syntax-prosody mapping, and there are no peculiar phenomena in terms of prosody. Figures 14 and 15 exemplify pitch contours produced by speaker 03m_J for A1 [+Acc, *ga*] and A2 [+Acc, *no*], respectively.

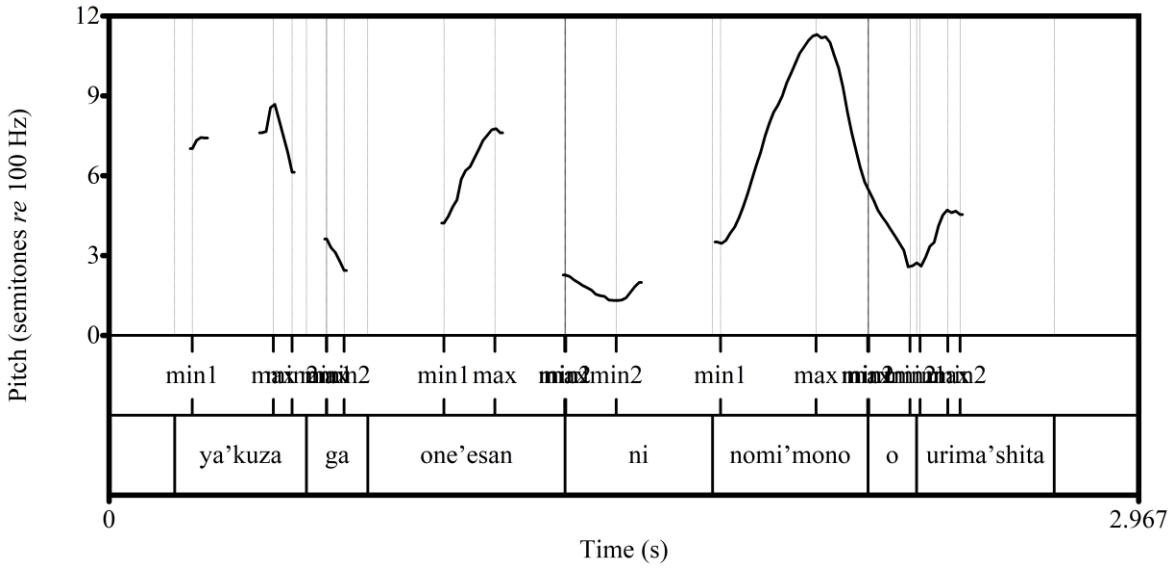


Figure 15: The pitch contours extracted from speaker 03m_J, eighth repetition for A1[+Acc, *ga*]

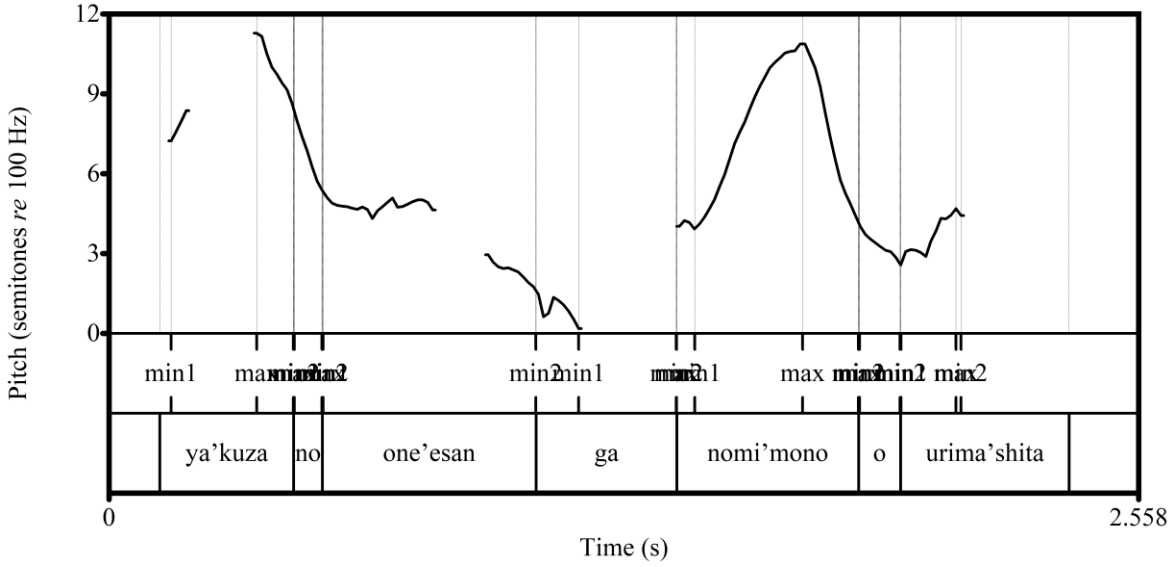


Figure 16: The pitch contours extracted from speaker 03m_J, eighth repetition for A2 [+Acc, *no*]

In Figure 15, an actual pitch contour for A1 [+Acc, ga] by speaker 03m_J is presented. For A1, there are three IPs; the first two words form an IP each and the last two words form one IP by looking at F0 drop for the beginning of each IP boundary. In Figure 16, comparing the F0 peaks of the first two words in A2 [+Acc, *no*], i.e., *ya'kuza* and *one'esan*, the effect of downstep

can be easily detected on *one'esan*, as the F0 peak of *one'esan* is much lower than that of *ya'kuza*. Since the F0 peak of *nomi'mono* both in A1 and A2 is nearly as high as the beginning of utterance, it can be interpreted that new IP boundaries are inserted before *nomi'mono*. This is as expected based on the syntax-prosody mapping. Looking at the last word *urima'shita*, although the contour is hardly observable due to the voiceless segments in the word, it at least shows the F0 compression on the word, which can be interpreted as a result of downstep. In other words, the last two phrases in A1 [+Acc, *ga*] and A2 [+Acc, *no*] are in the same IP as predicted. F0 rises are seen at the beginning of the utterance, between the particle *ga* and *nomi'mono* and between *nomi'mono o* and *urima'shita.* These rises indicate that there are initial lowering at AP boundaries. In Japanese, two accented words cannot form a single AP, but instead, each accented word consists of individual AP. Therefore, based on the observation, prosodic phrasing for A1 [+Acc, *ga*] and A2 [+Acc, *no*] are listed as follows:

A1{$_{IP1}$($_{AP1}$ ya'kuza ga)} {$_{IP2}$ ($_{AP2}$one'esan ni)}{$_{IP3}$($_{AP3}$ nomi'mono o)($_{AP4}$ urima'shita)}

A2 {$_{IP1}$ ($_{AP1}$ ya'kuza no) ($_{AP2}$ one'esan ga)} {$_{IP2}$($_{AP3}$ nomi'mono o)($_{AP4}$ urima'shita)}

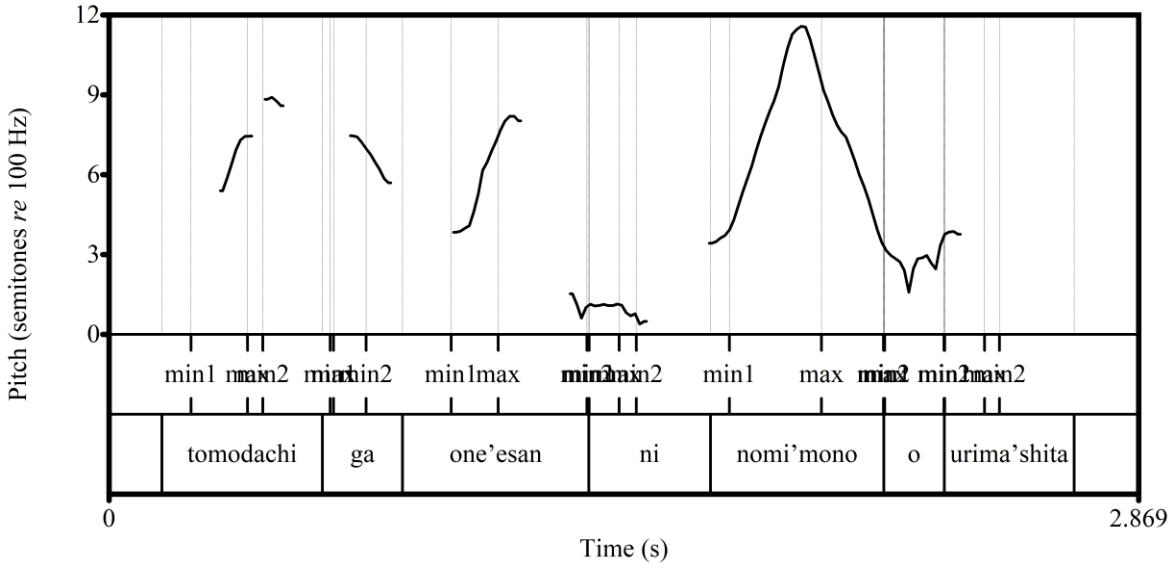


Figure 17: The pitch contours extracted from speaker 03m_J, eighth repetition for B1 [-Acc, *ga*]

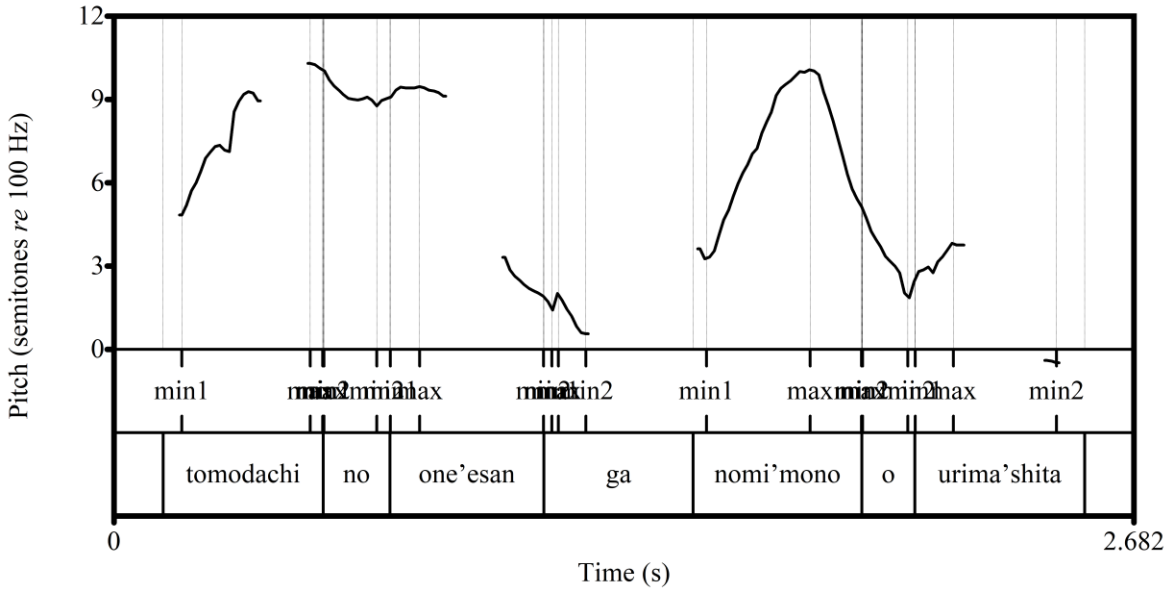


Figure 18: The pitch contours extracted from speaker 03m_J, eighth repetition for B2 [-Acc, *no*]

Figures 17 and 18 show pitch contours extracted from speaker 03m_J, the eighth repetition for B1 [-Acc, *ga*] and B2 [-Acc, *no*]. In B1 and B2, an abrupt fall in F0 cannot be observed for the first word, *tomodachi.* In other words, the speaker 03m_J pronounced *tomodachi* as unaccented as it is supposed to be. Looking at the beginning of the second word *one'esan* in B1, it is preceded by an F0 fall, and then a rise around the beginning of the word. This indicates that there is an initial lowering due to the AP boundary insertion between the particle *ga* and *one'esan*. In addition to AP boundary, there would be IP boundary between the first word *tomodachi ga* and the second word *one'esan ni* in B1 since there is a large fall. In contrast, B2 has a slight fall in F0 around the particle *no* and rises at the beginning of *one'esan.* That is, it can be said that there is an AP boundary. Since there is no large fall in F0 between *tomodachi ga* and *one'esan* unlike, in B1, where there is an IP boundary, it can be said that there is an AP boundary but no IP boundary between *tomodachi no* and *one'esan* in B2.

According to the observations above, the prosodic phrasing for B1 [-Acc, *ga*] and B2 [-Acc, *no*] by speaker03m_J can be listed as follows;

B1{$_{IP1}$($_{AP1}$ tomodachi ga)} {$_{IP2}$($_{AP2}$ one'esan ni)}{$_{IP3}$($_{AP2}$ nomi'mono o)($_{AP3}$ urima'shita)}

B2 {$_{IP1}$($_{AP1}$ tomodachi no) ($_{AP2}$ one'esan ga)} {$_{IP2}$($_{AP3}$ nomi'mono o)($_{AP4}$ urima'shita)}

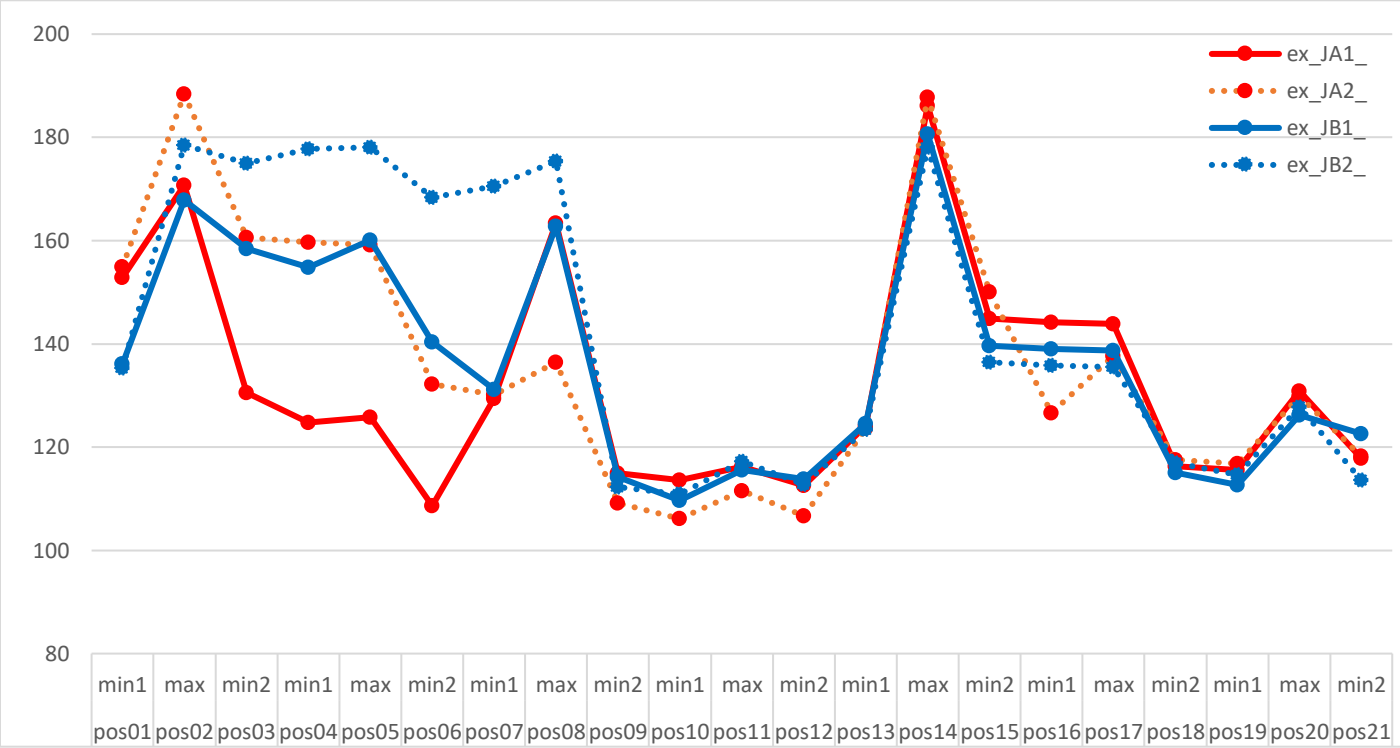


Figure 19: Line plots with mean F0 values from 21 measurement points by speaker 03m_J

Looking at the average of F0-maxima/minima for speaker 03m_J in Figure 19, the speaker produced pitch accent for all of the lexical items and downstep at *one'esan* (pos7-9) in A2 [+Acc, *no*] as predicted. As seen in the actual pitch contours, line plots with mean F0 maxima/minima also show a clear difference in pitch contours for the first words; *ya'kuza* and *tomodachi*. For A1 [+Acc, *ga*] and A2 [+Acc, *no*], which have the accented word *ya'kuza*, there are large F0 falls around position 2-3. This means that speaker 03m_J produced *ya'kuza* as accented. When looking at B1 [-Acc, *ga*] and B2 [-Acc, *no*] with the unaccented word *tomodachi*, there is no large fall around pos2-3, which indicates that speaker 03m_J produced *tomodachi* as unaccented. By comparing the F0 peaks of *one'esan* and *ya'kuza* in A2 [+Acc, *no*], it is evident that *one'esan* gets F0 compression compared to *ya'kuza*, which means that the effect of downstep appears on the part as predicted. As based on the prosodic mapping, IP boundaries seem to be inserted in all four stimuli, between *one'esan* and *nomi'mono*, since a pitch reset occurs and the height for *nomi'mono* is almost as high as the beginning of the utterance. An F0 compression can be seen on the verb *urima'shita* (pos19-21) in comparison to *nomi'mono o*, which shows the effect of downstep on the last word and the last two words are in the same IP as the preceding *nomi'mono* for all of the four stimuli. By comparing B1 [-Acc, *ga*] and B2 [-Acc, *no*], there is a difference in the height of F0 at the particles (*ga* for B1 and *no* for B2). The

particle *no* in B2 [-Acc, *no*] gets a higher F0 peak at the pos 4-6 than that of B1 [-Acc, *ga*]. B2[-Acc, *no*] forms a flat contour between pos2-8, which is predicted to form one IP itself. There is a small dip between *tomodachi no* and *one'esan ga,* which indicates an AP boundary between them. As to B1, at the particle *ga* (pos4-6) of F0 is dropped, which indicates an initial lowering at an AP boundary and IP boundary. An initial lowering in B2 is smaller than in B1, but it can show that there is an AP boundary. Therefore, the first two words in B2 are in two different APs. All of the prosodic phrasing produced by speaker 03m_J are as predicted.

In sum, prosodic phrasing produced by speaker 03m_J can be listed as follows:

A1{$_{IP1}$($_{AP1}$ ya'kuza ga)} {$_{IP2}$($_{AP2}$ one'esan ni)}{$_{IP3}$($_{AP3}$ nomi'mono o)($_{AP4}$ urima'shita)}

A2 {$_{IP1}$($_{AP1}$ ya'kuza no) ($_{AP2}$ one'esan ga)} {$_{IP2}$($_{AP3}$ nomi'mono o)($_{AP4}$ urima'shita)}

B1{$_{IP1}$($_{AP1}$ tomodachi ga)} {$_{IP2}$($_{AP2}$ one'esan ni)}{$_{IP3}$($_{AP2}$ nomi'mono o)($_{AP3}$ urima'shita)}

B2 {$_{IP1}$($_{AP1}$tomodachi no) ($_{AP2}$one'esan ga)} {$_{IP2}$($_{AP3}$nomi'mono o)($_{AP4}$urima'shita)}

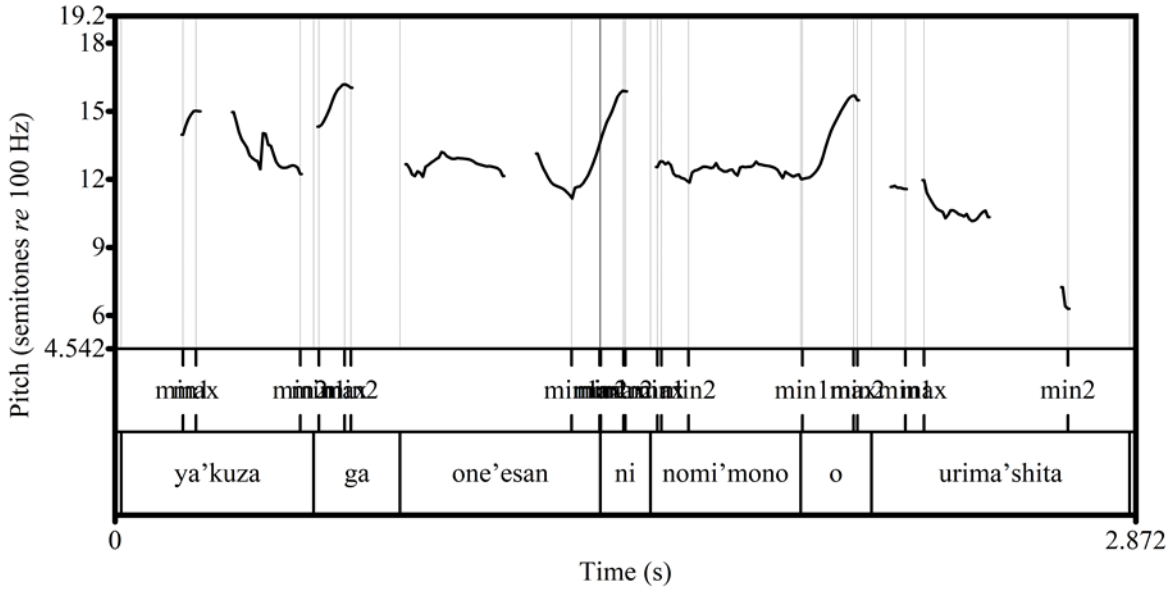


Figure 20: Pitch contours extracted from speaker 01f_J, eighth repetition for A1 [+Acc, *ga*]

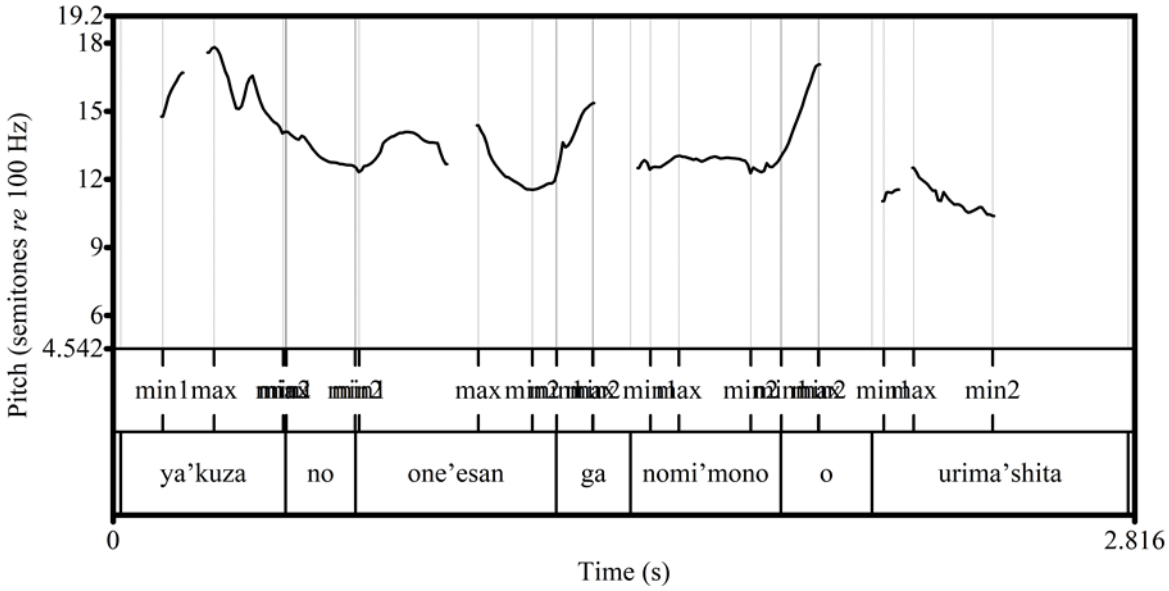


Figure 21: Pitch contours extracted from speaker 01f_J, eighth repetition for A2 [+Acc, *no*]

Now we move onto the next speaker 01f_J. The overall characteristics of the speech produced by speaker 01f_J are the strong tendency of assigning BPMs on AP boundaries. Figure 20 is an actual pitch contour produced by speaker 01f_J, the eighth repetition for A1 [+Acc, *ga*]. A large fall in F0 can be observed for the first word *ya'kuza*. In other words, speaker 01f_J produced a lexical pitch accent for the first word. In A1 [+Acc, *ga*], the nominative case particle *ga*, the locative case particle *ni*, and the accusative case particle *o* show a sharp rise due to BPM assignment. However, BPMs are not assigned to the genitive case particle *no*. As Igarashi (2015) assumes, the sentence medial BPM occurs at AP as a domain. There are AP boundaries between *ya'kuza ga* and *one'esan ni*, between *one'esan ni* and *nomi'mono*, and between *nomi'mono o* and *urima'shita.* As to downstep predicted on *one'esan* in A2 [+Acc, *no*], there is a slight F0 compression in a comparison between *ya'kuza no* and *one'esan ga*. This shows that there is the effect of downstep on *one'esan* in A2, as predicted. Therefore, the first two phrases are in the same IP. About the last two phrases in both A1 and A2, the verb *uri'mashita* seems to get slight F0 compression in comparison to the previous word *nomi'mono o*. Based on that, it would be reasonable to assume that the last two phrases in A1 and A2 are in the same IP. Prosodic phrasing for A1 and A2 produced by speaker 01f_J can be listed as follows:

A1{$_{IP1}$($_{AP1}$ ya'kuza ga)} {$_{IP2}$($_{AP2}$ one'esan ni)}{$_{IP3}$($_{AP3}$ nomi'mono o) ($_{AP4}$ urima'shita)}

A2 {IP1(AP1 ya'kuza no)(AP2 one'esan ga)} {IP2(AP3 nomi'mono o) (AP4 urima'shita)}

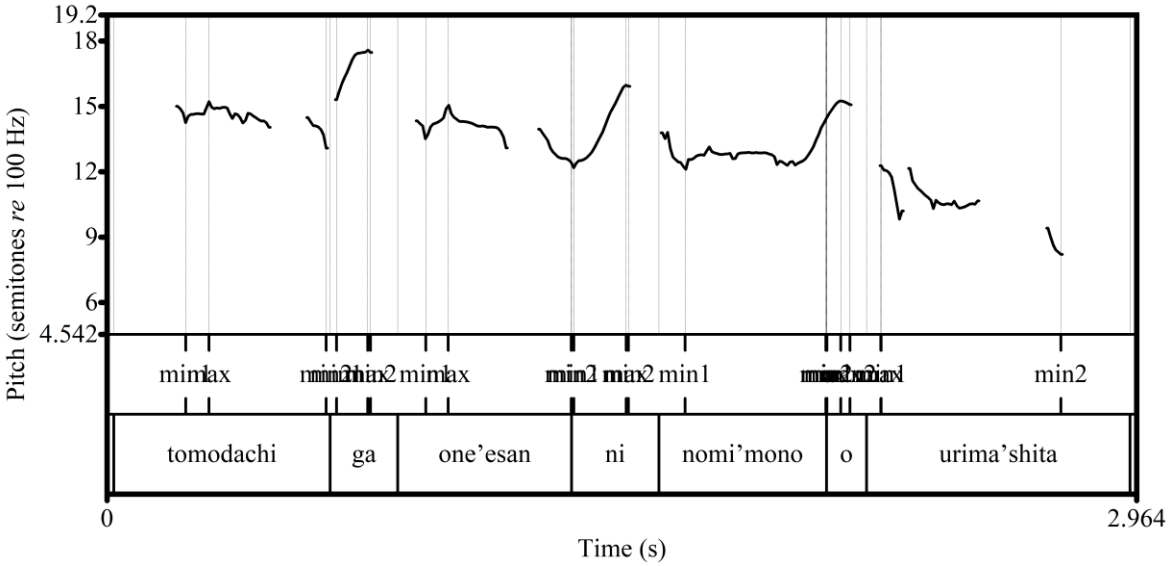


Figure 22: Pitch contours extracted from speaker 01f_J, eighth repetition for B1 [-Acc, *ga*]

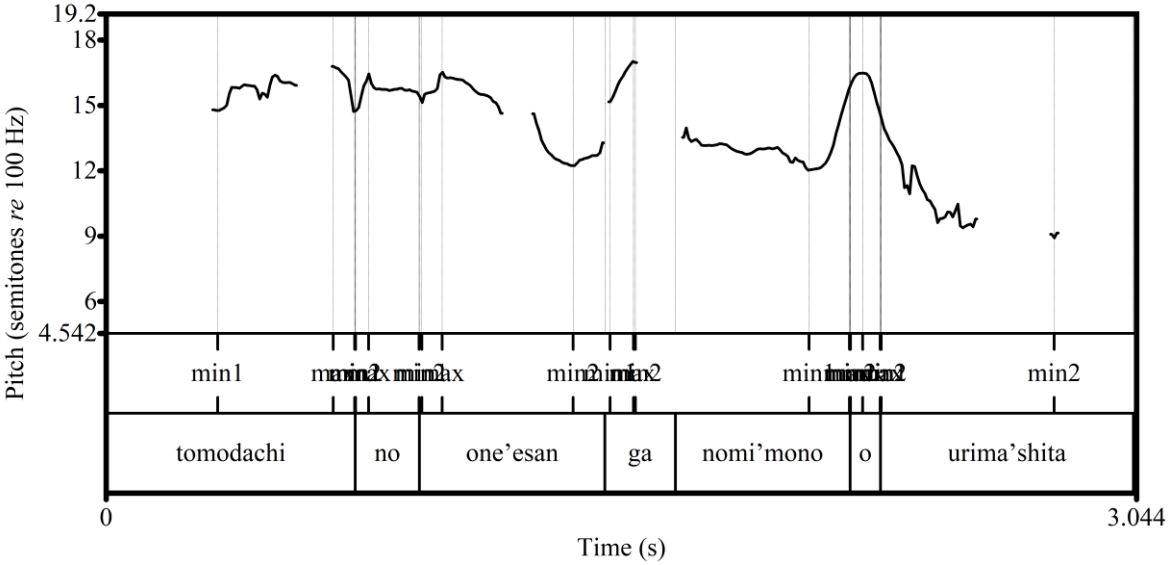


Figure 23: Pitch contours extracted from speaker 01f_J, eighth repetition for B2 [-Acc, *no*]

The lack of an abrupt fall in F0 on the first word *tomodachi* is evident in Figure 22 and 23. Speaker 01f_J produced *tomodachi* as unaccented, as it is supposed to be. Looking at B1[-

Acc, *ga*] and B2 [-Acc, *no*] produced by speaker 01f_J, there seems to be no initial lowering between *tomodachi ga* and *one'esan ni* in B1, and the lowering is also absent in B2. In both B1 and B2, the F0 range on *urima'shita* is compressed. Therefore, it can be assumed that *nomi'mono o* and *urima'shita* are in the same IP both in B1 and B2. Similar to A1 and A2, speaker 01f_J assigned BPM on the particles except for the genitive case *no* in B2 (pos 4-6), which means that there is no AP boundary between the particle *no* and *one'esan ga* in B2. For B1 [-Acc, *ga*] and B2 [-Acc, *no*] by speaker 01f_J, prosodic phrasing would be:

B1{$_{IP1}$($_{AP1}$tomodachi ga)}{$_{IP2}$($_{AP2}$ one'esan ni)}{$_{IP3}$($_{AP3}$ nomi'mono o)($_{AP4}$ urima'shita)}

B2 {$_{IP1}$($_{AP1}$tomodachi no)($_{AP2}$ one'esan ga)} {$_{IP2}$($_{AP3}$ nomi'mono o) ($_{AP4}$ urima'shita)}

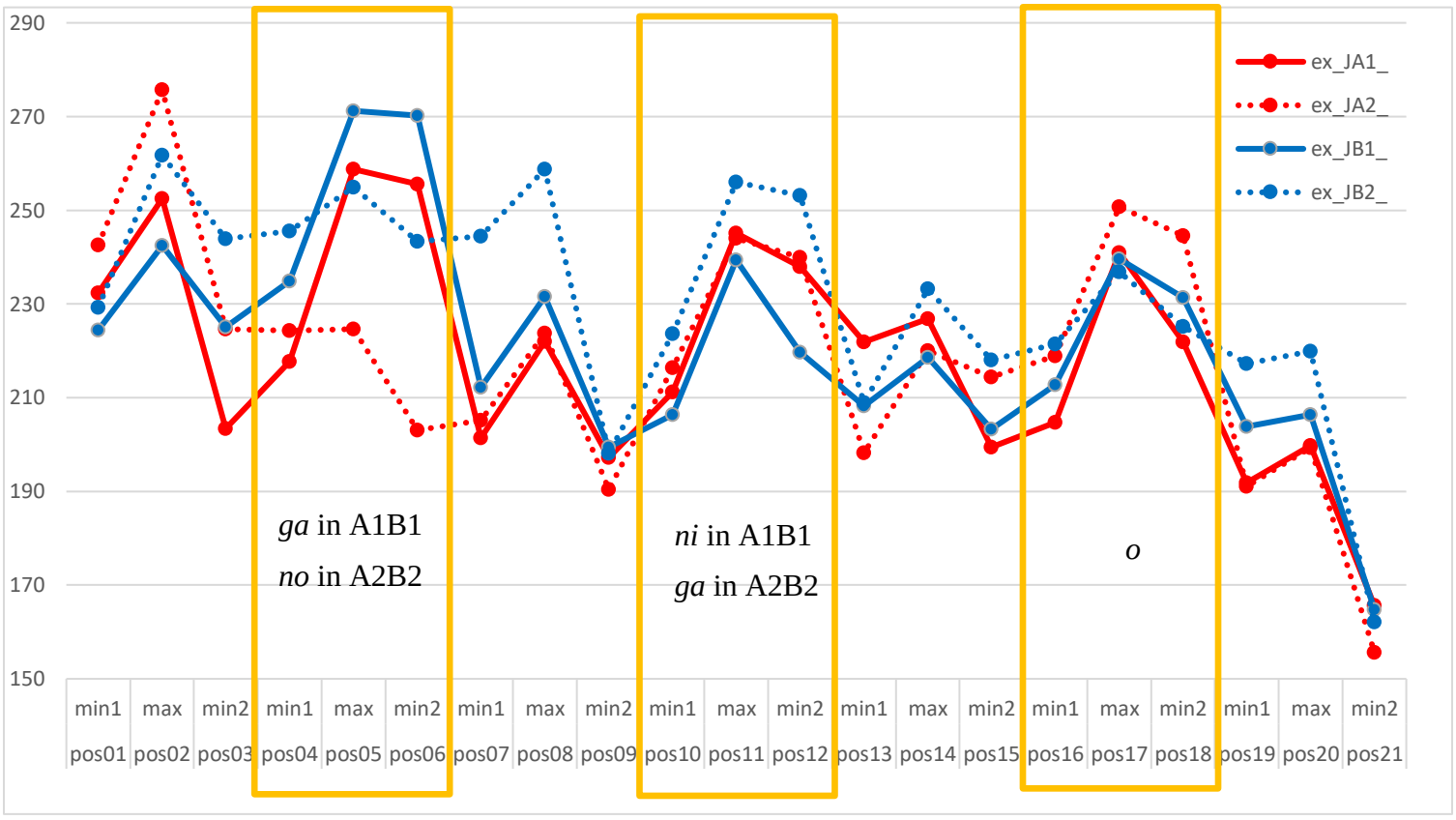


Figure 24: Line plots with mean F0 values from 21 measurement points by speaker 01f_J

Figure 24 illustrates the average of F0 for speaker 01f_J and shows the striking feature of speaker 01f_J, clear BPMs on the particles. As for the actual pitch contours in Figure 20-23, F0 compression is also observed at pos 7-9 *one'esan* in A2 [+Acc, *no*] by comparing the F0 peak to that of *ya'kuza* in A2. The comparison of the F0 peaks at *one'esan* in A2 and B2 can also prove

the effect of downstep in A2. In the speech of Speaker 01f_J, boundary pitch movement (BPM) are observed at particles; *one'esan-ni/ga* at pos10–12 in all four conditions and *o* at pos16-18 in all four examples. BPM cannot be seen on *no* at pos4-6 in A2 and B2, which are marked with orange squares. Looking at the prosodic phrasing for A1 [+Acc, *ga*], there are falls in F0 due to the initial lowering right before the beginning of *one'esan* (pos 7-9) and *nomi'mono* (pos13-15). Since there is F0 compression on *urima'shita* comparing to the F0 peak of *nomi'mono o* in all four stimuli, the last two phrases are assumed to be in the same IP.

In sum, speaker 01f_J has a strong tendency of assigning BPMs on AP boundaries, which is unique among the Japanese participants. As to the prosodic phrasing, speaker 01f_J produced as predicted based on the syntax-prosody mapping. The prosodic phrasing produced by speaker 01f_J can be represented as follows:

A1{$_{IP1}$($_{AP1}$ ya'kuza ga)} {$_{IP2}$ ($_{AP2}$ one'esan ni)}{$_{IP3}$($_{AP3}$ nomi'mono o) ($_{AP4}$ urima'shita)}

A2 {$_{IP1}$($_{AP1}$ ya'kuza no)($_{AP2}$ one'esan ga)}{$_{IP2}$($_{AP3}$ nomi'mono o)($_{AP4}$ urima'shita)}

B1{$_{IP1}$($_{AP1}$ tomodachi ga)}{$_{IP2}$($_{AP2}$ one'esan ni)($_{AP3}$ nomi'mono o)($_{AP4}$ urima'shita)}

B2 {$_{IP1}$($_{AP1}$ tomodachi no)($_{AP2}$ one'esan ga)}{$_{IP2}$($_{AP3}$ nomi'mono o)($_{AP4}$ urima'shita)}

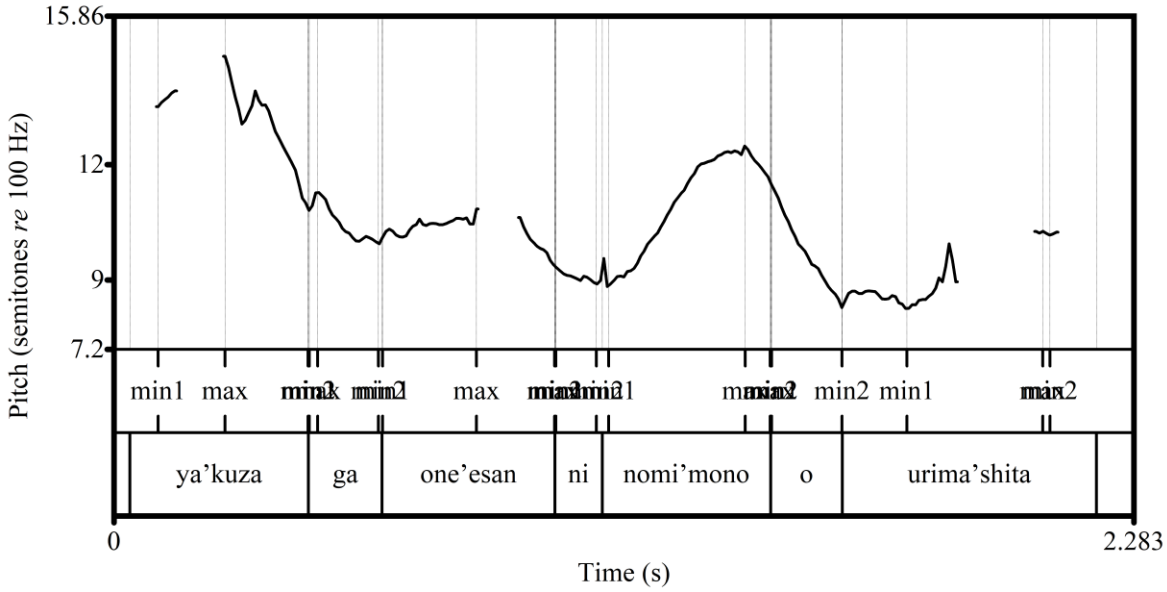


Figure 25: Pitch contours extracted from speaker 02f_J, eighth repetition for A1 [+Acc, *ga*]

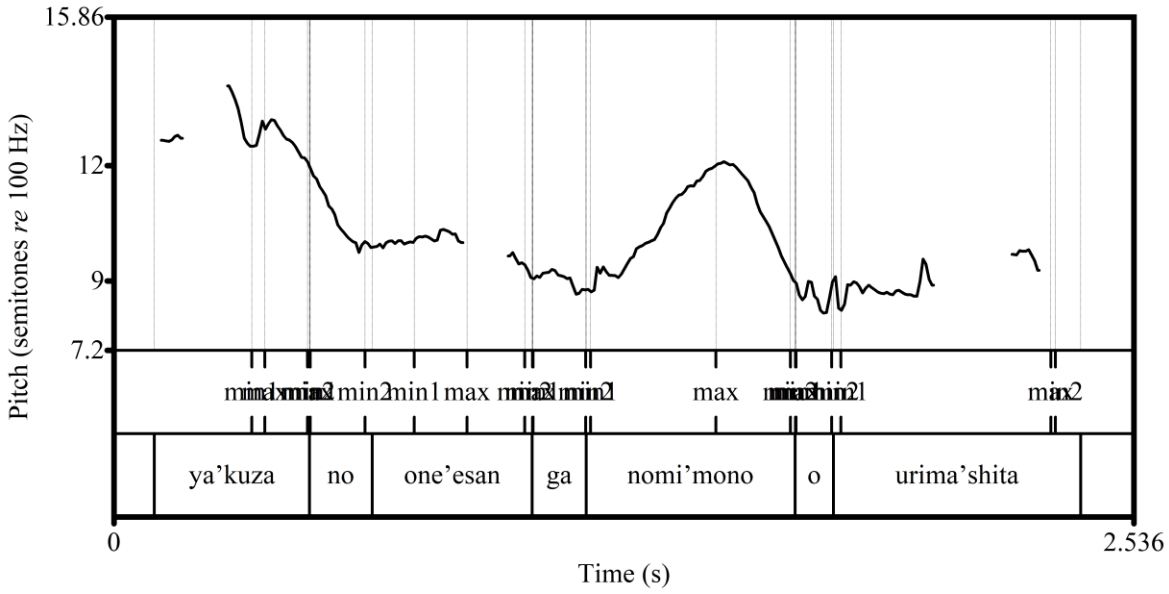


Figure 26: Pitch contours extracted from speaker 02f_J, eighth repetition for A2 [+Acc, *no*]

Moving onto the next speaker 02f_J, Figure 25 and 26 present actual pitch contours from speaker 02f_J, eighth repetition for A1 [+Acc, *ga*] and A2 [+Acc, *no*]. What stands out in the speech of speaker 02f_J is that downstep is unexpectedly found in A1, even though there is an XP boundary at the second word. In other words, the syntax-prosody mapping seems to be overridden by rhythmic effect.

Comparing these two figures, it is evident that the second lexical word *one'esan* gets F0 compression in both of the utterances showing the effect of downstep. Downstep in A2 on that part is predicted by the syntax-prosody mapping, while downstep in A1 is not predicted based on the syntax-prosody mapping. However, downstep in A1 can be explained by the absence of expected IP boundary between *ya'kuza ga* and *one'esan ni* due to the rhythmic effect. On the last phrase *urima'shita* in both A1 and A2 show the F0 compression showing the effect of downstep, which indicates that the last two phrases are in the same IP. In other words, speaker 02f_J inserts IP boundary only between *one'esan ni* and *nomi'mono o* in A1 and *one'esan ga* and *nomi'mono o* in A2. Between each phrase in A1 and A2, there are falls in F0 showing initial lowering due to the AP boundary. Thus, it can be said that speaker 02f_J seem to insert AP boundaries where they are predicted.

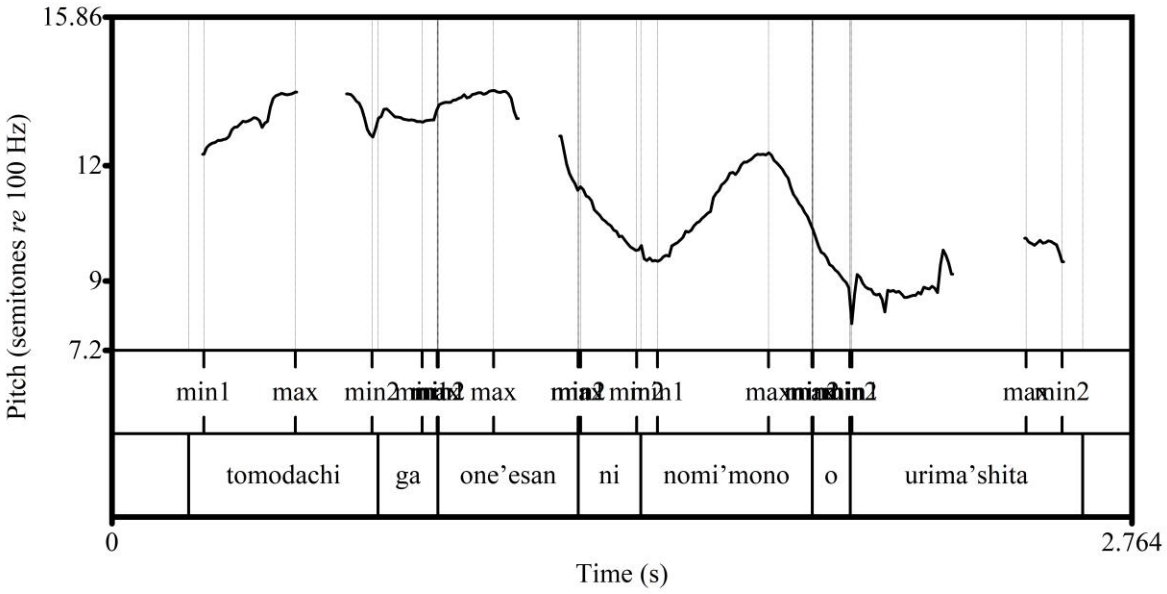


Figure 27: Pitch contours extracted from speaker 02f_J, eighth repetition for B1[-Acc, *ga*]

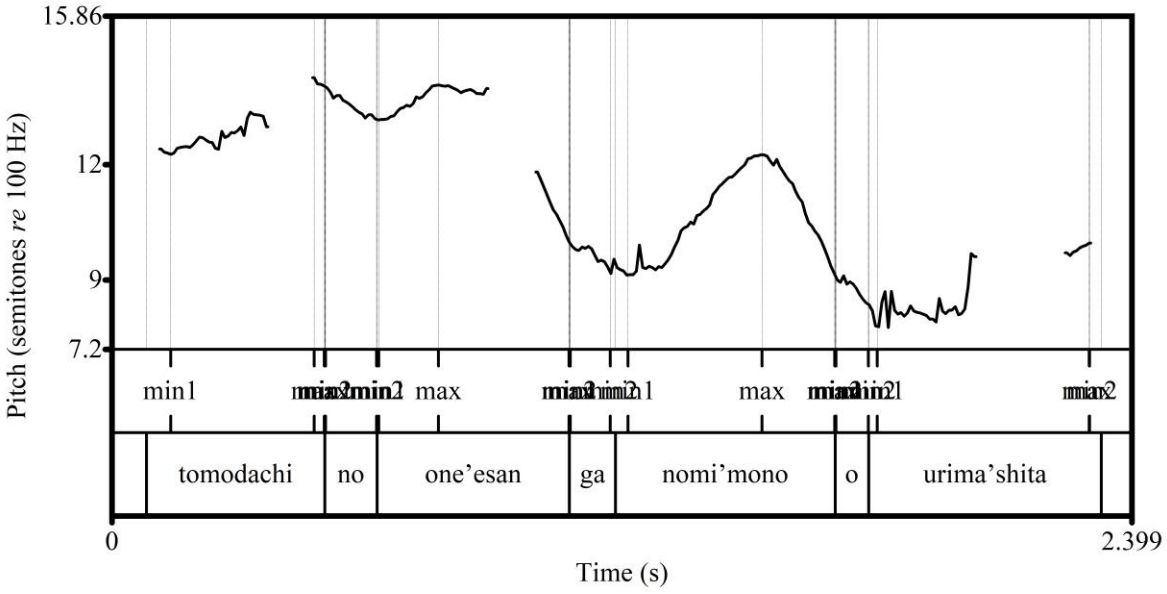


Figure 28: Pitch contours extracted from speaker 02f_J, eighth repetition for B2 [-Acc, *no*]

Actual F0 contours for B1 [-Acc, *ga*] and B2 [-Acc, *no*], eighth repetition produced by speaker 02f_J are presented in Figure 27 and 28. Looking at B1, the end-based theory predicts that there would be an IP boundary between *tomodachi ga* and *one'esan ni*. However, in the speech produced by speaker 02f_J, a large drop between the phrases is missing, which indicates that they consist of one IP. As to an AP boundary between *tomodachi ga* and *one'esan ni*, a slight fall in F0 can be observed. Thus, there seems to be an AP boundary between them. On the last verb *urima'shita* in both B1 and B2, the effect of downstep can be observed through the compression of the pitch range.

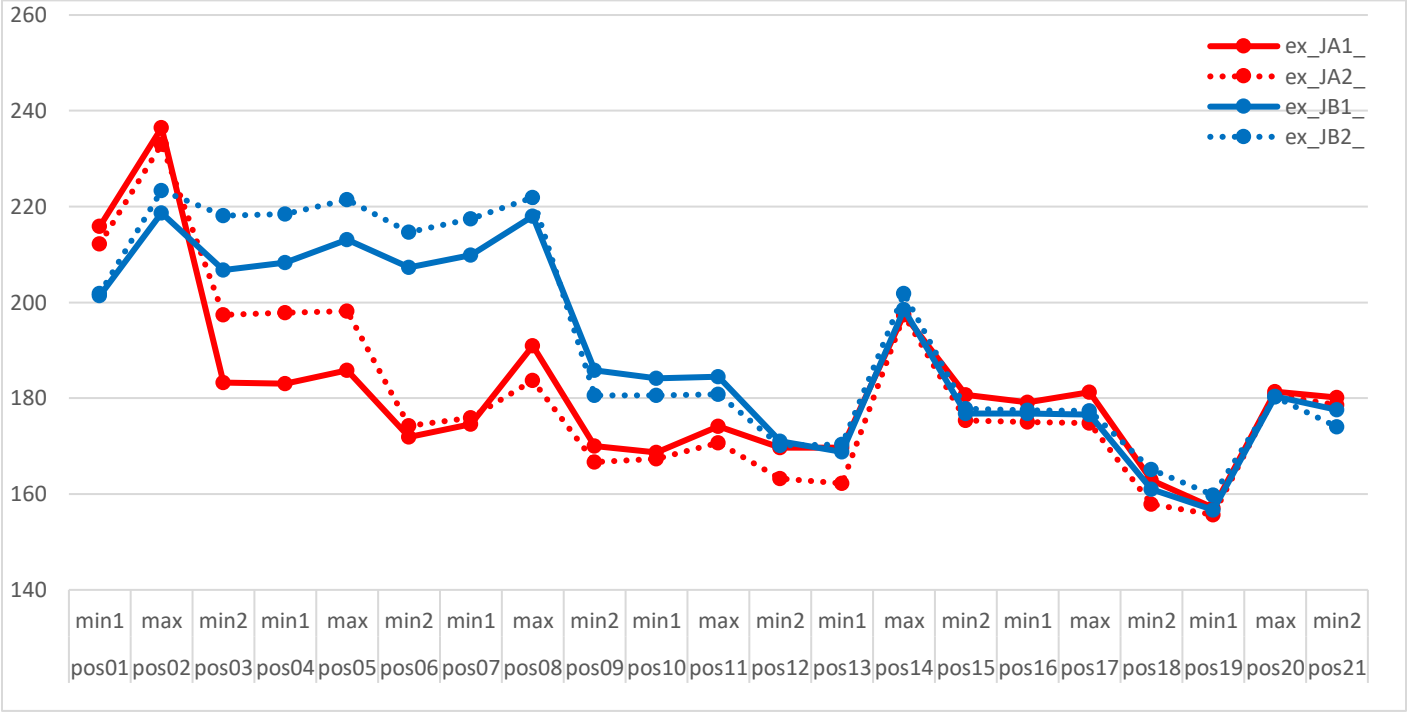


Figure 29: Line plots with mean F0 values from 21 measurement points by speaker 02f_J

Figure 29 presents the average of F0 at 21 measurement points for Speaker 02f_J the speaker. It primarily illustrates the same features in the speech of speaker 02f_J as in the actual pitch contours. Speaker 02f_J produced all lexical pitch accents and downstep for A2 accurately. An abrupt fall can be seen for the first word *ya'kuza* in A1 and A2, while a fall is missing from the first word *tomodachi* B1 and B2 (around pos1-3). As to A1 [+Acc, *ga*], downstep is observed, which is not predicted based on syntax-prosody mapping, but due to the rhythmic effect. Based on the syntax-prosody mapping theories, an IP boundary between *ya'kuza* and

one'esan in A1[+Acc, *ga*] is predicted, but is not produced by speaker 02f_J. Instead, the first two words are integrated into a single IP which allows downstep to occur at *one'esan* in A1 as well as A2. In sum, actual F0 contour and the average of F0 illustrate the same features and prosodic phrasing for speaker 02f_J can be listed as the following:

A1{$_{IP1}$($_{AP1}$ ya'kuza ga) ($_{AP2}$ one'esan ni)}{$_{IP2}$($_{AP3}$ nomi'mono o)($_{AP4}$ urima'shita)}

A2{$_{IP1}$($_{AP1}$ ya'kuza no)($_{AP2}$ one'esan ga)} {$_{IP2}$($_{AP3}$ nomi'mono o)($_{AP4}$ urima'shita)}

B1{$_{IP1}$($_{AP1}$ tomodachi ga)($_{AP2}$ one'esan ni)}{$_{IP2}$($_{AP3}$ nomi'mono o)($_{AP4}$ urima'shita)}

B2{ ($_{AP1}$ tomodachi no)($_{AP2}$ one'esan ga)} {$_{IP2}$($_{AP3}$ nomi'mono o)($_{AP4}$ urima'shita)}

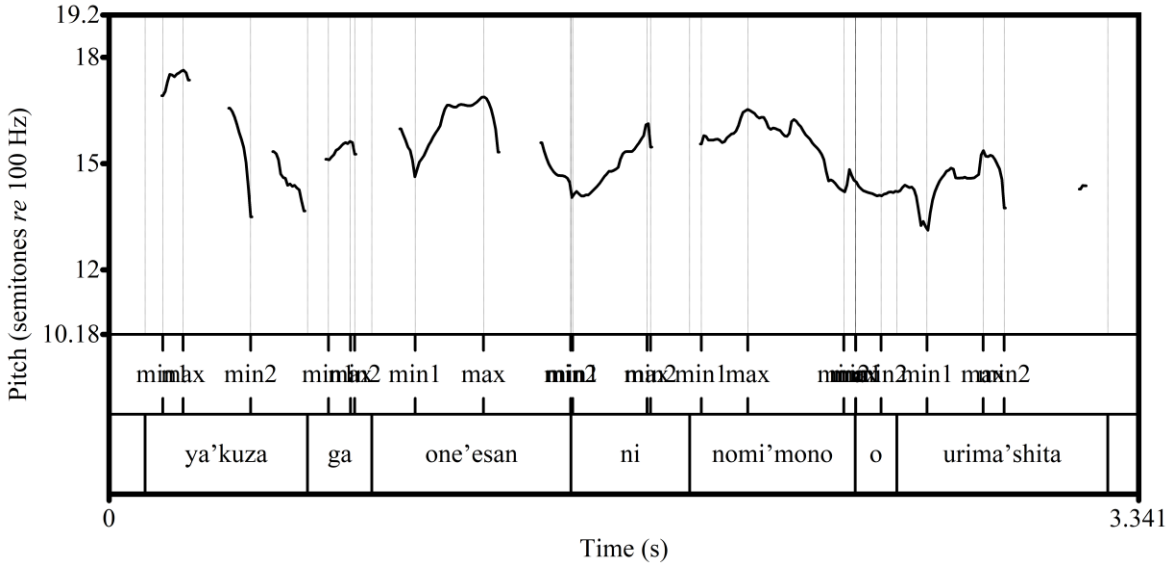


Figure 30: Pitch contour extracted from speaker 04f_J, eighth repetition for A1 [+Acc, *ga*]

Speech produced by speaker 04f_J is relatively monotonous compared to the other three Japanese native speakers. As for the researcher's perception of speaker 04f_J, the speech sounds natural and no noticeable foreign accent. As to the pitch accent assignment at the lexical level, there are abrupt falls on *ya'kuza, one'esan, nomi'mono* and *urima'shita*, which characterise pitch accents. As to the prosodic phrasing, fall in F0 are observed between *ya'kuza ga* and *one'esan ni*, and between *nomi'mono o* and *urima'shita.* The fall can be analysed as the insertion of AP boundaries. The effect of downstep is shown at *urima'shita,* which suggests that the last two words are in the same IP. Based on the observation, a prosodic phrasing for A1 produced by speaker 04f_J can be illustrated as follows;

A1{IP1(AP1ya'kuza ga) }{IP2(AP2one'esan ni)}{IP3(AP3 nomi'mono o)(AP4 urima'shita)}

Pitch contours produced by speaker 04f_J look similar in all other conditions, except the absence of pitch accent on tomodachi in B1 [-Acc, *ga*] and B2 [-Acc, *no*]. Therefore, here we do not take time to discuss the rest of conditions by looking at the actual pitch contours. Prosodic phrasing by speaker 04f_J can be summarised as follows;

A1{$_{IP1}$($_{AP1}$ ya'kuza ga)}{$_{IP2}$($_{AP2}$ one'esan ni)}{$_{IP3}$($_{AP3}$ nomi'mono o)($_{AP4}$ urima'shita)}

A2{$_{IP1}$($_{AP1}$ ya'kuza no)} {$_{IP2}$($_{AP2}$ one'esan ga)} {$_{IP3}$($_{AP3}$ nomi'mono o)($_{AP4}$ urima'shita)}

B1{$_{IP1}$($_{AP1}$ tomodachi ga)}{$_{IP2}$($_{AP2}$ one'esan ni)}{$_{IP3}$($_{AP3}$ nomi'mono o)($_{AP4}$ urima'shita)}

B2 {$_{IP1}$($_{AP1}$ tomodachi no)}{($_{AP2}$ one'esan ga)} {$_{IP3}$($_{AP3}$ nomi'mono o)($_{AP4}$ urima'shita)}

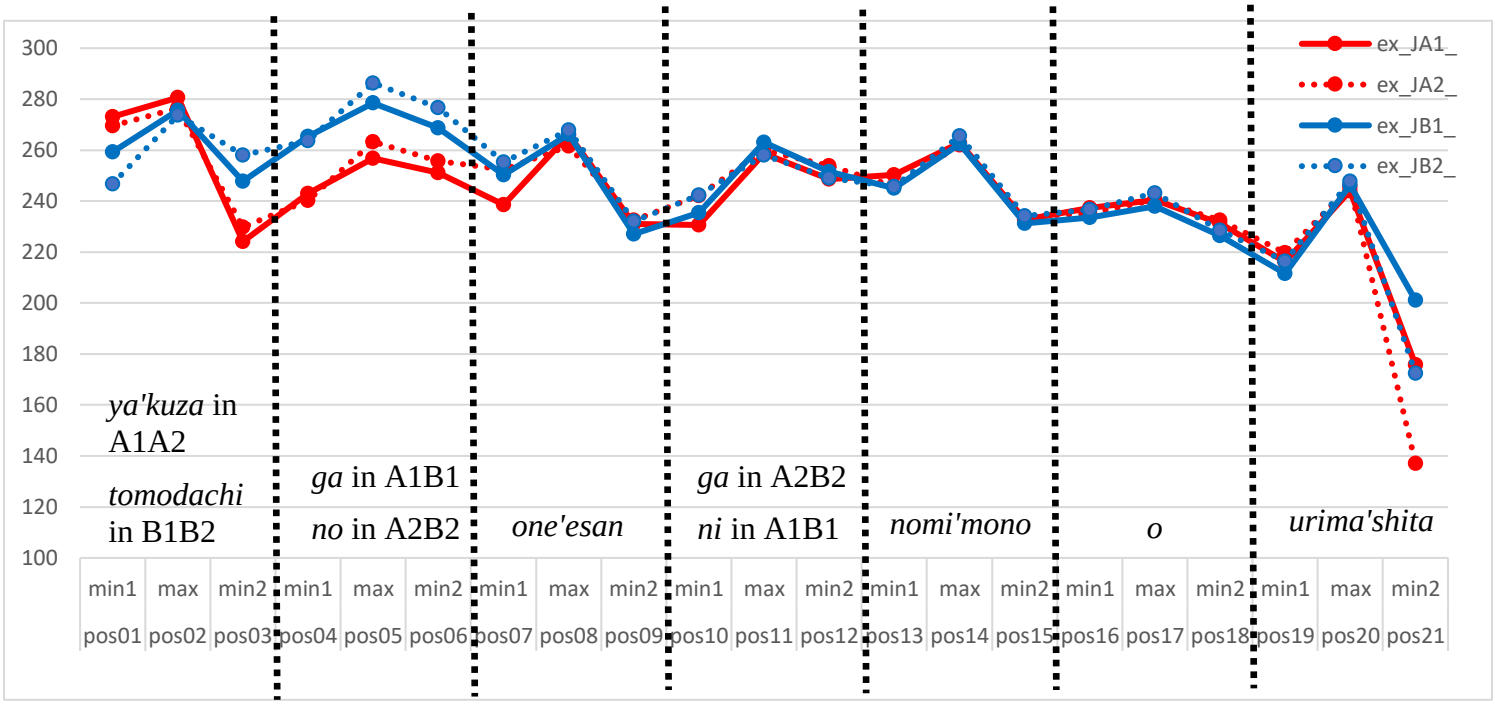


Figure 31: Line plots with mean F0 values from 21 measurement points by speaker 04f_J

Although her speech does sound native, speaker 04f_J's speech is relatively monotonous in comparison to the other Japanese native speakers. While overall, there is no clear contrast

between the conditions, F0 height for B1[-Acc, *ga*] and B2 [-Acc, *no*] are slightly higher than A1[+Acc, *ga*] and A2 [+Acc, *no*]. This indicates that the stimuli with *ya'kuza* are pronounced with a sharper fall in F0 than the other stimuli with *tomodachi*. As in the actual pitch contour, BPM on the nominative case marker *ga* in A2 [+Acc, *no*] and B2 [-Acc, *no*] are observed.

In summary, all of the Japanese native speakers assigned a lexical accent for all words properly. As to the phrasal level, the features of each speaker's speech are summarised in Table 12. If the F0 compression on *one'esan* in A2 [+Acc, *no*] is observed in both actual pitch contours and the average of F0 contour, ○ is noted as the effect of downstep is observed. If the speech of each speaker lacks the F0 compression, it is noted as × as a downstep does not occur. When BPM on particles occurs particle *no* and *ga*, ○ is written under the category of 'BPM on *no* and *ga* (pos4-6)'. If not, ×is written. In addition to ○, which particle gets BPM is noted. If both of the particles, *no* and *ga* (pos4-6), get BPM, ○ is noted under the category of 'BPM on *no* and *ga*'. If none of the particles gets BPM, × is written in Table 12.

Table 12: Features at the phrasal level produced by Japanese native speakers

speaker	downstep in A2 *one'esan*	BPM only *ga* pos4-6	BPM on *no* **and** *ga* pos4-6	BPM on ni **and** ga pos10-12	BPM on *o* pos 19-21
01f_J	○	○	×	○	○
02f_J	○ A2 and A1	×	×	×	×
03m_J	○	×	×	×	×
04f_J	× Hard to see	○	×	×	×

4.2 Swedish learners of Japanese: overview

There are two sections for Swedish learners of Japanese. First, the overall characteristics of speech produced by Swedish learners of Japanese at word level and phrase-level are explained.

Then, the next section 4.3 discusses the examples of individual cases. As same as for Japanese native speakers, the observation is based on actual pitch contours, the average of F0 contour obtained from each speaker and the researcher's perception who is a Tokyo Japanese native speaker.

Overall, most of the Swedish learners of Japanese produced errors in terms of lexical pitch accent for the first words; *ya'kuza* and *tomodachi*. The effect of downstep is also missing among the majority of the Swedish students. There are roughly three patterns commonly observed among the Swedish participants in terms of the overall shape of contours; the first group is the learners producing BPM on all of the particles. The second group is the Swedish participants who did not produce any difference in pitch contour for the four stimuli. The last group is the students produced a unique pitch contour, which cannot be found in the other students.

In Table 13, the errors of lexical pitch accent by Swedish learners of Japanese are noted with × accompanied by what kind of the errors they produced. Correct pitch assignments for each word are provided on the top of the table with apostrophes. The participants' correct pitch accent assignments are marked with ○. Whether each speaker's pitch accent assignment is correct or wrong is determined based on the observation of both actual pitch contours and, the average of F0 at relevant measurement points, as well as the author's perception as a Tokyo Japanese native speaker. The information of learners' Swedish is on the left of the chart, and Rik is for Rikssvenska, which means Standard Swedish in English. Skå is for Skånska, which is a Swedish variation mainly spoken in the Skåne area in Sweden. Norrland Svenska is a Swedish variation mainly spoken in the Norrland area in Sweden.

10 out of 14 Swedish learners of Japanese produced errors in the pitch accent for *ya'kuza*. The error pattern for *ya'kuza* is the same for 9 of the learners making mistakes, which is *yaku'za*. The exception is speaker 03m_S producing the word as unaccented. As to the pitch assignment for an unaccented word, *tomodachi*, 10 out of 14 learners made errors. The most frequent error for *tomodachi* is *tomoda'chi*, while one participant 14m_S produced *tomo'dachi*. Besides the first words *ya'kuza* and *tomodachi*, there are few errors in pitch accent assignment; two learners produced *one'esan* as unaccented *oneesan* (04f_S and 09f_S), and two learners produced *nomi'mono* as unaccented (03m_S and 04f_S). Overall, the learners pronounced the last three words correctly. As to the verb phrase *urima'shita*, uniform pitch accent assignment in the polite form *ma'su* and *ma'shita* would be one of the reasons why the learners produced it correctly.

Table 13: The errors of lexical pitch accent by Swedish learners of Japanese

Swedish variety	University	speaker	*ya'kuza*	*tomodachi*	*one'esan*	*nomi'mono*	*urima 'shita*
Skå	Lund	01m_S	× *yaku'za*	○ for B2 ×for B1 *tomoda'chi*	○	○	○
Rik+Skå	Lund	02m_S	× *yaku'za*	× *tomoda'chi*	○	○	○
Rik	Lund	03m_S	× *yakuza*	○	○	× *nomimono*	○
Skå	Lund	04f_S	○	○	× *oneesan*	× *nomimono*	○
Rik	Stockholm	05m_S	× *yaku'za*	× *tomoda'chi*	○	○	○
Rik	Stockholm	06f_S	× *yaku'za*	×*tomoda'chi*	○	○	○
Rik	Stockholm	07m_S	× *yaku'za*	× *tomoda'chi*	○	○	○
Rik Swedish L2	Stockholm	08m_S	× *yaku'za*	× *tomoda'chi*	○	○	○
Rik	Stockholm	09f_S	× *yaku'za*	○	× *oneesan*	○	○
Rik	Stockholm	10m_S	× *yaku'za*	× *tomoda'chi*	○	○	○
Rik	Lund	11f_S	○	× *tomoda'chi*	○	○	○

Norrland	Lund	12m_S	× *yaku'za*	× *tomoda'chi*	○	○	○
Rik+Skå	Lund	13f_S	○	× *tomoda'chi*	○	○	○
Rik	Stockholm	14m_S	○	× *tomo'dachi*	○	○	○

At the phrasal level, there is no F0 compression on the second lexical word, *one'esan* in A2 [+Acc, *no*] in comparison to *ya'kuza* in A2 [+Acc, ga] in their speech. In other words, the speech produced by 9 out of 14 Swedish learners of Japanese lack downstep in A2[+Acc, *no*] which is predicted to show downstep. Downstep produced by the remaining five learners is caused by either F0 compression on *one'esan* due to the BPM on the previous particles (*ga* and *no* in pos 4-6), as seen in speaker 02m_S and 04f_S. The other pattern of showing the effect of downstep is that all four stimuli show the effect, as in, for instance, speaker 10m_S, 11f_S, and 14m_S. All of them made mistakes to assign a pitch accent on an unaccented word *tomodachi* as tomoda'chi. Their errors at the word level meet the condition of causing downstep. Speaker 08m_S is a unique case where only A2[+Acc, *no*] and B2 [-Acc, *no*] show the F0 compression on *one'esan* indicating downstep. As well as speaker 08m_S, speaker 10m_S produced a unique pattern, in which only B1 [-Acc, *ga*] and B2 [-Acc, *no*] shows the effect of downstep on *one'esan*. Speaker 01m_S did not produce downstep in A2 [+Acc, *no*], but instead, the effect is observed in B1 and A1.

As to the BPMs on the genitive case particle *no* and the nominative case particle *ga* (pos4-6), the speech by six out of 14 learners show BPMs. Among these six learners, five learners put BPM on both of the particles, while one learner (speaker 12m_S) produced BPMs only on particle *ga*, which is the same feature as one native speaker, speaker 01f_J. On the locative case *ni* and the nominative case *ga* (pos10-12), six learners put BPM on both of the particles, and there are no cases in which learners produce BPM only either *ga* or *ni*. On accusative case *o* (pos 19-21), the speech by four out of 14 learners shows BPMs. If BPM is assumed to be inserted for APs boundaries as Igarashi (2015), it would indicate that six learners assigning BPM on both *ni* and *ga* inserted APs boundaries for the both particles.

Table 14: Features at the phrasal level produced by Swedish learners of Japanese

speaker	downstep in A2	BPM on *ga* only pos4-6	BPM on *no* **and** *ga* pos4-6	BPM on *ni* **and** *ga* pos10-12	BPM on *o* pos 19-21
01m_S	× A1 B1 get downstep	×	×	×	×
02m_S	○? due to BPM	×	○	○	○
03m_S	×	×	×	×	○
04f_S	○? due to BPM	×	○	○	×
05m_S	× downstep in B2	×	×	×	×
06f_S	×	×	○	○	×
07m_S	×	×	○	○	○
08m_S	○ All stimuli get downstep	×	×	×	×
09f_S	×	×	○	○	○
10m_S	× Only B1B2	×	×	×	×
11f_S	○ All stimuli get downstep	×	×	×	×

12m_S	×	○only particle *ga*	×	○	×
13f_S	×	×	×	×	×
14m_S	○All stimuli get downstep	×	×	×	×

As to the learners' patterns to produce prosodic phrasing, there are three types of students; the first group is the Swedish participants who did not produce any difference in pitch contours for the four stimuli. The second group is the learners producing BPM on all of the particles, as corresponds to AP boundaries. The last group is the learners who produced a unique pitch contour which cannot be found in the other students. The first two groups could be explained as general tendencies among Swedish learners of Japanese. About the last group with individual varieties, this group will be brought up in the next section 4.3.

Speaker 05m_S, 08m_S, 11f_S, 13f_S, and 14m_S consist of the first group of Swedish learners of Japanese producing similar contours for all of the stimuli named as 'the similar contours group' as follow. Figure 31-35 illustrates line plots of the F0 average for the similar contours group.

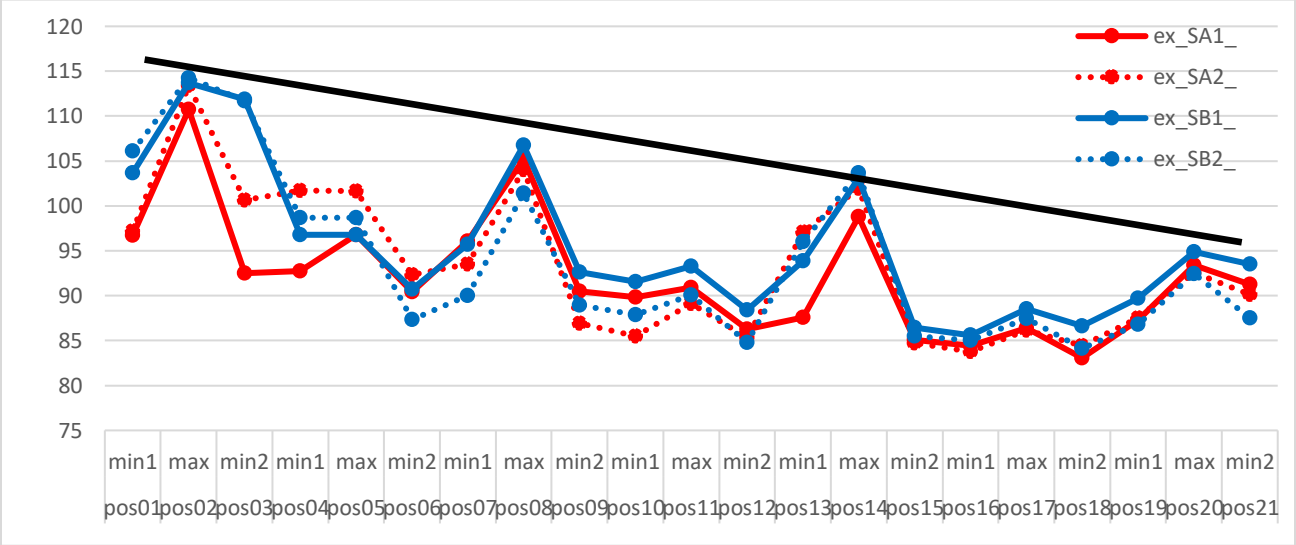


Figure 32: Line plots with mean F0 values from 21 measurement points by speaker 05m_S

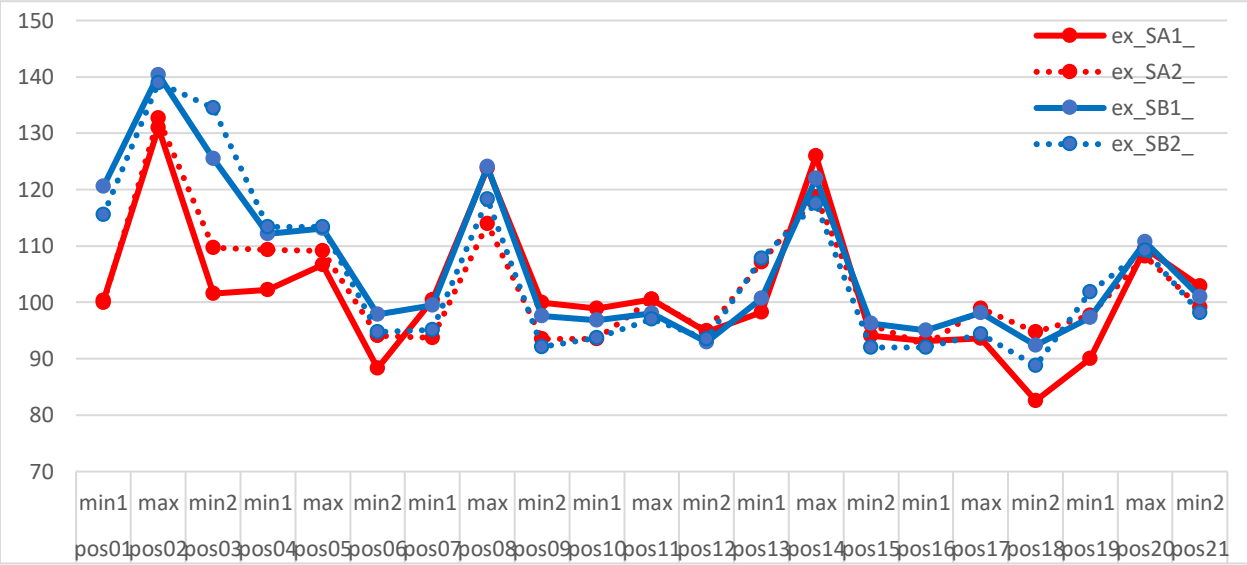


Figure 33: Line plots with mean F0 values from 21 measurement points by speaker 08m_S

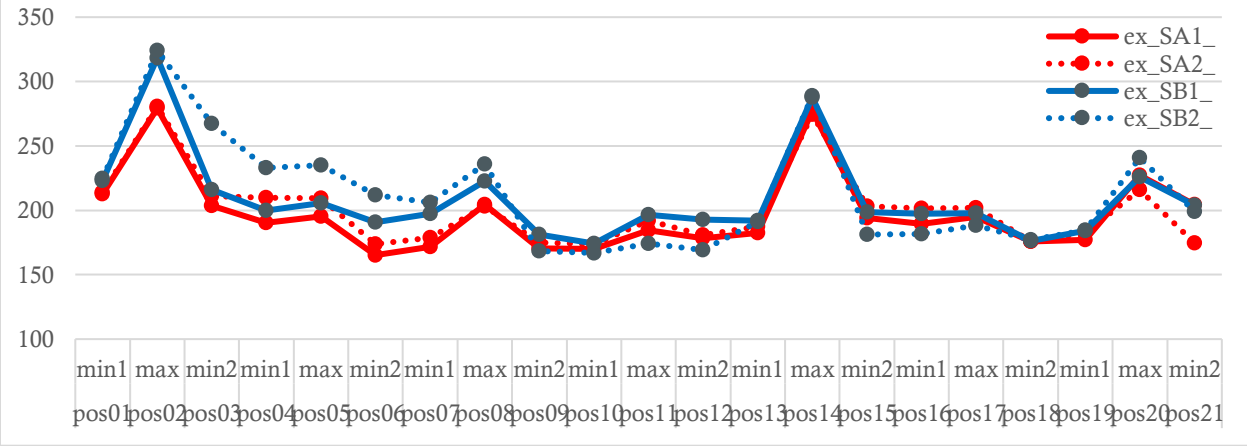


Figure 34: Line plots with mean F0 values from 21 measurement points by speaker 11f_S

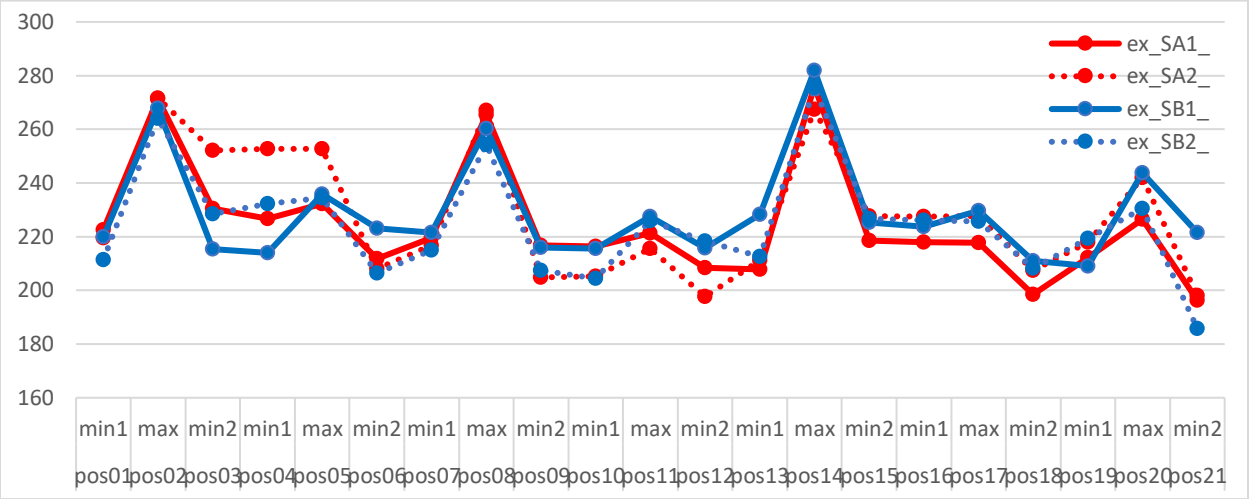


Figure 35: Line plots with mean F0 values from 21 measurement points by speaker 13f_S

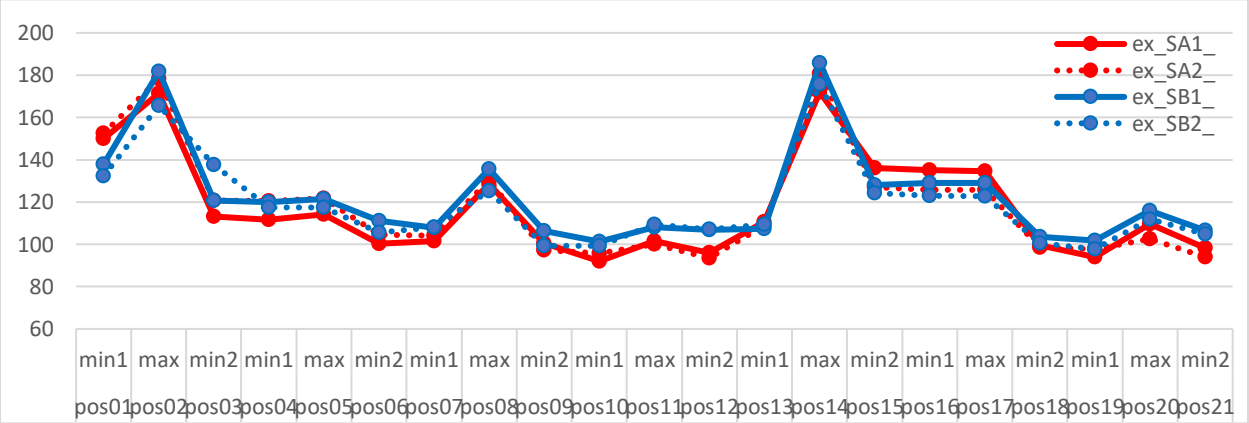


Figure 36: Line plots with mean F0 values from 21 measurement points by speaker 14m_S

Among the similar contours group, one of the common features is that the contours lack differences in the pitch contours despite the difference in the syntactic structures of the stimuli for the first two phrases. Their speech lacks the variety in terms of syntactic difference. However, the rises and falls due to the pitch accent of accented lexical words are depicted in Figure 32 to 36 by each speaker. Besides *tomoachi* in the speech of speaker 05m_S, the other speakers in the similar contours group pronounced *tomodachi* as an accented word *tomoda'chi*. Apparent rises and drops in F0 are seen at lexical words pronounced as accented; *ya'kuza* (pos1-3 in A1A2), *tomoda'chi* as an error (pos 1-3 in B1B2), *one'esan* (pos 7-9), *nomi'mono* (pos 13-15) and *urima'shita* (pos 19-21).

Regarding prosodic phrasing produced by the similar contours group, speaker 05m_S and 13f_S appear to show the same prosodic phrasing. Due to the limitation of space, here we only discuss speaker 05m_S as an example of prosodic phrasing as same as in speaker 13f_S. As shown in Figure 32, the F0 peaks of each lexical pitch accent show a gradual drop to the end of utterance in speaker 05m_S's speech. This phenomenon is termed as 'declination' which means a gradual decline of intensity and fundamental frequency as high subglottal energy at the beginning is used as an utterance continue (Lieberman 1967; Gussenhoven 2004). F0 compression on *one'esan* (pos 7-9) may be explained by the fact that there is a gradual decline due to declination, but not due to the effect of downstep. Since when more than one phrase containing an accented word, downstep is triggered, it can be analysed that each phrase produced by speaker 05m_S form IP on its own. Between each word (a lexical noun + a particle), F0 drops are found, which can be interpreted that there are AP boundaries inserted. Based on the observation of F0 average, prosodic phrasing by speaker 05m_S can be listed as follows (learner's errors in lexical pitch accent are underlined)

A1{$_{IP1}$($_{AP1}$ <u>yaku'za</u> ga)}{$_{IP2}$($_{AP2}$ one'esan ni)}{$_{IP3}$($_{AP3}$ nomi'mono o)}{$_{IP4}$($_{AP4}$ urima'shita)}

A2{$_{IP1}$($_{AP1}$ <u>yaku'za</u> no)} {$_{IP2}$ ($_{AP2}$one'esan ga)} {$_{IP3}$($_{AP3}$nomi'mono o)}{$_{IP4}$($_{AP4}$ urima'shita)}

B1{$_{IP1}$($_{AP1}$tomodachi ga)}{$_{IP2}$($_{AP2}$one'esan ni)}{$_{IP3}$($_{AP3}$ nomi'mono o)}{$_{IP4}$($_{AP4}$ urima'shita)}

B2 {$_{IP1}$($_{AP1}$tomodachi no)}{$_{IP2}$($_{AP2}$one'esan ga)}{$_{IP3}$($_{AP3}$nomi'mono o)}{$_{IP4}$($_{AP4}$urima'shita)}

Figure 37 to 40 provides actual F0 contours by speaker 05m_S for the four stimuli and confirms the observation as we discussed based on the average of F0 produced by speaker 05m_S. That is, declination and the prosodic phrasing listed above are also observed in the actual contours.

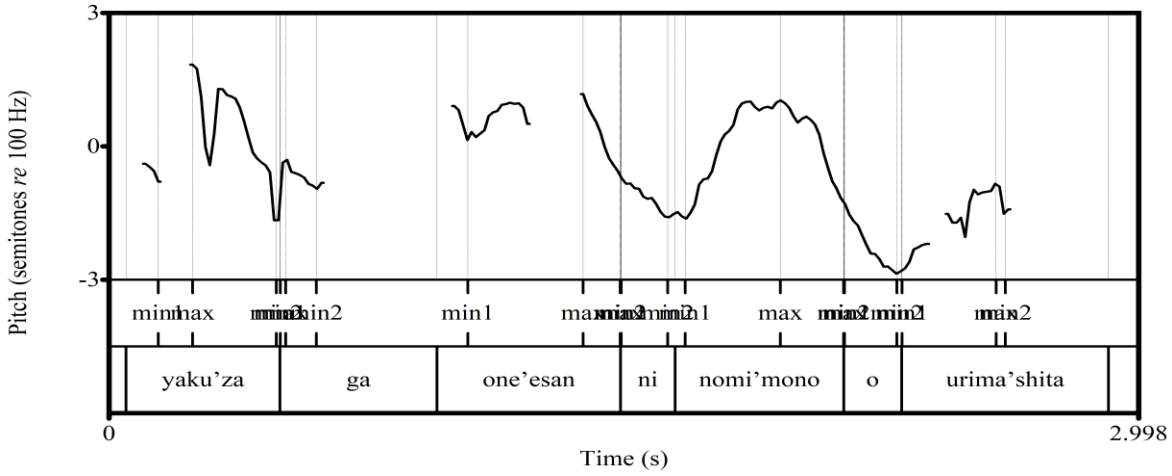


Figure 37: Pitch contour extracted from speaker 05m_S, seventh repetition for A1 [+Acc, *ga*]

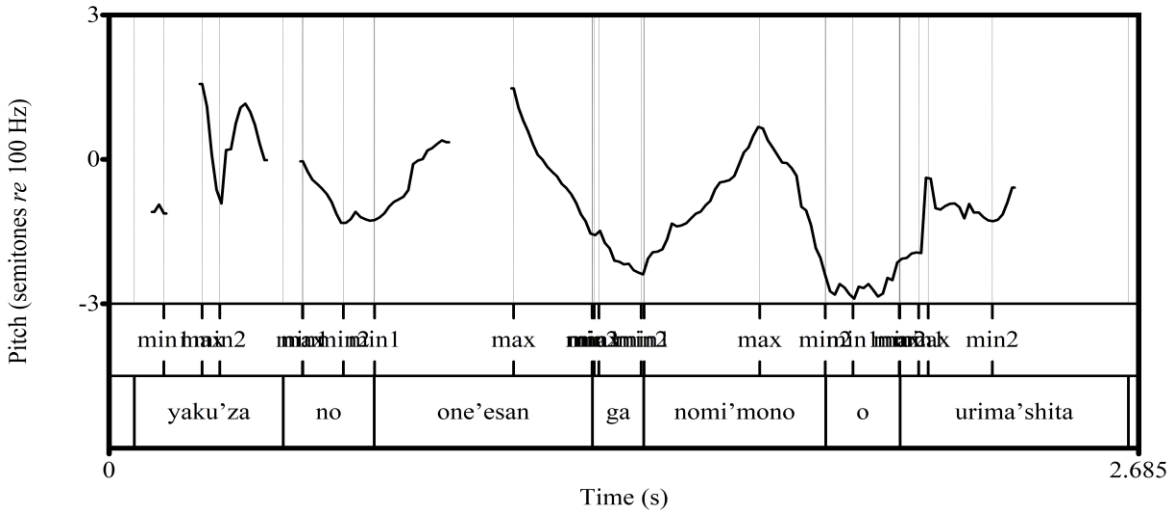


Figure 38: Pitch contour extracted from speaker 05m_S, seventh repetition for A2 [+Acc, *no*]

Speaker 05m_S's speech for A1 and A2 represented in Figure 37, and 38 respectively show declination from the beginning to the end of the utterance. With falls and followed rises between each phrase, it can be argued that there are AP boundaries for each phrase. The effect of F0 compression due to downstep cannot be found in speaker 05m_S, which indicates each phrase form own IP.

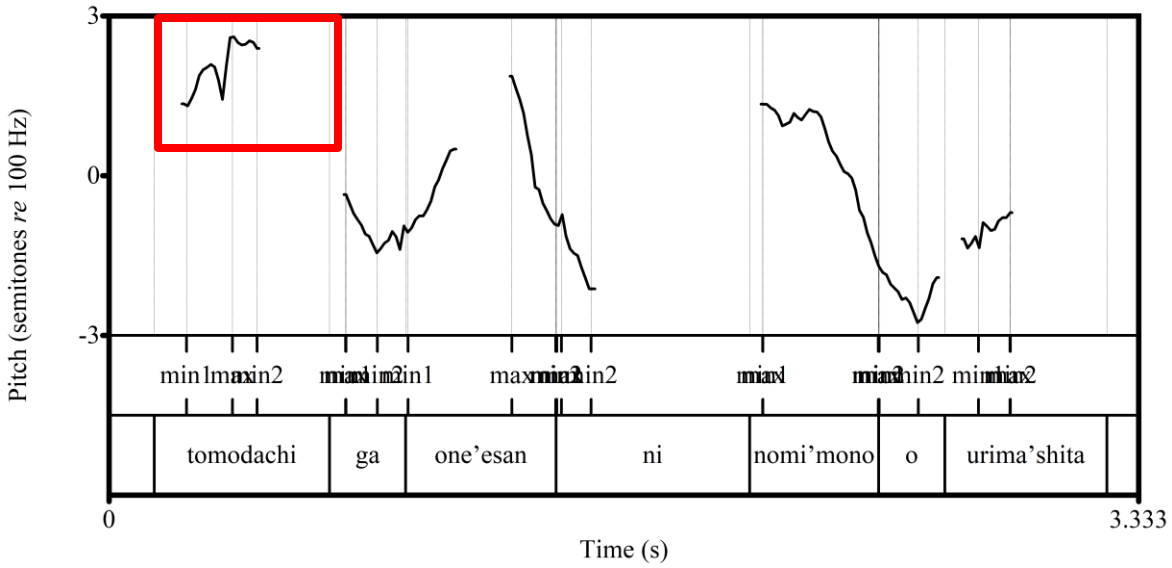


Figure 38: Pitch contour extracted from speaker 05m_S, seventh repetition for B1 [-Acc, *ga*]

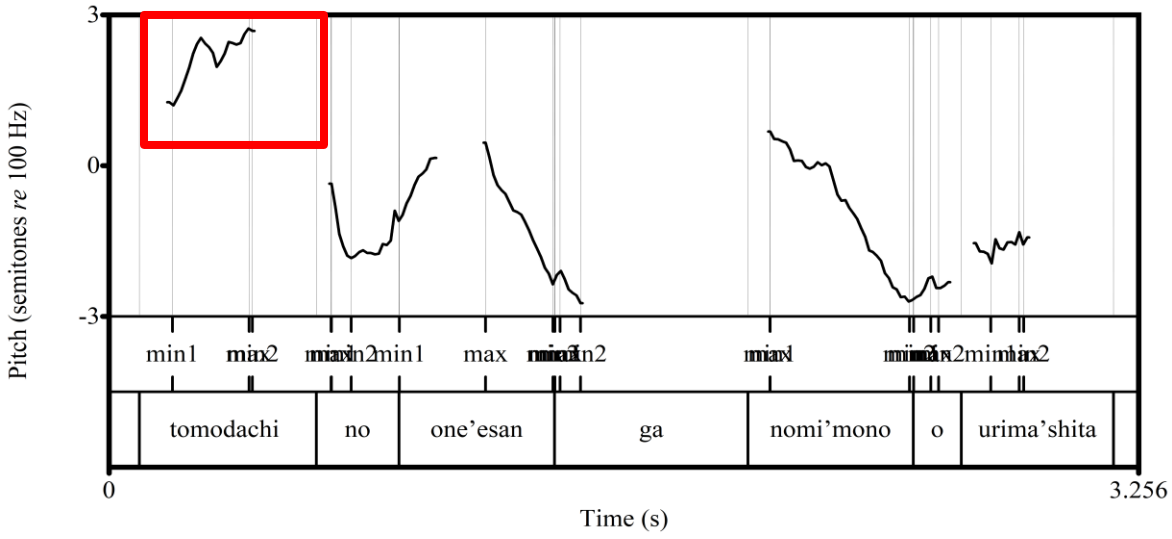


Figure 39: Pitch contour extracted from speaker 05m_S, eighth repetition for B2 [-Acc, no]

Figure 38 and 39 present actual pitch contours of B1 [-Acc, *ga*] and B2 [-Acc, *no*], seventh and eighth repetition, respectively. Looking at the first word *tomodachi* in both B1 and B2, a large fall in F0 cannot be found, but the F0 peak of the nominative case *ga* in B1 and the

genitive case particle *no* are realised very low. This can indicate that speaker 05m_S produced *tomodachi* as *tomoda'chi*[2] as an error. Only in B2, F0 compression on one'esan is found. That is, the effect of downstep is unexpectedly observed in B2 due to the lexical pitch accent error on *tomodachi* as *tomoda'chi*. Between each phrase, there are falls and followed rises, which show that there are AP boundaries. Additionally, there are significant drops between each phrase without the effect of downstep, and it can be said that the utterances produced by speaker 05m_Sconsist of four APs and IPs.

The rest of the speakers in the similar contours group, speaker 08m_S, 11f_S, and 14m_S share the same pattern of prosodic phrasing as follows, and as one of the examples, speaker 11f_S's production is discussed.

A1{$_{IP1}$($_{AP1}$ yaku'za ga)($_{AP2}$one'esan ni)}{$_{IP2}$($_{AP3}$ nomi'mono o)($_{AP4}$ urima'shita)}

A2{$_{IP1}$($_{AP1}$ yaku'za no)($_{AP2}$ one'esan ga)} {$_{IP2}$($_{AP3}$ nomi'mono o)($_{AP4}$ urima'shita)}

B1{$_{IP1}$($_{AP1}$tomoda'chi ga)($_{AP2}$ one'esan ni)}{$_{IP2}$($_{AP3}$nomi'mono o)($_{AP4}$urima'shita)}

B2 {$_{IP1}$($_{AP1}$tomoda'chi no)($_{AP2}$ one'esan ga)} {$_{IP2}$($_{AP3}$ nomi'mono o)($_{AP4}$urima'shita)}

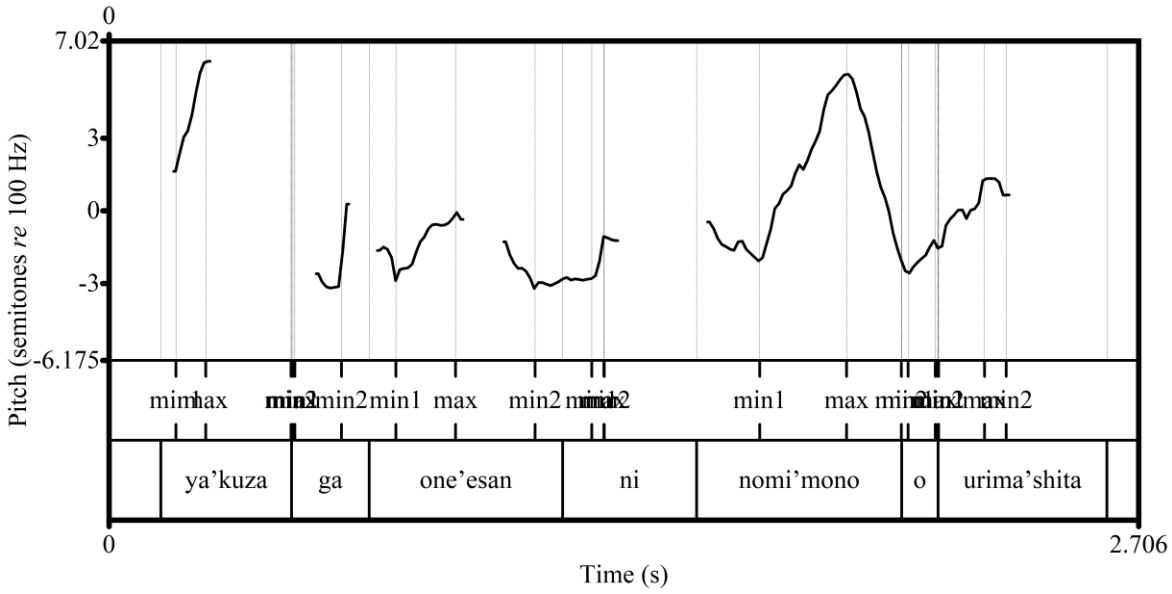


Figure 41: Pitch contour extracted from speaker 11f_S, eighth repetition for A1 [+Acc, *ga*]

[2] Speaker 05m_S is a Rikssvenska speaker, and the errors at *tomoda'chi* (marked with red squares in Figure 38 and 39) appear to be due to a negative transfer from accent 1 in Stockholm Swedish with a later pitch peak comparing to the same error of *tomoda'chi* produced by speaker 01m_S (speaker of Skånska) on page 72.

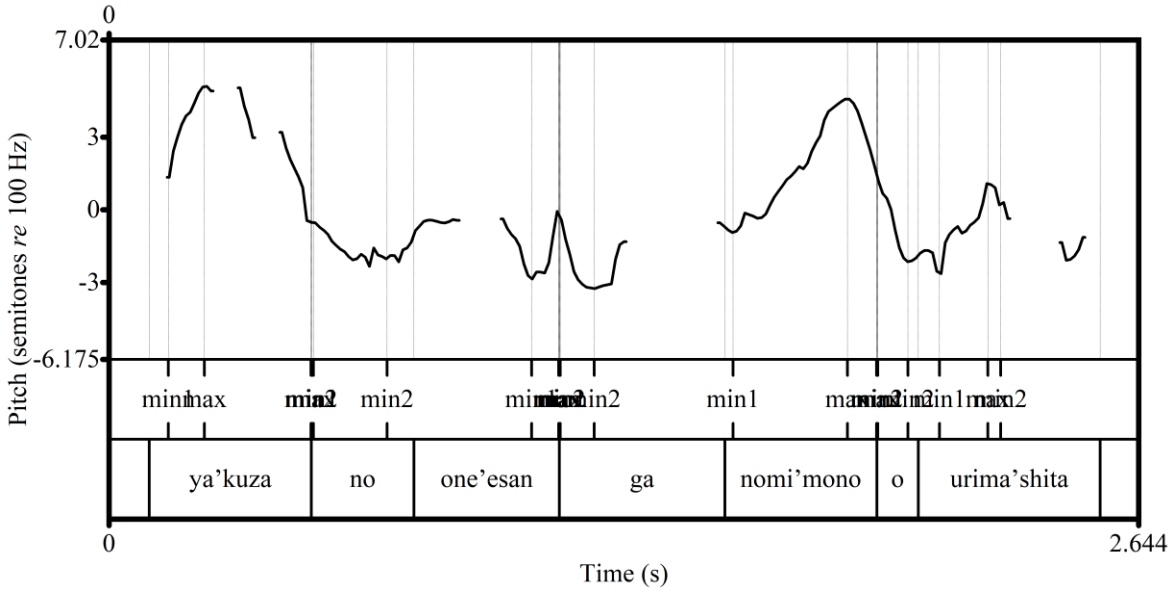


Figure 42: Pitch contour extracted from speaker 11f_S, eighth repetition for A1 [+Acc, *ga*]

Pitch contours for A1 [+Acc, *ga*] and A2 [+Acc, *no*] produced by speaker 11f_S are presented in Figure 41 and 42. One of the main features of speaker 11f_S is that F0 peak of *nomi'mono* is as high as the first word *ya'kuza.* This may be due to an insertion of a new IP boundary between *one'esan ga* in A1 *one'esan no* in A2 and *nomi'mono o*. In both A1 and A2, each word (one lexical word+a particle) starts with a fall and followed by the rise, which means that there is an initial lowering of new AP. Looking at the second word *one'esan* and the last verb *urima'shita,* there is clear F0 compression in comparison to the F0 peak of *ya'kuza* and *nomi'mono* respectively in both A1 and A2. The compression can be interpreted as the effect of downstep. Therefore, the first two phrases can be analysed in the same IP, and the last two phrases are in another IP.

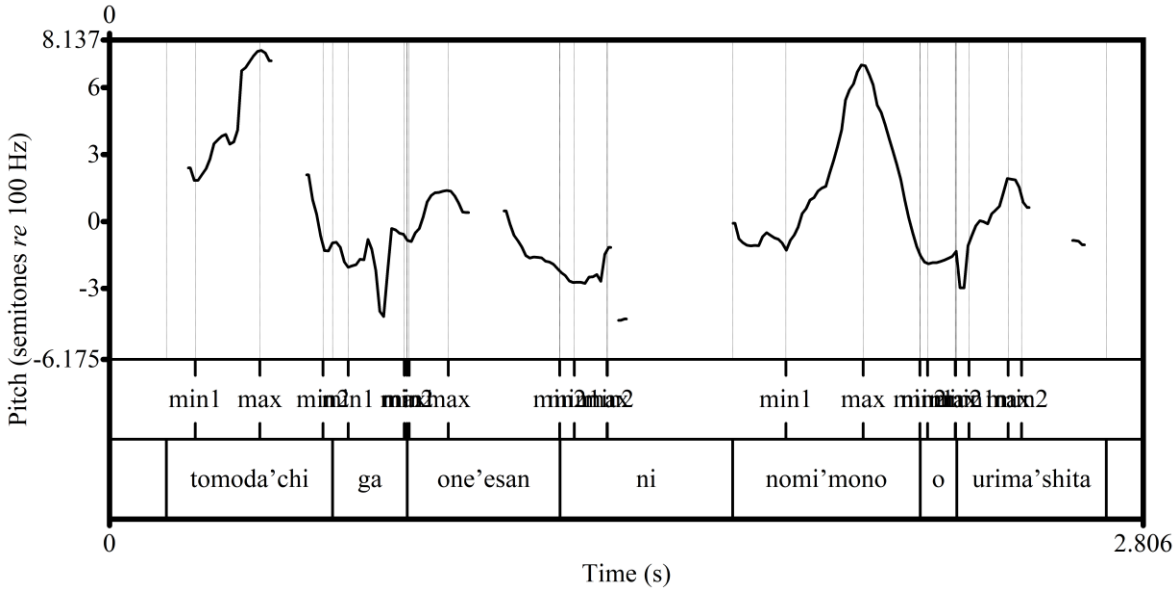


Figure 43: Pitch contour extracted from speaker 11f_S, eighth repetition for B1 [-Acc, *ga*]

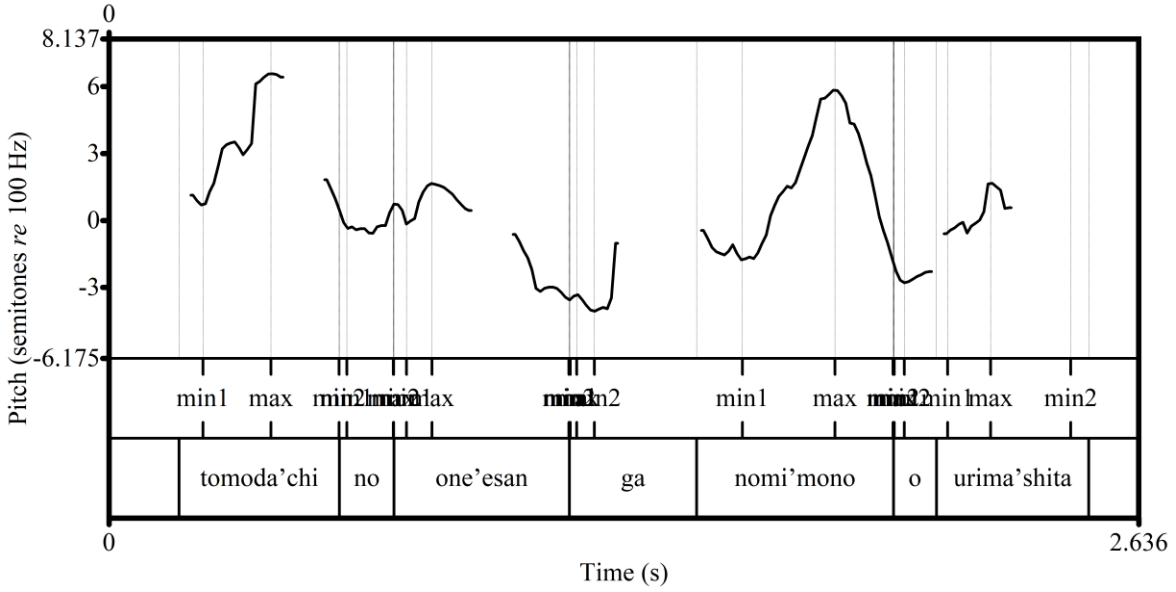


Figure 44: Pitch contour extracted from speaker 11f_S, eighth repetition for B2 [-Acc, *no*]

Figure 43 and 44 provide actual pitch contours produced by speaker 11f_S for B1 [-Acc, *ga*] and B2 [-Acc, *no*]. Looking at large falls in F0 in the first word, it was pronounced as an accented word as an error in both B1 and B2. Moving to the next word *one'esan* in both B1 and B2, it is clear that there is F0 compression on the word compared to the first word. This result can be analysed as the effect of downstep on *one'esan*, and the first two phrases are in the same IP. A similar F0 compression is observed at the verb *urima'shita*, which indicates that the last two phrases are in one IP and triggers downstep on the verb. Due to the limit of paper, we cannot discuss in detail on speaker 08m_S and 14m_S. However, their speech is observed to have the same prosodic phrasing as speaker 11f_S.

The second group is named as BPM group, where the Swedish learners of Japanese assign BPM on all of the particles as corresponds AP boundaries for all of the stimuli. Five Swedish learners, Speaker 02m_S, 04f_S, 06f_S, 07m_S and 09f_S belong to this group. As an example from the BPM group, line plots of the F0 average produced by speaker 07m_S is provided in Figure 45.

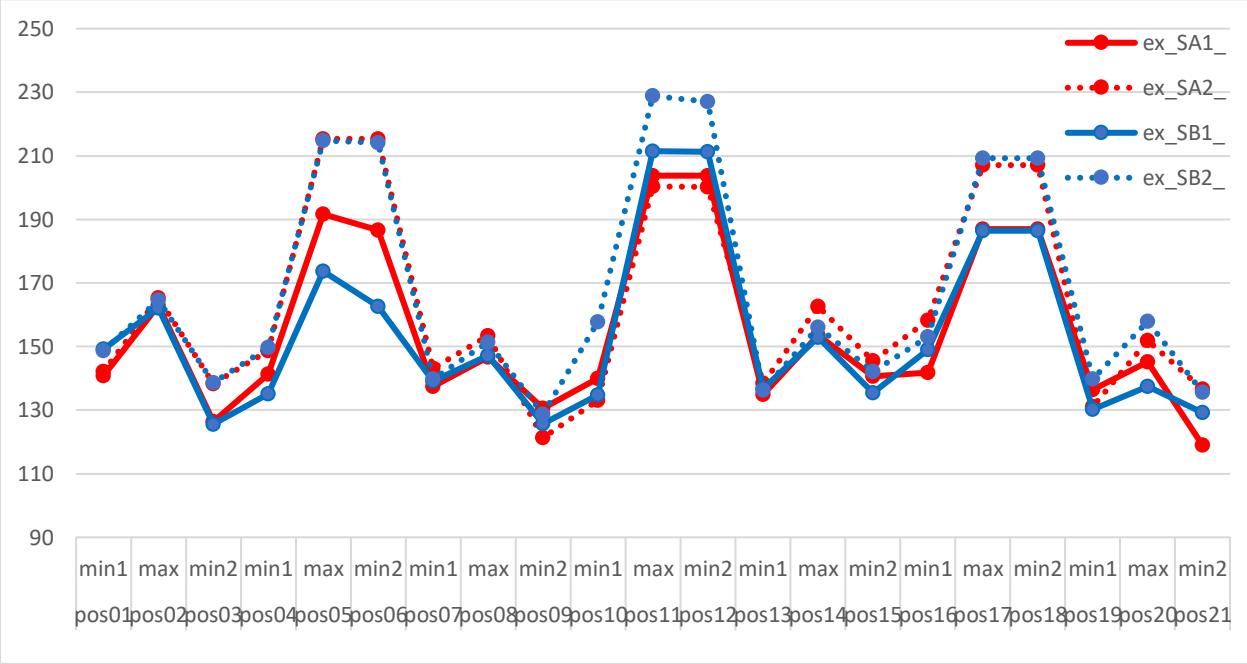


Figure 45: Line plots with mean F0 values from 21 measurement points by speaker 07m_S

What is interesting about the average of F0 contours by speaker 07m_S in Figure 45 is that there are evident BPM on all of the particles. If we look closely and compare between A1 B1 and A2 B2, F0 peaks of BPM for the nominative case *ga* (pos 4-6) in A1 B1 are lower than that of the genitive case *no* (pos 4-6) in A2 B2. This tendency can be seen on the accusative case *o* (pos16-18) as well.

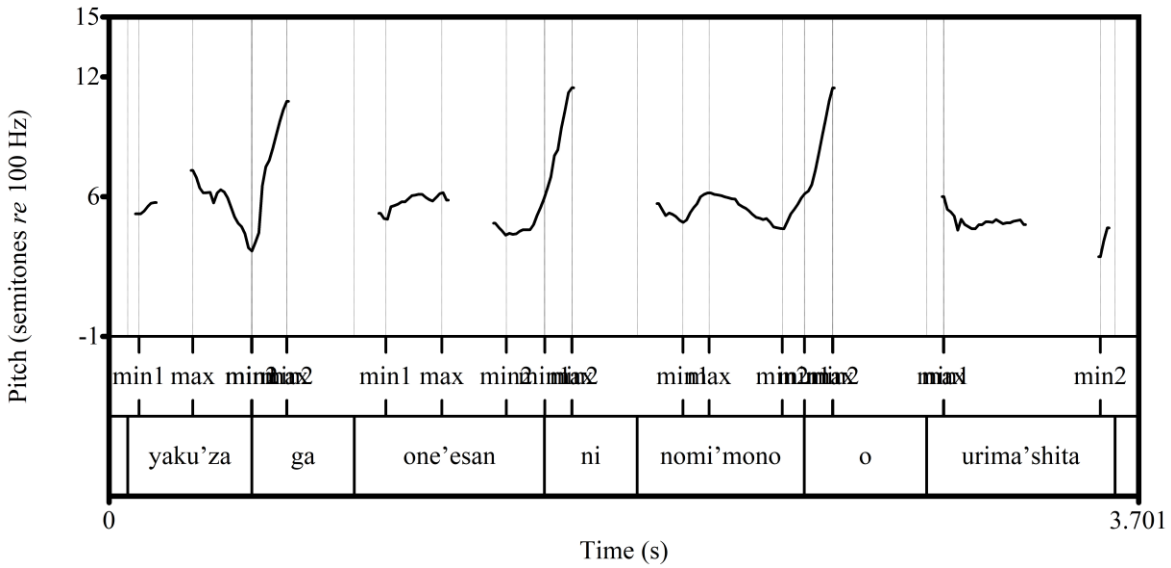


Figure 46: Pitch contour extracted from speaker 07m_S, seventh repetition for A1 [+Acc, *ga*]

If we now turn to Figure 46 an actual pitch contour for the seventh repetition for A1 by speaker 07m_S seems to support the observation based on the line plots of the F0 average represented in Figure 45. That is, large rises at the particles (*ga, ni,* and *o*) due to BPM are represented.

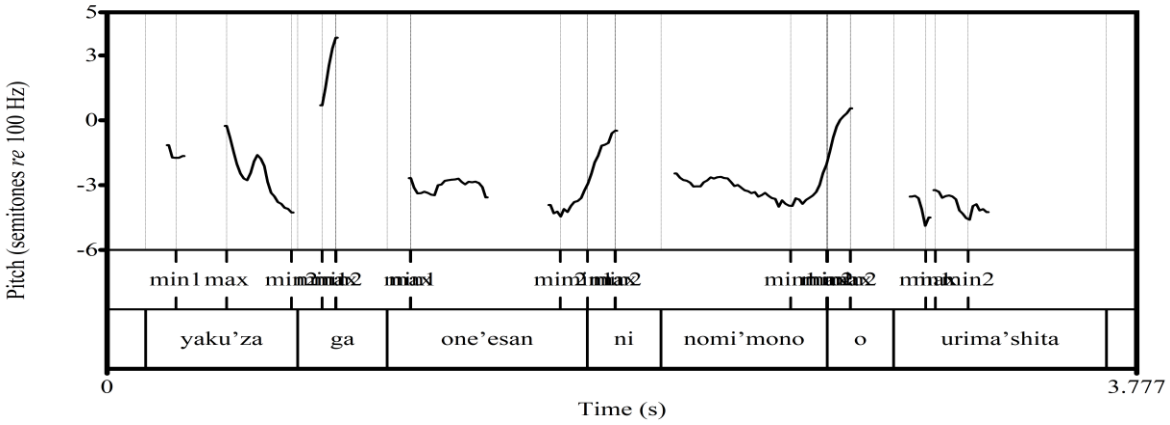


Figure 47: Pitch contour extracted from speaker 02m_S, seventh repetition for A1 [+Acc, *ga*]

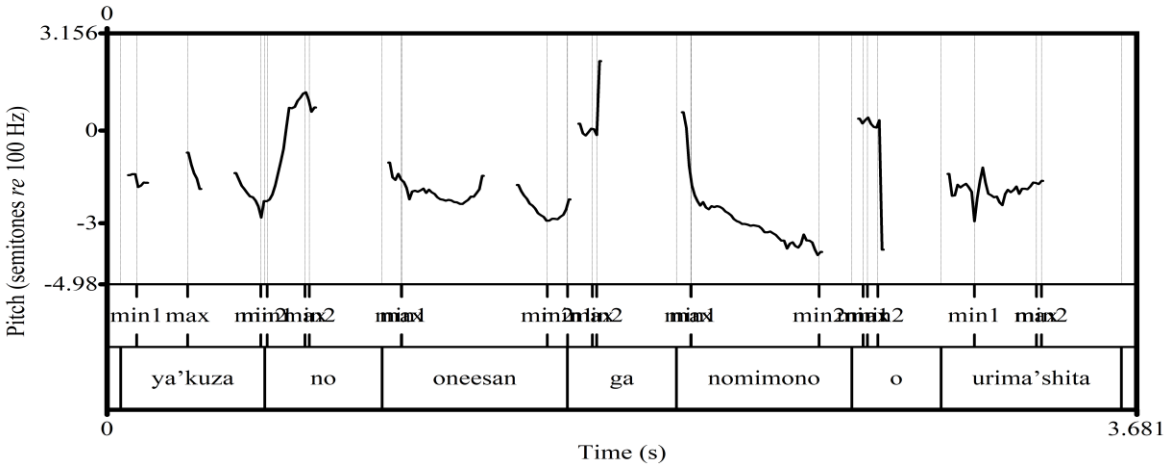


Figure 48: Pitch contour extracted from speaker 04f_S, seventh repetition for A2 [+Acc, *no*]

Figures 47 and 48 are actual pitch contours for A1, A2 by speaker 02m_S, and 04f_S, respectively. Both utterances are the seventh repetition. As same to speaker 07m_S, they show a large rise in F0 at the particles. Figure 49 and 50 below illustrate B1 and B2 production by speaker 06f_S and 09f_S. Although due to the limit of space, this paper cannot provide figures of all of their production by the speakers in the BPM group, all of the five students in the BPM group produced evident BPMs on all of the particles for all of the stimuli. Table 15 summarises the information of the two groups explained in section 4.2.

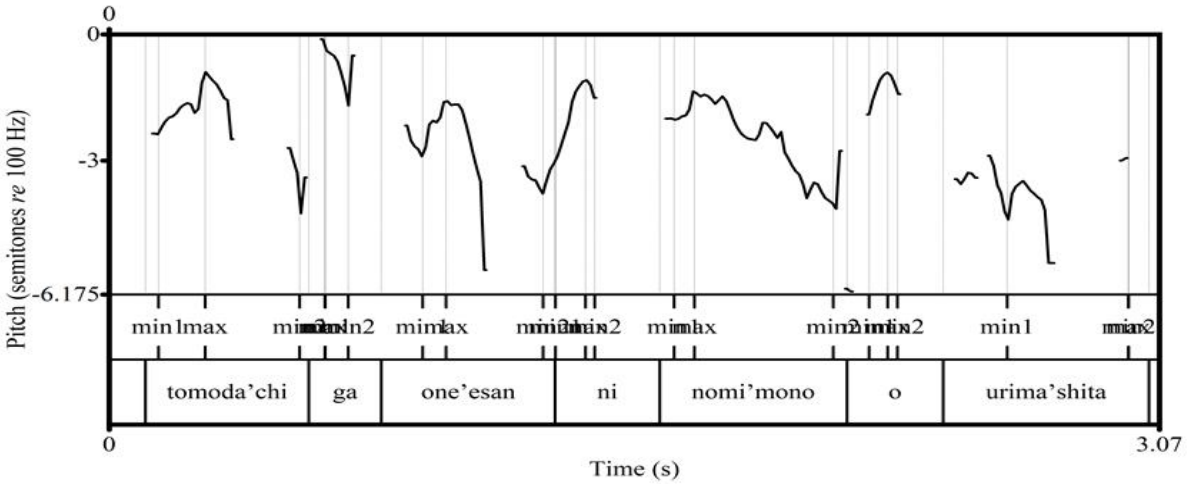


Figure 49: Pitch contour extracted from speaker 06f_S, seventh repetition for B1 [-Acc, *ga*]

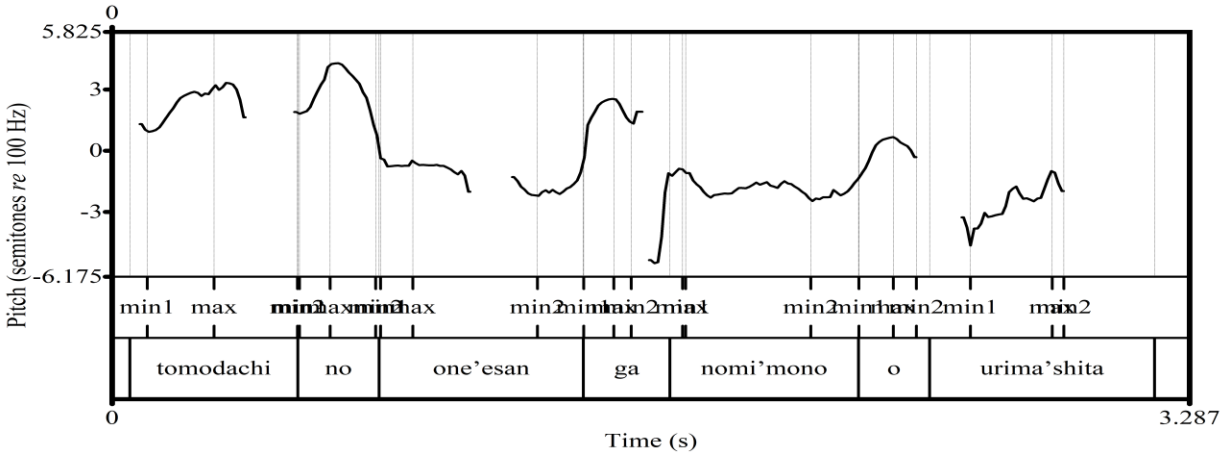


Figure 50: Pitch contour extracted from speaker 09f_S, seventh repetition for B1 [-Acc, *no*]

Table 15: Two groups of Swedish learners of Japanese based on their overall contours

The similar contours group	05m_S, 08m_S, 11f_S, 13f_S, 14m_S
The BPM group	02m_S, 04f_S, 06f_S 07m_S, 09f_S

Regarding prosodic phrasing produced by Swedish learners of Japanese in the BPM group, as Igarashi (2015) argues that BPM occurs at AP boundaries, all of the learners in this group seemed to divide each phrase into four different APs and consequently four IPs. Therefore, prosodic phrasing by the BPM group learners can be summarised as follows:

A1{$_{IP1}$($_{AP1}$ yaku'za ga)}{$_{IP2}$($_{AP2}$one'esan ni)}{$_{IP3}$($_{AP3}$ nomi'mono o)}{$_{IP4}$($_{AP4}$ urima'shita)}

A2{$_{IP1}$($_{AP1}$ yaku'za no)} {$_{IP2}$($_{AP2}$one'esan ga)} {$_{IP3}$($_{AP3}$nomi'mono o)}{$_{IP4}$($_{AP4}$ urima'shita)}

B1{$_{IP1}$($_{AP1}$tomodachi ga)}{$_{IP2}$($_{AP2}$ one'esan ni)}{$_{IP3}$($_{AP3}$nomi'mono o)}{$_{IP4}$($_{AP4}$urima'shita)}

B2 {$_{IP1}$($_{AP1}$tomodachi no)}{$_{IP2}$($_{AP2}$one'esan ga)}{$_{IP3}$($_{AP3}$nomi'mono o)}{$_{IP4}$($_{AP4}$urima'shita)}

In summary, the majority of Swedish students made errors in the lexical pitch accent assignment, particularly the first two words; *ya'kuza* and *tomodachi*. At the phrasal level, the lack of downstep on the expected phrase *one'esan* in A2 [+Acc, *no*] are commonly observed. Even among the learners producing downstep, most of them produce for all four stimuli or F0 compression due to BPM on the particle before *one'esan* or because of the errors in the lexical pitch accent prior to the expected part of downstep. As to the overall pitch contours produced by Swedish learners of Japanese, there are three groups in terms of tendency; the first group is namely 'the same contours group' where learners produced similar contours despite the syntactic difference in each stimulus. All of the learners in 'the same contours group' answered that they speak Rikssvenska. The second group is 'the BPM group', where learners assign BPM on all of the particles at AP boundaries. The last group is the rest of the students who do not have the tendency found among the two groups but have individual patterns of pitch contour. Prosodic phrasing produced by the similar contours group and the BPM group can be summarised as in Table 16. Some of the learners from the last group that has individual characteristics are discussed in the next section 4.3.

Table 16: Summary of prosodic phrasing by the similar contours and BPM group

group	speakers	prosodic phrasing
similar contours	05m_S, 13f_S	A1{$_{IP1}$($_{AP1}$ yaku'za ga)}{$_{IP2}$($_{AP2}$one'esan ni)}{$_{IP3}$($_{AP3}$ nomi'mono o)}{$_{IP4}$ ($_{AP4}$ urima'shita)}
		A2{$_{IP1}$($_{AP1}$ yaku'za no)} {$_{IP2}$ ($_{AP2}$ one'esan ga)} {$_{IP3}$($_{AP3}$ nomi'mono o)}{$_{IP4}$($_{AP4}$ urima'shita)}
		B1{$_{IP1}$($_{AP1}$tomodachi ga)}{$_{IP2}$($_{AP2}$ one'esan ni)}{$_{IP3}$($_{AP3}$nomi'mono o)}{$_{IP4}$($_{AP4}$urima'shita)}
		B2 {$_{IP1}$($_{AP1}$tomodachi no)}{$_{IP2}$($_{AP2}$ one'esan ga)} {$_{IP3}$($_{AP3}$ nomi'mono o)}{$_{IP4}$($_{AP4}$urima'shita)}
	08m_S, 11f_S, 14m_S	A1{$_{IP1}$($_{AP1}$ yaku'za ga)($_{AP2}$one'esan ni)}{$_{IP2}$($_{AP3}$ nomi'mono o)($_{AP4}$ urima'shita)}
		A2{$_{IP1}$($_{AP1}$ yaku'za no)($_{AP2}$ one'esan ga)} {$_{IP2}$($_{AP3}$ nomi'mono o)($_{AP4}$ urima'shita)}
		B1{$_{IP1}$($_{AP1}$tomodachi ga)($_{AP2}$ one'esan ni)}{$_{IP2}$($_{AP3}$nomi'mono o)($_{AP4}$urima'shita)}
		B2 {$_{IP1}$($_{AP1}$tomodachi no)($_{AP2}$ one'esan ga)} {$_{IP2}$($_{AP3}$ nomi'mono o)($_{AP4}$urima'shita)}
BPM	02m_S, 04f_S, 06f_S, 07m_S, 09m_S	A1{$_{IP1}$($_{AP1}$ yaku'za ga)}{$_{IP2}$($_{AP2}$one'esan ni)}{$_{IP3}$($_{AP3}$ nomi'mono o)}{$_{IP4}$ ($_{AP4}$ urima'shita)}
		A2{$_{IP1}$($_{AP1}$ yaku'za no)} {$_{IP2}$ ($_{AP2}$ one'esan ga)} {$_{IP3}$($_{AP3}$ nomi'mono o)}{$_{IP4}$($_{AP4}$ urima'shita)}
		B1{$_{IP1}$($_{AP1}$tomodachi ga)}{$_{IP2}$($_{AP2}$ one'esan ni)}{$_{IP3}$($_{AP3}$nomi'mono o)}{$_{IP4}$($_{AP4}$urima'shita)}
		B2 {$_{IP1}$($_{AP1}$tomodachi no)}{$_{IP2}$($_{AP2}$ one'esan ga)} {$_{IP3}$($_{AP3}$ nomi'mono o)}{$_{IP4}$($_{AP4}$ urima'shita)}

4.3 Swedish learners of Japanese: individual cases

This section concerns the last group of Swedish learners of Japanese who produced individually unique patterns. Speaker 01m_S, 03m_S, 10m_S, and 12m_S belong to this group. Speaker 01m_S is unique in a way that A1[+Acc, *ga*] B1[-Acc, *ga*] get the effect of downstep; in contrast, the first two phrases in A2 B2 form flat like contours. Speaker 03m_S produced plateau shape contours for the first two phrases for all of the stimuli with differences in F0 peak at particles a nominative case *ga* in A2B2 and a locative case *ni.* Speaker 10m_S produced higher F0 for A1A2 from particles *ga* in A1 *no* and *ni* in A2 after the first word to particles *ni* in A1 and *ga* in A2 in comparison to B1B2. Speaker 12m_S shows a strong tendency to assign BPM on particles besides a genitive case *no* and an accusative case *o.* Due to the limitation of space, here we discuss only speaker 01m_S and 12m_S.

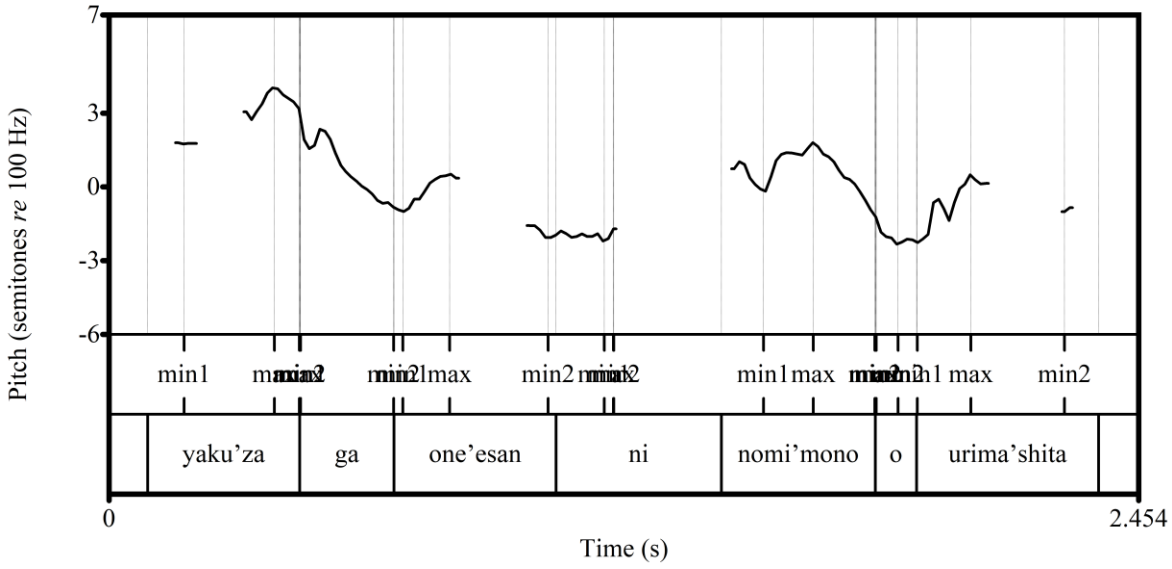


Figure 51: Pitch contour extracted from speaker 01m_S, eighth repetition for A1 [+Acc, *ga*]

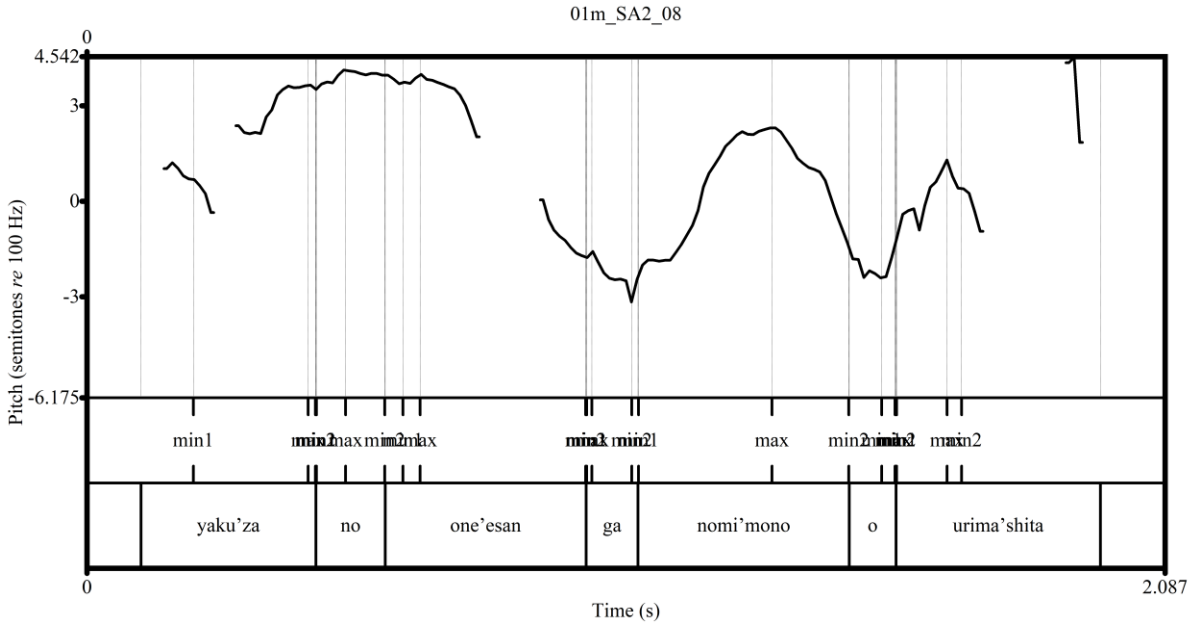


Figure 52: Pitch contour extracted from speaker 01m_S, eighth repetition for A2 [+Acc, *no*]

Figure 51 and 52 present A1 [+Acc, *ga*] and A2 [+Acc, *no*] production by speaker 01m_S, eighth repetition. Speaker 01m_S produced errors in terms of pitch accent assignment *ya'kuza* as *yaku'za.* Pitch accents of the other words are assigned correctly. On the second word *one'esan* in A1, the F0 range appears to be compressed in comparison to the first word *yaku'za*. Similar F0 compression is observed at the last verb *urima'shita* in A1 and A2. These compressions can be explained as the effect of downstep. Downstep in A1 is unexpected based on the syntax-prosody mapping. This unexpected downstep can be due to the rhythmic effect. The predicted downstep in A2 is missing in speaker 01m_S's speech. Falls and following rises are seen at the predicted AP boundaries in A1 but absent in A2, which means that speaker 01m_S inserted AP boundaries as predicted in A1 but not A2. In A2, it is possible to say that yaku'za no one'esan ga is in one AP although it contains two pitch accents in one AP and against the characteristic of AP that has only one accent in each AP. As to A2, an IP boundary is assumed to be between *yaku'za no* and *one'esan ga* since there is no effect of downstep. However, a large fall in F0 caused by an IP boundary is also lacking, and the first two words form a plateau shape contour rather. Based on the contour, here we analyse these two words as in one IP but no downstep (underlined part below) Prosodic phrasing of A1 and A2 by speaker 01m_S can be summarised as follows:

A1{$_{IP1}$($_{AP1}$ yaku'za ga)($_{AP2}$one'esan ni)}{$_{IP2}$($_{AP3}$ nomi'mono o) ($_{AP4}$ urima'shita)}

A2{$_{IP1}$($_{AP1}$ yaku'za no one'esan ga)} {$_{IP2}$($_{AP2}$ nomi'mono o) ($_{AP3}$ urima'shita)}

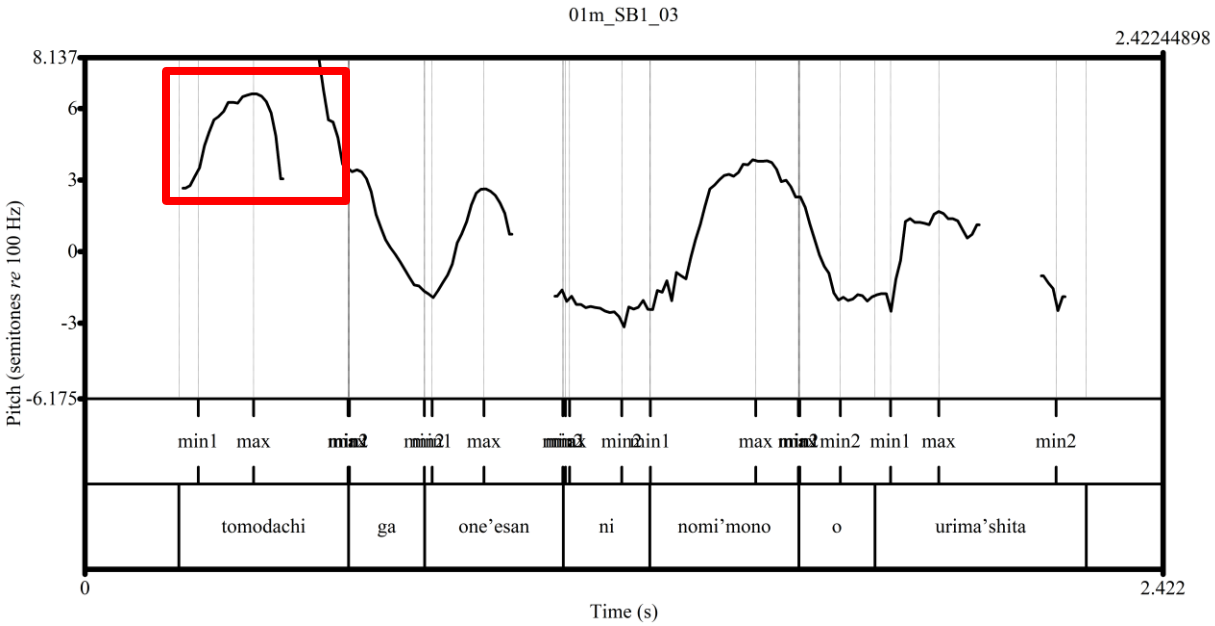


Figure 53: Pitch contour extracted from speaker 01m_S, third repetition for B1 [-Acc, *ga*]

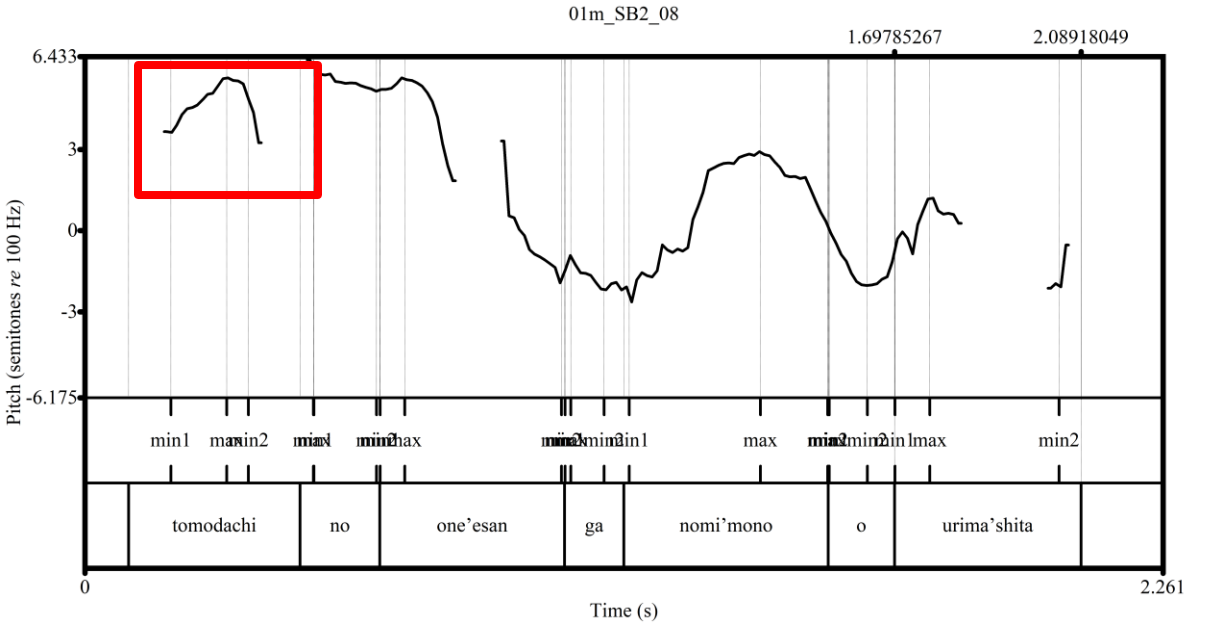


Figure 54: Pitch contour extracted from speaker 01m_S, eighth repetition for B1 [-Acc, *no*]

Figure 53 and 54 provide speaker 01m_S's third B1 [-Acc, *ga*] and eighth B2 [-Acc, *no*] production. The first unaccented word *tomodachi* is produced correctly in B2, but not in B1. In B1, *tomodachi* seems to be pronounced as *tomoda'chi*[3]. The F0 peak at *one'esan* is lower than the first word *tomodachi*, which can indicate that the effect of downstep unexpectedly appears on *one'esan.* According to the theory of syntax-prosody mapping, the first two words *tomodachi ga one'esan ni* in B1 is not the necessary condition to trigger downstep. However, in speaker 01m_S's speech, downstep seemed to occur. Similar compressions are seen at the last verb *urima'shita* in both B1 and B2. At the beginning of each word, falls and rises in F0 can be observed, which could be analysed as the insertion of AP boundaries. Based on the observation prosodic phrasing for B1 and B2 by speaker 01m_S are listed below. The part that cannot be explained based on the syntax prosody mapping is underlined in B1.

B1{$_{IP1}$($_{AP1}$tomoda'chi ga)($_{AP2}$ one'esan ni)}{$_{IP2}$($_{AP3}$nomi'mono o)}($_{AP4}$urima'shita)}

B2{$_{IP1}$($_{AP1}$tomodachi no)}{$_{IP2}$($_{AP2}$one'esan ga)} {$_{IP3}$($_{AP3}$nomi'mono o)}{$_{IP4}$($_{AP4}$urima'shita)}

To sum up, the most exciting aspect of the speech by speaker 01m_S is that A1 and B1 show the effect of downstep, while A2, which is an expected stimulus to get downstep and B2 show relatively flat shape contour. In other words, the overall contours vary depending on the syntactic difference between A1 B1 [±Acc, *ga*] and A2 B2 [±Acc, *no*].

The next case is speaker 12m_S, who shows a strong tendency to assign BPMs on particles except the genitive case particle *no* in A2 [+Acc, *no*] and B2 [-Acc, *no*]. Figure 55 and 56 illustrate speaker 12m_S's A1 and A2 production. As shown in Figure 55 and 56, it is very evident that BPMs are assigned to the particles except for genitive case *no* and accusative case *o*; nominative case *ga* and locative case *ni* get BPM. In other words, BPM is produced only at the left edge of XPs where IP boundaries are predicted. This pattern is parallel to one of the native speaker's data (speaker 01f_J). F0 compressions can be observed at the last verb *urima'shita* in both A1 and A2. The compressions could suggest that the last two phrases are in one IP both in A1 and A2 produced by speaker 12m_S. Prosodic phrasing for A1 and A2 by speaker 12m_S are:

A1{$_{IP1}$($_{AP1}$ yaku'za ga)}{$_{IP2}$($_{AP2}$one'esan ni)}{$_{IP3}$($_{AP3}$ nomi'mono o)($_{AP4}$ urima'shita)}

A2{$_{IP1}$($_{AP1}$ yaku'za no)($_{AP2}$ one'esan ga)} {$_{IP2}$($_{AP3}$ nomi'mono o)($_{AP4}$ urima'shita)}

[3] The errors at *tomoda'chi* made by speaker 01m_S marked with red squares in Figure 53 and 54 seem to have earlier pitch peaks in comparison to speaker 05m_S (Rikssvenska speaker) on page 61

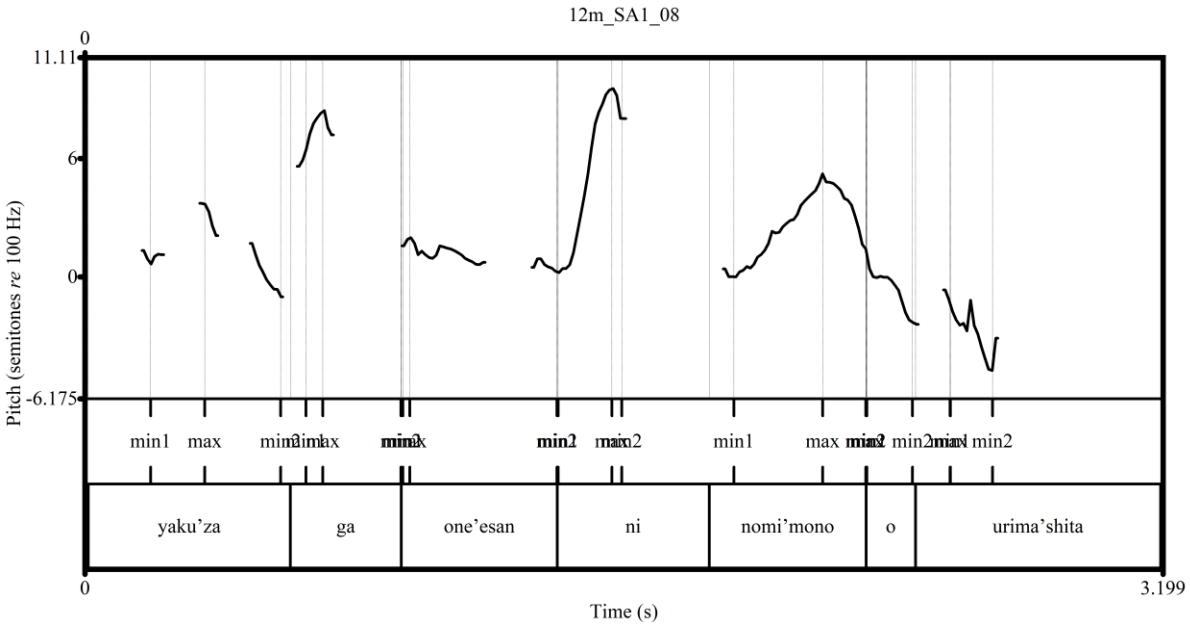


Figure 55: Pitch contour extracted from speaker 12m_S, eighth repetition for A1 [+Acc, *ga*]

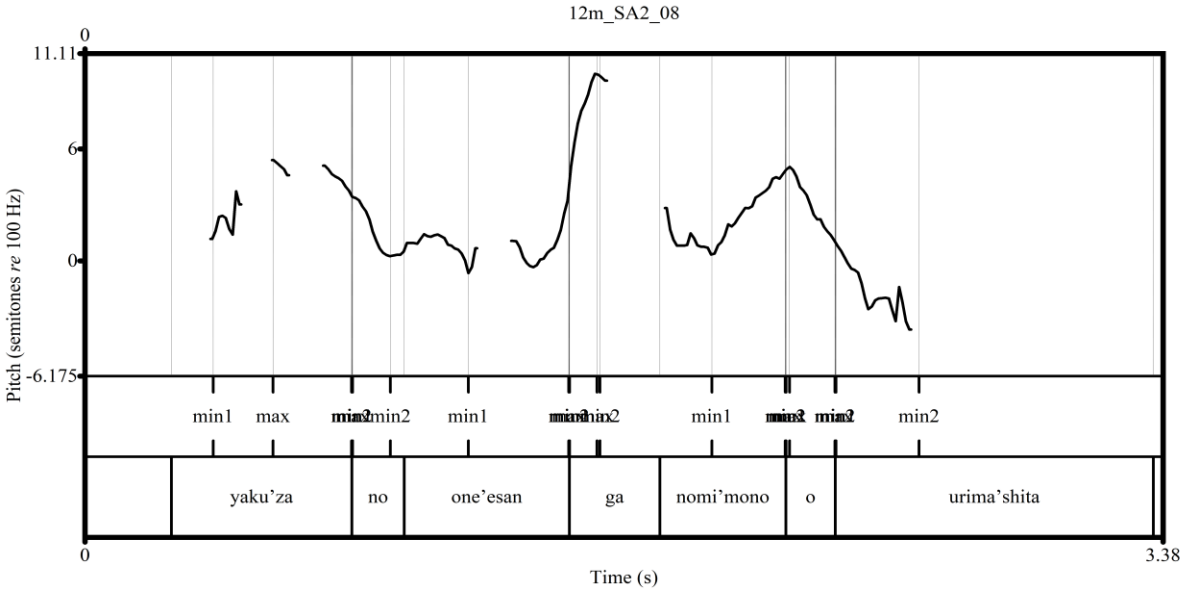


Figure 56: Pitch contour extracted from speaker 12m_S, eighth repetition for A2 [+Acc, *no*]

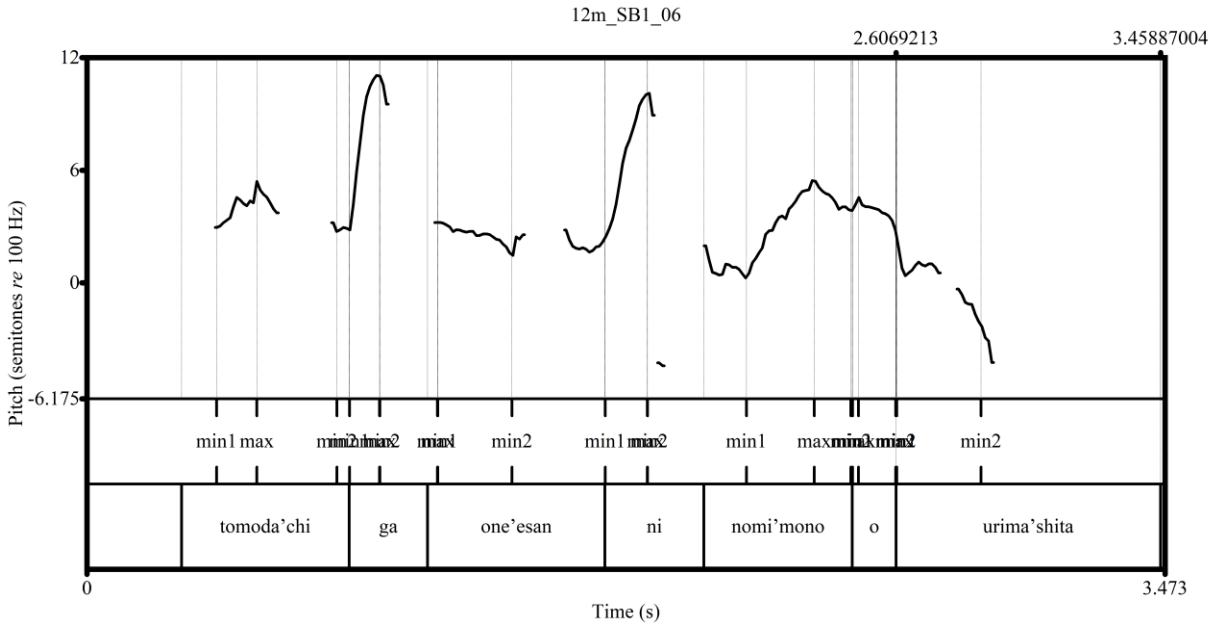


Figure 57: Pitch contour extracted from speaker 12m_S, sixth repetition for B1 [-Acc, *ga*]

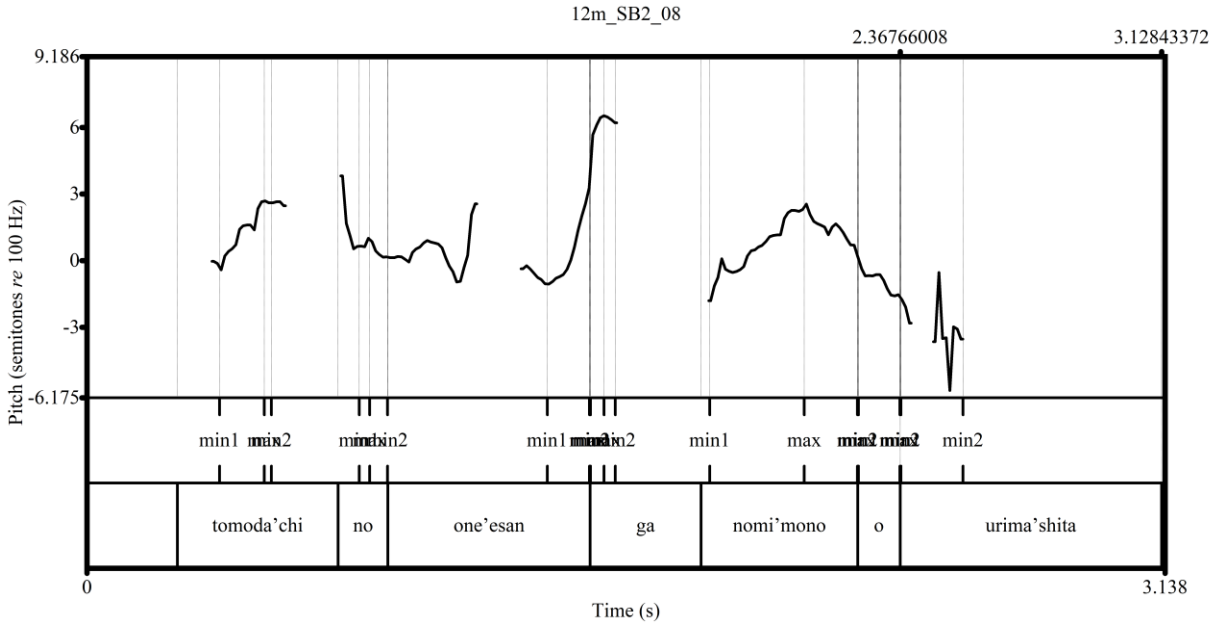


Figure 58: Pitch contour extracted from speaker 12m_S, eighth repetition for B2 [-Acc, *no*]

In B1 [-Acc, *ga*] represented in Figure 57 and B2 [-Acc, *no*] shown in Figure 58, speaker 12m_S produced an unaccented word *tomodachi* as *tomoda'chi.* The same tendency, as seen in A1 and A2, is observed on the particles *ga*, and *ni* showing BPMs. F0 peaks on the last verb *urima'shita* appear to be compressed in both B1 and B2, which indicates that the last two words are in the same IP and trigger the effect of downstep. Based on the contour shown in Figure 57 and 58, prosodic phrasing for B1 and B2 produced by speaker 12m_S is:

B1{$_{IP1}$($_{AP1}$tomoda'chi ga)}{$_{IP2}$($_{AP2}$one'esan ni)}{$_{IP3}$($_{AP3}$nomi'mono o)}($_{AP4}$urima'shita)}

B2{$_{IP1}$($_{AP1}$tomoda'chi no)($_{AP2}$ one'esan ga)} {$_{IP2}$($_{AP3}$ nomi'mono o)($_{AP4}$urima'shita)}

Overall, the data of pitch contour produced by speaker 12m_S shows the strong tendency of assigning BPMs on the following particles; nominative case marker *ga* and locative case marker *ni* and excluding genitive case *no* and accusative case *o*. In other words, speaker 12m_S appears to assign BPMs on the left edge of XPs, where IP boundaries are predicted.

In this section, we have discussed two speakers from the last group, where each student has an individually unique pattern. Speaker 01m_S produces downstep only in A1 and B1 [±Acc, *ga*] not in A2 and B2 [±Acc, *no*]. In speaker 12m_S speech, BPMs are observed at the only the left edge of XPs, on a nominative case particle *ga* and a locative case particle *ni*. Prosodic phrasing and unique features of speaker 01m_S and 12m_S are summarised in Table 17.

As to the correlation between the characteristics of Speech produced by Swedish students and their attributes, such as their university, what Swedish varieties they speak and learning experience, the result of this study seems not to show a clear correlation.

Table 17: Two speakers from the last unique group

Speaker 01m_S: downstep in A1 B1 [±Acc, *ga*]
A1{$_{IP1}$($_{AP1}$ yaku'za ga)($_{AP2}$one'esan ni)}{$_{IP2}$($_{AP3}$ nomi'mono o) ($_{AP4}$ urima'shita)}
A2{$_{IP1}$($_{AP1}$ yaku'za no one'esan ga)} {$_{IP2}$($_{AP2}$ nomi'mono o) ($_{AP3}$ urima'shita)}
B1{$_{IP1}$($_{AP1}$tomoda'chi ga)($_{AP2}$ one'esan ni)}{$_{IP2}$($_{AP3}$nomi'mono o)}($_{AP4}$urima'shita)}
B2{$_{IP1}$($_{AP1}$tomodachi no)}{$_{IP2}$($_{AP2}$one'esan ga)} {$_{IP3}$($_{AP3}$nomi'mono o)}{$_{IP4}$($_{AP4}$urima'shita)}
Speaker 12m_S: BPMs only on the left edge of XPs (nominative and locative case marker)
A1{$_{IP1}$($_{AP1}$ yaku'za ga)}{$_{IP2}$($_{AP2}$one'esan ni)}{$_{IP3}$($_{AP3}$ nomi'mono o)($_{AP4}$ urima'shita)}
A2{$_{IP1}$($_{AP1}$ yaku'za no)($_{AP2}$ one'esan ga)} {$_{IP2}$($_{AP3}$ nomi'mono o)($_{AP4}$ urima'shita)}
B1{$_{IP1}$($_{AP1}$tomoda'chi ga)}{$_{IP2}$($_{AP2}$one'esan ni)}{$_{IP3}$($_{AP3}$nomi'mono o)}($_{AP4}$urima'shita)}
B2 {$_{IP1}$($_{AP1}$tomoda'chi no)($_{AP2}$ one'esan ga)} {$_{IP2}$($_{AP3}$ nomi'mono o)($_{AP4}$urima'shita)}

Chapter 5 Discussion

In this chapter, the results are discussed concerning the two research questions and the findings of previous studies. The previous studies (Sakamoto 2010; Ogasawara 2015) researched the errors at word-level prosody, but not sentence level. In Tsurutani (2011) study, her analysis has an underlying assumption that Japanese native speakers produce similar contours. With this assumption, her discussion focus on how the errors produced by the learners deviate from the model speech of Japanese native speakers. As a reminder for readers, this study aims to examine the errors at both word and sentence prosody while showing that there are varieties in the pitch contours produced by Japanese native speakers. To guide the study, the two research questions were formulated in section 3.1, repeated here:

Research question 1: How is the interaction between syntax-prosody mapping and prosodic-well-formedness realised in the utterances produced by Swedish learners of Japanese in comparison to Japanese native speakers?

Research question 2: Can their errors be explained as a transfer from their L1, Swedish?

In addition to the two research questions, the findings of previous studies and predictions of learners' errors and difficulties will also guide the discussion. Where it is relevant, the previous studies are referred in relation to the findings of this study. To remind what we have previously discussed, the findings of previous studies and the predictions are listed below. Based on CAH (Lado 1957) and MDH (Eckman 1977, 1985), the following errors and difficulties concerning Swedish learners of Japanese are predicted:

-Assigning stress on the penultimate syllable of Japanese word, and associating a rise in F0

-Difficulties in acquiring unaccented words in Japanese

-Assigning H tone to mark phrase boundaries in Japanese when L tone is supposed to be assigned

Moreover, the previous studies on L2 prosodic acquisition discussed in section 2.3 reported English learners' frequent errors in Japanese are:

-Assigning a pitch accent to unaccented words or assigning it on a wrong location

-Assigning unnecessary H tone at the end of the prosodic word

-Lack of downstep

-Monotonous speech

This chapter consists of three sub-sections. The first two subsections discuss the two research questions. The last subsection is a summary of the discussion. In section 5.1, the speech of Japanese participants and Swedish learners of Japanese are compared in terms of how the interaction between the syntax-prosody mapping and prosodic well-formedness is depicted. In section 5.2, whether the errors produced by Swedish learners of Japanese can be explained is discussed.

5.1 RQ1: the interaction between syntax-prosody mapping and prosody well-formedness

The result of this study indicates that the pitch contours produced by Japanese native speakers correspond with the predictions made by the theory of syntax-prosody mapping and prosody well-formedness despite the varieties in the prosody of their production. The diversity in the Japanese native speakers' production is contrary to the previous study by Tsurutani (2011), where ten native speakers produced similar contours, and their speech is used as the model speech to compare with the production of English learners of Japanese. While the production by Swedish learners of Japanese lack the correspondence between syntax and prosody, and their cases cannot be fully explained based on the interaction between the syntax-prosody mapping and prosodic well-formedness.

In detail, the speech by speaker 03m_J clearly shows the correspondence between syntax and prosody. Although the contour by speaker 02f_J differs a lot from that of speaker 03m_J, it can be explained as the binarity effect. As Igarashi (2015) argues, assigning BPMs on AP boundaries is commonly observed among Japanese native speakers as we see in the speech by speaker 01f_J. According to the previous studies (Hara 1993; Kibe, Takeda, and Tanaka, et al., 2013; Hirachi 2008), this phenomenon is also termed as "high-rising intonation" (shiriagari intonation in Japanese). "high-rising intonation" has been initially spread from females in the Kanto area, which is around Tokyo. Therefore, speaker 01f_J can be explained as an example of "high-rising intonation." Although the speech produced by speaker 04f_J is rather monotonous in comparison to the other Japanese native speakers, Figures of actual pitch contours by speaker 04f_J

indicate that the speaker produced prosodic phrasings as predicted based on the syntax-prosody mapping.

About the sentence prosody generated by Swedish learners of Japanese, five learners from the similar contours group and five learners in the BPM group produced similar contours for the four stimuli. These findings may indicate that either the learners did not recognise the syntactic differences among the stimuli, or they did recognise but failed to reflect the syntactic difference on the prosodic level. The rest of the four Swedish learners of Japanese produced individually unique contours, and it can imply the learners' attempt to reflect the syntactic difference on prosody. However, their production still fails to reflect the syntactic difference of sentences on prosody correctly. For instance, the pitch contours produced by speaker 01m_S show the difference between the [±Acc, *ga*] condition and [±Acc, *no*], although the contours failed to create different contours due to the absence and presence of pitch accent on the first words, *tomodachi*, and *ya'kuza*. Speaker 01m_S produced an unaccented word *tomodachi* as *tomoda'chi*. Because of that, the effect of downstep was produced in B1 which does not match the prediction based on the theories. Also, as to prosodic phrasing in A2, speaker 01m_S produced one AP and IP for the first two accented words. However, in Japanese one AP can have one accented word at most. In othwer words, prosodic phrasing for A2 by speaker 01m_S is against the rule in Japanese prosodic phrasing. The speech by speaker 12m_S lacks variety in the overall contours depending on the four stimuli and assigned BPMs on the nominative case marker *ga*, locative case marker *ni*, but not on the genitive case marker *no*. A possible explanation for this could be that speaker 12m_S seemed to recognise the genitive case marker *no* connecting nouns. Speaker 12m_S managed to reflect the syntactic structure by not putting BPMs on the genitive case *no*.

As to the findings that one Japanese native speaker (speaker 01f_J) and six learners of Japanese (speaker 02m_S, 04f_S, 06f_S, 07m_S, 09f_S, and 12m_S) assigned BPMs on the particles, this finding broadly supports the Tsurutani study (2011). Tsurutani also reported that English learners of Japanese assigned pitch rise at the end of accentual phrases. Another possible explanation for assigning BPMs by learners of Japanese is that learners of Japanese might hear native speakers' speech with rises at the end of APs and attempt to copy the features since it is commonly observed among the young generation of native Japanese speakers in the Tokyo area. Although BPMs are observed both in the speech of one Japanese native speaker and six learners of Japanese, the difference between them is that the native speaker assigned BPMs only when there are syntactic boundaries while the five learners of Japanese put BPMs on all of the particles. (speaker 12m_S is the exception). If these findings can be generalised, assigning BPMs is commonly observed both in native Japanese speakers and learners of Japanese. Native speakers of Japanese recognise the syntactic differences of sentences, and hence are capable of distinguishing where to put BPMs. In contrast, learners of Japanese seem to assign BPMs no matter what syntactic structure sentences have.

As far as prosodic phrasing is concerned, overall, Swedish learners of Japanese produced prosodic phrasing for the accentual phrase correctly based on the syntax-prosody mapping, but not intonation phrase. The lack of downstep is observed among most of the learners also supports the assumption that the Swedish learners of Japanese could not produce boundaries for intonation phrase in Japanese. Not only errors at prosodic phrasing but their errors at lexical pitch accent assignment also influenced whether downstep occurs at parts where it should occur. This finding is consistent with Tsurutani's view (2011) that learners' errors consist of the combination of errors at word and sentence level. Another explanation for the lack of downstep among the learners would be because downstep occurs under different conditions in Japanese and Swedish. The learners might have an ability to produce the effect of downstep, but do not know the condition of downstep to occur in Japanese. According to Gussenhoven and Bruce (1999), downstep is observed in Stockholm Swedish, but the condition for downstep to occur in Stockholm Swedish is related to focus, not in the same condition of downstep in Japanese.[4]

Another error is that their speech does not reflect the syntactic difference between [*ga*] condition where the structure is 'SUBJECT *ga* INDIRECT OBJECT *ni*' and [*no*] condition where the structure is 'SUBJECT (modifier noun *no* modified noun) *ni*.' Based on the theory of syntax-prosody mapping, [*ga*] condition is predicted to have two intonation phrases. In contrast, [*no*] condition is predicted in one intonation phrase. Most of the learners separated the first two phrases in [*no*] condition. Some of the learners producing this error can be explained as the insertion of BPMs. Another explanation would be negative transfer from their L1, namely Swedish, that the learners assigned H tone to mark phrasings as it is in Swedish.

5.2 RQ2: Can be the learners' errors explained as a transfer from their L1, Swedish?

The learners' errors at the lexical pitch accent level can be explained as a negative transfer from their L1. Specifically, most of their errors in the first words *ya'kuza* and *tomodachi* are *yaku'za* and *tomoda'chi*. These errors can be explained that the learners assigned stress on the penultimate syllable of the words as it is the default location of stress in Swedish. Frequent errors at the unaccented word *tomodachi* are predicted based on CAH and linguistic markedness as well.

[4] In Stockholm Swedish, when a focused word with focal H is followed by an unfocused word, it forms a high plateau. After the plateau, any lexical H is subjected to get the effect of downstep (Gussenhoven and Bruce 1999)

In terms of prosodic phrasing, assigning a H tone at the end of the accentual phrase is observed among the majority of the learners. Swedish prosodic phrasing that roughly corresponds to the accentual phrase in Japanese would be the prosodic phrase, which mostly corresponds to a syntactic phrase. According to Myrberg (2013), H tone is an obligatory part of focal accent. It appears on the last word of prosodic phrase it is in. Although it is hard to simply compare Swedish prosodic phrases and Japanese accentual phrases, one possible explanation for the learners' tendency to assign H tone would be negative transfer from Swedish way of marking prosodic phrase to Japanese accentual phrase. It cannot confirm that negative transfer from their L1 is the only reason to cause the error. Having said that, it can be said that transfer is one of the causes triggering the error along with the other factors such as the tendency of assigning BPMs, which is not necessary be defined as errors. A Japanese native speaker (speaker 01f_J) also shows the tendency of assigning BPMs. In terms of perception of these H-tones on the particles, the speech produced by the learners does not sound natural and it sounds like putting boundaries unnecessary. Speaker 01f_J also put BPMs on the particles, but BPMs in her speech are not perceived as intense as that of the learners and it does sound natural.

In sum, learners' errors at lexical pitch accent seem to be caused primarily due to negative transfer from their L1. In contrast, their errors at the phrasal level appear to be caused by the combination of negative transfer and other factors.

5.3 Summary of the discussion

According to the current discussion, this study has demonstrated that the variety among Japanese native speakers is all within the predictions based on the theory of syntax-prosody mapping and prosodic well-formedness. At the same time, the diverse contours produced by Swedish learners of Japanese cannot be explained by the theories. The learners' errors at the lexical pitch accent seem to be explained as a negative transfer from their L1. As to their errors in terms of prosodic phrasing, a negative transfer could be one of the causes along with the other factors, such as their errors at lexical level triggering errors at the phrasal level and the tendency of putting BPMs.

As to suggestion for future study, it would be better that future studies of analysing the errors by learners of Japanese would take into account of Japanese native speakers' speech as the product of the interaction between the syntax-prosody mapping and prosodic well-formedness instead of assuming that their speech does not have variety and use their speech as model speech. Consequently, prosodic errors by learners of Japanese may be analysed with a focus on how their speech deal with the syntax and prosody, but not with a focus on how much their speech deviates from Japanese native speakers' 'model' speech. Thus, this study suggests that prosodic errors by L2 learners of Japanese would be defined whether the errors can be explained by the syntax-

prosody mapping and prosodic well-formedness, instead of the deviation from the speech by Japanese native speakers.

In this study, Swedish participants were recruited from Stockholm and Lund University and they speak two different Swedish varieties: Rikssvenska and Skånska. Due to the small number of the participants, the result of this study does not support any correlation between their speech in Japanese and their dialects. However, if a future study with a larger number of participants with different dialects was conducted, it might be able to provide interesting insights on the relation between L2 languaage learners' dialect in L1 and their speech in L2.

What this study has demonstrated could also contribute to Japanese language education. Based on the findings, it seems that Swedish learners of Japanese tend to make errors and have difficulties as follows:

- Putting stress on a penultimate syllable in Japanese words, and associating it with a rise in F0
- Frequent errors at unaccented words due to an unnecessary pitch accent assignment
- Difficulty in differentiating and reflecting syntactic structure to prosody, especially at the level of intonation phrase
- The strong tendency of assigning a H tones to where it should have L tone as prosodic marking
-BPMs on particles

If I could suggest improving the learners' prosody, explicit instruction on the learners' tendency to assigning stress might help the learners' understanding of their own production, and it may help to decrease their errors at the lexical level. Also, when introducing different grammatical structures, the fact that syntactic difference influences prosody can also be introduced to the learners. Practicing how to reflect the grammatical structure on prosody could also help to improve their production.

5.4 Limitations of the study

One of the practical limitations of the study is related to the participants. First of all, there is a huge disparity in number of Swedish learners of Japanese and Japanese native participants (14 Swedish learners and 4 Japanese native speakers). Since the researcher lived in Sweden and studied at Lund University, it was practically difficult to recruit Japanese native participants. In

addition to that, there is a diversity in where each participant comes from within Sweden and Japan. Some of the Japanese native speakers had lived abroad or areas where other Japanese dialects are spoken. The reason why the participants who had lived places besides the Tokyo area were recruited is because their speech did not show dialectal characteristics as far as the researcher who is a native speaker of Tokyo Japanese perceived and the practical difficulty to recruit Japanese native speakers. It is possible that having the Japanese native speaker who had lived other areas in Japan or abroad caused influences in the data set that the researcher could not notice. As to Swedish participants, some of them had lived in different places within Sweden besides the Skåne or Stockholm area. It would have been ideal to have more consistency of where each participant group had lived in Sweden and Japan in order to minimise possible influences.

As to the data analysis, the researcher decided to conduct it by each speaker and provided the detailed description. This enabled us to capture the characteristics of each speaker such as the tendency of assigning BPMs in all of the four stimuli, highlighting that some speakers assigned BPMs on particular particles, but made exceptions of not putting BPMs. However, concerning the comparative aspect of the first research question, the data analysis could have been conducted by each stimulus type. With this way, the analysis could have compared between how Swedish learners of Japanese and Japanese native speakers produced for the same stimuli.

I also acknowledge that the consideration of the theoretical background about intonation of Swedish including the discussion of Swedish regional varieties is not as rich as that of intonation of Tokyo Japanese. Although my specialisation is Japanese and the main focus of the thesis is Japanese, having wider range of the theoretical discussion on Swedish could have enriched the discussion and even might have brought different insights of the data.

Lastly, this study could not identify what causes influenced to the learners' errors particularly at the phrasal and sentence level. What this study could do is at least discussing possible factors behind the errors. For instance, the learners' errors in the lack of downstep could be caused by their errors at the lexical level, the lack of their understanding of the condition for downstep in Japanese, or the learners had the understanding of when or where downstep is supposed to be produced, but failed to realised in intonation. In order to identify what cause come into play, more carefully controlled experiment setting would be necessary.

Chapter 6 Conclusion

This thesis has examined L2 intonation by Swedish learners of Japanese with a focus on pitch accent to answer the following two research questions:

Research question 1: How is the interaction of syntax-prosody mapping and prosodic-well-formedness realised in the utterances produced by Swedish learners of Japanese in comparison to Japanese native speakers?

Research question 2: Can their errors be explained as a transfer from their L1, Swedish?

Unlike Tsurutani's study (2011), the speech data from Japanese native speakers for this study shows diversity in terms of pitch contours. This study does not compare the native speakers' speech and the learners' speech with a focus on the deviation from the native speaker. Instead, this study analyses in terms of how their speech shows the interaction between the theories of syntax-prosody mapping and prosodic well-formedness. Based on that, the first research question has been formulated. The second research question aims at exploring the errors of Swedish learners of Japanese that has not been studied well in the previous study. Moreover, the second research question attempts to examine the theories proposed by the previous studies on L2 prosodic acquisition in the case of Swedish learners of Japanese.

Four stimuli in Japanese were prepared to extract the speech of Japanese native speakers and Swedish learners of Japanese. Recording sessions were conducted where the participants were asked to read the stimuli loud. The stimuli were carefully formulated and contain the two conditions, which are two types of syntactic structures and accentedness of the first word. The vocabulary in the stimuli was chosen within the level of learners.

The findings of this study show that four Japanese native speakers produced four different shapes of contours, and the data analysis has demonstrated that their speech is within the theories of syntax-prosody mapping and prosodic well-formedness. The combination of these yields the diversity of contours among the native Japanese speakers. In contrast, the varieties of contours produced by Swedish learners of Japanese seems not to be within the theories of syntax-prosody mapping and prosodic well-formedness. Furthermore, most of the Swedish learners of Japanese made errors at the lexical level, which seems to trigger other errors at the phrasal and sentence level. The learners' errors at the lexical level seemed to be explained as a negative transfer from

their L1, Swedish. At the phrasal level, the learners' errors seem to be caused by the combination of a negative transfer and the other factors discussed in section 5.2.

These findings demonstrate the importance of analysing the speech by Japanese native speakers and L2 Japanese learners with an assumption that the speech consists of the interaction between the theories of syntax-prosody mapping and prosodic well-formedness. This perspective can capture the complicated aspects of L2 learners' errors in prosody and could enable us to learn more effective ways of teaching Japanese prosody.

References

Anderson-Hsieh, J., R., Johnson, R and Koehler, K., (1992). The relationship between native speakers' judgements of non-native pronunciation and deviance in segmentals, prosody, and syllable structure. *Language Learning*, **42**(4), 529-555.

Baker, C., (2011). *Foundations of bilingual education and bilingualism.* 5th edition. Bristol: Multilingual matters.

Banno, E., Ikeda, Y., Shinagawa, C., Takashiki, K.,(2011). *Genki 2 textbook.* Tokyo: The Japan times.

Bruce, G., (1977). *Swedish word accents in sentence perspective.* Ph.D. thesis, Lund University.

Bruce, G and Hermans, B., (1999). Word tone in Germanic languages. In: H. van der Hulst. ed. Word prosodic systems in the languages of Europe. Berlin: De Gruyter. pp. 605-658.

Eckman, F., (1977). Markedness and the contrastive analysis hypothesis. *Language learning* **27**(2), 315-330.

Eckman, F., (1985). Some theoretical and pedagogical implications of the markedness differential hypothesis. *Studies in second language acquisition* **7**, 289-307.

Fery, C., (2016). *Intonation and prosodic structure*. Cambridge: Cambridge University Press.

Frid, J., (2001). Swedish word stress in optimality theory. *Working Papers*, (Department of linguistics and phonetics, Lund University) **48**, 25-40.

Gussenhoven, C., (2004). *The phonology of tone and intonation*. Cambridge: Cambridge University Press.

Gussenhoven, G and Bruce, G., (1999). Word prosody and intonation. In: H. van der Hulst. ed. *Word prosodic systems in the languages of Europe*. Berlin: De Gruyter. pp. 233-271.

Gårding, E and Lindblad, P., (1973). Constancy and variation in Swedish word accent patterns. *Working papers*, (Phonetic laboratory, Lund Univeristy) **7**, 36-110.

Hara, K., (1993). Iwayuru “shiriagari intonation ni tsuite: sono onkyotekitokucho to insho no kankei. [The acoustic feature and sensory impression of what is called "high-rising" intonation in Japanese]. *Gengokenkyu* [study of language and culture]**11**, 61-71.

Hirachi, O., (2008). Atarashi onseihyogen no kenkyu: akusento intonaishon niyoru jyohohuka ni chumokushite [the study of new sound expression: with focus on information conveyed by prosody]. *Bunkakankyokenkyu* [study of culture and environment] **2**, 30-39.

Hirose laboratory, University of Tokyo., (n.d.). *Online Japanese accent dictionary*. [Viewed 6 March 2020]. Available from: http://www.gavo.t.u-tokyo.ac.jp/ojad/swd/pages/home

Igarashi, Y., (2015). Intonation. In: H. Kubozono, ed. *Handbook of Japanese phonetics and phonology.* Berlin: De Gruyer. pp. 525-568.

Ishihara, S., (2015). Syntax-phonology interface. In: H. Kubozono, ed. *Handbook of Japanese phonetics and phonology.* Berlin: De Gruyter. pp. 569-618.

Ito, J and Mester, A., (2007). Prosodic adjunction in Japanese compounds. *MIT Working Papers in Linguistics 55: Formal Approaches to Japanese Linguistics* **4,** 97-111.

Ito, J and Mester, A.,(2012). Recursive prosodic phrasing in Japanese. In: T. Borowsky, S. Kawahara, T. Shinya and M. Sugahara, eds. *Prosody matters: Essays in honour of Elisabeth Selkirk*. London: Equinox Publishers. pp. 280–303.

Ito, J and Mester, A., (2013). Prosodic subcategories in Japanese. *Lingua* **124**(1), 20–40.

Jun, S., (2005). *Prosodic Typology: The Phonology of Intonation and Phrasing.* Oxford: Oxford University press.

Jun, S., (2014). *Prosodic typology 2.* Oxford: Oxford University press.

Kawahara, S., (2015). The phonology of Japanese accent. In: H. Kubozono, ed. *Handbook of Japanese phonetics and phonology.* Berlin: De Gruyter Mouton. pp.445-492.

Kawakami, S. (1957). Jun'akusento ni tsuite [On quasi-accent]. *Kokugo Kenkyu* [Study of the National Language] **7**, 21–31.

Kibe,Y., Takeda,N., Tanaka, Y., Hidaka, M and Mitsui, H., (2013). *Hogengaku nyumon* [the introduction to Dialectology]. Tokyo: Sanseido.

Kubozono, H., (1988). *The organisation of Japanese prosody*. Edinburgh: University of Edinburgh dissertation.

Kubozono, H., (1993). *The organization of Japanese prosody.* Tokyo: Kurosio.

Kubozono, H., (1997). Prosodic structure of Japanese and its Acquisition. *Onseigengoigaku* [medical language study]. **38**, 281-286.

Lado, R., (1957). *Linguistics across cultures: Applied linguistics for language teachers*. Michigan: University of Michigan Press.

Ladd, D, R., (2008). *Intonational phonology,* 2nd edition. Cambridge: Cambridge University Press.

Lanz, L.A., (2003). *Japanese pitch-accent: cross-linguistic perceptions by speakers of stress- and pitch-accent languages*. M.A. Dissertation, University of Hawaii.

Lieberman, P, (1967). *Intonation, perception, and language*. Cambridge: MIT press.

Liberman, M., (1975). *The intonational system of English*. Ph.D. thesis, Massachusetts Institute of technology.

Maekawa, K., Kikuchi,H., Igarash, Y. and Venditti, J., (2002). X-JToBI: An extended J-ToBI for spontaneous speech. In: *proceeding of 7th International Conference on Spoken Language Processing*, ICSLP September 16-20 2002, Denver, United States. United States: International Speech Communication Association. pp.1545–1548.

Major, R., (2001). *Foreign Accent*. New York: Routledge

Major, R., (2008). Transfer in second language phonology: A review. In: E. Jette. And M. Zampini, eds. *Phonology and second language acquisition*. Amsterdam: John Benjamins Publishing Company. pp. 63-94.

Max Planck Institute for Evolutionary Anthropology, Department of Linguistics., (n.d.). Leipzig Glossing Rules: Conventions for interlinear morpheme-by-morpheme glosses. [online]. *Max Planck Institute for Evolutionary Anthropology Department of Linguistics*. [Viewed 5 May 2020]. Available from: https://www.eva.mpg.de/lingua/resources/glossing-rules.php

McCawley, J, D., (1968). *The phonological component of a grammar of Japanese*. Berlin: De Gruyter Mouton.

Mennen, I., (2015). Beyond segments: towards to a L2 intonation learning theory. In: E. Delais-Roussarie, Avanzi, M and Herment, S, eds. *Prosody and languages in contact: L2 acquisition, attrition and languages in multilingual situations*. Berlin: Springer-Velag. pp. 171-188.

Mesthrie, R., Swann, J., Deumert, A., Leap, W. L., (2009). *Introducing Sociolinguistics*. Edinburgh: Edinburgh University Press.

Mouri, T., Hirose,K. and Minematsu, N., (2003). Consideration on vowel durational modification for Japanese CALL system. *Proceeding. EUROSPEECH,* 3153-3156.

Myrberg, S., (2010). *The intonational phonology of Stockholm Swedish*. Ph.D. thesis, Stockholm University.

Myrberg, S., (2013). Sisterhood in prosodic branching. *Phonology*, **30**(1), 73-124.

Myrberg, S. and Riad, T., (2015). The prosodic hierarchy of Swedish. *Nordic Journal of Linguistics*. **38**, 115-147

Ogasawara, N., (2015). Analysis of Japanese pitch accent in L1 and L2 speech. *Concentric: Studies in Linguistics.* **41**(2), 63-89.

Pierrehumbert, J., (1980). *The phonology and phonetic of English intonation*. Ph.D. thesis, Massachusetts Institute of technology.

Pierrehumbert, J and Beckman, M., (1988). *Japanese tone structure*. Cambridge: MIT Press.

Poser, W. J., (1984). *The phonetics and phonology of tone and intonation in Japanese*. Ph.D. thesis, Massachusetts Institute of technology.

Riad, T., (2013). *The phonology of Swedish*. Oxford: Oxford University press.

Sakamoto, E., (2010). *An investigation of factors behind foreign accent in the L2 acquisition of Japanese lexical pitch accent by adult English speakers*. Ph.D. thesis, The University of Edinburgh,

Sato, T., (1995). Comparison of the influence of segmental information and prosody on the assessment of Japanese pronunciation. *Sekai no Nihongo kyoiku* **5**, 139-154.

Selinker, L., (1972). Interlanguage. *International Review of Applied Linguistics*, **10**, 209–231

Selkirk, E. O., (1986). On derived domains in sentence phonology. *Phonology Yearbook* **3**, 371–405.

Selkirk, E. O., (2009). On clause and intonational phrase in Japanese: The syntactic grounding of prosodic constituent structure. *Journal of the Linguistics Society of Japan* **136**, 1–39.

Selkirk, E. O., (2011). The syntax-phonology interface. In: J. Goldsmith, J. Riggle. and A.Yu. eds. *The handbook of phonological theory*, 2nd edition. Oxford: Blackwell. pp. 435-484.

Selkirk E. O and Tateishi K., (1991). Syntax and Downstep in Japanese. In: C. Georgopoulos and R. Ishihara. eds. *Interdisciplinary Approaches to Language.* Dordrecht: Springer. pp. 519-543.

Shinya, T., Selkirk, E.O. and Kawahara, S., (2004). Rhythmic boost and recursive minor phrase in Japanese. *Proceedings of the Second International Conference on Speech Prosody*, 183–186.

Tsurutani, C., (2007). Acquisition of geminate consonants by English speaking learners of Japanese. *Review of Asian and Pacific Studies* **32**, 25-41.

Tsurutani, C., (2008). *Pronunciation and Rhythm of Japanese as a second language*. Hiroshima: Keisuisha Publisher.

Tsurutani, C., (2011). L2 intonation: A study of L1 transfer in Japanese intonation by English-speaking learners. *Second language*. **10**, 79-102.

Uwano, Z., (1999). Classification of Japanese accent systems. In: S. Kaji, ed. *Proceedings of the Symposium 'Cross-Linguistic Studies of Tonal Phenomena, Tonogenesis, Typology, and Related Topics'*. Tokyo: ILCAA. pp.151–178.

Uwano, Z., (1989). Nihongo no akusento [Accent in Japanese]. In: M. Sugito, ed. *Koza nihongo to nihongo kyoiku 2: Nihongo no onsei, on'in (jo)* [Japanese and Japanese teaching 2: Japanese phonetics, phonology 1], Tokyo: Meiji Shoin. pp. 178–205.

Venditti, J., (2005). The J_ToBI model of Japanese intonation. In: J. Sun-Ah, ed. *Prosodic typology: The phonology of intonation and phrasing*. Oxford: Oxford University press. pp.172–200.

Venditti, J., Maekawa, K. and Beckman, M, E. (2008). Prominence marking in the Japanese intonation system. In: S. Miyagawa and M. Saito, eds. *The Oxford handbook of Japanese linguistics*, New York: Oxford University press. pp. 456–512.

Wiener, S and Goss, S., (2019). Second and third language learners' sensitivity to Japanese pitch accent is additive: An information-based model of pitch perception. *Studies in second language acquisition.* **41**(4), 897-910.

Zampini, L. M., (2008). L2 speech production research. In: J. G Hansen, L. M. Zampini, eds. *Phonology and Second Language Acquisition*, Amsterdam: John Benjamins Publishing Company. pp. 219-249.

Appendix A Informed consent for Swedish participants

Form of informed consent

I hereby give my consent to participate in a study in linguistics in Lund University in Sweden.

This study is about the analysis of Japanese speech produced by Swedish learners of Japanese.

For this purpose, I will be presented stimulus sentences in Japanese and read them aloud for recording. Also, I will fill in a questionnaire asking about my language learning experiences and language proficiency. The recording data and the answers to the questionnaire will be used only for research purposes. All of the data obtained from this study will be presented in a way that the participants cannot be identified through the data.

By my signature below, I certify that:

• I understand that my participation is voluntary and I may withdraw at any time without giving any explanation.

• I have received assurance that the data will be dealt completely anonymous.

• I understand that the results of the study will be presented at scientific conferences and in journal articles without revealing the identity of the participants.

• I received sufficient information prior to the experiment. On completion, I will have the opportunity to ask questions.

• I will receive compensation after carrying through the whole experiment.

This consent form is signed in two copies, one for me and one for project documentation.

_____ _____

Place and date Signature

For more information, you may contact Natsumi Goto na3416go-s@student.lu.se

Lund University Master's programme in language and linguistic Japanese specialisation

Appendix B Informed consent for Japanese participants

日本語母語話者インフォームドコンセント

ここにルンド大学修士課程言語学に関する研究の趣旨を理解し、研究に参加することに同意します。

この研究はスウェーデン人日本語学習者の発話分析に関する研究です。日本語母語話者の発話データはスウェーデン人学習者の発話と比較するために使用されます。

この研究のために用意された日本語の例文を録音のために読み上げます。また年齢や性別、出身地に関する質問に回答します。録音データと質問への回答は研究目的のためだけに使用されます。また本研究は被験者が特定されないように論文においてデータを取り扱います。

下記に署名することにより、以下のことを承諾します。

• 本研究への参加は任意であり、研究のあらゆる段階において参加を辞退する権利があります。辞退する際、理由を研究者に述べる必要はありません。

• 私はこの研究で得られるすべてのデータにおける匿名性を保証する旨の説明を受けました。

• 私は研究の成果が学術雑誌や学会で発表される可能性があることについて理解し、承諾します。また発表される際も被験者の匿名性が保証されるという説明を受けました。

• 実験の前に私は十分な説明を受けました。また実験終了後にも質問をする機会があることを理解しています。

• 私は実験終了後、謝礼を受け取ります。

このインフォームドコンセントは二枚組です。一枚は被験者へ、もう一枚は研究者が保存します。

_____ _____

場所と日時　　署名

研究に関する質問がある場合は以下に連絡してください。na3416go-s@student.lu.se

ルンド大学修士課程言語学日本語専攻　後藤夏実

Appendix C Questionnaire for Swedish participants

Questionnaire for Swedish students

Name _____

University/ Major / Year University: _____ **Major:** _____ **Year:** _____

Native languages you or your family use _____

Age when you started learning Japanese _____

(If you started learning Japanese but stopped learning for some period of time but started learning again, please write down years of studying and stopping and a year you started learning again

Ex) 2007-2010, 2012-up to now

Have you stayed in Japan for sightseeing? Please choose : yes /No

If Yes, please write down your three recent visits if you can remember

How long	**When**	**Places you visited**
Ex) 3 weeks	Last year	Kyoto, Osaka, Kobe

Have you visited Japan for study abroad or taking languages courses in Japan? (Including school trip to Japan) Please circle: Yes /No and write down the details

How long	**Where you studied**	**When**	**What you did during the stay**
Ex) 3 months	Tokyo University	Last year from September	Took language classes every day and cultural activity

How often do you use Japanese besides your university study?

Ex) reading manga or novels in Japanese, watching drama, anime, or films in Japanese, talking with friends or partners in Japanese, speaking Japanese at work, listening to Japanese music and so on. Please write down the details

What you do	How often

How often do you use Japanese with native Japanese speakers?

Only in classroom, with friends or partner or family? Please describe below

Have you studied other foreign languages? Please choose : <u>Yes /No</u>

If Yes, please fill in below

Languages	When	How long
Ex) Korean	2010-2012	2 years

Do you have any Japanese language certificate? ex) JLPT test, Kanken 漢検 etc <u>Yes/No</u>

If yes, please describe in details below

Test	When you took	Score/Grade

ありがとうございます！