PROCESSING OF INTONATION PATTERNS IN JAPANESE: IMPLICATIONS FOR JAPANESE AS A FOREIGN LANGUAGE

DISSERTATION

Presented in Partial Fulfillment of the Requirement for
The Degree Doctor of Philosophy in the
Graduate School of The Ohio State University

By
Sanae Eda, M.A.

*****

The Ohio State University
2004

Dissertation Committee: Approved by
Dr. Mari Noda, Adviser
Dr. Mary Beckman
Dr. Mineharu Nakayama

__________________________
Adviser
Department of East Asian
Languages and Literatures
ABSTRACT

Accent and intonation in foreign language instruction tend to be treated in a haphazard manner for two reasons: lack of research analysis that can be applied to the second language acquisition (SLA) situation; and absence of pedagogical materials that effectively train learners on this aspect of the target language. Japanese language instruction at all levels will benefit from both deductive and inductive instructions on accent and intonation at the level of utterance and discourse. Such a seemingly challenging task can be accomplished by using a computerized exercise that involves the effects of various accent and intonation patterns.

The experimental study presented in this dissertation examined the processing of three types of prosodically marked syntactic and pragmatic contrasts by native and non-native speakers of Japanese. Three results are salient. First, there was a statistically significant difference between the performance of native and non-native speakers in both the discrimination and the interpretation tasks. Second, while non-native speakers performed equally well in the discrimination task regardless of their proficiency, their performance in the identification task was strongly predicted by their proficiency. Finally, native speakers of the Tokyo dialect and speakers of other dialects performed differently on the interpretation task.
As an extension of the investigation, I propose ways to make the topic of accent and intonation a part of the curriculum at all levels of instruction. I will also demonstrate a computer program currently being developed to help Japanese students improve their communicative effectiveness by directing their attention to the effects of variation in accent and intonation patterns.
Dedicated to my father, mother, and sister.
ACKNOWLEDGMENTS

First of all, I would like to express my wholehearted thanks to my adviser, Professor Mari Noda. I met her in the summer of 1991, the first summer I spent in the United States, in a training program for teaching Japanese as a foreign language. This encounter had a profound impact on me, and I came to the Ohio State University in 1994 to be able to study with her. I cannot believe it has already been ten years! Arigatoo gozaimasita, Noda-sensei, for everything you have taught me and shared with me.

Thanks are also due to the committee members, Professor Mary Beckman and Professor Mineharu Nakayama. I am grateful not only for their extremely insightful and constructive comments on the subject matter, but also for their tireless support, encouragement, and patience. They read my chapters thoroughly and sent comments very quickly. I simply cannot find enough words to express my thanks.

I owe much thanks to Debbie Knicely. Without her friendly reminders, I would have missed so many deadlines. Especially during the final year when I lived away from campus, she kindly offered to make frequent trips to University Hall to turn in forms and take care of necessary business on my behalf. Xie xie, Debbie-san!

One of the experiments was supported in part by grants to the author by the following Ohio State units: the Center for Cognitive Science, the Department of East
Asian Languages and Literatures, and the Department of Linguistics. Support was also provided to Professor Mineharu Nakayama by the Office of Research. I thank Dr. Eleanor Jordan, Dr. Hideki Maki, Koichi Sawasaki, Misako Terashima, and Yuko Kuwai for their assistance in data collection.

Thanks are also due to Kiyoko Yoneyama and Hiroko Joy (Niioka) for their assistance in recording the stimuli. Satoko Katagiri and Pauline Welby gave constructive comments on early versions of the study. Brook Oberst, Cathy Rosen, and Tom Savas proofread the final manuscript.

I am indebted to Professor Mari Noda, Abhijit Varde and Jeff Borisch, who belong to a team developing a multimedia edition of *Japanese: the Spoken Language*. The design of the computer module for learning accent and intonation of Japanese presented in this dissertation has been greatly inspired by their comments.

I would also like to thank my family—my late father, mother, and sister, for their warm encouragement and support. I learned the joy of life from my family after my father was diagnosed with cancer and spent the last three months of his life in a hospice. He was full of positive energy until the last moment of his life and encouraged me to continue working on my degree; that I have been able to bring this work to a meaningful conclusion is especially gratifying because of him.

And, lastly, but certainly not least, I would like to thank my husband, Wen-shiang Chen, for his loving support. As I write this, he is next to me putting the finishing touches on his dissertation. I remember thinking I would be finished two or three years before him; now I think to myself, what happened? … and, isn’t he so amazing???
VITA

April 18, 1968 ..................... Born – Okayama, Japan

1991 ............................. B.A. English Education, Hiroshima University

Summer, 1991 ...................... Summer Intensive Training Program on the
  Teaching of Japanese as a Foreign Language,
  Exchange: Japan Program

1991 – 1994 ....................... EEP Lecturer / Graduate Teaching Assistant,
  University of Puerto Rico

Summers, 1992, 1994 ............ Instructor of Japanese, Language: Japan Program
  Gifu Prefecture, Japan

1994 ............................. M.A. English Education, University of Puerto Rico

1997 ............................. M.A. Japanese Pedagogy, The Ohio State University

Summers, 1995 – 2002 ............ Instructor of Japanese, Middlebury College
  Intensive Summer Program, Japanese School

1994 – 2003 ....................... Graduate Teaching and Research Associate
  The Ohio State University

2003 – Present .................... Visiting Instructor of Japanese
  Valparaiso University
PUBLICATIONS


FIELDS OF STUDY

Major Field: East Asian Languages and Literature
# TABLE OF CONTENTS

Abstract ................................................................. ii

Dedication .............................................................. iv

Acknowledgments ....................................................... v

Vita ........................................................................ vii

List of Tables ................................................................ xii

List of Figures .......................................................... xiii

Chapters:

1. INTRODUCTION ......................................................... 1

   1.1 Basic Concepts ......................................................... 2
      1.1.1 Goal of learning and teaching Japanese as a foreign language .... 2
      1.1.2 Intonation and lexical accent ........................................ 3

   1.2 Motivation .............................................................. 5

   1.3 Five inter-related areas of concern that are fundamental to the learning and teaching of less commonly taught languages (LCTLs) in the United States ................................................................. 7
      1.3.1 Lifelong language-learning career .................................. 8
      1.3.2 Expertise .................................................................. 9
      1.3.3 Culture-based learning and teaching ............................ 9
      1.3.4 Responsibility ......................................................... 10
      1.3.5 Local conditions .................................................... 11

   1.4 Organization of this dissertation .................................... 11

2. REVIEW OF LITERATURE ........................................... 13
2.1 The acquisition of intonation: comparison between first language acquisition and second (adult) language acquisition................................. 13
  2.1.1 Acquisition of intonation in children’s first language acquisition 17
    2.1.1.1 Production data: Acquisition of emphatic stress..................18
    2.1.1.2 Comprehension data: Early experiments.........................19
    2.1.1.3 Comprehension data: Focus on the comprehension of the pragmatics of intonation ........................................22
    2.1.1.4 Criticism on the methodology........................................26
  2.1.2 Acquisition of intonation by adult foreign language learners ...... 27
  2.1.3 Comparison between the two acquisition processes: Children acquiring their native language vs. adults learning a second (foreign) language ........................................................................ 30
  2.1.4 Acquisition of intonation by native speakers of other dialects of Japanese ................................................................................................................................................. 32
2.2 Limitations of teaching intonation of Japanese within the framework of “pronunciation” practice: old theories and practices ......................... 34
  2.2.1 Issue number one: too much emphasis on the consonants and vowels? .......................................................................................... 35
  2.2.2 Issue number two: pronunciation training takes up too much time? .........................................................................................................................37
  2.2.3 Issue number three: pronunciation taught at the beginning of the course … and that’s it? ............................................................................40
  2.2.4 Issue number four: it focuses on production only........................ 41

3. INTONATION SYSTEM OF JAPANESE LANGUAGE .......................... 43

3.1 The prosodic hierarchy in Japanese .................................................. 44
  3.1.1 Pitch Accent .................................................................................. 44
  3.1.2 Effect of phrasing: Syntactic function....................................... 47
  3.1.3 Effect of phrasing: Pragmatic Function ..................................... 50
3.2 Inventory of boundary pitch movements ......................................... 53
  3.2.1 Analysis of semantic properties associated with boundary pitch movements ............................................................... 55
  3.2.2 Contrast between InfoQ and IncredQ rises ......................... 60
  3.2.3 Contrast between InfoQ and Insisting rises........................ 63
  3.2.4 Rationale for the selection of experimental focus.................. 65

3.3 Interaction between boundary pitch movements and sentence final particles ......................................................................................................................... 65
  3.3.1 Further examination of SFP ne ............................................... 68
  3.3.2 Further examination of SFP yo ................................................. 75
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>A syllabus of a beginning level Japanese language course, utilizing the hiragana chart as an introduction to pronunciation</td>
</tr>
<tr>
<td>3.1</td>
<td>Summary of the semantic properties associated with the five boundary pitch movements presented in Figure 3.4</td>
</tr>
<tr>
<td>5.1</td>
<td>Activities for practicing intonation contrast using the sentence <em>soo desu ka</em> as an example</td>
</tr>
</tbody>
</table>
**LIST OF FIGURES**

3.1 Schematic representation of utterances with differing accent specifications ................................................................................................................... 46

3.2 The effect of downstepping in an utterance consisting of four accented accentual phrases ................................................................................................. 49

3.3 The effect of the placement of an intonation phrase boundary ...................... 49

3.4 An utterance consisting of four accented accentual phrases with pragmatic focus on the second accentual phrase, produced as a response to the question *Mari-chan wa nitiyoobi ni oniisan ni au n desu ka?* “Will Mary meet her elder brother on Sunday?” ......................................................................................................................... 52

3.5 An utterance consisting of four accented accentual phrases with pragmatic focus on the third accentual phrase, produced as a response to the question *Mari-chan wa nitiyoobi ni oneesan ni au n desu ka?* “Will Mary meet her elder sister on Saturday?” ......................................................................................................................... 52

3.6 Schematic representation of BPMs examined and characterized by Venditti and her colleagues (taken from Venditti, Maeda, and van Santen, 1998, p. 2) .............................................................................................................................. 56

3.7 Mean standardized (z) scores for each of the BPM types, on 8 semantic scales (taken from Venditti et al. (1998), p. 2) ........................................................................ 57

3.8 Pitch Contour of the Utterance with Information Question Rise .............. 62

3.9 Pitch Contour of the Utterance with Incredulity Question Rise................ 62

3.10 Pitch Contour of the Utterance with Information Question Rise (2) ........ 64

3.11 Pitch Contour of the Utterance with Insisting Rise ................................... 64

3.12 An utterance with √ intonation + *ne* .......................................................... 70
3.13 An utterance with ! intonation + ne.................................................................70

3.14 Interaction between sentence final particle ne and boundary pitch movement, reproduced from Sugito (2001), p. 11 .............................................74

3.15 An utterance with ✓ intonation + yo...............................................................76

4.1 Comparison between native speakers of Tokyo dialect and speakers of other dialects in the discrimination task .................................................86

4.2 Comparison between native speakers of Tokyo dialect and speakers of other dialects in the interpretation task ..................................................87

4.3 Comparison between native speakers and non-native speakers of Japanese in the discrimination task .............................................................89

4.4 Comparison between native speakers and non-native speakers of Japanese in the interpretation task ............................................................90

4.5 Non-native speakers’ performance in the discrimination task plotted against their Japanese proficiency test (JSKIT) scores ........................................92

4.6 Non-native speakers’ performance in the identification task plotted against their Japanese proficiency test (JSKIT) scores .....................................93

4.7 Non-native speaker scores in discrimination (top) and identification (bottom) tasks, plotted against their Japanese proficiency test (JSKT) scores ..................................................................................96

4.8 Percent correct discrimination of falling, insist rise (CC) and incredQ rise (QQ) in “same” condition .................................................................97

4.9 Percent of correct discrimination in conditions in which the tokens were presented as intonational minimal pairs .................................................98

5.1 Menu screen of the Accent and Intonation Module in the multimedia edition of Japanese: The Spoken Language .................................................130

5.2 A screen capture of the Intonation Module for a perception exercise .......132

5.3 A screen capture of the Intonation Module for a production exercise .......134
CHAPTER 1

INTRODUCTION

This dissertation explores two important and under-researched topics in second language acquisition: how speakers of Japanese perceive and process variation in intonation to disambiguate pragmatic and syntactic structures of utterances; and how native and non-native speakers of Japanese acquire the ability to process such cues. Two experiments were designed and conducted to examine the perception and control of intonation by native and non-native speakers of Japanese. A quantitative analysis of the data suggests that the acquisition of intonation involves two related yet distinct cognitive processes: the ability to perceive relevant differences and similarities in the pitch contours of utterances (phonetic level processing) and the ability to associate the pragmatic intentions of the speaker with the perceived pitch contours (pragmatic level processing). The data also implies that there is a gap between the acquisition of the first and second ability and also that the gap between these two levels of acquisition exists in parallel both in native language and second language acquisition.

More long-term objectives of this investigation are to provide practical suggestions for improving Japanese language curriculum design in order to better assist learners in becoming able to interact in and with the culture; and for developing
instructional materials that will most effectively facilitate this process. The scope of the present study will be limited to the context of teaching college level Japanese to native speakers of English in the United States. This necessarily implies that this study focuses on the learning process for cognitively mature learners with American cultural and linguistic background, embedded in a formal instructional environment. Furthermore, the discussion will be limited to the analysis of, and acquisition of intonation of, the so-called Standard Japanese, the variety of Japanese spoken as a first dialect by Japanese native speakers in the Tokyo area and only as a second or third dialect by speakers elsewhere in Japan.

1.1 Basic Concepts

1.1.1 Goal of learning and teaching Japanese as a foreign language

In accordance with the idea set forth in *A Framework for Introductory Japanese Language Curricular in American High Schools and Colleges* (Unger, et al., 1993, hereafter, the Framework), this study assumes that “the purpose of learning Japanese is … to become an informed foreigner who can function in Japanese society in a way that does not make Japanese feel uncomfortable or otherwise impedes the attainment of practical goals.” In other words, the purpose of learning Japanese is not to become a “native” in that culture but to become able to appropriately interact “in and with the culture” (Walker and McGinnis, 1996, p.5). The purpose of language instruction is defined in this dissertation based on this goal of language learning so as to provide the guidance necessary to facilitate this process. Placing emphasis on the performance of
culturally coherent behavior as the goal of learning has significant implications for instruction on intonation as well.

1.1.2 Intonation and lexical accent

Features superimposed on a string of smaller units of sound, are “characterized by the fact that they must be described in relation to other items in the same utterance” (Ladefoged, 1993, p. 15). The suprasegmental features such as pitch (whether the voice is perceived as high or low), length (duration), and intensity (degree of respiratory energy) are combined and perceived by the listener in terms of the rhythm and melody of the utterance.

The term “intonation” has been used in various senses in the fields of Japanese linguistics and Japanese language pedagogy (cf. Inoue, 1992 for further discussion on this point). In the narrowest sense, the term “intonation” is used specifically to refer to phrase final pitch configurations, which are observed in terms of rising or falling of fundamental frequency (f0\(^1\)) contours (cf. Ladd, 1992; Inoue, 1992; Ogawara, 1997). At the other end of the continuum, ‘intonation’ is used as a cover term that refers to all prosodic aspects of an utterance other than that of pitch accent (Fukuda, 1996). Such a broad definition of intonation will include aspects of speech such as the voice quality, which might indicate the speaker’s gender and age (Fukuda, 1996), speed of utterance, as well as duration of pauses between utterances (Higuchi, 1997; Imanishi & Akiyama, 1992; Sugito, 1989a).

---

\(^1\) Pronounced as “ef-zero.”
The definition of intonation in this dissertation will be something in-between these two ends of the continuum. I will use the term ‘intonation’ to refer to the prosodic ‘melody’ of an utterance that can be measured in terms of the pitch movements and communicates the speaker’s emotional state and attitude toward the topic under discussion (Maeda & Venditti, 1998; Venditti, Maeda, & Van Santen, 1998), pragmatic focus of an utterance (Sugito, 1989b), and provides cues for underlying syntactic structures (Kubozono, 1987, 1992; Misono, Mazuka, Kondo, & Kiritani, 1997; Uyeno, Hayashibe, Imai, Imagawa, & Kiritani, 1981; Venditti, 1994). According to this definition, the discussion of ‘intonation’ presented in this dissertation will include both the control of pitch range within an utterance (phrasing) and the examination of its phrase final pitch configuration. A more specific term, Boundary Pitch Movement (BPM), will also be used to refer specifically to the phrase final pitch configuration (Maeda & Venditti, 1998; Venditti et al., 1998).

In order to understand the creation of the melody of an utterance, one needs to also learn the function and realization of lexical accent in Japanese. Accent is an abstract marking singling out a particular phonological unit within a word so as to make that unit behave in a linguistically prescribed manner with respect to some aspect of the intonation contour. In Japanese, each lexical item is specified as accented or unaccented as a part of its lexical property.\(^2\) An accented word has at most one bitonal pitch accent associated with the lexically accented syllable, and this is realized as a sharp pitch fall. An unaccented word does not have any pitch accent associated with it. Accent in Japanese is

\(^2\) A detailed definition of lexical accent is provided in Chapter 3.
reflected in the pitch contour of lexical items, and thus participates in the creation of the intonation of the utterance. The locations of lexical accent is not predictable and they must be learned for each word as a part of its property (Tsujimura, 1996, p. 74). The subsequent discussion of intonation will consider the accuracy of phonological realization of lexical accent as an integral part of accuracy in intonation.

Throughout this document, the examination of intonation is limited to that of standard Japanese, spoken in the Tokyo area.

1.2 Motivation

In recent years, an increasing number of researchers have investigated the roles of prosody in spoken language communication. Intonation, or the ‘melody’ of sentence created by the variation in pitch modulation, as defined in the previous section, plays important roles in speech communication in both comprehension and production.

How would prosody facilitate more effective communication for learners of Japanese as a foreign language, then? This may appear to be a basic and straightforward question. However, in a narrow sense of communication, that is, if successful communication is defined as “getting the literal message across,” the prosodic differences may be treated as small differences in nuances. Considering the challenges of covering a large amount of material within a given amount of time, the answer to the above question of “how important are the communicative functions of intonation?” may at first glance seem to be “not very important.”

The acquisition of intonation is relevant and important to communicative success in the context of Japanese as a foreign language in many senses. First, and most
importantly, accurate use of intonation should be a part of the pragmatic competence of any Japanese language user, and learners of Japanese should be given opportunities to develop, and receive assistance in developing such competence. Walker and Noda (2000) discuss the implications of language as “performed culture.” If the goal of learning Japanese is to become able to function well in the target culture and to use the forms and communicative strategies of the target language, languages users must know the features of what Walker and Noda term “performance – when and where something is said or written, who initiates and receives the message, what the linguistic inventory is, and who might be observing” (p. 199). Examined in this framework, communication is not construed as simply “getting the literal message across.” Language learning, likewise, is not limited to the acquisition of lexical and syntactic knowledge, i.e., the list of rules for producing well-formed utterances that allows the learner to manipulate linguistic forms. Thus, it is of great importance that the language curriculum help the learners develop competence in both perceiving the implications of and producing intonation from the very beginning of their learning because learners need to understand the basic organizational structure of tonal events in Japanese in order to comprehend and utilize intonational variation as a communicative tool.

In addition to increasing the effectiveness of communication, acquisition of intonation will play a role in improving the perceived naturalness of the learners’ speech. Some studies on the production of Japanese by non-native learners have shown that prosodic qualities of their utterances had significant influence on the evaluation of ‘naturalness’ of speech (Mitsui, 1994; Sato, 1995; Suzuki, 1992). The results of
experimental studies such as those of Sato (1995) and Suzuki (1992) showed that prosodic cues, such as intonation, stress, and rhythm, had a stronger effect on the evaluation of naturalness of L2 Japanese than the segmental cues, such as the quality of distinct phonemes. Among the three prosodic cues, change of pitch patterns, intensity, and length of segments, pitch was found to have the most significant influence on the judgment of naturalness in both studies, the influence of the other two prosodic cues, i.e., intensity and the length of segments, being statistically non-significant. Mitsui (1994) concluded that intonation and timing of utterances were as important factors which should be targeted in instruction in order to increase the naturalness of Japanese spoken by learners of Japanese who had lived in Japan for one year as exchange students.

To date, little empirical study has attempted to characterize how native and non-native speakers of a language acquire the ability to perceive the differences in meaning conveyed by such intonational variation. The area of acquisition of intonation, therefore, deserves more attention. The current study examines how native and non-native speakers of Japanese learn to perceive the contrasting intonation patterns of Japanese and to accurately interpret their implication for utterance meaning.

1.3 Five inter-related areas of concern that are fundamental to the learning and teaching of less commonly taught languages (LCTLs) in the United States

Given the need to teach intonation, there is the question of how instruction and learning relate to each other, especially in the case of Japanese. Walker and McGinnis (1995) provide some guiding framework for the foreign language curriculum. While their primary focus is on the learning and teaching of less commonly taught languages
(LCTLs), and Japanese has since become one of the more commonly taught languages in the US, their framework continues to apply to Japanese (as well as to most language learning). In characterizing learning of a foreign language, they discuss five critical areas of concern.

1.3.1 Lifelong language-learning career

Walker and McGinnis (1995) assert that “learning to function more successfully within a culture, be it native or foreign, is a lifelong learning career” and that “the proper role of formal instruction is to enhance and sustain that career.” This means that, while only a limited time is available for formal instruction, learners will have to manage on their own lifetime’s worth of practice outside the classroom both during their enrollment in language courses and after completing a set of courses. Therefore, we need to take into consideration methods to help equip them with the tools they need to continue improving their communicative skills beyond the classroom.

How is it possible, then, to help learners become capable of successfully managing the time available to them for practicing the language beyond formal instruction: resolving the following two needs will contribute to answering this larger question.

First we need to identify what needs to be dealt with as a part of formal instruction. Second, we need to define how much time is available and use this to prioritize the topics to be presented formally as a part of the instruction. Once we have these two pieces of information, an appropriate amount of time can be allotted to each topic. How to manage the severe limitation of time is the key concern in teaching any
language in a formal setting. This exercise can be applied to identify how to incorporate intonation in the curriculum of instruction.

### 1.3.2 Expertise

“At any time, learners of a language are somewhere on a continuum from novice to expert either in general ability in the language or in a specified set of skills in the language” (Walker and McGinnis, p. 2). This implies that we should expect different performance from novices and experts. The learner’s knowledge about and ability to utilize intonation is not an exception: what novices are capable of handling will be different from what experts can do. This also implies that exercises designed for novices should be different from those designed for experts.

The questions that need to be answered in order to appropriately address this issue of expertise will include: what is a novice performance in the area of intonation? What is an expert performance? What kinds of exercises will be appropriate to assist the learners at different proficiency levels?

### 1.3.3 Culture-based learning and teaching

As discussed above, the intention of learners of less-commonly taught languages (LCTs) is to function in the LCTL culture. This being the case, Walker and McGinnis observe that, “the broader enterprise is that of intercultural communication and the development of linguistic skills is incidental to it” (Walker and McGinnis, 1995, p. 3). In other words, learning a LCTS involves more than the examination of linguistic structures of the language and memorization of patterns. Let us define the ultimate object of inquiry in learning LCTL as “meaning.” Meaning is not contained within a language; rather,
language is only one of many mediums with which learners can negotiate the meaning in an understood context in which communication takes place.

Communication “only works within an understood context, and a culture is the broadest identifiable context within which communicative abilities and language abilities operate” (Walker and McGinnis, 1995, p. 3). Language teachers need to be reminded of this when designing exercises for language practice.

1.3.4 Responsibility

“Learners will never learn enough in a classroom to finish a course successfully, and will never learn enough in a course to negotiate a foreign culture successfully” (Walker and McGinnis, 1995, p. 4). From early in their careers as learners of LCTLs, they need to be made responsible for managing their own learning. Expert teachers are, in contrast, responsible for guiding learners toward efficient management of learning, in addition to providing learners with the optimal conditions for learning. The responsibility of language teachers is to identify what will provide learners with the optimal conditions for learning within a course and for continuing to learn after they have completed the course.

Another issue that needs to be addressed is the degree of acceptance of foreign speakers’ performance in Japanese. There is a tendency in Japanese society to think of foreign-sounding Japanese as “cute.” This is readily apparent if one considers the type of non-native speakers who appear on TV talk shows—the more extreme their foreign
accent, the more popular these people are. On the other hand, if the intention of a learner of Japanese is to be a professional, then s/he needs to be trained accordingly. It is the responsibility of language instructors to inform the learners of the possible choices, the “prices” they will pay, and the consequent “gains.” In turn, learners are responsible for making informed choices.

1.3.5 Local conditions

The discussion in this dissertation is based on the recognition that “every locality has unique conditions” (Walker and McGinnis, 1995, p.4). Thus, the discussion presented in this dissertation on the teaching of intonation in general, and any specific suggestions will be successfully applied to the teaching only by adapting them to local conditions and by taking into consideration the needs that are specific to the locality.

1.4 Organization of this dissertation

The second chapter will review literature in the field of acquisition of intonation. I will also review current practice in the teaching of Japanese pronunciation and examine how prosody (accent and intonation) is treated. I will briefly explore issues of the use of computer in the context of language instruction. In order to discuss how to teach

3 A friend of mine shared with me an interesting observation he made. He went to Japan through the Japan Exchange and Teaching (JET) Programme, and taught English in a small town in Fukuoka Prefecture. When he arrived in the town, he was a “typical American,” as he referred to himself: he was a tall guy with blonde hair and blue eyes, and he did not speak a word of Japanese. His friends were very eager to teach him some Japanese words and phrases. When he started to feel comfortable to ask simple questions in Japanese, however, his friends started discouraging him from learning Japanese, which was a surprise to him. His friends would tell him he “should not bother to learn Japanese because it is too difficult.” In addition, when his friends (native speakers of Japanese) would teach him some phrases, they would pronounce Japanese so that it sounded like Japanese spoken by native speakers of English—not in the way they would themselves say it as native speakers of Japanese. It was clear to him that his friends wanted him to remain an American in terms of both physical appearance and linguistic ability (inability?).
intonation, we also need to have a good understanding of what we are trying to teach. The third chapter will, therefore, summarize the characteristics of the prosodic organization of well-formed Japanese utterances, focusing on the aspects that are relevant to the teaching of Japanese language to native speakers of American English. The fourth chapter presents an experimental study which examined the perception of intonationally contrasting utterances by native and non-native speakers of standard Japanese. The implications of the data obtained by the experiment to the teaching of Japanese intonation will be discussed. The final chapter will discuss pedagogical implications of the experimental study and suggest ways to address the acquisition of intonation in the teaching of Japanese language. Finally, suggestions for further studies will be made.
CHAPTER 2

REVIEW OF LITERATURE

2.1 The acquisition of intonation: comparison between first language acquisition and second (adult) language acquisition

This section will review the experimental literature relevant to understanding the acquisition of intonation. The review is not limited to Japanese intonation patterns; studies concerning both first language acquisition processes in children, and later language acquisition processes in adult foreign language learners will be included. I will then discuss the implications of the results of these experiments, addressing questions such as: How does the acquisition process compare at the various stages of acquisition? What similarities and what differences may exist between the two acquisition processes?

The acquisition of intonation both by children and by adult language learners is an area that has not been systematically studied to date in the sense that the available studies in the literature have not been performed with scientifically sound methodologies. Furthermore, to my knowledge, no existing studies have examined the acquisition of intonation with a focus on the acquisition of its pragmatic roles specifically in the
Japanese language but rather have tended to concentrate on English learners. Therefore, this set of questions will be answered based mostly on a review of literature that has examined the acquisition of English intonation. In assessing the implications of these experiments, emphasis will be placed on uncovering methodological inconsistencies in order to be able to perform a meaningful comparison of the results.

It is important to keep in mind that there are various differences between the acquisition of native language(s) by young children and the acquisition of second or foreign language(s) by cognitively mature adult learners. For instance, the process of native language acquisition takes place while the children are developing their more general cognitive capacity. In contrast, second and foreign language learners are already cognitively mature. Another important area in which native and second language acquisition processes differ is that adult second language (L2) learners are already a fluent native speaker of their first language (L1). Thus the examination of L2 acquisition needs to be concerned also about the possible influence of their L1 on the acquisition process. When there are similarities between their L1 and L2, the process may be greatly enhanced. When there are differences, the process may be hindered by such differences. For a complete review of the differences between L1 and L2 acquisition processes, see Sawasaki and Nakayama (2001).

Acquisition of intonation is built on something that is fundamental and universal, as the dimensions underlying the acquisition of prosodic controls are often simultaneously involved in various other contrasts that are fundamental to the
vocalization mechanism (Vihman, 1996). Such acoustic variables include, for instance, the control of pitch, duration, and intensity or amplitude.

In particular, there are two aspects of intonation that have been pointed out as universal across languages. The first aspect is the control of voice quality and pitch range manipulation for “paralinguistic” signaling of involvement and affect. Generally, an expanded pitch range is used to convey involvement and a raised pitch range to convey affect. The latter is often accompanied by soft, breathy voice qualities. In contrast, a narrow pitch range and pressed voice qualities can be associated with disapproval. It is not surprising, therefore, that this basic aspect of intonation is often exaggerated in speech directed to children, especially by caretakers addressing infants.

Another aspect of intonation that has been pointed out as universal across all languages is that it is used as a mechanism for grouping elements of continuous speech signals into clauses. A large number of researchers have examined the proposed “universals” on how prosodic units map onto syntactic units (Morgan & Demuth, 1996). Recognition of this universal aspect has led researchers to speculate that information available in adult speech may “contain clues to certain fundamental syntactic distinctions, … signaling the domains within which such analyses may be efficiently deployed” (Morgan, 1996, p. 2). Examination of infants acquiring Japanese as the native language, in this particular framework of analysis, has provided additional insight because the syntactic structure of Japanese and English are characterized by different parameters, particularly in their “Head Direction” and “Branching Direction” (Mazuka, 1996). Further, the view that cues to syntactic structure may be found in the prosodic
organization of adult speech has motivated researchers to examine the correspondence between prosodic groupings of utterances with their syntactic structures (Cutler, 1996), to question whether prosodic cues that closely correspond with underlying syntactic structures can serve as a bootstrapping mechanism (Mazuka, 1996; J. J. Venditti, S.-A. Jun, & M. E. Beckman, 1996) and to attempt to identify salient prosodic cues by paying close attention to the speech directed to infants (Fisher & Tokura, 1996).

Does this mean, then, that the process of the acquisition of intonation should be effortless because intonation has features that are universal and fundamentally the same across all languages?

While there is something universal about the voice qualities and the raised pitch range, languages differ dramatically in inventory of pitch movements at the phrase boundaries (boundary pitch movements, BPMs, hereafter) and meanings associated with them. For example, there is more than one type of rising intonation in Japanese, and one particular type if rising PBM indicates an emphatic assertion, not a question. This distinction certainly does not apply to English. (See Chapter 3.2 for a more detailed discussion on this point.) Likewise, while the chunking function in intonation seems to be universal and all languages that have been studied in a detailed way seem to have some prosodically marked unit whose boundaries tend to correspond to clauses, the relationships between these specific units in prosodic hierarchy, such as an “intonation phrase,” and smaller and more basic units of sounds, such as a syllable, are quite language specific. For example, Japanese, Korean and French have tonally marked unit, which, following Pierrehumbert and Beckman (1988), have been called “accentual
phrases” (Jun, 1998; Jun and Fougeron, 2002). However, languages such as English, Italian, and Chinese do not have such a unit. Speakers of Japanese as a second language therefore must learn to produce and parse this unit correctly and appropriately.

This section has discussed the universal nature of some aspects of intonation across languages, and some aspects that are specific to languages. The next section will examine the acquisition of intonation by young children, especially by focusing on their ability to use intonational cues to derive the pragmatic implications of utterances.

### 2.1.1 Acquisition of intonation in children’s first language acquisition

A review of literature on the acquisition of intonation in children reveals an interesting and intriguing anomaly: experimental data indicate that, while children are capable of producing adult-like intonation patterns at as early as three years of age, their comprehension of sentence-level prosody is acquired relatively late, some empirical data indicating that the performance of ten-year-olds was still significantly worse than adult performance. This is viewed as an ‘anomaly’ because it is generally the case that the development of receptive skills precedes the development of production skills. A number of studies have been done in an attempt to answer the question of why this anomaly exists, but no satisfactory explanation has been given.

In Section 2.1.1, we will first review some data which help us understand the development of production skills by children. Next, we will review some experimental data which shed light on the developmental process of comprehension of intonation in children, in Sections 2.1.2, 2.1.3, and 2.1.4.
2.1.1.1 Production data: Acquisition of emphatic stress

Hornby (1971) provides data to show the ability of children ages six, eight, and ten to use “contrastive stress” correctly. In his experiment, the subjects were shown a set of two pictures and heard a description that did not exactly match any of the pictures, and were asked to correct the description by telling the experimenter what he should have said. Hornby’s data show that in 93% of the cases, children used contrastive stress to emphatically mark the element being corrected, which lead Hornby to conclude that “stress … may be the most basic device employed for marking topic-comment relationship (p.1986). These data clearly indicate that children have developed the productive skills to use contrastive stress correctly by age six. In a previous work, using a similar picture description task, Hornby and Hass (1970) demonstrated that the ability to correctly assign contrastive stress was well established even earlier, in four -year-olds.4

MacWhinney and Bates (1978) examined the use of nine sentential devices, emphatic stress being one of the nine, by children at ages three, four, and five for conveying “givenness” and “newness,” providing additional data to support the early development of the productive skills in the acquisition of intonation. In their experiment, three sets of statistical analysis showed significant age effects, indicating that older children used more stress than younger children as a marker of contrast. However, a significant interaction between age and an increase in the adult-like use of emphatic stress was found in only three analyses out of nine possible sets in which the age variable could have had a significant effect. Altogether, the data is interpreted by MacWhinney

4 The results are reported in Cutler and Swinney (1986) as well as Cruttenden (1985).
and Bates to indicate that although a slight increase in the use of stress as a marker to
distinguish new information from given was observed, the acquisition of this device
(stress) was largely completed by age three (p. 552).

2.1.1.2 Comprehension data: Early experiments

Despite the demonstration of early development of the productive skills in the
acquisition of intonation, the existence of a curious “anomaly” has been repeatedly
pointed out: that young children’s acquisition of receptive skills in the comprehension of
prosody seems to lag behind their acquisition of productive skills. The data provided in
Hornby (1971) clearly demonstrate this discrepancy between the development of
productive and receptive skills. Using the picture-matching technique, Hornby (1971)
studied children’s (ages six, eight, and ten) ability to make use of various structural and
prosodic cues to comprehend the topic-comment distinction in sentences. Although the
data was obtained from the same set of subjects, the those who demonstrated the mastery
of the productive skills in correctly using the contrastive stress nonetheless performed at
only slightly above the chance level in the comprehension task when they were required
to choose the picture which contained the element being marked with the emphatic stress.

Investigating the relative importance of the three linguistic devices, prosody,
syntactic markers, and word order, Lahey (1974) studied four- and five-year-olds’
comprehension of complex sentences (coordinate, center-embedded relative clause and
right-branching relative clause) by having the children act out the sentences they heard.
The results indicate that while manipulation of syntactic markers (i.e., deletion of
function words and tense marking) had a significant effect on their performance,
elimination of the prosodic cues (i.e., use of list-like prosody in place of normal prosody) 
did not negatively influence on their comprehension of the sentences. This result has been 
taken by successive researchers as additional evidence to indicate that “these [prosodic] 
cues are not used to extract linguistic information before about five to six years of age” 
(Moore et al. (1989), p. 154) and that “sentence prosodic features, in particular sentence 
stress patterns … [are] very poorly processed by young children” (Cutler and Swinney, 
1986, p. 147).

The data we have examined so far clearly show that adult-like competence in 
processing the prosodic information does not fully develop before children are as old as 
ten years of age. This might make one wonder what it is that has caused them to perform 
poorly in these experimental contexts. Cutler and Swinney (1986) raised the question of 
whether or not children’s poor performance observed to date was due to the inadequacy 
of the measurement. Specifically, they suggest that “the failure to find certain prosodic 
processing effects in young children is due simply to the insensitivity of the tasks which 
were used: perhaps the children could process the prosodic structure, but could not carry 
the results of this processing through the necessary additional cognitive stages to the 
desired response” (pp. 148-149). In other words, it might be the case that children are 
actually capable of recognizing various types of intonation patterns, but they fail to take 
advantage of such prosodic cues simply because they have not yet learned to associate 
appropriate meanings with these intonation patterns.

In order to test their hypothesis, Cutler and Swinney (1986) used an on-line word-
monitoring task to test the ability of children between ages four and eight to make use of
the prosodic cues in detecting target words in a continuous speech. In this experimental paradigm, children were instructed to listen for a specified target word, and press a response key as soon as they detected the word in the sentence (or the passage) to which they were listening. In the test sentences and paragraphs, the target words appeared in the accented position half of the time and elsewhere in the other half. In the analogous on-line phoneme monitoring task, it has been established that adults can detect the target phonemes with a shorter response time when the target phoneme appears in an accented word, demonstrating that “adult listeners make very active use of the prosodic structure of speech” by “direct[ing] their attention toward the most important parts of incoming messages” (p. 149). Cutler and Swinney assumed that, since this task does not require the subjects to compute the pragmatic implications of the utterances being conveyed by the accent placement, this on-line monitoring task would be less cognitively demanding. If it is the case that children are capable of recognizing the accented words, they hypothesized that their performance might indeed be as good as that of adults in this experimental paradigm and they would be able to detect the target word with a shorter response time when the word was accented.

Cutler and Swinney compared the data they obtained from children with those obtained from adults. Contrary to the prediction formulated by Cutler and Swinney (1986), however, the children’s performance was not like that of adults even in this on-line monitoring task, which they had assumed to be a more direct measurement of children’s use of prosodic structure in comprehension. This led Cutler and Swinney
(1986) and subsequent researchers to the conclusion that children did not make active use
of the prosodic structure of speech.

Cutler and Swinney’s data, however, does not provide a definitive answer to the
question of whether young children were unable to use prosodic cues effectively. While
phoneme monitoring does not require the subject “to compute the pragmatic
implications” it is a task that requires reading fluency in an alphabetic writing system –
which younger children will not have. Therefore, the performance difference between
young children and adults may have been due to their lack of reading fluency.

2.1.1.3 Comprehension data: Focus on the comprehension of the pragmatics of
intonation

This section reviews some experimental works that employ methodological
strategies different from those of the studies discussed in the previous section. The
following studies attempt to measure children’s ability to comprehend the variations in
meaning conveyed by the intonation patterns by employing experimental approaches that
isolate the variables being tested.

Moore et al. (1993) conducted a series of two experiments to test the ability of
three, four, and five-year-old children to make use of lexical and prosodic cues in the
comprehension of relative certainty. This was done by having the children guess the
location of hidden objects on the basis of two statements they heard as clues to the
location. For example, when they were given clues from two puppets, one puppet saying
“I think it is in the blue box” and the other “I know it’s in the red box,” being able to find
the object in the red box was considered an indication of their competence in
comprehending the pragmatics of these belief terms. In the first experiment, the children heard a pair of two sentences which differed either in terms of the belief terms (know, think or guess) or of utterance final pitch contour (falling or rising). In the second experiment, lexical and prosodic cues were mixed and matched so that the lexical cues indicating a stronger degree of certainty were paired with a prosodic cue indicating a lesser degree of certainty, and vice versa. This was to test the interaction and relative importance of the two types of cues in comprehending the message when the two cues were in conflict.

Moore’s findings can be summarized in the following three points. The first point is, three-year-old subjects used neither lexical nor prosodic cues reliably. Their performance remained at the chance level even when the speaker’s differing degrees of certainty were conveyed by the different belief terms or the utterance final pitch pattern. Second, the availability of the prosodic cues helped the four-year-old children guess the location of the object at a level significantly higher than chance. However, in Experiment 2, in which two cues in conflict were provided simultaneously, the four-year-old children did not show evidence of sensitivity to the prosodic cues. Moore suggests that, taking into consideration the fact that children are just starting to show sensitivity to the lexical cues at this age, these lexical cues “take on a dominant role with respect to the comprehension of speaker certainty” and that “prosodic cues are ignored when in the presence of relevant lexical information” (p. 164). It is worth noting that his four-year-old children did seem to use the variation in sentence final pitch contour as cues to indicate the speaker’s degree of certainty. Lastly, although five-year-old children did not show evidence of sensitivity
to the prosodic cues in Experiment 1, the results of Experiment 2 show that they are not completely unaware of the implications of the prosodic cues. In Experiment 2, while they correctly responded to the cues conveyed by the know-think difference at a level significantly higher than chance, their performance deteriorated when the same pair of lexical cues appeared with prosodic cues which were mismatched in terms of the degree of certainty (“I know ...” with rising intonation, and “I think …” with falling intonation), indicating that the children’s confidence in the greater degree of certainty expressed lexically was undermined by the presence of the contradicting prosodic cues.

Lastly, more convincing data of the children’s comprehension of prosodic information are found in Cruttenden (1974) and Cruttenden (1995), which employed methodologies to examine children’s sensitivity to the prosodic cues independently of other variations such as syntax, lexical semantics, and word order. Cruttenden (1974) tested seven- and ten-year-old children’s ability to make use of intonation as a cue to predict the results of football games in British English at ages between seven and ten. In England, the result of a football game is reported in a fixed order (home-away), and the conventional use of the rising and falling-rising tone on the first score would give a strong cue to the listener in predicting the result (whether the game was home-win, draw, or away-win) before the second score is read. In this experiment, 28 children were asked to first listen to a reading of a list of results of football games in which the scores of the second teams were not given, and then guess the score of the second team. The results indicate that more than 50 percent of the children “show nothing approaching full competence” (p. 226). Indeed, out of 28 children, 11 children performed at the chance
level, and an additional 9 children at only slightly better than chance. Cruttenden (1974) also examined the possible influence of the degree of interest in football. The statistical analysis indicated that there is a lesser interaction between their interest in football and score ($P < 0.10$) than that between their age and score ($P < 0.001$). This result indicates that, while the children’s level of interest in football did not serve as a significant predictor of their task performance, their age clearly did, and that children who were seven had not yet developed a full competence in comprehending this particular use of intonation patterns.

Cruttenden (1985) tested ten-year-old children’s awareness of the different meanings conveyed by three types of intonation contrasts: “tone,” “nucleus placement,” and “grouping.” This was accomplished by having 20 ten-year-old children hear pairs of sentences which might (or might not) differ minimally in their intonation patterns (i.e., intonational minimal pairs), and a) respond to questions testing whether they accurately perceive them as being different (or the same) and b) match pictures which represented the messages of the sentences correctly. The data obtained from the children were compared with those obtained from twenty adults. Clearly, in all the test sentences, the children’s performance was much lower than that of the adults, showing that “their ability to make metalinguistic judgments about language is still very limited at this age” (pp. 656-657). Cruttenden’s (1985) data are valuable because the children in this experiment were tested on their comprehension of the pragmatic implications being conveyed by intonation in contextualized settings (however limited they may be). Thus, what the data tell us is, as summarized by Cruttenden (p. 657), that “when children are having to
interpret utterances in real life whose intonations are critical, they are probably not interpreting such utterances consistently, or in the same way as adults.”

2.1.1.4 Criticism on the methodology

Moore et al. (1993) offered various criticisms on the methodologies employed in the experiments mentioned in the previous sections. Their criticism can be summarized in the following three points. First, Moore et al. (1993) pointed out that “the methodologies employed are biased towards finding an advantage for lexical over prosodic information” (p. 155). For example, the authors comment that when children, as experimental subjects, are asked to act out the statements they have heard (cf. Lahey, 1974), they are in a way forced to attend more to the syntactic structure of the sentence or its meaning than to the prosodic aspect of the utterance. Likewise, the ability to identify a particular word (cf. Cutler and Swinney, 1986) cannot be considered an appropriate measurement of the children’s knowledge or the understanding of the pragmatics of intonation. Moore et al. thus commented that the reason prosody did not show any influence on the comprehension in such experiments may have been simply because the experimental tasks discourages children to pay enough attention to the prosody. Second, they pointed out that the children’s failure to attend to the prosodic cues might have been simply due to the difficulty of the tasks assigned to them. For instance, the sentences used in Lahey (1974) were complex in their syntactic structure, and the failure of the children to attend to the aspect of prosody might have been due to the limitation of their cognitive capacity (Moore et al., p. 160). Similarly, when forced to make a choice between two options, both of which contain errors, they may have been too involved with the semantics of the
sentences, neglecting the prosodic aspect of the utterances (cf. Hornby, 1971) Based on these observations, Moore et al. (1993) claim that “the future work on comprehension of prosody in young children should use tasks in which the requirement is to respond to the pragmatics, rather than the semantics, of the stimuli” (p. 160) by having them respond to the speaker’s intended meaning.

Taking these points into consideration, it seems reasonable to take Moore’s data as evidence to indicate that children are indeed starting to develop sensitivity to the variation of meaning under variation in intonation, and make use of such prosodic information in comprehending the messages being conveyed in the utterance. At the same time, as Cruttenden (1975) and Cruttenden (1984) have demonstrated, their comprehension skills are not yet sophisticated enough to be able to derive the correct pragmatic interpretations in context, and to understand the subtle implications being expressed by the use of various types of intonation, and this process of acquisition of these skills is apparently is not fully completed by the time they are ten years old.

2.1.2 Acquisition of intonation by adult foreign language learners

Cruz-Ferreira (1987; 1989) applied the experimental techniques used in Cruttenden (1985) to the examination of the comprehension of intonation in English and Portuguese by learners of English who are native speakers of Portuguese and learners of Portuguese who are native speakers of English respectively. Cruz-Ferreira (1987) provides a description of her experimental study, and based on the findings of the research, later purports a more general experimental procedure that can be used to test the comprehension of English intonation by learners of any native language. Cruz-
Ferreira (1987) correctly points out that the past studies on non-native use of intonation have been production studies, often limited to the examination of “foreign accent,” i.e., characterization of the learner’s utterances in terms of deviation from the native norm, and the less obvious phenomenon of the comprehension of intonation has been neglected. This observation applies to the studies on the acquisition of Japanese as a second language⁵ - no existing study has examined the development of comprehension ability with reference to intonation other than the examination of pitch accent by adult learners of Japanese.⁶ Her view of intonation as an “independent linguistic level” which, in each language, has its own system of meaning is clearly and explicitly expressed by Cruz-Ferreira (1989, p. 25): that intonation provides cues to convey the intentions of the speakers, and that it interacts with the lexical and grammatical structures of the utterance. I agree with this point, and the system of Japanese intonation will be examined in detail in Sections 2.2 and 2.3.

The experiment of Cruz-Ferreira was carried out in order to accomplish two objectives: 1) to compare native and non-native comprehension of intonational patterns in Portuguese and English, and 2) if the process of comprehension was different, then to answer why and how. More specifically, the researcher was interested in the way the intonation patterns of the subjects’ native language would interfere with the acquisition of the L2 intonation systems. The experimental participants heard a pair of two sentences,

⁵ The term “second language” and “foreign language” will be used interchangeably in this document, referring to any languages which are acquired by adults. Also, the notation L1 and L2 will be used to denote first and second languages, respectively.

⁶ Nishimura et al. (1996) studied the perceptual processing of lexical accent by American learners of Japanese.
which might (or might not) differ minimally in some aspect of intonation, and were first asked if the two sentences were the same or different (discrimination task). Participants were then presented with a set of three sentences written in their native language, and asked to perform either an identification or an interpretation task depending on their answer to the discrimination task. When subjects judged the two sentences to be the same in the discrimination task, they were asked to indicate which sentence best corresponded to the meaning of the sentence they had just heard (identification task). When subjects judged the two sentences to be different in the discrimination task, they were asked to “interpret” the two sentences by choosing one appropriate meaning for each sentence (interpretation task). The same set of three sentences was used as potential responses for both the identification task and interpretation task.

The results of the experiment reported in Cruz-Ferreira (1987) support her hypotheses, showing clearly that the non-native speakers were applying their knowledge of the L1 intonation system in interpreting the new intonation system.

In Cruz-Ferreira (1989), summarizing the two assumptions underlying the design of a test which seeks to assess the comprehension of English intonation by non-native speakers of English, she makes the following explicit statements:

a) that the “intonation patterns of any L2 cannot be adequately mastered by the learner otherwise than by systematic teaching,” and

b) that “the first difficulty faced by non-natives with the intonation of L2 lies in comprehension and not in production: in the first stages of learning, a faulty
comprehension of intonation will determine difficulties in production, and not vice-versa” (1989, p. 24).

In other words, she proposes that the acquisition of perceptive skills necessarily precedes the acquisition of productive skills, and that the discrepancy between L1 and L2 intonation systems will lead to inaccurate comprehension, which, consequently, will cause difficulties in production. This suggests the importance of obtaining empirical data to test the comprehension ability of the L2 learners and to examine the patterns of interference between their L1 intonation system and the target language. Such empirical data will help us propose a more “systematic” way of providing instruction to train the learners to develop competence in the accurate comprehension and production of L2 intonation patterns.

2.1.3 Comparison between the two acquisition processes: Children acquiring their native language vs. adults learning a second (foreign) language.

We have examined the experimental data to show how children acquiring their native language and adults learning a second (foreign) language develop their competence to accurately comprehend the intentions of the speaker. The differences in the two developmental processes may be summarized in the following points:

- **Order of acquisition:** As has been discussed above in detail, children seem to acquire the production skills before they have developed their full competence in their comprehension ability. In contrast, adult language learners will follow the opposite course of acquisition: their acquisition of comprehension skills will precede the acquisition of production skills, and Cruz-Ferreira (1987) suggests that the acquisition
of perceptive skills will be a necessary condition for developing skills to properly produce the L2 intonation patterns.

- **Influence of the intonation system of their L1:** In the adult L2 acquisition process, the acquisition of the L2 intonation system will be negatively influenced by their L1 system (interference) if and when the two systems differ from each other. This does not apply to the children’s acquisition of their native language.

- **Limitation of the cognitive capacity:** Children who are acquiring the intonational system of their native language are still in the process of acquiring other aspects of the system of the language as a whole, such as syntax and lexical semantics, as well as developing their general cognitive capacity. As Cruttenden (1987, p. 659) points out, this limitation in their cognitive capacity may lead to children’s inefficiency in allocating due attention to intonational cues as opposed to other cues when they are in conflict. This differs from the adult learners of foreign languages, who possess mature cognitive and analytical capacity.

- **Interaction between intonation and pragmatics:** For both children and adults, the biggest challenge appears to be to learn the interaction between an intonation pattern and its intended pragmatics. For children, this requires learning to associate the correct meaning to the intonation pattern. For adult L2 learners, this process is more complicated because it requires them to 1) learn to correctly identify as well as produce the new inventory of intonation patterns, 2) learn the interaction of the intonation and lexical-grammatical elements, and 3) develop competence in being able to make the right kinds of inferences. Especially when the target L2 is
linguistically as well as culturally distant from their L1, this process will require additional attention. For example, imagine a situation in which a native speaker of American English is learning Japanese as a second language, and developing the ability to comprehend the intention of the speaker, “disagreement” in particular, being conveyed by the intonation. In order for him to successfully comprehend his interlocutor’s “disagreement” by paying attention to the intonation, he would first need to learn how to perceive this intonation pattern correctly, and learn to associate this pitch movement with its intended meaning, “disagreement.” In addition, the context in which it is appropriate to explicitly express such disagreement would differ greatly in his home culture (America) and in the target culture (Japan). Therefore, learning to be able to perceive the intonation alone does not guarantee that he has developed the competence to use the cues conveyed by the intonation to accurately perceive the intention of the speaker.

2.1.4 Acquisition of intonation by native speakers of other dialects of Japanese

Little research has attempted to describe how native speakers of different dialects perceive and process each other’s utterances when there are intonational differences across dialects of the language. One of the rare studies, Kōri (1992) examined how the intonational contrasts of the Osaka dialect influenced the perceived degree of politeness and sincerity. His data is highly suggestive in understanding the acquisition of intonation in the native language, indicating that native speakers of other dialects are not fully assimilated to the implications of the so-called standard dialect, to the extent that they
react negatively to dialectal variations. The significance of the contribution of his data to the analysis of the results of this study will justify a detailed examination of his research.

Osaka, the second largest city in Japan, is located in the western part of Japan, and its dialect differs from the Standard Japanese in its vocabulary and intonation. Use of rising intonation with apologies (‘gomen ne’) and thanks (‘arigatoo’) is particularly common in Osaka, but not in the Standard Japanese. The subjects, men and women living in the vicinity of Osaka who were not users of this type of rising intonation themselves, evaluated the utterances ‘gomen ne’ and ‘arigatoo’ spoken with rising (Osaka) intonation and with Standard intonation. The evaluators were native speakers of neither Osaka nor Tokyo Japanese, but had more familiarity with Osaka, since they were living and working there. The data show that they tended to evaluate negatively utterances with the intonation patterns of the Standard dialect, the dialect that was less familiar to them, as “not sincere,” “cold,” and “strange.” Thus, the data indicate that native speakers of different dialects have different systems of making association between tune and sentence. As the result of the interaction between the intonation and sentence, speakers of different dialects arrive at different pragmatic implications such as being “insincere,” “cold,” and so on.

Kôri’s (1992) study focuses on responses to expressions that involve emotional reactions. The focus of his study, therefore, is different from the question of whether they are able to ‘figure out’ the literal meanings of utterances. It is significant that the data shows even native speakers of the Japanese language may associate different pragmatic
implications of to unfamiliar intonation patterns when they are speakers of different dialects.

This is a clear demonstration of the need for learners of Japanese to be systematically trained to acquire a system of intonation that is consistent. If the target language is standard Japanese, all the aspects of their speech should be consistent with its system, including the aspect of intonation. As Kôri’s data show, deviation from the expected norm may cause the user of the language to leave a negative impression, which may have a negative impact on his/her ability to form positive relationships with the interlocutors.

2.2 Limitations of teaching intonation of Japanese within the framework of “pronunciation” practice: old theories and practices

In this section, I will review instructional materials available for teaching intonation in Japanese. The aspect of ‘intonation’ is often addressed as a part of ‘teaching pronunciation’ (発音指導: hatsuon shidô). I have, however, several reasons to argue that the issue of ‘intonation’ should not be equated with ‘teaching pronunciation’ in a traditional sense without due consideration. When I say ‘teaching pronunciation in a traditional sense’ here, I am implying that there is a commonly shared notion of what ‘teaching pronunciation’ looks like. Having taught Japanese for over ten years in several educational contexts in the United States, I am also aware that the notion of ‘teaching pronunciation’ is often regarded unfavorably, and is rarely dealt with as a part of the Japanese curriculum.
I will therefore, briefly reexamine the practices commonly associated with ‘teaching pronunciation’ (発音指導: *hatsuon shidô*). The question of how to improve instruction on pronunciation was discussed in some detail by Eda (1996, pp. 87-91). One significant limitation of Eda’s earlier discussion was that it did not take into account the aspect of spra-segmental features, such as accent and intonation, in teaching pronunciation. The current discussion will therefore re-address this issue as an attempt to expand its scope to include all the areas of speech production, both segmental and spra-segmental, in the teaching Japanese.

Prosodic features of utterances are determined by the configuration of tonal events superimposed on a string of smaller units of sounds, and such tonal events are systematically governed by rules, as discussed in Chapter 2. Therefore, learners will benefit greatly from the knowledge and understanding of some essential aspects of speech at lower levels of the hierarchy in order to discuss intonation. What I want to say is that 1) intonation should not be regarded as an appendix to the teaching of pronunciation, that 2) teaching intonation will be effective if other aspects of the prosodic organization of Japanese are discussed as a prerequisite for understanding both the syntax and pragmatics of intonation, and 3) teaching intonation should focus on the learners’ acquisition of pragmatic competence.

### 2.2.1 Issue number one: too much emphasis on the consonants and vowels?

A brief review of textbooks designed to teach Japanese pronunciation reveals an overwhelming emphasis placed on providing instruction on accurate production of consonants and vowels, of which Japanese words and sentences are composed. The focus
of *hatsuon shidō* is on the analysis of articulatory gestures required to produce discrete sound units, and the production of combinations of those units that form words. This practice in itself is not a problem. The learners’ perception and production of consonants and vowels are influenced significantly by their native languages. For instance, native speakers of Chinese and Korean languages tend to have difficulty in distinguishing voiced and voiceless sounds, such as /p/ versus /b/, /t/ versus /d/ and so on. It will, therefore, be extremely useful for speakers of Chinese and Korean languages to receive focused instruction on accurate pronunciation of such consonants that would present difficulties. The same would apply to speakers of other languages. In addressing these points, it will be more effective to isolate the problem area and work on it directly.

One problem that arises as a consequence of too much emphasis on the accuracy of consonants and vowels is that teaching pronunciation has come to be equated with learning to produce meaningless sound units in isolation. In addition, because teaching pronunciation tends to be limited to the level of reproduction of sounds in isolation, instruction on accent and intonation has received marginal attention within the context of the teaching of pronunciation. For example, Fujito and her colleagues (1979) put accent and intonation in her textbook at the end as if it were an afterthought.

Another tendency that arises as a consequence of too much emphasis on the accuracy of consonants and vowels is a focus on the lower level of processing, i.e., perception and production of sounds divorced from ‘meaning.’ The textbooks for pronunciation treat pronunciation in isolation from everything else, first by presenting and analyzing the target unit of sound, which is then followed by a list of a handful of
words which contain the sound being introduced. Such exercises, however, are not linked to any kinds of exercises that involve communication between two speakers, i.e., exercises at the discourse level. It seems, in other words, that the current ‘pronunciation training’ is completely disconnected from the reason for good pronunciation. If the purpose of pronunciation training is assumed to be that correct pronunciation helps smooth communication, pronunciation exercises need to become more communicative rather than production of phonemes or of words in isolation.

2.2.2 Issue number two: pronunciation training takes up too much time?

The second ‘myth’ associated with pronunciation training is that it is as a process that takes an inordinate amount of time. The available materials for teaching pronunciation are designed (obviously enough) to deal with the aspect of pronunciation, and such an instructional design may be easily interpreted to imply that they should be used as textbooks for courses whose sole purpose is to teach correct pronunciation in Japanese. This has also been taken to imply that activities that focus on pronunciation cannot be combined with more communicative exercises.

A notion of ‘teaching pronunciation’ such as this has led teachers of Japanese to choose one of two approaches. One approach is to devote the first few weeks of instruction to teaching pronunciation. The following figure offers an example of this approach. It is a syllabus of a Japanese language course at the beginning level at a university in China, published on the Internet (Kaji, 2001). I will include the syllabus in its entirety to show the proportion of the time spent for the pronunciation exercises. Note that Kaji chooses to spend half of his course (13 hours out of 26) on practicing
pronunciation using the table of fifty moraic sounds in Japanese called 五十音図 (gozyuu-on zu).
<table>
<thead>
<tr>
<th>時間目</th>
<th>会話の導入（形容詞）</th>
<th>会話の導入（動詞）</th>
</tr>
</thead>
<tbody>
<tr>
<td>15 時間目</td>
<td>五十音図(13) ば行～会話</td>
<td>五十音図(19) みゃ行</td>
</tr>
<tr>
<td>16 時間目</td>
<td>会話～五十音図(14) きゃ行</td>
<td>会話～五十音図(15) しゃ行</td>
</tr>
<tr>
<td>17 時間目</td>
<td>会話～五十音図(15) しゃ行</td>
<td>会話～五十音図(16) ちゃ行</td>
</tr>
<tr>
<td>18 時間目</td>
<td>会話～五十音図(16) ちゃ行</td>
<td>会話～五十音図(17) にゃ行</td>
</tr>
<tr>
<td>19 時間目</td>
<td>会話～五十音図(17) にゃ行</td>
<td>会話(代名詞)～五十音図(18) ひゃ行</td>
</tr>
<tr>
<td>20 時間目</td>
<td>会話(代名詞)～五十音図(18) ひゃ行</td>
<td>会話～五十音図(19) みゃ行</td>
</tr>
<tr>
<td>21 時間目</td>
<td>会話～五十音図(19) みゃ行</td>
<td>日本語 10 品詞(1～5)～五十音図(20) りゃ行</td>
</tr>
<tr>
<td>22 時間目</td>
<td>日本語 10 品詞(1～5)～五十音図(20) りゃ行</td>
<td>日本語 10 品詞(6～10)～五十音図(21) ぎゃ行</td>
</tr>
<tr>
<td>23 時間目</td>
<td>日本語 10 品詞(6～10)～五十音図(21) ぎゃ行</td>
<td>会話～五十音図(22) じゃ・ちゃ行</td>
</tr>
<tr>
<td>24 時間目</td>
<td>会話～五十音図(22) じゃ・ちゃ行</td>
<td>会話(現在・過去)～五十音図(23) びゃ・びゃ行</td>
</tr>
<tr>
<td>25 時間目</td>
<td>会話(現在・過去)～五十音図(23) びゃ・びゃ行</td>
<td>テスト</td>
</tr>
</tbody>
</table>

Table 2.1: A syllabus of a beginning level Japanese language course, utilizing the hiragana chart as an introduction to pronunciation
Another reaction to the notion that teaching pronunciation takes an inordinate amount of time is rightly expressed by Aizawa (1991, 124) that “it is not worth spending the precious time for pronunciation.” I agree with her, that if teaching pronunciation needs to take up that much time, then it is not really worth it.

It is important to realize, however, that it is not the case that students cannot speak Japanese until they have mastered all the correct pronunciations of consonants and vowels. Nor is it the case that pronunciation needs to be taught to the exclusion of everything else. Using a textbook that is dedicated to teaching pronunciation does not mean that it cannot be used in combination with other textbooks. There are many ways to make exercises on pronunciation a part of communicative exercises. I will discuss how to do so in Chapter Five.

2.2.3 Issue number three: pronunciation taught at the beginning of the course ... and that’s it?

Observing that “the initial point of contact between the two languages is in the area of pronunciation and basic intonation patterns,” Hammerly asserts that “it will be of crucial importance to pay attention to aspects of pronunciation and basic intonation patterns at the early stages of learning, especially in a program that places emphasis on oral proficiency” (2001, p. 36) This is a view that is widely shared (Fujito et al., 1979; Nakanishi, 1987; Sakuma, 1964; Suzuki, 1991), and I also agree with his statement. Hammerly also comments that encouraging learners to use the language to communicate without paying proper attention to its pronunciation will only result in “disabling them
[learners], and likely permanently so” (Hammerly 2001, vii), thus arguing for the importance of providing explicit training in accurate pronunciation.

While this is an important and valid point, such a view should not be taken to mean that once pronunciation is discussed at the beginning of the course, it is never addressed systematically in class again. To the contrary, I argue that pronunciation, especially the aspect of “pronunciation” that involves learning patterns of contrasting intonation, is an integral part of the language that should be made a strategy for successful communication at all levels of instruction.

2.2.4 Issue number four: it focuses on production only.

As the fourth problem, I want to point out something that may not be very obvious, that such pronunciation training tends to focus only on the production of units of sound in isolation, and tends to ignore perception and interpretation. This commonly accepted practice is explicitly reinforced in the guidelines for Oral Proficiency Interview (OPI), published in 1986 by the American Council On the Teaching of Foreign Languages (ACTFL). These guidelines have gained widespread recognition and appreciation since their publication, and have been used as a standard for measuring learners’ functional competency. This publication continues to be regarded as one of the most influential publications dealing with the teaching of languages in the United States. In this document, “accuracy in pronunciation” is defined as the “ability to reproduce segmental and supra-segmental (pitch, stress, intonation) features of the language” (Swender, 1999, p. 25, emphasis is mine). This definition is problematic, since “accuracy in intonation” as defined by ACTFL does not include accuracy in the learners’ pragmatic
competence, i.e., the ability to understand the implications being communicated by prosodic variation. The definition does not deal with perception, and, therefore, it does not address interpretation, either. But note that these are testing guidelines, and not instructional guidelines. To the contrary, instruction on intonation needs to focus more on the aspect of pragmatics.
CHAPTER 3

INTONATION SYSTEM OF JAPANESE LANGUAGE

Two aspects of the prosodic organization of Japanese language are relevant to this study, namely, the effect of intonational phrasing and, second, the types and functions of pitch movements at phrase boundaries. This chapter will first discuss the prosodic structure of Japanese language, focusing on aspects such as the pitch accent and prosodic structure of a well-formed utterance in Japanese. This will be followed by the analysis of utterance final elements in Japanese. The functions of sentence final particles and their interactions with the utterance final pitch movements will be discussed in the final section of this chapter.

The term ‘pitch’ is used in this discussion to refer to the prosodic ‘melody’ of an utterance that can be measured in terms of the contour of the fundamental frequency (F0). Throughout this document, the examination is limited to the accent and intonation of standard Japanese, spoken in the Tokyo area. I will employ the framework of analysis initially proposed by Pierrehumbert (1988) and developed further by Venditti (1997), and assume that there is a hierarchical structure governing the realization of the intonation patterns.
3.1 The prosodic hierarchy in Japanese

3.1.1 Pitch Accent

The term ‘pitch accent’ is used in various senses in research in phonetics. The notion as explained in Beckman (1992) is that pitch accent is “some pitch event that can be represented phonologically by a tone level or a sequence of two tone levels,” which “is anchored within the word or phrase … at a designated tone-bearing-unit (syllable or mora)” (p. 187). Defined in this sense, both English and Japanese can be considered to have pitch accent, but accent is realized differently in these two languages in terms of how such accent is physically manifested. As Sugito (1989a) observes, while accent in Japanese is often described as takasa-akusento ‘(pitch) height-accent,’ English accent is described as tuyosa-akusento ‘stress-accent.’ This is due to the fact that, in English, the tone-bearing-unit, the syllable with which the accent is anchored, is cued acoustically by its intensity (i.e., it is louder) and duration (i.e., it is longer) in addition to the pitch modulation. This has the effect of marking the accented syllable as perceptually more salient than the rest of the utterance. In Japanese, in contrast, accent is realized simply as a sharp F0 fall immediately after the ‘accented mora,’ the mora with which the accent is anchored, and, unlike the English stressed syllable, the accented mora does not receive increased intensity and duration. Also, there are differences in where accents can occur. In Japanese, accent can occur on any type of syllable, whereas in English, syllables with

---

7 A similar definition of accent is proposed in Sugito (1989b) and Ladd (1992).

8 Also, terms such as ‘non-stress accent’ and ‘melodic accent’ are used by Ladd (1992, p. 200) to describe the characteristics of Japanese pitch accent.
reduced vowels cannot take accent. Keeping these differences in mind, I will hereafter use the term ‘accent’ and ‘pitch accent’ to refer to the Japanese accent.

In Japanese, each lexical item is specified as accented or unaccented as a part of its lexical property. An accented word has at most one bitonal pitch accent associated with the lexically accented syllable, and this is realized as a sharp pitch fall. An unaccented word, on the other hand, does not have any pitch accent associated with it. For example, consider the three schematic representations of the following utterances in Figure 3-1 by paying attention to the contrasting accent patterns.

(1)  \textit{hasi' o nuru}. (I paint a bridge.)
(2)  \textit{ha'si o nuru}. (I paint chopsticks.)
(3)  \textit{hasi o nuru}. (I paint the edge [of something].)

In the transcriptions, a diacritic (') is used to mark the mora with which the accent is associated, and is placed immediately following each accented mora. While sentences (1) and (2) begin with an accented noun, \textit{hasi'} (bridge) and \textit{ha'si} (chopsticks), sentence (3) begins with the unaccented noun \textit{hasi} (edge [of something]).
Most obviously, accent specifications help distinguish homophones, such as *ha'si* (chopsticks) and *hasi'* (bridge), or *a'me* (rain) and *ame'* (candy). Differences in accent specification also play a significant role in speech processing, as demonstrated by Cutler and Otake (1998) and Yoneyama (in progress). Cutler and Otake’s (1998) data show that native listeners of standard Japanese use pitch-accent information as soon as it becomes available at the word initial position, to quickly activate and select the candidates in spoken-word recognition tasks. 

An accentual phrase typically consists of a word and its postposition. This is the level above the word, and the lowest intonationally marked level in the hierarchy of
prosodic organization in Japanese. An accentual phrase may contain at most one accented word, or may contain no accented word. An accentual phrase can be made up of more than one word, and it is possible therefore to have two or more words bearing lexical accent within an accentual phrase. In such a case, the second and following accents will be “degenerated” to the extent that they are no longer detected, consequently resulting in an accentual phrase containing only one fully realized lexical accent (see Maekawa, 1994). An accentual phrase is delineated by low boundary tones, and also is characterized by a rise to high pitch on the second mora, or earlier if there is lexical accent on the first mora. In Figure 1, note that there is a pitch peak associated with the phrasal high on the second mora in sentences (2) and (3). These are properties of the phrases rather than of the lexical items. Note also that this peak is not fully realized in sentence (1) since the first mora in that sentence is specified to have a lexical accent. Thus we may say that while utterances (1) and (2) begin with an accented accentual phrase, utterance (3) begins with an unaccented accentual phrase.

3.1.2 Effect of phrasing: Syntactic function

Two or more accentual phrases can be combined to form an intonation phrase, at a higher level, above the accentual phrase, in the hierarchy of prosodic organization. This level is the domain of “downstepping.” Downstepping is a step-wise reduction in phrasal pitch range that is conditioned by the lexical accent specification of the accentual phrases that make up an intonation phrase. That is, the peak values in an accentual phrase (whether it is accented or not) will be lowered if it is preceded by an accented accentual phrase within the same intonation phrase. This is illustrated in Figure 3.2.
Downstepping is blocked by an intonational phrase boundary. This is illustrated in Figure 3.3. In the utterance in Figure 3.2, all four accentual phrases, each of which contains an accented word, are grouped into one intonation phrase, and the pitch peaks of the second through fourth accentual phrases are successively reduced by downstepping. The utterance in Figure 3.3 is the same string of words as the utterance presented in Figure 3.2. However, this utterance is made up of two intonation phrases. The arrow in Figure 3.3 indicates the boundary between the intonation phrases. From the perspective of a listener, this manipulation of the placement of intonation phrase boundaries may be signaled by a sudden expansion of pitch range. Moreover, this pitch range expansion is available for sequences with unaccented words as well as those with accented words. That is, it can occur not just to enhance the lack of downstepping, but also as the sole pitch-related cue to the boundary. Thus, it is this manipulation of the pitch range expansion that serves as a fairly reliable cue for disambiguating potentially ambiguous surface syntactic structures.
Figure 3.2: The effect of downstepping in an utterance consisting of four accented accentual phrases

Figure 3.3: The effect of the placement of an intonation phrase boundary
Note that, although the occurrence of downstepping is not independently specified as part of the higher-level utterance prosody (being determined by the lexical choice of accented or unaccented word), what is under the speakers’ control at the utterance level is which accentual phrases are grouped into one intonation phrase. In other words, the speaker can manipulate the grouping of phrases into a hierarchy of prosodic constituents in Japanese, which is manifested in variation in local pitch range. Accurate production and perception of pitch accent thus plays a significant role in sentence processing because they influence to the overall pitch contour of an utterance through its contribution to the downstepping and their interaction with the placement of intonation phrase boundaries.

Many studies have shown that native listeners of Japanese can make use of such prosodic cues. In particular, previous studies have identified reliable prosodic cues which correlate with the syntactic structure, and have shown that cues for distinguishing potentially ambiguous surface syntactic structures come from the variation in intonational phrasing, which is cued by an expansion of the overall pitch range (Kubozono, 1992; Misono et al., 1997; Venditti, 1994; J. Venditti, S.-A. Jun, & M. E. Beckman, 1996). Misono et al. (1997) further demonstrate that prosodic cues can influence the interpretation of a sentence to the extent that an interpretation that is favored for the prosody can sometimes override an interpretation favored for the semantics of the word string.

3.1.3 Effect of phrasing: Pragmatic Function

The same manipulation of phrasing and local pitch range variation is used also to indicate the focus of attention in an utterance. That is, a speaker can break an utterance
into intonation phrases, and the accentual phrase that begins the new intonation phrase
(which is marked by the pitch expansion) will be perceived as the item that receives the
pragmatic focus. In figures 3a and 3b, the arrows indicate the intonation phrase
boundaries. These utterances are made up of the same string of words and have the same
underlying syntactic structure. The shapes of their pitch contours differ, however,
because the speaker produced them as responses to two different questions, placing
pragmatic focus on different accentual phrases.
Figure 3.4: An utterance consisting of four accented accentual phrases with pragmatic focus on the second accentual phrase, produced as a response to the question *Mari-chan wa nitiyooobi ni oniisan ni au n desu ka?* “Will Mary meet her elder brother on Sunday?”

‘Mary will meet her elder brother ON SATURDAY.’

Figure 3.5: An utterance consisting of four accented accentual phrases with pragmatic focus on the third accentual phrase, produced as a response to the question *Mari-chan wa nitiyooobi ni oneesan ni au n desu ka?* “Will Mary meet her elder sister on Saturday?”

‘Mary will meet her ELDER BROTHER on Saturday.’
Note also that, unlike in English, this pitch range variation for signaling focus of attention in Japanese can target an unaccented word as easily as an accented one, and it thus is not necessarily accompanied by stress—an increase in loudness and duration. This might cause some difficulty for native speakers of English both in perception and production.

It is certain that speakers of Japanese may not rely solely on intonation to communicate syntactic and pragmatic contrasts. For instance, the pragmatic focus on doyoobi “Saturday” in the utterance presented in Figure 3.4 may be more explicitly brought into focus using different syntactic structures (au no wa doyoobi desu “it is Saturday that she’ll see her older brother”). It will also be the case that an utterance with four consecutive accent phrases is considered a long utterance, and therefore might be perceived as less natural. These points of criticism, however, should not be taken to nullify the roles played by the effect of contrasting intonation patterns.

3.2 Inventory of boundary pitch movements

Pitch movements at the intonation phrase boundaries at both sentence medial and sentence final positions can provide rich cues to understanding the speaker’s intention. I will use the term Boundary Pitch Movement (BPM, hereafter) to refer to the variations at intonation phrase boundaries. A commonly shared assumption underlying the examination of BPMs is that their characterization applies to any phrase boundaries, including both sentence medial and final positions (Maeda & Venditti, 1998; Venditti et al., 1998). For the sake of simplicity, however, the discussion of BPMs in this dissertation will be limited to the utterance final positions. The scope of this discussion is
limited to the examination of the intonation patterns of the standard Japanese spoken in
and around Tokyo, Japan.

Researchers who examined the prosodic features of Japanese tended to categorize
phrase final intonation patterns simply into two types, rise and fall (cf. Pierrehumbert and
The earliest exception to this generalization is Kawakami, who reported in work in the
1960’s, later collected and reprinted in Kawakami (1995), that there were five sub-
categories within the rising intonations at the phrase final positions. After identifying the
five “shapes” of pitch contours, which he believed to be categorically distinguished,
Kawakami (1995) discusses the meanings associated with those five BPMs.
Unfortunately, Kawakami’s observations were probably ahead of their time. Coming in
an era when most linguistic analyses of intonation were narrowly focused on phrase-level
accent patterns (e.g., Hiraguchi, 1968), his observations were essentially ignored until
Venditti and colleagues began to systematically investigate different rising BPMs (Maeda
& Venditti, 1998; Venditti et al., 1998) using a methodology pioneered by Gussenhoven
and colleagues.

Attempts to seek the interface between pragmatics and prosody, Venditti and
her colleagues (Venditti et al., 1998; Maeda et al., 1998) have empirically tested
intonationally contrasting utterances and demonstrated that pitch movements at the
intonation phrase boundaries can provide semantic cues to communicate the speaker’s
intention independently of the lexical and syntactic structures they interact with. In their
experiments, the five types of BPMs which they tested\(^9\) were indeed judged by native speakers of Japanese to have different semantic implications. The experimental study presented in Chapter Four of this dissertation examines the perception of two pairs of BPMs utilizing the categorization proposed by Venditti and her colleagues. In the following section, I will discuss her experimental studies and the data in detail.

3.2.1 **Analysis of semantic properties associated with boundary pitch movements**

Maeda et al. (1998) and Venditti et al. (Venditti et al., 1998) proposed that there were (at least) five BPMs that could be considered distinct in Japanese based on the shapes of their pitch contour (f0) and examined how they were judged differently by native listeners of standard Japanese on eight semantic scales. In a series of experimental studies, ten subjects, i.e., native listeners of Japanese who were naïve to the purpose of the experiment, were asked to listen to short utterances (for example *Naoya ni* “[this is] for Naoya,” with a dative case marker *ni* at the utterance final position) produced with intonation types exemplifying the five BPMs: incredulity question rise, information question rise, prominence-rending rise, insisting intonation rise, and rise0fall boundary movement. These five intonation types are schematically represented in 3.1 below (taken from (Venditti et al., 1998); diacritics are my own).

\(^9\) The five categories tested by Venditti and her colleagues were motivated independently of Kawakami’s work, and the categories examined in their work do not correspond exactly to the five categories proposed in Kawakami (1995).
Subjects were asked to listen to the utterances that minimally differed in the BPM pattern, and to rate each utterance in terms of how strongly they agreed or disagreed with eight statements (i.e., semantic scales) such as: The speaker … 1. will continue taking. (continue), 2. is insisting on the information. (insisting), 3. is simply confirming information. (confirming), 4. cannot believe the information. (disbelief), 5. is emphasizing the phrase itself. (emph-info), 6. is emphasizing the function of the phrase in the discourse. (emph-funct), 7. is explaining the information. (explaining), 8. sounds irritated. (irritated). The semantic scales in the experiment were written in Japanese, and each stimulus was rated on a five-point scale, varying from 1 (agree) to 5 (disagree).

Figure 3.7 and Table 3.1 summarize the semantic properties of each of the BPM types, on 8 semantic scales.
Figure 3.7: Mean standardized (z) scores for each of the BPM types, on 8 semantic scales (taken from Venditti et al. (1998), p. 2)
Table 3.1: Summary of the semantic properties associated with the five boundary pitch movements presented in Figure 3.4

<table>
<thead>
<tr>
<th>Question</th>
<th>Prominence-lending</th>
<th>Insisting</th>
</tr>
</thead>
<tbody>
<tr>
<td>incredQ</td>
<td>infoQ</td>
<td>prom</td>
</tr>
<tr>
<td>– continue</td>
<td>– continue</td>
<td>+ continue</td>
</tr>
<tr>
<td>– insisting</td>
<td>– insisting</td>
<td>– insisting</td>
</tr>
<tr>
<td>+ confirming</td>
<td>+ confirming</td>
<td>+ confirming</td>
</tr>
<tr>
<td>– disbelief</td>
<td>0 disbelief</td>
<td>– disbelief</td>
</tr>
<tr>
<td>0 emph info</td>
<td>– emph info</td>
<td>– emph info</td>
</tr>
<tr>
<td>– emph funct</td>
<td>– emph funct</td>
<td>– emph funct</td>
</tr>
<tr>
<td>– explaining</td>
<td>– explaining</td>
<td>– explaining</td>
</tr>
<tr>
<td>– irritated</td>
<td>– irritated</td>
<td>– irritated</td>
</tr>
</tbody>
</table>

Table 3.1 is based on their Figure 2 (p. 2), reproduced as Figure 3.7 on p. 55, but the subjects’ ratings of the five BPMs on each semantic scale are indicated only in terms of its direction (either + or –).\(^{10}\) I created the table in order to highlight the similarities and differences among the five BPMs. The pattern of shading in this table highlights which BPMs are grouped together by the response patterns on the different scales. Note that the first three BPMs share very similar characteristics, with the exception of the ratings on the first semantic scale. This observation motivated Venditti et al. (1998) to suggest that

\(^{10}\) Two responses are indicated with ‘0’ because the score was very close to 0 (zero).
the five tested intonation types should be grouped into three general categories; question, prominence-rending, and insisting.

More specifically, *insist* and *rise-fall* are sub-categories of ‘insisting type’ BPMs. They are both perceived by native listeners of Japanese as conveying the speaker’s intention to “continue talking” and to “insist on the information,” and were also rated as strongly “emphatic” by the native listeners. The two sub-categories of ‘question type’ BPMs, *incredulity question rise* and *information question rise*, are perceived as conveying the speaker’s intention of “confirming information.” The connotation of “insistent” and “emphatic” are not implicated at all by these two BPMs.

This study is important because all previous work on acquisition of Japanese as a foreign language had focused only on the contrast between rising and falling boundary pitch movements for question versus statement functions, as in the underlined utterances from a dialogue used in research by Ayusawa and her colleagues (Nishimura et al., 1996) on the acquisition of *gimonbun* (疑問文: “question sentences”).

[1] A: あれ、なに？
[2] B: どれ？
[4] B: あれ？それは電話。
[6] : そう。電話。電話の写真。
[7] : へえ。だれが作ったの？あなた、それとも山田さん?
Ayusawa and colleagues varied the target word among unaccented *denwa, gakkoo, okurimono*, final-accented *yama’, zibiki’,* initial-accented *ne’ko, ho’n, me’gane*, and so on, in order to understand the interaction between BPM and lexical accent type. However, they did not look at any other BPMs than these, which are probably best described as either the incredQ or infoQ rise (functioning as an echo question) versus the falling BPM of an ordinary declarative. The research by Venditti and her colleagues, however, demonstrates clearly Kawakami’s point that there are several different types of rising BPMs, which convey various different speaker intentions, not all of which can be captured by the notion of *gimonbun*. Clearly, then, research is needed to understand better this full variety of rising BPM patterns and to see how they are mastered in learning Japanese as a foreign language.

### 3.2.2 Contrast between InfoQ and IncredQ rises

Let us first examine the contrast between InfoQ and IncredQ rises. As for their semantic characteristics, while an utterance produced with InfoQ would be perceived as a question, IncredQ will be perceived as an expression of the speaker’s strong disbelief in what s/he has just heard. Their acoustic characteristics are shown in Figures 3.8 and 3.9. Figures 3.8 and 3.9 illustrate the contrast between the InfoQ rise and the IncredQ rise, respectively. In order for the illustration to be easily understood, I have placed in each
figure a thick dotted line to indicate the bottom of the speaker’s pitch range, and an arrow at the right most position to indicate the degree of the final rise. These two BPMs typically differ from each other in two respects. First, The IncredQ rise contour is characterized by a greater range of rise in the utterance final position than that of the InfoQ rise. In the two figures below, while the IncredQ contour had a rise rise of 200 Hz, the rise in InfoQ was about 160 Hz. In addition, IncredQ tends to be characterized by a lower F0 value at the beginning of the utterance. This trend is easily noticeable in Figure 3.9.
[hontoo ni taberare’ru no?]
“Is this really edible?”

Figure 3.8: Pitch Contour of the Utterance with Information Question Rise

[hontoo ni taberare’ru no???]
“Is this really edible???”

Figure 3.9: Pitch Contour of the Utterance with Incredulity Question Rise
3.2.3 Contrast between InfoQ and Insisting rises

Let us examine the contrast between InfoQ and Insisting Rise. While InfoQ is perceived as a question, an utterance ending with the Insisting rise will be perceived as an emphatic statement, not a question. Figures 3.10 and 3.11 illustrate the contrast between the InfoQ rise (Figure 3.10) and the Insisting rise (Figure 3.11). The difference between these two contours is more localized to utterance final position. While the f0 of InfoQ rises to the top of the speakers pitch range, as indicated by the solid arrow in Figure 3.10, the f0 of the Insisting rise reaches only to the mid point of the speaker’s pitch range, as indicated by the dotted arrow in Figure 3.11. Perceptually, the Insisting rise differs from InfoQ in that it is often accompanied by an abrupt glottal stop.
“Is this really the one from Nara?”

Figure 3.10: Pitch Contour of the Utterance with Information Question Rise (2)

“This is really the one from Nara!”

Figure 3.11: Pitch Contour of the Utterance with Insisting Rise
3.2.4 Rationale for the selection of experimental focus

These two contrasts were used as the stimuli for the experiment. I was interested in these two types of contrasts because I predicted that there would be some difference in the rate at which non-native speakers of Japanese acquire the ability to perceive and discriminate between these two types of contrasts. Since English language has only a two-way distinction, namely falling BMP and rising BMP, and thus does not utilize finer distinctions within the rising BMP, I predicted that the similarities between the rising shapes of InfoQ and IncredQ would present some difficulty to learners to discriminate between these two BPMs. In contrast, because the contrast between InfoQ and IncredQ differed in terms of the degree of rise, I expected that acquisition of the ability to discriminate these two might take place at a faster rate. The acquisition of the ability to distinguish between InfoQ and IncredQ was of interest, additionally, because the semantic properties associated with these rising BPMs influence the interpretation of the sentence final particles, ne. In the next section, I will briefly discuss the interaction between sentence final particles and BPMs.

3.3 Interaction between boundary pitch movements and sentence final particles

Examination of BPMs has significant implications also for the analysis of the functions of sentence final particles (SFPs, hereafter). An increasing number of researchers have focused on the interaction between sentence final particles and intonation and have attempted to explain the functions of SFPs in terms of their interactions with the intonation patterns. What are SFPs? Japanese language has a large
number of SFPs, such as $yo$, $ne$, $nee$, $ka$, $zo$, $sa$, $na$ (McGloin, 1990), which usually consist of one mora and are attached to utterances at the utterance final position to communicate certain nuances of the speaker. Researches have frequently pointed out the dialogue specific nature of SFPs, that they are used extremely frequently in spoken dialogues, and that their use in written text is limited to reported dialogues or letters and children’s literature emulating the spoken dialogues.

I will summarize the recent theories concerning SFPs with focus on two most widely used and therefore most well studied SFPs, $ne$ and $yo$. Note that, from this point on, I will use the plural form, SFPs, to refer only to these two particles unless otherwise indicated. The scope of the discussion will be limited to examination of the use of the SFPs at the sentence final position, not including the use of what Uyeno (Uyeno, 1971) terms the “interjective” use of SFPs at sentence medial positions.

The works on the SFPs published between 1975 and 1995 discussed the functions of SFPs mostly in two general frameworks of analysis: 1) some studies examined the functions of SFPs based on the attribution of information (Kamio, 1978, 1990, 1994, 1997; Masuoka, 1991; Oso, 1986) and 2) some others focused on the ‘communicative functions’ of SFPs (Izuhara, 1993; McGloin, 1990; Onoe, 1997). However neither the information attribution approach nor the communicative function approach was complete and sufficient in explaining all the functions of the SFPs. One of the most significant challenges of these two approaches was that they did not take into consideration the variations in intonation in utterances. Particularly, $ne$ shows apparent variation in meaning under variation in sentence final intonation, and such alternation of intonation
could negatively influence the acceptability judgment in discourse, other things being equal (Eda, 2000).

The past ten years has witnessed a significant shift in the focus of the analysis of SFPs. A number of researchers have attempted to understand the functions of *ne* and *yo* in light of their interaction with discourse and intonation patterns (Eda, 2000; Katagiri, 1997, 1998; Takubo & Kinsui, 1996, 1997). While there are differences in detail, they propose that intonation pattern carries certain semantic implications independently of the grammatical and lexical elements, SFPs being one of such elements, and that the meanings of utterances can be derived as a result of the interaction between the grammatical and lexical elements on the one hand and intonation patterns on the other. Examination of *ne* and *yo* thus focuses more on how their functions interact with the intonation patterns. The speaker’s intentions such as “question,” “expressing agreement,” “increasing rapport between the speaker and hearer,” etc., which hitherto had been seen as the functions of the SFPs, are derived as the result of such interaction. This point characterizes their works as significantly different from others which also discussed the functions of SFPs in relation to the intonation patterns, such as those of Oso (1986), Jorden (1987), and Hubbard (1992).

A proposal made by Eda (2000) may be regarded as an extension of the theories put forth by Takubo and Kinsui (1996; 1997) and Katagiri (1997; 1998) although it was motivated independently. Eda’s (2000) proposal utilizes the empirical data provided by Venditti and her colleagues, which were examined in the previous sections. The availability of the data allowed Eda (2000) to make a significant contribution to the
advancement of the analysis of SFPs because the analysis of SFPs had been limited due to the fact that they depended on the binary categorization of pitch patterns, rising or falling, which had tended to be based on impressionistic analysis.

Eda’s (2000) proposal for a new approach to the analysis of the SFPs consisted of two main elements. First, it assumes that the “meaning” of each use of *ne* and *yo* can be derived as the result of the interaction between the lexical properties of *ne* and *yo* on the one hand, and the sentence final intonation pattern, on the other. The lexical properties of *ne* and *yo* are defined as follows (173):

- *Yo* is used to indicate to the hearer that the speaker has already accepted the prepositional content of the utterance.
- *Ne* is a discourse marker which signals the speaker’s expectation / request for an immediate response from the hearer, and also indicates that the speaker is leaving some room for the hearer to make comments on the prepositional content of the utterance.

The above stated definitions were motivated by Takubo and Kinsui (1997), but an additional statement has been added to identify *ne* as a marker to signal the speaker’s request for an immediate response, as well as the speakers stance toward the prepositional content of the utterance. This addition will account for why the use of *ne* is obligatory in some contexts.

### 3.3.1 Further examination of SFP *ne*

Now let us examine two utterances with SFP *ne* which differ minimally in the aspect of intonation: *Hontoo ni Nara no na no ne (✓)* “It is really [the one] from Nara, right?” and *Hontoo ni Nara no na no ne (!)* “It is really [the one] from Nara, you see?”
A native speaker of Japanese would paraphrase the first utterance with full-rise intonation (√) as something like the following: the speaker thinks that the item under discussion is from Nara, but since she is not sure whether this is really the case, she wants the hearer to confirm her assumption. On the other hand the latter utterance with ! intonation will be interpreted as follows: the speaker and the hearer probably share the information, and the speaker is making a comment, assuming that the hearer would also be aware, that the item under discussion is from Nara. The F0 contours of these utterances are shown in Figures 3.12 and 3.13 respectively. A native speaker of standard Japanese who was naïve to the purpose of the study read the target sentences printed on a card. Before reading the sentences, she heard a brief description of the context. Each of the target sentences was thus produced in a carefully controlled context. The speaker was instructed to pay attention to the differences in the two contexts in which the two utterances were made, but no explicit instruction was given in regard to manipulation of the pitch modulation.
Figure 3.12: An utterance with √ intonation + ne

Figure 3.13: An utterance with ! intonation + ne
A brief examination would reveal that the variation in meaning of *ne* in these utterances could be attributed to the sentence final intonation patterns associated with them. Furthermore, observing the parallels between the two intonational minimal pairs, it is reasonable to predict that the speaker’s intentions such as “confirming the assumption” and “explaining, or making a statement,” are being communicated by the intonation patterns. This is strong evidence to show that functions such as “confirming” and “commenting,” which have been argued in the past to be functions of *ne* are actually obtained by the interaction between *ne* and the intonation patterns with which it appears.

Studies such as those of Inukai (2001), Moriyama (2001), and Sugito (2001) examined the characteristics of *ne* by focusing on the difference in intonation. Note that Inukai (2001), Moriyama (2001), and Sugito (2001) consider intonation to be an intrinsic property of the SFP *ne*; they do not necessarily propose to examine the functions of the SFP *ne* and intonation patterns independently of each other. Their view is reflected in their use of the term “intonation of the sentence final particle *ne*” (終助詞「ね」のイントネーション).11 Despite the difference in the framework, their findings are consistent with the argument presented in Eda (2000) that there are two clearly different uses of *ne* when it is “spoken with a high-pitch intonation” (corresponding to the use of *ne* with √ intonation) or “low-pitch intonation and short duration” (corresponding to the use of *ne* with ! intonation) (Inukai, 2001), and provide a window thorough which to appreciate the functions of BPMs.

11 Moriyama (2001), however, suggests that study of the SFP *ne* should be extended to examine how much of the ‘meaning of *ne*’ comes from the more general function of intonation itself, which he terms “the Intonation Constraint” (イントネーションの制約の一般化) (p.46-47).

71
Inukai’s study examined the use of *ne* with “low-pitch intonation and short duration” by 14 speakers (four male, ten female) of a dialect of Japanese spoken in Aichi prefecture. His study focuses on whether speakers of Aichi prefecture used this variant of *ne* with low-pitch intonation (his prediction was that its use was limited to speakers of the standard dialect) and whether there is a gender difference in its use (his prediction was it would be used mostly by male speakers). His data confirmed the first prediction; while all subjects could perform tasks to distinguish the implications of the two different types of *ne*, only four of the subjects could re-produce the utterances accurately. Three female subjects reported that they used it, but when asked to produce the utterance, they failed to do so with consistency. Inukai’s second prediction was not confirmed. While he had predicted that *ne* with low-rise pitch would be used mostly by male speakers, three of the four subjects who could produce this variance of *ne* were female.

Sugito termed the *ne* with √ rise as “*ne* that expects responses from the hearer” (聞き手の反応を求める「ね」) and the *ne* with ! rise as “*ne* to signal speaker’s self-consent” (自己確認の「ね」) (2001). Her choice of these terms clearly reflects her view of the functions of these instances of *ne*. Using a set of ten stimuli, she examined the perception of intonational contrast by native speakers of Japanese.\(^\text{12}\) The stimuli were created from an utterance *soo desu ne* “that is so, isn’t it” with artificially manipulated

\(^{12}\) Three subjects participated in this perception study. No detailed description was given about the subjects. It is not clear whether or not they were native speakers of standard dialect.
final pitch patterns ranging from a full rise to the top of the pitch range of the utterance (indicated as \(\text{\textcircled{1}}\) in Figure 3.21) to a non-rise (indicated as \(\text{\textcircled{11}}\)).

The top panel in Figure 3.14 shows the F0 contours of the set of stimuli (Set A) with the final rise starting at the onset of [n]; the bottom panel shows the F0 contours of the set of stimuli (Set B) with the final rise starting at the end of [n], i.e., at the onset of [e]. Each set of stimuli were presented in a random order, and three native speakers of Japanese performed a forced choice perception task by choosing if the item presented was “ne that expects responses from the hearer” or “ne to signal speaker’s self-consent.” The test was repeated 10 times for each subject.

The responses of the subjects are summarized on the right of the set of numbers using vertical lines in each panel. For both Set A and Set B, the stimuli with rise to a higher pitch peak were judged as “ne that expects responses from the hearer.” For each set of stimuli, there were items that received mixed evaluation (items 4, 5, 6 in Set A and 6, 7 in Set B).
Figure 3.14: Interaction between sentence final particle *ne* and boundary pitch movement, reproduced from Sugito (2001), p. 11
Sugito’s data implies there is a boundary at which the perception of meaning changes from one meaning to another. This question of whether there is a boundary for the two types of rising intonation comparable to the category boundary between two phonemes (such as /l/ and /r/) is an interesting and intriguing one. Moriyama’s (2001) data, for instance, suggested that even full rising intonation can be perceived as an utterance for what Sugito terms the “ne to signal the speaker’s self-content,” thus indicating that the interpretation is a matter of degree and not clearly divided into categories. Further research will be necessary to answer this very important and interesting question.

Moriyama (2001) also empirically examined the distinction between ne with high- and low-short-rises, but focused more on the instances of ne with falling intonation and the pragmatic implications of falling intonation. He pointed out that the “when [ne is] spoken with a falling intonation, it can be used as confirmation … but it sounds somehow rude because of the lack of ostensive marking of confirmation” (p. 284).

3.3.2 Further examination of SFP yo

The SFP yo appears to differ significantly from ne in one respect, that differences between yo with an infoQ rise and yo with a fall are not as easily captured in words as the analogous difference for ne. Figure 3.14 below shows the F0 contour of an utterance, hontoo ni nara no na no yo (√) “It is really [the one] from Nara, I am telling you.”
Comparing Figure 3.15 with Figures 3.12 and 3.13 (on page 69) provides evidence that while the SFP *yo* is procured with a full rising intonation, the utterance clearly conveys the sense that the speaker is “explaining” to the hearer the fact that the item under discussion is from Nara, and that the speaker is being emphatic about the information being communicated.

Tanaka and Kubozono (1999) point out that if *yo* is used with a falling intonation, it will be perceived as a statement that highlights the “speaker’s surprise, disappointment and/or frustration due to the difference of the opinion between the speaker and the hearer” (p. 122). For example, the two utterances in example 4 both mean that they have an appointment at eight o’clock, but (4b) will be more appropriate in a situation where hearer may have made a mistake about the time of the appointment and the speaker is thus pointing out the hearer’s mistake with an implication of annoyance.
The examination of interaction between SFPs and intonation is intriguing. Learners of Japanese will benefit from instructions on the use of SFPs especially in relation to intonation.
CHAPTER 4

THE EXPERIMENTS

4.1 Experiments

This chapter describes the procedures and results of two experiments, which examined the processing of three types of intonationally marked contrasts by native and non-native speakers of Japanese. The contrasts are chosen because, while they are similar in terms of their f0 shapes, their semantic properties are diametrically opposite of each other.

4.2 Research Questions

The first experiment was designed in order to address two sets of questions. The first set of questions concerns differences across groups of speakers of different dialects and languages:

A. Do native speakers of different dialects of Japanese perceive and interpret variations in intonation patterns in the same way?

B. Do native speakers of English perceive and interpret variations in intonation patterns in Japanese in the same way that the native speakers of Japanese do?
C. If they do not, what are the qualitative differences between the way native speakers of English and speakers of different dialects of Japanese process intonation?

The second set of questions concerns stages of acquisition by native speakers of English who are learning Japanese as a second language (L2), and any learning trends that can be observed:

D. Do those who have had longer exposure to Japanese demonstrate a greater degree of accuracy in their perception and interpretation of intonation patterns?

E. Are there any patterns of errors among L2 learners of Japanese?

F. Can such patterns be attributed to the differences between the prosodic structures of Japanese and English?

4.3 Hypotheses

The first, and most obvious prediction is that there will be qualitative as well as quantitative differences in the perception and interpretation of the variation of intonation contrasts by native and non-native users of Japanese. A related prediction is that two groups of native speakers of Japanese would exhibit identical or at least similar performances on the discrimination and identification tasks. Although the native speakers of other dialects of Japanese may not necessarily be native speakers of the standard dialect, they will have had considerable exposure to the standard dialect through the education system, television, and other means. Therefore, I predicted that they would have developed perceptive competence in comprehending intonation patterns. Finally, it
is predicted that non-native speakers who have had longer experience with the language will behave more similarly to the native speakers than those who are at beginning levels of learning Japanese, both in their perception and interpretation of intonation patterns.

4.4 Experiment 1

4.4.1 Materials

A total of 36 sets of utterances were used to test the processing of three different types of prosodic contrasts. Each set of stimuli consisted of a base sentence produced with two different intonational patterns. A female native speaker of standard Japanese with training in linguistics read each sentence four times, creating two tokens each of the two intonationally contrasting token types. Thus, a total of 144 tokens (36 base sentences \( \times \) 2 token types \( \times \) 2 repetitions) were recorded. All stimuli were recorded in one session, in a sound booth using a head mounted microphone. The utterances were digitally recorded on at 22.05 kHz and 16-bit precision. The researcher was present during the recording session to make sure that the intended contrasts were clearly audible in the stimuli.

The sentences are listed in the appendix. Most of the base sentences were created using vocabulary items and grammatical structures already familiar to all of the non-native subjects at the time of the study.\(^{13}\) The base sentences used to test the perception of prosodic phrasing were six syntactically ambiguous sentences and six pragmatically ambiguous sentences, consisting of four accentual phrases (APs). Special attention was

\(^{13}\) Two sentences were exceptional in that they contained vocabulary items that were not familiar to the L2 subjects who were enrolled in the beginning level Japanese course.
paid so that all accentual phrases were accented, i.e., contained a pitch accent. The speaker read the base sentences in two intonationally contrasting ways, once with an intonational phrase boundary between the first and second APs, and once with a phrase boundary between the second and third APs. The base sentences for testing the understanding of the functions of boundary pitch movements (BPMs) were pragmatically ambiguous, but no particular effort was made to control the accentuation of lexical items because the word sequence was the same for all utterance types in a set, so that if there was any effect of having an accented versus unaccented word just before the sentence final particle on the shape of the BPMs, this effect would be true for all three BPMs in any set.

The listener’s comprehension was tested by a forced-choice paraphrasing task, and a set of three paraphrases was created for each sentence. Two expressed the different meanings for the sentence when uttered with the two target prosodic patterns, and the third was a distracter that did not match either utterance. The researcher prepared the paraphrasing sentences first in English, and then translated them into Japanese. Three native speakers of Japanese checked the Japanese translations to ensure that the translations were faithful to the original English sentences, and sounded natural in Japanese.

Authorware (MacroMedia) was used to create the experimental module so that the timing of the presentation of audio and visual stimuli could be controlled.
4.4.2 Subjects

The experiment involved three groups of subjects: 36 adult learners of Japanese who were native speakers of English (L2), 20 adult native speakers of standard Japanese (NS-Tokyo), and 31 adult native speakers of Japanese who spoke other dialects (NS-Other).

Most of the L2 subjects were students enrolled in the Japanese language courses at The Ohio State University (1st, 2nd, 3rd and 4th year). They participated in this study as a part of their language courses, which emphasize the development of oral communication skills. Additionally, some students in the graduate program in Japanese volunteered to participate. One difficult task was how to objectively measure the proficiency levels of the L2 subjects. For instance, even among the first year students, language background varied considerably from those who started formal study of the language with no prior experience to those who had lived in Japan for some time before enrolling in the Japanese language course. It was obvious that the extent of formal study or enrollment in certain courses would not adequately capture the qualitative difference between ten weeks of classroom instruction and a ten-week stay in Japan with a Japanese family. Therefore, I used their scores on the Japanese Skills Test (JSKIT) as an index of their proficiency level. JSKIT is a test being developed by the National Foreign Language Center as a comprehensive proficiency test.

The native speaker subjects were recruited in two different locations. One group of subjects consisted of participants of a program for teaching Japanese as a foreign language, being held in a suburban area outside Philadelphia, PA. They had completed
their undergraduate education in Japanese universities, and had lived in the U.S. for six weeks prior to their participation of this study. They volunteered to participate in the study. The other group of subjects was composed of undergraduate students who were enrolled in a Japanese university, and studying at a branch campus located in West Virginia. The length of their stay in the U.S. ranged from three months to two years. Each subject received $10 for participating in this study.

All Japanese participants filled out a questionnaire. Since the data collected at two locations of the experiment did not exhibit significant difference, their data were combined, and then re-divided into two based on whether the subject was a native speaker of the dialect spoken in and around Tokyo area. Those who were born in either Tokyo or Kanagawa and had lived there for at least the first ten years of their life were categorized into this group. Speakers of all other dialects were grouped together as native speakers of other dialects. The latter group was comprised of individuals from a wide range of places including Osaka, Kyoto, and Nagoya.

4.4.3 Procedure

The 36 pairs of utterances were divided into two sets, so that each set consisted of three blocks of different types of prosodic contrasts. Within each set, the three blocks were presented randomly, and within each block, the six utterances were presented randomly. The presentation order of the sets, however, was controlled so that one half of the subjects heard set 1 first, followed by set 2, and the other half of the subjects heard the sets in a reverse order.
In each test, one half of the trials presented pairs of utterances that were different intonational contrasts, and the other half two tokens of the same type. In order to avoid a possible item effect, two sets of test modules were created so that sentences that were presented as intonational minimal pairs in one module could be presented in the other module as two tokens of the same type. Subjects were randomly assigned to the four testing conditions (two presentation orders and two testing modules).

Each subject was seated in front of a computer to take the test. All the visual stimuli were presented on the computer screen, while the audio stimuli were presented through headphones. The instructions and four practice sentences were also provided on the computer. Subjects were told to pay attention to the variation in the prosodic aspects of utterances, and were instructed to note the difference in ‘nuances’ as they listened. After listening to the instructions and trying four examples, they were asked to proceed at their own pace.

Subjects performed two tasks, a discrimination task and an identification task, successively on 36 pairs of utterances. First, they heard a pair of two utterances twice, and were asked whether or not the two sentences communicated the same intentions of the speaker (discrimination task). The stimuli were presented only aurally, and no visual cues were provided. Subjects were then presented with a set of three sentences written in their native language, and were asked to perform an identification task, i.e., choosing a paraphrase from the choices on the screen. They were given another chance to listen to the sentences before making their choices. If a subject judged the two sentences to be the same in the discrimination task, the identification task was to choose one of these
paraphrases. If a subject judged the two sentences to be different in the discrimination task, the same set of three paraphrasing sentences was used as potential responses for both utterances, with the subject first choosing one of the three sentences to describe the meaning of the first utterance, and then choosing one out of the remaining two sentences for the second.

The subjects' responses were scored in terms of the accuracy of perception and of comprehension, based on the number of correct responses in the discrimination task and in the identification task, respectively.

4.4.4 Results

Figures 4.1 and 4.2 compare the performance of the NS-Tokyo group against that of the NS-Others group for all 36 pairs of utterances. Each data point in the plot represents the subjects' responses to one pair of utterances averaged across all the subjects. Figure 4.1 plots accuracy on the discrimination task and Figure 4.2 plots accuracy on the identification task. If the responses of the two groups were identical, the points in these plots would cluster near the $y = x$ line, or at least be distributed randomly above and below the line. For the discrimination task, the data points appear equally spread on both sides of the line. A Wilcoxon signed rank test confirmed that the two groups of native speakers of Japanese did not differ significantly in their discrimination task performances ($Z=-0.707$, $p=0.24$).
Figure 4.1: Comparison between native speakers of Tokyo dialect and speakers of other dialects in the discrimination task
Figure 4.2: Comparison between native speakers of Tokyo dialect and speakers of other dialects in the identification task
Figure 4.2 compares their identification task performances. The average score for the NS-Tokyo group was higher than that of the NS-Others group for most of the 36 pairs of utterances, and the data points appear consistently below the $y = x$ line. A Wilcoxon signed rank test confirms that the two groups of native speakers of Japanese differed significantly ($Z = -2.91$, $p < 0.01^*$).\footnote{This is a by-items analysis, of course, and a more complete analysis would also provide an analysis by subjects. A comparison of histograms of scores by subjects, however, shows that the differences will be significant in a by-subjects analysis.} \footnote{While 0.05 is commonly used as the cutoff point in considering the criterion for significance, I will use a more conservative criterion of 0.01 because these are multiple analyses of the same set of speakers.}

Because of this difference between the two native speaker groups, the performance of the L2 group was compared with that of NS-Tokyo group only. In what follows, the term ‘NS group’ will refer only to the group of native speakers of Tokyo dialect.

Figures 4.3 and 4.4 compare the accuracy of perception of the L2 group against that of the NS group in their discrimination (Figure 4.3) and identification (Figure 4.4) tasks for 36 pairs of utterances. Similarly to Figures 4.1 and 4.2, each data point in the Figures 4.3 and 4.4 represents the subjects’ responses to one pair of utterances averaged across all the subjects. The maximum possible score was one and two for discrimination and identification tasks, respectively. In both figures, the data points tend to appear below the $y = x$ line, indicating that the value of $y$, representing the task performance of the L2 subjects, was consistently smaller than the value of $x$, representing the task performance of the NS subjects. A Wilcoxon test confirmed that these two subject groups differed
significantly in both discrimination task ($Z=0.64$, $p^* < 0.01$) and identification task ($Z=-2.55$, $p^*<0.01$).

Figure 4.3: Comparison between native speakers and non-native speakers of Japanese in the discrimination task
Figure 4.4: Comparison between native speakers and non-native speakers of Japanese in the identification task
Figure 4.5 and 4.6 summarize the performances of the 36 L2 subjects in discrimination (Figure 4.5) and identification (Figure 4.6) tasks. Each data point represents the total number of correct responses of the L2 subjects, plotted against their language proficiency test scores on the x-axis. The maximum score for their task performance (y-axis) was 36, and that of JSKIT (x-axis) was 100. A regression line fitted to all the data points in Figure 4.5 accounts for 0% of the data (R-squared=0.00). The correlation between the subjects’ performance of the identification task and their language proficiency as shown in Figure 4.6 was significant (R-square=0.20, p*<0.01).
Figure 4.5: Non-native speakers’ performance in the discrimination task plotted against their Japanese proficiency test (JSKIT) scores
Thus, the data indicates that learners of Japanese performed the discrimination task well regardless of their level of proficiency. In contrast, when they were asked to associate the intonational variation with the intended meaning of the utterance, their task performance had clear correlation with their level of language proficiency.
Experiment 2

The second study examined the non-native perception of Japanese intonation, with focus on the variation of BPMs. This study compared three types of rising BPMs (infoQ, incredQ, insist) and falling BPM. Especially of interest was a fuller exploration of the contrast between infoQ (full rise to the top of the pitch range) and insist (rise to the mid-point of the pitch range). Assuming that these three BPMs were different types of rising intonation, we hypothesized that they would present differing degrees of difficulty to NNS of Japanese. Specifically, it was predicted that insistQ would be more difficult for native speakers of English since there is no comparable distinction in English, thus requiring them to create a new category for the contrasting intonation patterns without using an analogy to English.

4.4.5 Materials

A native speaker of Tokyo Japanese recorded the stimuli following a procedure described in Section 4.3.1.

4.4.6 Subjects

A total of 32 native speakers of English participated in this study. All the subjects were participating in a 9-week intensive Japanese program in a total-immersion environment in the United States, and the experiment was conducted during the first two weeks of the program. All the participants took the same proficiency test (JSKIT) at the beginning and end of the program, except for the first year students, who did not take it at the beginning. Although this study was administered at the beginning of the language
program, we used the second set of the JSKIT scores for the analysis because about one third of the participants of the experiment were first year students. Since there was a significant correlation between the two sets of test scores (r=0.947, n=22, p<0.01*), we felt justified in using the second set of JSKIT scores.

4.4.7 Procedure

This study involved the same tasks as Experiment 1, but the format of the presentation was modified so that subjects performed the two tasks separately. They heard 36 pairs of utterances for discrimination; when they had completed the discrimination tasks, they heard 40 sentences which were not used in the discrimination task and performed identification tasks.

4.4.8 Results

The data from this study showed the same trend as the data in the previous study. Figure 4.7 shows scatter plots of the subjects’ scores in discrimination (top) and identification (bottom) against their JSKIT scores. Note the wide variance across the level of Japanese proficiency in the discrimination task (top). Although the regression line in this figure appears to have a slight down-slope, taking into consideration the sizes of variance, intercept and slope, we concluded that there was no learning trend in the subjects’ discrimination task performance.

As discussed in Section 4.4.4, while no correlation is found between accuracy in discrimination and language proficiency (r=-0.04, n=32, p>0.10), accuracy in associating appropriate interpretations with various BPMs is strongly correlated with the subjects language proficiency (r=0.59, n=32, p<0.01*).
Figure 4.7: Non-native speaker scores in discrimination (top) and identification (bottom) tasks, plotted against their Japanese proficiency test (JSKIT) scores.
Figure 4.8 shows the percent of correct discrimination of *falling* (*Fall*), *insist rise* (*CC*) and *incredulous rise* (*QQ*) in the condition in which the tokens were presented with another token of the same type (“same” condition). Error bars are placed to indicate 5% margin of error. The meaning of the falling intonation (Fall-Fall) was correctly identified less often than that of Incredulity Question (QQ-QQ), although the difference was not significant by a Wilcoxon signed rank test ($Z=-1.56$, $p=0.06$). The meaning of the Incredulity Question (QQ-QQ), in turn, was correctly identified less often than that of Insisting Rising (CC-CC), although the difference between the two conditions was not significant by a Wilcoxon signed rank test ($Z=-1.11$, $p=0.13$). We can see that the discrimination is quite good in this condition for all token types.

![Figure 4.8: Percent correct discrimination of falling, insist rise (CC) and incredulous rise (QQ) in the “same” condition.](image-url)
Figure 4.9 shows the percent of correct discrimination in the condition in which the tokens were presented as intonational minimal pairs. The data is plotted from left to right in the order of difficulty. Interestingly, the most ‘difficult’ pair turned out to be the contrast between *falling* and *insist*. The Wilcoxon signed rank test showed that, while there was a significant difference between F-CC and CC-Q conditions ($Z=-2.79$, $p^*<0.01$) and between F-Q and F-QQ conditions ($Z=-2.42$, $p^*<0.01$), there was no significant difference between CC-Q and F-Q conditions ($Z=-0.94$, $p=0.18$). This means that learners of Japanese tended to group falling intonation and insisting rise intonation.
together. Thus the data implies that, contrary to the original hypothesis, *insist rise* is more difficult to distinguish from, and more easily confused with, *falling* than it is with *infoQ rise*.

### 4.4.9 Analysis

The data from the NNS did not support my hypothesis that the similarity of F0 shapes of the two rising BMPs, *infoQ* and *insist*, and the lack of a comparable distinction in English, would make them difficult to discriminate from each other for native speakers of English. Contrary to our prediction, native speakers perceived *insist rise* as more similar to *falling* than to *infoQ rise*. There are two possible ways to account for this result. First, it might be due to the fact that *insist rise* has similar semantic functions as *falling* BPM, which makes them different from *infoQ*, or, secondly, it might be due to the F0 shape of *insist rise*, which reaches only to the mid-point of the pitch range. This is an issue that requires more exploration in the future, and such examination will need to address the question of to what extent the two factors, similarity of F0 shape on the one hand, and the functions of the BPM, on the other, play roles in the perception of intonation patterns.

### 4.5 Discussion

This study used discrimination and identification tasks to examine how three groups of listeners perceive and process the intonational structure of utterances in Japanese. There were two predictions. First, non-native speakers of Japanese should exhibit significant differences from native speakers, second, speakers of Standard
Japanese and of other dialects of Japanese should perform the tasks similarly to each other. The comparison between native and non-native performance in discrimination and identification tasks illustrated in Figures 4.3 and 4.4 provides strong support for the first prediction.

The native speakers of Japanese were consistently more accurate than non-native speakers of Japanese in perceiving the intonational differences and interpreting them. These results confirm the effects reported by Cruz-Ferreira (1987), who demonstrated that non-native speakers of English and Portuguese perceived and interpreted English and Portuguese differently from native speakers. As for the questions concerning the patterns of acquisition, are there any noticeable patterns of errors among non-native speakers? Can such patterns be attributed to the differences between the prosodic structures of Japanese and English? In order to answer these questions, several statistical comparisons were made, but no clear patterns of interference were observed. These questions, therefore, will need to be addressed in future studies.

The comparison between native speakers of Standard Japanese and of other dialects of Japanese illustrated in Figures 4.1 and 4.2 provides only a partial support for the second prediction. While the two groups of Japanese performed similarly in the discrimination task, speakers of Standard Japanese, on average, outperformed the speakers of other dialects in the interpretation task, suggesting that speakers of other dialects do not completely learn the meanings of BPMs and phrasal pitch range variation of Standard Japanese, despite daily exposure to it through the education system and mass media. Similar effects were reported in Kôri (1992), who examined how differences in
the intonational variations between Tokyo and Osaka dialects influenced the perceived degree of politeness and sincerity. It is important to note that, different from Kôri’s (1992) study, the data shown in the current paper only indicates that the native speakers of other dialects did not agree with the interpretations based on Standard Japanese. This may be due to the fact that subjects with diverse linguistic background (i.e., dialects of Japanese) were grouped altogether as “speakers of other dialects” in this experiment. Thus the data do not tell us how they differed from the native speakers of Standard dialect. Further study will be necessary to characterize the qualitative differences among native speakers of different dialects of Japanese.

The second goal of this study was to characterize how L2 speakers of Japanese acquired such abilities. The data presented in Figures 4.5 and 4.6 showed an interesting pattern of acquisition. While a clear learning trend is observed in the subjects’ identification task performance, such a learning trend is not observable in their discrimination task. Note the subjects’ high performance in the discrimination task. No apparent learning trend is observed in the L2 discrimination task performance due to the ceiling effect, that the subjects, even with low-level language proficiency, had already arrived at a high level of competence in the discrimination task. The data clearly indicate that non-native speakers of Japanese acquire these two aspects of language processing at different rates.

It is interesting to note that this apparent ‘anomaly’ seems to apply even to the native speakers of Japanese who speak different dialects of Japanese. Together, the results of this experiment suggest arguably that the two processes, recognizing
intonational cues on the one hand, and associating the correct interpretations with perceived intonational variations on the other, involve qualitatively different processes. Further research is necessary to determine how these levels of processing are acquired by native and L2 speakers of Japanese, and how they interact with each other in sentence processing.

To summarize, I have discussed three salient results obtained in the experiment. First, there was a statistically significant difference between the performance of native and non-native speakers in both the discrimination and the interpretation tasks. Second, among the native speakers of Japanese, native speakers of the Tokyo dialect and speakers of other dialects performed differently. Finally, while non-native speakers performed equally well in the discrimination task regardless of their proficiency, their performance in the identification task was strongly predicted by their proficiency. Together, these results support the idea that recognizing prosodic cues is easier than associating the correct interpretations with them. I have pointed out the parallel phenomenon discussed in first language acquisition studies, and proposed that acquisition of accurate production of prosodic effects and that of adult / native like competence in pragmatic use involve different cognitive processes.

4.6 Limitations of This Study

It is important to recognize the weaknesses of the study. I will identify two areas in which the current study should be improved to increase the applicability of the data to a more general context. The first of such areas is that the stimuli used in this experiment were prepared using utterances produced by only two speakers, both of whom were
female. This study, therefore, needs to be followed up by a similar experiment using speech samples obtained from native speakers of the Tokyo dialect, including both men and women of various ages, to see whether similar effect is obtained. Second, this experiment should be repeated to examine differences between native speakers of standard Japanese vs. speakers of other dialects. For the purpose of the experiment, I grouped the subjects in Experiment 1 as NS group and NNS group depending on whether they were native speakers of the standard dialect (NS group) or not (NS-Other group). As a result, the NS-Other group included speakers of various dialects. Another experiment will be necessary to examine the effect of their native dialect on performance in pragmatic processing.
CHAPTER 5

APPLICATION OF FINDINGS TO TEACHING JAPANESE AS A FOREIGN LANGUAGE

One of the more long-term goals of the present investigation is to improve Japanese language curriculum design. The empirical data gathered through experiments, described in Chapter 4, suggest three basic conclusions we can make about the acquisition by native and non-native speakers of Japanese of the ability to perceive intonation and to interpret the socio-pragmatic nuances characterized by the various intonation patterns: 1) the ability to perceive intonational variations and the ability to associate interpretations to such variations involve two different levels of processing; 2) second language (or second dialect) speakers’ control of the pragmatics of intonation develops at a slower rate than their ability to perceive intonational variations; and 3) native speakers of dialects other than the Tokyo dialect may not have the same pragmatic competence in implications of intonational variations in the standard dialect spoken around the Tokyo area.

In this final chapter, I will attempt to consolidate the theoretical research and practical application in a coherent manner; based on the conclusion above as pedagogical
rationale, I make practical suggestions for incorporating explicit and implicit instructions into the Japanese language curriculum that will facilitate the acquisition of Japanese intonation. It is hoped that such training will serve as an essential tool for learners of Japanese to help them become more effective in spoken language communication. The discussion will also include a design for an instructional material that will most effectively facilitate this process. In particular, I will present the design of a computer-assisted language-learning (CALL) program and discuss how such material can be effectively integrated and utilized in the curriculum.

Note that the current discussion is intended neither as the presentation of a single course whose sole objective is to teach Japanese intonation, nor as a statement of specific instructional procedures and methodologies. Rather, it assumes the existence of a well-established language curriculum, and focuses on ways to add to the curriculum elements that will enhance the learning of intonation. The analysis focuses on the teaching and learning of Japanese in universities in the United States. This means that the current discussion assumes learners who are cognitively mature, who possesses the native or fluent second-language knowledge of the English language, and uses examples based on comparisons between the phonological features of English and Japanese. However, the suggestions presented in this chapter should be applicable to other teaching / learning contexts as well, so long as the learners’ cognitive, linguistic and cultural backgrounds are taken into consideration in adapting them to the specific learning conditions.
5.1 Theoretical rationale into pedagogical practice

The experimental results presented in Chapter Four allow us to gain deeper insight and understanding of the acquisition of intonation in Japanese. Focusing on the listeners’ perspective, I have identified two related but distinct abilities that contribute to the development of competence in intonation: the ability to perceive the variations in intonation, on the one hand, and the ability to associate interpretations to such variations, on the other. From the speaker’s point of view, the two abilities can be contrasted as fluent production of a certain pitch pattern, on the one hand, and pragmatic competence for felicitous use of such pitch pattern, on the other. The data show that learners’ ability to associate intended meanings with the variations in intonation (pragmatic processing) develops at a slower rate than their ability to perceive variations in intonation (phonetic processing). The data also demonstrate that native speakers of Japanese who speak non-standard dialects may not have the same system of associating pragmatic implications as the speakers of standard Japanese when dealing with certain types of contrasting intonation. This indicates that the acquisition of the pragmatics of intonational variation is a process that develops through one’s exposure to, and experience with the target dialect.

5.1.1 Learners’ ability to accurately deal with intonation depends on their pragmatic and communicative competence.

The first, and most important factor to be taken into consideration in the discussion of instructional methods and materials is that instruction on intonation should be pointed toward the acquisition of pragmatic competence in interpreting the
implications of utterances, as well as in producing utterances with intonation patterns that will communicate the learners’ intentions in a felicitous context, rather than focusing on their ability to perceive and produce the pitch shape. Lack of concern for the pragmatics of intonation is the most significant shortcoming among all the existing pedagogical materials. They focus on dealing with the lower level of processing, i.e., perception and production of sounds and pitch patterns divorced from ‘meaning.’ My study strongly suggests, however, that exercises focusing only on phonetic processing will be insufficient to help the learners of Japanese acquire full competence in intonation. For instruction on intonation to be fully effective, it needs to direct the attention of the learners to both the variation in intonation itself (phonetic level processing) and to its association with contextual implications (pragmatic level processing).

The reason intonation exercises are frequently focused on mechanical reproduction of pitch patterns may be that prosodic aspects of utterances, including accent and intonation, have been addressed almost exclusively within the framework of teaching pronunciation (発音指導 hatsuon shido), and that teaching pronunciation has traditionally been equated with spending an inordinate amount of time learning to produce meaningless sound units in isolation. Textbooks that are designed specifically to teach Japanese pronunciation have tended to focus on the itemized articulations required to produce discrete sound units, such as vowels and consonants, in Japanese and combinations of those that form words (Fujito et al., 1979; Nakanishi, 1987; Sakuma, 1964; Suzuki, 1991), Tanaka (1999) being an exception to this tradition. This style of
teaching pronunciation is not a problem in itself, for that is, indeed, how one can improve his/her pronunciation. However, it is not sufficient for teaching intonation.

The data from the present study indicates that learners, especially at the lower level of proficiency, may be able to perceive or produce the intonation patterns accurately when they are instructed to do so, but this does not guarantee that they know the implications associated with such variations. The existence of this gap, indeed, may be the biggest challenge that we, as language instructors, may encounter in teaching intonation because whether the learner has acquired pragmatic competence in dealing with intonation is not as easily observable in their language behavior. In the case of production, the instructor may ask the learners to repeat a sentence, and if they are able to reproduce the prosodic patterns accurately, that is taken as a sign of acquisition. It does not mean, however, that the learners can reproduce the intonation pattern the next time in a felicitous context to communicate their intention. It may well be the case that the learners are even unaware that the intonation pattern in question was being used to communicate a certain intentions. The same is true for perception/comprehension. It may also happen that the learners can ‘interpret’ the utterances they hear in a communicative context, but there is no way for us to know if they are able to attribute the subtle implications to the varying intonation with which the utterance is made. It is highly likely that they are processing the utterance as a string of lexical item with certain underlying structural patterns, and the learners may not be using intonation as one of the ingredients for the composition of the ‘meaning’ of the utterance.
This point was discussed in Chapter 2, when examining the results of the experimental studies. When subjects were asked to pay attention to differences, their attentions would go primarily to the lexical and grammatical variation. It is highly likely that learners of Japanese may not start paying attention to intonation unless they are instructed to do so, especially to the aspect of pragmatics.

Learners of Japanese, assuming that they already have some training in and facility with the recognition of intonational variation, will benefit from exercises in which they are instructed to pay attention particularly to contrasting intonation patterns, and are asked to use intonation to create contrasting pragmatic implications. In the case of perception, they will be asked to ‘interpret’ the sentences. In the case of production, they will be asked to produce certain intonation patterns in appropriate contexts to communicate their implications. I will discuss this point in more details with examples in Section 5.2.

5.1.2 Both production and perception skills should be practiced.

As a consequence of the tendency to deal with accent and intonation exclusively in the context of “pronunciation” (発音指導 hatsuon sido), instruction on accent and intonation has tended to focus primarily on the production of accurate prosodic patterns. Such instructional materials are designed with an implicit assumption that teaching of pronunciation should focus on the accurate perception and production of the sound units of which Japanese words and sentences are composed. When developing new instructional materials for intonation, therefore, one should make sure that both aspects of
communication, i.e., perception and production, are practiced at the phonetic level and the pragmatic level.

5.1.3 **Intonation should be practiced using meaningful utterances in context.**

The data obtained from the empirical study demonstrates that the acquisition of skill in understanding the pragmatic implications associated with intonation patterns occurs at a slower pace than that of learners’ ability to perceive variations in intonation. This clearly indicates that the instruction on intonation should spend more time for the development of their pragmatic skills. In order for that to happen, it is essential that such exercises take place, using utterances that are meaningful in context, rather than using words and sentences in isolation.

The important point to be considered is not whether the items included in the exercises are words or sentences; the question is whether the items included in the exercises, both words and sentences, are being practiced as situated utterances. An utterance is defined in the glossary provided by an on-line resource provided by SIL (formerly known as Summer Institute of Linguistics) as a “string of speech found between breaths or pauses” which “involves particular intention … and production of a particular effect on the addressee as the result” (SIL). Moriyama (1997) discusses in detail the intonation patterns of one-word utterances, and points out that the pragmatics of one-word utterances in discourse, which do not have underlying syntactic structures, can only be communicated and interpreted by intonation. As he correctly points out, the importance of the roles played by intonation in one-word utterance communication cannot be overemphasized.
The pragmatic implications of intonation are practiced more effectively using utterances, when the learners employ units of expression such as words and sentences with the particular intention in making an utterance (production), or when an utterance directed toward them produces a particular effect as the result (perception). This is applicable to training of pronunciation of any kind, segmental or prosodic. The reason for including instruction on pronunciation is that correct pronunciation, including correct use of intonation, will help smooth communication. In pronunciation training, pronunciation is often isolated from everything else in the spoken language. In addition, as was discussed in Chapter 2.2.2, pronunciation in Japanese as a foreign/second language is frequently introduced in connection to, or perhaps even in order to, introduce the syllabaries. That may be a pedagogically sound strategy in that it is probably more effective to isolate a problem area and work on it first. Upon mastery of accurate pronunciation, however, the learners should be encouraged to learn to apply their new skills in the more realistic contexts of communication. That last step is often missing in instruction dealing with pronunciation and intonation. That even a word alone can be practiced as a word in isolation or as an utterance is important because the effect of these exercises will be different. To make their learning more contextualized, learners may be given opportunities to produce various kinds of reactions to match the implications communicated by the variations in intonation in otherwise identical utterances (Moore & Davidge, 1989; Moore, Harris, & Patriquin, 1993). Use of visual images (pictures of items in question or drawings depicting the situation) will enhance the process effectively.
This point will be further discussed in Section 5.4, in which I will describe the design and production of a computer module for teaching intonation.

5.1.4 **Intonation should be an integral part of language classes at all levels of instruction.**

Instruction on intonation within meaningful communicative and contextualized activities should be an integral part of language practice at all levels of instruction. This claim is most strongly justified by the fact that not all the functions of intonation can be introduced effectively to beginning level learners. For instance, contrasting intonation patterns cueing for varying underlying syntactic structures of utterances should be discussed when learners are already familiar with, or are being introduced to the syntactic structures with which such intonation patterns are used.

The use of intonation patterns in one-word utterances offers a good example. While citation forms, such as *iku* ‘go’ *kau* ‘buy’ are regularly used for listing verbs, it is important for the learners of Japanese to be aware that verbs in citation form may occur by itself as a complete sentence (Jorden & Noda, 1987, p. 14; Moriyama, 1997). Thus *Iku* (with a rising intonation) … *Iku.* (with a falling intonation) is a conversation which may be translated as a conversation between two persons: ‘Are you going?’ … ‘I am going.’ As Moriyama points out, because one-word utterances which do not have an underlying syntactic structures, its pragmatic implication can only be communicated and interpreted by intonation. As he correctly points out, intonation plays a crucial role in one-word utterance communication. The concept of speech style in Japanese is complex and important. Because speech in the distal-style, i.e., the **desu-masu** style, is most
generally acceptable for foreign adults, especially those who are not yet very proficient in the Japanese language, most textbooks that are widely used in the United States introduce the verbs in distal style before they introduce the so-called citation form (Jorden & Noda, 1987; Makino, Hatasa, & Hatasa, 1998; Tohsaku, 1999). The citation forms, as well as the casual speech style that utilize the citation forms for one-word sentences, are introduced much later. Thus the aspect of intonation becomes relevant and important when learners reach the intermediate level where learn citation forms of verbs, and therefore will be able to use the one-word utterance that use the citation forms.

This claim also warns against the tendency for intonation to be dealt with as a part of pronunciation at the beginning of the course, only to be neglected and forgotten altogether later, when learners could make better use of the information and practice targeted toward intonation. Instructional materials for teaching pronunciation have traditionally been designed with an implicit assumption that pronunciation must be taught as a separate item distinct from other aspects of the language, which consequently implies that any time spent on teaching pronunciation will mean time taken away from other activities.

Another assumption underlying the practice of providing pronunciation instruction only at the beginning is the idea that pronunciation can be perfected early on, and that there is no room for improvement later in the learning paths. While it is true that pronunciation needs to be learned early on in order to conduct any activities in spoken language, it is not necessarily the case that the early training is sufficient. Teaching and learning correct pronunciation using traditional methods would require a large amount of
time and mechanical practice, involving much repetition. This unfortunate misconception of “teaching pronunciation” has lead teachers to feel that teaching pronunciation in class is impractical, and therefore dispensable. Referring to the mechanical repetitions associated with teaching pronunciation, Otsubo (1989) asserts that it is not worth spending classroom time for such activities. Pointing out the importance of prioritizing the items to be dealt with in the language classroom, Aizawa (1991) observes, “it is hard to believe that the aspect [of pronunciation] deserves such priority” (125).

However, reaction against the idea of devoting valuable class time solely to pronunciation should not go so far as to exclude the topic of pronunciation from practical consideration entirely because pronunciation can be integrated into other meaningful communicative activities. In fact, work on pronunciation should take place as part of communicative activities. Teaching intonation is not an exception.

The treatment of intonation as an integral aspect of learning at all levels is related to the basic goal of learning a language, identified in Chapter One. The goal of language learning is not to “study” the language as a scholarly subject matter, but to become a functioning individual in the community in which the target language is used. Given that assumption, the importance of pragmatic competence becomes clearer. Moreover, pragmatic competence does not necessarily entail competent performance. One may have an intellectual understanding of how the intonation variations reflect different intentions, but making sense of another person’s utterance or expressing one’s intentions are matters of performance (Walker, 2000; Walker & Noda, 2000).
When and how should intonation be introduced and dealt with as a part of the language curriculum? I address these questions in Sections 5.3 and 5.4.

**5.1.5 Instructors of Japanese need to cultivate basic understanding of the syntax and pragmatics of intonation in Standard Japanese.**

Instructors of Japanese need to be trained in order for them to fully understand the pragmatic aspects of standard Japanese, especially if they are not native speakers of standard Japanese themselves. Examination of the processing of syntactically and pragmatically contrasting utterances by native speakers of Japanese, as discussed in Chapter Four, has provided data that suggest that the native speakers of dialect backgrounds other than Tokyo did not have Tokyo-speaker like patterns on the “comprehension” task, suggesting that native speakers of other dialects of Japanese may not have been completely assimilated to the pragmatics of standard Japanese. This implies that some Japanese language teachers need training in the pragmatics of the intonation of the standard dialect.

The same view is expressed by Komori (1987) in his textbook on Japanese accent that is intended for native speakers of Japanese. He points out that not all native speakers of Japanese have native intuition in the accent patterns of standard Japanese and explains that native speakers of Japanese will still need to receive proper training in Tokyo accentual distinctions if they are native speakers of different dialects, specially speakers of those *mu-akusento* ‘un-accented’ dialects, if they want to obtain jobs that require the ability to speak the standard dialects (e.g., TV announcers, actors, actresses) or learn to speak standard Japanese. His textbook includes a wide variety of exercises, and it is
noteworthy that he provides such a rigorous training program for those who are already native speakers of Japanese. This point would clearly apply to the instruction to develop pragmatic competence in dealing with intonational variations.

Second, instructors need to learn that variations in the prosodic organization of utterances can provide rich cues to help disambiguate syntactic as well as pragmatic ambiguities. As was discussed in Chapter 3, examination of intonation should not be equated with the impressionistic analysis of feelings of the speaker. Rather, the prosodic aspect of Japanese utterances is governed by its own structural organization, and there is a systematic correlation between the prosodic structure of an utterance and the variety of semantic and syntactic cues that are communicated by such variations. The knowledge of the prosodic organization of well-formed utterances in Japanese will allow Japanese instructors to predict, identify, and correct learners’ mistakes more systematically. Informed instructors will, in this way, be able to offer guidance to modify the speech patterns of learners more effectively.

In addition to the rules of intonation, instructors should also receive training to increase their effectiveness in providing effective feedback on the learners’ errors. Because intonation is a part of the spoken language at every level of proficiency, it is the aspect of learners’ language use to which teachers become accustomed most quickly, making errors more difficult for instructors of Japanese to detect. It has been observed that novice teachers participating in a training program at The Ohio State University tend to have difficulty detecting pronunciation errors, especially those involving prosodic
features. It is also the case that even experienced teachers tend to miss these errors if they have not been trained to detect them and have not made a conscious effort to do so. It may also be the case that intonation is seldom brought to the attention of instructors because, whereas there is a virtually never ending supply of structural and vocabulary elements which need to be taught, there are only a relatively small number of intonational variations, which lead to the allocation of a proportionally small amount of attention to the teaching of intonation. Based on the examination of the interaction between declarative and procedural memory elements in the acquisition of skills, Eda (1996) pointed out the role of instructional feedback in pronunciation as a factor influencing the rate of acquisition as well as the probability of successful learning. Assuming that her claim is applicable to the acquisition of intonation, the training of Japanese instructors should improve the likelihood of their providing effective feedback to the learners.

Lastly, instructors need to be trained in order to increase their awareness and sensitivity to issues around the teaching and learning of intonation. Because the pragmatic roles played by intonation have not been discussed sufficiently, there has been little information available for teachers until recently, and the teachers have remained unaware of the importance of prosody in communication. Training for the instructors of Japanese will contribute significantly to making up for the lack of resources available in the field, and to facilitating the process of increasing the body of knowledge in the field.

16 This observation is due to Professor Mari Noda, the director of the Summer Program East Asian Concentration (SPEAC) Program at the Ohio State University.
collectively to address and seek better ways of incorporating instruction on intonation
into curriculum.

A number of groundbreaking studies on Japanese prosody have been produced
within the last fifteen years, allowing us to start to examine this aspect of Japanese in a
more systematic manner. Unfortunately, however, discussion and examination of
teaching the prosodic aspect of spoken Japanese has thus far been limited to the
examination of lexical accent patterns, Japanese moraic rhythm, often with focus on the
acquisition of timing and durational contrast in special moraic syllables, such as geminate
consonants, moraic nasal, and long/short vowels, and has not extended much beyond the
Although intonation has been recognized as an aspect that deserves much attention by
many researchers (Hammerly, 1991; Jorden, 2000; Unger, Lorish, Noda, & Wada, 1993),
such discussions rarely go far enough to make concrete suggestions on how to deal with
intonation in a classroom, or to introduce it in the language curriculum in a planned and
orderly manner.

5.2 Additional Factors for Consideration on Memory

In addition to the suggestions based on empirical data, the formulation of
instructional curriculum requires consideration for the workings of human brain, that
performs, stores, and uses the various information related to intonation. In this section, I
visit the issue of memory in the context of Japanese as a foreign language. Specifically, I
refer to two frameworks of memory: compiling knowledge as performed stories (Walker
& Noda, 2000) and development of automaticity and attention (Nara, 2001), Then,
keeping these frameworks in mind, I summarize the pedagogical implications of the findings from the experiments and propose in sections 5.3 through 5.5 some preferred strategies for pedagogy, organization of the Japanese language curriculum, and classroom activities.

5.2.1 Language as a performed culture

Walker and Noda (2000) examine the role of memory in language acquisition from the perspective of “language as a performed culture.” A successful language learning process environment can be characterized as a place where learners create “a usable cultural memory that prepares them for future performances in the target culture” (193).

Treating language as performed culture has significant implications for pedagogical considerations in relation to the instructional methodologies and materials for Japanese intonation; it makes a strong appeal for the need to present the nuances of intonation in linguistically and pragmatically rich contexts. The notion of performed culture rules out learning of language as if they were a list of items to be placed into learners’ memory space, one followed by another. Presentation of what to be learned does not provide a sufficient condition for the creation of memory elements which will enable a learner of Japanese to acquire the ability to perform tasks in culturally appropriate context.

The notion of language as a performed culture can be applied to the acquisition of intonation. Suppose a learner has received instruction on the contrasting pragmatic implications associated with incredulity rise intonation. Suppose also that she has
practiced the intonation pattern, and now is able to produce the intonation pattern, fully aware of the pragmatic implications of such intonation pattern. However, the intellectual understanding of how to use the newly learned item is not enough to make her utilize the skill in a truly felicitous context. For instance, if the context were such that the learner has just been urged to partake of some food by his/her supervisor, it would be inappropriate for him/her to openly express doubt and ask if it is edible. A learner who is fully aware of the implication of each of the BPMs, and has developed the skills to produce the intended contour, would thus be given the opportunity to fully utilize their ability to make it a part of their communicative skills and strategies.

Thus, the examination of the acquisition of intonation in the framework of performed culture allows us to realize that true competence in intonation needs to be developed in context. Learners must be given the opportunity to build their stories, or personal memory of performances.

This idea is in keeping with Nara’s (2001) call for activities to enhance restructuring.

5.2.2 Automaticity and restructuring of processes

Nara (2001) provides a thorough review of literature on the issue of automaticity, and discusses the pedagogical implications for teaching Japanese as a foreign language. Automaticity, the degree to which the performance of a task is automatically executed, is a concept that is not limited to language learning. With repeated practice, one can execute intended tasks effortlessly, with speed and accuracy. Some examples of such tasks may be driving stick-shift cars, playing musical instruments, playing computer games, etc.
The degree of mastery of complex cognitive skills, i.e., whether there is a level of automaticity in the execution of certain skills, however, is not the only factor that determines the efficiency with which cognitive processes can be executed. Cheng (1985) argued that the difference in performance depends on whether or not, or how well, coordination and integration of skills is accomplished. She termed this process restructuring. Cheng gives the example of a piano player who must play a three-beats-per-count rhythm with one hand and a four-beats-per-count rhythm with the other. She points out that even after the player has practiced and mastered to the degree of automaticity the two tasks of playing the respective rhythms with each of the hands separately, coordinating the hands to play both rhythms simultaneously will be a difficult task. She thus claims that the acquisition of complex cognitive tasks involves not just the mastery of two sub skills and combining them, but rather a ‘new organization and integration of skills, which elevates the required efficiency to a new level of mastery (Cheng, 1985, 422; qtd. in Nara 2001, 248).

Describing this process in more detail, Nara (2001) asserts:

“… Through practice that builds and strengthens associative memory paths, learners reach a certain level of proficiency with a skill, and when they attempt to integrate it with another skill, they have to restructure and reorganize the two skills. When they have reorganized them successfully, they achieve a new level of performance efficiency previously unattained. Through more practice and restructuring, the cycle is repeated as the learners move closer to mastery” (248).

This view, as summarized by Nara, seems to accurately characterize some of my own observations in the classroom in fostering learners’ acquisition of accent and intonation in Japanese. For example, when I was pointing out a learner’s error in accent, I experienced the following: When asked to identify a building in response to a question
ano tatemono wa nan desyoo ka (what would that building be?), a learner replied ryo’kan* desu (it is a Japanese style inn), with an unnecessary lexical accent on the first mora of a word ryokan (a Japanese style inn), pronouncing it as HLL (H for ‘high,’ L for ‘low’ pitch on each mora). This word does not bear a lexical accent, and thus should be pronounced with a pattern of LHH. I stopped the interaction and corrected his pronunciation of the word ryokan. After some repetition of the target word, I had him say the whole sentence ryokan desu by having him repeat after the model I provided orally. He was able to repeat the sentence with accuracy. Satisfied, I asked him the original question, ano tatemono wa nan desyoo ka (what would that building be?), to which he excitedly responded ryo’kan* desu. Nara’s point on cognitive restructuring explains this very well. Although the student could produce the prosodic patterns by repeating the model, the moment his attention shifted to grammatical accuracy and the production of meaning, it seemed as if he had completely forgotten about the word accent. He apparently needed to “restructure” his cognitive resources so that he would be able to pay attention to the prosody while attending to the message he wanted to convey in the utterance.18

Thus, instructional methodologies and materials should also aim at the acquisition of sub-tasks and the integration of such sub-tasks as well.

17 The asterisk here indicates that this response has an error.

18 This point was first brought to my attention by Professor Noda, who has had experiences such as this.
5.2.3 Declarative vs. Procedural Knowledge Elements

Focusing on the acquisition of a new phonemic category in Japanese by a native speaker of English, Eda (1996) examined the roles of declarative memory elements (DMEs) within one prominent theory of cognitive skill acquisition, ACT-R, as the framework of analysis. The study investigated the effectiveness of providing explicit instruction on pronunciation declaratively, and concluded the following: 1) the availability of declarative memory elements will influence the rate of acquisition, 2) the availability of declarative memory elements will influence the probability of successful learning, 3) such an effect will be most relevant at the beginning stages of the acquisition process, 4) effective instructional feedback needs to be provided. Although the discussion was focused on the production aspect of language acquisition at the level of phonemes in Eda’s study, the framework of analysis based on the examination of the mechanism that underlies and produces human cognitive behavior should apply to the acquisition of intonation as well.

Declarative knowledge instruction was defined as “providing instruction that will become directly accessible to the learners as conscious recollection and can be brought to mind as a proposition or an image” (Eda, 1996, p. 20). Eda also states that declarative instruction is not limited to providing explanations. Use of hand gestures, diagrams, and charts representing the pitch patterns of the utterances can all be utilized as a part of declarative instruction. (See also Noda, 1998) For the instruction on intonation, for example, the instructor may use hand gestures to indicate the contrasting height of the utterance final pitch peaks. One way to use hand motions for pitch movement may be to

123
hold the left hand in front of the body with the palm facing up, and place the right hand over it with the palm facing down. The instructor may keep the left hand in the same place and indicate the difference of utterance final pitch peak by moving the right hand vertically up from low to mid, or from low to high position. Figures that are showing the pitch movement in utterances using the f0 tracking will serve as effective tools for the demonstration of intonation contrasts (Hirata, 1999; Tanaka & Kubozono, 1999)

5.3 Presentation and Sequencing of Instructional Materials

How can intonation be introduced as a topic in the classroom? Table 5.1 presents sample exercises showing how the effect of intonation patterns can be demonstrated and practiced, which can be easily utilized as a basis for developing instructional activities. In this series of exercises, soo desu ka ‘it’s as you said’ is used as a target sentence. If this is said with a normal falling tone at the end of the utterance, it indicates that the speaker has just learned something new, and that that piece of information has been accepted. In contrast, if it is said with a distinct rise in pitch at the end of the sentence, it is an indication of surprise and/or doubt or actual questioning; new information has been received, with interest and/or surprise and/or doubt on the part of the speaker (Jorden & Noda, 1987 p. 53)

These three sets of exercises involve the perception of intonation patterns, but they are intended to direct the learners’ attention to different levels of processing. In each set, item ‘a’ is intended to be the one to correspond to the utterance of Speaker B1, and item ‘b’ to that of Speaker B2, respectively. In Exercise 3 below, the learners are asked to
“react” to the input by selecting the behavior that would demonstrate their understanding of the differing implications of the two contrasting utterances.
Example Practice:

Speech Samples [Context: at a bookstore, two speakers, A and B, are looking at books.]

Speaker A: Kono hon ga omosiroi desu yo.

Speaker B1: Soo desu ka ∨
Speaker B2: Soo desu ka ♀

Exercise 1: Pay attention to the difference in intonation, and choose the item that best describes the type of intonation used by B1 and B2, respectively.

a. The sentence was said with a falling intonation.

b. The sentence was said with a rising intonation.

Exercise 2: In what situation would the two instances of 'soo desu ka' be produced? Choose the item that best describes the intention of Speakers B1 and B2, respectively.

a. The speaker has just been told something, and is expressing the state of mind that he has accepted the information.

b. The speaker has been told something, but is expressing his doubt as to whether that is actually the case.

Exercise 3: What do you think Speaker B will do following the two instances of 'soo desu ka'? Choose the item that best describes the action that Speaker B take following each utterance.

a. Accept the suggestion as is, and consequently may choose to purchase that book.

b. Being doubtful as to whether the book is really interesting, may not choose to purchase the book.

Table 5.1: Activities for practicing intonation contrast using the sentence soo desu ka as an example.
This series of exercises can be used to improve the learners’ knowledge about intonation. Learners should also have opportunities to produce these intended intonational contrasts in a controlled context before they are invited to engage in activities that require them to utilize these contrasts more spontaneously.

In designing an activity for teaching / learning intonation, I suggest five ordered steps: 1) awareness building, 2) development of lower level processing skills, 3) development of pragmatics processing skills, 4) development of pragmatic skills in communicative context, and 5) development of monitoring skills. In addition, as mentioned earlier (Section 5.1.2), it is important that both perception and production exercises are provided. Learners should be given opportunities to work on improving both their perceptive and productive skills, and whenever possible, the acquisition of perceptual skills should precede that of production skills. Ayusawa (2003) points out that whether learners improved perception skills will help them produce the patterns they are able to hear is a question for further investigation. While the learners’ improved production skills may not necessarily facilitate their acquisition of productive skills, their perceptive skills will allow them to accurately understand the intentions of the utterances they hear when they are engaged in exercises to improve production skills. Their improved perceptive skills will also help them monitor their own productive skills. This point will become crucially important when learners use computerized instructional material to practice intonation, which is likely to take place in the absence of a human instructor to monitor their utterances and provide necessary feedback.
The proposed sequence is abstract enough to allow for use with different contents. As activities are designed in spiral fashion, the contents differ depending on the level of instruction. These activities should be presented in spiral fashion throughout all levels of instruction by expanding upon the materials previously learned, which are therefore familiar to the learners. “We learn to hear by hearing, to speak by speaking, to interact by interacting,” observe Unger et al. (1993, p. 15). In the case of acquisition of intonation, if we want the learners of Japanese to learn how to use the Japanese intonation in context, they will do so by having ample opportunity to practice doing so. While there may be only a limited number of contrasting patterns in intonation for learners to acquire, ample opportunities for them to perceive, interpret, produce, and use these patterns in felicitous context will be required before they attain a significant level of mastery at phonetic and pragmatic levels of processing.

5.4 Development of Computer Aided Language Learning Material for Intonation

Currently, there is no commercially available instructional material that is designed to provide exercises on intonation. Hardly any materials are available that deal with pronunciation at a discourse level, and this is an area that deserves increased attention. In designing such instructional material, we need to prioritize and decide what can be included in the course, because the availability of reliable instructional material will allow intonation to be practiced outside class. Materials that support work outside class are important in Japanese especially when language learning is viewed as a life-long process that goes well beyond the limits of time available in a given curriculum. Availability of instructional material outside class will also allow the human instructors
of Japanese to focus on providing opportunities for learners to experience the aspect of intonation as they engage themselves in negotiating their intentions with other speakers of Japanese. In other words, language instructors will be able to focus more on assisting the learners to “reorganize their memory” (Nara, 2001) and to compile cultural memories by adding to their memory another set of possible behaviors in a given context (Walker & Noda, 2000)

Availability of computerized materials will allow human instructors to focus more on using intonation for communicative purposes. Ease of access will be another issue to take into consideration. Instructional materials that are made easily available to students will encourage and enable them to work outside the classroom. There is a need for consistency, as well, in providing constant feedback and modeling to the students.

Recognizing the lack of materials for teaching intonation, I am designing and developing a computerized module for teaching Japanese accent and intonation. The module is being developed for inclusion as part of multimedia collection to supplement Japanese: The Spoken Language, Parts 2 & 3. This set of multimedia materials is designed for intermediate and advanced level learners of Japanese; this particular module is intended to be used as review material. While I was responsible for the selection of the content of material, the design of the screen, definition and shapes of buttons, color scheme, and overall architecture of the interaction has been developed in collaboration with the project members.¹⁹

¹⁹ Mari Noda, Abhijit Varde, and Jeff Borisch.
This module will include activities to improve the learners’ perception and production of accent (Accent Module) and intonation patterns (Intonation Module). Note that, in contrast to the definition of “intonation” employed in this dissertation, I will be

---

20 Draft at the time of this writing in 2004.
using the term “intonation” in this module to refer only to the effect of the utterance final pitch configuration (BPMs). This decision was motivated by recognition of the need to present the behavior of accent and intonation in a simple and consistent manner. I gave lectures to learners of Japanese enrolled in beginning, intermediate, and advanced level courses using the material currently being implemented in the Accent and Intonation Module. One of the challenges that I encountered was the definition and use of terminology, i.e., accent and intonation. The responses of the students during and after the lectures indicated that they would readily associate the term ‘intonation’ with the utterance final pitch movement, hence the decision to use a term that is consistent with their previous experience with the terminology. For this reason, the issue of the overall contour of utterance and the effect of pitch range manipulation associated with intonation phrase boundaries (phrasing) will be introduced as “special accent activity.”

The Accent Module will present facts about the functions and characteristics of accent (AI-1), provide exercises for learners to practice producing lexical items with accurate accent patterns (AI-2, AI-3), and provide exercises for them to use the pitch range to control the phrasing within an utterance to perceive and produce prosodic cues for syntactic structure and pragmatic focus (AI-4, AI-5). Similarly, the Intonation Module will make a short presentation about the functions and characteristics of intonation (AI-6), and provide exercises for learners to practice perceiving and producing contrasting BPM patterns (AI 7-10).
Figures 5.2 and 5.3 are screen captures of Intonation Activity in which contrasting BPMs are presented and practiced using utterances that differ minimally in the aspect of intonation.

![Intonation Activity Screen Capture](image)

Figure 5.2: A screen capture of the Intonation Module for a perception exercise\(^{21}\)

\(^{21}\) Draft at the time of this writing in 2004.
In the perception exercise, one pair of utterances that differ minimally in the aspect of BPM will be presented as an intonational minimal pair. The use of the minimal pair is intended to highlight and thus bring to the attention of learners that the implication of an utterance may be altered by the different BPM with which it is said. A pair of pictures will be used to illustrate the difference in implication. The users will be able to hear each utterance as many times as they want by pressing the button indicated as “model” in the bar immediately above each picture. The users will be able to go to the production exercise by pressing a button indicated as “Practice Speaking” located in the bottom center of the presentation panel.
In the production practice screen, they will be able to hear the model, record their utterance, and then replay their utterance followed by a model utterance. Assuming that they have developed the ability to perceive the differences in BPMs, they will be able to compare their own utterance with the model to check the accuracy of their production.

In the previous section, I proposed the design of activities for accent and intonation using five ordered steps: 1) awareness building, 2) development of lower level processing skills, 3) development of pragmatics processing skills, 4) development of
pragmatic skills in communicative contexts, and 5) development of monitoring skills.

The Accent and Intonation Module presented in this section is not intended to cover all
development of monitoring skills.

The Accent and Intonation Module presented in this section is not intended to cover all
the five steps identified earlier. Rather, this module should be considered as a tool to
allow learners of Japanese to go through the first three steps on their own outside the
classroom.

The final two steps can then be addressed by the language instructors effectively
in the classroom. For example, if a learner has practiced the contrast between information
question rise and incredulity rise using the module, a classroom activity might include a
situation that necessitates the learners to choose to ask a question or express their
disbelief / doubt about the situation. In addition, the instructor should plan the activity to
create a context where use of such BPM is or is not an option.

5.5 Conclusion

This dissertation has investigated the process of acquisition of intonation by
native and non-native speakers of standard Japanese with focus on a) aspects of
intonation that correlate strongly with pragmatic focus and the underlying syntactic
structure of an utterance and b) interaction of pitch movements with the lexical and
grammatical elements of utterances at the utterance final positions. The final chapter
examined ways to apply empirical findings to the pedagogy of Japanese as a foreign
language. Many of the suggestions presented in this thesis are not final answers to the
question considered. Rather, they are to be recognized as carefully formulated hypotheses,
which need to be further investigated. In the following, I will identify thee areas that
require further research and suggest directions for further study.
As an extension of this dissertation research, I hope to complete the Accent and Intonation module. An important next step, thus, will be to empirically test the effectiveness of the use of computerized module for accent and intonation.

It may be necessary for the experimental study to be augmented by a series of classroom observations for the following purposes. First, such study may provide data to characterize the acquisition of accent and intonation qualitatively. Second, it would be interesting to obtain further information about the effect of feedback in a classroom, such as 1) what kind of feedback is actually made available; 2) what effects do different types of feedback have in a classroom setting; and 3) what are the really “instructional” sorts of feedback. Lastly, data obtained through actual classroom observation will also provide insight on how best to improve the behavior of language instructors in the classroom.

As Ayusawa (2003) states, the acquisition of accent and intonation by learners of Japanese as a foreign language is an area that will develop significantly (p. 47). One aspect of this area of study is the interaction of BPMs with grammatical and lexical elements of an utterance. Currently there is no coherent theory that explains the various functions and behaviors of SFPs, for instance. More careful examination will be necessary to establish a more comprehensive account of SFPs based on their interactions with BPMs.
APPENDIX A

Sound stimuli used for Experiment 1: Note that the stimuli were presented aurally, and were never presented in written format to the subjects. Categories and numbers assigned to each pair of utterances are given for reference. At the time of experiment, the stimuli were presented in a random order.

Category CC: InfoQ (√) vs. IncredQ (√!) - Can’t Believe it!!

1. 大丈夫ですか √ (b)
   大丈夫ですか √ ! (c)
   a. 話者は、自分は大丈夫だと言い張っている。
   b. 話者は、聞き手が大丈夫かどうか知りたい。
   c. 話者は、聞き手が本当に大丈夫だとは信じられないでいる。

   a. The speaker is insisting that she is feeling okay.
   b. The speaker is asking a question to find out if you are feeling okay.
   c. The speaker is incredulous: she does not think you are feeling okay.

2. 佐藤君が、結婚するの √ (a)
   佐藤君が、結婚するの √ ! (c)
   a. 話者は、自分は大丈夫だと言い張っている。
   b. 話者は、聞き手が大丈夫かどうか知りたい。
   c. 話者は、聞き手が本当に大丈夫だとは信じられないでいる。

   a. The speaker is insisting that she is feeling okay.
   b. The speaker is asking a question to find out if you are feeling okay.
   c. The speaker is incredulous
3. そんなに安かったの √ (b) そんなに安かったの √! (a)
   a. 話者は、聞き手が買った物が、そんなに安かったとは信じられないでいる。
   b. 話者は、聞き手が買った物が、そんなに安かったのかどうか知りたい。
   c. 話者は、聞き手が買った物が安かっただろうという自分の思いを確認している。

   a. The speaker is incredulous: she cannot believe the price was that low.
   b. The speaker asks a question to find out if the price was that low.
   c. The speaker wants to confirm her assumption that the price was that low.

4. 本当ですか √ (b) 本当ですか √! (a)
   a. 話者は、自分が今聞いたことが本当だとは信じられないでいる。
   b. 話者は、自分が今聞いたことが本当かどうか知りたい。
   c. 話者は、自分が今聞いたことが本当だろうという自分の思いを確認している。

   a. The speaker is incredulous: she cannot believe what you have just said.
   b. The speaker asks a question to find out if what you have just said is true.
   c. The speaker wants to confirm her assumption that what you have just said is true.

5. 田中さんが作ったんですか √ 田中さんが作ったんですか √!
   a. 話者は、田中さんがこのケーキを作ったのだと言い張っている。
   b. 話者は、田中さんがこのケーキを作ったのかどうか知りたい。
   c. 話者は、田中さんがこのケーキを作ったとは信じられないでいる。

   a. The speaker is insisting that Tanaka made the cake.
   b. The speaker is asking a question to find out from you if Tanaka made the cake.
   c. The speaker cannot believe Tanaka made the cake.

6. 子供も食べましたか √ (a) 子供も食べましたか √! (c)
a. 話者は、子供も食べたかどうか知りたい。
b. 話者は、子供も食べた、と聞き手に知らせている。
c. 話者は、子供も食べたとは信じられないでいる。

a. The speaker is incredulous: she cannot believe the price was that low.
b. The speaker asks a question to find out if the price was that low.
c. The speaker wants to confirm her assumption that the price was that low.

Category C: InfoQ (√) vs. IncredQ (✓) - Invitation vs. Negative Question

1. 明日のコンサート、行きませんか  √  (b)
   明日のコンサート、行きませんか  ✓ ！ (a)

   a. 話者は、聞き手がコンサートに行かないと知って、信じられない気持ちでいる。
   b. 話者は、コンサートに行かないか、と聞き手を誘っている。
   c. 話者は、コンサートに行かないか、と聞き手を誘っているが、実は、そのコンサートにはいってほしくないと思っている。

   a. The speaker is incredulous: she cannot believe that you are not going to the concert.
   b. The speaker is inviting you to the concert.
   c. The speaker thinks you are going to the concert, but she does not want you to.

2. このケーキ、食べませんか  ✓  (c)
   このケーキ、食べませんか  ✓！ (a)

   a. 話者は、聞き手がケーキを食べない様子なので、信じられない気持ちでいる。
   b. 話者は、聞き手がケーキを食べるだろうと思っているが、実は、そのケーキを食べてほしくないと思っている。
   c. 話者は、ケーキを食べないか、と聞き手を誘っている。

   a. The speaker is incredulous: she wonders why you are not eating the cake.
   b. The speaker is inviting you to eat the cake, but she does not want you to.
   c. The invites you are going to eat the cake.

3. 買いませんか  ✓  (a)
   買いませんか  ✓！ (b)

   a. 話者は、コンピューターを買わないか、と聞き手を誘っている。

139
b. 話者は、聞き手がコンピューターを買わない知って、信じられない気持ちでいる。
c. 話者は、聞き手がコンピューターを買うのか買わないのか知りたいがっている。

a. The speaker is inviting you to buy the computer.
b. The speaker is incredulous: she cannot believe that you are not going to buy the computer.
c. The speaker is asking whether you are buying the computer.

4. この本、読まない √ (b)
   この本、読まない √！ (c)

   a. 話者は、聞き手がこの本を読むだろうと思っているが、実は、その本を読んでもらいたくないと思っている。
b. 話者は、この本を読まないか、と聞き手を誘っている。
c. 話者は、聞き手がこの本を読まないと知って、信じられない気持ちでいる。

   a. The speaker thinks that you are going to read this book, but she does not want you to.
b. The speaker invites you to read this book.
c. The speaker is incredulous: she cannot understand why you are not reading this book.

5. このゲーム、しない √ (a)
   このゲーム、しない √！ (c)

   a. 話者は、このゲームをしないか、と聞き手を誘っている。
b. 話者は、このゲームをしないように聞き手を説得している。
c. 話者は、聞き手がこの本を読まないと知って、信じられない気持ちでいる。

   a. The speaker thinks that you are going to read this book but she does not want you to.
b. The speaker invites you to read this book.
c. The speaker is incredulous: she cannot understand why you are not reading this book.

6. 山田先生の授業、出ない √ (c)
   山田先生の授業、出ない √！ (b)
a. 話者は、聞き手山田先生の授業に出席しないよう、説得している。
b. 話者は、聞き手が山田先生の授業に出席しないと知って、信じられない気持ちである。
c. 話者は、山田先生の授業に一緒に出席しないか、と聞き手を誘っている。

a. The speaker persuading you to skip Prof. Yamada’s class.
b. The speaker is incredulous: she is surprised to hear that you are not attending Prof. Yamada’s class.
c. The speaker invites you to attend Prof. Yamada’s class with her.

Category D: Disambiguation of syntactically ambiguous sentences

1. 便利な|本と辞書を買いました (a)
便利な本と|辞書を買いました (b)

a. 話者は、便利な本と、便利な辞書を買った。
b. 話者は、便利な本と、辞書を買った。
c. 話者は、本と、便利な辞書を買った。

a. The speaker bought a book and a dictionary, both of which were convenient.
b. The speaker bought a convenient book, and a dictionary.
c. The speaker bought a book, and a convenient dictionary.

2. 駅の前の|本屋とレストランに行った (a)
駅の前の本屋と|レストランに行った (b)

a. 話者は、駅の前の本屋と、駅の前のレストランに行った。
b. 話者は、駅の前の本屋と、他のところにあるレストランに行った。
c. 話者は、他のところにあるレストランと、駅の前の本屋に行った。

a. The speaker went to a bookstore and a restaurant, both of which were in front of the station.
b. The speaker went to a bookstore in front of the station, and a restaurant.
c. The speaker went to a bookstore, and also to a restaurant in front of the station.

3. 佐藤さんと山下さんの|友達がきた (c)
佐藤さんと|山下さんの友達が来た (b)

a. 佐藤さんと、山下さんと、山下さんの友だちの三人が来た。
b. 佐藤さんと、山下さんの友だちの、二人が来た。
c. 佐藤さんと山下さんの友だちが、一人来た。

a. Mr. Sato, Mr. Yamashita, and Mr. Mamashita’s friend came. (Three people came.)
b. Mr. Sato and someone else who is Mr. Yamashita’s friend came. (Two people came.)
c. Someone who is a friend of both Mr. Sato and Mr. Yamashita came. (One person came.)

4. 京都で買ったラジオがなくなった (b)
京都で買ったラジオがなくなった (a)

a. 話者はが京都にいたときに買ったラジオが、なくなってしまった。
b. 話者が京都にいたときに、そこで自分で買ったラジオがなくなった。
c. 話者が何年か前に京都にいたときに買ったラジオが、なくなってしまった。話者は、ラジオは盗まれたのではなく、なくしたのだということを強調している。

a. The speaker lost a radio, which she had bought in Kyoto.
b. In Kyoto, the speaker lost a radio, one that she had bought for herself and not been given as a present.
c. The speaker lost the radio which she bought in Kyoto, and she emphasizes that it was lost, not stolen.

5. 孝志君と幸子ちゃんの先生に会いました (b)
孝志君と幸子ちゃんの先生に会いました (c)

a. 話者は、孝志君と幸子ちゃんと、孝志君と幸子ちゃんの先生の三人に会った。
b. 話者は、孝志君と、幸子ちゃんの先生の二人に会った。
c. 話者は、孝志君と幸子ちゃんを教えている先生に会った。

a. The speaker met Takashi, Sachiko, and their teacher. (She met three people.)
b. The speaker met Takashi, and Sachiko’s teacher. (She met two people.)
c. The speaker met the teacher who teaches both Takashi and Sachiko. (She met one person.)

6. 金曜日に作ったケーキを子供が食べた (a)
金曜日に作ったケーキを子供が食べた (b)

a. 話者が金曜日に作っておいたケーキを、子供たちが食べてしまった。
b. The children ate the cake that the speaker had made on Friday.

2. The speaker does not think that Kyoto has tasty sweets.

3. The speaker expresses her disbelief that Mr. Kato made cake.
緑の傘は見ましたが（c）

a. 話者は、緑の傘を見たことを不思議に思っている。
b. 話者は、緑の傘を見たが、他の色の傘は見なかったらしい。
c. 話者は、緑の傘は見たが、傘の他に緑色の物は見なかったらしい。

a. The speaker has seen a green umbrella (and is wondering about it).
b. The speaker has seen a GREEN umbrella (and implies that she has not seen a different colored one).
c. The speaker has seen a green UMBRELLA (and implies that the umbrella was the only green thing he saw).

4. 佐藤さんのお兄さんに電話しましたけど（b）
佐藤さんのお兄さんに電話しましたけど（a）

a. 話者は、佐藤さんのお兄さんに電話してみたが、実際に彼と話をしたわけではない。
b. 話者が電話したのは佐藤さんのお兄さんで、お姉さんではなかったということを強調している。
c. 話者は、佐藤さんのお兄さんに電話してみたが、彼からの返事はなかった。

a. The speaker CALLED UP Mr. Sato’s elder brother (but she did not speak to him in person).
b. The speaker called Mr. Sato’s elder BROTHER (but not his elder sister).
c. The speaker called up Mr. Sato’s elder brother (but perhaps got no reply).

5. 今日はドイツのビールが安いですよ（b）
今日はドイツのビールが安いですよ（a）

a. 話者は、ドイツ産のものがたくさんある中で、今日安いのはビールだということを強調している。
b. 話者は、ビールがたくさんある中で、今日安いのはドイツ産だということを強調している。
c. 話者は、ドイツのビールが安いのは今日だということを強調している。

a. Today, German BEER is on sale (but not other items that are also from Germany).
b. Today, GERMAN beer is on sale (but not the beer from other countries).
c. TODAY, German beer is on sale (but they may not be cheap tomorrow).
6. 明日は山下さんに会いたいんですけど (a)
明日は山下さんに会いたいんですけど (c)

a. 話者は、今日は山下さんに会えなかったので、明日は、彼に会いたいと思っている。
b. 話者は、明日、他の人ではなく山下さん本人に会いたいと思っている。
c. 話者、明日、電話で話すのではなく、山下さんに直接会いたいと思っている。

a. The speaker wants to meet Mr. Yamashita TOMORROW (because she couldn’t meet him today).
b. The speaker wants to meet MR. YAMASHITA tomorrow (not someone else).
c. The speaker wants to MEET Mr. Yamashita tomorrow (instead of taking to him over the phone).

Category S: InfoQ Rise ✓ vs. Insisting Rise !

1. 孝志君が飲んだの ✓ (a)
孝志君が飲んだの ! (b)

a. 話者は、孝志が飲んだのかどうか知りたい。
b. 話者は、孝志が飲んだということが信じられないでいる。
c. 話者は、孝志が飲んだのだと言い張っている。

a. The speaker is asking a question to find out form you whether Takashi drank.
b. The speaker is incredulous: she cannot believe that Takashi drank.
c. The speaker is insisting that Takashi drank.

2. 明日、映画が見たい ✓ (b)
明日、映画が見たい ! (a)

a. 明日映画が見たい、と話者は言い張っている。
b. 明日映画が見たいかどうか、話者は聞き手に尋ねている。
c. 明日映画が見たい、と話者や聞き手に伝えている。

a. The speaker is insisting that she wants to see a movie tomorrow.
b. The speaker is asking a question to find out from you if you want to see a movie tomorrow.
c. The speaker is making a simple statement that she wants to see a movie tomorrow.
3. この本、おもしろくなかった √ (c)
   この本、おもしろくなかった ! (b)
   a. この本は、あまり面白くなかったと話者は伝えている。
   b. この本は面白かったと話者は言い張っている。
   c. この本が面白かったかどうか、話者は聞き手に尋ねている。

   a. The speaker is making a statement that she did not think this book was interesting.
   b. The speaker is insisting that this book was interesting.
   c. The speaker is asking a question to find out from you whether this book was interesting.

4. こちらにもございます ! (a)
   こちらにもございます √ (b)
   a. 話者は、探しているものがこの店にあるかどうか聞き手に尋ねている。
   b. 話者は、聞き手が探しているものがこの店にあると言い張っている。
   c. 話者は、探しているものがこの店にあるということが信じられないでいる。

   a. The speaker is asking a question to find out if this store has [the item under discussion].
   b. The speaker is insisting that this store has [the item under discussion].
   c. The speaker is incredulous: she does not think that this store has [the item under discussion].

5. 病院には行きません √ (a)
   病院には行きません ! (c)
   a. 話者は、病院に行きたいかどうか聞き手に尋ねている。
   b. 病院には行きたくない、と話者は聞き手に伝えている。
   c. 病院には行きたくない、と話者は言い張っている。

   a. The speaker is asking a question to find out from you if you don’t want to go to the hospital.
   b. The speaker is informing you that she does not want to go to the hospital.
   c. The speaker is insisting that she does not want to go to the hospital.

6. これは直美ちゃんに √ (a)
   これは直美ちゃんに ! (c)
147

a. これは直美に渡すものなのか、話者は聞き手に尋ねている。
b. これは直美に渡すものだということが、話者には信じられないでいる。
c. これは直美に渡すものだと、話者は言い張っている。

a. The speaker is asking a question to find out from you if this is for Naomi.
b. The speaker is incredulous: she cannot believe that this is for Naomi.
c. The speaker is insisting that this is for Naomi.

Category SS: InfoQ Rise √ vs. Insisting Rise + ne！

d. 加藤さんは|ケーキを作りましたよ (b)
加藤さんはケーキを|作りましたよ (c)

e. 話者は、加藤さんが作ったのはケーキだったということを強調してい
る。
f. 話者は、ケーキは(買わないで)か父さんが自分で作ったということを
強調している。
g. 話者は、ケーキを作ったのは加藤さんだったということが信じられな
いでいる。

d. The speaker is emphasizing the fact that Mr. Kato made CAKE (because he
usually makes something else).
e. The speaker is emphasizing the fact that Mr. Kato MADE cake (because he
usually buys it).
f. The speaker expresses her disbelief that Mr. Kato made cake.

2. 京都は|お菓子がおいしいですよ (c)
京都はお菓子がおいしいですよ (b)

d. 話者は、京都にはあまりおいしいお菓子はないと思っている。
e. 話者は、京都でおいしいのはお菓子だけで、他のものはあまりおいし
くないと思っている。
f. 話者は、お菓子がおいしいのは京都だけだと思っている。

d. The speaker does not think that Kyoto has tasty sweets.
e. The speaker thinks that Kyoto has tasty SWEETS (implying that other kinds
of food is not so good in Kyoto).
f. The speaker thinks that KYOTO has tasty sweets (implying that it is not the
case with other cities).

3. 緑の傘は見ましたか （a）
緑の傘は見ましたが（b）

d. 話者は、緑の傘見たこととは思わず、不思議に思っている。
e. 話者は、緑の傘見たが、他の色の傘は見なかったらしい。
f. 話者は、緑の傘見たが、傘の他に緑色の物は見なかったらしい。

d. The speaker has seen a green umbrella (and is wondering about it).
e. The speaker has seen a GREEN umbrella (and implies that she has not seen a different colored one).
f. The speaker has seen a green UMBRELLA (and implies that the umbrella was the only green thing he saw).

4. 佐藤さんのお兄さんに電話しましたけど（b）
佐藤さんのお兄さんに電話しましたけど（c）

d. 話者は、佐藤さんのお兄さんに電話してみたが、実際に彼と話したわけではない。
e. 話者が電話したのは佐藤さんのお兄さんで、お姉さんではなかったということが強調している。
f. 話者は、佐藤さんのお兄さんに電話してみたが、彼からの返事はなかった。

d. The speaker CALLED UP Mr. Sato’s elder brother (but she did not speak to him in person).
e. The speaker called Mr. Sato’s elder BROTHER (but not his elder sister).
f. The speaker called up Mr. Sato’s elder brother (but perhaps got no reply).

5. 今日はドイツのビールが安いですよ（b）
今日はドイツのビールが安いですよ（c）

d. 話者は、ドイツ産のものがたくさんある中で、今日安いのはビールだということを強調している。
e. 話者は、ビールが多くある中で、今日安いのはドイツ産のだということを強調している。
f. 話者は、ドイツのビールが安いのは今日だということを強調している。

d. Today, German BEER is on sale (but not other items that are also from Germany).
e. Today, GERMAN beer is on sale (but not the beer from other countries).
f. TODAY, German beer is on sale (but they may not be cheap tomorrow).
6. 明日は山下さんに会いたいんですけど (c)
    明日は山下さんに会いたいんですけど (a)

d. 話者は、今日は山下さんに会えなかったので、明日は、彼に会いたいと思っている。
e. 話者は、明日、他の人ではなく山下さん本人に会いたいと思っている。
f. 話者は、明日、電話で話すのではなく、山下さんに直接会いたいと思っている。

d. The speaker wants to meet Mr. Yamashita TOMORROW (because she couldn’t meet him today).
e. The speaker wants to meet MR. YAMASHITA tomorrow (not someone else).
f. The speaker wants to MEET Mr. Yamashita tomorrow (instead of taking to him over the phone).
APPENDIX B

Instruction: After hearing each sentence twice, choose one statement that best describes what the speaker has said. The speaker (she) is talking to Akio, who is referred to as ‘he’ in the questions.

1. 孝志君と行くの！(D)²²
   A: The speaker is making a simple statement that she is going (to a party) with Takashi.
   B: The speaker is asking a simple question to find out if he is going (to a party) with Takashi.
   C: The speaker cannot believe that she has just heard, that he is going (to the party) with Takashi.
   D: the speaker is (emphatically) insisting that she is going (to a party) with Takashi.

2. これ、田中さんが作ったの√(B)
   A: The speaker is making a simple statement that Mr. Tanaka has made this.
   B: The speaker is asking a simple question to find out if Mr. Tanaka has made this.
   C: The speaker cannot believe that she has just heard, that Mr. Tanaka has made this.
   D: the speaker is (emphatically) insisting that Mr. Tanaka has made this.

3. アメリカに行きます．(A)
   A: The speaker is making a simple statement that she is planning to go to America.
   B: The speaker is asking a simple question to find out if he is planning to go to America.
   C: The speaker cannot believe that she has just heard, that he is planning to go to America.
   D: the speaker is (emphatically) insisting that she is planning to go to America.

4. 全部できました！(D)

²² Note these utterances were presented aurally, and never in a written format, to the subjects at the time of the experiment. The alphabet in the parentheses corresponds to the item that is intended as the best paraphrasing sentence in the list provided below. Note also that these items were presented in a random order at the time of experiment.
A: The speaker is making a simple statement that she finished (her work).
B: The speaker is asking a simple question to find out if he has finished (his work).
C: The speaker cannot believe that she has just heard, that he has finished (his work).
D: the speaker is (emphatically) insisting that she has finished (her work).

5. コンサートは八時半から√ (B)

A: The speaker is making a simple statement that the concert starts at 8:30.
B: The speaker is asking a simple question to find out if the concert starts at 8:30.
C: The speaker cannot believe that she has just heard, that the concert starts at 8:30.
D: the speaker is (emphatically) insisting that the concert starts at 8:30.

6. あの先生なんですか√ (B)

A: The speaker is making a simple statement that that person is a teacher.
B: The speaker is asking a simple question to find out if that person is a teacher.
C: The speaker cannot believe that she has just heard, that that person is a teacher.
D: the speaker is (emphatically) insisting that that person is a teacher.

7. 本当に奈良のなの！(D)

A: The speaker is making a simple statement that this object is really from Nara.
B: The speaker is asking a simple question to find out if this object is really from Nara.
C: The speaker cannot believe that she has just heard, that this object is really from Nara.
D: the speaker is (emphatically) insisting that this object is really from Nara.

8. 鈴木さんが結婚するの?? (C)

A: The speaker is making a simple statement that Mr. Suzuki is getting married.
B: The speaker is asking a simple question to find out if Mr. Suzuki is getting married.
C: The speaker cannot believe that she has just heard, that Mr. Suzuki is getting married.
D: the speaker is (emphatically) insisting that Mr. Suzuki is getting married.

9. 明日、出来ます√ (B)

A: The speaker is making a simple statement that she can finish it tomorrow.
B: The speaker is asking a simple question to find out if he can finish it tomorrow.
C: The speaker cannot believe that she has just heard, that he can finish it tomorrow.
D: the speaker is (emphatically) insisting that she can finish it tomorrow.

10. こんな辞書は使いません。 (A)

A: The speaker is making a simple statement that she does not use this kind of dictionary.
B: The speaker is asking a simple question to find out if he does not use this kind of dictionary.
C: The speaker cannot believe that she has just heard, that he does not use this kind of dictionary.
D: the speaker is (emphatically) insisting that he does not use this kind of dictionary.

11. 自分で出来ます?? (C)
A: The speaker is making a simple statement that she can handle the task by herself.
B: The speaker is asking a simple question to find out if she can handle the task by herself.
C: The speaker cannot believe that she has just heard, that she can handle the task by herself.
D: the speaker is (emphatically) insisting that she can handle the task by herself.

12. 孝志君が飲んだの√ (B)
A: The speaker is making a simple statement that Takashi finished this drink.
B: The speaker is asking a simple question to find out if Takashi finished this drink.
C: The speaker cannot believe that she has just heard, that Takashi finished this drink.
D: the speaker is (emphatically) insisting that Takashi finished this drink.

13. デパートに行きたい！ (D)
A: The speaker is making a simple statement that she wants to go to a department store.
B: The speaker is asking a simple question to find out if she wants to go to a department store.
C: The speaker cannot believe that she has just heard, that she wants to go to a department store.
D: the speaker is (emphatically) insisting that she wants to go to a department store.

14. これ違います. (A)
A: The speaker is making a simple statement that this is a mistake.
B: The speaker is asking a simple question to find out if this is a mistake.
C: The speaker cannot believe that she has just heard, that this is a mistake.
D: the speaker is (emphatically) insisting that this is a mistake.

15. そんなに高かったの?? (C)
A: The speaker is making a simple statement that the price was that high.
B: The speaker is asking a simple question to find out if the price was that high.
C: The speaker cannot believe that she has just heard, that the price was that high.
D: the speaker is (emphatically) insisting that the price was that high.

16. すみません. (A)
A: The speaker is making a simple statement that she is sorry.
B: The speaker is asking a simple question to find out if he is sorry.
C: The speaker cannot believe that she has just heard, that he is sorry.
D: the speaker is (emphatically) insisting that she is sorry.

17. 子供も食べました?? (C)

A: The speaker is making a simple statement that children also ate it.
B: The speaker is asking a simple question to find out if children also ate it.
C: The speaker cannot believe that she has just heard, that children also ate it.
D: the speaker is (emphatically) insisting that children also ate it.

18. コンサートのチケットを渡しました. (A)

A: The speaker is making a simple statement that she has given the concert tickets (to them).
B: The speaker is asking a simple question to find out if he has given the concert tickets (to them).
C: The speaker cannot believe that she has just heard, that he has given the concert tickets (to them).
D: the speaker is (emphatically) insisting that she has given the concert tickets (to them).

19. 本当です?? (C)

A: The speaker is making a simple statement that it is true.
B: The speaker is asking a simple question to find out if it is true.
C: The speaker cannot believe that she has just heard, that it is true.
D: the speaker is (emphatically) insisting that it is true.

20. これは直美ちゃんに！ (D)

A: The speaker is making a simple statement that this gift is for Naomi.
B: The speaker is asking a simple question to find out if this gift is for Naomi.
C: The speaker cannot believe that she has just heard, that this gift is for Naomi.
D: the speaker is (emphatically) insisting that this gift is for Naomi.

21. 本はもう返した. (A)

A: The speaker is making a simple statement that she has returned the book.
B: The speaker is asking a simple question to find out if he has returned the book.
C: The speaker cannot believe that she has just heard, that he has returned the book.
D: the speaker is (emphatically) insisting that she has returned the book.

22. この車、田中さんの√ (B)

A: The speaker is making a simple statement that this car is Tanaka’s.
23. 明日は来ません。 (A)
A: The speaker is making a simple statement that she will not come tomorrow.
B: The speaker is asking a simple question to find out if she will not come tomorrow.
C: The speaker cannot believe that she has just heard, that she will not come tomorrow.
D: the speaker is (emphatically) insisting that she will not come tomorrow.

24. 明日も来ます。 (A)
A: The speaker is making a simple statement that (this person) will come tomorrow, too.
B: The speaker is asking a simple question to find out if (this person) will come tomorrow, too.
C: The speaker cannot believe that she has just heard, that (this person) will come tomorrow, too.
D: the speaker is (emphatically) insisting that (this person) will come tomorrow, too.

25. こんな辞書は使いません！ (D)
A: The speaker is making a simple statement that she does not use this kind of dictionary.
B: The speaker is asking a simple question to find out if she does not use this kind of dictionary.
C: The speaker cannot believe that she has just heard, that she does not use this kind of dictionary.
D: the speaker is (emphatically) insisting that she does not use this kind of dictionary.

26. 日本語も分かります？？ (C)
A: The speaker is making a simple statement that he understands Japanese also.
B: The speaker is asking a simple question to find out if he understands Japanese also.
C: The speaker cannot believe that she has just heard, that he understands Japanese also.
D: the speaker is (emphatically) insisting that he understands Japanese also.

27. ここにもあります。 (A)
A: The speaker is making a simple statement that the item is available here, too.
B: The speaker is asking a simple question to find out if the item is available here, too.
C: The speaker cannot believe that she has just heard, that the item is available here, too.
D: the speaker is (emphatically) insisting that the item is available here, too.

28. これ、違います！ (D)
A: The speaker is making a simple statement that this is a mistake.
B: The speaker is asking a simple question to find out if this is a mistake.
C: The speaker cannot believe that she has just heard, that this is a mistake.
D: the speaker is (emphatically) insisting that this is a mistake.

29. ここにもあります√ (B)

A: The speaker is making a simple statement that the item is available here.
B: The speaker is asking a simple question to find out if the item is available here.
C: The speaker cannot believe that she has just heard, that the item is available here.
D: the speaker is (emphatically) insisting that the item is available here.

30. コンサートは七時半から！ (D)

A: The speaker is making a simple statement that the concert starts at 7:30.
B: The speaker is asking a simple question to find out if the concert starts at 7:30.
C: The speaker cannot believe that she has just heard, that the concert starts at 7:30.
D: the speaker is (emphatically) insisting that the concert starts at 7:30.

31. 忘れたの. (A)

A: The speaker is making a simple statement that she has forgotten (the message).
B: The speaker is asking a simple question to find out if he has forgotten (the message).
C: The speaker cannot believe that she has just heard, that he has forgotten (the message).
D: the speaker is (emphatically) insisting that she has forgotten (the message).

32. これ佐藤さんの?? (C)

A: The speaker is making a simple statement that this is Sato’s.
B: The speaker is asking a simple question to find out if this is Sato’s.
C: The speaker cannot believe that she has just heard, that this is Sato’s.
D: the speaker is (emphatically) insisting that this is Sato’s.

33. これ食べていいい?? (C)

A: The speaker is making a simple statement that it’s okay to eat this (sushi).
B: The speaker is asking a simple question to find out if it’s okay to eat this (sushi).
C: The speaker cannot believe that she has just heard, that it’s okay to eat this (sushi).
D: the speaker is (emphatically) insisting that it’s okay to eat this (sushi).

34. これは聡子ちゃんに！ (D)

A: The speaker is making a simple statement that she wants to give this to Satoko.
B: The speaker is asking a simple question to find out if he wants to give this to Satoko.
C: The speaker cannot believe that she has just heard, that he wants to give this to Satoko.
D: the speaker is (emphatically) insisting that she wants to give this to Satoko.

35. 車はもうあります！ (D)

A: The speaker is making a simple statement that she already owns a car.
B: The speaker is asking a simple question to find out if she already owns a car.
C: The speaker cannot believe that she has just heard, that she already owns a car.
D: the speaker is (emphatically) insisting that she already owns a car.

36. これは佐藤さんに. (A)

A: The speaker is making a simple statement that she made this for Sato.
B: The speaker is asking a simple question to find out if he made this for Sato.
C: The speaker cannot believe that she has just heard, that he made this for Sato.
D: the speaker is (emphatically) insisting that she made this for Sato.

37. そんな電話はありませんでした √ (B)

A: The speaker is making a simple statement that there was no such phone call.
B: The speaker is asking a simple question to find out if there was no such phone call.
C: The speaker cannot believe that she has just heard, that there was no such phone call.
D: the speaker is (emphatically) insisting that there was no such phone call.

38. 「すみません」?? (C)

A: The speaker is making a simple statement that she is very sorry.
B: The speaker is asking a simple question to find out if he is very sorry.
C: The speaker cannot believe that she has just heard, that he is very sorry.
D: the speaker is (emphatically) insisting that she is very sorry.

39. それはもう出来ました √ (B)

A: The speaker is making a simple statement that she has finished (the work).
B: The speaker is asking a simple question to find out if he has finished (the work).
C: The speaker cannot believe that she has just heard, that he has finished (the work).
D: the speaker is (emphatically) insisting that she has finished (the work).

40. アメリカに行きます?? (C)

A: The speaker is making a simple statement that she will go to America.
B: The speaker is asking a simple question to find out if he will go to America.
C: The speaker cannot believe that she has just heard, that he will go to America.
D: the speaker is (emphatically) insisting that she will go to America.


