FACTORS PREDICTING
NATIVE AND NONNATIVE LISTENERS’
EVALUATIVE REACTIONS TO
JAPANESE ENGLISH

DISSERTATION

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

By
Reiko Kachi, M.A.

* * * * *

The Ohio State University
2004

Dissertation Committee:
Professor Keiko K. Samimy, Adviser
Professor Larry E. Miller
Professor Alan R. Hirvela

Approved by

Adviser
College of Education
Copyright by

Reiko Kachi

2004

All right reserved
ABSTRACT

The World Englishes and English as an International Language (EIL) points of view have challenged the traditional schemes of applied linguistics research. Since English belongs to all the worldwide users, nonnative speakers’ perspectives should be acknowledged with as much attention as those of native speakers. Consequently, the current study investigated native and nonnative listeners’ evaluative reactions to and the intelligibility of Japanese English, a nonnative English variety. Japanese English was chosen as the target of investigation, due to the importance of English-medium international communication for Japanese people and its exemption from localized norms.

Quantitative as well as qualitative methods were employed for triangulation and complementarity. In the ex post facto study, listeners’ native language backgrounds (i.e., American English, Hindi, Mandarin, and Japanese), as well as the speakers’ proficiency levels, were hypothesized as the predictors of the evaluative reactions to and the intelligibility of Japanese English. Also investigated as possible factors were the listeners’ gender, familiarity with Japanese English, and exposure to nonnative speakers of English other than Japanese English speakers. Eighty-four undergraduate and graduate students recruited at an American university listened to extemporaneous speech. The listeners evaluated it on a semantic differential scale and also completed a word-for-word dictation
task for the intelligibility. For further qualitative interviews, thirteen participants were recruited from the *ex post facto* study.

Regression analyses indicated that the Hindi speakers were the most critical of Japanese English while the American participants were the most positive about it, especially about rudimentary speakers. Reactions from Chinese and Japanese participants turned out to be similar. Japanese English speech was more intelligible than native speech, at least to Chinese and Japanese listeners. Interviews revealed that the Hindi speakers were irritated by the choppiness of Japanese speech, while the other nonnative participants were generally more sympathetic. The power imbalance among different types of English users was also articulated. Based on the results of the study, implications were presented for applied linguistics research and English education, especially where international communication was the goal of instruction. Additionally, recommendations for further studies were included.
Dedicated to Hiroyuki and Takuma Kachi,

who have always been there for me,

and to Yoko and Masu Ono,

who set long-lasting role models in my life
ACKNOWLEDGMENTS

I wish to express my foremost gratitude to Dr. Keiko Samimy, my adviser, who has always believed in my ability to complete this dissertation, and has also been my emotional as well as intellectual mentor for five years.

I wish to thank Dr. Alan Hirvela and Dr. Larry Miller, who served on my candidacy examination committee and continue to serve on my dissertation committee, for their incessant support, encouragement, and patience in correcting my methodological and stylistic inaccuracies.

I also thank Professor Susan Sarwark, Dr. Joe Gliem, Dr. Donald Winford, and Dr. Diane Belcher for generously sacrificing their time when I needed their advice and expertise.

I am grateful to the undergraduate and graduate students who participated in my study, especially those who agreed to be interviewed. I am also thankful for the cooperation of the students and staff in Nunoike Gaigo Senmon Gakko, in Nagoya, Japan. My gratitude goes to my friends and acquaintances who let me take their speech samples for my study.

I also wish to express my appreciation of the intellectually stimulating experiences that I had with my classmates and friends at the Ohio State University, Jocelyn Hardman, Ling-Miao Yeh, Darcy Lear, Mary Ball, Seung Jung Kim, Jung-Mi Nam, Suyoun Kim, Mi-Kyung Baek, Carmen Chacón, Junko Otoshi, Chiho Kobayashi,
and many others. I am especially thankful for Pornsiri Muangsamai, with whom I was fortunate to share good times and bad times in our pursuit of Ph.D. degrees. She also taught me that an acute scholar does not always have to be contentious.

I am indebted to my personal friends, Azusa Uno and Junko Ishibashi, for patiently listening to my grumbling and for never losing confidence in me.

My thanks to all that I have unintentionally omitted in my acknowledgment of appreciation.

Last but not least, I wish to express my sincere gratitude to my family, Hiroyuki and Takuma. The very existence of Takuma brightened up my days when I was discouraged and depressed. Hiroyuki always supported me emotionally and financially, to whom I will feel indebted forever.
June 29, 1964 ............................... Born – Tokyo, Japan

1987 .......................................... B.A. Education, Utsunomiya University, Japan

1987-1989 ............................... English teacher

Oyama-Nishi High School, Tochigi, Japan

1991-1992 ............................... English instructor, AEON, Mie, Japan

1993 ........................................ M.A. TESOL, Saint Michael’s College

1994-1997 ............................... English lecturer

Nunoike-Gaigo-Senmon-Gakkō, Nagoya, Japan

1998- present ............................. The Ohio State University, Columbus, Ohio

PUBLICATIONS

Research Publication

FIELDS OF STUDY

Major Field: Education

Area of Specialization: Foreign and Second Language Education

Minor Fields: Quantitative Research Methods

Japanese Linguistics
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Abstract</th>
<th>ii</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dedication</td>
<td>iv</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>v</td>
</tr>
<tr>
<td>Vita</td>
<td>vii</td>
</tr>
<tr>
<td>List of Tables</td>
<td>xiii</td>
</tr>
<tr>
<td>List of Figures</td>
<td>xv</td>
</tr>
</tbody>
</table>

**Chapters:**

1. Introduction ............................................................... 1
   - Statement of the Problem ........................................ 8
   - Significance of the Study ..................................... 10
   - Objectives of the study, hypotheses and a research question ............................................. 12
   - Definition of Terms ............................................. 15
   - Basic Assumptions ............................................. 18

2. Review of the literature .................................................. 19
   - World Englishes .................................................. 21
     - Global spread of English .................................. 21
     - New varieties of English .................................. 22
     - English as an International Language (EIL) ............. 25
   - English Use and English Language Teaching in Japan ............................................. 28
     - Japanese English ............................................. 28
     - ELT in Japan and its impacts .......................... 30
     - English as an international language and Japanese users ............................................. 31
     - Summary ..................................................... 36
### 3. Procedures

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantitative Inquiry</td>
<td>63</td>
</tr>
<tr>
<td>Research site</td>
<td>67</td>
</tr>
<tr>
<td>Subject selection</td>
<td>67</td>
</tr>
<tr>
<td>Determination of sample size</td>
<td>68</td>
</tr>
<tr>
<td>Criteria for selecting participants</td>
<td>68</td>
</tr>
<tr>
<td>L1 backgrounds</td>
<td>68</td>
</tr>
<tr>
<td>English proficiency</td>
<td>69</td>
</tr>
<tr>
<td>Educational/vocational backgrounds</td>
<td>70</td>
</tr>
<tr>
<td>Gender and age</td>
<td>71</td>
</tr>
<tr>
<td>Ethical considerations and incentives</td>
<td>71</td>
</tr>
<tr>
<td>Main independent variables</td>
<td>72</td>
</tr>
<tr>
<td>Speakers</td>
<td>72</td>
</tr>
<tr>
<td>Speech materials</td>
<td>74</td>
</tr>
<tr>
<td>Outcome measures</td>
<td>77</td>
</tr>
<tr>
<td>Evaluative Reactions scale</td>
<td>77</td>
</tr>
<tr>
<td>Intelligibility test</td>
<td>80</td>
</tr>
<tr>
<td>Rival independent variables</td>
<td>81</td>
</tr>
<tr>
<td>Personal information questionnaire</td>
<td>81</td>
</tr>
<tr>
<td>Conditions of testing</td>
<td>82</td>
</tr>
<tr>
<td>Data analysis</td>
<td>83</td>
</tr>
<tr>
<td>Qualitative Interview Study</td>
<td>87</td>
</tr>
<tr>
<td>Participants</td>
<td>87</td>
</tr>
<tr>
<td>Interview questions</td>
<td>88</td>
</tr>
<tr>
<td>Data collection</td>
<td>88</td>
</tr>
<tr>
<td>Data analysis</td>
<td>90</td>
</tr>
<tr>
<td>Combining the quantitative and qualitative data</td>
<td>90</td>
</tr>
<tr>
<td>4. Results and Discussion</td>
<td>92</td>
</tr>
<tr>
<td>Quantitative <em>ex post facto</em> study results</td>
<td>92</td>
</tr>
<tr>
<td>Descriptions of Participants</td>
<td>92</td>
</tr>
</tbody>
</table>
Data analysis ........................................................................................................98
  Descriptive analysis of variables ........................................................................98
    Independent Variables ........................................................................................98
  Dependent variable: Evaluative Reactions .........................................................103
  Dependent variable: Intelligibility .........................................................................104
  Correlation Analysis ............................................................................................110
  Regression analysis .............................................................................................115
  Evaluative Reactions ............................................................................................118
    Initial regression analysis of evaluative reactions .............................................118
    Further regression analysis of evaluative reactions .........................................119
    Regression analysis of evaluative reactions to different proficiency levels ..........121
    Evaluative reactions to native speakers ..........................................................121
    Evaluative reactions to advanced Japanese speakers ......................................123
    Evaluative reactions to intermediate Japanese speakers ...............................124
    Evaluative reactions to rudimentary Japanese speaker ....................................125
  Summary of the evaluative reactions analyses .................................................126
  Intelligibility .........................................................................................................127
    Initial regression analysis of intelligibility .......................................................127
    Further regression analysis of intelligibility of speakers with different proficiency levels .................................................................131
    Intelligibility of native speakers ......................................................................131
    Intelligibility of advanced Japanese speakers .................................................132
    Intelligibility of intermediate Japanese speakers ..........................................133
    Intelligibility of rudimentary Japanese speakers .........................................134
  Summary of intelligibility analyses .................................................................135
  Interpretation of regression results ....................................................................136
  Discussion of quantitative study findings .........................................................138

Qualitative interviews findings ...........................................................................148
  Descriptions of participants ..............................................................................149
  Responses to the interview questions ...............................................................150
  What did you feel while you were rating this speaker? .......................................150
    Speaker 1 .........................................................................................................150
    Speaker 2 .........................................................................................................151
    Speaker 3 .........................................................................................................152
    Speaker 4 .........................................................................................................153
    Speaker 5 .........................................................................................................154
    Speaker 6 .........................................................................................................155
    Speaker 7 .........................................................................................................156
    Speaker 8 .........................................................................................................157
  What are accents? Do they bother you? ..............................................................158
  Have you found specific accents or specific language speakers more difficult to understand than others? Why? .................................................159
  Do you feel more comfortable speaking with NS or NNS, or doesn’t it matter? .................................................................160
What do you think about international usage of English? About nonnative speakers speaking English? ..................................................160
Recurring themes ...............................................................................162
  Perceived Intelligibility/Comprehensibility .......................................163
  Irritation .........................................................................................165
  Accents .........................................................................................168
  Criteria Alteration ..........................................................................172
  English as an international language? ..............................................175
  Power relationship between native and nonnative speakers
  (or among 3 Concentric Circles) ......................................................180
Reflection on qualitative interviews ..................................................184
Summary of qualitative interviews ....................................................185

5. Summary, Conclusion, and Pedagogical Implications .........................187
  Summary of the findings ...................................................................189
  Research hypothesis 1 (H1) ...............................................................189
  Research hypothesis 2 (H2) ...............................................................191
  Research hypothesis 3 (H3) ...............................................................193
  Research hypothesis 4 (H4) ...............................................................193
  Research Question for Qualitative Interviews .................................194
  Combining quantitative and qualitative findings ..............................195
  Conclusion ......................................................................................198
  Pedagogical implications ...............................................................201
  Recommendations for further research ..........................................203
  Limitations of the study .................................................................204

List of references ..............................................................................207

Appendices
  A. Speech transcriptions and rates of speech ....................................220
  B. Visual cue for narration ...............................................................226
  C. Evaluative Reactions scale ........................................................228
  D. Intelligibility test ........................................................................232
  E. Personal Information questionnaire .............................................236
  F. Informed consent form (for quantitative study participants) ..........240
  G. Informed consent form (for qualitative study participants) ..........243
  H. Scatterplots for metric data ........................................................246
### LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Comparison among listeners with various L1 backgrounds and speakers with different English proficiency levels</td>
<td>66</td>
</tr>
<tr>
<td>3.1</td>
<td>Cronbach’s alpha across eight speakers</td>
<td>79</td>
</tr>
<tr>
<td>3.2</td>
<td>The scales of measurements and descriptive statistics used in the data analysis</td>
<td>84</td>
</tr>
<tr>
<td>3.3</td>
<td>Magnitudes of association</td>
<td>85</td>
</tr>
<tr>
<td>4.1</td>
<td>Listeners’ native language (L1)</td>
<td>93</td>
</tr>
<tr>
<td>4.2</td>
<td>Descriptive statistics for age, gender, US residence, and TOEFL scores</td>
<td>95-97</td>
</tr>
<tr>
<td>4.3</td>
<td>Descriptions of the speakers (gender and English proficiency levels)</td>
<td>99</td>
</tr>
<tr>
<td>4.4</td>
<td>Listeners’ familiarity with Japanese English</td>
<td>101</td>
</tr>
<tr>
<td>4.5</td>
<td>Listeners’ exposure to foreigners (nonnative English speakers)</td>
<td>103</td>
</tr>
<tr>
<td>4.6</td>
<td>Descriptive statistics of Evaluative Reactions</td>
<td>104</td>
</tr>
<tr>
<td>4.7</td>
<td>Mean comparison among L1s (Evaluative Reactions)</td>
<td>105</td>
</tr>
<tr>
<td>4.8</td>
<td>Mean comparison (Reactions): L1 groups and speakers' proficiency</td>
<td>106</td>
</tr>
<tr>
<td>4.9</td>
<td>Descriptive statistics of Intelligibility</td>
<td>108</td>
</tr>
<tr>
<td>4.10</td>
<td>Mean comparison among L1s (Intelligibility)</td>
<td>109</td>
</tr>
<tr>
<td>4.11</td>
<td>Mean comparison (Intelligibility): L1 groups and speakers' proficiency</td>
<td>110</td>
</tr>
<tr>
<td>4.12</td>
<td>Correlations (associations) among variables</td>
<td>111</td>
</tr>
<tr>
<td>4.13</td>
<td>Correlation matrix of variables</td>
<td>117</td>
</tr>
<tr>
<td>4.14</td>
<td>Multiple regression results on Evaluative Reactions (Simultaneous)</td>
<td>119</td>
</tr>
</tbody>
</table>
4.15 Further regression analysis of Evaluative Reactions (Stepwise) ...............120
4.16 Regression analysis of Evaluative Reactions to native speakers (Stepwise) ...122
4.17 Regression analysis of Evaluative Reactions to advanced speakers ..........123
4.18 Regression analysis of Evaluative Reactions to intermediate speakers .......124
4.19 Regression analysis of Evaluative Reactions to rudimentary speakers ......125
4.20 Regression analysis of Intelligibility (Stepwise) ........................................128
4.21 Regression analysis of Intelligibility on selected IVs .................................129
4.22 Regression analysis of Intelligibility of native speakers .........................132
4.23 Regression analysis of Intelligibility of advanced speakers ....................133
4.24 Regression analysis of Intelligibility of intermediate speakers ...............134
4.25 Regression analysis of Intelligibility of rudimentary speakers ..............135
4.26 Summary of regression analyses .............................................................137
4.27 Demographics of qualitative interview participants ..............................149
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Conceptual model of the <em>ex post facto</em> study</td>
<td>65</td>
</tr>
<tr>
<td>3.2</td>
<td>Data analysis model of the <em>ex post facto</em> study</td>
<td>86</td>
</tr>
<tr>
<td>4.1</td>
<td>Responses to Speaker #4 accents across listeners’ L1</td>
<td>140</td>
</tr>
<tr>
<td>4.2</td>
<td>Responses to Speaker #7 accents across listeners’ L1</td>
<td>141</td>
</tr>
<tr>
<td>B.1</td>
<td>Visual cue for narration</td>
<td>227</td>
</tr>
<tr>
<td>H.1</td>
<td>Correlation between familiarity with JE and evaluative reactions</td>
<td>247</td>
</tr>
<tr>
<td>H.2</td>
<td>Correlation between familiarity with JE and intelligibility</td>
<td>248</td>
</tr>
<tr>
<td>H.3</td>
<td>Correlation between familiarity with JE and exposure to foreigners</td>
<td>249</td>
</tr>
<tr>
<td>H.4</td>
<td>Correlation between exposure to foreigners and evaluative reactions</td>
<td>250</td>
</tr>
<tr>
<td>H.5</td>
<td>Correlation between exposure to foreigners and intelligibility</td>
<td>251</td>
</tr>
<tr>
<td>H.6</td>
<td>Correlation between evaluative reactions and intelligibility</td>
<td>252</td>
</tr>
</tbody>
</table>
Masao is a 35 year-old Japanese businessman. He is working in a branch office of a Japanese automobile company in Brussels, Belgium. Since Masao does not speak Flemish nor French at all and his German is at the rudimentary level, he usually speaks with his Belgian colleagues in English. Masao is not confident in his English, although his grades in English as a school subject were always better than the class average, albeit by a narrow margin. It has been a month since he was transferred to this branch from Japan. Masao recollects. At first, he was impressed with how fluently his new colleagues spoke English, which was supposedly a foreign language to them as well. He was worried that they might not understand his rusty English, which he had not used for more than 10 years after his graduation from college. He was greatly relieved when he found that they seemed to understand him, even though minor, manageable miscommunications occurred daily. Now Masao is more concerned with his colleagues’ perception of him. Do they recognize him as the capable businessman that he is? Sometimes Masao cannot help but wish he could be speaking Japanese with them. In English, he is afraid, he must sound inexperienced, maybe childish or even stupid…..
Today, many nonnative speakers of English face situations similar to Masao’s, although he is fictional. A person who has learned English as a school subject one day finds him/herself in a circumstance where s/he has to use it to communicate with people whose mother tongue is different from his/hers. The need for the investigation of the acceptability and intelligibility of nonnative English has been widely recognized while the language has established its status as the most powerful international language, as Strevens (1980) predicted. The majority of research studies on these issues, however, have been conducted based on the native speakers’ judgements (e.g., Anderson-Hseih and Koehler, 1988; Anderson-Hseih, Jhonson, and Koehler, 1992; Derwing and Munro, 1997). In other words, native speakers have almost always judged the acceptability and intelligibility of nonnative speakers’ English. However, the rationality of integrating nonnative perspectives into the discussion on English as an international language is evident, when what “international” communication really means is deliberated. Nonnative speakers listen to other nonnative speakers as often as, or even more frequently than, they listen to native speakers, as in Masao’s case, since more international communication is conveyed between two (or more) nonnative speakers than between a native speaker and a nonnative speaker (Kachru, 1992; Matsuda, 2003).

Situations like Masao’s are gaining more prominence in proportion to the spread of English an international/world language (Smith, 1983; Nakayama, 1982, 1994; Brown, 1995; Widdowson, 1994). It is the language used for global cross-cultural communication (Smith and Rafiqzad, 1979), and enjoys its “dominance” (Tsuda, 1996; Swales, 1997), “hegemony” (Tsuda, 1999) and “power” (Pennycook, 2001) as an international lingua franca. Crystal (2003) estimates that the number of native speakers (NSs), who use
English as a mother tongue or first language (L1), is about 323 to 329 million, while a further 422 to 500 million nonnative speakers (NNSs) use English as an institutionalized language. He states that the NNS population who learns English as a foreign language is around 500 million to 1 billion, depending on the definition of English users. Simply put, NNSs of English outnumber NSs. As a result, English is now used in contexts where there are no NSs and this unprecedented spread of English involves practically every part of the world (Kachru, 1992). According to Crystal (1997), “Mother-tongue use by itself cannot give a language global status” (p. 3). In other words, the status of English as an international/world language depends on its use by NNSs, as well as NSs.

The spread of English has led to the diversification of the language, namely the development of localized nonnative norms, somewhat different from native norms in pronunciation, vocabulary, grammar, pragmatics and communication styles. It has created issues that had never been raised before. Two concerns relevant to the current study were verbalized by Strevens (1980) as the diversification of English progressed.

One was the decline of international intelligibility of English, which means that English might eventually become unintelligible among speakers of different varieties as the result of localization and nativisation. A good example of this concern is the argument by Prator (1968) criticizing the unintelligibility of Indian English. More recently, Crystal (1997) wrote, “Inevitably, the emergence of new Englishes raises the spectre of fragmentation --- the eventual dissolution of English into a range of mutually unintelligible languages (as happened when Latin gave rise to the various Romance languages, such as French, Spanish, and Italian, over 1000 years ago)” (p.134).
The second concern was the value judgement of *acceptability* of language, in other words, evaluative reactions to nonnative varieties. Urdang (1990) states that a listener’s “pre-judgement” of a speaker based on perceived accents is inevitable in any kind of spoken communication, and asserts that nonnative varieties are devalued in comparison with native counterparts. Therefore, Quirk (1990) and Urdang (1990) insist that the dominant models of English should remain native varieties, especially when international communication is the purpose of teaching/learning English.

The fact that a large quantity of research has been conducted on the acceptability and intelligibility of nonnative English varieties is proof that Strevens (1980) concerns were valid. However, the majority of the studies have been conducted from NSs’ perspectives. In other words, the judges and evaluators of nonnative English have been native listeners. Given that native speakers of English are and will always be important participants in international communication conveyed in English, to include native speakers’ perspectives in research is critical. This does not at all mean, however, that native judgements alone should be measured concerning nonnative varieties’ intelligibility and acceptability. The following viewpoints provide further evidence of the need to focus more attention on NNSs in global English communication.

The aforementioned diversification of English has fostered innovative, liberal, and postcolonial views of the language and the language users. They have shed new light on the issues from the ownership of the language to the appropriateness of norms/standards/models of English instruction, and the way English should be learned/taught and researched. As for the ownership of English, Widdowson (1994) proclaims that English now belongs not only to its NSs, but also to NNSs. Even though
the majority of lay persons might still believe that the ownership of English is in the hands of NSs (e.g., Matsuda, 2003), many scholars of the English language have recognized and accepted the viability of nonnative English varieties (Bamgbose, 1998).

The norms/standards/models of English have also been debated. First of all, to clarify the regional differences in English usage, Kachru (1985) classified English-speaking contexts into three concentric circles: (1) Inner Circle, where English is used as a mother tongue (e.g., the United States and the United Kingdom), (2) Outer Circle, where English is used as an institutionalized language (e.g., Bangladesh, Ghana, and India), and (3) Expanding Circle, where English is taught as a foreign language (e.g., China and Japan). Kachru (1985) made these distinctions, because the traditional NS versus NNS dichotomy, as seen in Quirk (1981) and Trudgill and Hannah (2002), could not capture the complexity of the current situation.

Kachru (1982, 1985 and 1986) claims that Outer Circle English varieties should be treated in their own right, in applied linguistics and English Language Learning (ELT), as legitimately as such major native varieties as American and British English. In other words, Indian English, for example, should not be judged based on the British norms, because it serves local pragmatics in India. The majority of Indian English users speak English for intranational communication purposes, that is, only among Indians. Indian people may claim that they are native speakers of Indian English, even though they may be nonnative speakers of British English. Sridhar (1994) and Kachru (1994) also emphasize the necessity of research on Outer Circle English varieties, while criticizing the Inner Circle-centered-ness of applied linguistics research and ELT. This line of
thought on varieties of English is generally called World Englishes (WE), or New Englishes points of view.

Although the discourse of World Englishes is powerful and convincing, and more importantly, emancipating for Outer Circle English users, it is not satisfactory for the Expanding Circle. In the attempt to raise the global recognition and acceptance of Outer Circle English varieties, the importance of the Expanding Circle has been minimized. Sridhar (1994) and Kachru (1994), for example, strongly demand research conducted in the Outer Circle but fail to mention the Expanding Circle. A good example of this type of prioritization of world users of English can be seen in the book entitled English around the World (Chesire, 1991) that includes research from both the Inner and Outer Circles, but excludes the Expanding Circle completely.

While describing Outer Circle varieties as institutionalized, Kachru (1992) calls an Expanding Circle variety a performance variety, influenced by individual learner differences rather than by local/communal norms. The lack of consistent local norms and the subsequent difficulty of codification make the distinction between Expanding Circle English varieties and the “learner language” (Ellis, 1994) difficult (Strevens, 1987/1992). As a result, the Expanding Circle has been left in the conventional frame of “nonnative speakers” and its varieties have been treated as interlanguage (Selinker, 1972), which is a developmental and incomplete form of the target language, that is, an Inner Circle variety. Consequently, less controversy has existed about the language norms to be instructed in the Expanding Circle, and few have questioned the validity of conducting research on Expanding Circle English solely based on native judgements.
However, the significant role of the Expanding Circle varieties in the investigation of English as an international language (EIL) has recently been acknowledged to some degree. Yano (2001) argues that the need for international communication connects all three concentric circles and conceptualizes the Expanding Circle not as merely a component of EIL but potentially the most informative among the three for the investigation of EIL. In the empirical research area, Seidlhofer (2001) conducted a study on NNS-NNS communication by means of discourse analysis, based on the belief that EIL should not be confused with English as a native language (ENL) of the UK or USA.

The importance of investigating English usage in the Expanding Circle as a variety, not as a deficient imitation of Inner Circle English, is also evident from the following facts. First of all, the majority of English learners in the Expanding Circle no longer study English as a foreign language in the traditional sense, that is, to absorb the culture of another nation (Baxter, 1980). Japanese people, for example, learn English as a tool for wider cross-cultural communication beyond use in Anglophone, or English dominant, countries. Many of them do not want to speak like Americans (Baxter, 1980). As such, Suzuki (1971) and Nakayama (1994) claim that Inner Circle cultures can be and should be detached from English education in Japan. Second, the number of English learners/users in the Expanding Circle surpasses that of the Inner and the Outer Circles separately, and the Expanding Circle is, literally, now ‘expanding’ (Graddol, Leith, and Swann, 1996; Kachru, 1997). Therefore, Expanding Circle English speakers are no longer the minority, at least in number, in global English communication. Third, the Inner Circle-centered ideologies of ELT (Pennycook, 2001), combined with so-called
Westernization of many countries in the Expanding Circle, has created a serious social, cultural and political imbalance (Lummis, 1976; Kubota, 1998; Tsuda, 1999), which will be discussed in the following chapter. ELT Professionals, as well as English applied linguists, are responsible for helping adjust the outmoded notions of the importance of NS English, and they need research conducted from the impartial view of language varieties rather than NS English as the norms or model.

Statement of the Problem

Based on the above discussions, the necessity of inquiring into an Expanding Circle English variety is clear, and the inquiries should include both native and nonnative judgements from all three concentric circles. In particular, such critical issues as acceptability and intelligibility should be examined from the broader perspectives. However, this type of research is relatively scarce.

Among the few studies that featured NNSs’ perspectives, one theoretical and methodological characteristic is commonly observed. The nonnative speakers from whom speech samples were taken were very advanced in English proficiency. For example, the speakers in the study of Smith (1992) were college/university teachers and those used in Munro (1995) were graduate students and faculty at a Canadian university.

Several explanations are possible for the choice of speakers. First, the research sites that were usually within the Inner Circle may have limited the recruitment of speakers. Most international college students recruited in Inner Circle universities were likely to be already quite advanced speakers of English. Second, when a study compared
NNSs from the Outer and Expanding Circles, the proficiency levels of the speakers had to be as equal as possible to control extraneous variables, but speakers’ average English proficiency might be much higher in the Outer Circle than in the Expanding Circle. Therefore, low-proficiency Expanding Circle speakers might have been intentionally avoided in studies. Third, many scholars actually consider EIL the language for international leadership, as Smith (1998) argues. The rationale may be that the nonnative population who (would) use English for international communication is limited to educated professionals, who are already fairly fluent in spoken English.

These very advanced speakers are, however, unlikely to represent the typical English users in the Expanding Circle. Ellis (1994) refers to the ultimate proficiency level of most Expanding Circle English users by stating, “classroom learners often fail to develop much functional language ability” (p. 228). Similarly, Stanlaw (1992) describes the English proficiency of ordinary Japanese people as follows, “…a typical Japanese student may study it (= English) for six to ten years; still few Japanese actually speak English well enough to converse with foreigners beyond a rudimentary exchange of greetings” (p. 181). If relatively low proficiency is a characteristic of Expanding Circle varieties, then it should be built into the research on the Expanding Circle varieties.

Today, less-advanced English users originated from the Expanding Circle seem to have more and more opportunities and needs to communicate with both native and nonnative English users, in both spoken and written forms, through increasing tourism, enterprise, and the advancement of Informational Technology. Yet, many years of residence in the Inner Circle countries to ‘acquire’ the language is not an option for most of the English users/learners in the Expanding Circle. Their exposure to Outer Circle
English may also remain limited. For these reasons, Expanding Circle countries may find it difficult to focus on developing international communicability of their English. Given these realities, the investigation of the international acceptability and intelligibility of less-advanced nonnative speakers from the Expanding Circle to native and nonnative listeners is imperative, but scarce. This study aims to fill the gap.

The current study investigated the international acceptability and intelligibility of Japanese English, namely, English spoken by native speakers of Japanese. Japanese English was chosen as the target language variety for investigation, because English is taught and learned in Japan almost exclusively for the purpose of international communication (Yano, 2003). Japanese English is indeed one of the varieties frequently heard in international contexts. Especially in Asia, the political, economical and cultural bonds among countries have recently become even stronger, and the importance of English as a lingua franca in Asia has been widely recognized in Japan owing to the studies of such scholars as Honna (1990). Furthermore, Japanese companies are present throughout Asia, so that Japanese English is a variety commonly spoken and heard in that region of the world. Consequently, mutual attitudes and intelligibility in English communication have increasingly become critical issues, including the Japanese context.

Significance of the Study

In order to nurture egalitarian internationalism through English communication and to fully enjoy its benefits, knowledge of NNS-NNS communication is crucial, but its investigation has just had its genesis. New English norms exclusively intended for
international communication have ardently been called for by scholars such as Suzuki (1971), Smith (1976), Quirk (1981), Nakayama (1982), and Yano (2001), but concrete norms have not been suggested yet, because of the lack of knowledge of NNS-NNS communication.

This study penetrates into the acceptability and the intelligibility of nonnative English speakers to other nonnative listeners and sheds light on their vital aspects of English as a world language. The findings in the study are expected to offer some insights for English applied linguistics and ELT. Furthermore, the findings of this study will be especially beneficial for ELT decision-makers in the Expanding Circle, who give credence to English for international communication, but have had no choice but follow the native norms due to the lack of research-based knowledge about EIL.

Even among Expanding Circle countries, the gravity of international communication in teaching/learning English makes Japan unique, while many other nations have started to use English for *intranational* purposes as well (Bamgbose, 1998; Yano, 2001). Furthermore, Yano (2001) states, “the users of English as a Foreign Language (EFL) who are relatively free from the localization of English might be able to contribute to create, maintain, and develop the globalized standards of English” (p. 119). Therefore, the current study offers insights not only for Japanese English users or the Expanding Circle but also for all those who utilize English for international communication.
Objectives of the Study, Hypotheses and Research Questions

The foremost purpose of this study is the examination of the evaluative reactions to and the intelligibility of Japanese English across native and nonnative listeners. In other words, the goal is to determine the degree to which listeners’ L1 backgrounds affect their attitudes toward and their comprehension of Japanese English. As mentioned in the statement of the problem, the variability of English proficiency levels among Japanese English users was also considered a potential factor that influences international communication.

The listeners were chosen from native speakers of American English, Hindi (Indian), Mandarin Chinese, and Japanese, respectively representing the Inner, Outer, Expanding Circles, and fellow countrymen. The listeners’ evaluative reactions to Japanese English were assessed on a semantic differential scale that was developed by the researcher based on existing scales. The intelligibility of the speakers was measured on a word-for-word dictation test, based on the recommendation by Subtelny, Whitehead, and Orlando (1980) that extemporaneous speech samples, rather than oral reading, should be used to assess intelligibility.

In the main framework, this is an *ex post facto* study, because main independent variables, namely, listeners’ L1 and speakers’ proficiency levels, could not be manipulated by the researcher. The following two hypotheses were formulated to be tested through the quantitative research.

**H1: The listeners’ evaluative reactions to Japanese English will be predicted by:**

(1) The listener’s L1 (American English, Hindi, Mandarin, and Japanese)
(2) The speaker’s proficiency

H2: The intelligibility of Japanese English will be predicted by:

(1) The listener’s L1

(2) The speaker’s proficiency

Three rival independent variables were chosen based on the review of related research. They were (1) listeners’ gender, (2) listeners’ familiarity with Japanese English, and (3) listeners’ exposure to NNSs of English (foreigners).

Coats (1986) found that, contrary to common belief, gender explains findings in sociolinguistic better than social-economical status. Gender difference was found to be significant among speakers in the study of Alford and Strother (1990) on native and nonnative listeners’ attitudes toward American regional accents. In the study of Cargile (2002) on the effect of information processing, gender again created a big difference. Listeners’ familiarity with a specific accent was found to be strongly related to their intelligibility scores (Gass and Varonis, 1984; Derwing and Munro, 1997). Listeners’ exposure to foreign languages and cultures has also been associated with their attitudes towards accented speech (Gass and Varonis, 1984, Anderson-Hsieh and Koehler, 1988; Matsuura, Chiba, and Yamamoto, 1992).

Based on the findings of the studies mentioned above, the following research hypotheses were formulated for each rival independent variable.

H3: The listeners’ evaluative reactions to Japanese English will be predicted by:

(1) The listener’s gender

(2) The listeners’ familiarity with Japanese English
(3) The listeners’ exposure to foreigners

H4: The intelligibility of Japanese English will be predicted by:

(1) The speaker’s and listener’s gender

(2) The listeners’ familiarity with Japanese English

(3) The listeners’ exposure to foreigners

In order to better understand NNS-NNS communication, however, qualitative data were additionally obtained. Therefore, the whole design of this study is what Creswell (1994) calls the *dominant-less dominant design*, in which a single dominant paradigm is chosen as the framework of the study with a small component drawn from the alternative paradigm. The advantage of this research design, according to Creswell (1994), is that “it presents a consistent paradigm picture in the study and still gathers limited information to probe in detail one aspect of the study” (p. 177).

The quantitative paradigm was chosen as the framework, since the main interest of this study was the *general tendency* of the participants based on their L1 backgrounds. The existence of numerous quantitative studies on nonnative speech from NSs’ perspectives has provided a firm enough foundation, to which this study aims to add a new piece of knowledge. On the other hand, the participants’ individual differences in, as well as the processes of, evaluative reactions and comprehension were also of great interest but relatively unexplored. The ontology of qualitative paradigm, that is, the assumption that reality is subjective and multiple, was judged more appropriate for the exploratory purposes. The scarcity of empirical descriptions of NNS-NNS
communication, in other words, the dearth of research on nonnative speech acceptability and intelligibility from the NNSs’ perspectives, also made a priori designs unfeasible.

In the qualitative phase of the study, the listeners were interviewed on their evaluation reactions. The research question to be answered through the interviews was:

**What are the perceived factors that influence the listeners’ evaluative reactions to and the intelligibility of Japanese English?**

This question was a grand tour question (Creswell, 1994), and thus considered as the guide to explore the processes of judgements by individually unique participants.

**Definition of Terms**

To ensure the clarity and consistency throughout this study, definitions of terms are listed below.

**Evaluative Reactions**: This term is used to represent affective domain of the listeners, namely, “what people feel, perceive and believe” (Miller, 2001, pp.23). It involves attitudes toward, responses to, impressionistic evaluation and judgement of, what they hear. Operationally, listeners’ evaluative reactions are measured on the Evaluative Reactions scale that was developed by the researcher. The scale is a summated rating scale from which a mean can be calculated and treated as interval data. The score ranged from 8 (extremely negative) to 80 (extremely positive).

**Exposure to foreign languages/ cultures**: The frequency with which a participant listened to English-speaking non-Americans, other than Japanese speakers, during
the past year, on average. The listeners indicated the frequency by choosing mutually exclusive categories from 1: almost never except for greetings and formulaic expressions to 5: almost every week or more frequently, on the personal information questionnaire. The obtained datum was treated as ordinal, since only one item was used for the measurement and therefore it was not a summated datum.

**Familiarity with Japanese English**: The frequency with which a listener listened to an L1 Japanese person speaking English, during the past year. The listeners indicate the frequency by choosing mutually exclusive ordinal numbers from 1: almost never except for greetings and formulaic expressions to 5: almost every week or more frequently, on the personal information questionnaire. The obtained datum was treated as ordinal, since only one item was used for the measurement and therefore it was not a summated datum.

**Gender**: The male-female dichotomy that a participant self-reported on the personal information questionnaire. The obtained data was nominal.

**Intelligibility**: The accuracy of word/utterance recognition (Smith 1992) by the listener, measured by the percentage of correct words in a dictation task that exactly match with the transcription. The data obtained by means of this test were treated as ratio.

**Japanese English**: L2 English spoken by L1 Japanese speakers. Unlike Outer Circle English users, people in the Expanding Circle do not develop localized models/norms. Nonetheless, as Yano (2001) describes, Japanese people inevitably indicate common linguistic and sociocultural characteristics when
they speak English. Tarone, Cohen, and Dumas (1983) call this tendency “transfer” from the native language, one of the most significant characteristics of interlanguage.

Listeners’ L1: A person’s first or native language. The definition of L1 is very controversial. For example, Richards, Platt, and Platt (1992) state that L1 is usually acquired in a person’s childhood, used in his/her family, and spoken in a country where the person lives, but also state that there are many exceptional cases. In this study, L1 is defined as the language that a person acquired early in his/her childhood and that s/he uses among direct family members. Operationally, a listener’s L1 was the language that s/he self-reported by answering the question “what is your native language?” on the personal information questionnaire. Due to the purposive sampling, L1 was supposed to be one of the following languages: American English, Hindi, Chinese or Japanese. The data were treated as nominal.

Speakers’ proficiency: Proficiency is a learner’s ability to use the target language for specific or general purposes, productively (i.e., speaking and writing) or perceptively (i.e., listening and reading). Operationally, a speaker’s proficiency was measured by the TOEFL (Test of English as a Foreign Language) scores. To facilitate comparisons, the computer-based TOEFL scores were converted into paper-based counterparts. When TOEFL scores were not available, the researcher converted the speakers’ scores on other proficiency tests, such as TOEIC (Test of English for International Communication) and EIKEN (jituyou eigo kentei shiken [the test of practical English], which is a Japanese standardized English
proficiency examination). In this study, the speakers’ proficiency varied in the range of 400 and below to 600 on paper-based TOEFL. Since TOEFL scores were blocked into three categories (advanced, intermediate, and beginning-level), the data were treated as ordinal.

Basic Assumptions

1. Attitudes (defined as evaluative reactions to nonnative speech) have directions and can be measured quantitatively by means of a semantic differential scale.

2. Participants respond as truthfully as possible and to the best of their knowledge in the questionnaires. Therefore, the responses are trustworthy.

3. All the participants (i.e., listeners) possess sufficient language skills and intellectual ability to understand the questionnaire written in English and to follow the instructions.
CHAPTER 2

REVIEW OF THE LITERATURE

Fishman (1971) disproves the view of language as a mere carrier of content, and argued that language itself is content. In other words, people receive messages not only from what other people say, but also from how they say it. Cargile and Giles (1996) state, “…language is a powerful social force that does more than convey intended referential information. For better or worse, hearers can react to linguistic and paralinguistic variation in messages as though they indicate both personal and social characteristics of the speaker” (p. 338). Edwards (1999) emphasizes the filtering effects of “attitudes” in the perception of language. He claims that people do not intake sensory data as they are, but that the data are always filtered. That is, group members share the social “filters,” which Edwards (1999) also calls “stereotypes” or “culture.” Ryan, Giles and Sebastian (1982) state, “in every society the different power of particular social groups is reflected in language variation and in attitudes toward those variations” (p. 1). Therefore, language attitudes have been of significant interest to sociolinguists and social psychologists (Edwards, 1982). The findings from language attitudinal studies have been utilized for decision making, from an individual’s accent correction to national language policies. In
general, the language variation of the dominant group is valued more favorably than that of the minority group, and speakers of a lower prestige language tend to sacrifice their native speech style for the mainstream one in order for social successes (Ryan, Giles and Sebastian, 1982).

Nonnative English has long been investigated within this framework. By deviating from accepted native norms in phonology, syntax, and lexicon, a speaker has been associated with being a tourist or an immigrant in a context where English is a native language. Raters have usually been native speakers (NSs), the dominant group with power. Researchers have found that foreign-accented speakers have generally been downgraded in comparison with NSs.

On the other hand, the discourse of English as an International Language (EIL) requires the expansion of the speech community of English. Nonnative speakers (NNSs) are now the majority of English users in number (Kachru, 1992), although the current power relationship between NSs and NNSs in communication might not be equal (Tsuda, 1997). Nonnative speakers’ attitudes toward other NNSs are as important as, if not more important than, those of NSs. Mutual intelligibility also becomes an crucial issue for communication among nonnative interlocutors.

In the following review of the literature, the discourse of World Englishes (WE) will first be deliberated, followed by the proposition of English as an international language (EIL). English use and English education in Japan will be debated, in relation to EIL and WE. Studies on native listeners’ evaluative reactions to nonnative speech will be reviewed, followed by studies on nonnative listeners’ attitudes. Finally, research on the intelligibility of nonnative English across native and nonnative listeners will be examined.
World Englishes

The Global Spread of English

English has become an international language, when the number of its users is considered. According to Pennycook (1995), approximately 700 million to one billion people use English today. Ammon (1992) estimated the number of global English speakers to be 1.5 billion. With the most inclusive definition of “users”, Strevens (1992) and Kachru (1996) estimated that more than two billion people are using English with some proficiency. Among this gigantic population of English users, the number of NSs accounts for a fraction. According to an estimate by Crystal (1988), the population of NSs of English is slightly more than 300 million. What makes the English speaking population so large is the number of NNSs.

Crystal (1997) argued, however, that the number of speakers is not the determining factor for the global status of the language. For example, almost 1.3 billion people speak Chinese (Mandarin) in Republic of China and Taiwan as their mother tongue or public language. But few would claim that Chinese is the most powerful international language (Graddol, Leith, and Swann, 1996). Who speaks the language matters, according to Crystal (1997); in other words, how much political power the speakers of the language have. For political, scientific, technological, business, and academic purposes, English is more widely utilized than any other language in the world (Crystal, 1988; Fishman, 1992; Ellis, 1994; Tsuda, 1997). Smith (1998) describes English as the language of international leadership, which means that English is used as a lingua franca by field leaders from hundreds of countries. The statement that English is the most
dominant language for high-stakes, diplomatic communication is indisputable, even for those who resist its dominance (e.g., Phillipson, 1992; Tsuda, 1999).

The dominance of English as a *lingua franca* for more general communication purposes is also beyond doubt (e.g., Trifonovitch, 1981; Campbell, Ekniyom, Haque and Smith, 1982; Crystal, 1988; Graddol, Leith, and Swann, 1996; and Ammon, 2001). English is learned as a second or foreign language in practically every part of the world, except for the parts in which it is being acquired as a first language, (Kachru, 1992). Graddol, Leith, and Swann (1996) maintain that what makes a language important in the eyes of the world is the extent to which the language is found *useful outside its original setting*. In sum, in light of the number of speakers, the leadership of the speakers, and the usefulness in nonnative situations, English is arguably the most powerful international language.

**New Varieties of English and Applied Linguistics**

English is unmistakably the most intensively researched language in Second Language Acquisition (SLA), applied linguistics and language teaching theories, thanks to its international status established rather rapidly, only during the last 40 years (Crystal, 1996). In research, NNSs have traditionally been treated as learners who struggle to acquire native English features, and their language was called *interlanguage* (Selinker, 1972) or *learner language* (Ellis, 1994). At the end of English language research studies, researchers have made suggestions as to how to be more like a native speaker.

On the other hand, the spread of English has produced many local varieties of English. That is, when English is spoken intensively and exclusively among local NNSs,
it is naturally influenced by the local users’ native language (L1) and culture, which results in a localized variety that is different from its original model (i.e., American and British English). In Strevens’ (1981) words, “… paradoxically, as this one single language expands, the diversity of forms within the total envelope of ‘English’ also increases: ‘more use of English’ is accompanied by ‘more different kinds of English” (p. 1).

Not surprisingly, countries where English was officially institutionalized and used as a second language initiated the movement of English research toward the recognition and acceptance of their English varieties as local norms, independent from native English. Kachru (1985) calls these countries the Outer Circle of English users, and has claimed that the Outer Circle has developed local pragmatic functions independent from British norms, which used to be considered the only “acceptable” norm in the colonial era (Kachru, 1965, 1981, 1982, and 1992). For Outer Circle English speakers, Kachru (1986) maintains, English is merely one of the additional languages they speak in accordance with situations and register. Therefore, pressing the monolingual native language view against Outer Circle English users is the marginalization of multilingualism and multiculturalism. Indian people are proud of and want to maintain the distinct features of their English variety, mostly in phonology but in grammar and vocabulary as well.

This liberal movement advocating World Englishes, however, met strong resistance. Prator (1968) criticized the promotion of Indian English by Kachru (1965) as the cause of international unintelligibility of English spoken by Indians. He cited King’ words (1971, cited by Mehrotra, 1982b), “English is going to diverge into a series of
Germanic languages as Latin did into the Romance languages” (p.2). When Smith (1976) proclaimed the detachment of English as an International Auxiliary Language (EIAL) from British and American norms and cultures, Prator (1978) again insisted that English learners do want to speak like NSs, especially in the countries where English is being taught as a foreign language, which Kachru (1985) called the Expanding Circle. Quirk (1981) shared the same opinion as Prator (1976) and claimed that the promotion of diversity in English would hinder the function of English as a tool for international communication.

The discussion of the norms/models of English research and instruction that arose with the diversification of the language seems to have settled in the following statements. For intra-national communication purposes (i.e., in the Outer Circle), local varieties might be more practical and appropriate (e.g., Mehrotra, 1982b; Widdowson 1982). For international communication, however, especially in the Expanding Circle, where there is no local model of English, either British or American standards, in other words, Inner Circle norms, should be taught.

The basis for the argument advocating the use of Inner Circle varieties as the model in the Expanding Circle has been twofold, as Strevens (1980) forecast. First, Outer Circle English will hinder mutual, namely international, intelligibility of English among speakers of different varieties (Prator, 1968; 1978). As a matter of fact, the progressive diversification among the Outer has been reported. For example, Strevens (1992) reported that Singaporean English and Malaysian English were more different from each other than they had been twenty years before. Mehrotra (1982a), who strongly advocates the independence of Indian English from native norms, admits that, “There is a great deal of
Indian English, particularly at the phonological and lexical levels, which is incomprehensible in varying degrees to the speakers of English outside the Indian subcontinent” (p. 74). The second advantage of native models to the Expanding Circle is its acceptability or the listeners’ evaluative reactions to the language. Inner Circle English varieties are more accepted and preferred, and therefore safer to learn (Strevens, 1980; Quirk, 1990). Urdang (1990) claims that English learners in the Expanding Circle do not want to learn language varieties that are socially downgraded.

**English as an International Language (EIL)**

The initial and fundamental discussion of World Englishes centered the Outer Circle, as seen above. The local norms in the Outer Circle are relatively stable and substantially codified, which are the prerequisites of being accepted as a new language variety, according to Bamgbose (1998). In comparison, Expanding Circle English varieties are less stable and hardly codified. Richards and Tay (1981) clearly differentiate the Outer Circle from the Expanding Circle, stating that the former has different norms from native English, while the latter has only different proficiency levels. Even Kachru (1997), who criticized the use of term “interlanguage” as the manifestation of monolingual Inner Circle language view in 1986, does not hesitate to describe the Expanding Circle as norm-dependent, and to call English used in the circle a *performance variety*. In short, the Expanding Circle seems to be left in the traditional definition of nonnative speakers as learners who struggle for an Inner Circle variety.

However, the contribution of the Expanding Circle varieties to the internationalization of English has been also discussed, though not as ardently as the
Outer Circle English, and recently taken up in research. Instead of dividing contexts based on geographical regions, Widdowson (1994) claimed that the situations should be distinguished according to register. Instead of “who speaks it,” “for what purpose it is spoken” should be considered. According to this typology, the Inner and Outer Circles may use English both for intra- and international communication purposes. On the other hand, the Expanding Circle may concentrate on international communication, which connects all three concentric circles (Yano, 2003). Yano (2001) claims that the Expanding Circle English speakers are the critical informants for English as an International Language (EIL), for they are free from the influence of localization of English. Seidlhofer (2001) argues that when English is used for international communication, it should be considered English as a Lingua Franca (ELF), and should not be confused with English as a Native Language (ENL) of the UK and USA. She has conducted research on how a NNS communicate with another NNS by means of ELF in the multilingual classroom.

Based on these arguments, Expanding English varieties should not be considered as a simple compilation of performance errors in imitating Inner Circle varieties. A new research paradigm is necessary in applied linguistics along with the traditional one that has investigated how nonnative English is different from its native counterpart. That is, the quest for the knowledge on the features of nonnative English that promotes and hinders crosscultural, essentially including NNS-NNS, communication has just been inaugurated.

The issues that are critical for the Expanding Circle, in other words, what language varieties are most appropriate for international communication in light of intelligibility and acceptability, are relevant to all concentric circles of English users. This
is because an English speaker, native or nonnative, needs to be understood by his/her
listener, again native or nonnative, with decent intelligibility, and evaluative reactions will
essentially exist when communication takes place. In other words, people in the Inner and Outer Circles who use English mainly for *intranational* communication purposes still maintain the possibility for having to use it for international communication, due to the intertwined international politics and the advancement of Informational Technology, global enterprise, and tourism. The investigation of the characteristics of English varieties that are relatively intelligible and acceptable for many is important for all those who use English for international communication purposes.

In summary, this section first described an issue raised by the diversification of English due to its global spread. The issue was the norms/models/standards of English in applied linguistics research and language education. Unlike the Outer Circle of English whose localized norms were recognized and institutionalized, the norms of the Expanding Circle remained to be Inner Circle English varieties. As a result, Expanding Circle English was treated as learner language, whose deviation from Inner Circle English was considered as nonnative errors that had to be corrected or eliminated. More recently however, a different view of the Expanding Circle has been recognized: it can offer important information and insights into understanding international communication through English.
What is “Japanese English”? Does such a thing as Japanese English exist? The answer is yes and no. Yes; it exists when the word is defined as English produced by a group in which members share their native language, in the same sense as Korean English is English spoken by people whose native language is Korean, as Smith (1976) stated. No; Japanese English does not exist in the same sense as American English and Indian English exist, because English spoken by native speakers of Japanese lacks localized forms/forms to be admitted as an independent language variety (Strevens 1992; Yano, 2001). Theoretically speaking, the English variety spoken by most L1 Japanese speakers should be exactly the same as its language model (i.e., American English in many cases). In other words, Japanese learners are supposed to speak American English, in the latter sense.

However, as Ellis (1994) stated, the majority of English learners in the Expanding Circle do not reach native-like proficiency, probably due to limited exposure to English and the learners’ lack of strong motivation to master the language. As for Japan, Stanlaw (1992) reported the surprisingly low fluency of English among Japanese people, who usually learn the language in school for six to ten years. The exact word “Japanese English” is used very derogatorily in Japan, mainly referring to the distinctive pronunciation traits that L1 Japanese transfer to L2 English. When the assumption that any deviations from the native model are shameful is removed, however, the very distinctive features that are commonly observed in English spoken by Japanese native speakers entitle their language variety “Japanese English.” In this study, the term
“Japanese English” is used to symbolize the Japanese speakers’ ownership of English, including their divergent levels of proficiency and common deviations from native norms.

In the typology by Kachru (1985), Japan belongs to the Expanding Circle, where English is neither the first nor second language. As seen in the discussion of World Englishes, the language standards/norms adopted in the Expanding Circle are less controversial than those in the Outer Circle. Many scholars such as Prator (1968, 1976), Quirk (1990), Richards and Tay (1981), Strevens (1981), and Kachru (1997) agreed that the standards in the Expanding Circle should follow the native varieties. Kachru (1997) conducted a survey study on university curricula in Japan and confirmed this tendency. Specifically, in Japan, British and American varieties of English are almost always taught, as well as their cultures with an emphasis on literature.

Given this, what benefits can the ideology of World Englishes offer to ELT in Japan? The benefits might be more attitudinal, rather than linguistic, in terms of drastic changes in language standards/models. Strevens (1992), as well as Cook (1999), stated that the ideology underlying World Englishes does not directly affect teaching approaches and methodologies. He stated, “It is rather the gradual sophistication in learning and teaching. English has now added a new element: awareness of the fact that most ESL/EFL today relates to NNS populations requiring English for their internal purposes, or for dealing with other NNS populations, without the presence or intervention of native speakers” (p. 40). Even though Japan is a country that is willing to follow the norms/standards of native models, the ideology of World Englishes might change the attitudes of Japanese learners toward English, English learning, and the global English speakers.
ELT in Japan and Its Impacts

In modern history, Japanese learners have struggled to master native varieties of English, namely, British English before WWII and American English after the war (Koike, 1993). The prewar Japanese government chose English as “the” foreign language to be taught for the citizens to absorb the knowledge of the advanced West (Koike and Tanaka, 1995). English was taught as the national/ethnic language of the UK, not as the language used for “international communication”. Therefore, unsurprisingly, hypothetical interlocutors in English learners were always British native speakers.

After WWII, the world map of global English users has been innovated as discussed in the section on World Englishes, while the Japanese language education has remained almost the same, except for the fact that the target changed from British to American English. Kubota (2002) argues that Japanese language education “pays insufficient attention to the increased ethnic and linguistic diversity at the local and global levels” (p.28). She points out that, under the name of internationalization accompanied by Japan’s economic expansion in 1980s, Anglicization actually took place. The slogan of foreign language education that “a Japanese citizen should learn at least a foreign language to become a cosmopolitan” was interpreted as that s/he must learn English, for it was the international language. Two different registers of English, that is, English as a tool for international communication and English as the national language of the UK and USA, were confused intentionally or unintentionally. As a result, Outer Circle varieties of English have almost completely been ignored in Japanese ELT (e.g., Kachru, 1997).

Dichotomous and essentialized views of East and West were created and reinforced, and Japanese people were obliged to feel the need for assimilation to the West. As a backlash
to the excessive degrading of Japanese identity, *nihonjinron* (theory of the Japanese, e.g., Ogasawara 1995) prevailed. *Nihonjinron* posited differences between Japanese and Western cultures as absolute, and was utilized to save the Japanese people’s hurt pride. Kubota (2002) argues that *nihonjinron* leads to nationalism. In short, instead of internationalization, Anglicization and nationalism were promoted in Japan during the era, and ELT played a considerably important role for this scenario.

Before “internationalization” became a buzz word, however, Suzuki (1971) had already expressed his concern about the special status of English in Japanese foreign language education. He stated that the attitudes such as “Everyone must speak English as native speakers do” in ELT might sow a seed of implicit ideology among Japanese learners that English native ways are supreme and universal in other aspects of life. Lummis (1976) also pointed out the hidden racism in English education in Japan, and the consequent admiration of the West and disdain of the non-West by Japanese people, which he named the “ideology of *eikaiwa* [English conversation (instruction)]”.

Therefore, the fervor for internationalization seems to have fortified the Japanese inclination for Anglicization, rather than created it. Moreover, English education in Japan did not only fail to correspond to the post-war postcolonial acknowledgement of nonnative English varieties, but significantly helped the West-, or Inner Circle-centered ideology to root in Japanese mentality.

**English as an International Language and Japanese Users**

Baxter (1980) stresses the importance of recognizing the fact that, for the majority of Japanese, English is not a foreign language but a language for international
communication with both native and nonnative speakers. Like Suzuki (1971) and Smith (1976), Baxter (1980) also declares that Japanese need not speak like Americans, while they might want to conform to native English varieties in vocabulary and grammar. At the level of pragmatics, Ellis (1994) concluded that even advanced learners can not reach the NS level. The perspective of World Englishes offers us another interpretation of this same result: NNSs will use English with their own pragmatics, which are different from that of NSs, but this should not necessarily be considered as a deficiency. What is necessary for international communication is that both native and nonnative speakers develop tolerating and accepting attitudes toward mutual differences.

English has become an even more important part of people’s lives in Japan than Baxter (1980) described, with the rise of Japan’s status in global commercialism and industry, and the progress of informational technology. In 1990s, some Japanese companies dispatched a policy stating that employees who failed to obtain high scores in such standardized tests of English as TOEIC or EIKEN would not be promoted, however competent they were in non-English-related criteria. Since then, surprisingly many books and magazines have been published related to English. For instance, _Aete Eigo Kouyougo Ron_ (I dare to insist on making English an official language in Japan) written by Funabashi (2000) stayed among best sellers for months. Funabashi stirred up Japanese people's anxiety about their internationally “low” English proficiency, and asserted that Japan should adopt English as a second or official language, so that Japan can maintain its global competitive power in economics and technology. In 1998, English became a compulsory subject in middle school for the first time in history (Kubota, 2002). It used to be an elective, even though 95% of Japanese middle schools offered only English as a
foreign language. In the spring of 2002, public elementary schools started English education, with great expectations and apprehension among parents. In short, English is definitely no longer just a foreign language in Japan.

In reality, however, as Nakayama (1994) asserted, official English education seems to be failing to reflect the change in purpose of learning English. Though many Japanese people need to learn the language for instrumental purposes, its teaching is still oriented to integrative motivation. Koike et al. (1993) reported that most university professors and high school teachers place their priority on teaching its (i.e., British and American) culture, including literature, rather than developing students’ communicative competence.

Even with the advocacy of Communicative Language Teaching (CLT), which places less emphasis on native-like behavior, a CLT textbook analysis shows that the target population to communicate with is usually NSs, or even when NNS-NNS communication is taken up in a model conversation, NNSs speak perfect native English (e.g., Richards, Proctor, and Hull, 1997). Although Nakayama (1994) advocates the promotion of “valid” Japanese English whose target is not limited to native-speaker models, but to various locally valid Englishes and Honna (1990) emphasized the importance of recognizing English as a lingua franca in Asia, they still represent a minority in English educators in Japan.

The prospect that nonnative English varieties will be acknowledged in Japanese ELT in the near future does not seem to be promising either. The National Curriculum Standard Reform removed the word “international understanding” from the purpose of foreign language education in 1998. Instead, the reform stipulated, “In order for students
to develop practical communicative competence in the target language, great emphasis will be placed on practice in situations where the target language is actually used” (MEXT, 1998). The situations where the target language is used are never specified. With whom the communication is attempted is never described, either. The decision of whether the situations are limited to those with native speakers in the Inner Circle or whether they include those with other nonnative speakers in more multi-national settings seems to be left to interpretation. When the modern history of ELT in Japan is considered, the emphasis on Inner Circle varieties is unlikely to be reduced.

Numerous scholars made suggestions to adjust the excessive admiration for native norms in English education in Japan. For instance, Suzuki (1971) suggested that *Englic*, a more simplified form of English for international communication, instead of British or American English, should be taught in school, since Japanese learners did not want to think nor to behave like English or American. *Englic* is the term Suzuki created to refer to the concept of a language shared with speakers of diverse backgrounds, which is similar to the English as an International Auxiliary Language (EIAL) by Smith (1976). In the same line, Nakayama (1982) advocates “multinationalization,” which provides learners with information on both native and nonnative linguistic and sociocultural characteristics. Okabe (1998) also suggested that ELT curriculums should include many nonnative cultures, not only Western cultures. In short, these scholars suggested that instruction on English/American cultures, including literature, should be reduced from ELT, at least in public education for general purposes. And instead cultures of other nonnative varieties should be introduced.
Kubota (1998) emphasizes that the Japanese learners of English should be equipped not only with communication ability, but also with a critical awareness of the power of English and Western discourse. She claimed that this is plausible if English education curriculum in Japan includes the Outer and Expanding Circle varieties as well as native varieties. She also asserts that ELT professionals need to abandon their apolitical disguise and teach students “critical multiculturalism.” Critical multiculturalism enables learners to see the culture of the target language critically. Standard English can/should be taught, according to Kubota (1999), if it is what students want, but it should be taught only when students can also develop a critical attitude toward the constructed discourse of ELT.

Tsuda (1986, 1998, 1999) has questioned the taken-for-granted hegemony of English in international communication. He asserts that its implicit paradigm has put the NNSs of English at a disadvantage, in comparison with NSs. Oda (1999) reports incidents in a Japanese language education organization that illustrates Tsuda’s concern. For instance, English is always used among the organization members as if it were natural, despite the fact that the organization is located in Japan and the majority of members are Japanese. Moreover, when a discussion grows less amicable, native members sometimes silence nonnative members by maliciously commenting on their nonnative use of English, which are unrelated to the discussion topic. Trifonovitch (1981) also observed the “lingua-centric” attitudes in supposedly international communication scenes. For example, at international conferences whose communication medium is English, some NSs seemed to judge the worthiness of presentations based on the accents of the presenters, not based on their content. Based on the inequality of power between NSs and NNSs in
English-medium communication as seen above, Tsuda (1997) ardently advocates the “ecology of language,” which promotes (1) the Right to Language for those who do not speak English, (2) Equality in Communication, and (3) Multilingualism and Multiculturalism. He encourages non-English speakers to fight the hegemony of English in international communication.

Ishii (1998), however, criticizes Tsuda’s argument as too unrealistic. He said that even if the hidden agenda, the ideological problems, and the inequity of power between NSs and NNSs are understood by the public, the value and prevalence of English will not diminish. Current international communication, which heavily depends on English, would not cease for a second. In the conclusion of his 1999 article, Tsuda, in fact, admits that he is very aware of the status of English as an indispensable *lingua franca*. He wants to challenge the knowledge and attitudes of people who unconditionally accept and welcome the hegemony of English. Thus, arguments made by Kubota (1999) and Tsuda (1999), in fact, accord with those of Strevens (1992) and Cook (1999). The issue of World Englishes that is relevant to the Expanding Circle is more about attitudes toward nonnative varieties rather than standards/norms of English to be taught.

**Summary**

In summary, there are numerous intellectuals who have shown concern over the special status given to English in Japanese education. However, for more three decades after Suzuki raised the question in 1971, similar debates have taken place in regard to the gap between the needs of English users in Japan and how it is taught in school. Whereas the majority of people hope to learn English as a tool for international communication,
ELT in Japan has placed sole emphasis on American and British language varieties and Western cultures. The very fact that the debates have continued for more than 30 years indicates that real-life English instruction in Japan has remained unchanged. Discourses of World Englishes and English as an International Language have potential to change the convention of Japanese ELT.

Attitudes toward Nonnative Speakers

Native Attitudes toward Nonnative Speech

Strevens (1980) predicted the “acceptability” of language variation would be an issue when and where English is used as an internationally common language. Rather than the “acceptability of language,” however, the term “language attitudes” has commonly been used in the field of sociolinguistics.

The investigation of language attitudes goes back as early as 1960, when Lambert, Hodgson, Gardner, and Fillenbaum asked their French and English-speaking Canadian listeners to judge ten speech samples read in French and English, on 6-point scales. The researchers’ expectation was that the listeners would evaluate their affiliated group higher, in other words, French-speaking listeners would rate French speech higher while English-speaking listeners would rate English speech higher. Against the prediction, however, both French and English-speaking listeners rated English speech more favorably. The researchers attributed this result to the minority group reactions of the French-speaking Canadian listeners. This study clearly showed that language was not neutral. People judge other people differently based on the language they speak.
A piece of often-cited research on nonnative speech was reported by Kalin and Rayko (1978). Two hundred and three Canadian college students listened to five nonnative (L1 Italian, Portuguese, Greek, Slovak and West African) and five Canadian speakers, and rated each speaker’s efficacy and reliability as a candidate for four different occupational positions (i.e., foreman, industrial mechanic, production assembler, and plant cleaner). The overall ratings of NNSs were low, but they were rated higher when they were considered to be candidates for the socially less respected, less skill-required job: plant cleaner. Kalin and Rayko (1978) explained their findings by means of the participants’ association of nonnative accents and immigrants’ social status. Judicial cases that have actually arisen in the USA around foreign accents (Sato, 1991; Lippi-Green, 1997) are consistent with the results of the laboratory studies. Foreign accented English has been related to the social and cultural periphery.

Mulac, Hanley, and Prigge (1974) conducted a factor analysis on the fifty semantic differential traits that had been used in language attitudinal studies. They found that the traits were loaded on three different dimensions and named them (1) Socio-intellectual status, (2) Aesthetic quality, and (3) Dynamism. They then chose the most representative 21 items from the stock and developed their own scale: the Speech Dialect Attitudinal Scale (SDAS). The SDAS was administered to 92 American participants. Spontaneous speech made by European-born speakers, as well as Americans, were recorded and used as speech materials. The speakers’ “foreignness” was determined by the judgements of eight Speech and Hearing majors, not by the participant listeners. The researchers found that the more “foreign-accented” a speaker was, the more downgraded s/he was. Neither the sex of speakers and listeners, nor the listeners’
age-occupational status affected the language attitudes. Mulac (1976) developed a shorter (12-item) version of this SDAS and secured high reliability and construct validity. He also concluded that the scale was equally suitable for the use of female and male listeners.

Zhan and Hopper (1985) also reviewed existing semantic differential scales for language variation. They criticized previous studies for failing to completely cover the traits that composed the dimensions, which made comparisons among study results difficult. They examined 121 items and reduced the number to 56 by taking redundancy and sensibility into consideration. These 56 items were administered to 572 American undergraduate students. The factor analysis with oblique rotation produced a new, three-dimensional, 30-item scale: the Speech Evaluation Instrument (SEI). As seen below, the SEI has been widely used to assess attitudes toward nonnative speech, but Zhan and Hopper (1985) originally calculated its psychometrics based on the attitudes toward various dialects of American English (i.e., black Southern, white Appalachian, and midwestern).

Varonis and Gass (1982) conducted several experiments to address native reactions to nonnative speech. In the first, they examined how NSs reacted when talking with strangers with and without foreign accents. The researcher found that the NSs switched to so-called “foreigner talk” once they recognized that they were addressed by nonnative speakers. They also found NSs reluctant to be involved in a conversation with supposed foreigners. In the second experiment, Varonis and Gass (1982) investigated the influence of the strength of foreign accents and the degree of grammatical accuracy on the native listeners’ evaluation of speakers’ pronunciation. The speakers with good pronunciation were rated “good” regardless of grammatical accuracy, while speakers with
strong accents were judged worse when they read ungrammatical sentences than when they read grammatical ones. The researchers attributed the result to the “comprehensibility” of the speech. If the speech has a strong accent and is ungrammatical, it becomes very difficult to understand, and the listeners’ evaluations will be significantly lower as a result.

Giles, Williams, Mackie, and Rosselli (1995) used the SEI (Zhan and Hopper, 1985) as one of their instruments. The researchers carefully trained their bilingual speaker in order to carry out a matched-guise study. The same speaker read two opposing opinions about the English Only Movement (EoM [sic]) with two different accents (American and Spanish). Listeners were asked to rate their mood, their own position on the EoM issue, and the comprehensibility of the speech samples. They were also requested to rate their own patriotic identity. As a result, Anglo listeners turned out to feel happiness and strong national identity when they heard an ethnically similar (i.e., Anglo) sounding speaker arguing against EoM. The listeners were less happy when they heard the same argument with a Spanish accent. Like its predecessors, this study found a tendency to downgrade Hispanic-accented English on the part of Anglo American listeners, but was unique in demonstrating strong influence of the topic on listeners’ attitudes.

Cargile and Giles (1998) conducted a similar study to investigate the attitudes of non-Asian American college students toward Japanese-accented English. The researchers hypothesized that the speech content to be a primary independent variable for predicting listeners’ evaluative reactions. The other variable was the degree of accent with four levels: standard American English (i.e., no Japanese accent), moderate Japanese accent, strong Japanese accent, and strong Japanese accent with lexical and structural inaccuracy.
Cargile and Giles (1998) found both independent variables had significant main effects on listeners’ scores on the SEI (Zahn and Hopper, 1985). As for the speech content, aggressive speech was rated more negatively than its non-aggressive counterpart. The more accented the speech, the less favorably it was judged. However, to the researchers’ surprise, a moderate Japanese accent was judged as well-respected as a standard American accent. Moreover, the listeners indicated no difference in rating a strong Japanese accent and a strong Japanese accent with lexical and grammatical inaccuracy. The researchers attributed these findings to the stereotypes that Americans hold for Japanese, for example, “intelligent” though not so “attractive,” and to the possibility that the strength of accent may have “overridden” other variables which might have impacted on judgements otherwise.

Cargile (1997) investigated how the sense of affiliation influences evaluative reactions to a non-standard variety of English. Ninety-seven (58 Anglo-, 39 Asian-American) college students participated in the study. They listened to fictional job interviewees with Standard American- and Chinese-accents and evaluated them on the modified SEI (Zhan & Hopper, 1985). Unlike the findings by Kalin and Rayko (1978), listeners judged the Chinese-accented English speaker favorable as a job applicant as the American English speaker. (In fact, one speaker performed both guises.) The overall reactions of Asian American students to the accented speech were more critical than those of their Anglo-American counterparts. Asian Americans evaluated a Standard American English speaker more positively than a Chinese-accented speaker in the dimension of attractiveness as a potential classmate. The conclusion that Cargile (1997) drew from the study is that the contexts need to be considered in language attitudinal studies.
Delamere (1996) explored native attitudes toward several different English varieties. She used the matched-guise technique with five speakers who were bilingual in English and one of the following languages: Farsi (Iran), Arabic, French, Malay, and Spanish. Each speaker read two types of texts: one with grammatical errors and the other without. Thirty NSs in Florida listened and judged the speakers’ performance on a semantic differential scale. The Farsi and Arabic speakers were evaluated higher when their speech did not contain errors than when their speech did include errors. In contrast, the French and Malay speakers were rated higher with errors than without errors. The Spanish speaker was rated low, whether his speech contained errors or not. Delamere (1996) explained these results by the Floridan listeners’ familiarity with, and stereotypes of, the nonnative accents utilized in her study. A pedagogical implication drawn from the findings was that the necessity of grammar correction in the English language classroom might depend on the learners’ L1s.

One tendency commonly observed among the language attitude studies cited above was the depreciation of nonnative language varieties. As Edwards (1999) puts it, these findings might represent the stereotypes that people hold toward “immigrants” and “foreigners,” if not xenophobia. Edwards (1999) also argues that listeners may evaluate a speaker’s linguistic and metalinguistic attributes that reflect the group to which the speaker is likely to belong. Trifonovitch (1981) suspects that “while English is an international language, unfortunately, most of its speakers are not international persons” (p. 211). Kachru (1981) also stated, “… Such (negative) attitudes are not essentially based on linguistic value judgement but various other factors play an important role, one being a native speaker’s fear of seeing his language disintegrated in the hands of (or shall we say,
on the lips of) nonnative users” (p.34). Whether from stereotypes about social minority
groups, or ignorance of NNSs’ co-ownership of English or fear of violation of their own
property (= English), it is apparent that many English native participants in preceding
studies did not consider English as an international language. They were likely to have
judged the native-like-ness of the nonnative speakers they heard in the experiments.

Negative attitudes toward nonnative language varieties is, however, not unique to
English-speaking societies, but rather, universal (e.g., see Cummingham-Anderson, 1997).
The majority of studies on native reactions to nonnative speech conducted in other
languages (e.g., Politzer 1978; Galloway, 1980; Chastain, 1981; Ensz, 1982; Gynan,
1985; and Schairer, 1992) were intended to identify the nonnative errors that irritated
native listeners the most. The researchers, thus, attempted to determine error gravity; the
most serious errors were the ones that irritated native listeners most. The assumption that
learners should avoid making errors that irritate native speakers has been taken for
granted in foreign language education. The question is whether this assumption should be
applied to English. It depends on how English is recognized: as a national/ethnic language,
or as an international language.

The extreme difficulty of drawing real responses from listeners has been pointed
out by Eisenstein (1983) in the research on language attitudes toward foreign accents.
Since no one desires to be considered prejudiced or xenophobic, listeners tend to give
socially desirable responses, not only on direct questionnaires but also in more disguised
experimental settings. Real native attitudes toward nonnative speech might be far more
negative than they appear in research studies.
Nonnative Attitudes toward Other Nonnative Speakers

In comparison with the number of inquiries investigating native attitudes toward nonnative speech, far fewer studies have been conducted appointing nonnative listeners as judges. This research tendency has significant implications for the power relationship between NSs and NNSs. On the one hand, theoretically, English is considered as an international language by eminent scholars such as Kachru (1981), Trifonovitch (1981), and Widdowson (1982). On the other hand, English is still treated in reality as a national/ethnic language. That is, all the rights to judge nonnative language right or wrong, good or bad, and acceptable or unacceptable are reserved for “native speakers’ intuition”. Kramsch (1993) mentioned that NNSs are frequently awed and intimidated by NS norms and NS judgements.

Ludwig (1982), in the course of scrutinizing existing attitudinal studies, drew attention to the difference between the judgements by native lay persons and nonnative teachers. However, in the view of English as a national/ethnic language, the nonnative listener judgements were to be corrected. Ludwig (1982) concluded that nonnative language teachers should approximate their judgements to those of native speakers.

If one believes in English as an international language, in which NSs and NNSs are equal participants in communication, however, then one has to attend to nonnative listeners’ attitudes toward other nonnative English varieties. The following are some examples of studies in which nonnative listeners judged nonnative speech.

Fayer and Krasinski (1987) were interested in both perceived intelligibility/comprehensibility and language attitudes of Puerto Rican ESL learners toward Spanish accented English. Eighty eight Puerto Rican and 40 native listeners
listened to seven L2 English speech samples produced by L1 Spanish-speaking intermediate learners with 3 levels of proficiency: high-, mid-, and low-intermediate. The Puerto Rican listeners were less tolerant toward the nonnative accent than the native listeners were. Among the factors examined, pronunciation and hesitations were reported to be the leading irritating factors.

Matsuura, Chiba, and Yamamoto (1992) conducted a study on the language attitudes of Japanese listeners toward nonnative accented English. The speakers were six international students studying in Japan, with the following nationalities: Malaysia, Chinese Malaysia, Bangladesh, Micronesia, Hong Kong, and Sri Lanka, and one NS control from the USA. They all read the identical passage. Ninety two Japanese college students listened to the tape and completed the task of semantic differentials on their attitudes toward the varieties. The result showed that the native variety was perceived more favorably than the nonnative varieties. This tendency was not related to the listeners’ proficiency level. In this study, the listeners’ motivation for learning English was also investigated. Those who studied English for integrated purposes (i.e., who were drawn to America and American cultures) tended to downgrade non-American English varieties more harshly, in comparison with learners with instrumental motivation (i.e., who learned English as a communication tool).

Chiba, Matsuura, and Yamamoto (1995) were motivated to further examine the relationship among the nonnative listeners’ attitudes toward nonnative accents, the listeners’ motivation to learn English, and the listeners’ familiarity with the accents. One hundred sixty nine Japanese college students studying English and international business in Japan listened to short paragraphs read by NSs and NNSs with three different accents
(i.e., 3 Japanese, 3 NSs, and 3 NNSs (from Sri Lanka, Hong Kong, and Malaysia). Listeners were to rate the speech samples on a semantic differential scale, and to answer the Attitude Scale Questionnaire, adapted from the SEI (Zhan and Hopper, 1985). The researchers found the following tendencies: (1) NS varieties were ranked higher than Japanese English and other nonnative varieties, (2) instrumentally motivated listeners were less harsh toward nonnative accents than those with integrative motivation, and (3) Japanese accents were perceived more favorably than other nonnative speech. From this last finding, Chiba and her colleagues concluded that familiarity does affect the language attitude of nonnative listeners toward nonnative varieties, but familiarity in itself is not the determining factor of attitudes. In order to change the negative attitudes toward nonnative Englishes, the researchers suggested, “advocating … the rationality of ideology behind the development of world Englishes appears to be required” (p. 85).

Dalton-Puffer, Kaltenboeck, and Smit (1997) examined 132 Austrian university students’ attitudes toward English varieties (British native, British near-native, American-native and Austrian-accented). Native varieties were again preferred to nonnative varieties. Like Chiba et al. (1995), Dalton-Puffer et al. (1997) found that the Austrian participants judged their own accent negatively. Between native varieties, British accent was rated more favorably. The researchers related this result to the status of British English as the model in Austrian ELT. Another significant finding in this study was that the more exposed to other cultures, the more accepting that person was to nonnative English varieties. The participants who had spent a substantial amount of time abroad were more willing when asked whether they would like to make friends with the accented speakers.
The review of literature on nonnative listeners’ attitudes toward other nonnative speakers’ speech seems to indicate that NNSs have perceptions approximately similar to NSs. They tend to downgrade nonnative accents, as much as or even harsher than NSs do. Does this mean that a certain language variety has the innate quality of being belittled universally?

One study offers an insight. Alford and Strother (1990) examined nonnative attitudes toward native English regional varieties. With the matched-guise technique, one male speaker and one female speaker read an identical passage with three different American regional accents (Northern, Southern, and Midwestern). Sixty six international graduate and undergraduate students and 31 domestic students at an American college listened to the six speeches, believing that they were listening to six different speakers, and rated them on a scale with 24 bipolar adjectives. The main findings were: (1) NNSs were able to distinguish regional accents. In other words, they reacted to different accents differently, and (2) NNS ratings showed different patterns from those of NSs, though both groups rated the northern accent lowest. The results of this study seem to suggest that negative attitudes toward certain accents are not immanent in accents nor in raters but are acquired as newcomers to a community gradually learn local cultures, including stereotypes.

Dalton-Puffer et al. (1997) inferred that the rather uniform tendency in studies of nonnative attitudes toward nonnative speech is explicable, when one considers the current ELT situations in which learners are constantly encouraged to appropriate native norms, including grammar and pronunciation. Trifonovitch (1981) also asserted that such NNSs have falsely identified with NSs rather than other NNSs and have taken on native
speakers’ derogative or condescending attitudes toward NNSs. If Trifonovitch is right, qualitative inquiry might be more suitable to address this question in order to delve into what is going on in people’s minds.

In summary, evaluative reactions to nonnative speech have largely been conducted with matched-guise technique, which enable researchers to concentrate on listeners’ reactions to nonnative accents. The studies under review indicate that nonnative accents are more downgraded than native accents and that nonnative listeners are more critical of nonnative accents than native raters are. However, not all nonnative accents are rated equally, that is, some are more downgraded than other nonnative accents. The difference seems to be strongly related to the stereotypes that are attached to certain groups in a given community.

Empirical Studies on the Intelligibility of Nonnative Speech

Studies on the intelligibility of nonnative speech existed long before the issue of World Englishes arose. Abercrombie (1949), in his recommendation for teaching pronunciation to foreigners/immigrants, maintained that the acquisition of native-like pronunciation for adult learners is very difficult, if not impossible, and thus, teachers should not expect “perfection.” He continued that teachers should rather select specific phonemic features to teach and those features should be selected based on “intelligibility.” He also asserted that research is necessary to determine which pronunciation features most seriously hinder the intelligibility of nonnative speech. Therefore, even before World Englishes, research on “error gravity” or “error hierarchy”
of nonnative language drew substantial attention among second language acquisition researchers, for mainly pedagogical reasons.

Intelligibility and Comprehensibility

The terms “intelligibility” and “comprehensibility” have been alternatively utilized in linguistics research studies (Nelson, 1982; Matsuura et al., 1999). The definitions of these words provided in the American Heritage Dictionary of the English Language (Picket, et al., 2000) follow,

intelligible: capable of being understood, and
comprehensible: readily comprehended or understood; intelligible.

“Intelligibility” and “comprehensibility” are treated as synonyms.

In order to avoid confusion, more explicated definitions by Smith (1992) are adopted in this study. According to Smith (1992), language understanding consists of the following three phases: (1) *intelligibility*: to recognize words/utterances; (2) *comprehensibility*: to seize the meaning of words/utterances (locutionary force); and (3) *interpretability*: to reason the meaning behind word/utterance (illocutionary force) by analogy.

According to the categorization by Smith (1992), studies such as Gass and Varonis (1984), Smith and Rafigzad (1979), Smith (1992), Munro and Derwing (1995), Derwing and Munro (1997), and Scott (1999), can be said to have investigated the intelligibility of nonnative varieties of English. Whole or partial sentence transcription was adopted as the measurement of intelligibility in these studies.

Varonis and Gass (1982), Smith and Basazza (1982), Smith (1992), Fayer and Krasinski (1987), and Derwing and Munro (1997) conducted research on
comprehensibility of nonnative speech. Listeners in these studies were required either to simply describe the perceived comprehensibility of speech on a scale, or to answer multiple-type content questions based on what they heard.

On interpretability, fewer research studies have been conducted, with some examples from Smith (1992) and Van der Walt (2000). In these studies, the listeners were asked to paraphrase, in their own words, what they heard.

In reality, however, these three phases might be difficult to clearly divide. As is commonly accepted in psycholinguistics research on reading, comprehension is not a linear process moving from decoding (i.e., intelligibility) to meaning-making (i.e., interpretability) (Goodman, 1968). In the interactive model that is widely accepted (e.g., Rumelhart, 1977), old information (schemata: already obtained world knowledge) and new information (intake through sensory organs) interact in a person’s mind and synthesize understanding. Therefore, all three phases of comprehension need to be briefly reviewed in the following section.

Studies on Intelligibility of Nonnative Speech to Native Listeners

On read materials.

In order to examine the genuine effect of foreign accents on intelligibility/comprehensibility, many researchers have adopted “reading”, as opposed to extemporaneous speech. By adopting the “reading” method, researchers can control the vocabulary, grammar, and content of the speech. The following are such studies that utilized “read” materials.
In the study by Lane (1963), foreign accent is treated as speech distortion, as well as masking and filtering. Four speakers with different L1 backgrounds (American, Serbian, Punjabi, and Japanese) read a list of words. 24 American university students listened and wrote down the words. The researcher found that nonnative speech was about 36% less intelligible than native speech in four different levels of masking and filtering. Two types of speech distortion (i.e., foreign accent vs. filtering and masking) do not interact.

Gass and Varonis (1984) tried to determine the effect of listeners’ familiarity with the topic and the variety of English, and the listeners’ scores on a dictation test. The results indicated that all the kinds of familiarity investigated affected the intelligibility scores of 142 American university students, while familiarity with the topic was the most powerful factor.

Anderson-Hsieh and Koehler (1988) were interested in the effect of speech rate on intelligibility. When compared with a NS, their L1 Chinese speakers were judged less intelligible when they read the given material at a faster rate. This tendency was stronger with a heavily accented speaker than a mildly accented one.

Munro and Derwing (1995a) reached a similar conclusion. In their study, however, instead of asking listeners to judge intelligibility, the actual processing time was measured. The Canadian NS listeners needed more processing time when they had to verify Mandarin-accented statements than they did when verifying Canadian- and American-accented speech. The researchers concluded that foreign accents require longer processing time. The pedagogical implication of these two studies is that NNSs, especially with strong accents, should speak more slowly to be understood.
In Anderson-Hsieh and Koehler (1988) cited above, the listeners reported that their comprehension was hindered more severely by prosodic than by segmental features. Anderson-Hsieh, Johnson, and Koehler (1992) examined the correlation between the ESL instructors’ holistic assessment and the linguistic analysis of each speech sample. They found the experts’ ratings were explained up to 89% by prosody, segmental features, and syllable structure, of which prosody was the most important.

With artificially filtered speech that is completely unintelligible, Munro (1995) found that native listeners still could detect foreign accents based only on non-segmental factors (i.e., stress, intonation, rhythm, speech rate, and voice quality). In a word, prosody, or non-segmental speech features, seems to play a significant role in listeners’ perceptions on both the comprehensibility and accentedness of heard speech.

Scott (1999) attempted to find the relationship between the intelligibility of L1 Brazilian Portuguese speakers and three independent variables: levels of foreign accents, noise, and contextual predictability of the speech. In Scott’s study, noise and predictability turned out to affect intelligibility, but the degree of foreign accents of the speakers proved to have the greatest influence on intelligibility.

Van der Walt (2000) was concerned with the international comprehensibility of South African Englishes. Since she wanted to examine speech samples as authentic as possible, she extracted meaningful chunks of discourse from TV and radio programs broadcast in South Africa. One hundred forty NSs and advanced NNSs rated these samples for comprehensibility and paraphrased what they had understood. It was found that South African varieties of English were highly comprehensible to both native and nonnative listeners.
On extemporaneous speech.

Studies using extemporaneous speech samples are rather few. Albrechtsen, Henriksen, and Færch (1980) extracted six speech samples of L1 Danish speakers with three levels of error density. 300 native listeners from different home regions, with various occupations, sexes, and ages listened to the speech samples, answered two comprehension questions, and evaluated the samples on a 5-point semantic differential scale. The subjective judgement was compared with an objective linguistic analysis of the speech samples. The findings of this study were: (1) listeners’ age seemed to affect the evaluations (younger listeners were more critical), and (2) the frequency of the use of communication strategies, such as literal translation from L1 and self-correction, negatively correlated with comprehensibility ratings. In other words, the more communication strategies were used, the lower the comprehensibility was judged. However, the use of communication strategies may not have induced lower comprehensibility. A speaker who used fewer communication strategies tended to make fewer grammatical mistakes in the first place, and also tended to have better pronunciation. Thus, this study revealed the difficulty of using extemporaneous speech, because the researchers could not identify the factors that caused incomprehensibility.

Munro and Derwing (1995b) examined linguistically trained NSs’ (whose majors were either linguistics or TESOL) perception of nonnative speech produced by advanced L1 Mandarin speakers. All the speech samples turned out to be highly intelligible, and the comprehensibility ratings were also high. However, the perceived accentedness ratings varied more widely. The researchers concluded that even though perceived accentedness is related to intelligibility and perceived comprehensibility, they are not the same. That is,
it is very possible that speech with a heavy accent is highly intelligible and comprehensible.

In 1997, Derwing and Munro chose intermediate ESL students from four different L1 backgrounds (Cantonese, Japanese, Polish, and Spanish) as speakers and linguistically untrained NS college students as listeners. The listeners were to transcribe each speaker’s extracted short phrases word-for-word, rate them both on comprehensibility and accentedness, and identify the speaker’s L1. The main findings were: (1) intelligibility test scores and comprehension ratings were highly related, (2) accentedness ratings were not related to intelligibility nor comprehensibility, and (3) the accuracy of identifying the speakers’ L1s depended on the listeners’ familiarity with the variety of L2 English.

Studies on Intelligibility of Nonnative Speech to Nonnative Listeners

The listeners or evaluators or judges of nonnative speech have customarily been native speakers. As was mentioned in Chapter 1, reasons why only NS judgements were used in the evaluation of nonnative speech are understandable. First of all, the belief that only NSs possess the ability (and the right) to distinguish what is comprehensible from what is not has been prevalent in Second Language Acquisition (SLA) research (e.g., Ellis, 1994). A Swedish scholar Olsson (1977) clearly stated, “the yardstick of intelligibility should be a verdict by native Englishmen and not an armchair speculation by a nonnative.” (p. 2-3) The ownership of English was hardly questioned until the concept of World Englishes entered the field. The absolute authority of NSs is now more
actively challenged probably under the influence of the interdisciplinary popularity of post-modernism in the United States (e.g., Pennycook 1995).

Secondly, as Smith (1992) pointed out, two common beliefs were tenacious in SLA: (1) If a listener understands a native variety, then s/he would easily understand nonnative varieties as well, and (2) if a speaker is understandable to native-listeners, then s/he would also be understandable to nonnative listeners. These assumptions are, however, clearly negated by Smith (1992).

Thirdly, in terms of a research design, including nonnative listeners as well as native-listeners in a study would increase the number of independent variables. Having two different groups (NSs and NNSs) also requires a larger sample size, which is a burden to many researchers. Nonnative listeners are also likely to be more heterogeneous (in their L1s naturally, and in numerous related attributes) than native listeners. Especially in research on language attitudes, individual differences have been reported to be large even among comparatively homogeneous native listeners. Controlling extraneous variables across native and nonnative listeners would conceivably be more challenging. As a result, far fewer research studies adopted nonnative raters. The following section reviews them.

Bansal (1966) examined how intelligible Indian English was to a total of 178 NS listeners from the UK and USA, and NNS equivalents from Germany, Nigeria, and India. She used extemporaneous speech taken from 24 Indians with different L1 backgrounds and 4 Received Pronunciation (RP) users. Listeners were required either to orally repeat or to write down what they had heard. The main findings of this study were: (1) both NS and NNS listeners (including Indians) reproduced RP more accurately than Indian English,
Indian listeners’ L1 backgrounds did not matter, and listeners from the Expanding Circle (i.e., German) seemed to have difficulty remembering what they had heard in the foreign language, probably because of their low English proficiency.

Wilcox (1978) hoped to examine which accent of English was the most comprehensible for 320 Singaporean college student listeners: American, British, Australian, or Singaporean. The results were mixed, but the researcher found general tendencies. The Singaporean listeners understood Singaporean-accented English the best, with British the second, followed by American and Australian English. Wilcox attributed the result to the familiarity of each accent to the listeners, but also mentioned that the speech rate might have affected the result. It turned out that the Singaporean speaker read the materials significantly slower while the American speaker read faster than the others.

Eisenstein and Berkowitz (1981) tried to empirically confirm their qualitative findings on the NNSs’ perceived comprehension of NS and NNS varieties of English. In their antecedent study, NNS participants, especially when their proficiency was at the beginning level, had claimed that NNS English was easier to understand than NS English. In the 1981 experiment, the researchers requested 58 adult ESL students with diverse L1 backgrounds to listen to Standard American, New York, and Philippine-accented English and to rate the comprehensibility of each accent. Contrary to the previous qualitative findings, the participants rated Standard American English as easier to understand than New York- and Philippine (Tagalog)-accented English.

In the aforementioned study, Fayer and Krasinski (1987) examined perceived comprehensibility of Puerto Rican English to Puerto Rican ESL learners and NSs of English. The results turned out to be: (1) Puerto Rican listeners understood some, but not
all, Puerto Rican speakers better than the NS listeners did, and (2) the proficiency levels of the speakers had similar effects on both NS and NNS listeners. Speakers with lower proficiency were judged more difficult to understand by both Puerto Rican and native listeners.

Smith and Rafigzad (1979) used speech samples produced by 9 speakers of educated Englishes, from Hong Kong, India, Japan, Korea, Malaysia, Nepal, the Philippines, Sri Lanka, and the USA. Speech samples were extracted from hypothetical extemporaneous lectures addressing educated fellow countrymen of the speaker. Speakers first wrote and then read the lecture scripts. 1,386 people educated in different disciplines from 11 countries listened to the samples, completed a cloze test and answered a Listening Comprehension Questionnaire. One finding was that neither the speech sample taken from a NS nor a fellow countryman was the most intelligible to everyone. A nonnative variety of English that was easy for NSs to understand turned out to be easy for NNSs as well. Though meaningful in terms of advocating the international intelligibility of educated L2 speakers, this study included two potential extraneous variables. (1) As the researchers themselves admitted, grammatical complexity was not controlled. Therefore, the structure of the speech made by the Japanese speaker, for instance, turned out to be indeed simpler than the less intelligible speech samples (e.g., by the NS speaker). (2) The speech was first written and then read aloud. The speakers might have paid more attention to their pronunciation than they would have in truly extemporaneous speech production.

Maybe because Smith was aware of the grammatical complexity differences among speech samples in the 1979 study, Smith and Bisazza (1982) controlled
vocabulary and grammar in their study. An Indian, a Japanese, and an American graduate student at the University of Hawaii, majoring in either ESL or linguistics, read the same materials: three sections of the Michigan Test of Aural Comprehension (MTAC). 207 college students from the Outer Circle (Hong Kong, India, and the Philippines), The Expanding Circle (Japan, Taiwan, and Thailand), and the Inner Circle (Hawaii) listened to the tape and completed three tasks. The tasks were: (1) to answer the MTAC questions as ordinary examinees; (2) to identify the speakers’ nationalities; and (3) to rate the comprehension difficulty with (a) speakers (Indian, Japanese, and American), and (b) test forms (A, B, and C). Results indicated that most listeners (except for Indians) found American English the easiest to understand and Japanese English second easiest. American people were not as good at judging their peer American’s identity as Filipinos were. Japanese listeners also failed to identify their fellow countryman. The researchers attributed the results to the listeners’ greater familiarity with American and Japanese accents, compared with Indian accents.

Smith (1992) then used extemporaneous speech again. The speakers were fluent postgraduate and graduate students at the University of Hawaii with nine national varieties of English (China, Taiwan, India, the Philippines, Japan, UK, Papua New Guinea, Indonesia, and USA). Five conversations between two speakers were recorded. The listeners consisted of three groups: (1) NNS: 10 Japanese college students, (2) NS: 10 American undergraduate students in Hawaii, and (3) Mixed: One NS and eight NNS linguistically trained graduate students at the University of Hawaii. The listening tasks were (1) 7th word cloze test, (2) 3 multiple-choice comprehension questions, and (3) three-phrase paraphrasing. The findings were: (1) the Japanese NNS group did the poorest
on all three tests; (2) the mixed group surpassed NS on the interpretability test, which the researcher claimed was the most important for communication; and (3) familiarity with the varieties affected the perceived comprehension ratings, but not the intelligibility test. Smith (1992) emphasized the fact that many NNSs were judged to speak Standard English, which supported the statement that one can speak Standard English even with foreign accents. The NS varieties were not the easiest to understand, nor were NS listeners the best at understanding other varieties of English.

Matsuura, Chiba, and Fujieda (1999) examined the intelligibility and comprehensibility of native varieties of English (American and Irish) to 106 Japanese college students. Three Irish English speakers and three American English speakers recorded extemporaneous self-introductions. The listeners were asked to (1) dictate 10 researcher-selected words from each passage, (2) rate the comprehensibility of each speech on a 7-point Likert-type scale, and (3) answer the researcher-made multiple choice comprehension check questions.

The researchers found that (1) the Japanese listeners’ intelligibility test scores did not correlate with their comprehensibility ratings, but did correlate with their proficiency levels, (2) the degree of listeners’ exposure to spoken English was not related to their comprehension check test scores, but it was related to the self-reported comprehensibility ratings, and (3) the listeners’ familiarity with a specific variety of English did not affect the intelligibility scores nor comprehension check test scores, but it did affect their self-reported comprehensibility ratings. In short, the Japanese listeners in this study seem to have judged the comprehensibility of native speech according to their familiarity with
spoken English and with the variety of English, not necessarily by how well they actually understood the speech.

More recently, Major, Fitzmaurice, Bunta, and Balasubramanian (2002) investigated the influence of the listeners’ first language on their comprehension of native and nonnative varieties of English. Their primary research question was “to what extent do native and nonnative listeners perform better on a listening comprehension test, when the material is read by a speaker who shares their native language?” A total of 400 listeners from four different L1 backgrounds (Chinese, Japanese, Spanish and standard American English) listened to eight speakers with the same L1 backgrounds reading lectures adopted from the TOEFL Listening Comprehension Trial Test, and answered multiple-choice comprehension questions.

The results were mixed when it came to the influence of the listeners’ L1. The main results were (1) the native listeners understood all the varieties better than the nonnative listeners, and (2) only the Spanish-speaking listeners comprehended their fellow countrymen’s English significantly better than others, including American English. The listeners’ familiarity with an accent failed to explain the Japanese listeners’ performance in this study. The Japanese listeners did not comprehend Japanese-accented English as well as Chinese-and Spanish-accented English, though all the other listeners (namely, Spanish, Chinese, and American English-speaking listeners) understood Japanese English better than Chinese-accented English. The researchers tentatively attributed this result to the Japanese listeners’ negative attitudes toward their own accent. Unfortunately, however, Japanese listeners’ attitudes toward the Japanese accent were not measured in the study.
In sum, the review of nonnative speech intelligibility research has indicated mixed results. Generally speaking, (1) nonnative speech is less intelligible (or requires more processing time) than its native counterpart, (2) the stronger the accent, the less intelligible a speaker becomes, (3) listeners’ familiarity with the topic and the variety of English influence the intelligibility, and (4) nonnative listeners do not necessarily understand their compatriots’ English better than that of other nonnative speakers. However, as seen above, several counterexamples have been found to these generalizations.

Even though the intelligibility of nonnative speech is lower than that of native speech, it is not very low in itself, in comparison with perceived comprehensibility. That is, listeners seem to be able to understand nonnative speech better than they say they can, while nonnative listeners tend to claim that they understand native English better than they actually do.

Conclusion of the Chapter

Many studies have been conducted on the intelligibility of and the attitudes towards nonnative speech for the following reasons. The ultimate purpose of language is communication based on mutual intelligibility. And no verbal communication is free from evaluative reactions to interlocutors’ linguistic and metalinguistic characteristics. As seen above, the majority of these studies adopted NSs as raters of NNSs.

On the other hand, nonnative English has been promoted as new English varieties, at least in theory, since the global spread of English created many localized
norms. Historically, the discourse of World Englishes has developed with the Outer Circle as its focus. Recently, however, a movement has occurred to recognize the Expanding Circle as a critical component of English users for international communication.

Japan belongs to the Expanding Circle, and Japanese English has quite intensively been studied as “learner language” that should be corrected based on the model language (i.e., Inner Circle varieties) norms. However, the purposes for Japanese people to learn English have greatly changed for the last half-century, and the Inner Circle centered ELT has seriously influenced the socio-cultural aspects of Japanese lives. The investigation of Japanese English as an international language variety is an imminent need to promote true internationalization of English.
CHAPTER 3

PROCEDURES

The research design of this study is a combined qualitative and quantitative design (Creswell, 1994). More precisely, this study has adopted the dominant-less dominant design, in which the quantitative paradigm provides its main framework while one small component is drawn from the qualitative paradigm. In other words, this study has combined an ex post facto study and a follow-up qualitative interview study for triangulation and complementarity. This chapter explains the procedures of the quantitative phase first and then of the qualitative phase.

Quantitative Inquiry

The purpose of ex post facto research is to explain and predict a phenomenon (Miller, 1998). Two phenomena are under investigation in this study. One is the evaluative reactions to Japanese English across listeners from four different L1 backgrounds. The other is the intelligibility of the nonnative speech to these listeners. Speech samples were taken from the speakers who varied in English proficiency levels.
The variables to be explained and predicted were the listeners’ scores on the Evaluative Reactions scale and the Intelligibility test, while the listeners’ L1 backgrounds and the speaker’s L2 proficiency were presumed as predicting factors. An *ex post facto* study was chosen because one of the main independent variables, listeners’ L1, was an attribute, and, therefore, could not be manipulated.

Based on the review of related literature, three rival independent variables were chosen: (1) listeners’ gender, (2) listeners’ familiarity with Japanese English, and (3) listeners’ exposure to foreign (other than Japanese) speakers of English. Other possible extraneous variables such as listeners’ age and educational level, and the content of the speech were controlled by making them constant as much as possible. The listeners’ characteristics chosen as rival variables were assessed by means of the Personal Information Questionnaire, administered after the Evaluative Reactions scale and the Intelligibility test.

The study is intended to explain/predict the variability of the dependent variable by using information of two independent variables when the predictive effects of rival independent variables were removed. The model of the relationship among the dependent variables and the main and rival independent variables is displayed in Figure 3.1.
The Dependent Variable: (cognitive)
✓ Intelligibility of Japanese English

The Dependent Variable: (affective)
✓ Listeners’ evaluative reactions to Japanese English

The Main Independent Variables
✓ Listeners’ L1
✓ Speakers’ proficiency level

The Rival Independent Variables
✓ Listeners’ gender
✓ Listeners’ familiarity with JE
✓ Listener’s exposure to foreigners

Figure 3.1: Conceptual model of the *ex post facto* study
The L1 comparison is illustrated in Table 3.1. Eighty listeners with four different L1 backgrounds (i.e., American English, Mandarin, Hindi, and Japanese) were going to listen to eight speakers with four different English proficiency levels (i.e., native, advanced, intermediate and beginning-level). Each proficiency level consisted of one male and one female speaker.

<table>
<thead>
<tr>
<th>LISTENERS</th>
<th>AMR (NS)</th>
<th>JAPANESE (NNS)</th>
<th>N=8</th>
</tr>
</thead>
<tbody>
<tr>
<td>LNS</td>
<td>AMR n = 20</td>
<td>♂</td>
<td>♂</td>
</tr>
<tr>
<td></td>
<td>MND n = 20</td>
<td>♂</td>
<td>♂</td>
</tr>
<tr>
<td></td>
<td>HND n = 20</td>
<td>♂</td>
<td>♂</td>
</tr>
<tr>
<td></td>
<td>JPN n = 20</td>
<td>♂</td>
<td>♂</td>
</tr>
<tr>
<td>N=80</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: AMR = American, MND = Mandarin, HND = Hindi, JPN = Japanese, ADV = advanced, INTER = intermediate, and BGN = beginning-level proficiency.

Table 3.1: Comparison among listeners with various L1 backgrounds and speakers with different English proficiency levels
Research Site

The participants were individually recruited on a voluntary base, among graduate and undergraduate students at a large public university in the Midwest of the United States. This research site was chosen because of the large population of international students with diverse L1s. Students in this university can be considered as international English users, since they were supposed to interact with non-Americans at least to a certain extent, due to the number of international students as well as of international faculty. Approximately 35,000 full-time undergraduate students and 12,000 graduate/professional students were enrolled in the university at the time. Around 4% (1,880) of student population was international, coming from 89 countries. Unfortunately, probability sampling seemed to be unpractical, for the participants had to be in a specific place in a specific time and the administration took approximately one hour.

Subject Selection

Determination of Sample Size

The ultimate goal of this study was to examine whether and to what extent the knowledge of listeners’ L1s helps predict their evaluative reactions to, and the intelligibility of, Japanese English. Whether and to what extent the speakers’ proficiency levels influence the listeners’ reactions and comprehension were also investigated. Multiple regression analysis was chosen for the data analysis, because it “predicts the changes in the dependent variable in response to changes in the independent variables” (Hair, Anderson, Tatham and Black, 1998, p.14).
The sample size that was necessary for conducting a multiple regression analysis was calculated based on the suggestions by Hair et al. (1998), from two standpoints. First, 15 to 20 observations per independent variable are desirable to secure the generalizability of results. The study design with five independent variables, including main and rival, required at least 75 subjects. Secondly, the statistical power must be considered. Statistical power (1- β: Type II error) is the probability of correctly rejecting the null hypothesis when it should be rejected, but is inversely related with α (Type I error): the probability of incorrectly rejecting the null-hypothesis. According to Hair et al. (1998, p165), the sample size of 50 with five independent variables can detect $R^2$ (coefficient of determination) = 23%, with a power of .80 at a significance level ($\alpha$) = .05. The sample size of 100 will detect $R^2 = 12\%$ with other conditions remaining the same. Considering these points, the sample size of 80 was chosen in this study, with 20 observations in four L1 groups (levels) of an independent variable. This sample size should be able to detect around $R^2 = .18$ (R = ± .42) with a power of .80 at a significance level ($\alpha$) = .05.

Criteria for selecting participants

L1 backgrounds.

The listeners’ native languages (L1s) were American English, Hindi (Indian), Mandarin, and Japanese, respectively representing the Inner, Outer, and Expanding Circles, and the fellow countrymen. American English speakers were chosen among other Inner Circle varieties because of the close political and economical relationship between Japan and the United States. Hindi was chosen because of the remarkable growth of India in such global markets as technology. Mandarin speakers, rather than Korean speakers,
were chosen from the Expanding Circle, because selected L1s should be unrelated to each other. Korean is widely believed to belong to the same linguistic family as Japanese. Japanese speakers were included, because there will be more opportunities for Japanese people to listen to their fellow countrymen speaking English as the global village becomes smaller and smaller. All the L1s happened to be from Asia, except for American English, which should be justifiable when the growing importance of English in Asia was considered.

**English proficiency.**

The rationale introduced in Chapter 1 for including less advanced speakers from whom speech samples are collected was relatively low ultimate English proficiency that the majority of Expanding Circle speakers attain. The discussion of the proficiency level in this circle naturally applies to listeners. Expanding Circle English users are supposed to be less advanced in the listening comprehension of English as well, in comparison with Inner and Outer Circle users. Therefore, ideally, listeners of various proficiency levels should have been included in the investigation.

However, based on the findings by Eisenstein and Berkowitz (1981) and Matsuura et al. (1999), nonnative listeners with low proficiency may not be able to cope with an intelligibility (dictation) test. The intelligibility scores of Matsuura et al.’s listeners failed to show significant differences between Irish and American English stimuli, even though the perceived comprehensibility ratings indicated that those listeners perceived American English easier to comprehend than Irish English. In order to control this floor effect of listeners’ English proficiency, at least a decent proficiency was
considered to be necessary in order to gain reliable intelligibility test data in this study. For this practical reason, nonnative listeners participating in this study were relatively advanced English users: international undergraduate and graduate students at an American university. The TOEFL score requirements for the admission to this university guaranteed relatively high English proficiency of the participants. Enrolled international students were supposed to have obtained a score of at least 527 on the paper-and-pencil based TOEFL (i.e., 197 on computer-based).

In reality, however, due to the shortage of L1 Japanese participants, the researcher had to include several Japanese participants who were in the pre-academic English courses. Consequently, the lowest TOEFL score of the participants was lower than 527. Prospective participants whose TOEFL scores were lower than 450 were excluded from the data analysis, but nonetheless English proficiency of some L1 Japanese participants was not very high. This unanticipated outcome influenced the later data analysis, which will be discussed in Chapter 4.

**Educational/vocational backgrounds.**

Recruitment was strictly limited to the mentioned university, to control the participants’ educational/vocational backgrounds. Participants were recruited from various disciplines. TESOL, Foreign and Second Language Education, and Linguistics majors were excluded, for they might be intensively trained to understand nonnative English. One exception was a Japanese male graduate student majoring in TESOL. He was included because of the shortage of L1 Japanese participants. Since his residence in the USA was one month at the time of the study, however, his training in the TESOL
program was considered as not taking effect yet. To minimize the influence from
residence in the USA, where students were assumed to have abundant opportunities to
have had contact with people from various nationalities, the listeners were limited to
those who had lived in the United States for less than two years. A sufficient number of
participants could not be recruited with the stricter residence restriction. The American
English native-speakers were recruited at the same university, in order to make the
educational level similar among the listeners. Non-Japanese who had lived in Japan for
more than a year were excluded.

**Gender and age.**

Listeners’ gender, a nominal variable, was dummy-coded and put into a set of
multiple regression analyses as a rival independent variable. The participants’ age could
not be restricted from 20 to 32 as originally planned, because of the difficulty of
recruitment. The final range of the participants’ age was from 19 to 38, except for one
participant, who was 66. The relationship between the participants’ age and dependent
variables will be explained in the correlation analysis in Chapter 4.

**Ethical considerations and incentives.**

All the personal information of the participants that the researcher needed to
obtain would remain strictly confidential, and the participants were informed of this
condition before they decided to participate. When the participants were recruited, they
were informed of the procedures of the study in a written form. Before the actual session
started, each participant was orally notified of his/her right to withdraw from the study at
any time with no disadvantage. When a session was completed, participants received $15 for compensation.

The proposal of this study was reviewed by the Office of Responsible Research Practices (ORRP) at The Ohio State University and given an exemption from the Behavioral and Social Sciences Institutional Review Board (IRB), with the protocol number 02E0409.

**Main Independent Variables**

**Speakers**

The speakers were selected based on their English proficiency, and consisted of four groups: (1) L1 Japanese speakers at the beginning level of L2 English proficiency, (2) L1 Japanese speakers with intermediate L2 English proficiency, (3) L1 Japanese speakers with advanced L2 English proficiency, and (4) L1 English speakers (i.e., native speakers) as the control group.

Speakers were expected to vary in many aspects, such as pronunciation (including L1 and L2 accents), delivery, quality of voice, grammatical accuracy, syntactic complexity, vocabulary sophistication, and discourse styles. No special manipulation of these variables was attempted in this study, however, except for the selection of the speech topic, which was expected to limit the range of vocabulary and syntactic complexity to some extent. This study did not aim to determine what linguistic factors impacted the listeners’ evaluative reactions and the intelligibility of the speech, but to see whether listeners’ L1 was a predictor of the dependent variables, as well as the perceived factors. The biggest interest was whether native listeners and nonnative listeners would
react differently to the same nonnative and native speech samples. Thus, naturally occurring speech was considered more appropriate than artificially controlled speech, such as matched-guise, which has been commonly used in language attitudinal studies.

The original plan was to recruit L1 Japanese speakers at a vocational/technical school of English in Japan. The rationale for using this population was as follows: First, speakers needed to have reached to a certain proficiency level to convey speech tasks. Therefore, junior and senior high school students were considered as unsuitable for this study. Second, students who learned English only for the purpose of passing university entrance exams were not appropriate either, for they might not necessarily be interested in communication in English. Only those who intended to use English for international communication would be concerned about their international intelligibility, as Smith (1992) maintained. Generally speaking, students in a vocational/technical college of English in Japan were strongly motivated to use spoken English in the future. Third, learners with a wide range of proficiency were easily accessed in this type of school: From very low to relatively advanced.

In order to make the speakers’ levels in English proficiency as different as possible, the following procedures were taken. First, the researcher obtained the name lists of three classes that represented three different proficiency levels at a technical/vocational school in Japan. Students’ TOEFL scores were requested with the name lists, but some students had never taken the test. In that case, approximate proficiency was estimated from other standardized qualification English tests, such as TOEIC and EIKEN (jituyou eigo kentei shiken [the test of practical English], which was a Japanese standardized English proficiency examination). Those who had lived abroad for
longer than a month in the past were excluded from the list. Then, teachers provided the names of the best three students in their class. Students were contacted individually and requested to take part in the study. When the recommended students were not available, substitutes were introduced. The speech by total of nine students was recorded.

However, when three ESL experts examined the proficiency levels of these speech samples in order to choose the most appropriate ones, a problem arose. None of the speakers was judged as advanced, even though two speakers had passed pre-1st grade in EIKEN, which was considered quite advanced in Japan. Therefore, the researcher had to seek unfailingly advanced Japanese speakers of English, and eventually included two speakers who had been in the USA for an extended period of time. These two speakers were judged advanced by three experts. The advanced speakers mainly learned English in Japan, staying there until they became at least 17 years old.

One male and one female native speakers, as a control group, were recruited so that they would match the Japanese counterparts. The conditions were (1) they should have high school diplomas but not university degrees (yet), (2) they should not be engaging in speaking professions, and (3) they should have minimum proficiency in any other language but English. Four speakers were recorded. Two were discarded because they did not follow the instructions.

Speech Materials

The speakers’ tasks were as follows: (1) reading a passage aloud and (2) narrating a series of cartoons (adapted from Derwing and Munro, 1997). The first task was done for only a diagnostic reason, and was never used in the study. In recording
sessions, speakers were given sufficient time to examine the cartoons carefully and to ask content and vocabulary questions to the researcher, for the limited English proficiency of some speakers were considered. To avoid the practice effect, as well as to relax the speakers, the researcher spoke to them exclusively in Japanese except for when she provided them with necessary vocabulary. Each narrative lasted approximately two minutes, and was used for evaluative reaction ratings as it was. Two sentences were extracted from each speech sample to be used in the intelligibility test. As a rule, third and fifth sentences from the beginning were extracted, but when they were too short or too long, the succeeding sentences replaced. These extracted sentences were re-recorded separately from the whole narration. The duration of the sentences ranged from 2 to 6 seconds.

All the 8 speakers’ performance was recorded on a PC by means of the PetitRec software (available at http://www.vector.co.jp/soft/win95/art/se246059.html?site=n). No participants or ESL experts who checked the speakers’ proficiency levels reported complaints about the sound quality.

The use of extemporaneous speech for intelligibility/comprehensibility analyses is not free from criticism, however. In this kind of speech, researchers can not control the effects of grammatical accuracy nor complexity and word choice on listeners’ judgement of intelligibility and comprehensibility, failing to identify the influence by each factor. Therefore, reading aloud has been the most frequently used technique to construct speech materials. However, reading aloud materials can provide data only on accent and speech rate. What distinguishes nonnative speech from native speech includes not only phonological but also lexical, syntactic, and, probably, pragmatic differences.
Thus, extemporaneous speech was adopted on the ground that it serves better to answer the research questions than speech materials prepared by reading pre-determined texts or word lists. Subtelny, Whitehead, and Orlando (1980) also recommended academic and clinical investigators of speech intelligibility to utilize spontaneous speech, instead of oral reading. The transcription and the rate of speech of each speech sample are indicated in Appendix A.

Cargile and Giles (1998) pointed out the importance of content of speech that affects the listeners’ evaluation of the speakers. They asked their speakers to read prose on political and non-political topics and compared the reactions to them. In this study, however, a simple story was chosen as the content. Speakers improvised a brief story looking at a series of cartoons. The cartoons were about a wife who left her husband and their baby at home to attend her friend’s wedding (Appendix B). Even if it might not be impossible, the beginning level nonnative speakers would have had a hard time to improvise speech on a controversial issue. On the other hand, telling a story was a more commonly required skill both for advanced and beginning L2 users, as well as for NSs.

Speech samples were presented to the listeners in such an order that no same sex nor same proficiency level would adjoin. The effect of presenting order of speech samples was not particularly taken care of, because all the listeners would listen to the samples in the identical order. Even if an order effect existed, it was assumed to affect all the listeners equally.
Outcome Measures (Dependent Variables)

Three instruments were developed by the researcher based on the review of related literature. The content validity of these measures was secured by consultation with a panel of experts [a Second Language Acquisition (SLA) expert, two English as a Second Language (ESL) experts: a sociolinguist, and a research methods expert]. Face validity was addressed in a field study with domestic and international students at the same university, who were similar to the listeners in the main study. The following are the descriptions of each instrument.

Evaluative Reactions Scale

Semantic differential scales with bipolar adjectives are commonly used in language attitudinal studies. Especially, the Speech Dialect Attitudinal Scale (SDAS) developed by Mulac and his colleagues (1974, 1976) and the Speech Evaluation Instrument (SEI) by Zhan and Hopper (1985) have been widely used. The validity and reliability of these scales were established with a large sample size. However, both the SDAS and the SEI had too many response items, 21 and 30, respectively. The revised versions had still too many, in comparison with the recommendation by Miller (2001). Miller (2001) recommended 8 or fewer items, because “respondents seem to lose their frame of reference to the concept when there are more than this” (p.95). Chiba et al. (1995) developed their 10-item scale, but failed to report their instruments’ validity and reliability. Furthermore, this new scale, as well as the SDAS and the SEI, included some items that seemed to be irrelevant for the use of international users of English. The seemingly inappropriate adjectives were: friendly-not friendly, sophisticated-not
sophisticated, confident-not confident, to name a few. A new instrument with a manageable number of cross-culturally relevant items had to be developed, with its reliability and validity clarified.

Seven pairs of opposite adjectives and nouns were chosen from existing instruments. The items are: clear content -- unclear content (Ohama, Gotay, Pagano, Boles, and Craven, 2000), easy to understand -- difficult to understand (Derwing and Munro, 1997; Ohama et al., 2000), good pronunciation -- bad pronunciation (Giles, Williams, and MacKie, 1995), smooth -- choppy (Cargile and Giles, 1998), pleasing --irritating (Giles et al., 1995, modified), complex -- simple (Giles et al., 1995), and with no accent -- with strong accents (Chiba et al., 1985; Derwing and Munro, 1997). One pair (effective communication -- ineffective communication) was newly added, in consideration of the significance of communicative skills that are strongly advocated in language education, especially ELT.

The number of intervals between the bipolar adjectives was chosen to maximize the variability of individual scores. Derwing and Munro (1999) conducted their study with a nine-point Likert-type scale, though not ten, and reported their lay participants could reliably respond with nine levels. With the number of intervals, 10, an even number, the researcher could force listeners to choose either side, positive or negative, with no neutral choice. Miller (2001) stated that attitudes are rarely completely neutral.

The adjectival phrases were presented in a random order with negative and positive values alternatively, in order to avoid response sets. In the phase of data analysis, negative values were reversed for calculation. A score for one item varied from 1 (extremely negative) to 10 (extremely positive). The sum of the raw scores for 8 traits,
with potential minimum 8 and maximum 80, was the Evaluative Reactions score for each individual listener. The final version of the Evaluative Reactions Scale was indicated in Appendix C.

Cronbach’s alpha was calculated for each speaker to ensure internal consistency, a type of reliability, based on the pilot study data (Table 3.2). Cronbach’s alphas are between .67 to .80, except for that of Speaker 4. Hair, Anderson, Tatham, and Black (1998) recommended the lower limit of acceptability of .70, but they also stated that alpha of .60 was acceptable with a small number of items. The unacceptable alpha of Speaker 4 and the relatively low alpha of Speaker 7 may indicate unreliability of this scale when used to evaluate native speakers. When used for non-native speech evaluation, the intended purpose of this scale, this measurement indicated acceptable reliability.

<table>
<thead>
<tr>
<th>Evaluative Reactions to:</th>
<th>Cronbach’s alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speaker 1 (advanced, male)</td>
<td>.78</td>
</tr>
<tr>
<td>Speaker 2 (beginner, female)</td>
<td>.67</td>
</tr>
<tr>
<td>Speaker 3 (intermediate, male)</td>
<td>.80</td>
</tr>
<tr>
<td>Speaker 4 (native, female)</td>
<td>.49</td>
</tr>
<tr>
<td>Speaker 5 (beginner, male)</td>
<td>.73</td>
</tr>
<tr>
<td>Speaker 6 (intermediate, female)</td>
<td>.78</td>
</tr>
<tr>
<td>Speaker 7 (native, male)</td>
<td>.67</td>
</tr>
<tr>
<td>Speaker 8 (advanced, female)</td>
<td>.73</td>
</tr>
</tbody>
</table>

Table 3.2: Cronbach’s alpha across eight speakers
Intelligibility Test

Dictation has been adopted as a measuring method from the study by Derwing and Munro (1997). Since Brodkey (1972) suggested dictation as a valid and reliable measure for intelligibility of accented speech, dictation, in other words, word-for-word transcription, has been commonly used in intelligibility studies (e.g., Gass and Varonis, 1984; Munro and Derwing, 1995b; Derwing and Munro, 1997).

Two alternative methods, a partial dictation and a multiple-choice comprehension test, were considered, but dismissed. A partial dictation test (i.e., Cloze test or C-test) such as the one used in the study by Smith and Rafigzad (1979) and Matsuura et al. (1999) would give away contextual clues to listeners, and thus, might examine their reading, or grammar, or vocabulary skills, instead of listening comprehension. A multiple-choice comprehension test, as seen in the research by Smith and Basazza (1982), Matsuura et al. (1999) and Major et al. (2002), would make inter-speaker comparison impossible, with extemporaneous speech. The word-for-word transcription was decided to better serve as a measurement of “accurate recognition of sounds and words,” which is the definition of “intelligibility” by Smith (1992).

For the intelligibility test, sixteen sentences were extracted as the target of dictation from the 8 speech samples that were recorded for the Evaluative Reactions test. Listeners were asked to write down every single word. The intelligibility test scores were calculated based on the percentage of words that listeners correctly transcribed from the heard sentences. The “correct” transcription version (Appendix A, with the sentences selected for the intelligibility test in the bold print) was carefully constructed by the researcher, by listening to the samples as many times as necessary. Another ESL expert
confirmed the accuracy of the transcription. Derwing and Munro (1997) used two scoring methods: including and excluding trivial errors. They found very small differences between the method accepting trivial errors and the method of exact word-matches that penalized trivial errors. In this study, scores were calculated in both methods, but because of the overall high intelligibility, the method that allowed trivial errors failed to produce variance in intelligibility scores. Therefore, scores obtained by the exact-match method were used in the analysis. The form of Intelligibility test is shown in Appendix D.

**Rival Independent Variables**

**Personal Information Questionnaire**

In order to control extraneous variables and estimate the effects of rival independent variables, listeners were requested to provide demographic and linguistic background information on a questionnaire administered at the end of listening sessions (see Appendix E). Salant & Dillman (1994) clearly stated that the participants’ responses to this type of questionnaire on personal attributes are quite reliable, compared with attitudinal measurements. The following items were on the questionnaire.

1. Age
2. Gender
3. Residence in the USA
4. Educational background (Undergraduate or graduate or others)
5. Major(s) at the OSU
6. Native speaker of American English or nonnative? (Questions 7, 8, and 9 are relevant only for nonnative speakers of English.)
(7) Native language
(8) Home language
(9) The TOEFL score rank (for nonnative listeners only)
(10) The degree of exposure to Japanese speakers of English
(11) The degree of exposure to other nonnative varieties of English

Three questions (i.e., Q2 on gender, Q10 on familiarity with Japanese English, and Q11 on exposure to foreign languages and cultures) were asked to quantify the rival independent variables. Other questions were asked in order to confirm that the listeners fulfilled the required qualifications and to describe the subjects more precisely.

**Conditions of Testing**

Data collection started on June 10th and finished on November 4th, 2003. This unexpectedly prolonged duration of data collection was due to the difficulty to find subjects. Each session was conducted in the following procedures.

One or two listeners participated in a session at a time. First, informed consent was obtained from each participant (See Appendix F). The listeners self-reported whether they had any kind of hearing problems. Then, the listeners were instructed that they were going to listen to eight narratives, each of which would last approximately two minutes, and that the narratives would explain the same series of cartoons, which were shown to the listeners. The visual clues were presented for two reasons. First, in real communication, interlocutors often share the knowledge about a broad context. Second, by providing some information about the content, the effect of speech presentation order should be mitigated. Each listener was provided with a headset that was directly attached
to the researcher’s computer, and the volume was set to his/her comfortable level. The first task was to rate the speech on the Evaluative Reactions scale. Before the eight speech samples, one practice sample was demonstrated so that the researcher could confirm that the participants had understood their task.

After the Evaluative Reactions scale administration, the intelligibility test was provided. The listeners listened to the tape, but this time to shorter extracted sentences. The listeners were instructed to transcribe word-for-word as accurately as possible when they heard each sentence. They were also informed that they would listen to each sentence only once. A practice sentence was presented before the real session. The listeners had as much time as they needed to complete the task before they proceeded to the next sentence.

After the transcription intelligibility test, the listeners were asked to answer the personal information questionnaire. Participants were allowed to ask questions to the researcher, if they had any. After the completion of the questionnaire, a few questions were asked orally. Then, the participants were thanked and received $15. The whole procedure took no more than one hour.

Data Analysis

SPSS for Windows version 11.5 was used to analyze the quantitative data. Data were first summarized and organized with descriptive statistics. Since the scales of measurements of main dependent variables were interval and ratio, the mean, median, and mode were the central tendencies to be reported. Standard deviations described variability.
Correlation analyses were used to explain the relationship between the dependent variables and independent variables. The scales of measurements and the correlations used to describe the relationship between variables are shown in Table 3.3.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Scale</th>
<th>X1</th>
<th>X2</th>
<th>X3</th>
<th>X4</th>
<th>X5</th>
<th>Y1</th>
<th>Y2</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1: Listeners’ L1</td>
<td>Nominal</td>
<td></td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>X2: Speakers’ proficiency</td>
<td>Ordinal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>K</td>
<td>K</td>
</tr>
<tr>
<td>X3: Listeners’ gender</td>
<td>Nominal</td>
<td></td>
<td>C</td>
<td>C</td>
<td>rpb</td>
<td>rpb</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>X4: Familiarity with Japanese English</td>
<td>Ordinal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>K</td>
<td>K</td>
</tr>
<tr>
<td>X5: Exposure to foreign cultures</td>
<td>Ordinal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>K</td>
<td>K</td>
</tr>
<tr>
<td>Y1: Evaluative Reactions score</td>
<td>Interval</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>P</td>
<td></td>
</tr>
<tr>
<td>Y2: Intelligibility score</td>
<td>Ratio</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

P: Pearson product–moment coefficient  
K: Kendall’s tau  
rpb: point-biserial correlation coefficient  
C: Cramer’s V

Table 3.3: The scales of measurements and descriptive statistics used in the data analysis

Among the independent variables, one variable named “speakers’ proficiency level” was unique. It was an attribute of speech samples (or speakers), while other independent variables were those of listeners. Examining the correlation between speakers’ proficiency level and other independent variables had no point. Therefore, the following table contains some blanks. The magnitude of the relationship was interpreted based on the scale by Davis (1971) as shown in Table 3.4.
<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>.70 to .99</td>
<td>Very strong association</td>
</tr>
<tr>
<td>.50 to .69</td>
<td>Substantial association</td>
</tr>
<tr>
<td>.30 to .49</td>
<td>Moderate association</td>
</tr>
<tr>
<td>.10 to .29</td>
<td>Low Association</td>
</tr>
<tr>
<td>.01 to .09</td>
<td>Negligible Association</td>
</tr>
</tbody>
</table>

Table 3.4: Magnitudes of association

In order to test the four hypotheses formulated in Chapter 1, a set of linear regression analyses were conducted. The data analysis model is indicated in Figure 3.2. A simultaneous regression analysis was conducted with all the independent variables first and the resulting model was examined for significance. Coefficients of determination ($R^2$) were reported, as they described the magnitude of the regression of the dependent variable explained by the linear combination of the independent variables. The explained variance of the dependent variables in the initial regression was $A + B + C$ in Figure 3.2. When there were some independent variables whose semipartial (part) correlation coefficients were relatively small, then a stepwise regression was administered to investigate which variables contributed to the regression model. The sample size of 80 should detect $R^2$ as small as .18. Since the second regression of intelligibility entered a rival independent variable first ($B + C$, in the figure below), the unique variance explained by the main independent variables was indicated as $R^2$ change ($A$, in the Figure 3.2).
The Dependent Variable: *interval*
- Listeners’ evaluative reactions to Japanese English

The Dependent Variable: *ratio*
- Intelligibility of Japanese English

The Main Independent Variables
- Listeners’ L1: *nominal*
- Speakers’ proficiency level: *ordinal*

The Rival Independent Variables
- Listener’s gender:  *nominal; dichotomous*
- Listeners’ familiarity with JE:  *ordinal*
- Listener’s exposure to foreign languages and cultures:  *ordinal*

Figure 3.2: Data analysis model of the *ex post facto* study
Qualitative Interview Study

The listeners’ individual point of view was of significant interest to the researcher. The researcher was interested not only in their test scores (i.e., product), but also in the process, in other words, what kinds of decisions the listeners made during the test administration. For instance, at what point did a listener fail to recognize speech sounds or words? Or on what basis did a listener make a judgement? On the advantages of qualitative inquiry over quantitative research, Denzin and Lincoln (2000) maintained that qualitative investigators could get closer to the individuals’ point of view through interviews. The data inexpressible in numbers might offer rich information for understanding an incident deeper. Thus, in this study, qualitative data were collected through interviews as a supplement to the ex post facto study.

Participants

Thirteen interviewees were selected purposively, to obtain diversity based on the data obtained in the quantitative inquiry: evaluative reactions measured by the scale, gender, and race (which was only possible with American listeners.) The original plan was to recruit four students from four different L1 groups. It was not implemented, however, because three participants were no longer available at the point of qualitative data collection.
**Interview Questions**

The overarching research question for the interviews was formulated as follows: **What are the perceived factors that influence the listeners’ evaluative reactions to and the intelligibility of Japanese English?** In order to delve into the participants’ perceptions, interview questions were formulated as follows based on the initial analysis of quantitative data and the observation of quantitative study participants during the sessions.

1. *What did you think or how did you feel while you were rating this speaker?*
2. *What are accents? Do they bother you?*
3. *Have you found specific accents or specific language speakers more difficult to understand than others? Why?*
4. *Do you feel more comfortable speaking with NS or NNS, or doesn’t it matter?*

**Data Collection**

Interviewing was chosen as the data collection method because it is “one of the most common and powerful ways in which we try to understand our fellow human beings” (Fontana and Frey, 2000, p. 645). Interviews in this study had to be face-to-face, because the interviewees needed to listen to the speech samples that they had heard in the quantitative inquiry for the second time, in order to describe what was on their mind in the first listening. According to Gilhooly and Green (1996), retrospective (recall) protocol can not be expected to be as accurate nor as complete as a concurrent protocol, which is
also called “thinking aloud,” because rationalization may affect the retrospective protocol. However, for L2 English speaking participants, although they were relatively advanced, the additional task of expressing their thoughts in a foreign/second language while listening to and evaluating English speech simultaneously would be excessive. As a trade-off, the researcher was willing to sacrifice some accuracy and completion of the data for the feasibility of the task. Qualitative data collection was carried out from August to October 2003.

The procedure was as follows. After signing the consent form (Appendix G), participants listened to the eight speech samples and rate it again. In order to minimize the effects of retrospective rationalization, the participants were not allowed to see their previous ratings. Then, the researcher asked open-ended questions from the interview questions. The listeners were requested to articulate what was on their mind when they rated speech on each of eight given traits (i.e., interview questions #1). Since the interviews were semi-structured, the interviewees’ spontaneous responses were encouraged and not all the interview questions were asked. Interviews were recorded with a portable tape recorder (SONY TCM-20DV) with the consent of the participants. Interviews lasted from about 60 to 90 minutes. Field notes were taken. Each participant was interviewed only once. Interviews were transcribed by the researcher by using Panasonic Standard Cassette Transcriber Model RR-830. The transcribed data consisted of more than 130 letter-size pages, single-spaced.

Fontana and Frey (2000) stated that qualitative interviews should be conducted in natural settings, rather spontaneously. Thus, they continued, the researcher’s prolonged field engagement is required. This study had to violate these assumptions. Since the main
purpose of these interviews was to probe the listener’s evaluative reactions to specific speech samples, the tape had to be replayed in the presence of the listener and the researcher. Therefore, the interview setting was quite formal, far from being spontaneous. Even though the researcher saw every interviewee only twice, however, the researcher believes sincere conversations were elicited from the majority of participants. Schwandt (1997) stated, “But the use of structured open- (and closed-) ended interviews is not uncommon in field studies where time is limited, where it is desirable to get some specific or focused information from a large number of people, or both.” (p. 74) This study falls into this category.

Data Analysis

In a qualitative inquiry, data collection and data analysis are conducted simultaneously (Creswell, 1994). Thus, the researcher tried to find patterns among a participant’s comments during and after each interview. What seemed to be key words were highlighted on transcripts and field notes, and marginal notes were taken. As raw data gathered, the researcher went back to the previously collected data and compared them with new ones, trying to find common patterns among participants. After reading the transcripts several times, recurring themes emerged. These themes were color-coded and labeled.

Combining the Quantitative and Qualitative Data

The “combined method” suggested by Creswell (1994) or the “methodological eclecticism” by Hammersley (1996) was adopted in this study. In a more precise typology
by Tashakkori and Teddie (1998), this study employed a *sequential quantitative-qualitative analysis*, by forming groups of people/settings on the initial basis of quantitative data and then comparing the groups on qualitative data (p 135). Based on the scores on the Evaluative Reactions scale in the quantitative study, the participants were categorized into three groups: very positive, average, and very negative in attitudes toward Japanese English. This attitudinal difference and the participants’ L1 were given priority when the interviewees were selected, even though other differences were taken into consideration as mentioned above. The reason for this selection of interviewees was to observe the maximum diversity among them. In other words, the researcher wanted to understand the differences in participants’ decision-making processes that distinguished critical raters from sympathetic ones. Did a critical L1 Hindi listener, for example, utilize different rationalizations when evaluating Japanese English from his/her L1 Mandarin counterpart?

After the quantitative and qualitative data were analyzed and interpreted independently, the findings were combined. Since the main framework of this study was quantitative, the combination was mainly conducted by interpreting “qualitative findings to explain quantitative results,” according to the recommendation of Tashakkori and Teddlie (1998). However, the “combined method” is beneficial not only for triangulation, by which any bias inherent in a single method can be neutralized to some extent, but also for complementarity, in which overlapping and different facets of one phenomenon may emerge, as Creswell (1994) and Hammersley (1996) state. Therefore, the qualitative findings that were not directly related to the explanation of quantitative results were also considered as an important layer of the phenomena under investigation.
CHAPTER 4

RESULTS AND DISCUSSION

The study aimed at investigating the factors that predict native and nonnative listeners’ evaluative reactions to and the intelligibility of Japanese English speakers with different levels of proficiency. The processes in which listeners made judgements in their evaluative reactions to each speaker were also explored. Therefore, the research design was twofold: the *ex post facto* quantitative study phase as the main research framework with the qualitative interview phase. This chapter comprises the analyses of collected data and the discussion of the results and findings. First, the outcome of the quantitative research is discussed, and then the findings of the interviews are reported. The integration of the quantitative and qualitative findings will be implemented in the following chapter.

**Quantitative *ex post facto* Study Results**

The variables to be predicted in the study were the listeners’ scores on the Evaluative Reactions scale and on the Intelligibility test. The main independent variables were the listeners’ native language (L1) and the speakers’ English proficiency levels (i.e.,
native, advanced, intermediate, and beginning-level). The rival independent variables were the listeners’ gender, familiarity with Japanese English, and exposure to foreigners. Listeners were recruited across four L1s (i.e., American English, Hindi, Mandarin, and Japanese) at a large public university in the Midwest of the USA.

Descriptions of Participants

The original plan was to recruit twenty students for each L1 group. For administrative reasons, however, slightly more participants were actually included in the L1 American English and Japanese groups. The final breakdown of participants’ L1 is indicated in Table 4.1.

<table>
<thead>
<tr>
<th>L1</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>22</td>
<td>26.2</td>
</tr>
<tr>
<td>Hindi</td>
<td>20</td>
<td>23.8</td>
</tr>
<tr>
<td>Mandarin</td>
<td>20</td>
<td>23.8</td>
</tr>
<tr>
<td>Japanese</td>
<td>22</td>
<td>26.2</td>
</tr>
<tr>
<td>Total</td>
<td>84</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Table 4.1: Listeners’ native languages (L1s)
Eighty-four graduate and undergraduate students majoring in 46 different fields of study participated in the study. Since the researcher proctored every administration of instruments to a maximum of two participants at a time, and reminded them of the importance of completing all the responses, no data were missing. All the data collected from 84 participants were considered valid and brought into analysis.

Among the 46 fields, six participants majored in Chemical Engineering, five in Industrial and Systems Engineering, five in Accounting, and four in Electrical Engineering. Majors in the social sciences, such as sociology, and in the humanities, such as English, were also common.

As shown in Table 4.2, the average age of the participants was 25.1 years, with the range from 19 to 66 years. The mean ages of L1 groups were 24.8 years (L1 = American English), 25.1 years (L1 = Hindi), 26.5 years (L1 = Mandarin), and 24.1 years (L1 = Japanese). The wide age range of L1 American English members in comparison with that of native speakers of other languages was due to only one participant, aged 66 years. The second oldest American participant was 34 years old.

Forty-eight participants (57.1%) were female, while 36 (42.9%) were male. Unfortunately, the gender ratio among L1 groups could not be made more similar.

Of 84 participants, 31 (37%) were undergraduate students, whereas 53 (63%) were graduate students. Except for the L1 American English group, graduate students outnumbered undergraduate students. Especially among L1 Hindi and Mandarin groups, as many as 90 to 95% of the participants were graduate students.

The nonnative speaking participants’ average length of residence in the United States was 13.1 months (1.1 years). The mean lengths of residence of L1 groups were
16.3 months (L1 = Hindi), 14.7 months (L1 = Mandarin), and 8.8 months (L1 = Japanese).

TOEFL scores were examined for the purpose of assuring sufficient English proficiency on the part of the listeners. However, statistics suggest L1 group differences. The TOEFL scores of many Japanese speaking participants were not as high as those of L1 Hindi and Mandarin listeners. The descriptive statistics of TOEFL scores of L1 Hindi and Mandarin participants were identical, but there were 8 Hindi speakers at the highest level (TOEFL 650 points and higher), while only 3 Mandarin speakers were at the same level.

<table>
<thead>
<tr>
<th>Variable</th>
<th>L1 breakdown</th>
<th>N</th>
<th>N</th>
<th>Descriptives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>84</td>
<td></td>
<td></td>
<td>Mean = 25.1 years, Standard Deviation (SD) = 6.41 Range = 19/66</td>
</tr>
<tr>
<td></td>
<td>English speakers</td>
<td>22</td>
<td></td>
<td>Mean = 25.5 years SD = 9.88 Range = 19/66</td>
</tr>
<tr>
<td></td>
<td>Hindi speakers</td>
<td>20</td>
<td></td>
<td>Mean = 25.1 years SD = 5.45 Range = 19/38</td>
</tr>
<tr>
<td></td>
<td>Mandarin speakers</td>
<td>20</td>
<td></td>
<td>Mean = 26.5 years SD = 3.38 Range = 21/33</td>
</tr>
<tr>
<td></td>
<td>Japanese speakers</td>
<td>22</td>
<td></td>
<td>Mean = 24.1 years SD = 4.95 Range = 19/38</td>
</tr>
</tbody>
</table>

(Table Continues)

Table 4.2: Descriptive statistics for Age, Gender, US residence, and TOEFL scores
Table 4.2: (Continued)

<table>
<thead>
<tr>
<th>Gender</th>
<th>84</th>
<th>Female = 48 (57.1%), Male = 36 (42.9%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>English speakers</td>
<td>22</td>
<td>Female = 14 (64%), Male = 8 (36%)</td>
</tr>
<tr>
<td>Hindi speakers</td>
<td>20</td>
<td>Female = 3 (15%), Male = 17 (85%)</td>
</tr>
<tr>
<td>Mandarin speakers</td>
<td>20</td>
<td>Female = 14 (70%), Male = 6 (30%)</td>
</tr>
<tr>
<td>Japanese speakers</td>
<td>22</td>
<td>Female = 17 (77.3%), Male = 5 (22.3%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>School</th>
<th>84</th>
<th>Undergraduate = 31 (36.9%), Graduate = 53 (63.1%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>English speakers</td>
<td>22</td>
<td>Undergraduate = 16 (73%), Graduate = 6 (27%)</td>
</tr>
<tr>
<td>Hindi speakers</td>
<td>20</td>
<td>Undergraduate = 2 (10%), Graduate = 18 (90%)</td>
</tr>
<tr>
<td>Mandarin speakers</td>
<td>20</td>
<td>Undergraduate = 1 (5%), Graduate = 19 (95%)</td>
</tr>
<tr>
<td>Japanese speakers</td>
<td>22</td>
<td>Undergraduate = 12 (55%), Graduate = 10 (45%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>US Residence</th>
<th>62</th>
<th>Mean = 13.1 months, SD = 7.65, Range = 1/24</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hindi speakers</td>
<td>20</td>
<td>Mean = 16.3 months, SD = 6.71, Range = 4/24</td>
</tr>
<tr>
<td>Mandarin speakers</td>
<td>20</td>
<td>Mean = 14.7 months, SD = 6.81, Range = 1/24</td>
</tr>
<tr>
<td>Japanese speaker</td>
<td>22</td>
<td>Mean = 8.9 months, SD = 7.51, Range = 1/24</td>
</tr>
</tbody>
</table>

(Table Continues)
The participants were also requested to account for the languages that they used on a daily basis. Ten (50%) L1 Hindi members spoke Hindi for their daily communication, while 4 (20%) used English more frequently. Among L1 Japanese participants, 16 (73%) used Japanese mainly and 6 (27%) spoke English more. Twelve (60%) L1 Mandarin members primarily spoke Mandarin while 6 (30%) spoke English. Other participants used two or three languages to almost the same extent. In short, the majority of nonnative participants in the study used their native language as the primary communication tool on a daily basis, while 20 to 30% of participants used English more frequently than their native language.
The relatively low tendency to use English for daily conversation might be related to the timing of the data collection. The data were collected mostly in summer, when participants were not taking as many courses as in other quarters. Understandably, however, the participants who responded that they used their native language as the primary communication tool stated that they were using English the second most frequently.

**Data Analysis**

**Descriptive Analysis of Variables**

**Independent variables.**

The main independent variables were the listeners’ L1, presented in the previous section, and the speakers’ proficiency levels, determined by their standardized English test scores and the ratings by ESL experts (see Chapter 3 for details). The following table (Table 4.3) displays the description of the speakers’ English proficiency levels and their gender. The speakers were presented in this order when the Evaluative Reaction scale was administered. The abbreviation of each speaker’s English proficiency level and gender is also provided below.
The rival independent variables were the listeners’ gender, which was reported in Table 4.2, familiarity with Japanese English, and exposure to foreigners. The familiarity with Japanese English was measured by Question #10 in the Personal Information Questionnaire (Appendix E). Question #10 is reiterated below:

(10) Do you know any Japanese persons who speak English?
   □ NO (Please go to Number (11))
   □ YES (Please answer the following question.)

   During the past year, on average, how often have you heard Japanese persons speak English? Check ONE box with the frequency that best describes your case.
1 = I have almost never heard Japanese persons speak English, except for simple and short greetings and responses.
2 = I might have heard Japanese persons speak English occasionally, but not on a regular basis.
3 = I have heard Japanese persons speak English on a regular basis, but less frequently than once a month.
4 = I have heard Japanese persons speak English almost every month or more often, but not as often as every week.
5 = I have heard Japanese persons speak English almost every week or more frequently.

The participants who answered “no” to the first question were given a score of 0, which means no familiarity with Japanese English. The scores ranged from 0 to 5 and were treated as ordinal data. Since this was not a summated scale, it could not be treated as interval.

Table 4.4 displays the distribution of the familiarity responses. Sixteen (73%) L1 Japanese participants stated that they heard Japanese people speaking English every week or more often. Participants from other L1 groups had less familiarity with Japanese English. Especially among L1 Hindi participants, nine (45%) responded that they never heard Japanese speak English.
<table>
<thead>
<tr>
<th>Native Languages</th>
<th>American</th>
<th>Hindi</th>
<th>Mandarin</th>
<th>Japanese</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Familiarity with Japanese English</td>
<td>0=None</td>
<td>4</td>
<td>9</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>1=almost never</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2=not regularly</td>
<td>3</td>
<td>4</td>
<td>9</td>
<td>1</td>
<td>17</td>
</tr>
<tr>
<td>3=less than once a month</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>4=every month</td>
<td>6</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>12</td>
</tr>
<tr>
<td>5=every week or more</td>
<td>6</td>
<td>3</td>
<td>3</td>
<td>16</td>
<td>28</td>
</tr>
<tr>
<td>Total</td>
<td>22</td>
<td>20</td>
<td>20</td>
<td>22</td>
<td>84</td>
</tr>
</tbody>
</table>

Table 4.4: Listeners’ familiarity with Japanese English

The listeners’ exposure to foreigners, defined as NNSs other than Japanese, was measured by Question #11 in the personal questionnaire. Question #11 is reproduced below:

(11) Do you know any non-native English speakers (= persons for whom English is not their native language but speak English), other than Japanese speakers?

□ NO
□ YES (Please answer the following question.)

During the past year, on average, how frequently have you heard non-native English
speakers, other than Japanese speakers? Check ONE box with the frequency that best describes your case.

☐ 1 = I have almost never heard non-native speakers speaking English except for simple and short greetings and responses.
☐ 2 = I might have heard non-native English speakers occasionally, but not on a regular basis.
☐ 3 = I have heard non-native English speakers on a regular basis, but less frequently than once a month.
☐ 4 = I have heard non-native English speakers almost every month or more often, but not as often as every week.
☐ 5 = I have heard non-native English speakers almost every week or more frequently.

Again, the participants who responded “no” to the first question were given a score of 0. The scores were treated as an ordinal scale.

The descriptive statistics are shown in Table 4.5. The table indicates that the majority (64% to 85%) of participants across L1 groups were exposed to nonnative English quite intensively. Compared with the familiarity with Japanese English, the participants’ exposure to foreigners was high in general.

Even though the term “foreigners,” adopted from the study by Gass and Varonis (1984), seems to be inclined toward American perspectives, it is only used for the convenience of analysis. Note that Question #11 was worded to include English users of the listeners’ own varieties. This is consistent with the expectations based on the responses for the Japanese in the previous question.
<table>
<thead>
<tr>
<th>Exposure to Foreigners</th>
<th>Native Languages</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>American</td>
<td>Hindi</td>
</tr>
<tr>
<td>0=none</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>1=almost never</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2=not regularly</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>3=less than once a month</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>4=every month</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>5=every week or more</td>
<td>14</td>
<td>17</td>
</tr>
<tr>
<td>Total</td>
<td>22</td>
<td>20</td>
</tr>
</tbody>
</table>

Table 4.5: Listeners’ exposure to foreigners (nonnative English speakers)

Dependent variable: evaluative reactions.

The evaluative reactions, a dependent variable, were defined as the scores on the Evaluative Reactions Scale (Appendix C). Table 4.6 indicates the descriptive statistics of the evaluative reactions to each speech sample. The mean scores are the averaged sum of the points that the listeners marked on eight traits in the Evaluative Reactions Scale. The potentially maximum score was 80 while the potentially minimum score was 8.

Table 4.6 reveals that both NSs (N♂ and N♀) were judged far more positively than the NNSs. The mean reactions scores given to Speaker 7 (N♂) and Speaker 4 (N♀) were 70.5 and 69.9 when the full score was 80. Two advanced NNSs (A♂ and A♀), with mean scores of 49.0 and 46.9 respectively, were judged more positively than two
intermediate NNSs (I♂ and I♀) with scores of 43.6 and 44.5, but the differences were small. The beginning-level NNSs (B♂ and B♀) were evaluated the least positively, with mean scores of 30.0 and 29.0. The standard deviations of the evaluative reactions to NSs and beginning level speakers were low (from approximately 6.5 to 7.8), compared with those of intermediate and advanced NNSs (from 9.6 to 11.9). Not surprisingly, native speakers and beginning level speakers induced more homogeneous responses from the listeners than the speakers whose levels were between the extremes.

<table>
<thead>
<tr>
<th>Speaker</th>
<th>Mean</th>
<th>Range</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speaker1 (A♂)</td>
<td>49.04</td>
<td>13-73</td>
<td>11.91</td>
</tr>
<tr>
<td>Speaker2 (B♀)</td>
<td>29.00</td>
<td>10-60</td>
<td>9.59</td>
</tr>
<tr>
<td>Speaker3 (I♂)</td>
<td>43.57</td>
<td>15-68</td>
<td>10.39</td>
</tr>
<tr>
<td>Speaker4 (N♀)</td>
<td>69.93</td>
<td>54-80</td>
<td>6.49</td>
</tr>
<tr>
<td>Speaker5 (B♂)</td>
<td>30.02</td>
<td>12-57</td>
<td>9.87</td>
</tr>
<tr>
<td>Speaker6 (I♀)</td>
<td>44.46</td>
<td>19-76</td>
<td>10.30</td>
</tr>
<tr>
<td>Speaker7 (N♂)</td>
<td>70.49</td>
<td>40-80</td>
<td>7.78</td>
</tr>
<tr>
<td>Speaker8 (A♀)</td>
<td>46.93</td>
<td>17-75</td>
<td>10.45</td>
</tr>
<tr>
<td>Average Reactions to all the Speakers</td>
<td>47.94</td>
<td>29.25-69.38</td>
<td>6.41</td>
</tr>
<tr>
<td>Reactions to JPN NNSs</td>
<td>40.51</td>
<td>20.00-66.33</td>
<td>7.75</td>
</tr>
</tbody>
</table>

Table 4.6: Descriptive statistics of Evaluative Reactions
The comparison of L1 group means of evaluative reactions to speech in general is shown in Table 4.7. Overall, L1 American English listeners’ evaluative reactions were the most positive and L1 Hindi listeners were the most negative with mean scores of 51.6 and 42.9 respectively. This finding can also be seen in the skewness of distributions of scores, which indicates that the L1 American English group’s median (the middle score) was smaller than the mean, while the L1 Hindi group’s median was larger than the mean. The lower kurtosis of the L1 Japanese listeners indicates the flatness of the distribution. They marked fewer extremes on the Evaluative Reactions Scale as compared with other listener groups.

Since this research is not experimental in design, the group mean difference cannot be calculated by such a statistic as ANOVA. A regression analysis will later examine whether the groups are significantly different in evaluative reactions.

<table>
<thead>
<tr>
<th>Native Language</th>
<th>Mean</th>
<th>N</th>
<th>SD</th>
<th>Range</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Minimum</td>
<td>Maximum</td>
<td></td>
</tr>
<tr>
<td>American English</td>
<td>51.58</td>
<td>22</td>
<td>6.38</td>
<td>41.8</td>
<td>69.4</td>
<td>.94</td>
</tr>
<tr>
<td>Mandarin</td>
<td>49.82</td>
<td>20</td>
<td>5.95</td>
<td>34.5</td>
<td>64.0</td>
<td>-.25</td>
</tr>
<tr>
<td>Japanese</td>
<td>47.14</td>
<td>22</td>
<td>5.15</td>
<td>37.8</td>
<td>58.1</td>
<td>.45</td>
</tr>
<tr>
<td>Hindi</td>
<td>42.90</td>
<td>20</td>
<td>4.84</td>
<td>29.3</td>
<td>47.6</td>
<td>-1.51</td>
</tr>
<tr>
<td>Total</td>
<td>47.93</td>
<td>84</td>
<td>6.41</td>
<td>29.3</td>
<td>69.4</td>
<td>.271</td>
</tr>
</tbody>
</table>

Table 4.7: Mean comparison among L1s (Evaluative Reactions)
Each proficiency level’s mean score on the Evaluative Reactions Scale by L1 group is displayed in Table 4.8. Across most proficiency levels, the L1 American English group was the most positive, followed by the L1 Mandarin and the L1 Japanese groups. The L1 Hindi group was the most critical. The only exception to this tendency was found in the reactions to the advanced nonnative speech. Mandarin speaking participants judged advanced NNSs more highly than American listeners did.

<table>
<thead>
<tr>
<th>Native Languages</th>
<th>Reactions to Native Speakers</th>
<th>Reactions to Advanced Speakers</th>
<th>Reactions to Intermediate Speakers</th>
<th>Reactions to Beginning Speakers</th>
</tr>
</thead>
<tbody>
<tr>
<td>American English</td>
<td>Mean</td>
<td>74.11</td>
<td>50.80</td>
<td>47.59</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>22</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>Std. Deviation</td>
<td>4.08</td>
<td>8.80</td>
<td>10.43</td>
</tr>
<tr>
<td>Hindi</td>
<td>Mean</td>
<td>65.05</td>
<td>40.26</td>
<td>39.63</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Std. Deviation</td>
<td>7.09</td>
<td>6.61</td>
<td>7.85</td>
</tr>
<tr>
<td>Mandarin</td>
<td>Mean</td>
<td>73.03</td>
<td>51.00</td>
<td>44.93</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Std. Deviation</td>
<td>5.43</td>
<td>7.78</td>
<td>8.20</td>
</tr>
<tr>
<td>Japanese</td>
<td>Mean</td>
<td>68.43</td>
<td>49.45</td>
<td>43.61</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>22</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>Std. Deviation</td>
<td>6.05</td>
<td>8.87</td>
<td>6.57</td>
</tr>
<tr>
<td>Total</td>
<td>Mean</td>
<td>70.21</td>
<td>47.99</td>
<td>44.02</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>84</td>
<td>84</td>
<td>84</td>
</tr>
<tr>
<td></td>
<td>Std. Deviation</td>
<td>6.70</td>
<td>9.088</td>
<td>8.73</td>
</tr>
</tbody>
</table>

Table 4.8: Mean comparison (Reactions): L1 Groups and speakers’ proficiency
Dependent variable: intelligibility.

The intelligibility test was designed to quantify the accuracy with which the listeners could transcribe what they heard on the Intelligibility test that was developed by the researcher. Sixteen sentences were extracted from the 8 speech samples that were used with the Evaluative Reactions scale. The number of words that exactly matched the correct transcriptions of the sentences (Appendix A) was calculated for each listener. Unlike the evaluative reaction scores, note that the intelligibility scores were indicated by means of percentage. Percentage was utilized rather than actual counts of words that were correctly transcribed, because each test sentence had a different number of words, as sentence length ranged from 11 to 20 of words, with the average number of 14.75 words, as shown in Table 4.9. Extracting sentences that contained a similar number of words from all speech samples was difficult, owing to the various rates of speech, via words per minute (wpm), of the speakers. Speaker 7 (N♂) spoke at the rate of 180 wpm, while Speaker 2 (B♀) at 36 wpm.

Table 4.9 represents the descriptive statistics of the intelligibility test results. Overall, NNSs were found to be more intelligible than NSs. The intelligibility of NNSs ranged from approximately 92.9% to 96.9%, in comparison to the range of 88.2% and 74.6% for NSs. Due to the almost perfect intelligibility of NNSs, their standard deviations were smaller than those of NSs.
The L1 group comparison of mean scores on the Intelligibility test is indicated below (Table 4.10). As could be predicted, the L1 American English and the L1 Hindi groups transcribed English sentences the most accurately. Interestingly, though a majority of the speakers were Japanese, the L1 Mandarin group had a higher transcription mean score than the L1 Japanese group. All the L1 groups show a negative skewness in the score distribution, which means their medians were higher than their means. The extremely negative skewness and high kurtosis (leptokurtic) of the L1 American-English group indicate that the shape of distribution is far from the normal curve and the scores concentrated to the upper limit, which is an indication of the ceiling effect. Contrary to the study of Matsuura et al. (1999) in which the intelligibility tests were too difficult for the participants to produce score variance, the intelligibility test utilized in this study might have been too easy for the L1 American English group to generate score variance.
<table>
<thead>
<tr>
<th>Native Language</th>
<th>Mean</th>
<th>N</th>
<th>SD</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>American English</td>
<td>97.68</td>
<td>22</td>
<td>2.80</td>
<td>87.6</td>
<td>100.00</td>
<td>-2.49</td>
<td>7.55</td>
</tr>
<tr>
<td>Hindi</td>
<td>94.49</td>
<td>20</td>
<td>3.19</td>
<td>86.2</td>
<td>100.00</td>
<td>-0.47</td>
<td>1.24</td>
</tr>
<tr>
<td>Mandarin</td>
<td>88.34</td>
<td>20</td>
<td>6.38</td>
<td>72.2</td>
<td>98.75</td>
<td>-1.12</td>
<td>1.62</td>
</tr>
<tr>
<td>Japanese</td>
<td>85.32</td>
<td>22</td>
<td>6.88</td>
<td>69.4</td>
<td>96.28</td>
<td>-0.81</td>
<td>1.02</td>
</tr>
<tr>
<td>Total</td>
<td>91.46</td>
<td>84</td>
<td>7.10</td>
<td>69.4</td>
<td>100.00</td>
<td>-1.10</td>
<td>1.16</td>
</tr>
</tbody>
</table>

4.10: Mean comparison among L1s (Intelligibility)

The means of the intelligibility scores of the four L1 groups on different proficiency levels of speakers are exhibited in Table 4.11. What strikes one as significant in the table below is that the NSs’ intelligibility among L1 American English and Hindi listeners was not very different from that of other speech, which was very high (more than 90%). L1 Mandarin and Japanese participants, on the other hand, especially Japanese listeners, did not transcribe native speech correctly. An English listening proficiency difference is surmised to exist among the listeners, which is consistent with the differences in TOEFL scores, reported previously.
## Native Languages

<table>
<thead>
<tr>
<th>Native Languages</th>
<th>Native Speaker Intelligibility</th>
<th>Advanced Speaker Intelligibility</th>
<th>Intermediate Speaker Intelligibility</th>
<th>Beginning Speaker Intelligibility</th>
</tr>
</thead>
<tbody>
<tr>
<td>American English</td>
<td>Mean</td>
<td>96.94</td>
<td>99.29</td>
<td>97.99</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>22</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>Std. Deviation</td>
<td>2.81</td>
<td>1.59</td>
<td>3.143</td>
</tr>
<tr>
<td>Hindi</td>
<td>Mean</td>
<td>91.21</td>
<td>94.29</td>
<td>96.54</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Std. Deviation</td>
<td>6.90</td>
<td>3.91</td>
<td>2.98</td>
</tr>
<tr>
<td>Mandarin</td>
<td>Mean</td>
<td>76.86</td>
<td>92.46</td>
<td>95.12</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Std. Deviation</td>
<td>10.24</td>
<td>4.94</td>
<td>2.93</td>
</tr>
<tr>
<td>Japanese</td>
<td>Mean</td>
<td>61.05</td>
<td>89.31</td>
<td>94.38</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>22</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>Std. Deviation</td>
<td>15.09</td>
<td>9.77</td>
<td>3.82</td>
</tr>
<tr>
<td>Total</td>
<td>Mean</td>
<td>81.39</td>
<td>93.86</td>
<td>96.0149</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>84</td>
<td>84</td>
<td>84</td>
</tr>
<tr>
<td></td>
<td>Std. Deviation</td>
<td>17.22</td>
<td>6.91</td>
<td>3.49141</td>
</tr>
</tbody>
</table>

Table 4.11: Mean comparison (Intelligibility): L1 groups and speakers' proficiency

### Correlation Analysis

Bivariate correlation analysis was chosen to investigate the magnitudes and directions of the relationship between two variables. In order to measure the relationship between two metric variables (i.e., ordinal, interval, and ratio), the assumptions of linearity and homoscedasticity needs to be confirmed (Warmbrod, 2000). Scatterplots were constructed for metric data and examined for non-linearity. Since no scatterplots indicated the threat of non-linearity and heteroscedasticity (see Appendix H), the correlation analyses were conducted. The results of analyses are shown in Table 4.12.
Table 4.12: Correlations (associations) among variables

<table>
<thead>
<tr>
<th></th>
<th>X1</th>
<th>X2</th>
<th>X3</th>
<th>X4</th>
<th>X5</th>
<th>Y1</th>
<th>Y2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native Language</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(X1) Cramer’s V</td>
<td>1.00</td>
<td>.49</td>
<td>.40</td>
<td>.28</td>
<td>.89</td>
<td>.86</td>
<td></td>
</tr>
<tr>
<td>Speakers’ Proficiency Level (X2)</td>
<td>Kendall’s tau</td>
<td>1.00</td>
<td></td>
<td></td>
<td>.93</td>
<td>-.62</td>
<td></td>
</tr>
<tr>
<td>Listeners’ Gender (X3)</td>
<td>Kendall’s tau</td>
<td>1.00</td>
<td>.30</td>
<td>.34</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Familiarity to Japanese English (X4)</td>
<td>Kendall’s tau</td>
<td>1.00</td>
<td>.15</td>
<td>.01</td>
<td>-.21</td>
<td>.05</td>
<td></td>
</tr>
<tr>
<td>Exposure to Foreigners (X5)</td>
<td>Kendall’s tau</td>
<td></td>
<td>1.00</td>
<td>-.14</td>
<td>-.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evaluative Reactions (Y1)</td>
<td>Pearson r</td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
<td>.00</td>
<td></td>
</tr>
<tr>
<td>Intelligibility (Y2)</td>
<td>Pearson r</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
<td></td>
</tr>
</tbody>
</table>

The magnitudes of relationships are described in the terms of Davis (1971), which is reiterated below.
<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>.70 to .99</td>
<td>Very strong association</td>
</tr>
<tr>
<td>.50 to .69</td>
<td>Substantial association</td>
</tr>
<tr>
<td>.30 to .49</td>
<td>Moderate association</td>
</tr>
<tr>
<td>.10 to .29</td>
<td>Low Association</td>
</tr>
<tr>
<td>.01 to .09</td>
<td>Negligible Association</td>
</tr>
</tbody>
</table>

Table 3.4: Magnitudes of association  (reiterated)

Participants’ native language (L1) was moderately associated with both gender (Cramer’s V = .49) and their familiarity with Japanese English (V = .40). L1 was also related to the participants’ degree of exposure to foreigners, but the relationship was low (V = .28). A very strong association was detected between L1 and evaluative reactions (V = .89), and L1 and intelligibility (V = .86).

A positive and very strong association (Kendall’s tau = .93) was found between the speakers’ proficiency level and the evaluative reaction scores. In other words, the higher the speaker’s English proficiency level was, the more favorably his/her speech was evaluated. The speakers’ proficiency level was also, but negatively in this case, correlated substantially with speech intelligibility (tau = -.62). The higher the speaker’s proficiency level, the less intelligible was his/her speech. This result seems to be strongly related to the native speakers’ low intelligibility among the L1 Japanese and Mandarin listeners.
Listeners’ gender was moderately associated both with their familiarity with Japanese English ($V = .30$) and their exposure to other NNSs of English ($V = .34$). In order to calculate the relationship between gender and the metric dependent variables, a point biserial correlation coefficient ($r_{pb}$) was calculated. The point biserial coefficient is a Pearson product-moment coefficient, utilized to describe the relationship between a dichotomous nominal variable and an interval or ratio variable (Warmbrod, 2000). Gender was coded as follows: female = 0 and male = 1. A low negative relationship ($r_{pb} = -.21$) was observed between listeners’ gender and their evaluative reactions. That is, female listeners evaluated speech more positively than male listeners did. The relationship between gender and intelligibility was positive and negligible ($r_{pb} = .05$). The intelligibility test scores of male listeners were higher than those of females, but the difference was minute.

The Kendall’s tau statistic between the listeners’ familiarity with Japanese English and their exposure to other NNSs indicates a low positive association ($\tau = .15$). A listener who was familiar with Japanese English was slightly more likely to have been exposed to other NNSs than a person who was not familiar with Japanese English was. The relationship between the familiarity with Japanese English and evaluative reactions was positive and negligible ($\tau = .01$). A low negative association ($\tau = -.14$) existed between the familiarity with Japanese English and Intelligibility scores. Participants with higher familiarity with Japanese English tended to transcribe the extracted sentences less accurately than others. This result is explicable when the fact that the familiarity with Japanese English was highest in the L1 Japanese group, whose mean TOEFL score was the lowest among the nonnative L1 groups.
The degree of exposure to foreigners had a low negative correlation with evaluative reactions (tau = -.14). Participants who were more exposed to foreigners judged speech less favorably. The exposure to foreigners also had a negative relationship with intelligibility, but the magnitude was negligible (tau = -.01).

Unexpectedly, very little correlation was observed between intelligibility scores and evaluative reactions (r = .00). Participants evaluated intelligible speech no more favorably than unintelligible speech.

In sum, only two independent variables indicated substantial to very strong associations with the two dependent variables. Listeners’ L1 was positively and very strongly related to both Evaluative Reactions and Intelligibility. Speakers’ proficiency level was positively and very strongly associated with Reactions, but negatively and substantially associated with Intelligibility. Among independent variables, moderate associations were found between listeners’ L1 and gender, between gender and familiarity with Japanese English, between gender and exposure to foreigners, and L1 and familiarity with Japanese English. Some of these seemingly unexpected associations can be explained by the almost opposite gender ratios among the L1 American English and the L1 Hindi participants.

One extraneous variable that could not be controlled by the research design has remained: Age. The listeners’ age was planned to be constant (20-32), but turned out to have a wider range: from 19 to 66. Thus correlation between the listeners’ age and dependent variables was calculated. The Person product moment coefficients indicated a negligible and positive correlation (r = .03) between age and evaluative reactions, and a low and negative correlation (r = -.14) between age and intelligibility. A tentative
stepwise regression analysis of intelligibility, with age as well as other independent variables, indicated that age was not a significant (.14) predictor.

Regression Analysis

The purpose of the regression analyses was to predict the two dependent variables, (1) evaluative reactions and (2) intelligibility, by using the information available on five selected independent variables: speakers’ proficiency, listeners’ native language, gender, familiarity with Japanese English, and exposure to other nonnative English. In other words, each independent variable was examined in terms of the amount of its contribution to the explanation of the variance in each dependent variable.

One independent variable, the speakers’ English proficiency, could not be treated in the same way as the other independent variables, since it was an attribute of the speakers, instead of the listeners’. Accordingly, a regression equation was sought for each speaker’s proficiency level and then compared with one another.

The basic relationship dealt with in any regression analysis is the linear correlation between the metric dependent and independent variables (Hair, Anderson, Tatham, and Black, 1998). Regression analysis is robust and can tolerate ordinal variables (Miller, Acron, Fullerton, and Maltby, 2002), but not nominal variables.

Nominal data, therefore, must be transformed before being put into regression. In this study, the listeners’ native language (L1) needed the transformation. When dummy-coding a non-metric variable with \( k \) categories, \( k-1 \) dummy variables are needed. L1 was dummy-coded into three dummy variables (D1, D2, and D3). In D1, L1 American
English participants were given 1 while all others were given 0. In D2, L1 Hindi members were given 1 while others received 0. In D3, L1 Mandarin listeners were given 1 while others were represented by 0. This indicator coding, in which categories are assigned either 1 or 0, allows the direct group comparison, for the coefficients are in the same units as the dependent variable (Hair, Anderson, Tatham, and Black, 1998). The L1 Japanese group received 0 in all dummy codes, and functioned as a reference category. Gender was also a nominal variable, but it was dichotomous. Dummy-coding of gender had already been executed in the initial data entry. Since the female listeners received 0, they served as the reference category.

A correlation matrix of variables was compiled to examine multicollinearity (Table 4.13). Note that all the variables are treated as metric data and the statistic in use is the Pearson-product moment correlation coefficient. Due to the positive and moderate relationship \(r = .48\) found between two dummy variables, Hindi-ness and gender, their collinearity was examined. Collinearity statistics for Hindi-ness and gender were tolerance = .77 and VIF = 1.3. A high value (near 1.0) in tolerance and a low value (lower than 10) in VIF indicate that multicollinearity is not a concern (Gliem, 2001). Thus, multicollinearity was concluded not to be a concern in this regression analysis.

Another important assumption for a multiple regression analysis was the independence of residuals, which means “error associated with any observation is not correlated with error associated with any other observations” (Gliem, 2001, p.4). Since the value of Durbin-Watson: \(d = 2.08\) manifests the independence of residuals, multiple regression analyses were conducted.
A series of multiple regression analyses was conducted to test the research hypotheses. The hypotheses to be tested were as follows:

**H1**: The listeners’ evaluative reactions to Japanese English will be predicted by:

1. The listener’s L1 (i.e., American English, Hindi, Chinese and Japanese)
2. The speaker’s proficiency

**H2**: The intelligibility of Japanese English will be predicted by:

1. The listener’s L1
2. The speaker’s proficiency
**H3:** The listeners’ evaluative reactions to Japanese English will be predicted by:

1. The listener’s gender
2. The listeners’ familiarity with Japanese English
3. The listeners’ exposure to foreigners

**H4:** The intelligibility of Japanese English will be predicted by:

1. The speaker’s and listener’s gender
2. The listeners’ familiarity with Japanese English
3. The listeners’ exposure to foreigners

The following section describes the regression analysis of the evaluative reactions first, followed by that of the intelligibility.

**Evaluative reactions.**

*Initial Regression Analysis of Evaluative Reactions*

The result of a regression analysis with the simultaneous model of evaluative reactions on all the selected independent variables is displayed in Table 4.14. Adjusted $R^2$ indicates that this model accounts for 20.4% of the variance in evaluative reactions. The null hypothesis ($H_0: R^2 \neq 0$) was rejected ($F = 4.55, p < .01$). The model can predict evaluative reactions.

The coefficient of determination $R^2$ indicates the portion of the variance accounted for by an independent variable. The small $R^2$ values, as well as the insignificant $p$ values, for listeners’ gender, familiarity with Japanese English, exposure to foreigners, and Mandarin-ness, imply the probability that these variables were not effective predictors of evaluative reactions. Therefore, a stepwise regression with the
entry level of probability of $F < .05$ was attempted in order to investigate the contribution of each independent variable.

<table>
<thead>
<tr>
<th>Variables</th>
<th>$R^2$</th>
<th>$B$</th>
<th>$t$</th>
<th>Sig. (p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D1 (American-ness)</td>
<td>.12</td>
<td>4.49</td>
<td>2.52</td>
<td>.01</td>
</tr>
<tr>
<td>D2 (Hindi-ness)</td>
<td>.19</td>
<td>3.82</td>
<td>1.74</td>
<td>.09</td>
</tr>
<tr>
<td>D3 (Mandarin-ness)</td>
<td>.03</td>
<td>2.75</td>
<td>1.40</td>
<td>.17</td>
</tr>
<tr>
<td>X2 (Gender)</td>
<td>.05</td>
<td>-.49</td>
<td>-.336</td>
<td>.74</td>
</tr>
<tr>
<td>X3 (Familiarity with JE)</td>
<td>.00</td>
<td>.01</td>
<td>.03</td>
<td>.98</td>
</tr>
<tr>
<td>X4 (Exposure to foreigners)</td>
<td>.01</td>
<td>-.21</td>
<td>-.41</td>
<td>.68</td>
</tr>
<tr>
<td>(Constant)</td>
<td>48.08</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Standard Error = 5.72
Adjusted $R^2 = .20$
For model: $F = 4.55; p = .001$

Table 4.14: Multiple regression results on Evaluative Reactions (Simultaneous)

Further Regression Analysis of Evaluative Reactions

Further regression analysis was conducted with the stepwise entry method in order to specify greater contributing predictors of evaluative reactions. Stepwise regression selects one independent variable from a group of variables that makes the
largest contribution to $R^2$ of the model at each step, until there is no significant variable left (Gliem, 2001). Table 4.15 displays the regression results. Two models were produced. The second model with two variables is discussed here. The model is significant ($F = 12.58, p < .001$) and can explain 22% of the variance in the dependent variable. The selected predictors are D2: Hindi-ness and D1: American English-ness. The third dummy variable, Mandarin-ness, was excluded since it was insignificant. $R^2$ Change, also called Semipartial (Part) Correlation, represents the correlation between an independent variable and the dependent variable when the linear effects of the other independent variables are removed from the independent variable under consideration (Gliem, 2001). $R^2$ Change of D1: American-ness shows that the unique variance explained when the effect of D2: Hindi-ness was controlled was only 4%.

<table>
<thead>
<tr>
<th>Variables</th>
<th>$R^2$</th>
<th>$R^2$ change</th>
<th>$B$</th>
<th>$t$</th>
<th>Sig. ($p$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>D2 (Hindi-ness)</td>
<td>.20</td>
<td>.20</td>
<td>-5.52</td>
<td>-3.58</td>
<td>.00</td>
</tr>
<tr>
<td>X1 D1 (American-ness)</td>
<td>.24</td>
<td>.04</td>
<td>3.16</td>
<td>2.12</td>
<td>.04</td>
</tr>
<tr>
<td>Constant</td>
<td></td>
<td></td>
<td>48.42</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Standard Error = 5.67
Adjusted $R^2 = .23$
For model: $F = 12.58; p < .001$

Table 4.15: Further regression analysis of Evaluative Reactions (Stepwise)
The regression equation of the model for the evaluative reactions is:

\[ Y (\text{Evaluative Reactions score}) = 48.42 \text{ (intercept)} - 5.52 \text{ (if Hindi = L1)} + 3.16 \text{ (if American English = L1)} \]

The interpretation of the regression model is as follows. Constant (48.4) is the expected mean score of Japanese participants’ evaluative reactions. If a listener is an L1 Hindi group member, the evaluative score is predicted to drop by 5.5 points. An American listener is expected to evaluate the speech 3.2 points higher than a Japanese listener would. The evaluation by an L1 Mandarin group listener would not be significantly different from that of a Japanese listener. Therefore, the first element of Hypothesis 1 (H1) is supported. The listener’s native language does predict his/her evaluative reactions. On the other hand, none of the rival variables, namely listeners’ gender, familiarity with Japanese English, and exposure to foreigners, can predict the evaluative reactions. Hypothesis 3 was rejected.

Does the L1 predictability (adjusted coefficient of determination: adjusted R\(^2\) = .22) of evaluative reactions, however, hold among different English proficiency levels of speakers? In the following section, a set of multiple regression analyses is administered to investigate the evaluative reactions to the four proficiency levels of the speakers.

### Regression Analysis of Evaluative Reactions to Different Proficiency Levels

- **Evaluative Reactions to Native Speakers**

  A stepwise regression analysis was conducted with the entry level of probability of F < .05. Three independent variables (American-ness, Hindi-ness, and Mandarin-ness)
were entered. The final model (Table 4.16) with three variables was significant ($F = 11.10, p < .001$). The adjusted $R^2$ indicates that 27% of the variance in reactions to the native speech can be accounted for by the three dummy variables. When a Japanese listener’s Evaluative Reactions score is 68.4, that of a Hindi-speaking listener will be 3.4 points lower. American English and L1 Mandarin group listeners’ reactions will be 4 or 5 points higher than the Japanese. In short, the evaluative reactions to the native speakers can be predicted only by listeners’ L1. Hypothesis 1 was supported, while Hypothesis 3 was rejected.

<table>
<thead>
<tr>
<th>Variables</th>
<th>$R^2$</th>
<th>$R^2$ change</th>
<th>B</th>
<th>t</th>
<th>Sig. (p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>D2 (Hindi-ness)</td>
<td>.19</td>
<td>.19</td>
<td>-3.38</td>
<td>-1.91</td>
<td>.06</td>
</tr>
<tr>
<td>D1 (American-ness)</td>
<td>.24</td>
<td>.05</td>
<td>5.68</td>
<td>3.29</td>
<td>.00</td>
</tr>
<tr>
<td>D3 (Mandarin-ness)</td>
<td>.29</td>
<td>.06</td>
<td>4.59</td>
<td>2.59</td>
<td>.01</td>
</tr>
<tr>
<td>Constant</td>
<td></td>
<td></td>
<td>68.43</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Standard Error = 5.74  
Adjusted $R^2 = .27$  
For model: $F = 11.10; p < .001$

Table 4.16: Regression analysis of Evaluative Reactions to native speakers (Stepwise)
Evaluative Reactions to Advanced Japanese Speakers

A stepwise regression analysis was administered for the Evaluative Reactions to advanced Japanese speakers (Table 4.17). The regression entered only one variable. The simple linear regression with D2 (Hindi-ness) as the independent variable was significant (F = 24.27, p < .001). The adjusted R suggests that 23% of the variance in evaluative reactions to advanced Japanese speakers can be explained only by whether the listeners’ native language is Hindi or not. Hypothesis 1 was supported and Hypothesis 3 was rejected.

Compared with a Japanese listener, who gives 50 points out of 80 to an advanced Japanese speech sample, an L1 Hindi group listener is likely to offer 10 points less. Regression analysis indicates that the L1 American English and L1 Mandarin listeners would not evaluate advanced Japanese speakers differently from the L1 Japanese listeners.

<table>
<thead>
<tr>
<th>Variables</th>
<th>R²</th>
<th>B</th>
<th>t</th>
<th>Sig. (p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1</td>
<td></td>
<td>.23</td>
<td>-10.14</td>
<td>-4.93</td>
</tr>
<tr>
<td>D2 (Hindi-ness)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>50.40</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Standard Error = 8.03
Adjusted R² = .23
For model: F = 24.27; p < .001

4.17: Regression analysis of Evaluative Reactions to advanced speakers
Evaluative Reactions to Intermediate Japanese Speakers

A stepwise regression with the entry level of probability of $F < .05$ again produced only one model (Table 4.18). D2: Hindi-ness was the only independent variable entered to predict the evaluative reactions to intermediate Japanese speakers of English. The model was significant ($F = 7.14, p < .01$), but the adjusted $R^2$ was very small. Only 7% of the variance in reaction to intermediate Japanese speech can be explained by listeners’ Hindi-ness. Nonetheless, since only listeners’ L1 can predict the evaluative reactions, Hypothesis 1 was supported and Hypothesis 3 was rejected. Interpretation: If a listener’s native language is Hindi, s/he is likely to evaluate the speech 5.7 points lower than a Japanese listener would (constant = 45.4 points).

<table>
<thead>
<tr>
<th>Variables</th>
<th>$R^2$</th>
<th>$B$</th>
<th>$t$</th>
<th>Sig. ($p$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1 D2 (Hindi-ness)</td>
<td>.08</td>
<td>-5.77</td>
<td>-2.67</td>
<td>.01</td>
</tr>
<tr>
<td>Constant</td>
<td></td>
<td>45.39</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Standard Error = 8.42
Adjusted $R^2 = .07$
For model: $F = 7.14; p < .01$

Table 4.18: Regression analysis of Evaluative Reactions to intermediate speakers
Evaluative Reactions to Rudimentary Japanese Speakers

Another stepwise regression was conducted in order to predict the evaluative reactions to beginning-level Japanese speakers of English (Table 4.19). Only one variable (American-ness) was entered and the model was significant ($F = 7.43, p < .01$). The adjusted $R^2$ indicates, however, that the knowledge of the listeners’ native language can only account for 7% of the variance of evaluative reactions to beginning-level Japanese speech. American listeners are expected to evaluate the beginning-level speech 5.8 points higher than Japanese listeners would. L1 Hindi and L1 Mandarin listeners would not evaluate the speech differently from L1 Japanese listeners. Hypothesis 1 was supported while Hypothesis 3 was rejected.

<table>
<thead>
<tr>
<th>Variables</th>
<th>$R^2$</th>
<th>$B$</th>
<th>$t$</th>
<th>Sig. ($p$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1 D1</td>
<td>.08</td>
<td>-5.83</td>
<td>2.73</td>
<td>.01</td>
</tr>
<tr>
<td>(American-ness)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td></td>
<td>27.98</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Standard Error = 8.63  
Adjusted $R^2 = .07$  
For model: $F = 7.43; p < .001$

Table 4.19: Regression analysis of Evaluative Reactions to rudimentary speakers
Summary of the evaluative reactions analyses

Research Hypothesis 3 (H3), which was “A listener’s evaluative reactions to Japanese English will be predicted by his/her gender, familiarity with Japanese English, and exposure to foreigners,” was rejected for overall tendency. These rival independent variables did not explain the variance in the evaluative reactions to native and nonnative speech. On the other hand, Research Hypothesis 1 (H1), “Reactions will be predicted by the listener’s L1 and the speaker’s proficiency,” was supported. The main independent variable, L1, turned out to be a predictor, while the constants in regression equations constantly decreased as the speaker’s proficiency became lower. Overall, American listeners tended to evaluate speech more highly than Japanese listeners, while L1 Hindi group listeners tended to judge speech more critically. L1 Mandarin group listeners evaluated speech slightly more favorably than Japanese, but the difference was not significant.

When the evaluative reactions to speakers with different proficiency levels were examined, only two variables (American-ness and Hindi-ness) were found to be predictors of the reactions. Twenty-seven percent of the variance in reactions to NSs were explained by the listeners’ American and Hindi-ness. For reactions to advanced and intermediate Japanese speakers, only the listeners’ Hindi-ness mattered, explaining 22% for the advanced and 7% for the intermediate Japanese speakers. For reactions to the beginning-level Japanese speakers, American-ness was the only predictor, but the explained variable was also as small as 7%.
Intelligibility.

Initial Regression Analysis of Intelligibility

A simultaneous regression with all the independent variables produced a significant model ($F = 13.64, p < 0.001$) with the adjusted $R^2 = .48$. However, again, many variables did not contribute to the regression. When a stepwise regression was attempted, four models were produced. The independent variables entered into the final model were the three L1 dummy variables and gender. The listeners’ familiarity with Japanese and exposure to foreigners were excluded. The model was significant ($F = 20.83, p < .001$), and explained 49% of the intelligibility variance (Table 4.20). Therefore, Hypothesis 2 was supported. Intelligibility can be predicted based on the listeners’ native language. L1 American English, Hindi, and Mandarin listeners will obtain significantly better scores on the intelligibility test than L1 Japanese listeners. Hypothesis 4 was partially supported. Among the rival independent variables, only gender can predict the variance in intelligibility scores. The intelligibility scores of a female listener will be 2.4 points higher than that of a male listener.
<table>
<thead>
<tr>
<th>Variables</th>
<th>$R^2$</th>
<th>$R^2$ change</th>
<th>$B$</th>
<th>$t$</th>
<th>Sig. (p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1 D1 (American-ness)</td>
<td>.28</td>
<td>.28</td>
<td>12.69</td>
<td>8.23</td>
<td>.00</td>
</tr>
<tr>
<td>X1 D2 (Hindi-ness)</td>
<td>.47</td>
<td>.19</td>
<td>10.66</td>
<td>6.05</td>
<td>.00</td>
</tr>
<tr>
<td>X1 D3 (Mandarin-ness)</td>
<td>.49</td>
<td>.02</td>
<td>3.19</td>
<td>2.03</td>
<td>.05</td>
</tr>
<tr>
<td>X2 Gender</td>
<td>.51</td>
<td>.02</td>
<td>-2.38</td>
<td>-1.86</td>
<td>.06</td>
</tr>
<tr>
<td>Constant</td>
<td></td>
<td></td>
<td>88.24</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Standard Error = 5.08  
Adjusted $R^2 = .49$

For model: $F = 20.82; p < .001$

Table 4.20: Regression analysis of Intelligibility (Stepwise)

A possibility of a specification error emerged at this point. One of the specification errors in regression analysis assumptions is “No relevant independent variables have been excluded” (Gliem, 2001, p.4). In the comparison of intelligibility score means across listeners’ L1 (Table 4.10), the order conforms to the listeners’ mean TOEFL score order. Although the listener’s TOEFL score was not originally included as an independent variable, Cramer’s V statistics indicates it is very strongly associated with listener’s native language ($V = -.73$). What is suspected is that the variance in intelligibility was caused not by the listeners’ L1, but by their TOEFL scores, which supposedly represent their English proficiency. Therefore, in further analyses of intelligibility, the listener’s TOEFL score was examined as an additional rival independent variable to the regression.
As stated in the section on data description, nonnative listeners’ TOEFL scores were blocked into 5 ranks. 5 = 650 points and above, 4 = 600–650 points, 3 = 550–600 points, 2 = 500–550 points, and 1 = 500 points and below. For native participants, rank 6 was created. Even though the blocked TOEFL score was ordinal data, it was treated as interval in the following regression analysis (Miller, Acron, Fullerton, and Maltby, 2002).

The result of the regression analysis of intelligibility (Y) and selected independent variables (Ds and Xs) are exhibited in Table 4.21. The full model of stepwise regression with an entry level of probability of F < .05 entered 4 variables: TOEFL score, American-ness, gender, and Hindi-ness. The full model is significant (F = 25.64, p < .001). Fifty-four percent of the total variance can be accounted for by the linear combination of the four independent variables. The Durbin-Watson statistic (d = 2.18) shows the independence of residuals. The other variables, namely Mandarin-ness, familiarity with Japanese English, and exposure to foreigners, were excluded from the regression, due to their being insignificant.

<table>
<thead>
<tr>
<th>Variables</th>
<th>R²</th>
<th>R² change</th>
<th>B</th>
<th>t</th>
<th>Sig. (p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOEFL score rank</td>
<td>.46</td>
<td>.46</td>
<td>2.52</td>
<td>3.74</td>
<td>.00</td>
</tr>
<tr>
<td>D1 (American-ness)</td>
<td>.51</td>
<td>.05</td>
<td>6.91</td>
<td>4.34</td>
<td>.00</td>
</tr>
<tr>
<td>X2 (Gender)</td>
<td>.54</td>
<td>.03</td>
<td>-2.74</td>
<td>-2.25</td>
<td>.03</td>
</tr>
<tr>
<td>D2 (Hindi-ness)</td>
<td>.57</td>
<td>.02</td>
<td>4.48</td>
<td>2.05</td>
<td>.04</td>
</tr>
<tr>
<td>Constant</td>
<td></td>
<td></td>
<td>79.07</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Standard Error = 4.80
Adjusted R² = .54
For model: F = 25.64, p < .001

Table 4.21: Regression analysis of Intelligibility with TOEFL
Based on the partial regression coefficients (B), the regression equation of the model becomes:

\[
Y \text{ (intelligibility score)} = 79.07 \text{ (intercept)} \\
+ 2.52 \text{ (TOEFL rank)} \\
+ 6.91 \text{ (if Hindi = L1)} \\
- 2.74 \text{ (if male = gender)} \\
+ 4.48 \text{ (if American English = L1)}
\]

This equation is interpreted as follows: A listener’s intelligibility score will be 79.07%, when each independent variable is zero, representing Japanese female listeners. When the TOEFL rank (from 1 to 6) rises by one, add 2.52% to the intercept (79.07%). If the listener is an L1 Hindi group member, add 6.91%. If the listener is a male, reduce the score by 2.74%. If the listener is an American listener, add 4.48% to the intelligibility score.

Both Hypothesis 2 and Hypothesis 4 were partially supported by the regression results when TOEFL score ranks were considered. Among L1 differences, Hindi-ness and American-ness can predict intelligibility, but Mandarin-ness cannot. L1 Mandarin listeners will not score better or worse than Japanese listeners. Among the rival independent variables, gender can predict intelligibility, but familiarity with Japanese English and exposure to foreigners cannot. The newly added variable, listeners’ TOEFL score rank, turned out to be the most powerful predictor. After removing the effects of TOEFL score rank and gender, the unique proportion of the variance explained by the listeners’ L1 was the sum of \( R^2 \) change of D1 and D2.

\[
(.05) + (.02) = .07
\]

Approximately 7% of the variance in the dependent variable can be explained by L1.
Again, another question arises from this overall tendency. Is this tendency still true for different proficient levels? A series of regression analyses was implemented to examine the predictive power of independent variables in the next section.

**Further Regression Analysis of Intelligibility of Speakers with Different Proficiency Levels**

- Intelligibility of Native Speakers

When a stepwise regression was conducted with the entry level of probability of $F < .05$, the full model contained 5 independent variables: TOEFL score, Hindi-ness, Gender, American-ness, and Mandarin-ness (Table 4.22). The model is significant ($F = 41.32, p < .001$), and can account for 71% of the variance in native speech intelligibility. The $R^2$ change suggests that the TOEFL score again is the most powerful predictor. Hindi-ness turned out to be the second most powerful predictor.

Not only the L1 Hindi and American English group members, but also the L1 Mandarin, tend to understand native speech significantly better than L1 Japanese members, even beyond what the TOEFL score difference alone can explain. This means that the L1 Hindi and L1 Mandarin groups with the same TOEFL rank will understand NSs better than the L1 Japanese group. The unique contribution of the listener’s L1 to the regression is 14%.

Again, for native speech intelligibility, gender was selected as a predictor. The negative value means that a male listener is expected to score 5.5 % lower than a female listener.
Table 4.22: Regression analysis of Intelligibility of native speakers

- Intelligibility of Advanced Japanese speakers

The stepwise regression with the entry level of probability of $F < .05$ produced only one model. Table 4.23 indicates the results. The model is significant ($F = 37.44, p < .001$), but explains 31% of the variance, which is significantly smaller than what percentage it explained in the variance of native intelligibility. Interestingly, only the listeners’ TOEFL rank can explain the variance in intelligibility for advanced Japanese speakers. All the originally selected independent variables failed to predict the intelligibility for these speakers. Therefore, both Hypotheses 2 and 4 were rejected. However, it is noteworthy that no significant difference was found among four L1 groups for understanding advanced Japanese speech. Even though the L1 Japanese group could
not transcribe native speech as accurately as the other groups, they could transcribe what advanced Japanese speakers said as successfully as others.

<table>
<thead>
<tr>
<th>Variables</th>
<th>R²</th>
<th>B</th>
<th>t</th>
<th>Sig. (p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOEFL score rank</td>
<td>.31</td>
<td>2.84</td>
<td>39.45</td>
<td>.000</td>
</tr>
<tr>
<td>Constant</td>
<td></td>
<td>81.77</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Standard Error = 5.76  
Adjusted R² = .31  
For model: F = 37.44; p < .001

Table 4.23: Regression analysis of Intelligibility of advanced speakers

- Intelligibility of Intermediate Japanese Speakers

Table 4.24 displays the stepwise regression analysis of Intermediate speakers’ intelligibility. Similar to the advanced speaker intelligibility, this regression also produced only one model with one variable: TOEFL score rank. This time, however, the variance that the TOEFL score can explain was even smaller at 19%. Therefore, 81% of variance in Intermediate speech intelligibility remains unexplained. Furthermore, the partial regression coefficient (B) indicates that the increase in the intelligibility score when the TOEFL rank increases by one unit gradually becomes smaller as the speakers’ proficiency
level becomes lower. The coefficient was 4.84 with NSs, 2.84 with advanced Japanese speakers, and here we see it at 1.14 with intermediate Japanese speakers. The predictability of intelligibility by TOEFL score rank was higher in native speech than in Japanese speech.

<table>
<thead>
<tr>
<th>Variables</th>
<th>R²</th>
<th>B</th>
<th>t</th>
<th>Sig. (p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOEFL score rank</td>
<td>.20</td>
<td>1.14</td>
<td>4.52</td>
<td>.000</td>
</tr>
<tr>
<td>Constant</td>
<td></td>
<td>91.14</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Standard Error = 3.14  
Adjusted R² = .19  
For model: F = 20.42; p < .001

Table 4.24: Regression analysis of Intelligibility of intermediate speakers

- Intelligibility of Rudimentary Japanese Speakers

The stepwise regression with the probability of F < .05 for the beginning-level Japanese speakers again produced one model with one variable (Table 4.25). But this time, the variable was not TOEFL score rank, but Mandarin-ness. L1 Mandarin listeners transcribed beginning-level Japanese speech 7.4% less accurately than L1 Japanese listeners. The intercept (constant) shows the almost perfect intelligibility of the
beginning-level Japanese speakers (96%). Because of the ceiling effect, little variance seemed to be observable.

<table>
<thead>
<tr>
<th>Variables</th>
<th>R²</th>
<th>B</th>
<th>t</th>
<th>Sig. (p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>D3 Mandarin-ness</td>
<td>.12</td>
<td>-7.42</td>
<td>-3.33</td>
<td>.001</td>
</tr>
<tr>
<td>Constant</td>
<td>96.33</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Standard Error = 8.69
Adjusted R² = .11
For model: F = 11.11; p < .005

Table 4.25: Regression analysis of Intelligibility of rudimentary speakers

Summary of intelligibility analyses

Unlike the evaluative reactions, intelligibility turned out to be closely related to TOEFL scores (English proficiency) of listeners. Therefore, the regression analysis was conducted including TOEFL score ranks as an independent variable. The regression results indicated that the TOEFL score rank was the most powerful predictor of speech intelligibility in general. Listeners’ L1 was the second predictor and uniquely explained 7% of the intelligibility variance.
When the speakers’ proficiency was considered, the intelligibility across L1 group listeners became more complex. In the intelligibility of native speech, the linear combination of listeners’ TOEFL score rank, L1 and gender worked as predictors and explained as much as 71% of the variance. Surprisingly, for the intelligibility of advanced Japanese speakers, TOEFL was the only significant predictor and accounts for 31% of the variance. For the intermediate Japanese speakers’ intelligibility, TOEFL was also the only interpreter but can explain only 19% of the variance. For the beginning-level Japanese speakers, TOEFL was no longer a significant predictor of intelligibility variance, but only Mandarin-ness explained 11% of the variance. L1 Mandarin listeners understood beginning-level Japanese speech 7.4% less than Japanese listeners did. Unlike the Japanese listeners in the study by Major, Fitzmaurice, Bunta, and Balasubramanian (2002), who comprehended peer Japanese speakers less accurately than listeners with different L1 backgrounds, the L1 Japanese group in this study understood their fellow countrymen as well as other L1 groups. They understood the beginning-level Japanese speakers even more accurately than L1 Mandarin listeners.

**Interpretation of regression results.**

A simple summary of predictors and non-predictors of evaluative reactions and intelligibility is indicated in Table 4.26. “Yes” means that the variable has turned out to be a predictor, while “no” means a non-predictor. The signs (±) in parentheses indicate the directions of variables’ contribution in the regression equations; + refers to positive and – refers to negative. The numbers in parentheses are the coefficients of determination of the variables.
Table 4.26: Summary of regression analyses

A set of regression analysis indicated that the listeners’ native language (L1) was a predictor of their evaluative reactions to nonnative English speech. All the four L1 groups evaluated advanced speech more highly than less advanced speech, but the degree was different depending on the L1. An advanced Japanese speaker should expect more critical reactions from an L1 Hindi listener than any other listener. (Hindi listeners are
critical of NSs as well.) A beginning-level Japanese speaker of English may be able to expect more sympathetic attitudes from American listeners than other listeners. Neither familiarity with Japanese speakers or exposure to foreigners seems to soften listeners’ evaluative reactions to Japanese English. Likewise, gender did not matter concerning the reactions to Japanese English. Male and female listeners evaluated nonnative speech similarly.

The most significant thing about the intelligibility of Japanese English is that it was highly intelligible to all the listeners. The average intelligibility of 6 Japanese speakers was 94.8%. Unlike evaluative reactions, however, listeners’ English proficiency measured by the TOEFL is more important than their L1 in predicting the intelligibility of Japanese English, except for the beginning-level speakers. The higher a listener’s TOEFL score is, the more likely s/he can understand advanced and intermediate Japanese English. In order to understand beginning-level Japanese speech, however, only L1 was concerned. The L1 Mandarin group could not understand the novice Japanese speakers as well as the L1 Hindi, American, and Japanese groups did. Although they failed to understand native speakers, the L1 Japanese listeners demonstrated good understanding of their peer Japanese speech, especially at the rudimentary level, comparatively speaking.

**Discussion of Quantitative Study Findings**

The research results supported Hypothesis 1, 3, and 4 fully or partially, and rejected Hypothesis 2. Not only the rejected hypothesis but also the supported ones have created a new question: Why did this research produce these results? In this section, possible reasons of the research outcomes are discussed. Eight key questions emerged in
the analysis of data and it would be best address these questions as a way of discussing the results.

*Why were the L1 Hindi listeners the most critical of Japanese English among the four L1 groups?*

As expected from the review of previous studies, this study found that the nonnative participants were more critical about other NSs’ speech than native participants (L1 American English listeners) were. Furthermore, the results clearly indicate that the L1 Hindi listeners, representing the Outer Circle, were more critical than the L1 Mandarin and L1 Japanese listeners, representing the Expanding Circle.

A possible reason of this tendency is the combination of the L1 Hindi participants’ high English proficiency and their perceived identification as speakers of a nonnative variety of English. Their high English proficiency, demonstrated by the TOEFL scores and the intelligibility test scores in this study, is understandable, considering the role that English played in the education they received in their home country. For the majority of the L1 Hindi participants, English was the medium of instruction throughout their education. Even though they used Hindi as a home language, they might have “acquired,” as well as “learned,” English. On the other hand, since they consider English their “second” language, they might believe that any second language speaker should be as good at English as they are, with proper training. If this is the case, then, it is no wonder that the L1 Hindi participants were more demanding when they heard less advanced nonnative speakers than the L1 Mandarin and the L1 Japanese participants, who must have become fluent in English relatively recently.
Why was L1 Mandarin-ness not a significant predictor of evaluative reactions?

The reason why L1 Mandarin-ness was excluded in the stepwise regression analyses of evaluative reactions to Japanese English was that the L1 Japanese group was set as the reference category. No statistically significant difference was found between the L1 Japanese and L1 Mandarin listeners in their reactions. In other words, L1 Mandarin and L1 Japanese groups reacted similarly to various proficiency levels of Japanese English. This finding is particularly meaningful when the fact that the listeners’ L1 were chosen from the three concentric circles (i.e., Inner, Outer and Expanding Circles). Participants from different circles demonstrated different tendencies in their evaluative reactions to Japanese English. That is, the listeners from the Inner Circle were the most positive in their evaluative reactions to an Expanding Circle variety of English, while the listeners from the Outer Circle were the most critical. The reactions of the listeners from the Expanding Circle were moderate in comparison with those of the Inner and Outer Circle listeners.

Why were nonnative listeners more critical, even toward NSs, than native listeners were?

As expected, listeners’ native language was found to be a predictor of evaluative reactions to Japanese English. The nonnative listeners were more critical of nonnative speech than native listeners were, which is consistent with the findings of previous studies (Ludwig, 1982; Fayer and Krasinski, 1987; Cargile, 1997). An unanticipated result was that the evaluative reactions to native English also varied across native languages. Hindi listeners were more critical than any other L1 group and Americans were especially tolerant.
The distinctive reaction score of the Hindi participants was strongly related to their responses to one speech trait: accent. While all the members from L1 American, Mandarin, and Japanese groups judged that the NSs had little to no accent (scale value: 8～10), 7 (35%) Hindi group members considered Speaker #4 (N♀) to have an accent (scale value: 1~5), and 9 (45%) members responded that Speaker #7 (N♂) did, as indicated in the following figures (Figure 4.1 and 4.2). The L1 Hindi participants might have interpreted “accents” in a different way from the other listeners did. The details will be discussed in the qualitative interview section.

Figure 4.1: Responses to Speaker #4 accents across listeners’ L1

![Figure 4.1: Responses to Speaker #4 accents across listeners’ L1](image)
One possible reason why Chinese and Japanese participants did not evaluate native speech as favorably as Americans did was that these two L1 groups might have avoided marking extreme scores on the scale. The scale had 10 points between two opposite adjectival phrases, but most Mandarin and Japanese-speaking participants marked on neither end, toward the center. As Cargile (1997) found with his Asian American participants in a language attitudinal study, the Asian culture that favors moderation and understatement might have influenced the L1 Mandarin and Japanese listeners’ decisions.
Why did female listeners evaluate nonnative speech no differently than male counterparts?

Male and female listeners were expected to react to nonnative speech differently, based on the findings in sociolinguistics (Coats, 1986). Even though female speakers are known to be more polite in language use and more positive about listening than male speakers, this study found that they were not different in attitudes toward nonnative speech. The bivariate correlation, however, revealed that female participants actually evaluated speech samples slightly higher than male counterpart ($r = -.21$). Due to the relatively small mean difference, the regression significance did not reach the priori $p < .05$. With a larger sample size, gender could have been a significant predictor. When the unbalanced gender ratio across four L1 groups in this study is considered, the conclusion about gender effects in language attitudes needs to wait for larger-scale research.

Alford and Strother (1990) and Cargile (2002) claimed that their listeners indicated less positive attitudes toward female speakers. However, the comparison among speakers’ gender was not attempted in this study because extraneous variables were difficult to control. The speakers were different not only in gender and English proficiency level, but in many other aspects.

Why did female listeners understand speech better than male listeners?

The review of literature suggested that the listeners’ gender would be a predictor of evaluative reactions. In this study, however, a set of regression analyses indicated that it was not a predictor of reactions, but of intelligibility.
The higher intelligibility score of women than men’s might be related to the measurement, not the actual listening comprehension skills. The speech samples that had the largest variability in intelligibility were taken from the NSs and the beginning-level Japanese speakers. The intelligibility of advanced and intermediate Japanese speakers was almost perfect across four L1 groups. When listeners found a certain sentence was difficult to write down, men tended to give up the task altogether, leaving the writing column blank, while women tried to write even a single word. Since the score was calculated in the exact-match method, a single word was counted into the score, and eventually made a difference.

*Why didn’t either familiarity with Japanese English and exposure to foreigners affect evaluative reactions nor intelligibility?*

Gass and Varonis (1984) reported that even though the familiarity with the topic is the most powerful independent variable, familiarity with the specific accent and exposure to nonnative speech in general have significant facilitating effects on intelligibility. Derwing and Munro (1997) found that their native listeners’ intelligibility scores and their familiarity with the specific accents were correlated. Matsuura, Chiba, and Yamamoto (1992), as well as Chiba, Matsuura, and Yamamoto (1995), attributed L1 Japanese listeners’ negative attitudes toward other nonnative English varieties to the unfamiliarity with the varieties. Based on the preceding research above, the familiarity with Japanese English and exposure to other nonnative speech were hypothesized to predict the evaluative reactions toward, and intelligibility of, Japanese English. However, the obtained data strongly suggested that neither variable affected the dependent variables.
A possible reason for the outcome of evaluative reactions is that the adjectival phrases that were used on the scale in this study were less personal or more detached, compared with the scales from the previous studies. The Evaluative Reactions scale that was utilized in the study deliberately avoided affective judgements such as attractive, intelligent, and educated, for they were considered inappropriate as the measurements of international communication. The only item representing affects was irritating – pleasing. The other 7 items were more detached, such as effective communication – ineffective communication, and clear content – unclear content. Therefore, the positive feelings such as solidarity nurtured by the familiarity with Japanese English and exposure to foreigners might not have been reflected on the score.

The intelligibility score, on the other hand, seems to be explainable by the ceiling effect. Just like the Singaporean speaker in the study of Wilcox (1978), Japanese speakers spoke significantly more slowly, in comparison with native speakers (see Appendix A for speech rates). Even if nonnative speakers were actually more difficult to understand, listeners had sufficient time and background knowledge about the topic to masticate and conjecture what they said. As the result, the intelligibility of nonnative speakers became extremely high. Therefore, the listeners with less familiarity with Japanese English or exposure to foreigners gained as high scores as those with more familiarity and exposure.

*Why did the L1 Japanese listeners not understand native speech beyond the extent that TOEFL scores could explain?*

Theoretically speaking, the TOEFL score should have been the sole predictor of
the intelligibility of NSs, for TOEFL score is supposed to represent how well an English learner understand spoken and written standard American English. In reality, the regression concluded that the listeners’ native language explained intelligibility variance, besides TOEFL score. Japanese participants’ intelligibility score was significantly lower than other three language groups, when TOEFL score was controlled.

One possibility is that some Japanese participants’ listening comprehension skills did not reach the threshold level to understand the provided native speech. If TOEFL score 575, for instance, was the border to distinguish successful comprehension from failure, the intelligibility score would not follow the normal curve. It would be closer to dichotomous. Since the overall TOEFL scores of Japanese participants were the lowest among the L1 groups, the small proficiency difference possibly produced a large test score gap.

The other explanation is that the TOEFL score does not represent only listening comprehension skills. It is calculated based on the sub-scores in listening, reading, and writing sections. Japanese participants might possibly have been more proficient in reading and writing than listening.

Yet another explanation is possible when the US residence was considered. The average lengths of residence were 16 months for Hindi, 15 months for Mandarin, and 9 months for Japanese speakers. Since TOEFL is usually taken prior to the university admission, Hindi and Mandarin speakers might have taken the test a long time before they participated in the study. Longer residence in the USA may have helped Hindi and Mandarin speakers to improve their already advanced English even more.

The last possibility is the score assigned to American participants. There is no
theoretical background that native listeners should be given score 6 in TOEFL score ranks. A possibility is that advanced Hindi listeners might be as proficient in English as L1 American English group members. Thus, the value that should have been assigned to Americans might be 5.2 or 5.8. Since no data on NSs’ English proficiency in comparison with NNSs are available, the value = 6 had to be chosen rather arbitrarily.

Why did the Japanese listeners understand Japanese speakers as well as, or even better than, other L1 listeners, even though they failed to understand native speech?

This result contradicts the tendency that Major et al. (2002) found. In their study, the Japanese listeners did not understand peer Japanese speakers as well as Chinese, Spanish, and American listeners did. The differences of the results in this study and that of Major et al. (2002) might be due to the difference in speech samples. In the study of Major et al. (2002), predetermined texts written in standard American English for a TOEFL listening comprehension test were read by foreign accented-speakers. The current study, however, utilized spontaneous speech to investigate the intelligibility. The Japanese listeners in this study were in the advantage to predict the peer Japanese speakers’ word choice, grammar usage, and organization of ideas, as well as pronunciation and accents.
Qualitative Interview Findings

In the second phase of the study, qualitative interviews were conducted in order to explore and understand the processes of participants’ judgements of native and nonnative speech. The overarching research question for the interviews was: **What are the perceived factors that influence the listeners’ evaluative reactions to and the intelligibility of Japanese English?** The subordinate interview questions are reiterated below.

1. *What did you think or how did you feel while you were rating this speaker?*
2. *What are accents? Do they bother you?*
3. *Have you found specific accents or specific language speakers more difficult to understand than others? Why?*
4. *Do you feel more comfortable speaking with NS or NNS, or doesn’t it matter?*

In this section, the summary of responses to the interview questions is reported first, and then recurring themes will be discussed.

Thirteen interviewees were recruited from those who participated in the quantitative study, based on their L1, gender, and their evaluative reactions (i.e., very critical, average, and very positive). One of the advantages of stratified semi-structured interviews is, according to Hendrick (1997), to allow the interviewer to explore unexpected responses. In the initial stage of the interviews, the interview questions were
asked, but as the interviews progressed, the interviewees were encouraged to express their opinions freely on related issues.

Descriptions of Interviewees

The brief demographics of the thirteen interviewees are indicated in Table 4.27. Three American, three Indian, four Chinese, and three Japanese students participated. Three more students were contacted but not available for the study, because some of them had already moved to other schools and others never responded to the researchers’ request.

<table>
<thead>
<tr>
<th>Name</th>
<th>Native Language</th>
<th>Gender</th>
<th>Age (years)</th>
<th>US residence (months)</th>
<th>TOEFL rank*</th>
<th>Major</th>
<th>Evaluative Reactions mean score ( /80)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>English</td>
<td>Female</td>
<td>26</td>
<td>--</td>
<td>---</td>
<td>City &amp; Regional Planning</td>
<td>57.4</td>
</tr>
<tr>
<td>A2</td>
<td>English</td>
<td>Male</td>
<td>22</td>
<td>--</td>
<td>---</td>
<td>Psychology</td>
<td>41.8</td>
</tr>
<tr>
<td>A3</td>
<td>English</td>
<td>Female</td>
<td>21</td>
<td>--</td>
<td>---</td>
<td>Public Administration</td>
<td>69.4</td>
</tr>
<tr>
<td>H1</td>
<td>Hindi</td>
<td>Male</td>
<td>24</td>
<td>23</td>
<td>4</td>
<td>Civil Engineering</td>
<td>33.9</td>
</tr>
<tr>
<td>H2</td>
<td>Hindi</td>
<td>Male</td>
<td>38</td>
<td>22</td>
<td>4</td>
<td>Agricultural Engineering</td>
<td>47.6</td>
</tr>
<tr>
<td>H3</td>
<td>Hindi</td>
<td>Male</td>
<td>24</td>
<td>24</td>
<td>4</td>
<td>Mechanical Engineering</td>
<td>29.3</td>
</tr>
<tr>
<td>M1</td>
<td>Mandarin</td>
<td>Male</td>
<td>23</td>
<td>8</td>
<td>4</td>
<td>Chemical Engineering</td>
<td>50.1</td>
</tr>
<tr>
<td>M2</td>
<td>Mandarin</td>
<td>Female</td>
<td>29</td>
<td>11</td>
<td>4</td>
<td>English</td>
<td>34.5</td>
</tr>
<tr>
<td>M3</td>
<td>Mandarin</td>
<td>Female</td>
<td>30</td>
<td>10</td>
<td>3</td>
<td>Statistics</td>
<td>64.0</td>
</tr>
<tr>
<td>M4</td>
<td>Mandarin</td>
<td>Male</td>
<td>27</td>
<td>11</td>
<td>4</td>
<td>Chemical Engineering</td>
<td>49.1</td>
</tr>
<tr>
<td>J1</td>
<td>Japanese</td>
<td>Male</td>
<td>27</td>
<td>20</td>
<td>3</td>
<td>Accounting</td>
<td>43.8</td>
</tr>
<tr>
<td>J2</td>
<td>Japanese</td>
<td>Female</td>
<td>22</td>
<td>1</td>
<td>2</td>
<td>Marketing</td>
<td>48.8</td>
</tr>
<tr>
<td>J3</td>
<td>Japanese</td>
<td>Female</td>
<td>19</td>
<td>24</td>
<td>2</td>
<td>Education</td>
<td>39.9</td>
</tr>
</tbody>
</table>

*TOEFL ranks: 1 = ～500, 2 = 500～550, 3 = 550～600, 4 = 600～650, 5 = 650 ~

Table 4.27: Demographics of qualitative interview participants
Responses to the Interview Questions

In order to illustrate the characteristics of interview data and the processes of interpretation of those data, the interviewees’ responses to the five interview questions are presented in this section. Although the initial plan was to attend to all the five questions, only the first question was constant through the interviews, while the other questions were implicitly or explicitly answered through conversations, due to the digressive nature of the semi-structured interviews. Since the responses and comments were quite diverse, only those mentioned unanimously by several interviewees are exemplified in the following section.

*What did you think or how did you feel while you were rating this speaker?*

The interviewees spontaneously chose the speech traits to discuss for this question. Among the eight traits (i.e., content clarity, intelligibility/comprehensibility, pronunciation, fluency, emotional reactions, accent strength, and communicative efficiency), some were mentioned by most of the interviewees, while others were hardly referred to. Therefore, the total of the responses does not make thirteen. In the following section, the background information about each speaker will be presented first and then the opinions of the interviewees on his/her speech will be described.

- Speaker 1 (♂)

Speaker 1 is a male advanced speaker. He had studied English mostly in Japan until he became 17, and then he went to live in the UK for one year. When the speech material was assembled, he was 20 and a freshman at the university that was the research site. Approximately at the middle point of the speech, Speaker 1 was suddenly baffled. His nervousness afterward was revealed through the uneven rate of speech and stutters.
Two characteristics of the speech of Speaker 1 stood out through the interviews. The most frequently mentioned was its disorganization of ideas or lack of logical flow. Interviewees commented that Speaker 1 talked about one thing, and in the next moment, about another, without a logical flow between the sentences. Six interviewees commented that the content of the speech was not clear for this reason. Three explained that it is the reason why they perceived the speech was choppy. The lack of flow confused some listeners and impacted the perceived comprehensibility of the speech. The perceived incomprehensibility irritated some interviewees. Therefore, the lack of logical flow of the speech induced negative reactions on numerous traits.

Good pronunciation was another significant characteristic of Speaker 1. He was very positively judged both on the traits of pronunciation and accents, especially pronunciation. When asked, the interviewees expressed two definitions of “good pronunciation.” One was the pronunciation that they could understand and the other was what was close to native pronunciation. Speaker 1’s good pronunciation also enhanced the perceived intelligibility of the speech. Six interviewees said the content was clear to them, for they understood each sentence that Speaker 1 uttered. Three interviewees (M3, M4, and J2) even said they perceived Speaker 1 was a NS of English, while others stated they detected his nonnative accent.

- Speaker 2 (B♀)

Speaker 2 is a female beginning-level speaker of English. The 20 year-old student at a technical/vocational school of English in Japan had had all her 6 year-long English education in the country up to that date. When her speech was recorded, Speaker 2 had just passed EIKEN pre-2rd grade, which can be roughly
conversed into 400~450 on the TOEFL. She had little confidence in her English skills and was obviously nervous during the recording.

Three characteristics appeared prominent in the interviews. Choppiness of the speech was pointed out by nine interviewees, out of thirteen. Four interviewees described the speech “slow” rather than choppy, stating that the choppiness of this speech was different from that of Speaker 1. Four interviewees said the choppiness (or slow rate) evoked their irritation, and two of them said it hindered the comprehensibility as well.

Six interviewees unanimously referred to the bad pronunciation and the simplicity of the speech. The pronunciation, however, did not necessarily lower the perceived speech intelligibility, because the message was very simple and also it was delivered very slowly. On the other hand, the simplicity, or the scarcity of detailed information about the story, influenced the perceived comprehensibility. Five interviewees said that they had hard time grasping the whole picture, in other words, what the speaker really wanted to say, because she did not say much.

Besides these characteristics, four interviewees referred to the nonnative identity of Speaker 2. A common statement was, “…but it was not irritating, because I thought she was a nonnative (or a Japanese) speaker of English. I knew she was struggling.”

Speaker 3 (♂)

Speaker 3 is a male intermediate speaker in his 40s. He completed his formal English education in Japan. However, he had been in the USA for his profession for almost 9 years, communicating with Americans in English almost everyday. Speaker 3 looked relatively confident and composed during the recording. The speech was a little shorter (less than 2 minutes) than others.
Perhaps because Speaker 3 was neither extremely advanced nor novice, the evaluative reactions to him were almost split into halves. Four interviewees claimed that his speech was choppy, but three said his was smooth, at least smoother than the previous speakers. Six interviewees said his pronunciation was good, but two said it was bad. Four interviewees stated that Speaker 3 had pretty good pronunciation but had a strong accent. Six interviewees emphasized the clear content, based on its high intelligibility. Three rated the speech low on the same trait, for the bad story development due to the lack of essential details. One possible interpretation of these contradicting reactions is that most interviewees found the speech by Speaker 3 pretty easy to understand but quite different from native speech in phonology and precision of descriptions. The interviewees who regarded intelligibility as important might have evaluated this speech highly. For the interviewees to whom the approximation of native norms was important, this speech might have left much to be desired.

- Speaker 4 (N♀)

Speaker 4 is a female native speaker of English. She was in her late 30s, with a high school diploma. She was a little nervous during the recording.

Two characteristics were frequently mentioned about Speaker 4. First, all the interviewees, except for one (J2), correctly suspected that she was probably a NS of English. All the interviewees, except for one (A1), found her speech pleasing to hear. The reasons for the pleasant-ness were diverse: smooth-ness, clear and familiar pronunciation, appropriate word choice, high intelligibility/comprehensibility, and correct and detailed description of the pictures.
Five interviewees, however, pointed out the speech was not optimally smooth. They said they sensed the speaker’s hesitation or indecisiveness. According to M2, Speaker 2’s problem was that she revised herself, not in grammar but in content, which confused M2.

- Speaker 5 (B♂)

Speaker 5 is a male beginning-level speaker. He was 19 and had just passed EIKEN pre-2nd grade when his speech was recorded. He had such a clear and rather high-pitched voice that two interviewees thought he was a female speaker.

Complexity and bad pronunciation were the most frequently mentioned characteristics in Speaker 5’s speech. A1 and H2 pointed out that the speaker tried to make larger sentences. M1 and M2 stated that he put something that was not there into his story, in other words, he speculated. As the result, five interviewees described this speech complex. Seven interviewees mentioned the speaker’s bad pronunciation. These two attributes made the speech difficult for 6 interviewees to understand. M4 stated, “He translated his language into English.” A1 and M4 said that this speaker tried to do something challenging, but unfortunately it was beyond his current English skills.

J1 and J2 commented that this speech was easy for them to understand, but J1 assumed that it would not be easy for other listeners. Thus, J1 stated, he rated the speech as difficult to understand on the Evaluative Reactions scale. Three interviewees pointed out the poor grammar of the speaker and added that it distracted them from concentrating on the content. Six interviewees mentioned that this speech did not irritate them, however, because they knew the speaker was nonnative (2), his voice was pleasing to the ears (2),
the listener was accustomed to rudimentary speech (1), and it was easy to understand (1). (The figures in parentheses refer to the numbers of interviewees who expressed these opinions.)

- Speaker 6 (I♀)

Speaker 6 is a female intermediate speaker. She was 23 and had some college education, although she belonged to a technical-vocational school of English at the time. She had passed the EIKEN pre-1st grade, which can be converted to approximately 500-540 points on the TOEFL, when the speech was recorded. Speaker 6 appeared relatively calm during the interview, but she might have concealed her nervousness, for she mispronounced some words that she would not have in other situations.

Similar to the other intermediate speaker (i.e., Speaker 3), the comments about the speech by Speaker 6 were split. For example, four interviewees said her speech was effective in communication, but 4 others said it was ineffective. Three interviewees said the speech was choppy while four others said it was smooth. Although 6 interviewees mentioned that the speaker had bad pronunciation, four said her pronunciation was good.

This discrepancy in the evaluative reactions can be explained when the speech characteristics were considered. The pronunciation of Speaker 6 was good in general. She might have had an identifiable foreign accent, but it did not hinder the whole intelligibility. The interviewees who answered that her speech was effective in communication seemed to be emphasizing this aspect. In the meantime, Speaker 6 did mispronounce some words; she said /siː/ for “she,” /bə rlz/ for “balls,” and /ænɪməlz/ for animals, which were unintelligible for some listeners. If this aspect drew attention, the speech was likely to be judged ineffective in communication.
A unanimous opinion about Speaker 6 was that she made significantly many grammatical mistakes. Among them, confusion of pronouns (i.e., replacing she with he, or her with his) were specifically pointed out by 3 interviewees. As J1 put it, the speaker monitored her grammar consciously and frequently corrected herself, which might have drawn the listeners’ attention even more. The perceived choppiness and smooth-ness of the speech were related to this speech attribute.

- **Speaker 7 (N♂)**

Speaker 7 is a male native speaker of English. He was an undergraduate student at the point of speech recording.

The evaluative reactions to Speaker 7 were homogeneous across traits expect for irritation. Five listeners, including three Americans, stated that they gave him full marks. They said that Speaker 7 described the pictures in details, very smoothly, by using familiar pronunciation, with no foreign accent. They thought that his speech was easy to understand, both in intelligibility and comprehensibility, and that it was very effective in communication. As the result, the speech was pleasing for them to hear. A1 commented that the way Speaker 7 spoke was typical of a young American, with whom she identified. A1 sensed the speaker’s friendly-ness from the little sloppy way that Speaker 7 pronounced words.

On the other hand, many nonnative listeners expressed their feeling of unease with this speech. Compared with the other NS (Speaker 4), who produced 110 words per minute, Speaker 7 spoke at a faster pace (180 words/minute). Five interviewees stated the speech by Speaker 7 was too fast for them to understand 100% of what he said. M4 and
J1 commented that they received the impression that the speaker was not considering the nonnative listeners’ comprehension. M4 and H2 said that extreme fluency irritated them. M3 pointed out that Speaker 7 did not pronounce words clearly, which made her comprehension more difficult.

- Speaker 8 (A♀)

Speaker 8 is a female advanced speaker, in her 40s. She had completed her English education up to the tertiary level in Japan, and came to the USA more than 10 years ago. She was occasionally working as an ESL aid at a local elementary school when the speech was recorded.

Most interviewees commented that they perceived the speech by Speaker 8 was the most effective in communication among NNSs’. Seven interviewees stressed the detailed descriptions and the logical flow of the speech, which made its comprehensibility high. M2 said that the speed at which Speaker 8 strung up sentences was sufficiently quick. Therefore, many interviewees stated that this speech was more pleasing to hear than other nonnative speech.

However, seven interviewees stated that Speaker 8 had a strong accent. Seven (partly overlapping with those who pointed out the strong accent) said her pronunciation was bad. A1 assumed that Speaker 1 probably learned English after she became older, because in comparison with her good grammar and vocabulary, her pronunciation was poor. Though the overall ratings of Speaker 8 were favorable, the pronunciation characteristic made one interviewee make the following comment. M1 said, “Can’t understand. Very difficult. Bad pronunciation and she spoke very fast. Even with the pictures at hand, I missed two or three sentences. It really irritated me. Unclear content,
ineffective communication. Choppy between sentences.” Low intelligibility seems to have affected his evaluation of all the other speech traits.

**What are accents? Do they bother you?**

This question was added to the list, because the researcher noticed that the ratings of accents and those of pronunciation were related, in the quantitative data collection. In comparison with the definition of good pronunciation, the interviewees’ interpretation of accents was not straightforward. The most common statement was “I can’t explain it well, but I know it’s there when I hear it.”

An interviewee stated that he never thinks about the definition of accents, and he refused to improvise one. For some interviewees, an accent was a set of phonological signs that revealed the speaker’s identity. Some people said that bad pronunciation and strong accents were almost identical, while one interviewee said that pronunciation errors were accidental and accents must be habitual. M2 and A1 stated that pronunciation was (the knowledge about) the way a sound is articulated, while an accent is the sounds themselves that were actually pronounced. One interviewee (H3) said that an accent was like a gene and that one could never get rid of it. Some interviewees said they sensed the speakers’ accents in a sound (e.g., /r/ substitution with /l/), some said in the movement from one sound to another, some said at the end of a word, (e.g., a vowel insertion after the final consonant of a word), and others said in the sentence intonation.

A1 said she was fond of English with foreign accents. Other interviewees also responded that accents did not bother them, unless they were so strong that they affected intelligibility. When the intelligibility was threatened, the interviewees said, they would
have to make extra efforts to understand the language before they reach the content-level comprehension, which would be irritating.

*Have you found specific accents or specific language speakers more difficult to understand than others? Why?*

Understandably, the languages of the speakers with whom the interviewees found difficulty in understanding depended on the exposure to and familiarity with the specific accents. Interviewees answered this question from their limited experiences.

For the three Hindi-speaking interviewees, East Asian (i.e., Chinese, Korean and Japanese) English was difficult to understand, due to the accents. For four L1 Mandarin and Japanese-speaking interviewees, Indian English was difficult, because Indian speakers tended to speak fast with accents.

Several interviewees did not fall into this pattern, however. J2 and A3 stated that they did not find any difference understanding various L1 speakers. A3 said some native accents were difficult for her to understand. J3 said she found the Somali accent difficult. M2 found Japanese and Korean speakers of English difficult to comprehend due to their accents. A2 added that he did not want to overgeneralize, after he commented that Asians were difficult. Although H2 had found a Korean speaker very difficult to understand when he came to the United States two years before, he added that he now had no difficulty understanding accented English from all over the world.
*Do you feel more comfortable speaking with NS or NNS, or doesn’t it matter?*

Four interviewees answered that they were more comfortable speaking with NSs, for such reasons as easier comprehension and less risks of communication breakdown. Three Hindi speakers asserted that they were equally comfortable with native and nonnative speakers, unless the nonnative interlocutor was not unintelligible to them. Five interviewees responded that they were more comfortable speaking with NNSs. They roughly attributed the discomfort to two reasons. One was the inconsiderate or sometimes even rude attitudes from native listeners, and the other was their own incapability of understanding native, especially rapid speech.

*What do you think about international usage of English? About non-native speakers speaking English?*

Some interviewees were asked this question at the end of the interviews. Not all the interviewees were inquired, mostly due to the time restraints. The responses split interviewees into two groups; those who believed that English was actually being used for cross-cultural communication globally and those who doubted its status as an international language.

As the question was open-ended, the responses were various. Four people said that English was an international language. A1 and H1 advocated the global spread of English. H1 said that no other language could be used and understood in 100 countries. A1 said English could make international communication smoother. H2 pointed out the fact that despite the spread of English as an international lingua franca, native speakers are not accustomed to nonnative speakers. J1 expressed his ambivalent feelings about
English. He said communicating with native speakers in English usually puts nonnative speakers in the disadvantage. Although he sometimes resented the unfairness, he also thought the current trend could not be changed. He said that English is and will be the most powerful international language.

Other interviewees did not consider English as an international language. Two nonnative interviewees commented that English was the language of the USA. Simply because they were in the US at the moment, they spoke English. They said they would speak local languages if they went to other countries. Another interviewee, as long as these two, stated that English could not be an international language because it is not spoken internationally at all. M1 pointed out that the Japanese people did not speak English when he visited Japan. M2 said that Europeans other than British people neglected her when she spoke to them in English. When M3 suggested to her Mandarin-speaking friends that they should speak English among them to improve their English skills, her friends turned down the idea because it felt too awkward.

An American interviewee articulated his discomfort when he heard the word “English as an international language.” For him, the word implied that other language speakers were obliged to use English and it sounded too ethnocentric. He said he was strongly against the idea. At the same time, he expressed his respect to those who spoke any foreign language, not only English.

A Japanese interviewee stated that she had read a newspaper article about “English as an international language,” in which Japanese English speakers were encouraged to have more confidence in their Japanese-style English. When asked whether she was ready to take the advice, J2 simply chuckled and said, “Theoretically speaking, it
should be that way…., but I have just arrived in this country. I just want to communicate with people without misunderstanding.”

Recurring Themes

The listeners’ perceptions and judgements of speech traits turned out to be intertwined in very complex ways. Sometimes they were in a causal relationship, sometimes they were almost identical, and other times they were contrary. I also found that the data had two different aspects. Since the interviews were semi-structured, especially the first part of the interviews was a priori, more of a quantitative interview. The latter half of the interviews was more like a natural conversation on related areas. Therefore, in one aspect, interviewees provided information about prepared questions and in another, their responses were more private and sometimes unexpected. The first three themes introduced below are the answers strictly to my research question while the last three themes were what emerged as the interviewees’ beliefs and perceptions concerning English.

In this section, I will especially focus on intelligibility, irritation, and accents, for these traits appeared salient through the first part of interviews. Then, spontaneously recurring themes will be discussed. These themes are “criteria alteration,” “English as an International language?” and “power relationship between native and nonnative speakers.”
Perceived Intelligibility/Comprehensibility

The item easy to understand – difficult to understand on the Evaluative Reactions Scale assessed perceived intelligibility/comprehensibility. Interviews revealed, as Smith (1992) described, that understanding had layers: sound/word recognition (intelligibility) and word/utterance meaning (comprehensibility). Naturally, interviewees in the study utilized two different criteria depending on which layer they were judging, intelligibility or comprehensibility. Since the task of speech material used in this study was simple description of cartoons, meaning behind word/utterance (interpretability) will not be considered here.

Several interviewees referred to bad pronunciation and high rate of speech as factors that affected the intelligibility. The word “strong accent” was frequently used as the synonym to “bad pronunciation.” When I asked what the pronunciation was, many interviewees responded that it was how a speaker articulated individual sounds (phonemes). Some interviewees did not say more than “If I can understand, that’s good pronunciation,” which was a circular argument, because they also said, “I didn’t understand this speech because of the bad pronunciation.”

A typical remark on the pronunciation and the rate of speech that affected intelligibility is that of M1’s about Speaker 8 (A♀). “I can’t understand what she’s (A♀) trying to mention. It’s too difficult, ‘cause she has a very bad pronunciation and she speaks very fast.” This comment by M1 surprised me, because even though many interviewees said that the pronunciation of A♀ was not good, he was the only one who said he did not understand at all. The general reactions to the speech by A♀ was
positive, for its relative smoothness and inclusion of details. For M1, the pronunciation probably worked as a threshold and its other good characteristics could not be enjoyed.

The comment by J1 also implied that he was thinking word-level intelligibility, not comprehensibility. Again, pronunciation is the key to the intelligibility.

J1: When a speaker’s pronunciation is bad, I need to figure out what word he says before I can understand the content of speech. It may take time, and sometimes I can’t reach the content comprehension. In my mind, I spell out the words the speaker pronounces. But when the estimate goes wrong, the speech becomes incoherent. Then, I have no clue what he’s saying.

J2 frequently mentioned the rate of speech that hindered intelligibility. She said on Speaker 1 (A♂) and Speaker 7 (N♂) were both “too fast for me to understand.” She said she thought Speaker 1, who was really a Japanese speaker, was a native speaker of English.

These statements agree with the findings in the study of Munro and Derwing (1995a). They found that foreign accents decelerate the native listeners’ word recognition. In other words, when listening to accented speech, listeners needed more time to process the incoming information. Consequently, Munro and Derwing (1995a) reported, the listeners claimed the speech was too fast.

On the other hand, some interviewees evaluated comprehensibility of the whole speech. Comprehensibility seemed to be strongly related to perceived content clarity, as “logical flow,” “cohesion,” and “overall message” were key words. H1 made the following comment about Speaker 1 (A♂),

H1: The content was not clear. He was talking about one thing, and then talking about something else. The flow wasn’t there. That’s why I say
its’ not very easy to understand what he was trying to say. ….No, I didn’t have any difficulty understanding his pronunciation.

Similarly, M2 stated about Speaker 2 (B♀) as follows:

M2: I didn’t have problem understanding each word. I understand each word. But it was the overall message that I couldn’t quite get.

These comments by H1 and M2 imply that the threshold of intelligibility (word recognition) was no more concern to them. Compared with the interviewees who concerned with word-level intelligibility, those who were estimating utterance comprehension tended to make more critical comments on the Japanese speakers.

Yet reluctantly I have to admit that some interviewees seem to have changed their criteria from time to time. They applied intelligibility for one speaker and then comprehensibility for another. M3, for example, stated about Speaker 1 (A♂), “His speech was easy to understand. But his logic was not clear. I was confused, so I marked it not very positively.” M2 later said about Speaker 2 (B♀), “She was extremely choppy. She pronounced word by word. That’s why she was very easy to understand.” In this case, M2 was concerned about the comprehensibility with the more advanced speaker, while he was mostly discussing word level intelligibility with the beginning-level speaker. Thus, even within one listener, comparison of his reactions to two speakers seems difficult.

**Irritation**

Irritation was assessed with the item of irritating – pleasing on the Evaluative Reactions Scale. As the following comment reveals, however, irritation seems to be evoked by various factors. H1 said, “Choppiness, pronunciation, and the flow of the
speech, how he presented the ideas. These were factors that were taken into account.
That’s why I thought it was irritating.”

Among all, choppiness, or lack of fluency/smoothness, was the most frequently
reported reason of irritation. Interviewees asserted that there were two kinds of
choppiness. For example, Speaker 1 (A♂) was fast when he was articulating sentences,
but he took long pauses between sentences in the attempt to find proper things to say.
Probably from the pressure not to have unnatural pauses, he sometimes jumped from one
topic to another without transitions, which seemingly gave some interviewees the
unpredictable impression. As the following interaction indicates, they interpreted it as
choppiness. A2 and I were talking about Speaker 1 (A♂).

Reiko: (By checking A2’s responses on the Evaluative Reactions Scale) This is
rather clear but not extremely clear in content, right?
A2: Yeah. I think it was just because it was a little choppy. Maybe I should
have put it in the choppy section. I mean I understood what was going on,
but it seemed all over the place.
Reiko: Can you describe his choppiness more in detail?
A2: Sure, like just going from one sentence, going from like, um, the baby was
doing this and that and going to some completely radically different
subject, you know. That’s what I mean by choppy.

M2 stated also about Speaker 1, “His sentences were very choppy, a bit disorganized, and
I found that to be irritating.”

The choppiness of less advanced speakers such as Speaker 2 (B♀) and 5 (B♂)
was related to the scarcity of words. As the speech transcriptions (Appendix A) indicate,
the number of words uttered during the 2 minute-long speech by these two speakers was
extremely small. Words came out of their mouths off and on. Frequent repetition of the
same words and phrases was also observed in the speech by Speaker 5 (B♂). Even though
this type of choppiness sometimes induced a positive reaction such as “the pronunciation was not good, but understandable, because she spoke very slowly,” many listeners reported irritation.

Typical negative reactions were as follows. J3 said, “The most irritating speech among all was this (Speaker 2: B♀). Too choppy. She should speak faster. I didn’t understand what she wanted to say.” H1 also said about the same speaker, “She was irritating, because she was not clear and she was taking long breaks between sentences, breaks and pauses.” A2 said, “Her speech was incredibly choppy. Sorta difficult to understand.” Irritation caused by extremely slow speech seems to be linked to the short-term memory limitation. M2 commented on Speaker 2 as, “there was so many delays that by the time she gave me the last word in the sentence, I couldn’t remember what she said in the first.” Similar to the case of intelligibility, however, the boundary between what is acceptable and what is not depended on the individual perception. A3 made the following comment on the speech by Speaker 2 (B♀) that caused irritation among other interviewees.

A3: She was easy to understand. She’s definitely enunciating, pronouncing every word, and you can tell that she’s thinking about the words before she’s saying those words. But it makes it very choppy. Because you know the part of speech, you’re waiting for the verb. And it’s given eventually. And then you know it’s the end of the sentence.

Even though A3 realized the choppiness, she declared that listening to Speaker 2 was definitely pleasing. Thus, choppiness did not necessarily evoke irritation. Note that such reaction as A3’s was, however, comparatively uncommon.
Whether it was caused by choppiness or bad pronunciation, reduced intelligibility and comprehensibility seem to be the crucial factor of irritation. J1 said, “If I have to make special efforts to understand the person, it’s irritating, because I have to endure the temptation to express my frustration.” M2 put her irritation criterion into words as follows.

M2: If I have to work hard to understand the speech, it irritates me. Because I always feel that it’s the speaker’s responsibility to deliver meaning well, so if the speaker is not delivering meaning well, then I have to work very hard to understand, and then I do like, mmmmmmmmm… it’s not good.

Conversely, high intelligibility and comprehensibility induced positive reactions. J2 said about Speaker 8 (A♀), “Even though she had an accent, she was easy to understand. So she was pleasing.”

Voice quality was another factor influencing irritation. Two American female interviewees claimed that voice quality was the primary cause for their irritation or pleasing-ness. M1 also said the voice of Speaker 3 was irritating, because “his tone was a bit lower.”

Less commonly reported sources of irritation were as follows. J2 mentioned that pause fillers were irritating, because they carried no genuine information. M3 claimed that there were certain words and phrases that caused more irritation than others. M2 stated that grammatical inaccuracy also irritated her.

Accents

Perceived accents were gauged with the item of with strong accent --- with no accent on the Evaluative Reactions Scale. Accent seems to be a very complex and maybe
highly controversial concept. It is complex, because as the quantitative data indicated, the responses were divided into two extremes. For example, nine out of 20 Hindi-speaking participants evaluated Speaker 7 (N♂) as having a strong accent, while a large number of participants in other L1 groups responded that he had no or little accent.

I asked the L1 Hindi listeners who attributed “a strong accent” to Speaker 7 (N♂) to clarify their criterion for the accents, through email. Some replied, “An accent is what reveals the speaker’s identity, nationality, where he is from.” A Hindi-speaking male graduate student (H3) who responded that three Japanese speakers (#2: B♀, #5: B♂, and #6: I♀) had no accent on the scale explicating his criterion. “Their speech is so choppy that I can’t hear their accents. Since they speak word by word, they do not show accents. Accents can be perceived only in speech with some fluency.” Presumably, for those who rated NSs with strong accents might have thought the speakers’ identity as Americans was obvious from their accent.

In retrospect, if the item had read *with strong foreign accent – with no foreign accent*, less variability in responses might have been observed. My original intention was, however, to understand how people think and feel when they hear native and nonnative speech. By keeping the item rather ambiguous, I hoped to see whether what listeners associated with a concept would vary across their native languages. It did.

Accent might be a controversial concept as well, because it is related to politics, as Kalin and Rayco (1978) and Lippi-Green (1997) have claimed. After completing the Evaluative Reactions Scale, H3 said, “American people may think Indians have accent, but Indians think Chinese have strong accent.” Though he made this comment lightheartedly, it tongue-tied me for a while. I was struck by the idea, “So, there IS a
hierarchy among the Inner, Outer and Expanding Circles.” However, I did not hear such a straightforward remark as this from other interviewees. The hierarchical view might be unique to H3.

In the first round of the quantitative research session, no participant claimed that s/he could not or would not rate the speakers’ accents. However, not many interviewees gave me explicit answers when I asked, “What’s an accent? How is it different from bad pronunciation?” in the face-to-face interviews. The most common answer was that bad pronunciation and heavy accents were almost the same, but not exactly the same. When I probed by asking how they were different, many interviewees replied that they never thought about it seriously. When I pointed out that the listeners themselves had given different scores on pronunciation and accents, some were surprised and some were offended, as observed in the following interaction between A3 and me. We had just heard the speech by Speaker 3 for the second time.

Reiko: What about the pronunciation and accent with this speaker?
A3: …(chuckles) the pronunciation and accents. But I think I don’t know if I totally grasped what you mean by accent. Will you give me the definition of accent? That’s the thing I’m having a hard time grasping. Do you know what I mean, I don’t mean to make it difficult or anything. I just feel kind of
Reiko: Right, right, right. But you have said she has a strong accent, right?
A3: Oh, at least I thought, yeah, but I’m just second guessing myself and I’m backtracking because I’m not exactly sure I don’t feel comfortable, like, I mean strong accent as just like, from mine?
Reiko: Yeah.
A3: Mine?
Reiko: From your definition.
A3: My definition, I don’t have my definition.
Reiko: So, if someone’s speech is different from yours, you think that the person has a strong accent?
A3: See, I don’t like that. It’s a very leading question. It makes me feel, I don’t really like that at all.
Apparently this poor interviewee felt trapped. On the other hand, I was totally confused why he avoided answering directly. Many nonnative interviewees asserted, without hesitation, that a (foreign) accent was how a nonnative speaker’s speech differed from native speech, due to the influences of the speaker’s native language. My question was how he could rate the accent of the speaker if he had not had some kind of criteria. My best guess was that some American people must be very sensitive to political correctness.

Initially, I thought judging accents would be more neutral and therefore easier for interviewees to articulate their opinions on, compared with explaining their irritation. However, the distinction between pronunciation and accents was not clear among the interviewees, or/and some people were not comfortable discussing accents. What the interviewees maintained unanimously, however, was that an accent was not a big deal unless it interfered with intelligibility.

H3 took an example from my English to explain the difference between pronunciation and accents. He said, “Your pronunciation is good, but you still have an accent. I can understand you perfectly.” I asked some interviewees a question, “Do you want to speak like a native speaker, if you can?” No one replied that they did. On the other hand, no interviewee said, “I’m proud of my accent,” either, as I was expecting to hear from L1 Hindi interviewees, based on what I had read about World Englishes. They seemed to have accepted accents as a matter of fact. This tendency may be related to the many interviewees’ identity as graduate students. They were mature, highly educated, confident professionals in the making. Immigrants like the ones in the study of Norton (1995) might have commented differently on accents.
**Criteria Alteration**

In the course of describing how each speech sample was evaluated, the majority of interviewees verbalized their habit of utilizing different yardsticks to evaluate the communicative effectiveness of native and nonnative speech. The interviewees claimed that they changed their point of reference once they identified the speaker as nonnative, based on their accents in most cases, or their grammar and fluency in some cases. Since the real identities of the speakers were not disclosed in the study, the interviewees made their comments on the assumption that they had detected the speakers’ identity correctly.

H2, who had stayed in Japan and China for several months in the past, expressed his expectations about English spoken by Japanese and Chinese as follows. We were discussing Speaker 2 (B♀).

**H1:** And it was not irritating, because I could make out she is a foreigner, she’s not a native speaker.

**H1:** Because out of eight points, for really full prize, she was not suited to my judgement, so that’s why I don’t feel that she was doing this communication very effectively.

**Reiko:** But your judgement was not so negative, was it?

**H1:** Not so negative. Because I’m judging at the same time, I’m trying to take into consideration the kind of cultural background that she has, and the country where she’s coming from, like Chinese or Japanese. They don’t get so much of accomplishment of speaking English at home.

The following conversation was held with M3, who repeatedly told me that he hardly became pleased or irritated by listening to any kind of speech.

**Reiko:** So you thought this speech was neither pleasing nor irritating.

**M3:** Yeah. Can you give me an example of pleasing or irritating?

**Reiko:** Well, some people are easily irritated, because of choppiness,
for example. Some people find it really very irritating, because people tend to wait for the next word, “what are they going to say?” Some people can’t wait.

M3: OK, OK.
Reiko: But it doesn’t occur to you?
M3: I think, rather at the beginning I know he’s a foreigner. So, it’s not, if he say word by word, it’s very difficult for a foreigner to speak a native language. It’s understandable. So, I can understand.

In the following excerpt, M4 implied that she was very sympathetic toward nonnative speech. In this specific case, the speech that she had just heard was very slow (Speaker 2: B♀). Notice that M4 mentioned her own identity as a nonnative speaker immediately after she was asked to explain her reactions to the speaker.

Reiko: This is speech sample #2. And M4, why did you rate as you did?
M4: Because I’m a nonnative speaker, too, so I think although she is somewhat hesitant to speak, but I think I try to understand, because she speaks slowly, so I can understand what she tries to tell me.

J2 rated Speaker 2 (B♀) more pleasing to hear than Speaker 1 (A♂), who was more advanced in proficiency. In the preceding conversation, J2 stated that she had thought the first speaker was a native speaker of English. Therefore for Speaker 1, she did not feel the sympathy she felt for Speaker 2.

Reiko: On the scale of pleasing-irritating, you judged this speech (Speaker 2) was less irritating than the previous speech. Why?
J2: Well, in my heart, I thought I could understand the speaker’s feelings because I thought she was a Japanese. (laughter) I wasn’t irritated, because I was cheering the speaker, who tried to speak well. Since we are from the same country, I might have been less critical with her.

J1 also stated that he changed the way he comprehended or interpreted speech, once he recognized that the speaker was Japanese. Note in the following conversation that
he changed his manner of cognition, not attitudes. According to him, he could understand Japanese speakers of English better than other native or nonnative listeners would.

Reiko: In what case, do you judge someone’s pronunciation as bad?
J1: Well, it depends on the speaker’s nationality. I can understand Japanese mentality. I know, from my own experience, that they pronounce some sounds in certain ways, since they can’t pronounce them right. I usually can tell what they really want to say.

Reiko: Do you understand Japanese speakers of English as smoothly as you do native speakers?
J1: Um, no. I understand English native speakers in English. When I hear a Japanese person speaking English, I comprehend him in the mixture of Japanese and English. Not in complete sentences, but I pick some English words from the speech while I translate what he says into Japanese.

Reiko: Do you mean, even when he speaks English?
J1: Yes. It also depends on the speaker’s proficiency level. Perhaps a Japanese speaker comes up with an idea in Japanese and then translates it into English, I believe. So, I usually think of the original Japanese. When I’m not sure of what the speaker wants to say, I would reason it by analogy based on the original Japanese. So, I believe I understand the Japanese speaker more than what he actually speaks out. For I can predict, thanks to my background knowledge.

Reiko: What do you do when the Japanese speaker’s proficiency is high?
J1: The higher the proficiency, the more I switch my understanding to the way I’d understand a native speaker. The key is the balance.

In contrast to nonnative listeners, American interviewees did not particularly articulate their compassion toward the NNSs. However, they expressed their rather complex feelings toward the peer NSs. A1 made the following comment about Speaker 4 (N♀).

A1: No, uh-uh. Um, but she sounds educated. She sounds kind of bitchy, too. No, that’s not frank, but that’s why I put it irritating. It’s like I can tell, but that’s ‘cause I’m American so I know that.
This interviewee also emphasized her American identity when she evaluated another American (Speaker 7: N♂).

A1: well, I was thinking about… like I said, I’m American, I can tell, I can read American people very well. And like, he’s young, and he’s confident. He’s funny. And reason I can tell is, yeah, his voice, the way he talks, he’s really easy with his words. And there’s just, I know people like that. There’re certain kinds of people. He’s a very confident, popular, man. Young.

A1: Yeah. I mean, this is young American. This is what who I am. So you know, I can identify with it. And yeah, it’s good for you to keep that in mind, cause maybe some people would not even have that feeling. But there’re some people that you want to be around. They’re really charming. And he is that.

The fact that this interviewee did not make similar comments on NNSs might imply that she was using different criteria when evaluating native and nonnative speakers.

*English as an International Language?*

At the end of the interviews, I asked the interviewees what they thought of the international use of English, or simply English as an international language. The responses split into two groups; those who believed that English was an international language, and those who thought it did not exist.

First let us discuss the believers. A1 was optimistic and welcomed the current situation. Her view of English is related to her world view. Since America has been the center of the world, for good or for bad, so its language inevitably would be.

Reiko: What do you think about the international usage of English? I mean almost everyone speaks English to some extent all over the world?
A1: Do I think it’s a good thing?
Reiko: Yeah.
A1: Yeah, I feel like, it’s inevitable. It’s going to happen.
Reiko: Why is it inevitable?
A1: Because it just, it seems to me like, it’s spreading, like so much. And I think the United States, it’s the biggest English speaker, and the United States is the combination of different people from the world. And there are people like me, white, Caucasian, you know, Protestant. But, then you have just the combination of the rest of the world. Everybody settles around the United States, to a certain extent. They might not think that, but that’s just the way it is. And so, the United States is English. The other thing is that, we were talking about language, my fiancé and I. And he said it’s basically, you can judge a culture by how well you can communicate with each other, and English is like, you can communicate a lot with it. And I think that it’s like a revolution, you know like, communication will further yourself and further your culture. And English can just help that.

Another American interviewee, A2, was cautious about how his remarks as a majority group member could be perceived by minorities.

Reiko: What do you think of international usage of English, or English as an international language?
A2: It’s so tricky. I don’t like that idea. I don’t really like that idea. I think it’s possibly taking away from one’s culture.
Reiko: You don’t like the idea taking away…..?
A2: Making English the international language, like every country and all that. That’s what you’re saying, right? I don’t know. It feels very ethnocentric. It just feels like you’re trying to push something on another, I just don’t feel like … taking away one’s culture. I don’t know. I don’t know how to articulate either.
Reiko: Mmm.
A2: But I don’t. I don’t agree with that.
Reiko: I see. But for example, we can communicate with each other now, because I speak English to some extent, right?
A2: Right. Sure.
Reiko: So, it’s a kind of necessary evil, or something?
A2: That every person or every country should speak English?
Reiko: … OK, so you’re against the idea that everyone should speak English?
A2: Yeah. I’m against that idea.
Reiko: But aren’t you happy that you’re a native speaker of English?
A2: I don’t know anything else, so I guess I can’t really say that I’m not.
Apparent from the above interaction is that A2 and I were discussing different concepts. As discussed in the previous chapters, I had my own view of English as an international language in which every user is an equal stakeholder and should be respected. On the other hand, A2 interpreted the concept in such a way that English (as American language) is (or should be) used internationally. From his remark, I sensed the confusion and bemusement of someone who was born with privileges. Being a white English-speaking American man was not his fault. Yet he might have been criticized for these traits.

J1, a nonnative speaker, shared the concept with A2 but was more cynical about the international use of English. He believed that English is the most powerful international language and that it will continue to be, but he was not frustrated by the concept.

J1: English-speaking countries have power, in many senses, in this modern world. I mean, economically and politically. Since the USA and UK are the strongest, American and British people don’t think they need to learn the local languages even when they visit foreign countries…. Reiko: Do you think it’s inevitable? Or are you angry about it?
J1: I’m pissed. But I also think it can’t be helped. So, when I talk with an American, and when he says, “Huh?” I would say in my mind, “OK, then, you come to Japan and speak Japanese. You can’t, can you?” I’m just trying to calm myself down. (laughter) But, then again, no matter how indignant I am, the big trend won’t change.

Next, let us consider the interviewees who believed that English was not an international language. Their argument is grounded on the fact that people do not speak English in the Expanding Circle. M1 drew an example from his experience in Japan.

Reiko: What do you think about international use of English?
M1: You have to look how to define “the international language”. ‘Cause I speak English, ‘cause I’m now in America. It’s not because it’s an international language. I mean, if I’m in China or Taiwan, maybe I ask people first if they speak Chinese or not. And if I’m in Indonesia, I’ll ask
whether they speak Indonesian or not. So, since I’m in America, I speak English. But for international language, maybe I think, most countries use English as their language.

Reiko: Most countries?
M1: Yeah, but actually I think Chinese is most wide spread.
Reiko: Chinese is the most wide spread language? Do you mean, the population?
M1: The population, yeah.
Reiko: What about when you go to Europe, for example? What language would you use?
M1: I visited Japan, before.
Reiko: Uh-huh?
M1: And I spoke English there.
Reiko: Because?
M1: Because English was more popular.
Reiko: Than? Chinese?
M1: Yeah.
Reiko: And you don’t speak Japanese, do you?
M1: I know some words, but. And then even if I used English, I feel I didn’t really have effective communication with them.
Reiko: Because? Why? Because their English was not good enough?
M1: Yeah. Of course, it depends on the country, who you talk to.

In the excerpt above, M1 first said the number of NSs of English was bigger than that of any other languages, then corrected himself by mentioning that the language with the largest native speaker population was actually Chinese. When I asked about the situation where he had no command in local languages, he mentioned Japan, where he could not communicate in English anyway. Similarly, M2 referred to her experience in Europe where people would not use English with her.

M2: I don’t think English is an international language at all. (laughter)
Reiko: So your English is the language of …? …of the US. I wouldn’t use it anywhere else. Of course, if I go to England, it’s English, but it’s a different kind of English. It’s very different. They call things differently. (She goes into examples.) There are many countries in Europe where, because I’ve traveled so much, there are so many countries in Europe where English is not recognized. If you speak English, people would ignore you.
Reiko: You mean you had to speak the domestic language?
M2: Yes, yes. Even if it’s broken or not understandable at all. If you go to
France, you have to try to speak French, because if you say “Parlez vous anglais?” “Do you speak English?” people would say, “Non”. They’d immediately say no, even though they can, because they don’t want to embarrass themselves. So they would rather see you try to speak broken French, than they try to speak broken English and embarrass themselves. And it’s the same in Switzerland.

The term “English as an international language” seems to imply to this group that people all over the world are at least bilingual. Since the reality is not so, English is not an international language to them. M1 and M2 both empathized that they spoke English just because they were in the United States for the time being. Communicating with other NNSs in English in the United States, M2 stated when asked, was the last resort. She was not comfortable when she spoke English with NNSs. She spoke English with NNSs only when English was the only common language among them.

I wrote in my field notes later, “But isn’t it (communicating in English, since it is the only common language) exactly what’s happening in cross-cultural communication? Don’t we call it English as an international language?” When two speakers share a native language, they would naturally speak it. When they do not, English tends to be the language used, because it is the most widely learned second or foreign language. As a matter of course, I was not in the place to argue the definition of English as an international language with my interviewees. I learned, through the interviews, that the discourses of World Englishes or EIL were far from being accepted. They were unknown.
The interviews were not intended to probe into the issue of the power relationship between native and nonnative speakers. However, comments referring to it emerged frequently in the interviews, as J1’s comment in the preceding section indicated. I also asked the interviewees which they felt more comfortable speaking with NSs or with NNSs. The responses can be roughly divided into three groups.

Four interviewees (A2, M1, M2, and M3) answered that they were more comfortable speaking with NSs. As clearly stated in the previous section, English was the language of the USA for M1 and M2, so naturally they said they wanted to speak English only with Americans. M1 stated it was the way he could improve his English, and M2 added that most NNSs “freaked out” with her English proficiency, which truly amazed me during the interview. A2, an American, simply said that he was more comfortable with someone like himself. M3 responded, “it depends.” If a native speaker did not speak too fast and if M3 could understand what the native speaker said, he would be more comfortable with a native speaker. M3 expressed his concern that the risk of communication breakdown would be doubled when speaking with NNSs. In case his English was not sufficient, a native speaker would be able to infer the intended message, but he doubted a nonnative speaker would. These interviewees were quite straightforward in their logic.

Three Hindi speakers (H1, H2, and H3) asserted that they were equally comfortable with native and nonnative speakers, unless the nonnative interlocutor was unintelligible to them. They said they received some criticism of their English, especially
about speed. H1 said that he was frustrated right after he came to the USA. Eventually, however, he learned that the criticisms did not have to be taken as offensive and came to think critical feedback could be bi-directional. H2 shared his experience as follows.

H2: I’ve learned from my American friends. Like they do not hesitate in asking me, “can you repeat that again?” and now I’m in the habit of asking the same thing to them, ”Can you say that again, please?”

Reiko: So, your American friends ask you to repeat the same thing again?

H2: Yes, yes, when they don’t understand me. And initially, I was very hesitant to ask them to repeat. But now I’m comfortable asking them. Because now I know that it is the same for them not to be able to understand me, as I pronounce something differently from the way they’re used to hearing. And the same thing applies to me, I also find them speaking about something very differently, so I ask them.

The interviewees who were comfortable speaking with both native and nonnative speakers happened to be Hindi-speakers. I sensed their pride in their English in their remarks, and I wished all the NNSs could feel the same way as they did.

The third group (J1, J2, M4, and A1 and A3) responded they were more comfortable speaking with NNSs. Their comments made me ponder the politics of language. J2 and M4 repeatedly indicated their lack of confidence in their English skills. The following is the quotation from J2’s remarks.

J2: Well, I’m still comfortable speaking with international students. Because, they tend to use simple English, for English is not their native language. Sometimes their pronunciation confuses me, but overall, they try to speak hard and they try to listen to me very eagerly. Besides, when I speak with a native speaker, well, I’m not saying I’m ashamed of my poor English, but I feel a bit of inferiority. It’s OK if the native speaker is determined to understand me. But, I get panicked when they say “huh?” at McDonald’s, for example. Well, I don’t know…

M4 made the following comment on Speaker 7 (N♂).
Reiko: This was speech sample #7, what did you think?
M4: He must be a native speaker. And, but I think maybe in my opinion, maybe there is an intervening factor, because for an international student like me, I cannot catch up with every word when a native speaker is speaking. Because English is their language and you see it everyday, so they speak so fluently, and maybe they combine some words together while speaking. So sometimes I got a problem in listening. Again, their wording,… maybe their use of some words, it’s not enough for me to look, but I’m not so familiar with. So, although I can catch the main idea of their speaking, but,…mmm…. It’s my problem.

Such words as “inferiority” and “my problem” clearly indicate that these nonnative interviewees perceived English as “their” (i.e., native speakers’) language as opposed to “mine.” Is it a coincidence that both interviewees were from the Expanding Circle?

J1’s response was a little more complicated. In order to understand his remarks, some knowledge about the reason why he was interested in English in the first place might be necessary. He had participated in a home-stay program in high school, but could not communicate with his host family due to his limited English skills. He kept his regret for a long time after this experience, and had worked hard to master English since then. His impression of American people was very positive until he came to the United States to pursue his degree.

J1: It’s easier for me to comprehend native speakers, but I’m more comfortable speaking to nonnative speakers.
Reiko: What does it mean?
J1: If I just listen, I just have to understand. They talk, they do all the output. If I talk, however, I have to make them understand. Then, many native speakers take a certain attitude once they find out I’m not a native speaker. Their basic stance is “This guy doesn’t speak English. I won’t understand…” Even if I try to make myself understood hard, they’ve already half given up on me. On the other hand, nonnative speakers understand my feelings, which I just assume, but I feel they can understand my English quite well.
Reiko: Did you feel that way even when you were in Japan?
J1: Well, in Japan, we have few opportunities to speak with native
speakers. Most native speakers we interact with in Japan tend to be fond of Japan or Japanese culture, or to be learning the Japanese language themselves, so they know how difficult it is. We seldom speak with native speakers who have never been out of Ohio. So it was completely a different story. It was not until I came here that I knew that many Americans look down on those who can’t speak English.

J1’s remark can not be minimized as paranoia, when the following comments by NSs are taken into account. A3 shared her experience with peer Americans who would not understand NNSs.

A3: I had an ECOn teacher. This was quite interesting. I don’t know where he was from, but I used to talk to him like everyday. But in class, he had a heavy accent. So, he was teaching ECOn and he was like a Ph.D. student. He had been here for like, five, six years but he still had a heavy accent. There were people back of the classroom yelling up to the front of the classroom, I would say, “I can’t understand what you’re saying. We can’t understand you.” Now that’s rude, first of all. And second of all, I’m just like, if you’re back there talking, you wouldn’t understand what he’s saying. And being a person that I am, I was just like, he said this, and if you stop talking, you could understand him. And we could too. Because we can’t hear him, when you have people in the back whose voices are carrying you can’t hear them. And like, you can hear them after class, saying like, “oh, he can’t even speak English,” stuff like that. I think it’s like, the need to assimilate-kind of concept. If they need to concur, in order for me to understand what you’re doing, you need to say what I say, and you need to speak how I speak. And I don’t think that.

A1 made a comment about her interaction with international students in a similar situation as A2.

Reiko: Maybe some of the nonnative speakers you talk with are very, maybe rudimentary, or at the beginner or novice level, right? And when you talk to them, what do you feel?
A1: Um, I don’t know. I feel happy, because I have control, which is really bad, but it’s like something I admit. It’s like I can help them and it makes me feel powerful. I also feel so bad for them, because like, if you’re in class, I just met this girl, she’s from China. And she’s just been here like a month, and she has no family here. And
she’s in class, and my teacher’s talking, and she’s talking too fast.
And I know she can’t understand her. You know, and it makes me
feel really bad for her. But when I’m talking to her I’m just happy,
you know. Because A: like they’re listening to me. And B: like I get
help them. If they’re nice, then I’ll be more happy.

On the contrary to the unspecified Americans in J1’s example, A1 was very enthusiastic
to help out her classmate with limited English proficiency. However, the underlying
construct seems to be the same across the episode of J2, A2, and A1. Native speakers hold
power while NNSs throw themselves at their mercy.

Reflection on Qualitative Interviews

Due to the qualitative nature of this section of the study, the researcher could not
remain an impartial, objective observer of the phenomena. The first person pronoun will
be intentionally used to refer to the researcher in this section.

Fontana and Frey (2000) rightfully wrote, “Increasingly, qualitative researchers
are realizing that interviews are not neutral tools of data gathering but active interactions
between two (or three) people leading to negotiated, contextually based result” (p. 646).
These words exactly express what I, as the interviewer, experienced during and after the
interviews. My identities, as a nonnative speaker of English, an Asian, a Japanese, a
woman, a Ph.D. student, all influenced the interviews inevitably, as well as my being a
peer student in the same university as the participants. The fact that I am apparently older
than most of the interviewees may have affected some of them. On some occasions, the
interviewees’ hesitation, discomfort, or unwillingness to discuss certain issues was fairly
obvious, and I assumed it was because of me as an interviewer. A young American
Caucasian man, for instance, would have obtained completely different data from
interviews with the identical interviewees. In other words, the following findings could have been possible by the unique interactions between my interviewees and me.

As mentioned in the methodology chapter, Japanese participants were interviewed in Japanese, since I shared their first language. Other participants were interviewed in English. Fontana and Frey (2000) pointed out the difficulty of “understanding the language and culture of the respondents,” and as a nonnative speaker of English, I ironically realized how difficult it was to communicate delicate details/nuances with other speakers. Complex, layered meanings were sometimes replaced by ordinary, trite expressions. Compared to what I could obtain through my interviews with L1 Chinese and Hindi and even American English participants, the data from Japanese interviewees were much richer. It would be fair to mention that all the nonnative participants in these interviews were quite advanced speakers of English. I had also believed I had no problem communicating in English, until I conducted these interviews.

Summary of Qualitative Interviews

In order to probe individual listeners’ perception of Japanese English, face-to-face interviews with thirteen participants were conducted. The formulated research question was “What are the perceived factors that influence the intelligibility of, and listeners’ evaluative reactions to, Japanese English?” Findings were reported with 6 eminent themes: (1) intelligibility, (2) irritation, (3) accent, (4) criteria alteration, (5) English as an International language? and (6) power relationship between native and nonnative speakers.
“Understanding” speech was composed of two levels: word-level recognition (intelligibility) and whole speech grasp (comprehension). The perceived factors for intelligibility were the pronunciation and speed. For comprehensibility, cohesion and “the flow” were emphasized. Irritation was caused by many factors. Among them, choppiness was the most irritating trait for some. Pronunciation was also a factor, when it hindered intelligibility of the speech. Accent turned out to be a complex concept, but many interviewees articulated that it was not very important unless it interfered with intelligibility.

Participants changed their judging criteria based on the speakers’ (perceived) identity as native or nonnative speakers. Many nonnative interviewees voiced their sympathy toward NNSs. American interviewees seemed to look deeper into native speech. The majority of interviewees did not consider English as an international language, based on the argument that it was not spoken so internationally. For many interviewees, probably except for Hindi speakers, English was the language of America. Thus Americans were believed to have power, for good or for evil.

One thing to be remembered is that all the interviewees in this study were seeking degrees in an American university. Their perception of English as the American language might come from this research site. Nevertheless, finding that the international students who are exposed to other NNSs to a considerable extent conceived English as the American people’s language was quite surprising to me.
CHAPTER 5

SUMMARY, CONCLUSION,
AND PEDAGOGICAL IMPLICATIONS

English has been called an international language for decades. More and more international communication is carried out by means of English, increasingly in situations where no native speakers (NSs) are present. The number of nonnative speakers (NNSs) of English worldwide surpasses that of native speakers. In fact, Seidlhofer (2001) recently insisted that English as an international language should be treated and examined separately from English as an ethnic/national language of the UK and USA.

Yet, in reality, research in English applied linguistics and English language pedagogy have been conducted according to a long-established assumption. Nonnative English has been investigated from native perspectives, based on native norms. For example, studies on attitudes toward, and intelligibility of, nonnative speech have traditionally concerned how native speakers feel about, and how well native speakers understand, nonnative speech. This study aimed to contribute to the knowledge about English as an international language from the perspectives of nonnative users.
The current study has combined quantitative and qualitative approaches to understand the factors that predict native and nonnative listeners’ evaluative reactions to and the intelligibility of Japanese English. The importance of including nonnative listeners as evaluators and of utilizing speech samples taken from less advanced Japanese speakers were argued for within the framework of World Englishes and English as an International Language in the review of literature.

In the quantitative \textit{ex post facto} study, the listeners’ native languages and the speakers’ proficiency levels were set as main independent variables, while the listeners’ gender, familiarity with Japanese English and the exposure to other nonnative speakers were proposed as rival independent variables. Three types of instruments (Evaluative Reactions Scale, Intelligibility Test, and a personal information questionnaire) were developed and administered to 84 undergraduate and graduate students recruited from four different native language groups (American English, Hindi, Mandarin, and Japanese). The participants listened to and evaluated, on an eight-item, ten-point scale, eight speech samples gathered from six Japanese speakers of English with three different proficiency levels, and two American English speakers. The participants then wrote down 16 sentences extracted from the speech samples word for word, for the measurement of the intelligibility of the speech. The participants also provided information about their age, gender, major academic field, residence in the USA, TOEFL score, familiarity with Japanese English and degree of exposure to foreigners. Data were collected and statistically analyzed with SPSS.

In the qualitative interview study, thirteen participants were purposively selected from those who participated in the quantitative study, based on their evaluative reaction
scores, native languages, and gender. In a face-to-face interview, each interviewee listened again to the speech samples that were used in the quantitative study for the second time, and described how they judged them on the eight traits of the Evaluative Reactions scale. The interviews were semi-structured and the interviewees were encouraged to express their opinions on related topics. All the interviews were recorded on audio-tapes and transcribed with a transcription machine. Then the data were coded into significant themes and findings.

Four research hypotheses and one research question were formulated to facilitate the investigation. The summary and interpretations of research findings will be presented hereafter, by confirming or rejecting the hypotheses in the quantitative study and addressing the overarching research question in the qualitative phase. Then, quantitative and qualitative research findings are combined to better understand the phenomena. Conclusions and pedagogical implications will follow, leading the recommendations for further research. This chapter concludes with the limitations of the study.

Summary of the Findings

Research Hypothesis 1 (H1)

H1: The listeners’ evaluative reactions to Japanese English will be predicted by: (1) listener’s L1 (i.e., American English, Hindi, Chinese and Japanese) and (2) speaker’s proficiency (i.e., advanced, intermediate, and beginning-level).

H1 was confirmed. A stepwise regression analysis entered two dummy variables of L1 for a regression model. The proportion of variance in evaluative reactions that was
explained by listeners’ Hindi-ness and American-ness were approximately 22%. The L1 Hindi group rated speech more critically than the reference group (= L1 Japanese listeners), while the L1 American English listeners evaluated speech more favorably. The L1 Mandarin group was not significantly different from the L1 Japanese group in the evaluative reactions.

Descriptive statistics indicated that the speakers’ proficiency levels were reflected in the evaluative reactions to them. The higher the speakers’ proficiency, the more favorably his/her speech was evaluated by the listener groups. Yet, the difference in evaluative reaction scores between advanced Japanese speakers and intermediate Japanese speakers was small when compared to the difference between the native speakers and the advanced Japanese speakers.

When observed more closely, listeners from four different L1 groups reacted differently to speakers with different levels of proficiency. Compared with the Japanese listeners, L1 Hindi participants evaluated native speakers more critically, while American and Chinese listeners rated them more positively. In the evaluation of advanced Japanese speakers, L1 Hindi listeners were predicted to be more critical than any of L1 Japanese, American English, and Mandarin counterparts. Although the proportion of variance explained was relatively small (7%), L1 Hindi participants again turned out to be more critical than listeners from the other three groups in evaluative reactions to intermediate Japanese speakers. As for novice-level Japanese speakers, American listeners rated them significantly more favorably than other listeners did.
Research Hypothesis 2

H2: The intelligibility of Japanese English will be predicted by: (1) listener’s L1 and (2) speaker’s proficiency.

H2 was partly supported. The initial regression analysis indicated that the intelligibility scores obtained by the L1 American and Hindi groups were higher than L1 Japanese group. Statistically speaking, the L1 Mandarin listeners did not perform differently from L1 Japanese on the intelligibility test. When the listeners’ TOEFL score ranks were considered, however, a stepwise regression entered the listeners’ gender as a predictor as well as American-ness and Hindi-ness. In other words, if the TOEFL score rank was controlled, women obtained higher scores than men did, while L1 American English and Hindi listeners performed better than L1 Japanese and Mandarin counterparts. The proportion of variance in intelligibility scores explained by the four variables, namely listeners’ TOEFL score rank, American-ness, Hindi-ness, and gender, was approximately 54%.

The latter half of the H2 was rejected. That is, the intelligibility of Japanese speakers did not follow the proficiency levels. Advanced speech was not necessarily easier to understand than rudimentary speech. Descriptive statistics shows that the intelligibility of native speakers was actually lower than that of Japanese speakers among L1 Japanese and Mandarin listeners.

A closer look at the relationship between the speakers’ proficiency levels and the listeners’ traits revealed the following tendencies. The listeners’ TOEFL score ranks, L1s, and gender were predictors of the intelligibility of native speech. The participants with higher TOEFL score ranks understood native speakers better than the participants with
lower ranks, understandably. The L1 American, Hindi, and Mandarin listeners could transcribe native speech more accurately than L1 Japanese listeners did. Women understood native speech better than men did.

As a predictor of the intelligibility of advanced and intermediate Japanese speakers, only the listeners’ TOEFL score rank turned out to be significant. Listeners from all four L1 groups similarly understood advanced and intermediate Japanese speakers when the TOEFL scores were taken into account. The listeners’ TOEFL score ranks, however, failed to predict how accurately they would transcribe the beginning-level Japanese speech. Only one dummy variable turned out to explain the variance in the intelligibility. The intelligibility scores of the novice-level Japanese speakers was significantly lower among the L1 Mandarin listeners, in comparison with other L1 group members.

At first glance, the results of the intelligibility test did not make sense. Listeners’ L1 and gender, besides TOEFL scores, influenced the intelligibility of native speakers, while solely TOEFL scores predicted the advanced and intermediate Japanese speakers’ intelligibility. However, when the performance of the L1 Japanese participants is examined more deeply, the following pattern can be observed. For some reason, the L1 Japanese participants in this study were not as capable of comprehending native speech as other L1 listeners. However, the L1 Japanese listeners understood advanced and intermediate Japanese speakers as efficiently as any other listeners. They even understood novice-level Japanese speakers better than L1 Mandarin listeners. Comparatively speaking, English spoken by fellow countrymen proved to be easier for Japanese listeners to understand.
Research Hypothesis 3

H3: The evaluative reactions to Japanese English will be predicted by: (1) listener’s gender, (2) listeners’ familiarity with Japanese English, and (3) listeners’ exposure to foreigners

H3 was rejected. A set of stepwise regression analyses of evaluative reactions excluded all the rival independent variables: listener’s gender, familiarity with Japanese English, and exposure to foreigners. They failed to predict listeners’ evaluative reactions to native or nonnative speakers. The possible reasons for this result were discussed in the previous chapter.

Research Hypothesis 4

H4: The listeners’ intelligibility of Japanese English will be predicted by: (1) listener’s gender, (2) listeners’ familiarity with Japanese English, and (3) listeners’ exposure to foreigners

H4 was partially supported. A stepwise regression entered listeners’ gender as a significant predictor of native speaker intelligibility. However, a set of regression of Japanese speaker intelligibility excluded all the rival independent variables. None of these, listener’s gender, familiarity with Japanese English, and exposure to foreigners, predicts the intelligibility of Japanese English. This result contradicts that of Gass and Varonis (1984), who concluded that familiarity with the language variety and foreigners in general enhances intelligibility. The extremely high intelligibility scores in this study were the possible reason for the discrepancy, as was discussed in the previous chapter as well.
Research Question for Qualitative Interviews

What are the perceived factors that influence listeners’ evaluative reactions to, and the intelligibility of Japanese English?

In order to probe individual listeners’ perceptions of Japanese English, face-to-face interviews with thirteen participants were conducted. Through one hour-long interviews, participants explicated the criteria that they used for evaluating the speech samples. Besides the perceived factors influencing the most salient traits on the scale (intelligibility, irritation, and accent), three themes emerged through the interaction between the participants and the researcher. The themes were “criteria alteration”, “English as an International language?” and “power relationship between native and nonnative speakers.”

Perceived factors that affected the listeners’ responses to the item easy to understand – difficult to understand depended on how the listener interpreted “understanding.” When the listeners were concentrating on word-sound recognition (intelligibility), the pronunciation and rate of speech were reported to be influencing elements. When the listeners interpreted “understanding” as the semantic grasp of utterances (comprehension), content clarity, cohesion and “the flow” of speech were emphasized as the influencing factors.

Irritation was caused by various reasons. Among them, choppiness was the most frequently reported irritating trait. Pronunciation was also a factor, when it hindered the intelligibility of speech.

Accent turned out to be a complex concept, probably because accents were not specified as foreign accents on the scale. Whereas most participants interpreted it as
foreign (in this case, non-American) accents, a substantial number of L1 Hindi listeners construed “accents” as the indicator of the speakers’ nationality, and consequently attributed strong accents to American speakers.

Some participants expressed discomfort in discussing accents. On the other hand, many participants articulated that the accent was not very important unless it interfered with intelligibility.

The interviews also revealed that the participants did not evaluate native and nonnative speakers in the same way. Once the interviewees recognized the speaker was nonnative, they claimed that they adjusted to his/her choppiness and grammatical inaccuracy. The majority of interviewees did not consider English an international language. For them, probably except for L1 Hindi participants, English was the language of America. Consequently, some nonnative interviewees stated that they were uncomfortable communicating with other nonnative speakers in incomplete English. Some others expressed the frustration they felt about the unequal relationship between native and nonnative speakers of English.

Combining Quantitative and Qualitative Findings

Since the current study has adopted a dominant-less dominant design (Creswell, 1994), the integration of the findings from quantitative and qualitative studies is attempted by interpreting “qualitative findings to explain quantitative results,” according to the recommendation of Tashakkori and Teddlie (1998).
The \textit{ex post facto} research revealed the general tendency that the L1 Hindi listeners were more critical than the other listeners when they evaluated nonnative speech. When I conducted face-to-face interviews, part of this question was answered. First of all, the L1 Hindi participants expected a certain amount of information conveyed in a speech, that is, in this study, detailed descriptions of pictures. They were not looking for native-like pronunciation, but they wanted as much fluency. Too slow or choppy speech was not acceptable, even if English was not the speaker’s native language.

In the discussion on World Englishes, Kachru (1985, 1986) insisted that Indian English should be considered as the locally established variety of English as American or British English. Accordingly, I anticipated some L1 Hindi participants would claim their native speaker-status in Indian English, but it did not happen in this study. Hindi-speaking participants in the study considered themselves NNSs, and probably for this reason, they did not hesitate to criticize poor quality of the speech by Japanese English speakers.

In comparison, the L1 Mandarin and Japanese participants were more sympathetic toward the less advanced NNSs. Even though the majority of participants had demonstrated their high English proficiency by test scores like the TOEFL, their self-confidence in the language was more vulnerable than that of Hindi-speakers. Therefore, Mandarin and Japanese group listeners could have identified with the Japanese speakers. They probably recognized the speakers’ hesitation, reluctance, and exertion to force themselves to speak English to accomplish the given task.

On the other hand, the quantitative study demonstrated that the L1 Mandarin and L1 Japanese participants were still more critical of Japanese English than the L1 American participants were, which seemingly contradicted the interview findings. The
key seems to be the norms/standards of English that the L1 Mandarin and L1 Japanese listeners had received in their home countries. In their hearts, they might have changed the criteria of their judgements of nonnative speech due to the feeling of compassion and solidarity, but the norms that the participants believed what English should be like might have too deeply permeated their minds. In other words, for many L1 Mandarin and Japanese listeners, American English that was familiar through their English education had established the unshakable standard on which they could judge the quality of speech. The L1 Mandarin and L1 Japanese listeners often commented that they judged a certain sample as poor because it did not sound like true (which means native) English.

The L1 American English listeners showed very accepting attitudes toward beginning-level Japanese English in the quantitative study. The interview findings suggested three interpretations. One is that the American participants were honestly grateful for any foreigners who took the trouble to communicate in “their” (i.e., American people’s) language, since the participants’ general notion of English as an international language did not include its localization nor nativisation. Another interpretation is that American listeners might unconsciously have condescended to NNSs. “Poor English” might satisfy their sense of justice, in other words, their desire to side with the socially weak. The other possibility is that some American participants were more sensitive about discrimination against foreign accents than listeners from the other countries, due to the effects of multicultural education and political correctness.

The quantitative study produced another seemingly incomprehensible result: a speech sample with high intelligibility was not necessarily rated high in the evaluative reactions, while another speech with low intelligibility was judged very positively in the
reactions. The interviewees clearly explained this result. Some of the Japanese English speech samples were almost unintelligible but could be transcribed after the listener contemplated the context. The listener had to make efforts to harvest a certain amount of information. In this case, the perceived comprehension lowered, and the evaluative reactions went toward negative. On the contrary, some portion of speech by a NS was unintelligible to many nonnative speakers, but the whole speech provided enough information to satisfy the listeners. Even though some information was lost, as the original amount was abundant, the nonnative listeners perceived it was good communication, therefore, good speech. The other interpretation is related to the ownership of English. As some nonnative interviewees articulated, many L1 Mandarin and L1 Japanese participants might have thought that they “should” be able to understand the speech since it was American English, and that if they did not understand it, it was their fault rather than the speaker’s.

Conclusion

Divided by the classic native and nonnative dichotomy, the current study has produced findings that are in accordance with previous studies. Both native and nonnative listeners evaluated native speakers more favorably than nonnative speakers (Kalin and Rayko, 1978; Varonis and Gass, 1982; Cargile, 1997; Fayer and Krasinski, 1987; Matsuura, Chiba, and Yamamoto, 1992; Chiba, Matsuura, and Yamamoto, 1995; Dalton-Puffer, Kaltenboeck, and Smit, 1997). Furthermore, nonnative listeners were more critical in judging nonnative speakers than native listeners were (Fayer and Krasinski, 1987).
However, this study has succeeded in exhibiting the differences in evaluative reactions to and intelligibility of Japanese English, among listeners from the Inner, Outer, and Expanding Circles. In evaluative reactions, Outer Circle users (i.e., the L1 Hindi listeners in this study) were the most critical of Japanese English, while Inner Circle users (i.e., the L1 American English listeners) were the most positive about it.

On the other hand, listeners’ English proficiency (i.e., TOEFL score rank in this study) seems to be the most powerful predictor of their intelligibility scores, but listeners from all the concentric circles transcribed Japanese English with quite high accuracy (from 92.9% to 96.9%). After the linear effect of the TOEFL score rank was removed, the concentric circle of the listener’s origin did not seem to matter, except for the intelligibility of beginning-level Japanese English, which was significantly lower to the L1 Mandarin listeners than to the other listener groups.

However, when the difference between the intelligibility of native speech and Japanese speech to the L1 Japanese listeners is taken into consideration, the study results seem to suggest that a peer English variety might be easier to understand than other nonnative varieties. Unlike the study results of Major et al. (2002), the L1 Japanese listeners in this study understood beginning-level Japanese English better than the L1 Mandarin listeners, even though the L1 Japanese listeners did not comprehend native speech as accurately as the other L1 groups. The L1 Japanese interview participants articulated not only that they were empathetic toward the Japanese speakers in the affective domain, but also that they could “better understand” Japanese English cognitively because they could infer what the speakers really wanted to say, based on their linguistic and cultural background knowledge.
In conclusion, the current study has demonstrated that the English users from the three concentric circles are different from one another in evaluative reactions to and comprehension of an Expanding Circle variety. Therefore, research results with Outer Circle English users may not be directly applicable to the Expanding Circle, nor the other way around. Consequently, in research related to English as an international language in applied linguistics and ELT, users from all the Inner, Outer, and Expanding Circles should be involved.

Another insight can be drawn from the study. Ryan, Giles and Sebastian (1982) stated, “in every society the different power of particular social groups is reflected in language variation and in attitudes toward those variations” (p. 1). With the assumption that English users for international communication construct a type of large language community, this study suggests the “power imbalance” among the three concentric circles of English users, both through quantitative and qualitative findings.

The argument about “linguistics imperialism” by Phillipson (1992) and Tsuda (1999) haunted the researcher, especially during the face-to-face interviews. Unsparking convictions of Japanese English or apologetic comments on one’s own English such as “It’s my fault (not to be able to understand native speakers),” or “because of my poor English” made the researcher ponder whose benefits the current spread of English was serving. It may not be the result of the conspiracy of the Inner Circle to spread their language worldwide for their benefits, but has apparently created political inequity in communication among international English users. Without the intentional intervention of language education, English may be perceived as the language of the UK and USA by the majority of lay persons. The global attitudes toward nonnative, especially Expanding
Circle, varieties may continue to be negative even in the future, long after the familiarity with their English makes their intelligibility acceptably high.

Pedagogical Implications

Even though the qualitative interviews strongly suggested that individual listeners were so unique in their history, culture, and beliefs that foretelling exact personal reactions would be impossible, the quantitative data clearly showed knowing the listeners’ native language would provide a clue to a certain attitudinal predisposition. For example, according to the regression equation obtained through the _ex post facto_ study, an intermediate Japanese speaker should anticipate more critical attitudes toward his/her English from a L1 Hindi listener than L1 American, Mandarin, and Japanese listeners. The L1 Hindi listeners in this study responded that the choppiness was the most irritating factor. In other words, fluency is a key to more positive reactions from them. Although controlling one’s rate of speech according to the interlocutor’s native language sounds a little unrealistic, English educators can emphasize developing fluency (or more precisely, the amount of information one can convey per minute). If a nonnative speaker is advanced, on the other hand, s/he needs to speak in an orderly fashion, because the L1 Hindi listeners seem to regard cohesion important for comprehensibility. Needless to say, more knowledge about nonnative interlocutors’ perceptions would be essential for the manifestation of English as an international language.

While our current knowledge about nonnative speakers’ perceptions is still limited, observing interlocutor’s verbal and nonverbal reactions could be more
emphasized in English classes. Also, learners need to learn that a request for repetition and clarification should not be taken as insulting, neither to nor from the interlocutors.

Differences between native and nonnative speakers need to be reevaluated from the international/intercultural perspective as well. Medgyes (1992, 1994), for example, stated that a “deficient” command of English of nonnative teachers could be their advantage over native English-speaking teachers. The findings from this study will add a new asset to the nonnative teachers of English. Their speech perception is likely to be closer to that of other nonnative speakers than of native speakers. Therefore, nonnative teachers might be able to train their students more appropriately for international communication, than native teachers who are possibly too easy-going.

More importantly, this study suggests that nonnative speakers need more opportunities to develop positive attitudes towards other nonnative speakers. Alptekin (2002) pointed out that the pedagogical model solely based on native speaker norms is not only utopian and unrealistic but also constraining in teaching English as an international language. According to him, one of the criteria necessary to be considered in an alternative pedagogical model is the communicative competence that can equip learners with an awareness of difference. This study has manifested that the native and nonnative listeners are aware of differences at least between Japanese and American English. What is critical in the new pedagogical model seems, then, more accepting attitudes towards the differences.

Especially nonnative speakers need to learn to respect and appreciate themselves and other nonnative speakers, regarding what a difficult task they are accomplishing. Otherwise, simply increasing the exposure to other nonnative speakers might not help, for
exposure itself does not ameliorate the attitudes toward them as this study indicates, as well as Matsuura et al. (1992). The fact that native speakers of English are not the sole interlocutors that the learners are going to deal with should be more highlighted in the early stage of English education. Healthy and amicable attitudes toward nonnative-nonnative communication need to be intentionally nurtured in class, in order to modify the current overemphasis on native English-speaking interlocutors.

Recommendations for Further Research

The primary purpose of this study was to investigate whether listeners from three concentric circles would understand and react to Japanese English differently. Identifying speech traits that provoked certain reactions among listeners was not attempted. Therefore, comments such as “the replacement of /l/ with /r/ induces the sense of strong accent among 23% of L1 Mandarin listeners” can never be made. If certain linguistic characteristics need to be specified to improve intelligibility of and evaluative reactions to a variety of English, many variables need to be rigorously controlled, including the voice quality, gender, pronunciation, rate of speech, vocabulary, and grammatical accuracy. Such error gravity studies with nonnative listeners will be pedagogically helpful in EIL education.

The episode of J1, a Japanese participant in the qualitative interview, suggests that attitudes or evaluative reactions to certain language speakers might be impacted and changed by new experiences. In order to better understand the complex phenomenon, longitudinal, especially panel studies are recommended, in which the changes in attitudes
and listening comprehension skills of the same sample are investigated over an extended period of time, hopefully starting the observation before the participants come into an English speaking environment.

In this study, only American, Chinese, Indian, and Japanese listeners were included. Replications of the study with other L1 speakers from the Inner, Outer, and Expanding Circles will help deepen the understanding of the cross-cultural communication through English. Furthermore, the participants in this study were undergraduate and graduate students at an American university. Other populations such as younger learners in their home countries or professionals who habitually utilize English for international communication should be examined and compared with the university students in the Inner Circle utilized in the current study. An investigation into English teachers’ attitudes to English varieties across three circles will also be meaningful, for the purpose of teacher education.

Similarly, only a spoken language variety was investigated in this study. Granted that communication consists of two modes, spoken and written, and with the recognition of importance of global cyber literacy in English, linguistically and culturally diverse readers’ evaluative reactions to and the intelligibility/comprehensibility of nonnative varieties of written English will be of great practical, as well as theoretical, interests.

Limitations of the Study

This study has some limitations, including internal validity, external validity, and reliability. First of all, as all the *ex post facto* studies are doomed, selection of subjects can
be a threat to internal validity. The participants from four different L1 groups were likely to be different in other areas than L1. For example, compared with other L1 groups, L1 Hindi group had more members who were majoring in various types of engineering. Many constituents in L1 Mandarin and Hindi groups held the Graduate-Assistant, Teaching-Assistant, and Research-Assistant-ships at the university. With the restriction of time and budget, these facts had to be ignored, but might have influenced the research results. Another uncontrolled extraneous variable is the listeners’ attitudes towards Japan and Japanese culture, or toward East Asia and East Asian cultures. The majority of participants in the interviews identified the nonnative speakers, who were actually Japanese, as East Asian. The general attitudes that participants held towards East Asians might have influenced their evaluative reactions to the speakers.

Instrumentation might have been a threat to the external validity. First, international communication, by its nature, is interactive. Even though a reasonable extent of intelligibility is unfailingly the basis/prerequisite for interactive communication, uni-directional listening to narrative descriptions of a sequence of pictures, a method used in this study, is unlikely to occur in reality.

Secondly, while criticizing the native-speaker-oriented convention in English applied linguistics and ELT from the view of World Englishes and English as an International Language (EIL), this study has adopted the TOEFL scores as a yardstick of both the speakers’ and listeners’ English proficiency. Since the TOEFL is unmistakably a test of American English, its usage is theoretically contradicting the basic standpoint of the researcher. The TOEFL was used simply because no similarly standardized and widely-administered test exists to measure examinees’ proficiency in EIL. The issue of
assessment demands more attention with regard to the EIL proficiency. When an appropriate assessment for EIL is developed in the future, naturally English users from all three concentric circles will sit for the test side by side and the direct comparison will be facilitated.

Another threat to the validity is the choice of traits in Evaluative Reactions Scale. Eight speech traits were selected from the review of literature, and the content validity was secured by the panel of experts. However, the grammatical accuracy and rate of speech, not included among the traits, were frequently mentioned through the qualitative interviews. Some important traits might have been left out from the instrument.

On the other hand, the average scores in the intelligibility test were extremely high. The ceiling effect was suspected to have produced the small variance in intelligibility scores on Japanese speakers. A possibility that the selection of extracted sentences caused the result can not be negated.

Non-probability sampling of subjects in this study threatens external validity. Not only that the study results can not be generalized beyond the students in the current university, they even should not be generalized beyond the participants themselves. The participants were volunteers, so they may have been different from those who did not volunteer to participate in the study.

When the qualitative interviewees listened to the speech samples for the second time, they sometimes demonstrated quite different reactions from what they did in the ex post facto study. Though the direct comparison was impossible, for different procedures were taken in two studies, a threat to reliability as consistency of measurement was speculated.
LIST OF REFERENCES


Oda, M. (1999). English only or English plus? The language(s) of EFL organizations. In G. Braine (Ed.), Non-Native Educators in English Language Teaching (pp. 105-121). Mahwah, NJ: Lawrence Erlbaum Association, Inc.


One day the wife left home to go to a friend’s wedding. She’s wearing necklace, and a dress, um, and she has kind of rose-kind of thing in here. And she looks beautiful, an’ she looked beautiful. And the wife asked his husband to do housework, like, ah, cleaning houses and cooking, and babysitting. Um, **well, now the husband started to cook.** And ah, while he was waiting water to boil, he was playing with his kid. Um, … you know, like, **he showed some toys to ah, toys to his kid,** and playing with um, soccer ball. And ah, now, um, (pause) now water came out of the bin, the fried pan, and he rushed to stop it, he rushed to turn the switch off. And now he was ready to cook. He tried to cook something, like noodle, or, ah, I don’t know, but, yeah, he’s trying to cook something. Now babies, baby cried again back there. And, you know, husband can do only one thing. Trying to turn off and trying to cook, or playing with baby. Playing with his kid. And now, um, you know, he’s so tired. He was trying to do so many things. He cooked, he played with his baby, and ah, he wanted to has his lunch, dinner. So he’s tired, he slept. But the baby was so powerful, the baby was awake. And, baby was awake, and ah, now his wife came home. You know, she looked so surprised. She was, you know, she was expecting her husband to do house work, and cooking, and taking care of baby. But he slept and baby was awake. So, well, … they happens to divorce, or something? (laughter)

- 285 words
- 2 minutes 45 second
- 104 words/minute

One day the wife left home to go to a friend’s wedding. (pause) The husband say, say good by to his wife… he, **he stayed home with … his … baby,** he played with … his … baby. He … he was cooking, ah, he was cooking, <sigh>, baby is … crying. Um, **the pot very, the pot is very boiling.** Mmm, when she arrived at home, he is very, …he was very tired. He couldn’t, he couldn’t cook well. (pause)... She, she surprised.

- 85 words
- 2 minutes 20 seconds
- 36 words/minute
Speaker 3  I♂

One day the wife left home to go to friend’s wedding. The husband was told by wife, the wife, ah, to take care of his son, ah, their son, and ah,… (pause), he ah, **he thought he, ah, he could do that.** And, then, ah, he try to, tried to play with her, with, with him, with toy, and, also he [had], during the playing with him, **he also have to prepare the dinner.** And so he have to, he has to um, take, ah, take, ah, he has to[ok] ah, prepare cooking, and the h-, son started crying. And finally ah, he was tired and slip down. And when wife [t] ah, came back home from the wedding, ah, the room was so, ah, so messed. And, the son is playing, still, but ah, the husband was sleeping.

- 139 words
- 1 minute 51 seconds
- 75 words/minute

Speaker 4  N♀

One day, the wife had a wedding to go to and she decided that she would go by herself and leave her husband with the baby. She was going to a Japanese wedding, and **he said that, the husband said that he could take care of the baby.** So he would entertain the baby with, playing with his toys. He has a truck. And **they had the … All the toys out. There was a basketball, and a teddy bear.** And the husband decided that, while he was playing with the toys and the baby, he would put, ah, cook the, cook dinner for the baby. And, so, as they were playing, the food is on the stove, and it starts to boil over. So he runs into the kitchen. The husband runs into the kitchen and leaves the baby in the other room, as he’s crying. And it looks as though, he’s cooking rice and the rice’s boiling over onto the stove. And he’s frantic and he doesn’t know what to do. So he grabs the pan without a hot-pan-holder, and burns his hand. He ends up having a mess in the kitchen. There’s food everywhere, there’s bottles and dirty dishes. And he is exhausted. And he ends up laying on the floor in the living room with the baby. And the wife comes home and can’t believe the mess, and the clean up that she’s gonna have to do. (2:12)

- 241 words
- 2 minutes 12 seconds
- 110 words/minute
Speaker 5  B♂
One day the wife left home to go to a friend’s wedding. The husband, husband thought, “I can, I can do, I can do it anything, mother’s work. So first, the husband, the husband plays his son. It is so good, but his cooking skill is bad, not so good. He went to, he went to stove ah, fire. He thought, he thought, “I, I do not, I do not cook, so, so I try it again.” But his son, his son is angry and crying. Husband, husband very surprised, but he couldn’t do any, anything. So, the husband very tired. He didn’t do anything. Then her wife came, came home and very surprise. But the baby was so happy. He plays his father very much.

- 126 words
- 2 minutes
- 63 words/minute

Speaker 6  I♀
One day the wife left home to go to a friend’s wedding. She asked to her husband um, to care her child, their child. First, the husband, the husband have confidence to care his child. First he played with her, with his baby using toys, like stuffed animals and soccer balls. And in the evening, he started to cook, and he, while he is in kitchen, the baby started to cry. But … he doesn’t, didn’t use to cook, and child, and to care the baby, so he became upset. And he couldn’t stop the baby crying. After that, the wife came back to home but the husband tired, tired and fell asleep. And the kitchen became very messy. But the baby, but the baby looks very happy and the wife was very surprised.

- 134 words
- 1 minute 58 second
- 68 words/minute
One day, the wife left home to go to a friend’s wedding. And she looks very dressed up. She’s excited it’s probably a good friend of hers. She’s probably glad to leave the house too, leave the child at home with her husband. Her husband seems pretty happy. I mean, I think he thinks it’s a good time for him and the son to bond. So, he gets the truck and he’s playing with the sons. His chance for some alone time with his son, to play with him. He brought a soccer ball also, and to do a lot of guy things, so they can have their male bonding. And, being a guy, he can only concentrate on one thing at a time. So there’s a pot in the kitchen that’s beginning to boil over. So, he suddenly runs into the kitchen, and, which upsets his son. Son wants to play some more. It appears he is not having a good time at all, the husband is getting pretty frantic. So he’s trying to do two things at once, which he obviously can’t do. So he’s trying to take care of the pot and entertain his son at the same time. But, um, so, he’s probably wondering what his wife’s doing right now. She’s probably having a really good time, while he’s sitting home, trying to take care of all these things, probably realizing how difficult of the time that is. To be a mother, to take care of the son, ‘n to take care of the house. So, um, eventually, the father appears that he's worked really hard in the kitchen and it looks destroyed. And, there’s stuff coming down the countertops. There’re bottles everywhere. The sink is full and messy. The floor is a mess. The living room looks to be, a little more destroyed than it was previously. And so the wife comes home, and, it appears that she did go to the wedding she has a rose sitting in her purse, or in her bag. She comes in. She’s pretty shocked at what she sees. She’s surprised at how much the house can get torn up. So, the father’s on the floor, the baby appears to have had a good time, so he sets his mom.
One day the wife left home to go to a friend her, a friend’s wedding. The husband had never … spends the baby, spend with the baby alone at home, but he told the wife that he could handle it, handle everything. So, he said to the wife, “Don’t worry,” and “Have a good time.” And the wife left. An’ first, the husband started, started preparing the lunch for them. And he tried to make a stew for them. And he put everything in a pan and put on the stove. And while they were waiting for the stew it’s done, was done, the father was playing with the baby, with a lot of toys. But the husband had never cooked before, so he didn’t, actually, he didn’t know what to do. And, and he, he didn’t know this pan was overflowing. And when the husband noticed that the pan was overflowing, he ran into the kitchen. And he tried to take care of the stew. While the husband was in the kitchen, the baby started crying. So he didn’t know what to do. So anyhow, he, they ate lunch. And after that they played together. But when the wife came home, they were so exhausted. The baby was, the husband was very exhausted. The baby was still playing, but the husband was falling into sleep.
APPENDIX B

VISUAL CUE FOR NARRATION
B.1: Visual cue for narration
Name _________________________

Instruction:
The purpose of this session is to measure your reactions to the eight speech samples you are going to hear. Each sample will last approximately two minutes. There will be a pause before you hear the next sample. Please rate each sample on the set of scales below. There are eight pairs of adjectives with ten spaces between them. Please make sure you understand all the adjectives.

Questionnaire

Sample <Example> is/has/uses:


Here is how you place your check marks on the scale:
If you find that one adjective exactly represents your reactions to what you hear, you should place your check mark as follows:


If you think that the adjective very closely represents what you feel about the sample, you should place
your check mark as follows:


If you think that the adjective *closely* represents what you feel about the sample, you should place your check mark as follows:


If you think that the adjective *somewhat* represents what you feel about the sample, then you should place your check mark as follows:


If you think the adjective *very slightly* represents what you feel about the sample, then you should place your check mark as follows:


IMPORTANT

1. Place your check marks in the middle of the spaces, not on the boundaries at the colons.
2. Be sure you check every scale for every speech sample ---- do not skip any.
3. Never put more than one check mark on a single scale.

Sometimes it may seem as though you have had the same item before on the questionnaire. This will not be the case, so please do not look back and forth through the items. Do not try to remember how you checked similar items earlier in the questionnaire. Make each item a separate and independent judgement. *Work at a fairly high rate of speed* throughout the questionnaire. We want your immediate impressions, your immediate reactions to the samples. On the other hand, please try your best to give us your true impressions.

Let us try an example.
<table>
<thead>
<tr>
<th></th>
<th>Sample Number 1 is/has/uses:</th>
</tr>
</thead>
</table>
APPENDIX D

INTELLIGIBILITY TEST
Name _____________________

Instruction:
You are about to hear 16 short speech samples, consisting of one or two sentences. (Don’t worry, they will last only for three to six seconds.) Please listen to the samples carefully and write down every word you hear as much as you can. You will hear each sample only once.

Example _________________________________________________________________
_________________________________________________________________
_________________________________________________________________

Sample 1
_________________________________________________________________
_________________________________________________________________
_________________________________________________________________

Sample 2
_________________________________________________________________
_________________________________________________________________
_________________________________________________________________

Sample 3
_________________________________________________________________
_________________________________________________________________
_________________________________________________________________

Sample 4
_________________________________________________________________
_________________________________________________________________
APPENDIX E

PERSONAL INFORMATION QUESTIONNAIRE
Name ____________________

**Questionnaire**

Please answer the following questions about yourself. Check (√) in the box, or fill in the space provided.

(1) Age ______________ years
(2) Sex □ male □ female
(3) How long have you been in the United States in total?
   For ______________ years and ______________ months
(4) Are you an undergraduate student or a graduate student?
   □ undergraduate □ graduate □ other (______________)
(5) Major(s) at The Ohio State University _____________________________
(6) Are you a native speaker of American English?
   □ YES (Please go directly to Question (10) on the next page)
   □ NO (Please answer the following questions)
(7) What is your native language? Please choose ONE.
   □ Mandarin □ Hindi □ Japanese
   □ Other (Please specify: _____________________________)
(8) What language do you speak most often on a daily basis?
   □ English □ Mandarin □ Hindi □ Japanese
   □ Other (Please specify: _____________________________)

What language do you speak occasionally?
   □ English □ Mandarin □ Hindi □ Japanese
   □ Other (Please specify: _____________________________)
(9) Provide information on your best TOEFL score. Was the format Paper-Pencil or Computer-Based?
   □ Paper and Pencil (PP)
   □ Computer-Based (CB)
   To which of the following categories does your best TOEFL score belong?
□ 650 or higher on Paper (PP), or 280 or higher on Computer (CB)
□ Higher than 600 PP or 250 CB but lower than 650 PP or 280 CB
□ Higher than 550 PP or 213 CB but lower than 600 PP or 250 CB
□ Higher than 500 PP or 173 CB but lower than 550 PP or 213 CB
□ Higher than 450 PP or 133 CB but lower than 500 PP or 173 CB
□ 450 PP and 133 CB and below

(10) Do you know any Japanese persons who speak English?
□ NO (Please go to Number (11))
□ YES (Please answer the following question.)

During the past year, on average, how often have you heard Japanese persons speak English? Check ONE box with the frequency that best describes your case.

□ 1 = I have almost never heard Japanese persons speak English, except for simple and short greetings and responses.
□ 2 = I might have heard Japanese persons speak English occasionally, but not on a regular basis.
□ 3 = I have heard Japanese persons speak English on a regular basis, but less frequently than once a month.
□ 4 = I have heard Japanese persons speak English almost every month or more often, but not as often as every week.
□ 5 = I have heard Japanese persons speak English almost every week or more frequently.

● In the following space, please provide comments, if any, on your familiarity with Japanese English.
Do you know any non-native English speakers (= persons for whom English is not their native language but speak English), other than Japanese speakers?
☐ NO
☐ YES  (Please answer the following question.)

*During the past year, on average*, how frequently have you heard non-native English speakers, other than Japanese speakers? Check ONE box with the frequency that best describes your case.

☐ 1 = I have almost never heard non-native speakers speaking English except for simple and short greetings and responses.
☐ 2 = I might have heard non-native English speakers occasionally, but not on a regular basis.
☐ 3 = I have heard non-native English speakers on a regular basis, but less frequently than once a month.
☐ 4 = I have heard non-native English speakers almost every month or more often, but not as often as every week.
☐ 5 = I have heard non-native English speakers almost every week or more frequently.

● If you know the native languages of the people referred to above, please write them down:

What languages: _________________________________________

● In the following space, please provide other comments, if any, on your contact with non-native English speakers, other than Japanese.

※ This is the end of the questionnaire. Thank you very much for your cooperation!!!
APPENDIX F

INFORMED CONSENT FORM
(FOR QUANTITATIVE STUDY PARTICIPANTS)
Informed Consent Form  
The Ohio State University

Title of the Project:  
Evaluative Reactions to, and Intelligibility of, Non-Native English Speech  
Among Non-Native Listeners

Description:

1. Purpose: The study in which you will be participating is intended to investigate how well listeners with various native language backgrounds (American English, Chinese, Hindi, and Japanese) understand English spoken by non-native speakers, and how the listeners evaluate it.

2. Procedures: If you agree to participate in this study, you will be asked to attend a session, in which you will:
   (1) listen to eight narrative speech samples, each of which lasts approximately 2 minutes,
   (2) rate the samples on a bipolar adjective scale,
   (3) listen to sentences extracted from each speech sample you have listened to in task one for the second time,
   (4) write down every word you heard, and
   (5) provide your demographic information on the questionnaire.

3. Duration: This session will last for approximately one hour, including the explanation of procedures.

4. Compensation: You will be paid $15.00 in cash when the session is completed.

5. Withdrawal: You can withdraw the consent of participation at any time during the study, without any penalty or prejudice.

6. Your written responses are confidential. No one, except for the researchers whose names are given, has access to them. All the data will be discarded six months after the project is completed.
* If you have any further questions or concerns, please contact the following persons:

Representative Researcher: Reiko Kachi
Doctoral Candidate
Foreign and Second Language Education
The Ohio State University
Home Phone: (614) 889-1037
Email Address: kachi.1@osu.edu

Principal Researcher (Adviser): Keiko K. Samimy
Associate Professor
Foreign and Second Language Education
The Ohio State University
Office Phone: (614) 292-7597
Email Address: samimy.2@osu.edu

Consent For Participation in Research

I consent to participating in the research entitled: Evaluative Reactions to, and Intelligibility of, Non-Native English Speech Among Non-Native Listeners.

Keiko K. Samimy, Principal Investigator, or her authorized representative Reiko Kachi has explained the purpose of the study, the procedures to be followed, and the expected duration of my participation. Possible benefits of the study have been described.

I acknowledge that I have had the opportunity to obtain additional information regarding the study and that any questions I have raised have been answered to my full satisfaction. Furthermore, I understand that I am free to withdraw consent at any time and to discontinue participation in the study.

Finally, I acknowledge that I have read and fully understand the consent form. I sign it freely and voluntarily. A copy has been given to me.

______________________________  __________________________
Participant’s Signature        Date

______________________________  __________________________
Researcher’s Signature         Date
APPENDIX G

INFORMED CONSENT FORM
(FOR QUALITATIVE STUDY PARTICIPANTS)
Informed Consent Form
The Ohio State University

Title of the Project:
Factors predicting non-native speakers’ evaluative reactions to non-native speech

Description:
1. Purpose: The study in which you will be participating is intended to investigate how
listeners judge the comprehensibility and accented-ness of English as an international
language.

2. Procedures: If you agree to participate in this study, you will be asked to attend a
session, in which you will:
   (1) Listen to eight narrative speech samples that you have already listened to in the
       previous session. A pause will be placed anytime necessary.
   (2) Answer questions regarding the speech samples and other things. Please be
       reminded that all your reactions are tape-recorded.

You can express any questions or concerns after listening to each sample, but the
researcher might not necessarily be able to answer them.

3. Duration: This session will last for approximately 45 minutes to one hour, including
   the explanation of procedures.

4. Compensation: You will be paid $15.00 in cash when the session is completed.

5. Withdrawal: You can withdraw the consent of participation at any time during the
   study, without any penalty or prejudice.

6. Your responses are confidential. No one, except for the researchers whose names are
given, has access to them. The recorded responses will be deleted six months after the
project is completed.
* If you have any further questions or concern, please contact the following persons:

Representative Researcher: Reiko Kachi  
Doctoral Candidate  
Foreign and Second Language Education  
The Ohio State University  
Home Phone: (614) 889-1037  
Email Address: kachi.1@osu.edu

Principal Researcher (Adviser): Keiko K. Samimy  
Associate Professor  
Foreign and Second Language Education  
The Ohio State University  
Office Phone: (614) 292-7597  
Email Address: samimy.2@osu.edu

Consent For Participation in Research

I consent to participating in research entitled: Factors Predicting Non-Native Speakers’ Evaluative Reactions To Non-Native Speech

Keiko K. Samimy, Principal Investigator, or her authorized representative Reiko Kachi has explained the purpose of the study, the procedures to be followed, and the expected duration of my participation. Possible benefits of the study have been described.

I acknowledge that I have had the opportunity to obtain additional information regarding the study and that any questions I have raised have been answered to my full satisfaction. Furthermore, I understand that I am free to withdraw consent at any time and to discontinue participation in the study.

Finally, I acknowledge that I have read and fully understand the consent form. I sign it freely and voluntarily. A copy has been given to me.

____________________________________  ____________________________
Participant’s Signature                  Date

____________________________________  ____________________________
Researcher’s Signature                   Date

245
APPENDIX H

SCATTERPLOTS FOR METRIC DATA
Figure H.1: Correlation between Familiarity with Japanese English and Evaluative Reactions

Familiarity with Japanese English

Kendal’s tau = -.01
Figure H.2: Correlation between Familiarity with Japanese English and Intelligibility
Figure H.3: Correlation between Familiarity with Japanese English and Exposure to Foreigners

Kendall’s tau = .15
Figure H.4: Correlation between Exposure to Foreigners and Evaluative Reactions
Figure H.5: Correlation between Exposure to Foreigners and Intelligibility

Kendall’s tau = -0.01
Figure H.6: Correlation between Evaluative Reactions and Intelligibility