THE ROLE OF MOTIVATION IN SECOND LANGUAGE PRONUNCIATION

Tao-Chih Wen, B.A.

Thesis Prepared for the Degree of

MASTER OF ARTS

UNIVERSITY OF NORTH TEXAS

August 2005

APPROVED:

Jenifer Larson-Hall, Major Professor
Patricia Cukor-Avila, Committee Member and
Chair of the Linguistics Division
Jiyoung Yoon, Committee Member
Brenda Sims, Chair of Graduate Studies in English
James T. F. Tanner, Chair of the Department of
English
Sandra L. Terrell, Dean of the Robert B. Toulouse
School of Graduate Studies
This thesis investigates the phonological ability of exceptional second language (L2) learners of English and their levels of motivation. This study is the first of its kind to do a large-scale examination of L2 learners whose first languages (L1s) do not belong to the same Indo-European language family as English. Fifteen non-native speakers (NNSs) of English filled out a questionnaire and produced four speech samples, including a picture description task, paragraph reading task, sentence reading and word reading task. Fifteen native speaker (NS) controls also produced the same speech samples. Four NSs judged all participants’ accents. Six NNSs scored as highly as NSs on some of the speech segments using a 2-standard deviation (SD) cut-off point. There was no significant correlation between their scores on pronunciation and motivation.
ACKNOWLEDGEMENTS

A special thank to Dr. Jenifer Larson-Hall, my major advisor, and Dr. Patricia Cukor-Avila and Dr. Jiyoun Yoon for their inspiration, time, and patience. I would also like to thank all my family, especially my husband, Ethan, for their support and for allowing me to pursue my dream. Finally, thanks to my friends Jamie, Rachel, David, and Steve for their help and encouragement.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>ii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>v</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>vi</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Background to the Problem</td>
<td>1</td>
</tr>
<tr>
<td>Review of the Literature</td>
<td>2</td>
</tr>
<tr>
<td>Motivation</td>
<td>6</td>
</tr>
<tr>
<td>Research Questions and Hypotheses</td>
<td>11</td>
</tr>
<tr>
<td>METHODOLOGY</td>
<td>12</td>
</tr>
<tr>
<td>Pre-Screening Interview</td>
<td>12</td>
</tr>
<tr>
<td>Participants</td>
<td>13</td>
</tr>
<tr>
<td>Experimental Group</td>
<td>13</td>
</tr>
<tr>
<td>Control Group</td>
<td>14</td>
</tr>
<tr>
<td>Judges</td>
<td>15</td>
</tr>
<tr>
<td>Speech Samples</td>
<td>15</td>
</tr>
<tr>
<td>Questionnaire</td>
<td>18</td>
</tr>
<tr>
<td>Procedure</td>
<td>19</td>
</tr>
<tr>
<td>Post-Test Interview</td>
<td>20</td>
</tr>
<tr>
<td>Data Editing</td>
<td>20</td>
</tr>
<tr>
<td>Rating</td>
<td>22</td>
</tr>
<tr>
<td>RESULTS</td>
<td>24</td>
</tr>
<tr>
<td>Non-Native Speakers (NNSs) Performance on Four Speech Samples</td>
<td>24</td>
</tr>
<tr>
<td>Picture Description</td>
<td>24</td>
</tr>
<tr>
<td>Paragraph Reading</td>
<td>24</td>
</tr>
<tr>
<td>Sentence Reading</td>
<td>24</td>
</tr>
<tr>
<td>Motivation Scores</td>
<td>26</td>
</tr>
<tr>
<td>Correlation Tests</td>
<td>26</td>
</tr>
</tbody>
</table>
DISCUSSION ..................................................................................................................................31
  Research Question Analysis .................................................................31
  Profiles of 15 NNSs .................................................................32
  Motivation Questionnaire ..........................................................40
CONCLUSIONS AND IMPLICATIONS..............................................................................42
  Conclusions.................................................................42
  Limitation of the Study .................................................................42
  Ideas for Future Research .................................................................43
  Implications.................................................................44
APPENDICES
  A.  PICTURE DESCRIPTION TASK ..............................................46
  B.  PARAGRAPH READING TASK ..................................................48
  C.  SENTENCE READING TASK ..................................................50
  D.  WORD READING TASK ..........................................................52
  E.  QUESTIONNAIRE .................................................................54
REFERENCES.................................................................................................................................60
### LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Characteristics of the 15 NNSs tested</td>
<td>13</td>
</tr>
<tr>
<td>2. Gender and age of the 15 NSs tested</td>
<td>14</td>
</tr>
<tr>
<td>3. NNSs average score (over four judges) on four speech samples</td>
<td>24</td>
</tr>
<tr>
<td>4. NS and NNS group average scores and S.D. on four tasks</td>
<td>25</td>
</tr>
<tr>
<td>5. NNSs score on motivation as a whole and on individual field</td>
<td>26</td>
</tr>
<tr>
<td>6. Pearson correlation between NNSs score on speech samples and their motivation</td>
<td>27</td>
</tr>
<tr>
<td>7. Pearson correlation between NNSs score on four speech samples and their age, LOR, AOA</td>
<td>29</td>
</tr>
<tr>
<td>8. Ranking of 15 NNSs on motivation scores</td>
<td>33</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Screen of computer program used for judges to rate picture description, paragraph reading and sentence reading tasks</td>
<td>22</td>
</tr>
<tr>
<td>2.</td>
<td>Screen of computer program used for judges to rate word accuracy</td>
<td>23</td>
</tr>
<tr>
<td>3.</td>
<td>Scatterplot of motivation and picture telling accent score</td>
<td>27</td>
</tr>
<tr>
<td>4.</td>
<td>Scatterplot of motivation and paragraph accent score</td>
<td>28</td>
</tr>
<tr>
<td>5.</td>
<td>Scatterplot of motivation and sentence production accent score</td>
<td>28</td>
</tr>
<tr>
<td>6.</td>
<td>Scatterplot of motivation and word production accuracy</td>
<td>29</td>
</tr>
</tbody>
</table>
INTRODUCTION

Background to the Problem

It is very difficult to achieve a native-like accent in a second language (L2) if a learner begins to learn the language as an adult. Although there have been several studies which have specifically targeted exceptional learners and have found non-native speakers (NNSs) who ‘pass’ as native speakers (NSs) (Bongaerts, Planken and Schils, 1995; Bongaerts, Summeren, Planken and Schils, 1997; Bongaerts, Mennen and Slik, 2000; Ioup, Boustagui, Tigi and Moselle, 1994; Moyer, 1999), studies which test normal populations of learners have not found any learners whose first age of intensive exposure was above age 12 who have been rated as completely native-like for phonology (Asher & Garcia, 1969; Oyama, 1976; Thompson, 1991; however, Flege, Munro, & MacKay, 1995, testing 240 Italian immigrants to Canada, did find some exceptions). In fact, Long (1990) proposed that learners above age 12 would not be able to achieve a native-like accent. Some researchers discount the findings of exceptional learners by saying that native-like pronunciation needs to be found in more naturalistic language samples (DeKeyser and Larson-Hall, 2005; Long, 2005). In this thesis, I proposed to test the production abilities of exceptional learners on a wider variety of measures than was used by Bongaerts et al. (1995) and Moyer (1999), and also to test the proposition that exceptional learners can be found among a population of English learners whose first languages (L1s) are outside the Indo-European family. Additionally, I tested the hypothesis that motivation may be the essential factor responsible for accounting for exceptional performance. This hypothesis has not previously been tested among a group of exceptional learners.
Review of the Literature

Can L2 learners whose age of arrival (AOA) is after age 12 achieve a native-like accent? This question has long been debated among researchers. As early as 1967, Lenneberg proposed a critical period, which says that “Foreign accents cannot be overcome easily after puberty.” Since then research regarding AOA and pronunciation (Asher and Garcia, 1969; Flege, Munro & MacKay, 1995; Moyer, 1999; Oyama, 1976; Piske, MacKay & Flege, 2001; Suter, 1976; Thompson, 1991) has concluded that, in general, global pronunciation in an L2 will be rated more nativelike for those who begin learning the L2 at a young age.

Long (1990) proposed a clear definition for the range of the critical period. After reviewing several studies regarding the ultimate attainment of L2 phonology for younger and older L2 learners, he concluded that “a native-like accent is impossible unless first exposure is quite early, probably before 6 in many individuals and by about age 12 in the remainder. Very high standards can be attained starting later, of course, but not, it seems, native-like standards” (Long, 1990, p. 266).

Since then, the notion that there was a critical period between the ages of 6 and 12 for L2 learners has been challenged by other researchers (Bongaerts, Planken and Schils, 1995; Bongaerts, Summeren, Planken and Schils, 1997; Bongaerts, Mennen and Slik, 2000; Flege, Munro and MacKay, 1995; Ioup, Boustagui, Tigi and Moselle, 1994; Moyer, 1999). L2 learners whose initial AOA was after the age of 12 were often defined as late L2 learners. Flege, Munro and MacKay’s study (1995) provided some evidence that late L2 learners can still achieve global pronunciation scores within the NS range. They tested 240 native Italian speakers whose AOA ranged from 2 to 23 years. These participants’ production of five English sentences that were not focused on sounds especially difficult for Italian speakers of English were examined and rated by
10 native English speakers. The rating scale ranged from no foreign accent (top), medium foreign accent (middle), to strongest foreign accent (bottom). Flege used a sliding level scale to record minute variations in the judge’s ratings, going from 0 (strongest foreign accent) to 255 (no foreign accent). The results revealed that six percent of the 120 participants whose AOAs were between 12 and 16 years were rated native English speakers; however, the exact AOAs of the six percent were unknown. This study also found that no NNSs whose AOAs were after 16 years of age could achieve a NS rating.

A more in-depth case study on an exceptional learner was conducted by Loup, Boustagui, Tigi and Moselle (1994). They recruited two late learners, Julie and Laura, along with two other NNSs and 3 NSs of Arabic. Julie and Laura are both NSs of English. Julie’s AOA in Arabic was 21, Laura about 24. The former did not receive any formal instruction about Arabic while the latter did. Julie was rated as a NS by eight out of 13 judges on a dichotomous scale where judges had to decide if the speaker was a NS or not for a spontaneous speech task where she had to describe her favorite recipe. Laura was rated as a NS by six out of 13 judges. This study suggests that there may be some exceptional late L2 learners with an AOA after the age of 12 who can still achieve a near-native accent.

Another example comes from Bongaerts, Planken and Schils (1995). They recruited a group of 10 Dutch learners of English whose AOA was after age 15 and who were recommended by their tutors or colleagues at colleges, as well as five NSs of English as controls. The participants were asked to talk in English about a topic for three minutes, read a short English text, 10 English sentences, and a list of 25 English words which were selected to include most English phonemes. Four NSs of English who had no experience in rating pronunciation were recruited and asked to rate each of the 27 speakers. The results showed that a group of 10 advanced L2 learners
passed as NSs of English on all four speech samples. However, methodologically there was a problem with this study because not all of the native-speaking controls were judged to be native-like. Bongaerts et al. (1995) hypothesized that different backgrounds of judges and speakers may have confounded the results. Obviously, the fact that NSs were not judged highly in some cases lessens our confidence in the overall results. Bongaerts et al. (1995) stated that the NSs were from southern England or the Midlands, and may have had a detectable regional accent. The L2 learners, on the other hand, were trained to speak Received Pronunciation (RP), a “prestigious variety of British English…which is the variety spoken by, for example, most news-readers and program-presenters on British national radio and television” (Bongaerts et al., 1995, p. 45). The judges came from York, in northern England. Therefore, it is possible that the group of the advanced L2 learners received higher scores because their accent was regarded as more neutral than the group of NSs. The accent of the group of NSs may have had a “slight regional accent” which the judges “were not so familiar with” (Bongaerts et al., 1995, p. 46)

In their next study, Bongaerts et al. (1997) revised the problem that had caused the NSs to be ranked poorly in the first study. They recruited a group of 11 NSs of Dutch whose AOAs were after age 12, (though the exact AOAs were unknown) and who were recommended by university-based English-as-a-foreign-language (EFL) experts. Nine of the 11 NSs of Dutch taught English at either a Dutch university or a Dutch teacher-training institute. They read six sentences in English and thirteen native English-speaking judges rated them on a scale of one to five. The results revealed that five out of 11 Dutch learners scored within the NS range; three out of the five successful late learners consistently achieved this range on all six sentences. Their study indicated that late L2 learners whose AOAs was after age 12 could pass as NSs, at least some of the time, on a sentence-length segment.
In their fourth study (2000), Bongaerts et al. tested the possibility that the result of their third study (1997) could be applied to L2 learners whose L1 was not so closely related to their L2. They recruited 30 NNSs of Dutch whose AOA was between the age of 11 and 34 and who spoke 11 different languages as well as 10 NSs of Dutch as controls. All participants read six sentences which included all Dutch phonemes and 21 NSs of Dutch rated their accents on a 5-point scale. The results showed that two late learners, one a NS of German and the other a NS of English, were within the NS range. The rest of the NNSs of Dutch, especially those whose L1s were not so closely related to Dutch such as Turkish, Czech, Spanish, Italian, Armenian, and Berber, failed on the sentence reading task.

As seen in the above studies, there were a few late L2 learners who exceeded Long’s prediction (1990), passing as NSs of their L2 with regard to pronunciation. However, even in the strongest case (Bongaerts et al., 1997) this was only for sentences. I wanted to explore whether exceptional learners could pass for NSs in a less constrained speech sample, and also whether it would be possible for learners of English whose L2 was not so closely related to their L2 (as in the case of Dutch and English) to also pass as NSs. In Flege et al.’s study (1995), the two languages that they tested were Italian and English, which are both Indo-European languages. The target languages that Bongaerts et al. (1997) focused on were Dutch and English, which belong to the same Germanic language family. So far, the only experiment that has reported successful late L2 learners who came from a different language family from their L2 was the study by Ioup et al. (1994). The only problem with the Ioup et al. study is that, because their subject Julie was rated simply as native or non-native, it is difficult to say that Julie ‘passed’ as a NS with all of the judges, although she came very close. Most recent studies have used a Likert scale for nativeness,
which means we can look at the mean scores and variation for NSs and see whether NNSs achieved a score within that variation.

Motivation

What is the factor that might account for exceptional L2 learners achieving native-like scores past the critical period? I wanted to explore the factor of motivation. However, in the literature on this subject motivation is considered important, but little empirical evidence exist to support a strong role for motivation (Klein, 1995; Bongaerts et al., 1995, 1997). Also in the literature, motivation does not have a major role in L2 pronunciation (Purcell and Suter, 1980), but is given a very small (less than five percent) or non-significant role (Oyama, 1976; Thompson, 1991; Dekeyser and Larson-Hall, 2005, in a review of the literature) in explaining variation in scores on phonological measures.

However, for those who can pass as NSs, there may be an emotional factor for them to persist in pursuing a near native or native accent. Bley-Vroman (1989) also provided a similar suggestion for the role of affect: “The central role of affect in foreign language learning is absolutely indisputable.” (Bley-Vroman, 1989, p. 49) The role of affect is clearly defined as motivation by researchers such as Tremblay and Gardner (1995). As early as 1959, Gardner and Lambert had suggested that L2 achievement was related not only to language aptitude, but also to motivation. Another researcher, Klein (1995), proposed three factors that contribute to the different performances of late L2 learners, among them ‘propensity,’ which means “the different motivations that push a learner forward in the acquisition of the mother tongue and a L2” (p. 261). Propensity was regarded as the most influential factor on the differential performances of late L2 learners.
In an empirical study, Suter (1976) tested, among 20 other factors, the importance of motivation as predictors for pronunciation accuracy. Five of his factors, economic motivation, social prestige motivation, integrative orientation, cultural allegiance and strength of the speaker’s concern about his pronunciation of English, were related to motivation, according to Gardner’s definition of motivation (1985). He recruited 61 NNSs of English who came from four different L1 backgrounds including Persian, Arabic, Japanese and Thai. He interviewed these participants and asked them to fill out a questionnaire. In addition, he asked them to mimic novel sounds that he produced to get a score for their aptitude on mimicry, and to take a personality test to get their score on extroversion-introversion. Then he asked them to produce free speech about a holiday in their country. The first two minutes of the free speech were cut out and fourteen NSs of English judged their accent on a 6-point scale. Then he conducted a statistical correlation to compare NNSs pronunciation scores and their responses to other tasks. The results showed that L1, the strength of the speaker’s concern about pronunciation—part of motivation—and the amount of conversation carried on at work and at school with NSs of English were the significant factors that account for L2 learners’ pronunciation.

Four years later, Purcell and Suter (1980) conducted a more detailed statistical analysis on Suter’s data (1976). After applying correlation, factor analysis and regression analysis to examine in detail the 20 predictors in Suter’s study (1976) that might explain late L2 learners’ pronunciation accuracy, they concluded that only four of them were “useful in accounting for the variance of subjects’ pronunciation accuracy scores” (Purcell and Suter, 1980, p. 282). Those included were L1, accounting for 41% of the variance, aptitude for oral mimicry, 14.1%, length of residence (LOR, including years in an English speaking country and months residing with a NS of English), 11.4%, and strength of concern for pronunciation accuracy, 7%. The first two predictors
were beyond the learners’ control and the last two predictors were under the learners’ control.
Nevertheless, although LOR was decided by learners themselves, they make the decision based on their career plan or life plan. Therefore, the only factor that learners have control over and is directly related to pronunciation is the strength of concern for pronunciation accuracy.

Bongaerts et al. (1995, 1997) came to similar conclusions on the importance of motivation as Purcell and Suter (1980). The advanced Dutch learners of English in these studies shared some common background experiences. First, all participants in the experimental group were highly advanced late L2 learners who received “a large amount of mainly British English input from both native and non-NSs of the language. Second, “they also received intensive training in Received Pronunciation (RP)” (1997, p.454) when they were freshmen at the university. Third, they all regarded it as very important to not have a Dutch accent when speaking English. However, instruction seemed to lose its importance in their fourth study. In this study, when checking the two successful late learners’ backgrounds, Bongarets et al. (2000, p. 306) reported that while “one of them had received some instruction in the pronunciation of Dutch, the other one had received no such instruction at all….Both expressed a strong personal and professional interest in achieving very good pronunciation.”

Similar implications on the role of motivation arise from Moyer’s study (1999). She examined 24 late NNSs of German, whose AOA was after age 11. Their L1 was English, and they were instructors of German at a university. They were regarded by the researcher as highly motivated learners and speakers because of their profession. She had them read a list of 24 words, eight sentences, and a paragraph, as well as give some personal information in free speech. Four judges who were NSs of German rated their accents. Then she conducted correlation and regression analysis on the data. The results revealed that the mean score of all four speech samples
across 24 non-NSs was significantly different from the NS mean score while professional motivation was the most significant factor, accounting for 41% of the variance in NNSs pronunciation. There was only one learner who passed as a NS on all four speech samples, while the rest of the advanced learners failed to do so. This successful late L2 learner had an AOA of 22 and had received formal instruction of German for only five years. He self-reported that he had “a strong desire to acculturate and to sound German” (p. 98). The rest of the NNSs who failed expressed in the interview that “perfect pronunciation was neither realistic nor necessary for overall fluency.” (p. 88)

Therefore, because motivation is one of the few factors that individual learners have firm control over and based on the above empirical research, I decided to explore whether motivation might be a more important factor than some researchers have thought (Oyama, 1976; Thompson, 1991; DeyKeyser & Larson-Hall, 2005). It might be better illuminated if a more rigorous psycholinguistic test instrument were brought to bear on the issue. To substantiate this hypothesis, it is necessary to look into studies specifically about motivation before constructing an instrument to measure it.

Motivation study is one field of L2 acquisition that tries to investigate the factors that contribute to L2 learners’ performance. Earlier motivation studies were linked to the field of socio-psychology, of which Gardner’s studies were the representative. Gardner (1985, p. 50) proposed that “motivation involves four aspects, a goal, effortful behavior, a desire to attain the goal, and favorable attitudes toward the activity in question.” He also said that “individual differences in motivation are reflected in the latter three aspects listed above.” According to this definition, Gardner developed a questionnaire called the Attitude/Motivation Test Battery (AMTB). Gardner’s AMTB has several good points. First, it was extensively tested. Second, it
follows principles of sound questionnaire design. For example, it uses multi-item scales, which means that the questionnaires contain a group of questions that focus on the same phenomenon, not just one or two questions about a single area. Multi-item scales can “maximize the stable component that the items share and reduce the extraneous influences unique to the individual items” (Dornyei, 2003, p. 34). Third, it uses multiple-choice items, which are easy for learners to answer (Dornyei, 2003). However, when I examined Gardner’s AMTB (1985), I found two disadvantages. First, questions in some sections were limited to students who were still taking language classes. The questionnaire failed to focus on really advanced L2 learners who do not need any more formal instruction or pronunciation practice. Second, these sections targeted L2 learners’ motivation towards general language learning. It was not appropriate for my need to measure L2 learners’ motivation for pronunciation specifically. In order to construct a rigorous test of motivation, I adapted his AMTB for learners who were not in school.

The motivation factors that I chose to measure were based on Scherer’s model (1984) of stimulus appraisals, recommended by Schumann (1997). Scherer (1984) suggested that how people process stimulus events may affect their emotional reaction and argued that “the differentiated emotional states that we label anger, fear, joy, or sadness, for example, are the result of the successive outcomes of a series of stimulus evaluation checks (SEC).” The list of SECs includes novelty, intrinsic pleasantness, goal/need significance, coping potential, and norm/self compatibility. Schumann (1997) adopted this notion and changed its name to stimulus appraisals. He further applied stimulus appraisals to language learning and argued that “emotional reactions influence the attention and effort devoted to learning, and… patterns of appraisals may underlie what has been considered motivation in SLA” (Schumann, 1997, p. 8). L2 learners evaluate events about learning materials, environment, teachers, classmates, and themselves, based on the
above five stimulus appraisals. If the stimulus events, e.g., language learning, make them feel motivated through the screening of appraisals and the motivated behavior persists after a large amount of time, it will finally lead to success in language learning.

Research Questions and Hypotheses

This study addresses the following research questions:

1. Can we find speakers whose L1 is non-Indo-European with an AOA of 12 or more who pass for NSs in English on both constrained and more spontaneous speech samples?

2. Is motivation a key factor in explaining the success of those late L2 learners who pass as NSs?

My hypothesis is that when motivation is measured with a more rigorous and psycholinguistically valid instrument than has hitherto been used in this area, there will be a correlation between motivation and pronunciation scores.
METHODOLOGY

Pre-Screening Interview

In order to find a suitable group of exceptional learners, I prescreened 28 non-native speakers (NNSs), seven of them by phone and the rest of them through interviews in person. These candidates were located by recommendations from friends and acquaintances. The candidates were prescreened to ensure they fit the criteria that their first languages (L1s) would be outside of the Indo-European family, that their age of arrival (AOA) was 12 or older and that they were willing to participate in the study. Another purpose for the prescreening interview was to choose those whose English pronunciation was with undetectable regional accent. Because I did not want the NNSs that I interviewed to be aware of this fact, I asked them to provide some general information that I required in the first part of my questionnaire such as age, AOA, length of residence (LOR), and experiences in learning English pronunciation. While the potential participants gave me their information I made an informal decision as to whether they sounded like native speakers (NSs). As I am not a NS myself, this could not be a perfect judgment. However, it was sufficient for my purpose of recruiting a group of speakers whose accent would then be more rigorously tested by NS judges.

Sixteen NNSs passed the prescreenings and were subsequently recorded after the interviews. However, subject 106 was recruited without any interview because she claimed to have a native-like accent. Nevertheless, her actual pronunciation when recorded was found to be quite accented. Therefore, this participant was dropped and the final number of non-native participants was 15.
Participants

*Experimental Group*

Fifteen NNSs of English (three male and 12 female, age range from 21 to 55, mean age 28.3; AOA from 12 to 25, mean AOA 17.7) whose L1 was Mandarin, Japanese or Korean comprised the experimental group (see Table 1 for a summary of their characteristics). Their age of first-time learning English (AOL) ranges from six to 18, mean AOL 10.5. The English classes they had before formal English instruction at age 12 were about grammar, reading, or vocabulary. Classes were about an hour each time, five times a week. Some of them continued the classes until they received formal English instruction while others only went to classes for a period of time ranging from a couple of months to one or two years. They were advanced late second language (L2) learners who spoke English every day either at school or at work. Eight of them are in the workforce and the rest of them are students at colleges in the Dallas/Fort Worth metroplex area. They all came to the U.S. to study, either for junior high, senior high, or college. All NNSs in this group received $15 for their participation, and the experiment took approximately one hour.

Table 1. Characteristics of the 15 NNSs tested (AOA= age participants first came to U.S., AOL= age participants first began to study English before formal instruction; LOR= length of residence that participants have lived in U.S. since they came here).

<table>
<thead>
<tr>
<th>ID</th>
<th>Gender</th>
<th>L1</th>
<th>Age</th>
<th>AOA</th>
<th>AOL</th>
<th>LOR</th>
<th>Degree earned or working on</th>
<th>status</th>
</tr>
</thead>
<tbody>
<tr>
<td>102</td>
<td>F</td>
<td>Cantonese</td>
<td>23</td>
<td>19</td>
<td>7</td>
<td>4</td>
<td>MA student</td>
<td></td>
</tr>
<tr>
<td>103</td>
<td>M</td>
<td>Cantonese</td>
<td>23</td>
<td>20</td>
<td>6</td>
<td>2.5</td>
<td>BFA student</td>
<td></td>
</tr>
<tr>
<td>104</td>
<td>F</td>
<td>Korea</td>
<td>55</td>
<td>20</td>
<td>18</td>
<td>35</td>
<td>BA working</td>
<td>working</td>
</tr>
<tr>
<td>105</td>
<td>F</td>
<td>Taiwanese</td>
<td>22</td>
<td>14</td>
<td>10</td>
<td>8</td>
<td>BA student</td>
<td></td>
</tr>
<tr>
<td>107</td>
<td>F</td>
<td>Mandarin</td>
<td>32</td>
<td>20</td>
<td>6</td>
<td>13</td>
<td>MA working</td>
<td>working</td>
</tr>
<tr>
<td>108</td>
<td>F</td>
<td>Mandarin</td>
<td>21</td>
<td>19</td>
<td>12</td>
<td>2</td>
<td>BA student</td>
<td></td>
</tr>
<tr>
<td>109</td>
<td>F</td>
<td>Mandarin</td>
<td>35</td>
<td>14</td>
<td>12</td>
<td>20.5</td>
<td>PH. D. working</td>
<td>working</td>
</tr>
<tr>
<td>110</td>
<td>F</td>
<td>Mandarin</td>
<td>30</td>
<td>25</td>
<td>9</td>
<td>5.5</td>
<td>MA working</td>
<td>working</td>
</tr>
<tr>
<td>111</td>
<td>F</td>
<td>Mandarin</td>
<td>24</td>
<td>18</td>
<td>10</td>
<td>6</td>
<td>BA student</td>
<td></td>
</tr>
</tbody>
</table>

(table continues)
Table 1. (continued)

<table>
<thead>
<tr>
<th>ID</th>
<th>Gender</th>
<th>L1</th>
<th>Age</th>
<th>AOA</th>
<th>AOL</th>
<th>LOR</th>
<th>Degree earned or working on status</th>
</tr>
</thead>
<tbody>
<tr>
<td>112</td>
<td>F</td>
<td>Japanese</td>
<td>22</td>
<td>18</td>
<td>13</td>
<td>4</td>
<td>BA student</td>
</tr>
<tr>
<td>113</td>
<td>M</td>
<td>Mandarin</td>
<td>23</td>
<td>12</td>
<td>9</td>
<td>10</td>
<td>BS student</td>
</tr>
<tr>
<td>114</td>
<td>F</td>
<td>Japanese</td>
<td>30</td>
<td>23</td>
<td>10</td>
<td>7</td>
<td>MA working</td>
</tr>
<tr>
<td>115</td>
<td>F</td>
<td>Taiwanese</td>
<td>25</td>
<td>17</td>
<td>12</td>
<td>8</td>
<td>BA working</td>
</tr>
<tr>
<td>116</td>
<td>M</td>
<td>Mandarin</td>
<td>30</td>
<td>24</td>
<td>10</td>
<td>6.5</td>
<td>MA working</td>
</tr>
<tr>
<td>117</td>
<td>F</td>
<td>Korea</td>
<td>30</td>
<td>13</td>
<td>13</td>
<td>17</td>
<td>BA working</td>
</tr>
</tbody>
</table>

Control Group

Fifteen NSs of English (nine male and six female, age range from 18 to 55, mean age 24.2) were recruited to serve as a comparison to the NNSs. Table 2 gives their gender and age.

According to Long (1990, p. 274), it is necessary to have “NS baseline data if any gaps in the competence of advanced NNSs are to be revealed.” These 15 participants were students at the University of North Texas (UNT). None of them had linguistic experience. They were prescreened by a NS of English to have no identifiable regional accent because the previously discussed study by Bongaerts et al. (1995) had indicated that using controls and judges with identifiable regional accents might lead to problems in accurately judging NS status. The controls participated in this study voluntarily, without any pay. The production tasks took approximately 15 minutes for them to complete.

Table 2. Gender and age of the 15 NSs tested.

<table>
<thead>
<tr>
<th>ID</th>
<th>Gender</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>201</td>
<td>F</td>
<td>20</td>
</tr>
<tr>
<td>202</td>
<td>F</td>
<td>20</td>
</tr>
<tr>
<td>203</td>
<td>M</td>
<td>23</td>
</tr>
<tr>
<td>204</td>
<td>M</td>
<td>21</td>
</tr>
<tr>
<td>205</td>
<td>M</td>
<td>25</td>
</tr>
<tr>
<td>206</td>
<td>M</td>
<td>25</td>
</tr>
<tr>
<td>207</td>
<td>F</td>
<td>26</td>
</tr>
<tr>
<td>208</td>
<td>F</td>
<td>19</td>
</tr>
<tr>
<td>209</td>
<td>M</td>
<td>55</td>
</tr>
<tr>
<td>210</td>
<td>M</td>
<td>20</td>
</tr>
<tr>
<td>211</td>
<td>F</td>
<td>29</td>
</tr>
<tr>
<td>212</td>
<td>M</td>
<td>20</td>
</tr>
<tr>
<td>213</td>
<td>M</td>
<td>23</td>
</tr>
<tr>
<td>214</td>
<td>F</td>
<td>18</td>
</tr>
<tr>
<td>215</td>
<td>M</td>
<td>19</td>
</tr>
</tbody>
</table>
Judges

Four NSs of English, two male and two female, served as judges for this study. According to Piske, MacKay and Flege (2001), so far there is no definite conclusion on how many judges are needed to have a reliable measure on the degree of foreign accent, and normal numbers of judges range from 1 to 85. Piske MacKay and Flege (2001) proposed that if the degrees of foreign accent are within a small range, it is better to have more judges to detect the differences. The number of judges in this study is at the lower end of this range, but this was due to the fact that I could not pay judges for their time. All of the judges in this experiment were UNT linguistics graduate students who had at least one year of linguistic experience. According to McLaughlin (1978, pp. 58), “The linguist is far more likely to perceive an accent than the ordinary person.” The results of Thompson’s study (1991) also confirmed this claim. She mentioned “experienced raters were more reliable in estimating the degree of accentedness across samples than were inexperienced raters.” (p. 199) All four judges participated in the study voluntarily without pay.

Speech Samples

All native subjects were tested in a quiet room at the university, but only three NNSs were recorded in the same room. The rest of them, because they lived far from campus, were recorded in their homes or offices. The recording conditions were not as quiet as that at the university. If this difference had an effect, it would be only to make it more difficult for the NNS to pass as NS. In addition, there is no guarantee that their house or office is a quiet place suitable for recording.

Four speech samples were obtained from each participant. They were, in the order that participants performed them:

1. A free speech task using a series of pictures.

2. A paragraph reading task.
3. A sentence reading task.
4. A word reading task.

The reason for ordering speech samples from picture description to a list of words was to progress from a focus on global pronunciation to pronunciation on specific sounds so that subjects would not pay attention to specific sounds until they were reading the word list. Bongaerts et al. (1995, p. 37) concluded that picture description requires “the least degrees of monitoring,” while reading a list of words requires “the highest.” Researchers differ drastically about which task participants will perform better on. Bongaerts et al. (1995) suggested that NNSs would perform better on constrained tasks such as word list or sentences because they could monitor their pronunciation better. On the contrary, the results of Oyama’s (1976) and Thompson’s (1991) studies revealed that NNSs performed better on more unconstrained speech tasks such as free speech. The reason for this phenomenon may be that in free speech, participants can avoid sounds that are difficult for them to pronounce or choose synonym to replace words that contain those difficult sounds. In the end, I decided to follow Bongaerts et al. (1995).

The picture description task was used in order to get more spontaneous speech samples than could be collected through paragraph reading, but the speech sample was constrained to have only six pictures for participants to describe so that every speaker would use basically similar words and grammar. This is the type of speech sample that DeKeyser and Larson-Hall (2005) and Long (2005) proposed as a more rigorous test of NS status. The actual pictures used in the task were taken from Nunan (1999) (used with permission; see Appendix A for details). For picture descriptions, there were a series of six pictures on a sheet of paper. Subjects were asked to describe each picture in at least two sentences to make them into one story or several different
stories. Subjects could prepare in advance by writing down their sentences or just take some time to think about it.

The paragraph, sentence and word list reading tasks have been used previously in many studies of non-native pronunciation (Flege et al., 1995b; Bongaerts et al., 1995; Bongaerts et al., 1997; Moyer, 1999). The paragraph reading tasks used (see Appendix B for details) were adopted from Reader’s Digest (used with permission) to include every sound that had been chosen for the word reading task. For paragraph reading, subjects were presented three paragraphs on a sheet of paper and asked to look over them before reading.

The sentence reading task (see Appendix C for details) was designed by a NS of English to form a coherent sentence that included words that had been chosen for the list reading. Subjects saw 12 sentences in random order on the computer screen and were asked to look over them before reading.

When creating the word task, I chose words from Cook (2000) which he hypothesized would be difficult for Mandarin, Japanese and Korean learners to pronounce. The words included three contrasts which are not found any of the three languages (see Appendix D for details).

1. Tense /i/ and lax /ɪ/. Neither Mandarin, Japanese nor Korean has a contrast between /i/ and /ɪ/. Speakers of all three languages pronounce both as an /i/ but distinguish them by length. In other words, tense /i/ is realized per conventional English pronunciation as a diphthong [ɪj] but lax /ɪ/ is realized as a monophthongized [i]. Therefore, the word sit [sɪt] is realized as [sɪ] (Cook, 2000).

2. Voiced and voiceless word-final contrast. Mandarin has no contrast between voiced and voiceless final consonants. Therefore, many Mandarin speakers just do not pronounce the consonant found in coda position (Cook, 2000). Contrary to Mandarin speakers, Japanese speakers insert epenthetic vowels to form consonant-vowel (CV) syllables on final consonants. Therefore, What time is it? sounds like [wato taimu izu ito] (Cook, 2000, p. 178) Korean speakers both epenthesize final consonants and devoice final voiced consonants. Therefore, dog may sound like either [dοgu] or [dοk] (Cook, 2000, p. 191).
3. Word-final /l/ and /r/. In Mandarin, there are no word final /l/ and /r/. For a word like *fail* Mandarin speakers would pronounce it as [fəʊ] (Cook, 2000, p. 175). Speakers of the three languages often pronounce /r/ at the end of a word just as an [əː]. Therefore, the word *car* is pronounced as [kəː] (Cook, 2000, p. 176, 179, 192).

Before reading words, participants were presented with the list of 18 words on the computer screen and asked to read over them first. The NSs did not take the time to read over the list, while the NNSs normally spent a little time preparing for it.

**Questionnaire**

Participants filled out both a background questionnaire and a specific questionnaire concerning motivation. The background questionnaire asked about age, gender, L1, degree earned or working on, other languages spoken, LOR, AOA, and AOL. Motivational questions were rated on a scale of 1 to 10 (1= strongly disagree, 10= strongly agree). As mentioned previously, these were developed from Scherer’s model (1984), which examined subjects’ stimulus appraisals on novelty (questions 1 and 4), pleasantness (question 5, 6, 7, 11, 21, 26 and 30), coping potential (question 14, 15, 17, 18, 19 and 20), goal/need significance (question 9, 10, 12, 22 and 29), and norm/self compatibility (question 2, 3, 8, 13, 16, 23, 24, 25, 27 and 28). I adapted 10 questions of Gardner’s Attitude/Motivation Test Battery (AMTB) and revised them into seven questions in my motivation questionnaire. Then I classified the seven questions into the five motivation areas that I chose to investigate. I designed the rest of the questions in the motivation questionnaire according to the five fields of motivation. The following are some example questions. A full list of the questions can be found in Appendix E.

**Novelty:**

I want to learn more about English because it is a medium for me to get to know the world.

English is not important for me to learn about because it is not necessary in the world.
Pleasantness:

I like to watch TV, listen to the radio, listen to songs or watch movies in English.

I like to speak English.

Coping potential:

I pay careful attention to how Americans pronounce words.

I spend most of the time speaking English at work or school.

Goal/need significance:

Pronouncing like a NS of English is important for me.

Better pronunciation helps me in my career plan or study.

Norm/self compatibility:

I prefer to stay with the group of people who speak the same native language as I do.

I spend most of the time with Americans.

There were a total of 30 questions on the motivation questionnaire. The questionnaire included positively and negatively worded questions to test if subjects were consistent in answering the same questions. Later, scores on negative questions were transformed into scores on positive questions to make sure all the questions were investigated in the same direction (e.g., if the participant rated 1 on a negative question like “I don’t like to speak English”, it was transformed into a 10, a 2 into a 9, and so on.)

Procedure

First, all participants filled out a consent form and the NNSs also filled out the background and motivation questionnaires, which took about 10 minutes. Next, I recorded the four speech
samples. The recording took about 15 minutes. After that, the NNS took an aptitude test which took about 35 minutes.¹

Post-Test Interview

While participants were taking the aptitude test, I examined their responses to the questionnaire. If there was any information or rating that I felt was unclear (e.g., the respondent gave a low score on a positively worded question), I interviewed the respondent again after the test. However, I did not record the interview and asked the respondents different questions according to their answers to the motivation questionnaire. Then I wrote down their explanations under the questions that they rated.

Data Editing

The speech samples were recorded directly onto a desktop or a laptop computer using a high quality digital microphone with a 96KHz sample rate and 16 bit resolution. The data were normalized to 100% by using a sound-editing software program.

One eight to ten second segment which was free from noticeable pauses or grammatical errors was edited out of the picture descriptions task for all 30 speakers. According to Bongaerts et al. (1995), eight seconds is enough for NSs to reliably judge the existence of a foreign accent. The decision of where to edit was made by listening to the entire file, then going back to the beginning of the task and choosing the first sentence that was error free. This process was done by the author, a NNS of English. Unfortunately, no recheck was done by a NS on the part that was selected from each participant’s picture-describing speech to make sure that there were no grammar mistakes. A later check revealed that there were four NNSs and two NSs who had grammar mistakes in their speech. This may cause judges to recognize these NNSs more easily and mistake those NSs as NNSs and wider variation in NSs performance.

¹ The results of the aptitude test will be reported in a future study by Dr. Jenifer Larson-Hall.
The first sentence of the paragraph reading which did not contain noticeable pauses, stuttering, or mis-pronunciations was also cut. I started with productions from the second paragraph, picking the first sentence which fit my criteria, as described in the preceding paragraph. Clearly, having disfluencies would make it easier for judges to pinpoint a non-native accent. According to Oyama (1976), features other than accents, such as prosodic aspects, may affect accent rating, although judges are told to pay attention to foreign accents only.

For sentence reading, only six of the original 12 sentences were used because there was stuttering or mispronunciation in some of the sentences for either NSs or NNSs. The six sentences chosen were the same for each speaker.

For word reading, all 30 speakers read 18 words, but only six of the original 36 words were used for the reason mentioned above. These six words were also the same for each speaker. In future research, I would follow the method used by other researchers (Flege et al., 1995) and target one special part of the reading and have participants read it over until it sounded fluent on the recording.

The voice recordings from the picture descriptions (30), paragraph readings (30), sentence readings (180) and word readings (180) were randomized through a computer program\(^2\). The order in which the segments were played was randomized. For the picture descriptions and paragraph readings each judge would hear 30 speakers’ voices, but each judge heard these in a different order. For sentence and word readings, all 180 sentences and words (six for each speaker, all 30 speakers) appeared in completely randomized order so that no order effects would be obtained from the judges.

\(^2\) The computer program is written by Ammon Larson of Microsoft and is available upon request from Dr. Jenifer Larson-Hall.
Rating

The ratings were conducted on a computer in a quiet room. Judges were told that they would hear both NSs and NNSs of English and they were going to rate their accent. The judges completed a rating practice before actually rating the picture descriptions, paragraph readings, and sentence readings. They also completed a rating practice for word accuracy because the rating methods were different between word rating and the other three. The judges heard two speakers in the rating practice, neither one of which was a part in the actual test.

After practicing, judges did the formal rating. For the picture, paragraph and sentence ratings the judges rated the accent on the computer screen on a scale of one to seven (1=strong foreign accent, 7=no foreign accent), as shown in Figure 1. After judges heard a speaker’s voice on one of the speech samples they used the number buttons on the keyboard to rate the speaker. The computer would not broadcast the next speaker’s voice until a score was given to the previous speaker.

![Figure 1](image)

Figure 1. Screen of computer program used for judges to rate picture description, paragraph reading and sentence reading tasks.
In the word list task, the judges rated the speakers’ accuracy on words. On the left and right sides of the computer screen judges saw a minimal pair and four icons: definitely L, possibly L, definitely R, possibly R, as shown in Figure 2.

![Figure 2. Screen of computer program used for judges to rate word accuracy.](image)

L meant the left side of the computer screen and R the right. If the judges thought the word that the speaker said was definitely the word on the left or right, they clicked the icon “definitely L” or “definitely R” and the speaker got one point. If the judges were not sure, they clicked “possibly L” or “possibly R” and the speaker got zero points. The next sound would not appear if the judge did not click the icon with the mouse. Every 50 sentences or words, judges were given a choice to take a break or resume the judging.
RESULTS

Non-Native Speakers (NNSs) Performance on Four Speech Samples

Data were analyzed using Statistical Package for the Social Sciences 13th version (SPSS 13). I adopted the standard used in Flege, Munro and MacKay (1995) for judging native-like performance, which is that if a NNS can achieve a score within two standard deviations (SD) of the NS’s mean, 99.5% of all native speaker (NS) variance, then the NNS would be considered to ‘pass’ as a NS. The results are shown in Table 3.

Table 3. NNSs average score (over four judges) on four speech samples. Shaded areas indicate subjects who achieved within NSs’ norm (maximum score on the picture, paragraph and sentence task is 7; maximum score on the word task is 1).

<table>
<thead>
<tr>
<th>ID</th>
<th>Picture</th>
<th>Paragraph</th>
<th>Sentence</th>
<th>Word</th>
</tr>
</thead>
<tbody>
<tr>
<td>102</td>
<td>6</td>
<td>2</td>
<td>6.33</td>
<td>0.92</td>
</tr>
<tr>
<td>103</td>
<td>3.5</td>
<td>3.75</td>
<td>3.67</td>
<td>0.71</td>
</tr>
<tr>
<td>104</td>
<td>4.75</td>
<td>6.5</td>
<td>3.04</td>
<td>0.58</td>
</tr>
<tr>
<td>105</td>
<td>6.75</td>
<td>2.75</td>
<td>4.96</td>
<td>0.83</td>
</tr>
<tr>
<td>107</td>
<td>4.5</td>
<td>3.5</td>
<td>4.96</td>
<td>0.79</td>
</tr>
<tr>
<td>108</td>
<td>5.5</td>
<td>3.25</td>
<td>4.5</td>
<td>0.75</td>
</tr>
<tr>
<td>109</td>
<td>3.75</td>
<td>3.75</td>
<td>2.5</td>
<td>0.67</td>
</tr>
<tr>
<td>110</td>
<td>6.5</td>
<td>4.25</td>
<td>4.5</td>
<td>0.96</td>
</tr>
<tr>
<td>111</td>
<td>4.25</td>
<td>2.25</td>
<td>2.54</td>
<td>0.67</td>
</tr>
<tr>
<td>112</td>
<td>4.5</td>
<td>2.75</td>
<td>2.96</td>
<td>0.63</td>
</tr>
<tr>
<td>113</td>
<td>3</td>
<td>7</td>
<td>3.13</td>
<td>0.71</td>
</tr>
<tr>
<td>114</td>
<td>3.25</td>
<td>4.5</td>
<td>6.8</td>
<td>0.96</td>
</tr>
<tr>
<td>115</td>
<td>6.75</td>
<td>2.75</td>
<td>5.17</td>
<td>0.96</td>
</tr>
<tr>
<td>116</td>
<td>5.75</td>
<td>4.25</td>
<td>4.21</td>
<td>1</td>
</tr>
<tr>
<td>117</td>
<td>5</td>
<td></td>
<td></td>
<td>0.83</td>
</tr>
</tbody>
</table>

Picture Description

The NSs mean range from four judges was 6.49-7. Three out of fifteen NNSs, participants 105, 110, 115, achieved within this range. Their scores were 6.75, 6.5, and 6.75.

Paragraph Reading

The NSs’ mean range was 6.74-6.85. One NNS, participant 115, achieved this range. Her score was 7, which was as perfect as many NS scores.

Sentence Reading

The NSs’ mean range was 6.64-6.78. Again, participant 115 achieved within this range.

Word accuracy. The NSs’ mean range was .89-.93. Five NNSs, participant 102, 110, 114, 115, 116, achieved this range.
As we can see, there was one NNS, participant 115, who consistently achieved within the NS’s range of mean score on all four speech samples.

However, the mean score of 15 NNSs performance on all four speech samples was not within the range of the NSs’ and their scores vary a lot more than NSs, as shown in Table 4.

Table 4. NS and NNS group average scores and S.D. on four tasks (maximum score on the picture, paragraph and sentence task is 7; maximum score on the word task is 1).

<table>
<thead>
<tr>
<th>status</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>picture</td>
<td>NNS</td>
<td>15</td>
<td>4.9167</td>
</tr>
<tr>
<td></td>
<td>NS</td>
<td>15</td>
<td>6.9000</td>
</tr>
<tr>
<td>paragraph</td>
<td>NNS</td>
<td>15</td>
<td>3.9500</td>
</tr>
<tr>
<td></td>
<td>NS</td>
<td>15</td>
<td>6.9500</td>
</tr>
<tr>
<td>Sentence Ave</td>
<td>NNS</td>
<td>15</td>
<td>4.2720</td>
</tr>
<tr>
<td></td>
<td>NS</td>
<td>15</td>
<td>6.9108</td>
</tr>
<tr>
<td>Word Accuracy</td>
<td>NNS</td>
<td>15</td>
<td>.79713</td>
</tr>
<tr>
<td></td>
<td>NS</td>
<td>15</td>
<td>.96967</td>
</tr>
</tbody>
</table>

An independent-sample t-test was conducted to compare the mean scores for each of the four speech samples between the NS and NNS groups. Because there are four related measures, the t-test will be considered significant if found at the p<.0125 level. Levene’s test for equal variances was significant, so the t-test results with equal variances not assumed was used. There were significant differences in all cases: t(15)= -6.04, p=.00 (picture description task); t(14)= -7.64, p=.00 (paragraph reading task); t(14)= -7.07, p=.00 (sentence reading task); t(16)= -4.68, p=.00 (word reading task).

I calculated effect sizes to measure the magnitude of the differences between NSs and NNSs. For the picture task, eta squared = .57, which is a large effect size (Cohen, 1988) and suggests that 57% of the variance in pronunciation scores is explained by the categorization into NS and NNS groups. For the paragraph reading, eta squared=.68, again, a very large effect size. For the sentence reading, eta-squared=.64, and for the word accuracy, eta-squared=.44. This
means that not only were differences significant, but the magnitude of the differences was also large.

Motivation Scores

Next, I calculated scores on motivation based on the NNSs responses to the motivation questionnaire. The results shown in Table 5 suggest that most of these participants’ scores on motivation as a whole and on individual fields of motivation were fairly high.

Table 5. NNSs score on motivation as a whole and on individual field (average score for the questions in each area given; maximum score in each case is 10).

<table>
<thead>
<tr>
<th>ID</th>
<th>102</th>
<th>103</th>
<th>104</th>
<th>105</th>
<th>107</th>
<th>108</th>
<th>109</th>
<th>110</th>
<th>111</th>
<th>112</th>
<th>113</th>
<th>114</th>
<th>115</th>
<th>116</th>
<th>117</th>
</tr>
</thead>
<tbody>
<tr>
<td>Novelty</td>
<td>8.50</td>
<td>9.5</td>
<td>9</td>
<td>10</td>
<td>9.5</td>
<td>10</td>
<td>8.5</td>
<td>8</td>
<td>8</td>
<td>9.5</td>
<td>10</td>
<td>9</td>
<td>8.5</td>
<td>7.5</td>
<td>7.5</td>
</tr>
<tr>
<td>Pleasantness</td>
<td>8.43</td>
<td>8.71</td>
<td>9.71</td>
<td>8.43</td>
<td>8.71</td>
<td>9.71</td>
<td>8.14</td>
<td>7.86</td>
<td>7.86</td>
<td>7.86</td>
<td>7.86</td>
<td>9</td>
<td>8.29</td>
<td>7.43</td>
<td>4.86</td>
</tr>
<tr>
<td>Coping potential</td>
<td>7.83</td>
<td>8.33</td>
<td>8.17</td>
<td>8.33</td>
<td>8.33</td>
<td>8.17</td>
<td>7.5</td>
<td>6.17</td>
<td>7.67</td>
<td>8.17</td>
<td>8</td>
<td>8.5</td>
<td>8.83</td>
<td>6.83</td>
<td>6.5</td>
</tr>
<tr>
<td>Goal/Need significance</td>
<td>9</td>
<td>8</td>
<td>10</td>
<td>9.4</td>
<td>9.8</td>
<td>9.6</td>
<td>8.2</td>
<td>7.4</td>
<td>8.2</td>
<td>8.6</td>
<td>8.8</td>
<td>10</td>
<td>9.4</td>
<td>6.4</td>
<td>5.4</td>
</tr>
<tr>
<td>Norm/Self compatibility</td>
<td>7.3</td>
<td>6.6</td>
<td>9.6</td>
<td>8.8</td>
<td>7.2</td>
<td>9.4</td>
<td>7.4</td>
<td>7.5</td>
<td>7.4</td>
<td>8</td>
<td>9.1</td>
<td>7.3</td>
<td>8.2</td>
<td>6.2</td>
<td>5.7</td>
</tr>
<tr>
<td>motivation mean</td>
<td>8.03</td>
<td>7.87</td>
<td>9.37</td>
<td>8.8</td>
<td>8.37</td>
<td>9.3</td>
<td>7.8</td>
<td>7.33</td>
<td>7.73</td>
<td>8.17</td>
<td>8.6</td>
<td>8.5</td>
<td>8.5</td>
<td>6.73</td>
<td>5.73</td>
</tr>
</tbody>
</table>

Correlation Tests

After collecting NNSs scores on pronunciation and motivation, I looked for correlations between the two. There was no correlation between scores on the four tasks and overall motivation scores. Neither was there a correlation between the four tasks and the five individual fields, as shown in Table 6. This meant that neither motivation as a whole nor any individual field of motivation was associated with late second language (L2) learners’ performance on any of the speech samples.
Table 6. Pearson correlation between NNSs score on speech samples and their motivation (r value given; no r values were significant at the p<.05 level).

<table>
<thead>
<tr>
<th></th>
<th>Motivation mean</th>
<th>Novelty</th>
<th>Pleasantness</th>
<th>Coping potential</th>
<th>Goal/need significance</th>
<th>Norm/self compatibility</th>
</tr>
</thead>
<tbody>
<tr>
<td>Picture</td>
<td>-0.05</td>
<td>-0.25</td>
<td>-0.06</td>
<td>-0.2</td>
<td>-0.08</td>
<td>0.08</td>
</tr>
<tr>
<td>Paragraph</td>
<td>0</td>
<td>-0.27</td>
<td>-0.1</td>
<td>0.03</td>
<td>0.02</td>
<td>0.09</td>
</tr>
<tr>
<td>Sentence</td>
<td>0</td>
<td>-0.27</td>
<td>-0.03</td>
<td>-0.07</td>
<td>-0.01</td>
<td>0.1</td>
</tr>
<tr>
<td>Word</td>
<td>-0.33</td>
<td>-0.42</td>
<td>-0.2</td>
<td>-0.26</td>
<td>-0.21</td>
<td>-0.4</td>
</tr>
</tbody>
</table>

Scatterplots examined the covariance of NNSs overall motivation mean scores (motmean) and their pronunciation scores on picture, paragraph, sentence and word reading tasks, as shown in Figures 3-6 (15 dots = 15 NNSs). If there is a significant correlation, we should be able to detect a linear relationship or trend among the data points. The graphs visually illustrate that there was no correlation between any of the speaking tasks and the motivation mean score among the NNSs.

![Figure 3: Scatterplot of motivation and picture telling accent score.](image-url)
Figure 4: Scatterplot of motivation and paragraph accent score.

Figure 5: Scatterplot of motivation and sentence production accent score.
I also examined whether or not age, length of residence (LOR), or age of arrival (AOA) had any significant correlation with the 15 NNSs performance. The results, as shown in Table 7, revealed that there was no correlation. This suggests that neither age, LOR, nor AOA affected late L2 learners’ performance on any of the speech samples.

Table 7. Pearson correlation between NNSs score on four speech samples and their age, LOR, AOA (r value given; no r values were significant at the p<.05 level).

<table>
<thead>
<tr>
<th></th>
<th>Picture</th>
<th>Paragraph</th>
<th>Sentence</th>
<th>Word</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>-0.11</td>
<td>-0.13</td>
<td>-0.23</td>
<td>-0.25</td>
</tr>
<tr>
<td>LOR</td>
<td>-0.16</td>
<td>-0.04</td>
<td>-0.22</td>
<td>-0.42</td>
</tr>
<tr>
<td>AOA</td>
<td>0.17</td>
<td>-0.17</td>
<td>-0.03</td>
<td>0.42</td>
</tr>
</tbody>
</table>

When I examined the AOA of the three who passed as NSs on the picture descriptions, their AOAs were as follows: participant 105, 14; participant 110, 25; participant 115, 17. This is
counter-evidence to Long’s conclusion (1990) that after age 12, the possibility of a NNS passing as a NS with regard to pronunciation does not exist anymore.

On paragraph and sentence reading only participant 115 achieved within the NS range. On word accuracy, the five NNSs AOAs are as follows: participant 102, 19; participant 110, 25; participant 114, 23; participant 115, 17; participant 116, 24. We can see that none of the NNSs who achieved within the NS range on word accuracy reported an AOA earlier than the age of 12.
DISCUSSION

Research Question Analysis

Hypothesis 1: We can find late second language (L2) learners whose ages of arrival (AOAs) are after the age of 12 and whose first languages (L1s) are outside the Indo-European language family, and who can pass as native speakers (NSs) of English with regard to pronunciation.

My first hypothesis was confirmed by the results of the study. Three of the non-native speakers (NNSs) passed as NSs on the picture description task, one of the NNSs on the paragraph and sentence reading task and five of the NNSs on word accuracy. Participant 115 was consistently identified as a NS on all four speech samples. In participant 115 we have found a late starter who seems able to ‘beat’ an age restriction on pronunciation accuracy. This is so even though 115’s language background is Mandarin.

In sum, the data from this study would suggest that we can find late L2 learners whose AOAs are after age 12, whose L1s are outside the Indo-European language family, and who can pass as NSs of English on pronunciation. However, even among those who are informally judged to sound like NS (by me and others who recommended these speakers to me), only a few can pass as NS in a detailed and rigorous test of their pronunciation, and only one was able to do so consistently on all four measures.

Hypothesis 2: There is a correlation between motivation and successful late L2 learners’ pronunciation.

My second hypothesis was not supported by the results of the study. The results revealed that none of the fields of motivation correlated with NNSs performance on any of the speech samples to a significant level, which meant that motivation did not associate with late L2 learners’
pronunciation. Nevertheless, failure to find correlation between them may be due to the fact that most of these NNSs have high motivation, and thus I did not find a wide variation in their motivational levels. In order to examine this hypothesis of a correlation between motivation and pronunciation accuracy more completely, it would be good to collect data from 15 more ‘average’ NNSs. This might give me the variation that would be necessary to find a significant correlation. This remains for future research.

In order to shine more light on the situations of my extraordinary learners I will present some detailed information about all of my learners.

Profiles of 15 NNSs

Although motivation may not correlate with NNSs scores on pronunciation, some valuable information can possibly be gleaned from a consideration of individual learners. Table 8 is a chart that ranks participants according to their scores on motivation as a whole and on each individual field. For example, on the novelty score, participants 105, 108 and 113 got the highest scores and were ranked number 1. Participant got the second highest score and was ranked number 4. First, I will consider those individuals who passed for NS on some or all of the tasks. It is worthy to note that four of the six participants who passed as NSs on all or some part of the tests (participants 110, 114, 115 and 116) are not students anymore and work in an environment with mostly or only NSs.
Table 8. Ranking of 15 NNSs on motivation scores

<table>
<thead>
<tr>
<th>Rank</th>
<th>Novelty</th>
<th>Rank</th>
<th>Pleasantness</th>
<th>Rank</th>
<th>Coping Potential</th>
<th>Rank</th>
<th>Goal/Need Significance</th>
<th>Rank</th>
<th>Norm/Self Compatibility</th>
<th>Rank</th>
<th>Motivation Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>105</td>
<td>1</td>
<td>104</td>
<td>1</td>
<td>115</td>
<td>1</td>
<td>104</td>
<td>1</td>
<td>104</td>
<td>1</td>
<td>104</td>
</tr>
<tr>
<td>2</td>
<td>108</td>
<td>1</td>
<td>108</td>
<td>2</td>
<td>114</td>
<td>2</td>
<td>108</td>
<td>2</td>
<td>108</td>
<td>2</td>
<td>108</td>
</tr>
<tr>
<td>3</td>
<td>113</td>
<td>3</td>
<td>114</td>
<td>3</td>
<td>103</td>
<td>3</td>
<td>107</td>
<td>3</td>
<td>113</td>
<td>3</td>
<td>105</td>
</tr>
<tr>
<td>4</td>
<td>103</td>
<td>4</td>
<td>103</td>
<td>3</td>
<td>105</td>
<td>4</td>
<td>108</td>
<td>4</td>
<td>105</td>
<td>4</td>
<td>113</td>
</tr>
<tr>
<td>5</td>
<td>107</td>
<td>4</td>
<td>107</td>
<td>5</td>
<td>107</td>
<td>5</td>
<td>115</td>
<td>5</td>
<td>115</td>
<td>5</td>
<td>115</td>
</tr>
<tr>
<td>6</td>
<td>112</td>
<td>6</td>
<td>102</td>
<td>6</td>
<td>104</td>
<td>7</td>
<td>102</td>
<td>7</td>
<td>110</td>
<td>7</td>
<td>107</td>
</tr>
<tr>
<td>7</td>
<td>104</td>
<td>8</td>
<td>115</td>
<td>9</td>
<td>113</td>
<td>8</td>
<td>113</td>
<td>8</td>
<td>109</td>
<td>8</td>
<td>112</td>
</tr>
<tr>
<td>8</td>
<td>114</td>
<td>9</td>
<td>109</td>
<td>9</td>
<td>113</td>
<td>9</td>
<td>112</td>
<td>9</td>
<td>111</td>
<td>9</td>
<td>102</td>
</tr>
<tr>
<td>9</td>
<td>102</td>
<td>10</td>
<td>110</td>
<td>10</td>
<td>102</td>
<td>10</td>
<td>109</td>
<td>10</td>
<td>102</td>
<td>10</td>
<td>103</td>
</tr>
<tr>
<td>10</td>
<td>109</td>
<td>11</td>
<td>111</td>
<td>11</td>
<td>111</td>
<td>11</td>
<td>114</td>
<td>11</td>
<td>114</td>
<td>11</td>
<td>109</td>
</tr>
<tr>
<td>11</td>
<td>115</td>
<td>12</td>
<td>113</td>
<td>12</td>
<td>109</td>
<td>12</td>
<td>103</td>
<td>12</td>
<td>107</td>
<td>12</td>
<td>111</td>
</tr>
<tr>
<td>12</td>
<td>110</td>
<td>13</td>
<td>112</td>
<td>13</td>
<td>116</td>
<td>13</td>
<td>110</td>
<td>13</td>
<td>103</td>
<td>13</td>
<td>110</td>
</tr>
<tr>
<td>13</td>
<td>111</td>
<td>14</td>
<td>116</td>
<td>14</td>
<td>110</td>
<td>14</td>
<td>116</td>
<td>14</td>
<td>116</td>
<td>14</td>
<td>116</td>
</tr>
<tr>
<td>14</td>
<td>116</td>
<td>15</td>
<td>117</td>
<td>15</td>
<td>117</td>
<td>15</td>
<td>117</td>
<td>15</td>
<td>117</td>
<td>15</td>
<td>117</td>
</tr>
</tbody>
</table>

Participant 115 is 25 and a NS of Mandarin. Her age of first-time learning English (AOL) was age 13 when she went into junior high school in Taiwan. At that time, English was taught five hours a week, focusing on grammar and reading, with a NNS as the teacher. Four years later, at age 17, she and her family immigrated to the US. That was her AOA of English. This participant passed as a NS on all four of my production tasks. In addition, she was tested at her home. This revealed that the difference between conducting a recording at home or in the lab was not that large of a disadvantage for a very good NNS. She definitely had high levels of motivation, as shown in Table 8. Her score on coping potential was also the highest of all the NNSs. Because I regarded her pronunciation as the best among the NNSs sounds that I had recorded at that time, I asked her during the post-test interview how she identified herself. Obler claimed that to sound like a NS, “not only must one be willing to sound like someone from another culture, but one must
be willing to give up the protection that being foreign confers, since NSs may make allowances for grammatical errors when the speaker is obviously not a NS, and thus the person is protected from sounding foolish.” (1989, p. 152) Therefore, I asked participant 115 if she minded being considered an L2 learner and she said “yes.” She did not want NSs to tolerate her grammar mistakes or mispronunciations more just because she was an L2 learner. Her answer seemed to be line with Obler’s claim. I also asked her if she had the characteristics of the Geschwind cluster such as twinning, allergy or left-handedness like Ioup et al.’s (1994) Julie. She denied. She had no twinning history in her family, no allergies except for specific medicines, and she was right-handed. She spent most of her time with NSs and had no Mandarin friends. The only time that she would speak Mandarin was with her parents, whom she did not see often. Although she was born in Taiwan, and she recognized it as her origin, she did not volunteer this information to NSs. This also reflected her vague national identity.

Participant 110 is 30 and a NS of Mandarin. Her AOL was nine when she was in elementary school. She only studied grammar and vocabulary for one or two hours a week. She came to the U.S. at age 25 to get an master of arts (MA) degree. She passed as a NS on the picture description and word accuracy. On the motivation questionnaire she showed an interest in being with NSs. On question 13 (“I prefer to stay with the groups of people who speak the same native language as I do.”) she rated 2 (strongly disagree). She also expressed in the pre-screening interview that it was very important not to stick to the L1 group if one wanted to pronounce the L2 well. However, she felt little need for pronunciation accuracy compared to the other four NNSs. She gave a rating of 6 (almost no opinion) on question 10 (“Pronouncing like a NS of English is important for me,”) and 7 (mildly agree) on question 12 (“Better pronunciation helps me in my career plan or study.”)
Participant 105 is 22 and also a NS of Mandarin. She arrived in the U.S. at age 14 with her family. Her AOL was age 10 when she was in elementary school. The class focused on grammar, vocabulary, and reading. It was taught only one hour each time, five times a week by a Mandarin teacher. Participant 105 passed as a NS on the picture description. In the pre-screening interview she showed a strong interest in English, which can be seen in her independent-reading of Shakespeare and classic novels like Jane Eyre. In addition, she had strong intentions of assimilating into American society, which showed in her ratings on questions on norm/self compatibility. For example, she rated 10 (strongly agree) on question 8 (“Better pronunciation helps me to get along with Americans better.”), question 23 (“Better pronunciation helps me to incorporate into American society better.”), and 9 (strongly agree) on question 27 (“I have American friends whom I talk with on a daily basis.”)

Participant 102 is 23 and a NS of Mandarin. She arrived in the U.S. at age 19 to study. Her AOL was age seven. The English class she took was taught one hour each time, five times a week. The class focused on grammar, reading comprehension and vocabulary. She passed as a NS on word accuracy. She can speak six different languages (Cantonese, Mandarin, Malaysian, Hakka, Hokkien, and English). She self-reported in the post-test interview that she had every intention to speak the way a particular language was spoken. However, she received a lower rating in the area of norm/self compatibility. She rated 5 on question 8 (“Better pronunciation helps me to get along with Americans better.”), question 13 (“I prefer to stay with the groups of people who speak the same native language as I do.”), question 23 (“Better pronunciation helps me to incorporate into American society better.”) and 6 on question 24 (“I spend most of the time with Americans.”) Her explanation was that she accepted herself as a foreigner and had no preference for NSs of English. She spent 60 percent of the time speaking English and the rest of her time speaking other
Participant 114 is 30 and a NS of Japanese. She came to the U.S. at age 23 for her Master’s degree. Her AOL was age 10 when she was in elementary school, but she only studied conversation one hour a week with a Japanese teacher. She passed as a NS on word accuracy. She scored the highest on pleasantness and goal/need significance; coping potential was next to the highest. This means that her feelings about English or having great pronunciation were positive, that she had goals to improve her pronunciation, and that her ability to deal with events related to pronunciation was quite good. However, she received a lower rating in the area of norm/self compatibility. She rated 5 on question 13 (“I prefer to stay with the groups of people who speak the same native language as I do.”) Her explanation was that she had no preference for NSs and that it depended on the situation. The most important explanation came from the pre-test interview. She admitted that she did not think she could lose her foreign accent and that she did not work on her pronunciation, at least not after she came to the U.S.. The psychological factor reflected in her speech coincides with Tremblay and Gardner’s claim (1995, p. 507) that “self-confidence in the language learning context is usually assessed with measures of perceived proficiency at the time of testing.”

Participant 116 is 30 and a NS of Mandarin. He came to the U.S. at age 24 to pursuing his Master’s degree. His AOL was 10 when he was in elementary school, but he only studied for five hours a week, focusing on grammar and vocabulary with a Mandarin teacher. He passed as a NS
on word accuracy. In addition, he was the only one among 15 NNSs who received a mean rating on word accuracy as perfect as NSs. He scored almost the lowest on each of the fields of motivation. In the pre-test interview, he reported that he did not work on his pronunciation, but interestingly, some of his American colleagues thought he was born in the U.S. until they knew that he came from China. He also had the lowest score on norm/self compatibility. Although he rated 8 (agree) on question 24 (“I spend most of the time with Americans.”), he rated 4 (mildly disagree) on question 27 (“I have American friends whom I talk with on a daily basis.”). However, he did not make an effort to assimilate himself into American society. One may wonder what made him succeed in word accuracy since he seemed to have low motivation in pronunciation accuracy. The most important information came from the post-test interview. He mentioned that he regarded himself as an innate good language learner, which means that he had an aptitude for acquiring pronunciation. This, which will be discussed in implications, may be the reason for his partial success on pronunciation.

I will discuss the rest of the NNSs in the order that they were tested. Participant 103 is age 23 and a NS of Mandarin. His AOL was age 6, but the class focused on only grammar and vocabulary and was only one hour a week. Consequently, he received little spoken input in English. His AOA to the U.S. was age 20. He is a student and works part-time on campus. Although he received high scores on novelty, pleasantness and coping potential, he gave low scores on questions about norm/self compatibility. For example, he rated 6 on question 13 (“I prefer to stay with the groups of people who speak the same native language as I do.”) and question 24 (“I spend most of the time with Americans.”) and rated 4 (disagree) on question 27 (“I have American friends whom I talk with on a daily basis.”). This revealed that he spoke English when he worked but he did not make effort to assimilate himself into American society.
Participant 104 is age 55 and a NS of Korean. Her AOL was age 18 and AOA 20. She received almost the highest rating on all of the motivation fields except coping potential and novelty. She gave extreme scores (1 or 10) on 25 out of the 30 questions, which may explain why she scored the highest among all NNSs on three motivation fields. Undoubtedly, she has high motivation, but she did not score within the NS range in her pronunciation. Her case revealed that it is not sufficient to have only high motivation to achieve native-like pronunciation.

Participant, 107 is age 32 and a NS of Mandarin. Her AOL was age 6 but the class was irregularly taught and focused on vocabulary and short conversation. Her AOA was age 20. She scored high on four motivation fields but received lower ratings on norm/self compatibility. In answering questions based on norm/self compatibility, she rated 5 on question 13 (“I prefer to stay with the groups of people who speak the same native language as I do.”) and 24 (“I spend most of the time with Americans.”) and rated 4 (disagree) on question 27 (“I have American friends whom I talk with on a daily basis.”) This revealed that she did not make effort to assimilate herself into American society.

Participant 108 is age 21 and a NS of Mandarin. Her AOL was age 12 and AOA age 19. She scored the highest on novelty and pleasantness and her mean score on motivation as a whole was second to the highest. She definitely has high motivation. However, she did not pass as a NS even though her pronunciation score ranked number six. What is worthy of consideration is that her length of residence (LOR) was 2 years, the least among all NNSs. Piske et al. (2001, p. 199) also suggested that “additional experience in the L2 may well lead to less foreign-accented L2 speech.” Although there are several studies that found LOR not significant in accounting for L2 learners’ pronunciation, they did not consider the factor of motivation. It is possible that with high motivation and longer LOR, L2 learners can decrease their foreign accent, although only for the
first few years. Piske et al. (2001) also suggested that LOR is not a significant factor in decreasing foreign accent if L2 learners have been exposed to the L2 environment for a long time.

Participant 109 is age 35 and a NS of Mandarin. Her AOL was age 14 and AOA age 12. She did not score high on motivation as a whole or in each individual field. In answering questions based on pleasantness, she rated 5 on question 21 ("I like to imitate how Americans pronounce English on TV, radio or movies.") In answering questions based on norm/self compatibility, she rated a 5 on question 24 ("I spend most of the time with Americans.")

Participant 111 is age 24 and a NS of Mandarin. Her AOL was age 10 but it was a conversation class that focused on vocabulary. The class was less than two hours each week. Her AOA was age 18. She did not score high on motivation as a whole or on each individual field. In answering questions based on goal/need significance she rated 6 on question 10 ("Pronouncing like a NS of English is important for me."). Her explanation is that she thinks to understand others and make herself understood are more important.

Participant 112 is age 22 and a NS of Japanese. Her AOL is age 13 and AOA age 18. She did not score high on motivation as a whole or on each individual field either. In answering questions based on goal/need significance she rated 5 on question 22 ("Better pronunciation has nothing to do with my work or study.") Her explanation is that her major is accounting so it does not matter if she has better pronunciation.

Participant 113 is age 23 and a NS of Mandarin. His AOL was age nine but the class focused on vocabulary and grammar for only one hour per week. He received low scores on pleasantness and coping potential, although high scores on novelty, norm/self compatibility and motivation mean. His AOA was age 12 when he came to Canada, but he did not really assimilate into Canadian society. In the pre-screening interview he said he spent most of the time with
Mandarin friends when he was in Canada. The situation did not change until he came to University of North Texas (UNT) to get a Bachelor of Science (BS) degree. Therefore, the real age that he felt comfortable to communicate in English was at least 18. His pronunciation score was not high, in fact, it was among the bottom of the group.

Participant 117 is age 30 and a NS of Korean. Both her AOL and AOA were age 13. She scored almost the lowest on motivation as a whole and in each individual field. In the post-test interview she said that she did not like to watch TV programs in English. “That makes me headache,” she jokingly said. She also mentioned that she preferred Korean to English. In answering questions based on norm/self compatibility she rated a 9 (strongly agree) on question 13 (“I prefer to stay with the groups of people who speak the same native languages as I do.”) This rating is indeed in line with her answer in the post-test interview.

Motivation Questionnaire

The appropriateness of the questionnaire for the study of motivation is worthy of examination. The reason that I designed questions myself was because I could not find motivational questionnaires that were suitable for some of the advanced learners in this study who had ceased formal instruction. But the post-test interview revealed that there are some problematic questions regarding some fields of motivation that do not cover successful late L2 learners. In answering questions based on one field of motivation, coping potential, four out of six participants (participant 102, 105, 110 and 114) rated either ‘almost no opinion’ or ‘strongly disagree’ on question 15 (“I seek chances to speak English.”) They said they had plenty of chances to speak English, so they did not need to seek chances. On question 18 (“When Americans tease me for my pronunciation, I feel a little embarrassed but it is ok”) two out of six participants rated either ‘almost no opinion’ or ‘strongly disagree.’ Participant 110 said she did
not feel embarrassed at all while participant 116 said he had never been teased so far. On question 20 (“If Americans tease me for my pronunciation, it won’t discourage me from learning English; on the contrary, I will do my best to improve in order to change their opinions”) participant 105 rated ‘almost no opinion.’ She said she did not agree with the latter part of the statement because she did not live for others. In answering questions based on the last field of motivation, norm/self compatibility, participants 110, 114, 115 and 116 all rated ‘almost no opinion’ on question 25 (“I like to make friends with Americans”) because they all said they liked to make friends, not necessarily with Americans. On question 28 (“Americans respect me more if I have better pronunciation”) participants 102, 110, 114, 115 and 116 rated ‘almost no opinion’ or ‘disagree.’ They said pronunciation may play a role in the initial stage, but characteristics or behaviors were the reason they received respect from Americans.

For future research, I would consider to discard question 28 and revise the other questions as follows to make them more reflective of actual thinking of successful NNSs. These new questions will be more accurate and inspire either high or low ratings.

Question 15: I have many chances to speak English.

Question 18: When Americans tease me for my pronunciation, I either feel a little embarrassed or don’t feel embarrassed at all.

Question 20: If Americans tease me for my pronunciation, it won’t discourage me from speaking English.

Question 25: I would like to get to know Americans better.
CONCLUSIONS AND IMPLICATIONS

Conclusions

The results of the study indicate three conclusions. First, we can find late second language (L2) learners whose first languages (L1s) are outside of the Indo-European family to pass as native speakers (NSs) of English on pronunciation. This study is unique because three of the 15 non-native speakers (NNSs), whose L1s do not belong to the Indo-European family, passed as NSs on the picture description task. No previous study (Bongaerts et al., 1997; Bongaerts et al. 2000; Flege et al., 1995) has found NNSs who pass for NSs on such an unconstrained speaking task (although Ioup et al.’s, 1994, Julie, comes close).

Second, motivation may have an effect on the levels of pronunciation accuracy which NNSs can achieve, but some NNSs who also have high levels of motivation may not be able to achieve as high a level of pronunciation accuracy. Therefore, we cannot say that motivation is sufficient enough to allow L2 learners to pass as NSs, although several tendencies have been noted among all those who were able to pass as NSs which may indicate that motivation did have a role in L2 pronunciation.

Third, to investigate successful L2 learner’s motivation, we do need a new motivation questionnaire, especially for covering those who are at the workplace or have ceased formal education. A different motivation questionnaire for testing L2 learners’ motivation on pronunciation from the past motivation studies or L2 studies was created for this experiment.

Limitation of the Study

Based on the small sample of participants, the conclusions are limited to the current study. To substantiate the conclusion that late L2 learners whose L1s are outside of the Indo-European family can pass as NSs of English on pronunciation we need to find more successful L2 learners.
We should also use more judges, both experienced and inexperienced, to rate their accents. In addition, the motivation questionnaire is a new one. Therefore, the results of the findings that this questionnaire revealed are also limited to the current study.

Ideas for Future Research

There are four ideas for future research that I gleaned from this study. First, the profiles of the NNSs in this study showed that many of these successful late learners were in the workplace as opposed to college. This may be an important factor for other researchers if they want to investigate successful late L2 learners.

Second, the successful case of participant 115, who passed as a NS on all speech samples, was an encouraging example. Nevertheless, her self-report in the post-test interview that the frequency of her L1 use was substantially less than her L2 use aroused my attention. This phenomenon is in line with Piske et al.’s findings (2001). They reported that their native Italian learners of English “who continue to speak their L1 frequently have significantly stronger foreign accents in English than do individuals who speak their L1 infrequently” (Piske et al., 2001, p. 209). To achieve native-like proficiency on pronunciation does one have to make a choice between his/her L1 and L2, or third language (L3)? This question is left for future research.

Third, the case of participant 116 also caught my attention. He claimed to have the aptitude to pronounce language well. In the pre-test interview he said that he could not recall working on his pronunciation. In the post-test interview he said that even before he was exposed to the L2 environment, when he was still working in his native country, he already exhibited similar pronunciation to the one he has now. Therefore, it would be an interesting study if we could investigate more about each learner’s aptitude in pronunciation and compare it with their pronunciation score.
Fourth, we should consider the inadequacy of the past motivation questionnaires for participants of the current study. L2 learners who have ceased formal target language instruction may need different motivation questionnaires from those who are still taking language classes. In addition, English-as-a-second-language (ESL) and English-as-a-foreign-language (EFL) learners may need different motivation questionnaires, especially because of the target-language or non-target language society they are in. A new motivation questionnaire is created in this study. Future researchers may consider adopting it to investigate their advanced L2 learners’ motivation.

Implications

There is a pedagogical implication from this study. Although pronunciation instruction seems to have little influence on successful late L2 learners’ pronunciation, and this study found no significant correlation between motivation and pronunciation scores, motivation is always mentioned in successful late L2 learners’ self-reports to account for part of their success in pronunciation (Bongaerts et al., 1997; Bongaerts et al., 2000; Moyer, 1999). Therefore, ESL and EFL teachers can play a role in inspiring their students’ motivation, although only to a certain level. A teacher can increase students’:

1. **curiosity** in speaking the target language (novelty), for example, by recording their own speech
2. **interest** in speaking it (pleasantness), for example, by holding a speech contest
3. **need** for speaking it (goal/need significance), for example, by relating the target language to students’ future goals
4. **ability** to deal with different situations in the target language (coping potential), for example, by using role play activities
5. **chances** to contact NSs of the target language (norm/self compatibility), for example, by interviewing a NS of the target language

If a teacher can do these things, students will be motivated to speak the target language. Although motivation may not be sufficient enough to improve NNSs pronunciation to the point
where it is indistinguishable for NS, at a basic level such as the willingness to speak, motivation is indispensable. Only when students are motivated to speak, is there a possibility for them to improve their pronunciation. Therefore, ESL and EFL teachers should dedicate to how to utilize the teaching materials, the environment that students live in, reflections of the teachers’ self-learning, group activities, and finally students’ own personalities to inspire students’ motivation.
APPENDIX A

PICTURE DESCRIPTION TASK
When my mother-in-law was widowed and came to live with us, I added her to my list of burdens. Determined to show her what a great wife and mother I was, I angrily scrubbed pots while this sad woman who wanted to help sat, doing nothing.

In time I saw that her willing hands could free me from the bottomless ironing basket. When she took the children to the park, I could read a book or make an uninterrupted phone call.

Later Grandma’s help enabled me to become involved in volunteer and political work. Doing something about the state of a world I’d long been complaining about gave me a more positive outlook.
1. The boy took the cap off before praying.
2. We caught a cab at Seventh and Vine.
3. Water had started to leak into the cellar.
4. Jeff was chosen for the league all-star team.
5. This teacher has the least experience of any.
6. The little boy gave Santa his list of toys.
7. Playing hide and seek in that park is fun.
8. The sick baby wouldn’t stop crying.
9. When you feel unwell it’s best to rest.
10. Without this plan, we fear a budget deficit.
11. He took a sip of water from the well and gagged.
12. The little girl wants to wear the prettiest dress.
APPENDIX D

WORD READING TASK
1. cap/ cab
2. beat/ bead
3. leak/ league
4. loop/ lube
5. bet/ bed
6. tack/ tag
7. sleep/ slip
8. least/ list
9. seek/ sick
10. team/ Tim
11. sheep/ ship
12. Streep/ strip
13. wall/ war
14. fill/ fear
15. well/ wear
16. mole/ more
17. Mel/ mare
18. bowl/ bore
APPENDIX E

QUESTIONNAIRE

(7 QUESTIONS ADAPTED FROM GARDNER’S AMTB)
Questionnaire

Part I. General information

ID number _______ Age _______ Gender: Male Female

Native language ________________ Degree earned or working on: BA BS MA PhD

Other languages spoken ______________________________ Length of residence in the US _____

Age of first arrival in the US ______ Age of first acquisition of English ____________

1. If you started to learn English before age 12, how old were you then? How many hours a week did you spend learning it? What things did you study (grammar, vocabulary, etc)?

2. Are there any specific sounds or patterns of intonation that have been difficult for you? (For example: question intonation, stress, etc.)

Part II. Motivation Questionnaire

1. I want to learn more about English because it is a medium for me to get to know the world.

2. I don’t think how NSs of English pronounce has anything to do with my pronunciation.

3. I am not afraid of having an American identity because I either like to be an American.

4. English is not important for me to learn about because it is not necessary in the world.

5. I feel happy if people tell me that I have great pronunciation.

6. The majority of my experience of learning English has been pleasant.
7. I like to watch TV, listen to the radio, listen to the songs or watch movies in English.

8. Better pronunciation helps me to get along with Americans better.

9. I want to pronounce as best as I can.

10. Pronouncing like a NS of English is important for me.

11. I seldom watch or listen to any kind of English programs.

12. Better pronunciation helps me in my career plan or study.

13. I prefer to stay with the groups of people who speak the same native language as I do.

14. I spend most of the time speaking English at work or school.

15. I seek chances to speak English.

16. I am worried if I don’t have foreign accent, I will lose my ethnic native language identity.

17. I pay careful attention to how Americans pronounce words.

18. When Americans tease me for my pronunciation, I feel a little embarrassed but it’s ok.
19. I avoid speaking English if Americans tease me for my pronunciation.

20. If Americans tease me for my pronunciation, it won’t discourage me from learning English; on the contrary, I will do my best to improve in order to change their opinions.

21. I like to imitate how Americans pronounce English on TV, radio or movies.

22. Better pronunciation has nothing to do with my work or study.

23. Better pronunciation helps me to incorporate into American society better.

24. I spend most of the time with Americans.

25. I like to make friends with Americans.

26. I don’t like to speak English.

27. I have American friends whom I talk with on a daily basis.

28. Americans respect me more if I have better pronunciation.

29. I don’t care about my pronunciation.

30. I like to speak English.
REFERENCES


