All Theses and Dissertations

2012-03-15

Elicited Imitation Testing as a Measure of Oral Language Proficiency at the Missionary Training Center

Sara E. Moulton
Brigham Young University - Provo

Follow this and additional works at: https://scholarsarchive.byu.edu/etd

Part of the Educational Psychology Commons

Elicited Imitation Testing as a Measure of Oral Language Proficiency

at the Missionary Training Center

Sara Moulton

Dr. Randall Davies, Chair
Dr. Peter J. Rich
Dr. Charles Ray Graham

Department of Instructional Psychology and Technology

Brigham Young University

April 2012

ABSTRACT

Elicited Imitation Testing as a Measure of Oral Language Proficiency

at the Missionary Training Center

Sara Moulton
Department of Instructional Psychology and Technology

This research study aimed to create an alternative method of measuring the language proficiency of English as a Second Language (ESL) missionaries at the Missionary Training Center (MTC). Elicited imitation (EI) testing was used as this measure of language proficiency and an instrument was designed and tested with 30 ESL missionaries at the MTC. Results from the EI test were compared with an existing Language Speaking Assessment (LSA) currently in use at the MTC. EI tests were rated by human raters and also by a computer utilizing automatic speech recognition technology. Scores were compared across instruments and across scoring types. The EI test correlated highly with the LSA using both scoring methods providing initial validity for future testing and use of the instrument in measuring language proficiency at the MTC.

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Introduction

Missionaries for the Church of Jesus Christ of Latter-day Saints are often asked to learn a new language in which they will proselyte. They typically come to study at the Missionary Training Center (MTC) for a period of time ranging from 3 weeks to 12 weeks. During such time, language training is a crucial element of their learning.

Missionaries' levels of oral language proficiency is an important measure of their preparation to enter the mission field. The current primary method of measuring oral language proficiency at the MTC is by use of the Language Speaking Assessment (LSA)—a computerized audio-response assessment that elicits spontaneous speech from the missionaries. The LSA is designed to accommodate the high number of missionaries needing direct feedback on their oral language skills. Trained raters listen to missionaries' audio recorded responses to computer prompts specifically designed to elicit important language skills in a missionary context. These raters then score and give on-line feedback to individual missionaries to help the missionaries improve their language proficiency.

Statement of the Problem

This rating process, however, presents some problems for the missionaries and for the MTC as an institution. Ratings are subject to rater errors and variability in scoring. Also, the rating process is costly and time intensive. Although the LSA can potentially be used as an institutional measure of how well missionaries as a whole learn languages at the MTC, this method of measuring missionaries' language proficiency is primarily used only to give feedback to individual missionaries.

Another problem encountered at the MTC related to language testing is that the system that currently exists to place English as a Second Language (ESL) missionaries into appropriate language levels upon entering the MTC is based on oral interviews. Missionaries enter the MTC at differing proficiency levels and some need to be advanced to a shorter stay at the MTC while others need to spend more time in training. These adjustments in training times affect changes in schedules at the MTC as well as entrance dates into the mission field. If a more efficient method of measuring language could be implemented prior to entering the MTC, many of these problems could be eliminated. An instrument that is more economical and objective that is also valid and reliable would be a welcome improvement to the language training and assessment process at the MTC.

The Elicited Imitation (EI) test, for example, is a computerized test that measures oral language proficiency by having the subject hear and repeat utterances of varying syllable length in the target language. It is economical and simple to administer and rate. Such a measure would not replace the current Language Speaking Assessment given to missionaries at the MTC. Rather, it could be given as an additional assessment to aid in providing an institutional measure of overall language proficiency.

Statement of Purpose and Research Questions

The purpose of this research was to examine, develop and validate an Elicited Imitation test that will reliably and accurately measure ESL missionaries' spoken language proficiency. This study addressed the possibility of using EI tests as an institutional measure of language proficiency at the MTC. This study was guided by the following research questions:

1. How do scores on EI tests compare with those obtained using the LSA?

2. To what degree can automatic speech recognition (ASR) technologies
   facilitate scoring the EI tests in order to eliminate rater intervention?

Literature Review

It is important to understand both the history of language testing at the MTC as well as the origin and current research on EI testing in order to illustrate the relevance of this project.

Language Testing at the MTC

In the late 1970s, administration at the MTC identified the need to implement some form of language testing in order to determine how well missionaries were learning their new proselyting languages. At that time, the Foreign Service Institute (FSI) had an oral language interview that was used to determine language proficiency. A representative from Educational Testing Services (ETS) came to the MTC and trained raters on administering and scoring the FSI test to missionaries. Over time, this process became expensive and difficult and presented challenges in adapting to the unique context of missionary work. In the early 1980s the MTC stopped using this test and even stopped formal language testing altogether (C. Ott, personal communication, February 13, 2012).

A decade later, the MTC developed an in-house measure of language proficiency and began administering it to missionaries at the MTC. It followed the tradition of a basic in-person oral interview conducted by trained interviewers but was viewed as more appropriate for missionaries as it incorporated a missionary context. This new instrument

was called the Modern Language Performance Test (MLPT) and engaged missionaries in oral language tasks of increasing difficulty. The interviewers rated the missionaries based on a scoring rubric with criteria designed to match the test. This test was used throughout the 1990's but again, the MLPT turned out to be time intensive, costly, and typically only a few missionaries were sampled to participate in this type of language testing. Additionally, the content of the MLPT became antiquated with a change in the missionaries' curriculum that involved moving away from set "discussions" that they were to teach. At the same time, new technologies in language learning and testing were also emerging. It was determined that a more efficient and practical instrument be created in order to test language proficiency at the MTC (C. Ott, personal communication, February 13, 2012).

In 2004, the most recent measure of language proficiency at the MTC was developed. It utilized a Web-based on-line assessment (LSA) that allows a missionary to record audio responses to written prompts in their target language. The LSA was designed by a team of language experts at the MTC and grew out of the criteria used to rate missionaries on the MLPT (C. Ott, personal communication, February 13, 2012). The LSA has served as the "gold standard" of language testing at the MTC, primarily because it is closely based on guidelines set up by the American Council on the Teaching of Foreign Languages (ACTFL) and the scoring rubric they use to rate Oral Proficiency Interviews (OPIs) (see http://www.actfl.org/files/public/ACTFLProficiencyGuidelines2012_FINAL.pdf).

The LSA is currently in use at the MTC and has provided valuable insight into missionaries' language abilities. This instrument has also allowed more missionaries to participate in language testing, but still has the limitation of needing human raters to listen to and rate all of the missionary recordings. In spite of these limitations, language testing has a prominent place at the MTC and measures of proficiency have changed and improved over the years.

Recently, however, the MTC administration has expressed interest in finding an additional language proficiency measure that could reach even more missionaries—both in terms of administering the test and getting informative feedback from the test. A search for an additional test with increased efficiency and accuracy in rating language has led MTC researchers to investigate elicited imitation testing.

Elicited Imitation Testing

Elicited Imitation tests have the "learners listen to and repeat, to the best of their ability, utterances of varying lengths and complexities in the language being acquired" (Graham, Lonsdale, Kennington, Johnson, & McGhee, 2008, p. 1604). The underlying assumption of EI tests is that language learners with high oral language proficiency can recall and repeat longer and more complex sentences than those who have lower oral language proficiency.

Specific uses of elicited imitation tests have varied over the years. Early elicited imitation research was conducted by Fraser, Bellugi, & Brown (1963) and involved language learning in children. Other uses of these types of tests emerged decades later and have included studies involving English as a second Language (ESL) morphology,

bilinguality, and the effects of age on language learners, to name a few (see Grigg, 1986; Verhoeven, 1994; Scott 1994). These later studies have revived an interest in how elicited imitation testing works and how widely it might be utilized.

Bley-Vroman and Chaudron (1994) pioneered this renewed interest in research regarding fundamental issues surrounding elicited imitation. In doing so, they have carefully analyzed the necessary conditions under which EI testing may be determined to be a valid measure of oral language proficiency. Their research has also focused on what conclusions can be drawn about a second language (L2) learners' oral proficiency based on their ability to imitate L2 utterances. Based on their research, Bley-Vroman and Chaudron concluded that there is sufficient evidence that "the more you know of a foreign language, the better you can imitate the sentences of the language. Thus, EI is a reasonable measure of global proficiency" (p. 247).

Even more recent research provides evidence that elicited imitation testing can tap into *implicit* knowledge of language structure and function. Erlam (2006), for example, acknowledges criticisms of an elicited imitation approach (e.g., McDade, Simpson, & Lamb, 1982, who claim that elicited imitation is merely a function of rote recall of a stimulus); yet she demonstrated "that the design of an elicited imitation test can largely determine to what extent it is either a measure of a learner's internal language system or a measure of his/her ability to imitate given stimuli verbatim" (p. 467). Erlam is pioneering ways in which to create greater meaningfulness of the elicited imitation stimulus in test design. Additionally, her research has provided greater evidence that

elicited imitation testing is actually measuring significant implicit learning that gives greater insight into one's "true" language proficiency.

Researchers at BYU have recently regained interest in utilizing elicited imitation testing. They have used cutting edge research and procedures to validate elicited imitation as a method of language testing and have probed questions regarding factors affecting how tests are administered and scored (see Graham, et. al, 2008). They are also developing elicited imitation tests in several languages to continue their research (D. Lonsdale, personal communication, December 28, 2010).

Automatic Speech Recognition (ASR) in Rating of Elicited Imitation (EI) Tests

Because of the nature of elicited imitation testing, the tests themselves have the additional benefit of utilizing alternative methods of scoring. These tests may be scored by hand, similarly to the LSA, by individual trained raters. In such cases, rater subjectivity may be reduced due to the simplified process of scoring EI tests (i.e., rating each syllable as either correct or incorrect reduces subjectivity). Additionally, the use of Automatic Speech Recognition (ASR) technologies is also a possibility in rating EI tests and would eliminate the need for human raters altogether. If deemed sufficiently reliable and valid, the use of ASR scoring methods could greatly reduce the time and cost it takes to evaluate oral language proficiency.

Adaptations must be made in terms of test content in order to use elicited imitation testing at the MTC because the environment that missionaries encounter differs significantly from a traditional academic environment. Specifically, vocabulary unique to the missionary curriculum must be used in elicited imitation testing at the MTC.

Thompson (2005) conducted a master's thesis at the MTC to find the 500 most common vocabulary words necessary for ESL missionaries to learn in order to teach effectively. Accompanying this basic word list was a dictionary of definitions and lists of those 500 words in context. Using these 500 common missionary words in an elicited imitation test at the MTC would ensure that vocabulary and context are related to what missionaries are expected to learn at the MTC and would also provide initial face validity for the instrument.

Validation of the Instrument

If an elicited imitation test is to be used at the MTC, it must reliably demonstrate that it is a valid of measure of actual oral language proficiency. One way of ensuring validity for a new instrument is by comparing it with an existing instrument that already demonstrates high construct validity (Cronbach, 1970). However, "this procedure is helpful only if the test used as criterion is accepted as meaningful and important" (Cronbach, 1970, p. 122). Data gathered from the LSA have been analyzed by in-house researchers over the past seven years and validations of that instrument have also been conducted to ensure the meaningfulness and importance of that instrument (C. Ott, personal communication, February 13, 2012). Because of this, the LSA is likely the best measure by which to judge the validity of the EI test.

Method

In order to create and validate an instrument to help fill the need describe above, the following methods will be employed.

Development of an Elicited Imitation Instrument

In this project, items for an EI test were developed with vocabulary and context appropriate for missionaries at the MTC. Specifically, items were selected using the dictionary of the 500 most common English words at the MTC. In my role as a research specialist at the MTC, I categorized each sentence containing the word into syllable groups. Dr. Ray Graham, professor of Teacher Education at Brigham Young University, suggested using items that fell into 4 categories: 6-8 syllables (level 1 difficulty), 10-12 syllables (level 2 difficulty), 16-18 syllables (level 3 difficulty), and 21-22 syllables (level 4 difficulty). A total of 184 sentences were developed for this test. These items were then randomly divided into five different equivalent versions of the test. Each test contained approximately 35 items with eight or nine items from each difficulty level. Ten more items were added to each form of the test making the total number of items on each test form approximately 45. These latter items have been previously validated by the BYU Pedagogical Software Speech Technology (PSST) group but do not contain religious vocabulary missionaries would be expected to know.

After writing the items, audio recordings of each of the sentences were made and later input into the EI test delivery module. All items were recorded by one female. Only items tested from Form 1 of the test were used in this study. See Appendix A for a list of these items and their respective syllable counts.

Participants

In order to answer the evaluation questions listed above, Form 1 of the items on the EI test and the LSA were administered to English as a Second Language (ESL)

missionaries at the MTC. A stratified random sample of the missionaries was used (according to language proficiency level upon entering the MTC). These strata were determined by oral interviews conducted by the staff train the ESL missionaries at the MTC. Missionaries fit into one of three categories: beginning, intermediate, and advanced. A total of 36 missionaries were selected for participation in this project: 14 were beginning level non-native English speakers, 10 were intermediate, and 12 were advanced. The missionaries came from a variety of backgrounds and represented 15 different native languages.

Data Collection and Procedures

Each missionary took both the EI test and the LSA. These tests were completed at varying stages throughout the missionaries' stay at the MTC. Missionaries were randomly assigned which assessment they took first and completed both tests during one testing period. A test proctor assured that each missionary who took the tests followed testing procedure as outlined in the study.

Data Analysis

In order to determine overall proficiency scores, strict procedures were followed. The LSAs were divided among three trained raters who assigned scores according to a pre-existing rubric (see Appendix B). These raters were selected due to their expertise in language proficiency ratings. They were key players in creating the rubric and in training others on how to effectively complete LSA ratings. Raters listened to various speech samples recorded by each missionary and assigned a rating on four separate subscales: pronunciation, vocabulary, grammar, and fluency. LSA scores range on a scale from one

to seven on each of the four subscales. The overall score was calculated by averaging the scores for each of these four subscales.

EI tests were hand rated by one trained rater. Due to the large amount of time required for rating EI tests by hand and the cost of training raters, it was determined by MTC administrators that hand rated tests would only be rated once. Each item on the EI test was previously divided into its syllabic constituents and displayed in web-based rating interface. While listening to each recorded item, the rater entered a score of either 1, if the missionary got the syllable correct, or 0, if the missionary got the syllable incorrect. Missionaries received either a 1 or 0 score for each syllable in each item. Overall scores on the EI test were determined two ways: first, by calculating the number of syllables correct out of the total number of syllables for each item to determine a "percent correct score" for that item; secondly, the "percent correct scores" of all of the items were averaged together for each missionary.

Elicited Imitation tests were rated again using ASR technology based on the language model used at BYU. This model utilized the Sphinx ASR engine (Lee, 1989), which originated at Carnegie Mellon University and was based on the New York Times Annotated Corpus. Overall scores were again determined by the "percent correct scores" for each of the items individually, and for each missionary in the study. The ASR scoring system utilized a pronunciation guide that included acceptable variations in pronunciation for each syllable.

The overall scores on each of these measures were then correlated to determine reliability across the various measures of oral proficiency. Each subscale of the LSA was

also correlated with the overall scores of the EI. Additionally, the BYU group's items were correlated with the new MTC items to determine if there were any differences in scores based on vocabulary. This was also done because the previously determined validity of the ten BYU group's items would help determine the validity of the new MTC items. In addition to correlated measures of oral proficiency, item means of the BYU PSST group's items were compared with item means of the new MTC items. Lastly, a matched-pairs t-test was used to compare differences in percent scores between the hand rated EI test and the ASR scored EI test.

Results

The results here will help to first, establish the validity of the EI instrument and second, compare human rating methodologies with the use of ASR scoring.

Although 36 missionaries participated in the testing, data from six missionaries were removed from the final analysis of the results. Five of the missionaries whose data was not included, experienced technology problems with the EI and recordings were either not saved to the database, or were not audible enough to grade. The other missionary, whose data in not included, stopped responding to items approximately half way through. Data from the remaining 30 missionaries are represented in this analysis.

Elicited Imitation and LSA Assessment Correlation

The first research question in this study involved the relationship between the LSA and the EI test. The correlation between the LSA and the hand-scored EI test was quite high ($r=0.83$) with the corresponding relationship between the LSA and the ASR-

scored EI only moderately strong (r=0.61). Human EI scoring and ASR EI scoring also correlated quite well (r=0.82) (see Figure 1 below).

MTC and EI Validated Item Comparison

The BYU group's EI items fared nearly equally as well as the MTC EI items. Item means for the BYU items compared to the MTC items were not significantly different for the hand-scoring or the ASR scoring (p= 0.459 and p= 0.351, respectively). Additionally, individual missionary scores for the BYU items and the MTC items were strongly correlated (r=0.91), both findings suggesting these items seem to measure an equivalent construct (see Figure 2 below).
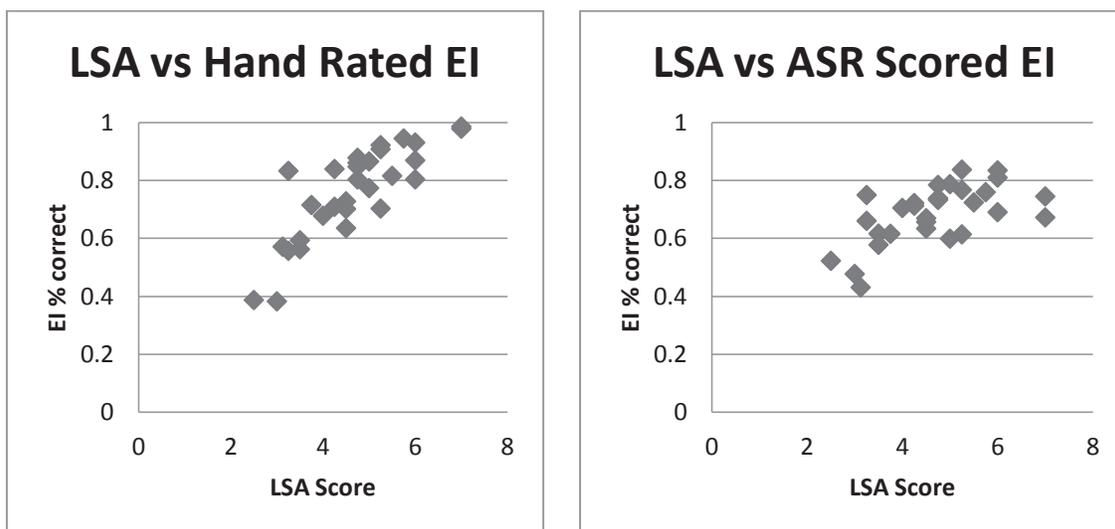
*Figure 1.* Correlation of scores between LSA and EI test.
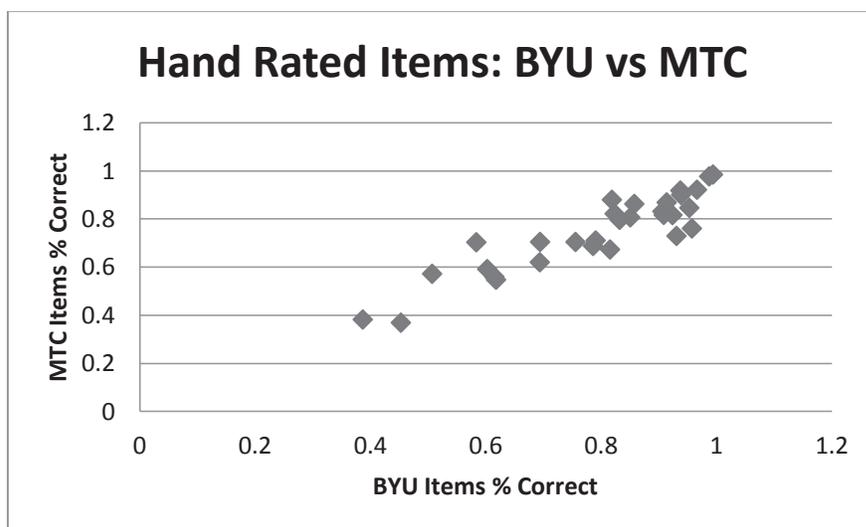
Figure 2. Correlation of scores of Hand Rated MTC items and BYU items.

Hand Rated and ASR Scoring Comparisons

The LSA scores showed a strong correlation with the hand rated BYU items and the hand rated MTC items (r=0.80 and r=0.81, respectively). Comparisons of LSA subscales to both the overall scores of human-scored EI and the overall scores of the ASR-scored EI showed an additional moderately strong relationship. The correlations for each of the sub scores of the LSA and the human-scored and ASR-scored EI tests are presented in Table 1 below.

Table 1

Correlations for LSA subscale and EI Test Results.

	Hand-scored EI	ASR-scored EI
Pronunciation	0.66	0.57
Fluency	0.78	0.67
Vocabulary	0.83	0.62
Grammar	0.74	0.51

Lastly, although the hand rated EI and ASR-scored EI correlated strongly (r=.83, see Figure 3 below), the matched pairs t-test revealed that the hand rated EI test scores and ASR-rated EI test scores differed significantly (p < .001).

Human EI v. ASR EI Scoring
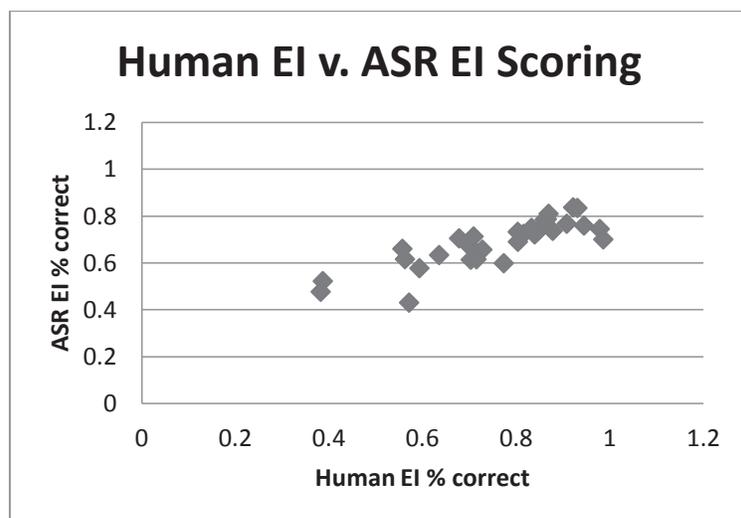
Figure 3. Correlation of scores of Hand Rated EI items and ASR scored EI items.

Discussion and Conclusions

The results of this study indicate several important points that need to be considered.

Summary and Interpretation of Results

Initial testing of the EI instrument for use at the MTC provides encouraging results. From the data, it is evident that there is a relatively strong relationship between what is being measured in the LSA and what is being measured in the EI. Previous validation of the LSA suggests that both are either directly or indirectly measuring a similar aspect of oral language proficiency.

Additionally, comparisons with the previously validated BYU EI items suggest that the newly written MTC items perform just as well, if not better, in measuring missionaries' language ability. A specific vocabulary set worked as well as a generalized vocabulary set. This finding, in particular, is encouraging for the field of language

testing in general. It demonstrates that EI testing can incorporate vocabulary from a specific or unique corpus, depending on the desired context, and still provide a valid estimate of oral language proficiency.

The ASR scoring mechanism did an adequate job in providing moderately predictable results of language proficiency. However, the matched paired t-test revealed that it is not as reliable as the rater scoring method. The scatter plot reveals that the ASR results are particularly a problem at the higher end of the LSA scores. Inquiry into why this might have been the case revealed several problems with recording mechanisms and storing audio files (e.g., audio files being clipped, poor recording devices, background noises, etc.). Additionally, errors were later discovered with the pronunciation guide used as the basis for ASR scoring.

Improving the recording devices used and refinement of the ASR may strengthen the results of that instrument as a means of rating language proficiency. Refinement of the ASR would include revising the pronunciation document of acceptable utterances for each syllable, and ensuring that the ASR can account for the pronunciation of the numerous native language (L1) backgrounds of the missionaries at the MTC. New headsets should also be purchased at the MTC in order to ensure high-quality recordings of missionary utterances on the test.

Contributions

The benefits of using EI testing at the MTC include reducing cost and increasing efficiency of language proficiency testing and scoring for missionaries. Additionally, the use of an EI test could be implemented before missionaries even arrive at the MTC

allowing missionaries to be placed in appropriate levels of training upon arrival. This could greatly reduce scheduling problems and changes in length of ESL missionaries' MTC training. The results of this study provide evidence of improving language testing at the MTC in significant ways.

On a broader scale, this project contributes to the body of literature on the use of EI testing. Because this type of testing is cost-effective and efficient, it could be used in various settings, such as junior high or high schools. Teachers in these settings typically shy away from oral language testing because of the lack of efficiency of such testing (i.e., an oral interview), which creates an inability to assess each student in a reasonable time frame. The use of an EI test in such an environment could provide a reasonable alternative for assessment, especially because current methods most likely lack any kind of speaking assessment.

The results of this study are promising and the potential uses of EI testing are far-reaching. Not only can we help solve issues with testing missionaries' language abilities at the MTC, but we can also provide insight into helping the language testing community in general. Such an idea could revolutionize our ability to measure spoken language proficiency as never before.

Limitations

Weaknesses surrounding this project include on-going debates on the nature of EI testing and whether it is enough to determine actual language proficiency. Challenges to its validity also include an argument against the initial face validity of the instrument. Although the body of research on this topic continues to grow, some critics still point to a

lack of assessment of conversational ability in an EI test (e.g., McDade, Simpson, & Lamb, 1982; Vinther, 2002). Those who would implement this at the MTC need a basic understanding of why and how EI can measure language proficiency and then see how the instrument compares against other instruments. This instrument may also need to be paired with a more traditional type of assessment (e.g., the LSA) to enhance the results of the test.

Issues with technology are other potential weaknesses of this study. Poor quality audio recordings or non-working equipment were seen to affect the scores of several participants in this study. Also, issues with pronunciation of learners from so many different L1 backgrounds seemed to affect the ASR scores.

Implications for Future Research

Although our initial testing provides some promising results, more testing needs to be done in order to strengthen the claims and results of this study. The remaining items that were developed must also be tested and an item analysis should be conducted for each of the items. If possible, the number of items administered on the EI may also be reduced. Ideally, this would allow shorter test periods for participants while still maintaining the validity of the test.

Additionally the results obtained using the ASR technology need to be stronger. Because this is a high-stakes assessment at the MTC, which could determine the amount of training a missionary receives, the correlation between an ASR-scored EI test and the LSA should ideally be around r=0.9. Improvements in technology and the language model used for the ASR could help improve that relationship.

One other possible improvement for the ASR scoring mechanism could include aligning the ASR library and pronunciation guide to the type of English being spoken by the missionary (e.g., British English vs. American English) especially if the missionary had any previous English training. Future research could also include the question of whether the ASR model should be changed based on the L1 of the speaker. The possibility of using a different language model for the ASR should be considered.

References

ACTFL (2012). ACTFL proficiency guidelines. Retrieved from

      http://www.actfl.org/files/public/ACTFLProficiencyGuidelines2012_FINAL.pdf

Bley-Vroman, R. & Chaudron, C. (1994). Elicited imitation as a measure of second-

      language competence. In E.E. Tarone, S. Gass & A.D. Cohen, Research

      methodology in second-language acquisition, Hillsdale: Lawrence Erlbaum. 245-

      261.

Cronbach, L. J. (1970). *Essentials of psychological testing* (3rd ed.). New York: Harper &

      Row, Publishers, Inc.

Erlam, R. (2006). Elicited imitation as a measure of L2 implicit knowledge: An empirical

      validation study, *Applied Linguistics, 27*(3), 464-491.

Fraser, C., Bellugi, U., & Brown, R. (1963). Control of grammar in imitation,

      comprehension, and production. *Journal of Verbal Learning and Verbal*

      Behavior, 2(2), 121-135.

Graham, C., Lonsdale, D., Kennington, C., Johnson, A., & McGhee, J. (2008). Elicited

      imitation as an oral proficiency measure with ASR scoring. Retrived

      from http://repository.dlsi.ua.es/242/1/pdf/409_paper.pdf

Grigg, T. J. (1986). The effects of task, time, and rule knowledge on grammar

      performance for three English structures. *University of Hawai'i Working Papers*

      in ESL, 5(1), 37-60.

Lee, K.F. (1989). Automatic speech recognition: The development of the SPHINX

      system. Boston: Kluwer Academic Publishers.

McDade, H., Simpson, M., & Lamb, D. (1982). The use of elicited imitation as a measure of expressive grammar: A question of validity. *Journal of Speech and Hearing Disorders,* 47, 19-24.

Scott, M. L. (1994). Auditory memory and perception in younger and older adult second language learners. *Studies in Second Language Acquisition, 16*(3), 263-281.

Thompson, C. (2005). 500 essential English words for ESL missionaries (Unpublished master's thesis). Brigham Young University, Provo, UT.

Verhoeven, L. T. (1994). Transfer in bilingual development: The linguistic interdependence hypothesis revisited. *Language Learning, 44*(3), 381-415.

Vinther, T. (2002). Elicited imitation: A brief overview. *International Journal of Applied Linguistics, 12*(1), 54-73.

Appendix A

Items used in Form 1 of MTC EI test	Syllables
Covenants bring blessings	6
Sin causes feelings of guilt	7
He wants us to become like him	8
God is the source of all truth	7
Sunday is a day for worship	8
Prophets taught His gospel	6
To be forgiven, we must accept Christ	10
Mercy is one of the attributes of God	11
You had the power to choose before you were born	12
We must avoid harmful drugs in any form	11
Through the grace of God, you can be saved from your sins	12
You can know Joseph Smith is a prophet	10
God loves us and will help us make right choices	11
We are commanded to love our neighbors	10
Answers to our prayers come in many ways	10
Agency is an eternal principle	11
We should treat our bodies with respect and reverence	12
The Lord commands us not to use wine and strong drinks	12
Will you begin reading the Book of Mormon from the beginning?	16
When we accept callings, we are sustained publicly in a Church meeting	18
Revelation is communication from God to His children	16
The eternal glory we receive will depend on our faithfulness	17
He has given us this experience on earth so we can learn and grow	18
The scriptures explain how the sacrament is to be administered	17
Heavenly Father is happy when you get on your knees and ask for help	18
As children of God, we have a special relationship with Him	16
If we want to increase our faith in Jesus Christ, we must work at it	17
Only through the Savior's grace and mercy can we become clean from sin	17
Our purpose in life is to find lasting peace, joy, and happiness	16
Pride can prevent us from having the Holy Ghost with us all the time	17
A veil would cover our memories, and we would forget our heavenly home	18
Revelation is given according to the needs and circumstances of the people	22
You will find that the key to happiness is to work for the happiness of others	21
The Holy Ghost is our Heavenly Father's messenger and is a special gift to us	22
Our purpose in this life is to have joy and to prepare to return to God's presence	21
*Have you eaten all your bread without any honey at all	15
*He cut himself while he was shaving	9
*He should have walked away before the fight started	13
*How well had he played the guitar before he began playing in the band	18
*If she listens she will understand	9
*Joe writes poetry	5
*They played games	3
*This book is so interesting I can't stop reading it	9
*When I was a teenager I would go to town every day	15
*Which of these doesn't taste like fried chicken	10

*These items are from the BYU PSST group

Appendix B

LSA Rubric:

Language Speaking Assessment Criteria

1	2	3	4	5	6	7

* NOTE: A 'NR/0' rating signifies that the speech sample is not ratable due to a non-language related cause (for example, technological problems, not enough speech, etc.).

Pronunciation *(Correctness of sounds; stress and intonation; foreign accent)*

Pronunciation is usually unintelligible	Difficult to understand at times; frequent errors in sounds, stress, or intonation; accent inhibits communication or is irritating	Fairly easy to understand; some errors in sounds, stress or intonation; accent attracts attention but does not inhibit communication	Can be understood without difficulty; accent is not pronounced

Fluency *(Rate and flow of speech)*

Speech is so slow or so fast that communication does not occur	Speech is slow enough (i.e. frequent or long pauses and fillers) or fast enough to cause discomfort to the listener; listener may feel obligated to help out	Rate of speech does not impede communication; occasional unnatural pauses and fillers do not distract significantly from the message	Rate and flow of speech are usually natural and facilitate communication

Vocabulary *(Use of words and expressions required for the situation)*

Vocabulary is inadequate to communicate intended ideas; often lacks even common, basic words and expressions	Uses some situation-specific vocabulary, but often lacks words and expressions needed to convey complete ideas; sometimes uses the wrong words or uses the same words repeatedly	Uses an adequate range of situation-specific vocabulary; words and expressions are sometimes imprecise, but speaker finds a way to convey intended meaning	Uses a broad range of appropriate and precise words and expressions needed to convey intended ideas; no searching for words

Grammar *(Application of language rules to generate correct forms and sentence structure)*

Does not use language rules; speech consists mainly of individual words strung together, with no regard for correct forms	Uses a limited range of language rules; as many errors as correct forms	Correctly uses a fairly broad range of language rules most of the time; uses grammar that is clearly required by the situation	Consistently and correctly uses an extensive range of language rules as required by the situation; few if any errors, even in less common or complex forms