Authenticity of "Language Town" as an Innovation in Assessing Learners' Speaking Ability: Moving towards a Virtual Language Town (VLT)*

Marzieh Souzandehfar

Assistant Professor in TEFL, Department of Translation Studies, Jahrom University

Seyyed Mohammad Ali Soozandehfar*

Assistant Professor in TEFL, Department of Teaching English as a Foreign Language, University of Hormozgan. (Corresponding author)

Abstract

Authenticity has always been the concern of test developers in the history of second language assessment. This study was an attempt to investigate the authenticity of the present researchers' innovative idea of "Language Town" as a method for assessing learners' speaking ability. To this end, a simulated town was designed like a real town in an outdoor space of about 400 square meters. The participants in this study were 31 undergraduate students of Translation Studies at Jahrom University who were taking the 4-credit course of Conversation 2. First, the students' speaking ability was measured based on the IELTS testing system. Then each student was sent to the Language Town with a few definite missions determined in advance by the teacher; e.g. ordering food in the restaurant, depositing/withdrawing money in a bank, etc. Using IELTS band descriptors, each student's speaking ability was measured by two raters in both tests. Then the correlation between the two sets of scores obtained from the IELTS test and the one in the Language Town were calculated. Using open-ended questions, a survey was also conducted to extract the students' attitudes towards the Language Town. The results of the statistical analyses showed a weak correlation (0.36) between the two sets of scores. The survey also revealed that almost all the students were interested in and had positive views towards Language Town as an authentic method of assessment. A Virtual Language Town (VLT) could be a solution to the practicality problems of the Language Town.

Keywords: Language Town, Authentic Assessment, Speaking Ability

Introduction

The major problem of teachers and learners of a second language, particularly those who deal with the language for the purpose of using it in the real context of the target language community, is the lack of authenticity of the teaching and testing methods. As a result, such approaches do not prepare the students appropriately for the real and challenging tasks which, in addition to syntax, language functions, fluency and coherence, pronunciation, range of lexis, demand other vital context-dependent competencies, such as pragma-linguistics, sociolinguistics, etc. In fact, the history of language testing is, to a large extent, the history of attempts to bridge the gap between tests and real-life language use. It is the history of progress towards more authenticity in language testing (Ingram, 2003, p. 4).

Considering the IELTS test as one of the most recent developments in the field of language testing, second language teachers have a tendency to use this test for assessing their students' general competence, including their speaking ability. However, different studies have revealed that in spite of an overall positive response to the test, the results of the IELTS tests cannot truly predict the learners' success in the real context of the target language community (Allwright & Banerjee, 1997; Cotton & Conrow, 1998; Hill, Storch, & Lynch, 1999; Kerstjen & Nery, 2000; Moore & Morton, 1999). This situation is even exacerbated when it comes to testing the learners' speaking ability. For example, Paul (2007) suggested that language production at a micro level similar to that in IELTS Speaking Test tasks is not necessarily an indicator of overall language adequacy at a macro level. Most directly, Rea-Dickins, Keily, and Yu (2007) found the Speaking Module to be a poor predictor of test takers' future academic performance. Rea-Dickins et al. (2007) added that even when students' scores on the sub-skills of the test fulfill a program's admission requirements, students are found to lack critical thinking and evaluative skills.

Furthermore, Ducasse and Brown (2011) indicated that the structured nature of the IELTS speaking interviews appears to limit the

ability of interviewers to elicit a broader range of interactional functions, even with more proficient speakers. They argued that as candidates are only required to respond to direct questions or propositions, no evidence of their ability to participate actively in oral discussions – to find a way to share and discuss their own knowledge or ideas, express their opinions, and challenge, support, or evaluate others' contributions – was gathered. According to Ducasse and Brown (2011), studies concerned with the predictive validity of IELTS have generally found little or no significant relationship between IELTS scores and subsequent general and academic performance. This lack of relationship is even more evident in IELTS Speaking Tests.

According to Pinner (2016), the problem is directly related to the issue of authenticity of language tests, including the IELTS. In one of his most recent works, Pinner (2016) replaced the 'classic' definition of authenticity with a reconceptualized version, which, as he claims, is more inclusive to other varieties of English. He poses the 'paradox of authenticity' arguing that authenticity is not something absolute, but "rather relative to the learner and their unique and individual beliefs" (p. 1). He states that in the real world, learners' motivations, in interaction with multiple contextual factors, make the task of communication quite complex and challenging; a condition which hardly exists in current language tests. Consequently, such tests are not authentic enough to guarantee the learners' successful task accomplishment in the real world. He tries to discuss authenticity in light of emergent theories of language acquisition such as chaos/complexity theory and dynamic systems approaches and consequently, introduces the Authenticity Continuum, which is a framework for treating authenticity as a "socially mediated" and "contextually dependent dynamic process of investment". Based on this definition, one can easily recognize the fact that the designers of language tests, including the IELTS, have largely ignored the important factors of context and society in the administration of their tests. In fact, most of the tests are administered in artificial situations which lack the necessary conditions of the real world contexts.

Concerning the authenticity limitations of the IELTS tests, particularly the speaking module, this study was an action research to practice the present researcher's innovative idea of "Language Town" as an authentic method to replace IELTS speaking test for assessing learners' speaking ability in her Conversation classes.

Language Town

Language Town is a simulated model town which is designed like a real town consisting of different places (e.g. a bank, a restaurant, shops, hospital, bus station, etc.), with some real primary facilities in each location and some proficient English language speakers in charge of each location. These proficient speakers play the role of attendants who interact with customers to fulfill their needs, just like what happens in the real world. The area of this town could be variable depending on the available space. The most important thing to consider is to design the town as real as possible to give the examinee the feeling of an authentic context. Each examinee is given a few missions, such as shopping, ordering food in a restaurant, depositing money in a bank, or other similar tasks that an individual is required to do in the real world. Those examinees that succeed to accomplish the task through effective communication, are considered as successful language users. The missions that are assigned to the students are selected based on the needs analysis carried out in advance.

Objective of the Study and Research Questions

This study was intended to investigate the authenticity of the present researchers' idea of "Language Town" as a method for assessing learners' speaking ability. More specifically, this study tried to answer the following questions:

1. Is there any relationship between the learners' speaking scores obtained through Language Town method of assessment and those achieved using the IELTS testing system?

2. What are the examinees attitudes towards Language Town as a method of assessing speaking ability?

Design of the Study

This study used a mixed method to evaluate the authenticity of Language Town as a method for assessing learners' speaking ability. In the first phase of the study quantitative research was carried out to determine whether there is any correlation between the IELTS scores and Language Town scores. In the second phase of the study, a survey was conducted to elicit the participants' feelings and attitudes towards Language Town.

Participants

The participants in this study were 31 undergraduate freshman students of Translation Studies (11 males and 20 females) at Jahrom University who were taking the 4-credit course of Conversation 2 with the present researcher. It was the second semester of the academic year.

Instruments

IELTS speaking module questions were used to assess the learners' speaking ability in the first phase of the study. Different sets of questions were used for different participant to avoid practice effect.

IELTS Band Descriptors was used to assess the students' performance on both speaking tests (Appendix A).

Open-ended questions were used in the survey to elicit the learners' attitudes towards Language Town as a method for assessing their speaking ability (Appendix B).

Data Collection

In the first phase of the study, the students' speaking ability was evaluated based on the IELTS testing system in two sessions. Conversation 2 is a 4-credit course which is held two sessions a week. The assessment took about 4 hours in each session. In this phase each student was interviewed by the researcher using the questions of the three parts of the IELTS test. In addition to the researcher who rated each student, another rater who was trained in advance was present to evaluate the students' performance using IELTS Band Descriptors. The average score of the two raters were calculated for each student to represent their IELTS speaking score. In the following week, Language

Town method was used to assess their speaking ability. As a result, during the two sessions in the following week, each student was sent to the Language Town, which was designed in advance. Each student was given a few definite missions determined by the teacher; e.g. shopping, ordering food in a restaurant, depositing/withdrawing money in a bank, chatting and discussing a topic with friends, talking to university professors in their office, etc. In the present study, proficient senior students were appointed as attendants in different locations of the Language Town. Each student's speaking ability was measured by two raters who followed the students in each location in the town, observed them, and based on the IELTS Band Descriptors, evaluated their communicative competence through interaction with the attendants in each location. Just like the IELTS scores, the average scores of the two raters gained from the Language Town method were calculated for each student to represent their speaking ability.

In the second phase of the study, a survey was conducted to extract the students' attitudes towards the Language Town as an authentic method for assessing their speaking ability. To this end, a set of open-ended questions were designed in advance to be asked from the students. The first two questions asked about the examinees feelings towards their Language Town experience. And the second two questions focused on the authenticity of the tasks the examinees were assigned to accomplish. The questions were written in Farsi (the students' mother tongue) to avoid any probable misunderstanding. Also the students were asked to write their answers in Farsi to let them express their actual attitudes freely and comfortably. The examinees wrote their answers to the questions on sheets of papers which were collected for later analysis.

Data Analysis
Pearson Correlation was conducted between the two sets of scores obtained from the IELTS test and the Language Town method.

The answers to the survey questions were read and analyzed carefully to find about the students feelings and attitudes towards Language Town approach. Also the students' opinions towards the

authenticity of the Language Town were elicited through the analysis of the answers.

Results and Discussion

Regarding the first question of the study, whether there is any relationship between the learners' speaking scores obtained through Language Town method of assessment and those achieved using the IELTS testing system, Tables 1 and 2 provide the answer. Having a look at the mean scores of the tests in Table 1, one can immediately notice that the mean of Language Town scores (4.8387) is smaller than that of the IELTS scores (5.8548).

Table 1

Descriptive Statistics

	Mean	Std. Deviation	N
IELTS	5.8548	.91464	31
LGTOWN	4.8387	1.59367	31

Furthermore, Table 2 shows a weak correlation index (0.36) which is significant at the 0.05 level. This means that there is not a strong relationship between the participants' IELTS speaking scores and their scores obtained via the Language Town.

Table 2

Correlations

		IELTS	LGTOWN
IELTS	Pearson Correlation	1	.361[*]
	Sig. (2-tailed)		.046
	N	31	31
LGTOWN	Pearson Correlation	.361[*]	1

Sig. (2-tailed)	.046	
N	31	31

*. Correlation is significant at the 0.05 level (2-tailed).

One possible interpretation for the lower mean score of the Language Town test compared to that of the IELTS test might be the fact that Language Town tasks were more difficult for students to accomplish than the IELTS speaking module questions. This confirms Pinner's (2016) argument that in the real world, learners' motivations, in interaction with multiple contextual factors, make the task of communication quite complex and challenging; a condition which hardly exists in current language tests, including the IELTS tests. In fact, excluding the context and simply sitting in front of the examiner and answering a set of questions in IELTS tests is far easier than being situated in a real context which demand more complex and multidirectional interactions both with different participants and the real objects present in the environment.

On the other hand, the weak correlation in Table 2 can imply that IELTS tests cannot truly predict the learners' success in the real context of the target language community. This is exactly in line with the previous studies (Allwright & Banerjee, 1997; Cotton & Conrow, 1998; Hill, Storch, & Lynch, 1999; Kerstjen & Nery, 2000; Moore & Morton, 1999; Paul, 2007; Rea-Dickins, et al. 2007) which indicated a low predictive validity for the IELTS tests. One good justification for the weak correlation between the two sets of scores could be what Ducasse and Brown (2011) argued that the structured nature of the IELTS speaking interviews appears to limit the ability of interviewers to elicit a broader range of interactional functions, even with more proficient speakers. In fact, since candidates are only required to respond to direct questions or propositions, there is no evidence of their ability to participate actively in oral discussions, such as finding a way to share and discuss their own knowledge or ideas, expressing their opinions, and challenging, supporting, or evaluating others' contributions. This is also confirmed by the lower mean of the Language Town scores

compared to the IELTS scores, showing the fact that Language Town tasks probably demanded a higher-order and more complex cognitive ability compared to those in the IELTS test.

Furthermore, the results of the survey revealed that almost all the students were interested in and had positive views towards Language Town as an authentic method of assessment. For example, one of the students stated, "testing in a … mmm… real context like Language Town is more exciting and …mmm… more meaningful." At the same time, they argued that the tasks in the Language Town were more challenging to accomplish. For example, when one of the participants was asked why he thought Language Town tasks were more demanding, he answered, "in the IELTS speaking test, the only thing we …mmm… had to do was to answer the questions the interviewer asked… but in the Language Town I have to speak with a person, act physically, and work with real objects to complete the tasks… so I need more concentration." Some other students argued that since it felt like a real context and that they had to interact with different people, speaking and communication were more difficult in the Language Town compared to the IELTS test. In fact, most of the answers implied the fact that the multiple contextual factors involved in the Language Town had made it more challenging and demanding.

On the whole, considering the results of both the correlation and that of the survey, one can infer that the Language Town has provided an authentic context which, as Pinner (2016) stated, replaced the 'classic' definition of authenticity with a reconceptualized version, which, as he claims, is more inclusive to other varieties of English. In fact, Language Town created a real context in which the learners' unique and individual beliefs were involved in doing the tasks. Language Town, due to its similarity to the real world situations, entangles learners' motivations, in interaction with multiple contextual factors and consequently, makes the task of communication quite complex and challenging; a condition which hardly exists in other language tests, like IELTS. In fact, a testing context like the Language Town can introduce tasks which are, as

Pinner (2016) argued, "dynamic", "socially mediated", "contextually dependent", and in one word, authentic.

Limitations of Language Town Idea

One obvious limitation of the Language Town idea as an authentic method for assessing the speaking ability is its low practicality, especially for high-stake tests. The need for a large space, decoration, various objects, and proficient attendants in each location are both expensive and time-consuming. In fact, it is not economical to develop such a Language Town for every speaking test and by every institute independently. However, one possible solution to this problem is to replace current language labs in schools, universities, and language institutes with small Language Towns. That is, we can devise one large, fixed, and permanent Language Town as a center specifically allocated to testing learners' speaking ability. In this way, the Language Town will be designed and equipped once and used by different groups of learners who should reserve the Language Town in advance to be used for their testing purposes.

Another idea, which could be even more economical than those mentioned above, is to develop a Virtual Language Town (VLT). This is what can be used easily in high-stake tests as well. In the following section, an imaginary VLT has been described and the way it functions has been elaborated.

Virtual Language Town (VLT): An economical solution to current inauthentic language tests

Wherever it is too dangerous, expensive or impractical to do something in reality, Virtual Reality (VR) is the answer (Virtual Reality Society, 2017). Considering the practical limitations of the Language Town, VLT can be considered as an innovation in the field of language testing. VLT is a virtual interactive Language Town utilizing virtual reality technology and computer programs which place the examinees in simulated 3D locations with smart virtual participants with whom test takers can communicate and be evaluated by automatic and intelligent raters based on predetermined rubrics and criteria defined for the program. In fact, VLT can make use of VR technology to place

language test takers in different locations and situations to interact with various virtual characters in each place. This is exactly what happens in interactive computer games. This can be done through collaboration among TESL, linguistics, and computer program experts.

One can think of an examinee standing in front of a 3D screen which displays a Language town. She starts walking in the streets looking for a bank nearby. She finds the bank, enters the building, takes a turn, and goes to the virtual banker. She starts communicating with the banker to do her banking affairs. After she completes her mission in the bank, she comes out of the building and since she feels a little bit tired, she goes to a coffee shop for some drink. She sits at the table, picks the menu and selects her favorite drink. Then she asks the waiter to come to her table and write down her order. After having her drink she asks for the bill and pays the bill. These are some sample missions that are given to the examinee to be accomplished within a predetermined amount of time to test her language use competence and in particular her speaking ability. The VLT is programmed in such a way that only those examinees who can communicate effectively can complete each task successfully and be considered as competent language users. Different levels of competencies can be defined in advance for the program and each candidate, based on their knowledge, can be assigned to their appropriate levels. Depending on the institute's goals, VLT can be programmed for both prognostic and evaluation of attainment tests. Furthermore, VLT can be administered simultaneously for a large number of students depending on the number of 3D screens available in the testing center.

Conclusion

Regarding the results of the present study, Language Town, especially its VLT version, can be considered as an authentic and economical method for testing all language skills, including the speaking ability. In fact, Language Town can replace language labs in different institutes including schools and universities to solve the problem of current language tests which suffer from inauthenticity and low predictive validity. Being situated in a simulated context with those characteristics

of the real world, both the examiners and examinees can much better evaluate one's competencies in a second or foreign language. Language Town method of assessment is especially useful for those teachers and learners of a second language, who in particular deal with the language for the purpose of using it in the real context of the target language community. Considering vital context-dependent competencies, such as pragma-linguistics and sociolinguistics, in addition to the knowledge of syntax, language functions, fluency and coherence, pronunciation, range of lexis, Language Town can prepare learners appropriately for the real and challenging tasks they encounter in the real world.

References

Allwright, J., & Banerjee, J. (1997). *Investigating the accuracy of admissions criteria: A case study on a British university*. Lancaster: Institute for English Language Education, Lancaster University.

Cotton, F., & Conrow, F. (1998). An investigation of the predictive validity of IELTS amongst a sample of international students studying at the University of Tasmania. In S. Wood (Ed.), *IELTS Research Reports, 1*, 72-115. Canberra: IELTS Australia, Pty Ltd & British Council.

Ducasse, A. M. & Brown, A. (2011). The role of interactive communicationin IELTS speaking and its relation to candidates' preparedness for study or training contexts. *IELTS Research Report, 12*, 1-26.

Ingram D.E. (2003). Methodology in the new millennium: Towards more authenticity in language learning and assessment. Paper to the First International Conference on pedagogies and learning, *New meanings for the new millennium*, University of Southern Queensland,Toowoomba, 1-4 October, 2003.

Ingram, D. & Bayliss, A. (2007). IELTS as a predictor of academic language performance, Part 1, in *IELTS Research Reports Volume 7*, IELTS Australia, Canberra and British Council, London, 137-203.

Kerstjen, M., & Nery, C. (2000). Predictive validity in the IELTS test. In R. Tulloh (Ed.), *IELTS Research Reports* (Vol. 6), (pp. 85-108). Canberra: IELTS Australia, Pty Ltd.

Moore, T. & Morton, J. (1999). Authenticity in the IELTS Academic Module Writing Test: A comparative study of Task 2 items and university assignments. *EA Journal*, 2, 64-106.

Paul, A. (2007). IELTS as a predictor of academic language performance, Part 2. In P.

McGovern & S. Walsh, *IELTS Research Reports,* (Vol. 6) (pp. 205-240). Canberra: IELTS Australia, Pty Ltd.

Rea-Dickins, P., Keily, R., & Yu, G. (2007). Student identity, learning and progression: The affective and academic impact of IELTS on 'successful' candidates. In P. McGovern & S. Walsh, *IELTS Research Reports,* (Vol. 6) (pp. 56-136). Canberra: IELTS Australia, Pty Ltd.

Appendices

Appendix A

Band	Fluency and coherence	Lexical resource	Grammatical range and accuracy	Pronunciation
9	• speaks fluently with only rare repetition or self-correction; • any hesitation is content-related rather than to find words or grammar • speaks coherently with fully appropriate cohesive features • develops topics fully and appropriately	• uses vocabulary with full flexibility and precision in all topics • uses idiomatic language naturally and accurately	• uses a full range of structures naturally and appropriately • produces consistently accurate structures apart from 'slips' characteristic of native speaker speech	• uses a full range of pronunciation features with precision and subtlety • sustains flexible use of features throughout • is effortless to understand
8	• speaks fluently with only occasional repetition or self-correction; hesitation is usually content-related and only rarely to search for language • develops topics coherently and appropriately	• uses a wide vocabulary resource readily and flexibly to convey precise meaning • uses less common and idiomatic vocabulary skilfully, with occasional inaccuracies • uses paraphrase effectively as required	• uses a wide range of structures flexibly • produces a majority of error-free sentences with only very occasional inappropriacies or basic/non-systematic errors	• uses a wide range of pronunciation features • sustains flexible use of features, with only occasional lapses • is easy to understand throughout; L1 accent has minimal effect on intelligibility
7	• speaks at length without noticeable effort or loss of coherence • may demonstrate language-related hesitation at times, or some repetition and/or self-correction • uses a range of connectives and discourse markers with some flexibility	• uses vocabulary resource flexibly to discuss a variety of topics • uses some less common and idiomatic vocabulary and shows some awareness of style and collocation, with some inappropriate choices • uses paraphrase effectively	• uses a range of complex structures with some flexibility • frequently produces error-free sentences, though some grammatical mistakes persist	• shows all the positive features of Band 6 and some, but not all, of the positive features of Band 8
6	• is willing to speak at length, though may lose coherence at times due to occasional repetition, self-correction or hesitation • uses a range of connectives and discourse markers but not always appropriately	• has a wide enough vocabulary to discuss topics at length and make meaning clear in spite of inappropriacies • generally paraphrases successfully	• uses a mix of simple and complex structures, but with limited flexibility • may make frequent mistakes with complex structures though these rarely cause comprehension problems	• uses a range of pronunciation features with mixed control • shows some effective use of features but this is not sustained • can generally be understood throughout, though mispronunciation of individual words or sounds reduces clarity at times
5	• usually maintains flow of speech but uses repetition, self correction and/or slow speech to keep going • may over-use certain connectives and discourse markers • produces simple speech fluently, but more complex communication causes fluency problems	• manages to talk about familiar and unfamiliar topics but uses vocabulary with limited flexibility • attempts to use paraphrase but with mixed success	• produces basic sentence forms with reasonable accuracy • uses a limited range of more complex structures, but these usually contain errors and may cause some comprehension problems	• shows all the positive features of Band 4 and some, but not all, of the positive features of Band 6
4	• cannot respond without noticeable pauses and may speak slowly, with frequent repetition and self-correction • links basic sentences but with repetitious use of simple connectives and some breakdowns in coherence	• is able to talk about familiar topics but can only convey basic meaning on unfamiliar topics and makes frequent errors in word choice • rarely attempts paraphrase	• produces basic sentence forms and some correct simple sentences but subordinate structures are rare • errors are frequent and may lead to misunderstanding	• uses a limited range of pronunciation features • attempts to control features but lapses are frequent • mispronunciations are frequent and cause some difficulty for the listener
3	• speaks with long pauses • has limited ability to link simple sentences • gives only simple responses and is frequently unable to convey basic message	• uses simple vocabulary to convey personal information • has insufficient vocabulary for less familiar topics	• attempts basic sentence forms but with limited success, or relies on apparently memorised utterances • makes numerous errors except in memorised expressions	• shows some of the features of Band 2 and some, but not all, of the positive features of Band 4
2	• pauses lengthily before most words • little communication possible	• only produces isolated words or memorised utterances	• cannot produce basic sentence forms	• Speech is often unintelligible
1	• no communication possible • no rateable language			
0	• does not attend			

Appendix B

سوالات بخش نظرسنجی:

۱. احساس خود را در مورد تجربه شهرک زبان توصیف کنید.

۲. بطور خاص و با ذکر دلایل توضیح دهید آیا نظر شما نسبت به این شیوه ی ارزیابی مثبت است یا منفی؟

۳. کدام شیوه ارزیابی را بیشتر منطبق بر نیازهای ارتباط شفاهی خود در دنیای واقعی می دانید، روش آیلتس یا شهرک زبان؟ با ذکر دلیل توضیح دهید.

۴. از نظر شما کدام شیوه ارزیابی دشوارتر و چالشی تر می باشد، آیلتس یا شهرک زبان؟ با ذکر دلیل توضیح دهید.