Language Proficiency and Identity: Developing a Structural Equation Modeling (SEM) of Identity for Iranian EFL Learners

Amir Mirahmadi Kia, Ph.D. Candidate, Department of English Language, Qeshm Branch, Islamic Azad University, Qeshm, Iran
amir.mirahmadikia@yahoo.com
Seyyed Ayatollah Razmjoo, Department of Foreign Languages and Linguistics, Shiraz University, Shiraz, Iran
arazmjoo@rose.shirazu.ac.ir
Shahram Afraz, Assistant Professor, Department of English, Qeshm Branch, Islamic Azad University, Qeshm, Iran
a.sh32@rocketmail.com

Abstract
This study was an endeavor to develop a model of identity among Iranian EFL learners. To achieve this end, a multiphase design was implemented. Initially, it attempted to investigate different factors of identity to propose and validate a model. Thus, 120 EFL learners studying in different English language institutes in Iran were randomly selected, and 36 learners were interviewed about their views of their identity in the qualitative phase. After extracting six factors of identity, including: second language acquisition and social status, cultural attachment, Persian language adhesion, pronunciation posture, technology involvement, and language identity, and second language knowledge, a questionnaire was constructed which reflected these factors. Then, in the quantitative phase, the questionnaire went through an exploratory factor analysis for the sake of validity. After its validity and reliability were corroborated through a pilot study with 20 learners, it was distributed among 120 EFL learners. Besides, Structural Equation Modeling (SEM) analysis was run to confirm that the final proposed model enjoyed validity for future research. To do so, the confirmatory factor analysis was run, and the model of identity was developed. Eventually, the possible relationship between 120 EFL learners' identity and their English language achievement scores were examined, and the results of this phase indicated that there was a significant and positive relationship between learners' identity and their English language achievement scores. The findings of this study can enhance awareness among English teachers, materials developers, and syllabus designers to equip themselves with the updated techniques to handle the possible challenges that may occur in EFL learning contexts.

Keywords: Identity, language proficiency, EFL context, Structural Equation Modeling (SEM)

Introduction

In recent years, in order to enhance the learner's communicative competence in the second language (L2), the application of a pragmatic component has been considered as an essential and influential aspect (Hinkel, 1999; Kasper, 1997; Rose & Kasper, 2001). Furthermore, many types of research about Inter-Language pragmatics to examine L2 learners' communicative competence have been carried out (Rose and Kasper 2001, 3-8). According to Kasper and Rose (2002), L2 learners reinforce the capabilities to understand and perform an action in a target language using L2 pragmatic capabilities (p.5). In 2001, Rose and Kasper proved that conducting an instructional intervention to train the learners' acquisition of the competence to communicate in the target language appropriately is essential and potential. However, learners who want to communicate in

the target language and try to find a correct link between their mother language and target language (TL) may have some competency shortages. Relatively, little pragmatic- based instruction create some difficulties in L2 negotiations for the learners. This study intended to address this lack and echoed the need to add pragmatics to the TL instruction.

In the current world of language learning, identity is receiving increasing attention. In the last decade, many scholars in the filed acknowledged the need for investigating the role of identity in language learning (Pavlenko & Lantolf, 2000; Schwartz, 2005). Riley (2007) believed that the way language shaped and was shaped by identity was a key topic within sociolinguistics. Many different factors and elements procreate people's identity, consisting of social, territorial, linguistic, cultural, political, religious, and ethnic related issues. For the same reason, identity and language learning were related to each other. Scientists and researchers were interested in the second and foreign language learning, language education, sociological and cultural aspects of language learning were significantly focused. The expression of self and an individual's identity is one of the essential and chief functions of language; however, linguists have significantly neglected this issue. Although the expression of self and individuals' identity is one of the principal functions of language, linguists have primarily neglected it, especially in foreign language learning. On top of that, religious identity is a factor that fails to provide for in research on the relationship between identity and language learning. Besides, Islamic identity is a deeply rooted element that has touched different aspects of Iranian language learners' lives.

Kinginger in 2004 had a case study about foreign language and learner identity. In that study, she reported a survey conducted on one of her students, Alice, who was studying at an American university. Alice`s field of study was the French language. The study was longitudinal. Alice's age was older than average among others, and to pay her tuition, she worked several jobs. However, she was a little disappointed with studying the French language in her college.

Meanwhile, Alice created a romantic image of France, which was somewhat unrealistic, according to her foreign language study in college. Travel brochures and television shows represented the image that was similar to Oz or wonderland; therefore, these similarities look like Alice's image of France. While Alice shared her own culture, she also expected the French people to be cultural, friendly, and accepting as she dreamed of making friends with the natives and learning their cultures. Alice's behavior toward language learning, her imagined communities of French language users, her accessibility to social networks at home and abroad, and her language learning usage as a source of coherence and lessons in persistence were dealt with by the study (p. 223). In this study, the specification of a renegotiated identity according to Alice's behaviors to clarify where such behaviors came from and how they were changed over time was determined by the author. Alice's imagination about the communities of French language speakers, her role within them, and the emblematic capital she would obtain via these efforts were extended by the questions. Besides, the kinds of community and the way she approached them were concerned by the author.

Among the most recent studies done on identity and L2 language in Iran, the one by Razmjoo and Izadpanah (2012) on the relationship between L2 literacy (reading and writing) and identity processing styles and also the one by Razmjoo (2010) on aspects of identity and EFL learners` achievement are among the notable ones**.** In the former one, the authors dealt with identity processing styles or the ways people are predisposed towards identity. In the latter, the author dealt with the aspects of identity, which are mainly various aspects of self-concept according to the argument put forward by Sussman (2000). The focus of the present study is developing a model of identity for Iranian students or, putting it more simply, the relationship between students' language proficiency and the identity model. Accordingly, more research is

needed to find out how the general identity of Iranian language learners is associated with foreign language learning. The findings of such studies would pave the way for a more in-depth analysis of aspects of identity and the way they facilitate or hinder foreign language learning. According to Day (2004), in learner education, one important goal was to prepare learners who were informed and flexible to manage the imposed changes in the curriculum and education policies while trying to understand issues such as learners` sense of educational aim, practices, learner identity, and agency.

Interestingly, as Beijaard (1995) asserted, high school (HS) learners` identity was essentially linked to their subject area, which was the case of the present study of HS EFL learners who faced imposed changes, requirements, and policies in Iran. The top-down strategy implemented has affected learners' professional identity in different ways, but positively or negatively? It was necessary to evaluate the impact on learners' identity before embarking upon a new program. Therefore, the following questions would be considered to achieve the proposed objectives.

Q1. How identity formed in EFL language learners of the institutes?
Q2. What model can possibly be formulated through SEM to explain identity among Iranian EFL learners?
Q3. What is the relationship between language proficiency and the identity model?

Literature Review

In 2019, Rezaei et al., in a study, conducted a survey of Iranian language teachers in Iran. Based on these data, a paradigmatic identity model is presented. The results of the confirmatory factor analysis showed that this model is suitable and has eight components. These findings also showed that old and experienced teachers have the highest level of Iranian cultural identity. In addition, the results showed that female teachers have a strong Iranian cultural identity compared to their male counterparts, and teachers in different fields of study with different mother tongue and university degrees do not differ much in terms of cultural identity.

In 2018, Khatib et al., conducted a study aimed at highlighting the main components of Iranian cultural identity and confirming the above components through structural equation modeling analysis (SEM). After analyzing the structural equation modeling (SEM) via AMOS, 22 was performed to test the model and the interaction between the components. The SEM results confirmed the presence of five factors. To achieve these goals, the researchers examined extensive local and international literature on language, culture, and identity. Based on literature and counseling with a 30-member group of undergraduate and graduate English language learners. Four university professors in sociology, an Iranian-speaking EFL cultural identity model with six components, were hypothesized: (nationality, religion, art, language, and Farsi Literature, Media, and Globalization). After analyzing the structural equation modeling (SEM) via AMOS 22 was performed to test the model and the interaction between the components. The SEM results confirmed the presence of five factors.

A study has been conducted by Mohammadi et al., in 2018, with the aim of examining the relationship between the cultural and social identity of Iranian learners and EFL learners. Data normality was assessed based on normal Skewness and Kurtosis tests, so the Pearson correlation coefficient was used in this study. The results showed a significant negative relationship between students' learning and their EFL learning (-678). It will be useful for curriculum designers, learners, their parents, and language teachers.

A study conducted by Razmjoo and Mavaddat (2015) sought to examine how justice judgments, outcomes, and identity orientations are related. To this end, the investigation used a structural equation modeling, intending to find out the possible mediating impact of identity orientation on the relationship between justice judgment and outcomes. Seventy-four students participated in this study. They were majoring in TEFL in the Department of Foreign Languages and Linguistics at Shiraz University. The participants were selected based on convenience sampling. The following three questionnaires were distributed and filled out by the participants: 1-distributive and procedural justice judgments, 2-rule compliance and outcome satisfaction, and 3-personal and social identity orientations. The data was collected and analyzed through descriptive statistics, correlation, and structural equation modeling. The findings of the study revealed a significant positive correlation between procedural justice and rule compliance.
Moreover, there was a significant correlation between distributive justice and outcome satisfaction. Based on the developed structural equation model, justice judgments were found only directly to impact the outcomes. Finally, identity did not have any mediating impact on the causal relationship between the two.

Sabatin (2013) examined the impact of students' cultural awareness on their learning English. Moreover, it aimed at probing the possible significant differences between the participants' performance on reading comprehension in terms of their gender and general English proficiency. The findings of the study showed significant differences between participants who are culturally aware and those who do not have any knowledge in terms of their performance on reading comprehension. Moreover, the results showed no statistically significant differences between male and female participants who have the cultural background knowledge and those who do not know about their performance on reading comprehension. The participants' scores of language proficiency revealed significant differences between subjects who have the cultural background knowledge and those who do not know when it comes to their performance on reading comprehension. The researcher concludes that teachers need to activate two kinds of prior knowledge when teaching a new topic: subject knowledge and cultural knowledge.

Razmjoo and Izadpanah (2012) addressed any relationship between identity processing styles and Persian EFL learners' L2 literacy. The sample of this study consisted of 160 advanced EFL learners (109 girls and 51 boys). The study made use of the following instruments: 1-a simulation of the TOEFL IBT test, aimed at measuring the subjects' L2 literacy-related proficiency; 2- a reading test; 3- a writing task; and 4- the identity style inventory (ISI-4) to measure the participants' identity styles. The findings showed that the identity processing style makes an essential contribution to L2 literacy-related proficiency. A slight positive relationship was observed between the informational processing style, L2 literacy, and L2 writing.
On the contrary, a negative and weak correlation was found between normative style and these components. They cannot significantly predict the learners' success or failure concerning overall L2 literacy. The authors of this study recommend the integration of identity surveys into L2 systems to help educators and teachers to teach L2 literacy-related proficiency more effectively.

Methodology
Participants
Considering the total population of EFL learners in Iran, a sample of 120 students were randomly selected to participate in the study. The sampling procedure was conducted in two parts. The following provide demographic information of the sample in the qualitative and quantitative sections. In this first section of the study, 36 EFL language learners were randomly selected to participate in an in-depth interview. The sample was selected from different provinces

of Iran: Azarbayjan Gharbi (3), Azarbayjan Sharghi (3), Fars (3), Hormozgan (3), Khorasan Razavi (3), Bushehr (3), Tehran (3), Karaj (3), Isfahan (3), Mazandaran (3), Sistan and Baluchistan (3) and Kerman (3). All the participants were randomly selected from different English Language Institutes in these provinces. The research assistants visited the institutes and conducted the interview under the supervision of the institute supervisor. Twenty-one learners were female, and 12 learners were male. Their age ranged from 18 to 23. Their proficiency ranged from elementary to advance based on the language proficiency test. In this second phase of the study, 120 EFL learners were randomly selected from different parts of Iran. The sample was selected from different provinces, including Fars (9), Hormozgan (10), Khorasan Razavi (9), Bushehr (10), Tehran (13), Karaj (9), Isfahan (11), Mazandaran (12), Sistan and Baluchistan (10), Kerman (15), Azarbayjan Gharbi (5), Azarbayjan Sharghi (7). Sixty-nine of the students were female, and 51 of them were male. Their age ranged from 18 to 23. Their proficiency ranged from elementary to advance based on the language proficiency test.

Instruments

This study employed some instruments to obtain data to answer research questions that were postulated in chapter one. The selection of a proper methodology was one of the most essential and crucial parts of the research study. The tool that would be used for this study was an interview which its reliability could be examined using Alpha Cronbach, and its validity was presented by Factor analysis. For more examination of its validity, TOEFL PBT could be used. To collect the data, the questionnaires would be distributed among the learners. For data analysis, it would be checked how each component of the identity would be related to proficiency. Acquiring proficiency in a particular language was assumed to be a cline and was "the degree of control one has over the language in question" (Hamayan & Damico, 1991, p. 41). One of the definitions of language proficiency referred to it as the degree of skill with which a person could use a language, such as how well a person could read, write, speak, or understand language could be contrasted with language achievement, which described language ability as a result of learning. However, language proficiency was a commonly used term and understood superficially by ordinary people as one's facility used a specific language, particularly in speaking and writing, offering a definition for this type of proficiency was not as simple as many belief. The common aspect among all the language proficiency theories was the fact that they had all incorporated grammatical competence into their framework as a crucial feature of language proficiency, and they tried to highlight the importance of grammatical and lexical knowledge in obtaining proficiency in the target language, in this case, English. By testing these skills, the researcher checked the grammatical and vocabulary knowledge and word power of the participants in writing skills. Therefore both semantic and syntax level of students was checked. Moreover, for analyzing the collected data, multiple regression analysis would be applied.

Data Collection Procedures

The data collection process was done through three steps. At first, the TOEFL PBT proficiency test was used to evaluate the EFL learners' skills. TOEFL was used as the test of language proficiency in this study. It was administered in a paper-and-pencil format, measured test takers' ability to use and understand English in a classroom setting at the college or university level. It evaluated reading, listening, and writing skills and knowledge of English conventions through four test sections: listening comprehension, structure, and written expression, reading comprehension, and the Test of Written English. The total timing for this test is 140-150 minutes, which includes 140 questions and one essay topic. In the second stage, the

oral interview was used to be asked the ELF learners from 12 different zones of Iran. Based on the present study, the interview questions were designed to relate the EFL learners to this study, to have the SWOT Analysis, and after that, to gain in-depth learners` statements and ideas. To have more analysis, their opinions about the six main factors were recorded. The interview questions were given to them in advance; therefore, EFL learners can read them and have enough preparation to answer them and organize them. Taking notes while interviewing was a chance given to them. The interview would continue until no new ideas or opinions were added to the initial EFL learners` sentences. When the interviewer figured out this fact, he did not continue asking more questions about that phase, and he would go on with the next phase to design the identity questionnaire. Credibility, transferability, dependability, and confirmability are the four rigor standards existing in qualitative research (Ary et al., 2006). Therefore, in order to have proper qualitative research, it has to follow the mentioned items. Ary et al., (2006) introduced two strategies to have secure credibility; thus, it was applied in this research. In the first strategy, at the end of the data collection procedure, the transcribed and video recorded data were given back to the interviewees to correct any mistakes they had during the interview session. This strategy is called a member check. In the second strategy, the interviewees` world was experienced by the readers via direct quotations- verbatim. This strategy is called a low-inference description. In the present research, dependability was applied, which was the second rigor of qualitative research. Documentation strategy was implied in order to have more references and work in the future. Finally, all EFL learners were given the identity questionnaire. They had to answer all questionnaire questions; thus, the grounded theory approach was used to analyze all data collected from the identity questionnaire.

1. Qualitative Research Phase

In the qualitative phase of the study, an interview was conducted with 30 EFL language learners. The interview consisted of 17 questions investigating language learners' identity. The interview took 25 minutes for each learner, and the language was Persian to help lower-level learners express their meaning correctly. With the permission of participants, the interviews were recorded using a second recorder (usually the interviewer smartphone). Immediately after the interview, the recordings were transcribed to ensure the reliability of the information provided. The current study employed grounded theory to analyze data in the qualitative section and develop a model of language learners' identity. For this purpose, three coding procedures were followed thoroughly. First, in the open coding phase, the data were broken down into parts to compare and contrast. Some words, phrases, and statements given by the participants were selected as delivering valuable material for the purported method. The process started with details of more general concepts. The researchers gave more questions on these concepts and searched the data more intensively. The general concepts were put into different categories to give a meaningful picture of the analyzed data. Second, in the axial coding phase, the researcher investigated each category, considering who, what, where, and why. The purpose was to investigate and describe the explicit relationship between categories. This process continued to find the core categories. It is the category with the numerous and most substantial links to the other categories, considering the variations and exceptions. The researcher searched for something that can connect all the categories to provide a coherent story. The last stage of coding was selective coding. Among different categories and their links provided in the axial coding section, the researcher searched for the core category with the highest links to guide further data collection.

2. Quantitative Research Phase

According to McMillan and Schumacher (2001), a questionnaire is a common technique for collecting data, and it is used to elicit reactions, attitudes, or beliefs. Moreover, a questionnaire is an instrument commonly used to elicit information about a variety of different issues (Richards, 2001). In the quantitative stage, the researcher decided to apply a developed questionnaire for collecting data on the perspectives of the participants. That is to say; Identity Questionnaire was designed by the researcher and used in this study to measure some categories after conducting interviews. The variables and scales derived the data analyzed in the qualitative section. In order to develop a reliable and valid questionnaire, the researchers went through the following steps.

A. Item accumulation

The categories constructed in the grounded theory stage were used to construct a pool of items. Moreover, to compare and contrast different questionnaires and benefit from their advantages, the literature was examined extensively. Firstly, a pool of 50 items was constructed, which included short and straightforward items using natural language. The items pool was sent to 5 experts in applied linguistics for further analysis. Due to some overlap, 15 items were removed after analysis by the experts.

B. Rating scale development

This study used a Likert-point scale for rating of items in the questionnaire. The points ranged from "strongly disagree, disagree, no idea, agree and strongly agree," and there was no reverse scoring. The purpose of this scale was to provide an accurate picture of respondents' answers and provide a normal distribution of the data.

C. Demographic information section

This section included information like age, gender, language proficiency, province, city, language institute, and length of studying English. It was put at the beginning of the questionnaire for warm-up.

D. Piloting

After designing the questionnaire, it was distributed to 70 ELT postgraduate students (30 males, 40 females) in two ways. It was disseminated to students via self-administered questionnaires and a Web-based survey when the number of self- available administered questionnaires did not reach the optimal count. The self-administered questionnaires were given to the EFL learners of different provinces according to the available sampling procedure. In other words, selecting the participants was based on available sampling in which the participants (EFL learners) were chosen based on their availability. It is one of the effective forms of nonprobability sampling, which was applied in the current research. Probability sampling is ideal, but in practice, available sampling would be all that is available to the researcher (Ary et al., 2014).

Finally, 65 students returned their questionnaires or/and sent back the filled questionnaire through email to the researchers. However, the optimal number of participants was 70. This number was considered for the pilot study because the survey relied on some hypothetical categories of concepts, and as a yardstick for each factor, 40 participants should, at least, have been assumed. As a solution to collect less-costly and economic data, the researchers used an online distribution of data via a Web-based survey. In fact, the researcher could immediately gather optimal data using an online survey with the help of current and previous colleagues to

send the link of the online-survey to their students. As soon as 40 responses were received, and the number of respondents reached 70, they aborted the survey so that no more participants could submit their responses. The goal at this stage of the study was to guarantee whether the current instrument was good enough in terms of face, content, and hypothetical constructs to be employed in the primary research. Besides, the internal consistency of the survey was another factor the designers of the survey wanted to reassure. Eventually, some changes to the wordings, item omission, and the like were carried out to the original questionnaire. It was launched for the primary research with 120 EFL learners of language institutes.

Data Analysis Procedures

Three coding procedures were applied in the current research introduced by Strauss and Corbin (1998). These coding procedures are open, axial, and selective. At first, transcribing the EFL learners` statements occurred. Then, via applying the mentioned coding procedures, the EFL learners` sentences' data was codified. In the initial step, open coding was used to codify the data to gain shared concepts and form categories. In the next step, to have larger categorizations, the obtained categories were worked on similar concepts. In the final step, "themes," which was a more comprehensive categorization, was attributed. To measure the internal consistency of the questionnaire in this study, Cronbach's Alpha coefficient was utilized. Cronbach's Alpha showed that the internal consistency of the whole questionnaire was 0.84, and for the six subscales (i.e., the six components of language identity) in the questionnaire, it was estimated to be 0.78, 0.66, 0.75, 0.79, 0.82, and 0.84, respectively. In the present study, Structural Equation Modelling (SEM), a multivariate analysis technique for exploring causality in models and the causal relations among variables, was run. SEM is rooted in the positivist epistemological belief that was cobbled together out of regression analysis, path analysis, and confirmatory factor analysis. SEM is used as a confirmatory technique to test models that are conceptually derived a priori or test if a theory fits the data. SEM shows the relationship between latent variables: the components of language identity in this study, and the observable variables, that is, the items in the questionnaire generated for each of the components in language identity construct. SPSS software (version 22) was used for data analysis. Exploratory and confirmatory factor analyses were used for validation in the SPSS (version 22) and PLS 3. Analytical tests were used to investigate the relationship between the variables of the study.

Codification of the Data

Maykut and Morehouse (1994) believed that each qualitative research's central part is the data analysis procedure. It can make a big difference between qualitative research and a quantitative one. The data were codified based on the three mentioned coding procedures introduced by Strauss and Corbin (1998). As Ary et al., (2006) mentioned, these pieces of data, categories, and themes are three steps to be referred to.

Open Coding (Step 1): The units of meanings, opinions, occurrences, behaviors, and so on as the data were read and reread to find the repeated statements by most of the interviewees. A wide range of codes appeared in this step, and some of them had to be removed later by the researcher.

Axial Coding (Step 2): Relating categories to their subcategories have a process in which a category link at the level of properties and dimensions introduced by Strauss and Corbin (1998). The development of some categories took place in the open coding procedure. In this procedure,

some categories would be changed later. After that, a category would be formed by bringing the pieces of data on the same topic together.

Some categories in the present research appeared. In this phase, the reader became familiar with the interviewees` opinions and statements, and of course, their world by reading the EFL learners` direct quotations. The English language and social status were two topics for the theme extracted from the EFL learners` statements. The repeated topics about talking, learning, and so on were stated by the EFL learners.

After marking the pieces of data, the second five categories that appeared were facts concerning cultural attachment. They believed that Iranian students are engaged to strive for improved performance. More importantly, in their views, Persian history and culture are highly respected, and the Persian language is taken superior over foreign languages. However, they believed that speaking in the English language is a value in their society.

The third five categories that emerged were related to Persian language adhesion that the interviewees regarded English as a threat to Persian culture and language. Also, they resisted that English values are destroying other cultures. Moreover, the interviewees emphasized on mastering Persian skills and the knowledge of the Persian language was necessary for national unity.

The next two categories referred to pronunciation posture. They were proud of speaking Persian with an English accent, and also they preferred the American English accent to the British.

Another received reports in the interview were technology involvement and language identity, facts about using technology and computer. The study participants reported that using computer empower their self to achieve language learning goal, and they were able to understand materials better. In their ideas, technology improves their "ego" and self-confidence that individuals must function in the English language.

The EFL learners had the last shared concept in their interviews. It was labeled as second-language knowledge; that is, facts about English language learning. They stated that learning English would be useful in the future, helped the growth of the mind, and enriched their native language and mutual understanding at a global level. Besides, they would like to learn the spoken form of foreign languages that mastering a second language grant high self-esteem to them and could estimate their development.

Selective Coding (Step 3): Assigning the categories to design the broader category or theme in the third coding procedure called selective. In this research, six comprehensive concepts or themes were extracted and form the main categorization. These six themes were second language acquisition and social status, cultural attachment, Persian language adhesion, pronunciation posture, technology involvement and language identity, and second language knowledge.

These three coding procedures took the systematicity in the data place; open, axial, selective. The aim was to design a comprehensive image of the identity for the Iranian EFL context. The model was formed by the categories known from the interviews, and the building blocks were the sub-components inside the categories. The categories were intermingled, and each one can affect the other one; by the way, each category stands by itself. In addition to the six main themes/factors of identity for the Iranian EFL context, the results highlighted 30 categories of the model indicated in the following table.

Table 1. *Main Themes and Categories of a Preliminary Model of Identity for EFL Context*

Theme	Category
Second language acquisition & Social Status	Talking English gives respect and prestige. English language makes one's personality more impressive. English language is essential for promoting communication development in groups. Learning English offer feeling of being appreciated.
Culture Attachment	Iranian students are engaged to strive for continuously improved performance. Iranian people are rewarded for excellent performance. Communicating in English language is considered as value in our society. Persian language is taken superior over foreign languages. Persian history and culture is highly respected.
Persian Language Adhesion	. The spread of English in Iran is a threat to Persian culture and language. . The set of values that comes with English is destroying other cultures. . Mastering 4 skills (speaking, listening, reading, writing) in Persian is priority. . The knowledge of Persian language is necessary for national unity. . Persian poems, novels and written scripts always captivate me.
Pronunciation Posture	. Speaking American English accent is much better than British. . Feeling proud of speaking Persian with English accent is regarded classy.
Technology Involvement & Language Identity	. Computer use empower one's self to achieve his/her language learning goal . By using computer, I am able to understand materials better and be understandable as well . Technology helps me to improve my 'ego' . My self-confidence grows by employing high-tech devices in 2nd language acquisition . To be up-to-date, people must be able to function in English language . I see technology in accordance with reflecting Iranian language learners manifestation (ability)

Second Language Knowledge	. What I learn during English class will be useful to me in future . The English learning really helps me develop as a 'person'. . The 2nd language learning enriches our native language (in this case, Persian) . 2nd language learning add o mutual understanding as a global level. . Additional language knowledge helps the growth of mind. . I would like to learn only spoken form of foreign languages. . Mastering 2nd language grant a high self-esteem to learner. . I can estimate my development of 2nd language acquisition by myself.

Results

Multiple Linear Regression Model

In the regression model, we intend to investigate the influence of the six components of linguistic identity on the scores of students participating in the PBT test. However, in the following table, we analyze the regression model regardless of the intercept on the components.

Table 2. *Model Summary*

Model	R	R Square[a]	Adjusted R Square	Std. Error of the Estimate
1	.994	.989	.988	7.195

a. For regression through the origin (the no-intercept model), R Square measures the proportion of the variability in the dependent variable about the origin explained by the regression. This CANNOT be compared to R Square for models, which include an intercept.

The value of R square (R5) is 0.989, 9.891 of the variation in PBT test scores depends on the identity component, so only about %1 of the test scores are not included in the model. Therefore, the estimated values of the PBT test scores are accurate.

Considering the acquisition of the second language and social state (A%), cultural dependency (A5), Persian language correlation (A3), status and pronunciation state (A2), the relationship between technology and linguistic identity (A2) second language knowledge (A6) and the PBT test score (MARK), the proposed regression model will be as follows:

$MARK = \hat{\beta}5A52 \ \beta3A32 \ \beta2A22 \ \beta6A6$

$MARK = \hat{=} -\%8.022 \ A5358.532 \ A32 \ 23.2\%3A22 \ \%3.923A6$

In the regression model, we intend to investigate whether the six components of linguistic identity influence the scores of students participating in the PBT test?

Since the test score cannot be negative, the constant reported in the model with intercept is as follows:

Table 3. *Constant Reported in the Model*

Coefficientsª

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	-163.580	31.993		-5.113	.000
	Technology Involvement & Language Identity	-22.180	10.747	-.392	-2.064	.041
	Pronunciation Posture	49.229	5.848	.698	8.418	.000
	Persian Language Adhesion	-12.495	5.829	-.205	-2.144	.034
	2nd Language acquisition & social status	24.685	9.029	.275	2.734	.007
	Culture Attachment	-7.250	3.879	-.111	-1.869	.064
	2nd Language Knowledge	26.680	4.735	.608	5.635	.000

a. Dependent Variable: Mark

As noted, although all the components and even the intercept are significantly different at significant levels, since the estimated value of the source is equal to -%63.280 and we do not have a negative score in the test, so the table of coefficients is not correct, theoretically. This means that if the learners disagree entirely with all the components, the test score will be negative, which is practically impossible. Therefore,

Table 4. *ANOVA Test*

ANOVAª,ᵇ

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	530595.315	4	88432.552	1724.868	.000
	Residual	5844.685	116	51.269		
	Total	636440.000ᶜ	120			

a. Dependent Variable: Mark
b. Linear Regression through the origin
d. This total sum of squares is not corrected for the constant because the constant is zero for regression through the origin.

In the analysis of variance (ANOVA), the regression model is significant at the level of 0. %, and the value of SIG is almost equal to zero. Also, the square error is minimal compared to the regression square error. So the residual values in the model are small.

Table 5. *Coefficients of Regression Model*

Coefficientsª,ᵇ

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	Technology Involvement &Language Identity	14.213	8.895	.811	1.598	.113
	Pronunciation Posture	43.513	6.342	2.789	6.861	.000
	Persian Language Adhesion	-28.235	5.468	-1.619	-5.164	.000
	2nd language acquisition &Social Status	-10.126	6.552	-.546	-1.545	.125
	Culture Attachment	-18.057	3.593	-1.183	-5.025	.000
	2nd Language Knowledge	13.943	4.449	.736	3.134	.002

a. Dependent Variable: Mark
b. Linear Regression through the origin

The coefficients table of the regression model shows that the relationship between technology and linguistic identity and second language acquisition and social state are not significant and are excluded from the model. Therefore, the linguistic identity components of the relationship between technology and linguistic identity and second language acquisition and social state do not affect the test score. The rest of the coefficients are significant at the level of 0.0%, and only the second language coefficient is significant at the level of 0.02.

On the other hand, the coefficient of correlation between Persian language and cultural dependency is negative; this means that the increase in the mentioned components decreases the test score of the learner. In other words, the correlation component of the Persian language and dependency had a negative effect on the PBT test score. The other components, namely status, state, pronunciation, and knowledge of the second language, are positive coefficients that will increase with the number of these components. The largest coefficient is related to the component of status and pronunciation; therefore, the status and pronunciation of learners have the most positive effect on the PBT test score. Also, the smallest coefficient is related to the correlation of Persian. It means that the most negative effect on the score is the correlation between Persian and English.

Structural Equation Method (SEM)
Structural equation modeling is a general and very powerful multivariate analysis technique of the multivariate regression family. In other words, it is an extension of the linear model that allows the researcher to analyze a set of regression equations and the relationships between variables simultaneously. The reason for the widespread use and popularity of this technique among researchers is that it overcomes the difficulty of analyzing the relationships between variables in human research in addition to providing a quantitative method to test the theory and contrary to the linear models used in traditional methods, (such as multiple regression), it can also estimate the measurement error. A structural model consists of two parts: • Measurement model: The measurement model defines the relationship between explicit and implicit variables.

●Structural model: The structural model specifies how the factors are interrelated. Generally, structural equation modeling has many applications in various studies due to its multiple capabilities and overcoming the limitations of traditional methods. The general steps that researchers need to follow to implement this approach include:

Study formatting,

●Control over the assumptions of the implementation in the structural equation model,

●Review, coordinate and modify the model,

●Proper interpretation of findings.

Structural equation method (SEM) examines the implicit relationship between dependent and response variables. This method is used when variables are not normally related to each other.
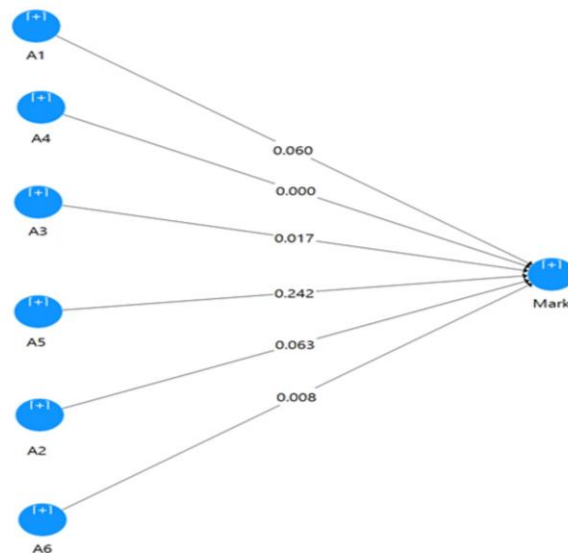

Figure 1. *Structural Equation Model*

Like the regression model, the effect of the technology coefficient factor on the PBT test score is not significant. Nevertheless, the coefficients of second language acquisition, social state, and cultural dependency are significant at the level of 0. %. Also, the effect of the Persian language correlation coefficient on the score of 0.02 and the effect of the coefficients of pronunciation status and second language proficiency on the PBT test score is significant at the level of 0.0%. Accordingly, the pronunciation status had the strongest, and the cultural dependency component had the weakest effect on learners' scores, meaning learners whose status and pronunciation was stronger than other learners scored significantly better on the test scores. On the other hand, although cultural dependency varies among individuals, cultural dependency among learners does not significantly influence their scores because the effect of the technology coefficient factor on PBT test score is not significant, so familiarity and relationship of technology learners do not increase or decrease their score in PBT test.
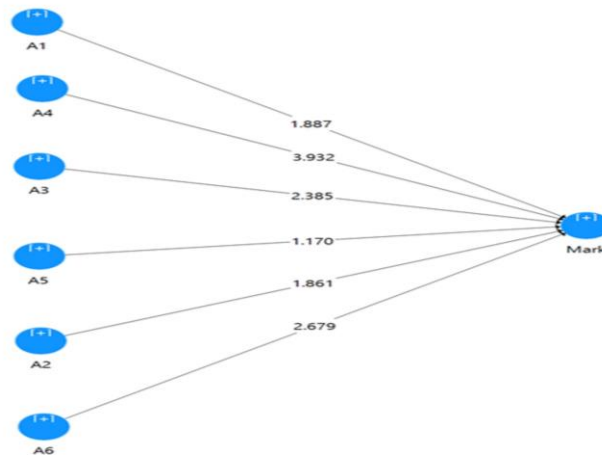
Figure 2. *Structural Equation Model*

Since the T-value in all components except technology is more significant than %.62, the statistical hypothesis of the research confirms the significant relationship between components of linguistic identity and PBT test number. Based on the T-VALUE, the strongest relationship is between the A2 (the status of pronunciation) and the PBT test score, which confirms the P-VALUE values. However, the second language acquisition and social status component had the least weak association with the PBT test score. The T value of the A2 was also lower than the PBT score, not significant, and it is equal to zero. Moreover, also, the effect of second language acquisition, social status, and cultural dependency on the test score is the same, and our hypothesis is as follows:

H0: There is no significant relationship between the two groups.

In all cases, the hypothesis H0 is rejected, meaning that there is a significant relationship between the two groups; this means a significant relationship between all components and test score. The two components of second language acquisition, Social status and cultural dependency at level %, Persian correlation at level 0.02, and the state of pronunciation and second language knowledge, reject the H0 hypothesis at level 0.0%. Now, we try to investigate the effect of component coefficients on the test score.
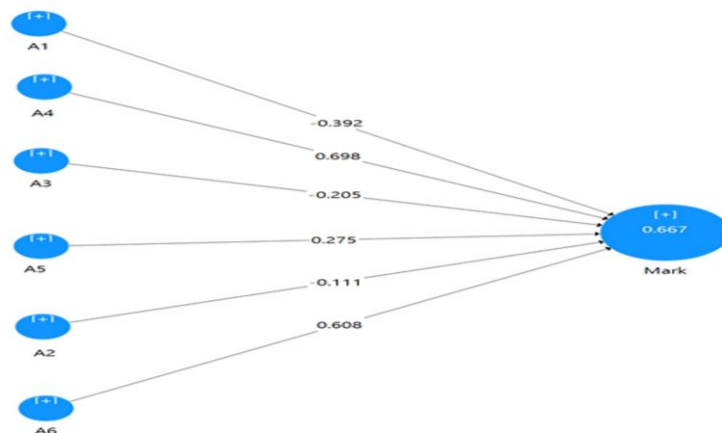

Figure 3. *Structural Equation Model*

Moreover, the structural equation model is as follows:

$$\widehat{MARK} = \hat{\beta}_0 + \hat{\beta}_1 A1 + \hat{\beta}_2 A2 + \hat{\beta}_3 A3 + \hat{\beta}_4 A4 + \hat{\beta}_6 A6$$

$$\widehat{MARK} = 0.667 - 0.392\ A1 - 0.111\ A2 - 0.205\ A3 + 0.698\ A4 + 0.608\ A6$$

Proposed Model

In this research, two methods of multiple linear regression and structural equation were used to test the scores and components. In the regression model, we found that, despite the high value of R^2 and the acceptability of the residuals, two components of the model were omitted, and the effect of the two components of second language acquisition and social state and the relationship between technology and linguistic identity was not significant. Also, in the covariance matrix, the Persian language correlation component was correlated with only two components. Moreover, the structural equation method is used when there is a hidden relationship between the independent variables and response. In this method, the effect of technology and linguistic identity on the model was not significant and was excluded from the model.

On the one hand, there is a hidden relationship between the components, and the PBT test score and two variables have a strong relationship. According to the obtained data, the existence of two components, the existence of a hidden relationship between the components, the structural equation model is the most appropriate and efficient model for this research. Therefore, the following model is selected as the final model:

$MARK = 0.6623 0.395\ A\%30.\ \%\%\%\ A53\ 0.502\ A32\ 0.698\ A220.608\ A6$

Table 6. *Relationship between Identity and Language Achievement Scores*

Factors	Correlation Test	Achievement Score
Second language acquisition & Social Status	Pearson Correlation	0.434
	P-Value	0.000
Culture Attachment	Pearson Correlation	0.576
	P-Value	0.000
Persian Language Adhesion	Pearson Correlation	0.106
	P-Value	0.000
Pronunciation Posture	Pearson Correlation	0.218
	P-Value	0.017
Technology Involvement & Language Identity	Pearson Correlation	0.634
	P-Value	0.249
Second Language Knowledge	Pearson Correlation	0.637
	P-Value	0.000

The last part of this study deals with the correlational phase which was carried out to figure out whether there was any possible relationship between identity and language achievement of EFL learners. To carry out this phase, the Pearson Correlation Coefficient was run. According to table 6, there is a significant relationship between the variables of the study and that is because of P-Value being less than 0.05, except "technology involvement & language identity". Thus, all the factors of identity (except one) have a positive relationship with language achievement scores that means if identity gets improved, language achievement scores will increase and vice versa.

Discussion

In this study, the linguistic components were examined one by one. Second language acquisition creates a special social status for each person and changes the personality of the individuals. Although this component was omitted in the regression model, we found that it had a hidden effect, and in the structural equation model, it had a significant negative coefficient. The people who learn a new language become culturally dependent and may change or transform all or part of their culture, behavior, and habits. The effect of this component on both models was negative, which means that as the cultural dependency increased, the score of the learner decreased. It is clear that this factor reduces PBT test scores. How to pronounce sentences and words in any language is very important, and a person who has just learned a second language can be more successful in the process of having a good accent. Therefore, this component has a very positive effect on the PBT test scores of learners. Many people think that using modern gadgets can improve learning English language, but in the analyzed researches in this study, this was utterly rejected. The use of technology does not help improve language learning and test scores. Like any science in which we need to learn its concepts and terms, language science has a significant role in learning it; PBT is accepted. Therefore, second language knowledge is accepted as an influential component of PBT test scores.

The results of the present study are in line with a study carried out by Pullen (2011) to probe the relationship between identity and perception of pronunciation of non-native speakers of English in an EFL context. The researcher made a comparison between the information related to language background, perception ratings, and the identity and pronunciation scores. Then findings showed that time when learning English starts as well as residence of three or more months abroad had a significant relation with both the identity and pronunciation scores. Moreover, in 2019, Rezaei et al., in a study, conducted a survey of Iranian language teachers in Iran. Based on these data, a paradigmatic identity model is presented. The results of the confirmatory factor analysis showed that this model is suitable and has eight components. These findings also showed that old and experienced teachers have the highest level of Iranian cultural identity. Besides, the results showed that female teachers have a strong Iranian cultural identity compared to their male counterparts, and teachers in different fields of study with different mother tongue and university degrees do not differ much in terms of cultural identity. The results of the present study are not supported by some previous studies. A study has been conducted by Mohammadi et al., in 2018, with the aim of examining the relationship between the cultural and social identity of Iranian learners and EFL learners. Data normality was assessed based on normal Skewness and Kurtosis tests, so the Pearson correlation coefficient was used in this study. The results showed that there was a significant negative relationship between students' learning and their EFL learning (-678). It will be useful for curriculum designers, learners, their parents, and language teachers. Also, Razmjoo (2010) carried out a study to determine the possible effect of identity aspects on the achievement of Iranian learners who were learning EFL at a language

institute in Shiraz. The findings showed no significant relationship between L2 achievement and the aspects of identity. Put it another way, the aspects of identity cannot predict language achievement in the Iranian context. As for the demographic factors, the only gender was found to account for two aspects of identity, i.e., personal and relational identities.

Conclusion

Out of the one hundred and twenty students surveyed, the majority was women. We selected these 120 language learners from 12 provinces of the country, of which learners from Tehran had the most and learners from Tabriz the least. They are between 18 and 23 years old, and most of the participants are young people over 20 years old. Two levels of Advanced and Upper-intermediate has been investigated. Learners' scores on the PBT Test range from 43 to 98. While these scores do not differ much between males and females, these scores were quite different between the two language levels. Although the number of Advanced participants in the test was higher and the Upper_intermediate was lower, their scores were much higher, indicating that the test score was not related to the learner`s level. The ration of language learners in both sexes is the same, indicating that gender has no effect on increasing or decreasing the numbers of partners at the language level. The only available discarded data belongs to the 98 scores in the Upper_intermediate group. After descriptive studies of gender, level, province, and language score of the learner, then the relationship between the components was tested, and most of the components were correlated with each other. However, the Farsi language correlation component was an exception. The existing correlations are all positive, and in one direction; that is, as the components increases, the other component would also increase. Then, the study examined the relationship between the scores of language learners participating in the PBT test with six components of linguistic identity, acquisition of second language and social status, cultural affiliation, Farsi language correlation, status and loss of state, the relationship between technology and linguistic identity and second language knowledge. In the regression model, the value of the detection factor was acceptable and high, and the hypothesis of the suitability of the model was also confirmed. However, it was found that the two components did not affect the test score.

In the structural equation model, the hidden relationship between the components was very high. Only the component of the relationship between technology and identity was left out of the model. The reported P-VALUE and T-VALUE values indicate the existence of the effect and the latent relationship between the independent and variable response variables; therefore, according to the analysis, the component of the relationship between technology and linguistic identity is not considered as one of the components of linguistic identity. This component does not affect the language score of students. The components of second language acquisition and social status, cultural dependence, and Farsi language correlation have the opposite effect on language learners; On the other hand, the status and pronunciation and knowledge of the second language have a significant impact on the students' scores. This means that as the learner becomes culturally independent, or his or her second language becomes Farsi, his or her PBT score decreases. Good accent and pronunciation help learners improve their PBT scores. The more you know about the language you are learning, the easier you will have language learning, so your scores will be better. According to the model of structural equations presented and the effect of the coefficients of each component, learners should focus on the cultural dependence and correlation of their Farsi language without camels; because this causes they try to increase their skills in pronouncing the language and its knowledge.

Pedagogical Implications

Teachers should be informed regarding how identity and foreign language learning are connected. In teacher training programs, different aspects of language teaching should be highlighted, and enough information regarding the direct and indirect way of identity change among students and teachers should be given to the would-be language teachers. It is suggested that teachers' tone should be impartial toward foreign culture, and there should not be any sign of complete approval or disapproval of foreign culture. It is also recommended that in discussing identity in the classroom the purpose should be awareness-raising rather than praising a particular norm or behavior. Some implications that can be directly related to the Iranian context of language teaching are listed below:

●Language teachers need to have adequate knowledge of identity and language learning. They need to know precisely what aspect of identity would affect the learners' motivation in learning a foreign language positively or negatively.

●Language teachers also need to be equipped with techniques and strategies to deal with challenges of conflict between identity and foreign language learning.

●It is recommended that the issue of identity be part of the whole language teaching curriculum. The language teaching curriculum should be sensitive to identity issues. Such sensitiveness would include sensitive materials and also sensitive methodology to different aspects of language learning.

References

Ary, D., Jacobs, L. C., & Razavieh, A. (2006). *Introduction to research in education* California: Thomson Thomson.

Beijaard, D. (1995). Teacher`s Perior Experiences and Actual Perceptions of Professional Identity. *Teachers and Teaching,1* (2), 281-294.

Beijaard, D., Meijer, P. C., & Verloop, N. (2004). Reconsidering research on teachers' professional identity. *Teaching and Teacher Education, 20(2)*, 107–128.

Hamayan, E. V., & Domico, J. S. (1991). Developing and Using a Second Language. In E. V. Hamayan & J. S. Damico (Eds.), *Limiting bias in the assessment of bilingual Children.* 1, 39-75.

Hinkel, E. (1999). *Culture in Second Language Teaching and Learning*. Cambridge: Cambridge University Press.

Kasper, G. (2006). 'Beyond repair: Conversation analysis as an approach to SLA', *AILA Review, 19*, 83-99.

Kasper, G. and Dahl, M. (1991). 'Research methods in interlanguage pragmatics', *Studies in Second Language Acquisition, 13*, 215-247.

Kasper, G. and Rose, K. (1999). 'Pragmatics and SLA', *Annual Review of Applied Linguistics, 19,* 81-104.

Kasper, G. and Rose, K. (2003). *Pragmatic Development in a Second Language.* (*Language Learning Monograph Series, 3*. 12-125.

Kasper, G. (2006) 'Beyond repair: Conversation analysis as an approach to SLA', AILA Review,
19: 83-99.

Kasper, G. and Dahl, M. (1991) 'Research methods in interlanguage pragmatics', Studies in Second Language Acquisition, 13: 215-247.

Kasper, G. (2006) 'Beyond repair: Conversation analysis as an approach to SLA', AILA Review, 19: 83-99.

Kasper, G. and Dahl, M. (1991) 'Research methods in interlanguage pragmatics', Studies in Second Language Acquisition, 13: 215-247.

Kasper, G. (2006) 'Beyond repair: Conversation analysis as an approach to SLA', AILA Review, 19: 83-99.

Kasper, G. and Dahl, M. (1991) 'Research methods in interlanguage pragmatics', Studies in Second Language Acquisition, 13: 215-247.

Kasper, G. (2006) 'Beyond repair: Conversation analysis as an approach to SLA', AILA Review, 19: 83-99.

Kasper, G. and Dahl, M. (1991) 'Research methods in interlanguage pragmatics', Studies in Second Language Acquisition, 13: 215-247.

Khatib, S. (2011). Applying the reader-response approach in teaching English short stories to EFL students. *Journal of Language Teaching and Research, 2* (1), 151-159.

Kinginger, C. (2008) 'Language learning in study abroad: Case studies of Americans in France', *The Modern Language Journal,* 92, 1-124.

Maykut, P., & Morehouse, R. (1994). *Beginning Qualitative Research: A philosophical and practical guide*. London: The Flamer Press.

McMillan, J. H., & Schumacher, S. (2001). *Research in education: A conceptual introduction* (5th Ed.). New York: Longman.

Mohammadi, H & Izadpanah, S. (2019). A Study of the Relationship between Iranian Learners' Sociocultural Identity and English as a Foreign Language (EFL) Learning Proficiency. *International Journal of Instruction, 12*, 53-68.

Pavlenko, A., & Lantolf, J. (2000). *sociocultural theory and second language learning.* New York: Oxford University Press.

Pullen, E. (2011). *The relationship between identity and pronunciation of non-native speakers of English in an EFL setting.* Thesis M.A. (English). The Program of Teaching English as a Foreign Language Bilkent University, Ankara.

Riley, P. (2007). *Language, Culture, and Identity*. London: Continuum.

Razmjoo, S. A. (2010). Language and identity in the Iranian context: The impact of identity aspects on EFL learners' achievement. *The Journal of Teaching Language Skills (JTLS), 2*, 99-121.

Razmjoo, S.A., & Izadpanah, M. A., (2012). On the Relationship between L2 Literacy (Reading   and Writing) and Identity Processing Styles of Iranian Advanced EFL Learners. *Journal of Research in Applied Linguistics, 3* (2), 2-22.

Razmjoo, S. A., & Mavaddat, R. (2015). On the relationship between justice judgments, outcomes, and identity orientations among Iranian EFL learners: A structural equation   model. *Iranian Journal of Language Teaching Research, 3*, 101-117.

Rezaei, S. & Bahrami, A. (2019). Cultural Identity among Iranian English Language Teachers. *International Journal of Society, Culture and Language, 1*, 67-82.

Richards J. C., & Farrell, T. S. C. (2005*). Professional Development for Language Teachers: Strategies for Teacher Learning*. Cambridge: Cambridge University Press.

Richards, J. C., & Rodgers, T. S. (2014). *Approaches and Methods in Language Teaching.* Cambridge: Cambridge University Press, New Edition.

Richards, J. C., & Rodgers, T. S. (2001). *Approaches and methods in language teaching.* Cambridge: Cambridge University Press.

Richards, J. C., and Schmidt, R. (2010). *Longman dictionary of language teaching and applied linguistics*. 3rd ed. Essex, UK: Pearson Education.

Richards J. C., Platt J., and Weber H. (1985). *Longman Dictionary of Applied Linguistics London.* Harlow U.K: Longman Addison- Wesley, 1986, p.289.

Ross, S. J. and Kasper, G. (2013) *Assessing Second Language Pragmatics.* Basingstoke: Palgrave MacMillan.

Sabatin, I. M. (2013). The Effect of Cultural Background Knowledge on Learning English Language. *International Journal of Science Culture and Sport, 1*(4), 22-32.

Schwartz, S. J. (2005). A new identity for identity research: Recommendations for expanding and refocusing the identity literature. *Journal of Adolescent Research, 20*, 293-308.

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Thousand Oaks, CA: Sage.