



Iranian EFL Learners' Familiarity with Reduced Forms in Spoken English based on their Proficiency Level

Rouhollah Zarei

English Department, Yasouj University, Yasouj, Iran

E-mail: rzarei@yu.ac.ir

Mitra Doostkam (Corresponding author)

Yasouj University, Iran

E-mail: mdoostkam90@yahoo.com

Received: 18-05-2014

Accepted: 01-07-2014

Published: 01-11-2014

doi:10.7575/aiac.ijalel.v.3n.6p.68

URL: <http://dx.doi.org/10.7575/aiac.ijalel.v.3n.6p.68>

Abstract

Reduced forms are thought to have a significant influence on understanding spoken language. They are important aspects of spoken language which have sometimes been undervalued in material development, teaching, and testing. This problem is further aggravated in an Iranian EFL setting. In an attempt to address this issue, this study set out to examine the familiarity of Iranian EFL learners with reduced forms as well as the relationship between the proficiency level of language learners and their familiarity with reduced forms. To this end, a test of RFs on listening comprehension was developed based on the literature and findings of a pilot study. Afterward, 306 English language learners were selected from two Iranian cities of Shiraz and Ahwaz based on multi-stage cluster sampling to participate in the study. The results of this study indicate that Iranian EFL learners are not familiar with RFs in spoken language. However, in contrast to earlier findings, there was a significant difference between the performance of advanced and intermediate learners based on their familiarity with reduced forms.

Keywords: Iranian EFL Learners, Reduced Forms, Spoken English, Proficiency Level

1. Introduction

Students do not recognize words they know while listening (Chen, 2002; Goh, 2000). Moreover, Sun (2002) stated that students are not able to segment speech, which makes listening difficult. Rosa (2002) states that language learners develop their listening and speaking skills based on this adapted English speaking style. On the contrary, when they face natural spoken language, speak with native speakers or listen to real-life English in movies and news, they discover that native speakers seem to speak fast while EFL students are unable to recognize word boundaries, words, and phrases.

Most of Iranian language teachers are not native English speakers and their emphasis is mainly on teaching grammar and vocabulary knowledge. Therefore, they frequently try to use articulated language to convey information more completely whereas native speakers speak in a way in which words seem joined together.

Even advanced EFL students have problem in interpreting natural spoken language since they have learnt their English through their eyes. EFL learners try to understand each sound precisely while native speaker instead of focusing on their output, focus on the message. Therefore, according to the above mentioned reasons, the present researcher intended to examine the familiarity of Iranian EFL learners with reduced forms, and the effect of their proficiency level on their familiarity with these forms in listening comprehension. Thus, the following research questions were posited in this study:

1. How familiar are Iranian EFL learners with reduced forms in spoken English?
2. What is the effect of EFL learners' proficiency level on their familiarity with reduced forms in listening comprehension?

The present study is significant in that it is one of the first attempts to identify learners' familiarity with reduced forms, the effect of their proficiency level on their familiarity with these forms.

2. Method

This study applied a quantitative method to examine the familiarity of EFL learners with reduced forms and the effect of their proficiency level on recognizing these forms in listening comprehension. The main instrument in this study was a specific listening test on reduced forms. To construct the test, both ideas extracted from the literature and experts were incorporated into the first draft. The test aimed to investigate the effect of learners' proficiency level on their familiarity with reduced forms in listening comprehension.

It is worth mentioning that the first step in designing a test is construct formation, i.e. defining a construct and its components, through capitalizing on relevant theories and empirical findings (Cronbach, 1971). In addition, the framework of a test must represent relevant components and exclude irrelevant components (Fulcher & Davidson,

2007), and the theoretical and practical grounds of the test determine test components. Hence five components of reduced forms, namely schwa, contraction, elision, linking, and assimilation built up the assessment framework of the listening test. These components were extracted from Rosa (2002) and Brown (2006).

The present researchers contented themselves with 53 out of the 72 available items due to time limitation of language classrooms. The materials applied in this study were chosen from a wide variety of sources including interviews, movies, series, talks, etc. in which instances of 5 types of reduced forms could be found. Since in an EFL context like Iran, having limited access to native speakers, the present researcher tried to use materials which were representative of natural or pseudo-natural spoken language.

According to literature on listening comprehension, it is of crucial importance that listening materials be representative of genuine speech and they should have the characteristics of the target language (Buck, 2001; Hughes, 1989). There are three reasons for choosing listening materials items: First, they must be representative of natural spoken language; second, the items were geared to the proficiency level of the participants; third, the test started with recognition items to assist the participants in overcoming psychological inertia which was an artifact of the administration conditions. It also lowers the anxiety level of the participants and ameliorates their performance on the test (Krashen & Terrell, 1983).

As already mentioned the original test was lengthy and took 50 minutes to complete. Examinees were unable to answer questions because it imposed a heavy burden on their short term memory. Furthermore, the test became unmanageable due to time constraints put by language institutes and memory limitation because of cloze test with every 7 items deleted. Besides, item analysis was done to ensure the validity and reliability of the measure.

It is worth noting that the test underwent several pilot studies, but the last pilot study was mentioned here. For example, Part E of the preliminary test composed 3 cloze tests which, based on the opinion of the experts, underwent some alterations. They opined that the cloze test format was not appropriate for listening test since it imposed a great cognitive burden on the short term memory of the listeners. On the other hand, the passage with approximately 220-250 words including every 7th words deleted was too long to be included in a listening test for each talk.

On the other hand, contrary to Nation and Newton (2008), if every 15th word is deleted, it may raise the possibility that the deleted words do not measure what the test purports to be measured and put the validity of the test under question. Therefore, the present researchers modified the last three talks, and the form changed to fill in the blank, and the deletions was determined based on the purpose of the test and the result of the pilot study.

According to literature on the effect of short term memory and listening comprehension, understanding spoken language in L2 requires more mental energy and time in comparison to first language (Magiste, 1979; Marsh & Maki, 1978). Accordingly, as the mental processes work less effectively for a second language, the listeners encounter "cognitive deficit" in understanding spoken language (Cook, 1991, p. 71).

Therefore, the load for learners to comprehend a message will be intensified. Further, short term memory for understanding aural stimuli in L2 becomes overloaded, and "it causes words to be purged before they can be organized in L2 patterns and then interpreted" (Rivers & Temperly, 1978 as cited in Ohata, 2006, p. 22). Therefore, the listeners will not able to remember all the information. Thus the current researcher modified Part C and E consisting 3 excerpts from different talks.

Hence the lengths of the 3 excerpts of Part C and E were determined to be between the ranges of 1.26 to 26.15 to ensure its face validity (Flowerdew & Tauroza, 1995; Schmidt-Rinehart, 1994). Moreover, the present researcher tried to adopt items between the ranges of 125 to 250 wpm, i.e. words per minute which is applied for measuring speech rate differed in increments of 25 wpm. According to Foulke (1968), speech rate between the ranges of 125 to 250 wpm does not affect listening comprehension. Therefore, all test items were extended in this range.

Additionally, accent is considered as an important variable in assessing listening comprehension. In addition to the validity of the test, it affects acceptability, intelligibility, and measurement issues of test among stakeholders (Harding, 2011, p. 22). However, the decision on selecting the accent depends on the purpose and the context in which it will be applied (Buck, 2001, p. 162). He also adds:

The choice of accent depends on the purpose of the test, and the context in which it is used... To a considerable extent, it depends on what test users are prepared to accept. From a measurement perspective, the most important thing when using an unfamiliar accent is that it is equally unfamiliar to all. (Buck, 2001, p.162)

Harding (2011) reiterates that applying an accent which is familiar to some but unfamiliar to others acts as a threat to the construct validity of a test and there is a possibility for test bias. It is crucial for a test to choose a homogenous accent for a test since applying a diverse accent raises the possibility that some test-takers to be advantaged over the others on condition that they are very familiar with that accent, while those test-takers, who are not familiar with the topic, face difficulty in understanding the test. Therefore, the present researcher adopted "orthodox approach" to speaker accent in the listening test in which speaker accents were limited to inner-circle varieties of English language (Harding, 2011, p. 20). The inner-circle contains the native English speaking countries such as the United Kingdom, the United States of America, and Canada (Kachru, 1985).

On the other hand, in most language institutes in Shiraz and Ahwaz, the American accent is applied. In other words, the textbook they studied was based on American accent including Interchange, Summit, American Files, etc. In addition, the students' accent and most of their teachers' was American. Therefore, the present researcher chose American accent in listening assessment. Besides, one American native speaker was recruited to tell the instructions. The native speaker was a male TEFL expert, an MA holder, and worked as a teacher in language institutes and he was an IELTS examiner.

The instructions of the test were said by him in a sound-insulated room in a recording studio under supervision of the present researcher.

The length of pauses between items is of crucial importance in processing spoken language. In testing listening comprehension, it is important to make a choice on the whole time of the test and time intervals between items. Several studies have emphasized that it would be helpful for examinees to give time to preview the stems of the question, for it helps examinees exploit metacognitive strategies through planning as well as turning their attention to the related parts of the text (Buck, 1995; Littlewood, 1981; Mendelsohn, 1995; Thompson, 1995; Ur, 1984; Vandergift, 1999). Previewing test items improves the performance of examinees in tests of listening comprehension (Yanagawa & Green, 2008).

On the other hand, the time intervals between items must be appropriate, neither too short nor too long. If the time interval between items is too long, it will bring about communication breakdown (Phillips, 1985). On the contrary, if the time interval is short, it will be difficult for the listener to retain the information in their mind. The listening test on reduced forms contained four parts. Prior to answering items in each part, examinees were given 10-second pause to have a look at all the items. In part A of the listening test on reduced forms, 8-second pause was inserted between each item. The time interval in this part of the test was determined based on the results obtained from the pilot study.

In part C, since the passages were long and they were in the form of recognition items, the time interval was determined on the basis of the results of pilot study. Therefore, 20 second pauses inserted after Talk One and Two, but 30-second pause were embedded after Talk Three.

On the other hand, in Part D, the "second-per-letter" method was adopted for each blank in items 34 to 53 in which the length of pauses between items containing blanks was determined based on the deleted letters in each blank. For example, a 16-second pause was allotted for an item in which 16 letters were deleted (Zahedi, 1997, p. 34). Since in this part the examinees were required to write full forms in the blanks, the letters of the words in their full forms were considered in this method. This method is more effective than "spell out letters" proposed by Oller (1979) in which the length of a passage is as much as the deleted portion when it is spelled out twice. Oller (1979)'s "spell out the letters" method was too subjective to apply because the length of each pause would be considerably influenced by the speakers' speed of speech. Table 1 demonstrates the length of time allocated to each blank.

Table 1. The Time Assigned to each Blank in Second-per-letter Approach

Item	Second-per-letter approach
34	9
35	11
36	8
37	8
38	7

In addition, in Part E, the "second-per-letter" method was adopted for each blank in items 39 to 45. Contrary to Part D, Part E consisted of three separate talks, but the length of pauses was determined on the basis of the letters of the whole blanks in each talk. Table 2 illustrates the length of time assigned to each blank.

Table 2. The Time Assigned to Each Blank in Second-per-letter Approach

	Items	Second-per-letter approach
Talk 1	39	10
	40	7
Talk 2	41	6
	42	13
	43	10
Talk 3	44	7
	45	16
	46	10
Talk 4	47	16
	48	6
	49	11
	50	8
	51	10
	52	13
	53	12

Items 1 to 33 were recognition items, hence the scoring procedures of these items were binary, i.e., either correct or incorrect. The scoring procedures of the items 34 to 53 are also binary as well. Consequently, the examinee gets no credit for one word as such because it would not satisfy the blank. In addition, examinees were not penalized for incorrect responses. As a consequence, one score was assigned for the correct answer to each item.

2.1 Initial Piloting

In order to develop the test, the present researcher drew on extant literature on teaching reduced forms (e.g., Brown & Hilferty, 1986; Bowen, 1976; Norris, 1993 & 1995; Weinstein, 1982), assessing connected speech (Brown & Kondo-Brown, 2006; Celce-Murcia et al., 1996), etc.

The designing of the test of reduced forms was conducted from November 18, 2012 to May 21, 2013. The test was presented to a panel of experts, and they were asked to give their feedback and suggestions on the test of reduced forms in general and its constituent items in particular. The panel composed of 3 university professors of English teaching and testing, 2 male and 1 female. The test was scrutinized by expert panelists for its face and content validity. They commented on the general appearance of the items, the clarity of instruction as well as the suitability of the test items.

As a consequence, some items were modified or discarded from the test. For example, according to expert panelists a part was added at the beginning of each talk in Part E of the test which served as a foundation for examinees to better understand the texts. Moreover, some items were deleted on the basis of experts' opinions, for they measure more than one aspect of reduced forms in one item simultaneously. Therefore, the present researcher eliminated those items from the test since they would act as a threat to validity of the test. In this stage, the rationale for retaining or eliminating an item is conceptual rather than statistical.

Then final piloting was undertaken to delineate the procedures adopted for conducting the study, and to examine the psychometric characteristics of the measure including validity and reliability. In this stage, the present researcher also investigated the clarity of instructions, scrutinized whether the time allocated to each part is sufficient for the examinees to answer each item, and examined other administration procedures. Furthermore, language learners were asked to express their opinion about any item not interesting to them.

Ninety seven language learners having the same representativeness with the target participants were recruited based on convenient sampling to participate in the pilot study from two cities under study, 45 of whom were advanced language learners and 52 of whom were intermediate language learners from Shiraz and Ahwaz language institutes.

Contrary to the initial piloting, the reason for including or excluding items was statistical with the aim of boosting the reliability and validity of the test. Before the test administration, the participants were provided with a brief explanation about the purpose of the test along with the way participants were required to answer questions in each part. Pilot study was also undertaken to determine the time allotment to each item or each part as well. The whole time for completing the test was 14 minutes and 13 seconds. The result of the pilot study demonstrated that 96% of participants finished the test during the allotted time to each item.

On the other hand, the decision to either discard or retain items rested on their IF. In other words, items with IF ranged from 0.3 to 0.8 were kept, for they were perceived as appropriate items of the test utilizing the whole construct (Brown, 1989; Fulcher, 1997). IF of the test items is demonstrated in Table 3.

Table 3. Item Analysis

Item No.	No. of correct responses	IF	Item No.	No. of correct responses	IF
1	36	0.37	19	33	0.34
2	34	0.35	20	12	0.12
3	33	0.34	21	51	0.53
4	19	0.19	22	48	0.5
5	14	0.14	23	31	0.32
6	41	0.42	24	28	0.29
7	45	0.46	25	35	0.36
8	30	0.31	26	29	0.3
9	21	0.22	27	29	0.3
10	53	0.55	28	49	0.51
11	62	0.64	29	22	0.23
12	51	0.53	30	18	0.19
13	23	0.24	31	29	0.3
14	38	0.39	32	34	0.35
15	31	0.32	33	17	0.17
16	29	0.3	34	45	0.46
17	32	0.33	35	30	0.31
18	30	0.31	36	34	0.35
			37	39	0.4

N=97

As a result of item analysis 8 items were discarded from the test including items 4, 5, 9, 13, 20, 29, 30, and 33. The discarded item, i.e. 4, 30, and 33, were related to the schwa section, and items 5, 9, 13, and 20 to contraction, elision, linking, and assimilation sections respectively. On the other hand, item 29 was a production item measuring contraction and assimilation simultaneously. It seemed that they were too difficult for language learners to answer. Although item analysis statistics such as IF is basically applied in norm-referenced tests, they can also be exploited in low- and medium-stake criterion-referenced tests with a fair degree of confidence (Shrock & Coscarelli, 2007, p. 256).

Therefore, the final test contained 29 items of which 17 were recognition and 12 were production items. Table 4 illustrates the sections included in the test and items allotted to each section (see Appendix).

Table 4. Sections and item numbers in listening test on reduced forms

Schwa	1, 2, 3, 18, 25
Contraction	4, 5, 6, 16, 17, 19, 29
Elision	7, 8, 9, 20, 23, 27
Linking	10, 11, 12, 21, 24
Assimilation	13, 14, 15, 22, 26, 28

Part A composed 15 recognition items, and the examinees were required to underline the word or phrase they had heard. The recognition items were bolded so the examinees' attention would not be distracted, or they mistakenly underline other parts of sentence. Part B was made up of 2 recognition items. These items were different from items in Part A in a way that they were contextualized. Accordingly, 5 production items in the form of fill in the blank were incorporated in Part C; on the other hand, two talks were included in Part D with blanks which were required to be filled by examinees.

To examine the validity of the test of reduced forms, PBT TOEFL test of listening comprehension was administered. Consequently, counter balanced test design was adopted to minimize ordering effect on performance of examinees on the test. The participants were divided into two halves. The first half took test of reduced forms; then, they took the PBT TOEFL listening test. At the same time the administration procedure of the test for the second half of examinees was permuted; in other words, they took the PBT TOEFL listening test first; afterward, they took the test of reduced forms. Then correlation was calculated between the two tests (0.809). Since the PBT TOEFL listening test composed of only recognition items, the same process was undertaken with IELTS listening comprehension which consisted of both production and recognition items.

The correlation between these two tests was calculated. There was a significant positive correlation between these two tests (0.798). On the other hand, the reliability of the test was calculated through Squared-error loss agreement index (Bachman, 1990) as it was a criterion-referenced test, and the reliability of 0.9799 was accrued. In conclusion, the result of the pilot study indicated that the test had appropriate validity and reliability to be exploited in the main phase of the study. Furthermore, Angoff method was adopted to determine the cut-off score for the purpose of identifying if Iranian EFL learners were generally familiar with reduced forms in spoken language.

To achieve consistent ratings based on Angoff method, 5 to 10 judges are needed to set the cut-off score (Sanju, Haque, & Oyebo, 2006). As a result, a panel of 5 TEFL experts was recruited based on purposive sampling to serve as judges and set cut-off score for the test of reduced forms in spoken language. The rationale for adopting purposive sampling procedure was that they must have a set of predefined features (Ary, Jacobs, & Razavieh, 1990), i.e. the university teachers should have knowledge of language testing, and they must also be TEFL experts. Afterward, the nature of the study was expounded to these expert judges. They were ensured that their answers would remain confidential and anonymous. In addition, their participation in the study was voluntary. The demographic information of the judges is illustrated in Table 5.

Table 5. Demographic information on the judges

	Frequencies	Percentages
Gender		
Female	2	40%
Male	3	60%
Education level		
Ph.D.	5	100%
Teaching experience		
6-10 years	1	20%
More than 10 years	4	80%
Age		
25-34	1	20%
35-44	2	40%
45-54	2	40%

Their age ranged from 25 to 54, and the majority of the judges were male (60%). Their teaching experiences ranged from 6 to more than 10 years; moreover, all of them had expertise in testing.

The panel of judges was completely briefed about the objective of the test of reduced forms, conceptualizing the definition of minimally qualified students, and procedures for rating test items based on Angoff method. All the judges were required to have the content knowledge of the test they were going to make decision about.

It is worth mentioning that, expert judges were given the test items along with a Test Item Rating adopted from Wheaton and Parry (2012). They were required to rate each item individually and independently and to give their ratings as percentages utilizing increments of five, such as 60, 90, 75. Their ratings must not exceed 95% or be less than 25%. Prior to rating, the definition of "minimally qualified performer, i.e. one who has least amount of education and experience to perform the task" was clarified for the panel of the expert judges (Wheaton & Parry, 2012, p. 5). Accordingly, they must rate each item by predicting how a minimally competent performer would answer each item. Then the Average Percentage Correct was computed, and the average score of 67 were obtained.

2.2 Main study

Once the test of reduced forms in listening comprehension was validated, it was employed in examining the familiarity of Iranian EFL learners with reduced forms. Moreover, it was exploited to investigate the effect of Iranian EFL learners' proficiency level on their familiarity with reduced forms in listening comprehension.

The participants were recruited from language institutes across Shiraz and Ahwaz. The initial sample consisted of 306 participants based on Cochran formula, 178 female and 128 male, with a mean age of 26 years at the intermediate and advanced levels who were categorized on the basis of those institutes placement tests. The number of participants in each proficiency group and the number of tests they took are illustrated in Table 6.

Table 6. Number of participants in each proficiency group and number of tests for each group

Proficiency level	Number	Measure
Intermediate	150	1
Advanced	156	1

In this study, cluster sampling was utilized in which the selected unit instead of an individual is a group of individuals who are naturally occurring together (Ary, et al., 2010, p. 154). Therefore, levels 5, 6, 7, and 8 were determined as the intermediate proficiency level and levels 9, 10, 11, and 12 as the advanced proficiency levels. The sample size in this study was based on statistical population (about 1500 people). According to the Cochran's formula, given the number of language learners in each city, the calculation is equal to 306. Two hundred and six respondents from Shiraz and 100 respondents from Ahwaz are selected. The multi-stage cluster sampling was used for sampling. Hence all language institutes in cities of Shiraz and Ahwaz were identified in different areas, and the clusters of institutes in different areas of both cities were selected. Then individuals were randomly selected from these institutes.

3. Results and Discussion

In this phase of the study, Angoff method was adopted to set the cut-off score for the purpose of identifying if Iranian EFL learners were generally familiar with reduced forms in spoken language. As was mentioned earlier, the expert judges were required to rate each item individually and independently and to give their ratings as percentages utilizing increments of five, such as 60, 90, 75. Their ratings must not exceed 95% or be less than 25%. Prior to rating, the definition of "minimally qualified performer, i.e. one who has least amount of education and experience to perform the task" was clarified for the panel of the expert judges (Wheaton & Parry, 2012, p. 5). Accordingly, they must rate each item by predicting how a minimally competent performer would answer each item.

Afterward, the standard deviation for each item was computed, and items having a standard deviation above 10 must be revised. Since the piloting had been undertaken previously, it can be observed that the standard deviation of all the items were lower than 10. Then the Average Percentage Correct was calculated, and the average score of 67.27273 was obtained. This figure, then, rounded to specify the cut-off score of 67. Three hundred and six students took the test; afterward, the cut-off score was determined. The reliability of the test items, at first, was assessed using Cronbach's alpha which was 0.726 which is rounded and the reliability of 0.73 was accrued. However, the test was criterion-referenced in nature, Squared-error loss agreement index should be computed, and the reliability of 0.9799 was obtained which was rounded and the figure 0.98 was accrued.

Table 7. Reliability of RFs Test

Number of cases	N of Items	Squared-error loss agreement index
306	44	0.98

Since the Squared-error loss agreement index for computing reliability of the test items was 0.98, the test had a high

reliability. Furthermore, Pearson product moment correlation coefficient was exploited to measure the reliability of the judges' ratings, as shown in Table 8.

Table 8. Pearson product moment correlation coefficient among Judges' ratings

		Expert 1	Expert 2	Expert3	Expert4	Expert5
Expert 1	Correlation	1	0.714	0.863	0.751	0.692
	sig		0.000	0.000	0.000	0.000
Expert 2	Correlation		1	0.800	0.807	0.655
	sig			0.000	0.000	0.000
Expert 3	Correlation			1	0.824	0.719
	sig				0.000	0.000
Expert 4	Correlation				1	0.762
	sig					0.000
Expert 5	Correlation					1
	sig					

Therefore, reliability among judges' ratings was 0.939 ($p=0.000$), indicating a significant, positive relationship among judges' ratings. This figure was rounded and the inter-rater reliability of 0.94 (Table 9) was obtained to demonstrate agreement among judges.

Table 9. Cronbach's alpha (Coefficient Alpha) of Judges' Ratings

Number of Cases	No. of Judges	Cronbach's Alpha
306	5	0.939

The results, as shown in Table 11, indicates whether Iranian EFL learners are familiar with reduced forms applying one-sample t-test. The mean of the judges' ratings to each question for setting cut-off score was considered 30, and it was compared with the mean of examinees' scores to show whether they are familiar with reduced forms in spoken language. The mean obtained from examinees' scores was 18.76 which is less than the supposed measure of 30. The magnitude of differences demonstrated that there was a significant difference ($t=-27.42$, $p=0.000$) between the mean of the judges' ratings and the mean of the examinees' scores. Therefore, Iranian EFL learners are not generally familiar with reduced forms in spoken language.

This study set out with the aim of determining the familiarity of Iranian EFL learners with reduced forms in spoken language. A strong relationship between reduced forms and listening comprehension has been reported in the literature (e.g., Brown & Hilferty, 1986; Ernestus, et al., 2002; Ito, 2001; Rosa, 2002; Weinstein, 2001). It is worth mentioning that this finding corroborates previous research into reduced forms in spoken language which found that Iranian EFL learners have little familiarity with reduced forms.

Table 10. Familiarity of Iranian EFL Learners with RFs on SL

variable	Test Value = 30					
	N	Mean	Mean Difference	Std. Deviation	t	Sig
Score	306	18.76	-11.23	5.85	-27.42	0.000

On the other hand, the familiarity of advanced EFL learners with reduced forms was computed through one-sample t-test illustrated in Table 11. As mentioned earlier, the mean of the judges' ratings was 30; then, this mean was compared with the mean of the advanced level examinees' scores. The mean of the advanced EFL learners was 20.45 which was significantly different from the judges' average rating ($p=0.000$; $t=-17.71$).

Table 11. Familiarity of Advanced Iranian EFL Leaners with RFs on SL

variable	Test Value = 30					
	N	Mean	Mean Difference	Std. Deviation	t	Sig
score	150	20.45	-9.55	5.39	-17.71	0.000

The following table (Table 12) illustrates the result of the familiarity of Iranian intermediate learners with reduced forms in spoken language through one-sample t-test. The mean of intermediate level students was 17.14. The obtained mean for intermediate learners (17.14) indicates that it is less than the given measure (30). With respect to the amount of $t=-22.43$; $p=0.000$, this difference was significant. Hence Iranian intermediate learners were not familiar with reduced forms in spoken language.

Table 12. Familiarity of Iranian Intermediate Learners with RFs in SL

variable	N	Mean	Test Value = 30			
			Mean Difference	Std. Deviation	t	Sig
score	156	17.14	-12.85	5.84	-22.43	0.000

An examination of the results of the test on reduced forms indicated two outliers. In comparison to other participants, two participants got unexpected scores of 3 and 28, respectively. Consequently, these participants were excluded from the study. Therefore, the final sample of the study includes 306 participants. Table 13 reports the descriptive statistics of the listening test on reduced forms.

Table 13. Basic statistics for the scores of participants

Measured variable	Mean	Std. error	Std	Skewness	Std. error	Kurtosis	Std. error
Score	18.7663	.40969	5.85159	.348	.170	-.324	.339

No. of cases=306

Results of the test of reduced forms demonstrate maximum variation among the participants. Since this aspect of listening has been neglected in English classrooms, participants were not familiar with reduced forms completely; therefore, in spite of all modifications made in the test content, it was still too difficult for the participants to answer.

Furthermore, independent-samples t-test was conducted to compare the scores of advanced and intermediate Iranian EFL learners' familiarity with reduced forms. There was a significant difference between the scores of advanced ($M=20.4500$, $SD=5.39243$) and intermediate level EFL learners ($M=17.1474$, $SD=5.84272$); $t(201.666) = 4.19$, $p=0.000$ (two-tailed). The magnitude of the differences between the means (mean difference= 3.30260 , 95% CI: 1.75122 to 4.85397) was moderate ($\eta^2=0.079$).

Table 14. Group Statistics

	Level	N	Mean	Std. Deviation	Std. Error Mean
Score	advanced	150	20.4500	5.39243	.53924
	intermediate	156	17.1474	5.84272	.57293

Table 15. Independent-samples t-test

	Levene's Test for Equality of Variances		t-test for Equality of Means							
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference		
								Lower	Upper	
score	Equal variances assumed	.783	.377	4.191	202	.000	3.30260	.78802	1.74879	4.85640
	Equal variances not assumed			4.198	201.666	.000	3.30260	.78678	1.75122	4.85397

The results, as shown in Table 15, indicated that Iranian EFL learners were not familiar with reduced forms in spoken language. However, the findings of this part of the study do not support the previous research. This result differs from

Henichsen's 1984 study in which the intelligibility of input was examined in the presence of connected speech by applying the Sandhi-Variation Exercise demonstrating that there was no significant difference between the performance of low and high ESL learners. In contrast to earlier findings, the result of the study indicated that there is a significant difference between the performance of advanced and intermediate learners concerning reduced forms. However, both of these groups have little familiarity with reduced forms in spoken language.

4. Conclusion

The results indicated that Iranian EFL Intermediate and advanced learners are not familiar with reduced forms in spoken language in general. These forms were generally undervalued in language classes; correspondingly, they are not covered in the listening sections of textbooks appropriately. They are not incorporated in teacher training courses too, for it is expected that they can be learned by osmosis. Therefore, unlike their significant influence on understanding listening comprehension as well as natural spoken language, they have not received the attention they deserve.

Although it is claimed that listening comprehension is not considered a passive skill any more, the current situation in language classes shows that all the evidence is to the contrary. Further, listening is taught through repetition with the emphasis mostly on top-down processing which requires listeners infer or guess the meaning, and the most of class time devoted to teaching grammar and vocabulary. Therefore, it can be inferred that in Iranian EFL context reduced forms can be viewed as the Cinderella aspect of a Cinderella skill.

Designing audio materials containing reduced forms can contribute language learners to develop their listening comprehension. On the other hand, due to the importance attached to assessing listening comprehension, the findings of the study would lend support to designing listening test, especially on reduced forms.

One of the issues that emerges from these findings is instructions language learners receive concerning reduced forms which are limited to what is presented to them in the grammar sections of the language books, for example in the form of contracted forms, or in pronunciation parts which are addressed only briefly and superficially.

Although importance has been attached to integrate both bottom-up and top-down processing in comprehending spoken language, the emphasis of books is mostly on top-down processing, e.g. listeners are asked to get the main idea of what they heard. On the other hand, all skills are not given equal attention while teaching should be designed such that all skills grow equally. A reasonable approach to tackle this issue could be to incorporate teaching of reduced forms in listening sections of the books and their workbooks.

Since this study, to the best of the researcher's knowledge, is the first attempt on administering a test of reduced forms on spoken language in an Iranian EFL context, the results will set the scene for more extensive research in this area. Hence the results of this study will be of benefit for test developers, since it offers useful tips in designing a test in the area of listening comprehension in general and reduced forms in particular.

An important question arising from this study might that with which aspects of reduced forms make listening comprehension more difficult for language learners to understand. Although there is a growing body of research in the area of teaching reduced forms, little attempt has been exerted to design materials by materials developers to act as a guidance in developing tests of reduced forms in spoken language.

4.1 Limitation of the study

The main limitation of the study may be related to administration problem. It was not possible for the present researcher to administer a placement test because the language institutes were under time limitation. Therefore, the present researcher had to content herself to the placement test conducted by those language institutes and the consultant with those institute managers and language teachers. For example, students at levels 5,6,7,8 were as intermediate and students at levels 9, 10, 11, 12 were as advanced level students.

References

- Ary, D., Jacobs, L. C., & Razavieh, A. (1990). *Introduction to Research in Education*. Philadelphia, PA: Hold, Rinehart, and Winston, Inc.
- Bachman, L. F. (1990). *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press.
- Bowen, J. D. (1976). Current Research on an Integrative Test of English Grammar. *RELC Journal*, 7, 30-37.
- Brown, J. D. (2006). *Authentic Communication: Whyzit Importan' ta Teach Reduced Forms?* Shizuoka, Japan: Tokai University College of Marine Science (p. 13-24).
- Brown, J. D., & Hilferty, A. (1986a). Listening for Reduced Forms. *TESOL Quarterly*, 20(4), 759-763.
- Brown, J. D., & Hilferty, A. (1986b). The Effectiveness of Teaching Reduced Forms of Listening Comprehension. *RELC Journal*, 17, 59-70.
- Brown, J. D., & Hilferty, A. (1989). *Teaching Reduced Forms*. Gendai Eigo Kyoiku January: 26-28.
- Brown, J. D., & Kondo-Brown, K. (Eds.). (2006a). *Perspectives on teaching connected speech to second language speakers*. Honolulu, HI: University of Hawai'i, National Foreign Language Resource Center.
- Brown, J. D., & Kondo-Brown, K. (2006b). Testing reduced forms. In J. D. Brown, & K. Kondo-Brown, (Eds.), *Perspectives on teaching connected speech to second language speakers* (pp. 247-264). Honolulu, HI: University of

Hawai'i, National Foreign Language Resource Center.

Buck, G. (2001). *Assessing Listening*. Cambridge: Cambridge University Press.

Chen, S. W. (2002). Problems in Listening Comprehension for Learners of EFL. *Studies in English Language and Literature*, 10, 57-70.

Cook, V. D. (1991). *Second language learning and language teaching*. London: Edward Arnold.

Cronbach, L. J. (1971). Test validation. In R. L. Thorndike (Ed.), *Educational measurement* (2nd ed., 443–507). Washington, DC: American Council on Education.

Ernestus, M., Baayen, R. H., & Schreuder, R. (2002). The recognition of reduced word forms. *Brain and Language*, 18, 162-173.

Flowerdew, J., & Tauroza, S. (1995). The Effect of Discourse Markers on Second Language Lecture Comprehension. *Studies in Second Language Acquisition*, 17, 435–458.

Foulke, E. (1968). Listening Comprehension as a Function of Word Rate. *The Journal of Communication*, 18, p. 198-206.

Fulcher, G. (1997). Text Difficulty and Accessibility: Reading Formulae and Expert Judgment. *System*, 25(4), 497-513.

Fulcher, G., & Davidson, F. (2007). *Language Testing and Assessment*. New York: Routledge.

Goh, C. C. M. (2000). *A Cognitive Perspective on Listening Comprehension Problems*. *Language, Culture and Curriculum*, 13, 137-152.

Harding, L. (2011). *Accent and Listening Assessment*. New York: Peter Lang.

Henrichsen, L. E. (1984). Sandhi-variation: A Filter of Input for Learners of ESL. *Language Learning*, 34, 103-126.

Hughes, A. (1989). *Teaching for Language Teachers*. Cambridge: Cambridge University Press.

Ito, Y. (2001). Effect of Reduced Forms on Input-Intake Process. *Second Language Studies*, 20(1), 99-124.

Kachru, B. (1985). Standards, codification and sociolinguistic realism: The English language in the outer circle. In R. Quirk & H. G. Widdowson (eds.), *English in the world: Teaching and learning the language and literatures*. Cambridge: Cambridge University Press, 11–30.

Krashen, S. D., & Terrell, T. D. (1983). *The Natural Approach: Language Acquisition in the Classroom*. Oxford: Pergamon.

Magiste, E. (1979). The competing linguistic systems of the multilingual: A developmental study of decoding and encoding processes. *Journal of Verbal Language and Verbal Behaviour*, 18, pp. 79-89.

Marsh, L. G., & Maki, R. H. (1978). Efficiency of arithmetic operations in bilinguals as a function of language. *Memory and Cognition*, 414, pp. 459-64.

Mendelsohn, D. (1995). Applying learning strategies in the second/foreign language listening comprehension lesson. In: Mendelsohn, D., Rubin, J. (Eds.), *A Guide for the Teaching of Second Language Listening*. Dominie Press, San Diego, 113–131.

Morley, C. (2007). Listening: Top down and bottom up. Retrieved September 25th, 2007, from <http://www.teachingenglish.org.uk/think/articles/listening-top-down-bottom>.

Norris, R. W. (1993). Teaching Reduced Forms: An Aid for Improving Lower-level Students' Listening Skills. *Fukoka Women's Junior College Studies*, 46, 49-56.

Oller, J. W. (1979). *Language Tests at School*. London: Longman Group Limited.

Rosa, M. (2002). *DON' CHA KNOW? A Survey of ESL Teachers' Perspectives on Reduced Forms Instruction*. University of Hawai'i at Mānoa.

Sanju G, Haque M & Oyeboode F. (2006). Standard setting: comparison of two methods. *BMC Med Educ* 6(46).

Schmidt-Rinehart, B. C. (1994). The effect of topic familiarity on second language listening comprehension. *Modern Language Journal*, 78(2), 179-189.

Shrock, S., & Coscarelli, W. (2007). *Criterion-referenced Test Development: Technical and Legal Guidelines for Corborate Training*. San Fransisco, CA: John Willey and Sons.

Slobin, D. (1973). Cognitive Prerequisites for the Development of Grammar. In C. Ferguson & D. Slobin (Eds.), *Studies of Child Language Development* (pp. 175–208). New York, NY: Academic Press.

Sun, K. (2002). Investigation of English Listening Difficulties of Taiwan Students. *Selected papers from the eleventh international symposium on English teaching/fourth Pan Asian conference* (pp. 518–526). Taipei: Crane Publishing Co.

Rivers, W. M., and Temperley, M. S. (1978). *A practical guide to the teaching of English as a second or foreign language*. New York: Oxford University Press.

Rosa, M. (2002). *DON' CHA KNOW? A Survey of ESL Teachers' Perspectives on Reduced Forms Instruction*. University of Hawai'i at Mānoa.

Taylor, L. (2006). The Changing Landscape of English: Implications for Language Assessment. *ELT Journal*, 60(1),

51-60.

Thompson, I., 1995. Assessment of second/foreign language listening comprehension. In: Mendelsohn, D., Rubin, J. (Eds.), *A Guide for the Teaching of Second Language Listening*. Dominie Press, San Diego, pp. 31–58.

Ur, P., (1984). *Teaching Listening Comprehension*. Cambridge University Press, Cambridge.

Vandergrift, L., (1999). Facilitating second language listening comprehension: acquiring successful strategies. *ELT Journal* (53), pp. 168–176.

Weinstein, N. (1982). *Whaddaya say? Guided Practice in Relaxed Spoken English*. Culve City, CA: ELS Publications.

Weinstein, N. (2001). *Whaddaya say? Guided Practice in Relaxed Speech*. New York: Addison Wesley Longman.

Wells, S.C. & Wollack, J.A. (2003). *An Instructor's Guide to Understanding Test Reliability*. Testing & Evaluation Services publication, University of Wisconsin. Retrieved January 4, 2006 from <http://www.wiscinfo.doit.wisc.edu/exams/Reliability.pdf>

Wheaton, A., & Parry, J. (2012). Using the Angoff Method to Set Cut Scores. *Users Conference* (pp. 1-7), New Orleans, LA.

Yanagawa, K., & Green, A. (2008). To show or not to show: The effects of item stems and answer options on performance on a multiple-choice listening comprehension test. *System* (38), pp. 107-122.

Zahedi, H. (1997). *Dicto-Phrase: A New Listening Comprehension Test* (Unpublished master thesis). Shiraz University, Shiraz.

Zare-ee, A. (2007). The Significance of Pauses in EFL Listening Comprehension Tests. repo.uum.edu.my/3215/1/Ab2.pdf

APPENDIX

Name:

The listening test consists of four parts. You will hear each sentence, conversation, or lecture once only. During the test you may not skip questions and come back to them later. Try to answer all questions that you hear on this test. Before you listen, you will have 10 seconds to have a look at the questions.

Part A

Directions: Listen and circle the word, phrase, or sentence you here.

1. It was not as Gwen expected it to be. /It wasn't as good as I had expected it to be.
2. She's talking about it at the end of a lesson. / She spoke about it at the end of the lesson.
3. They went for a walk. /They went for work. 4. She could've been more considerate. /She could be more considerate.
5. They **shouldn't tell/have told** the police anything.
6. She **d have been/would be** more than happy to help.
7. The postman said it arrived. /The postman said it had arrived.
8. How do you go to school? / How do you get to school?
9. I might have left my keys in the office. /Emma left my kids in the office.
10. The retreat places/There are three places.
11. Draw all the flowers. / There are all the flowers.
12. There is a comma after that. /Does it come an "R" after that?
13. When do/did you go shopping?
14. Joe travels much. / Where do you have lunch?
15. Would you have lunch? / Where do you have lunch?

Part B

Directions: In this part, you will hear three separate talks. Underline the word, phrase, or sentence you hear.

Talk One

Todd: what do you do with your dog when you travel? Who takes care of your dog?

Greg: Well, we have very good friends who love to have him. In fact, sometimes my friends offer before we go. They know **(16) they want to go/we're going**, "Hey, can we take care of Dancer?" and they take good care of him.

Todd: So, have you ever actually traveled with your dog?

Greg: Actually, yes. Well, when we travel in Japan, he travels with us almost everywhere we go. He sits in the car and sometimes we camp and sleep in the car with him or we stay in tents and he travels with us ... but **(17) we don't want to travel/we've also traveled (contraction)** overseas with him.

Part C

Directions: Listen and fill in the gaps with the word/words you hear. WRITE NO MORE THAN THREE WORDS for each answer and use full forms not contracted forms e.g. "I have" not "I've".

18. I always figure, she can make it better than me.
19. The hospital won't give out any information.
20. He sat next to me on the New York City subway.
21. I have been studying your equilibrium. The one you wrote here, at Princeton.
22. The couple is going like this you know, shrinking.

Part D

Directions: In this part of the test you will hear two separate talks. Fill in the blanks with the word, phrase, or sentence you hear. The talks will not be repeated. Use full forms not contracted forms of the word, phrase, and sentences e.g. "They are" not "They're".

Talk One

I started Improv Everywhere about 10 years ago when I moved to New York City with an **(23)**..... acting and comedy. Because I was new to the city, I didn't have access to a stage, so I decided to create my own in public places. So the first project we are going to take a look at is the very first No Pants Subway Ride. Now this took place in January of 2002.

Talk Two

What I posit and what positive psychology posits is that if we study what is merely average, we will remain merely average. Then those positive outliers, what **(24)**..... is come into a population like this one and say, why? Why is it that **(25)**..... you are so high above the curve in terms of your intellectual ability, athletic ability, musical ability, creativity, energy levels, your resiliency in the face of challenge, your sense of humor? Whatever it is, instead of deleting you, what I **(26)**..... is study you. Because maybe we can glean information **(27)**..... to move people up to the average, but how we can **(28)**..... average up in our companies and schools worldwide. The reason this graph is important to me is, when I turn on the news, it seems like the majority of the information is not positive, in fact **(29)**..... Most of it is about murder, corruption, diseases, and natural disasters.