The Development and Validation of Willingness-to-Listen in L2 (WTL) Scale

Ahmet Selçuk Akdemir
Ağrı İbrahim Çeçen University

Abstract

This study aims to develop and validate a willingness to listen (WTL) scale for second language (L2) contexts. There are several studies on willingness to communicate (WTC) in L2. However, most of these studies are based on the speaking dimension of communication. As an ignored dimension, listening in L2 is underestimated in WTC research. The study is based on the fact that listening, along with speaking, composes an important part of communication in L2 contexts. In order to develop a valid and reliable WTL in L2 scale, a series of analyses has been carried out after creating the items for the scale. The sample group is composed of 335 students for the first application, and 97 students for the second application. In order to assess the reliability and validity of the scale, Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) have been conducted. Also, Cronbach's α, Guttman Split-Half, and Spearman-Brown coefficients have been calculated. Consequently, WTL in L2 scale is

developed as a Likert-type scale with 19 items which constitute a four-factor structure. WTL in L2 is a valid and reliable scale in determining willingness to listen structure in various L2 contexts.

Keywords: willingness to listen, L2 listening, listening skills, communicative competence

Introduction

Contemporary conceptualisation of second language (L2) acquisition entails the development of communicative competence as well as language skills acquisition. L2 listening constitutes an important part of communicative competence. Learners are expected to be equipped with listening and speaking abilities as their communicative competence increases. However, research has shown that some L2 learners remain silent though they have a high level of communicative competence, while others have better communicative performance with limited competence (Baghaei, Dourakhshan, & Salavati, 2012). This contradicts the primary reason for language learning which is 'to be able to use language to communicate' (MacIntyre & Charos, 1996).

The importance of individual differences in L2 competence and performance has been verified by several studies (Dörnyei, 2005; Dörnyei & Skehan, 2003; Ehrman, Leaver, & Oxford, 2003; Grey, Williams, & Rebuschat, 2015; Skehan, 1991). The affective variables such as attitudes, motivation, and anxiety in L2 are among the most frequent topics of research on individual differences in L2. Along with 'motivation', the term 'willingness' as a personality factor in L2 has long been identified in several studies (Ellis, 1994; Dörnyei, 2005; Lightbown & Spada, 2006).

Review of Related Literature

L2 Willingness to Communicate (L2 WTC)

Willingness to communicate (WTC), a recent addition to the affective constructs, is emerging as a useful concept in individuals'

communicative preferences and tendencies in L2 (Yashima, 2002). The concept was first developed for first language (L1) communication referring to individuals' tendencies to engage in communication in the L1, when given a free choice (McCroskey, 1992; McCroskey & Baer, 1985; McCroskey & Richmond, 1987) and was applied to L2 communication (MacIntyre & Charos, 1996). L2 WTC is defined as 'a readiness to enter into discourse at a particular time with a specific person or persons, using a L2' (MacIntyre, Clément, Dörnyei, & Noels, 1998, p. 547). As WTC found its way in L2 research as an independent background variable, several studies have been conducted to investigate its relationships with communicative competence and performance in L2 (Baghaei et al., 2012; Bektaş & Çetinkaya, 2009; Hashimoto, 2002; Öz, Demirezen, & Pourfeiz, 2015; Peng, 2007) self-confidence (Clément, Baker, & MacIntyre, 2003; Yashima 2002), motivation and attitudes (MacIntyre, Baker, Clément, & Donovan, 2003; Ryan 2009), international posture (Yashima 2002, 2009), personality (Ghonsooly, Khajavy, & Asadpour, 2012; MacIntyre & Charos 1996), age and gender (MacIntyre, *Baker*, Clément, & *Donovan*, *2002*), L2 communication frequency (MacIntyre, Baker, Clément, & Donovan, 2002; MacIntyre et al., 2003; Yashima, Zenuk-Nishide, & Shimizu, 2004), ambivalence in immersion classes (MacIntyre, Burns, & Jessome, 2011), the impact of computer assisted language learning activities on WTC (Reinders & Wattana, 2015), cultural stereotypes (Wen & Clément, 2003), interactional and situational classroom context (Cao, 2011; Cao & Philp, 2006) and classroom environment (Alishah, 2014; Peng & Woodrow 2010; Şener, 2014). In their situational model of L2 confidence and affiliation, MacIntyre et al. (1998) propose WTC as the primary goal of language teaching.

In order to measure WTC in L2 contexts, several scales have been developed, most of which are based on McCroskey and Baer's (1985) WTC scale, which was originally developed for L1 communication research (MacIntyre & Charos, 1996; Baker & MacIntyre, 2000; MacIntyre et al., 2003; MacIntyre et al., 2002; Hashimoto, 2002; Yashima, 2002; Clément et al., 2003; Yashima,

Zenuk-Nishide, & Shimizu, 2004; Ghonsooly et al., 2012). Using the Rasch model, Weaver (2005) has developed a measure of L2 learners' WTC in speaking situations and writing situations in Japanese EFL classrooms. There are also some WTC measures having the dimensions of inside and outside classroom situations (MacIntyre, Baker, Clément, & Conrod 2001; Peng, 2007; Peng, 2013; Ryan, 2009; Peng & Woodrow, 2010). As the original form of WTC is mainly focused on speaking and writing abilities, most of these scales have the dimensions of speaking and writing. Though listening skill is accepted as an important component in L2 communication (Harmer, 2001; Richards & Renandya, 2002), interestingly L2 WTC scales underestimate listening as a dimension to be measured in learners' willingness to communicate. Indeed, the original form of WTC, developed for L1, assesses how a person is willing to speak in various situations. Roberts and Vinson (1998) argue that a great deal of researchers studying WTC 'have focused on the sending of information, not the reception of that information' (p. 46). All twenty items of McCroskey and Richmond's (1985) WTC measure a person's willingness to talk to others, not to communicate. However, communication is not restricted to speaking skills. This is valid also for L2 communication. Several studies point to L2 listening as the counterpart of speaking in L2 communication (Dunkel, 1991; Kurita, 2012; Murphy, 1991; Rost, 2001; Vandergrift, 2007). Obviously, nearly all L2 WTC adaptations also ignored listening as a component of communication. The L2 WTC scale Peng (2007) used in her study includes only three items for the listening part of communication (the items dealing with listening are: #23, #26, #27). In the same study, there are eight items for speaking, seven items for reading and nine items for writing. Khatib and Nourzadeh (2015) have developed and validated a WTC questionnaire for instructional settings. The items in the Instructional Willingness to Communicate (IWTC) questionnaire are based on an extensive literature review as well as interviews with experts and language teachers. After Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) conducted

in EFL settings, the final form of IWTC is found to have 27 items under six factors. Yet, IWTC ignores listening. Only three items out of 27 exist to measure the listening part of willingness to communicate in an EFL setting (the items dealing with listening are: #12, #19, and #22).

L2 Willingness to Listen (L2 WTL)

As a subsidiary component of WTC, willingness to listen (WTL) in L2 has not been studied adequately. There are some studies dealing with the issue in terms of L1 communication. Roberts and Vinson (1998) developed the first WTL scale based on the WTC scale developed by McCroskey and Richmond (1985). According to Roberts and Vinson (1998), 'a WTL scale would be similar in many ways to what McCroskey and Richmond, among others, have developed to measure willingness to communicate, but would vary in several important ways' (p. 46). They mention the very basic distinction of a WTL scale as focusing on listening rather than speaking. Designed as a self-report, the WTL scale, developed by Roberts and Vinson (1998), also aimed to determine the basic constructs of L1 communication with a broader focus on listening. Their WTL scale seems to be restricted to listening in a natural context. Nearly all items measure listening to a friend or a person encountered in a natural setting. However, listening is not restricted to everyday communication. Listening, especially L2 listening, is quite different from listening in a persons' L1. An L2 WTL scale is expected to have items for classroom listening as well as listening to daily conversations in the target language. The WTL scale developed by Roberts and Vinson (1998) has three dimensions: the speaker, the content, and the environment. It lacks an important dimension: the listener. Their scale disregards one of the two agents in communication. A modified version of their scale was prepared by Richmond and Hickson (2001). It was another self-report scale with 24 items. Their scale was presented to be used in a public speaking classroom. The items in both of the scales are designed to assess WTL in a L1 context. They focus on the speaker, the content, and the environment. There is no

item related to linguistic patterns such as rhythm, intonation, or accent. However, it is undeniable that any L2 context, which is expected to consist of native speakers and non-native speakers, should be considered in terms of the linguistic patterns of the communication. Thus, neither Roberts and Vinson's preliminary scale nor Richmond and Hickson's modified version is suitable for WTL in L2.

There is a research gap in L2 WTL in the relevant literature. There is no scale to determine language learners' WTL in L2, though 'willingness to listen' is defined as an important concept in communication (Roberts & Vinson, 1998). Determining L2 learners' WTL is quite important regarding the fact that L2 listening is a complex process with several dimensions and it needs to be measured in terms of other personality factors such as motivation, apprehension, anxiety etc.

The aim of this research is to fill this gap in the literature by developing and validating a L2 WTL scale in order to determine L2 learners' willingness to listen. The development and validation process has three steps: a) creating an item pool based on extensive literature review along with expert opinion, b) conducting a pilot study with the first draft, c) main study with the final version of WTL.

Method
Context

The research was conducted in the English language teaching (ELT) department of a public university in Turkey. The status of English in Turkey is that of a foreign language as there is no English-speaking community in the country. English is taught as a foreign language in all public and private educational institutions. Language teachers are also educated by non-native English speaking teachers. Listening skills in L2 are practiced and developed through formal instruction during language teaching. The study was conducted with the participation of advanced language learners who were also teacher candidates at the ELT department.

Participants

The participants of this study were 335 students who were in the second, third and fourth year (mean age: 22.3, age range: 17–26) of the English Language Teaching (ELT) department of a Turkish university's education faculty for the pilot study; and 96 students who were in the second and third year (mean age: 21.8, age range: 19–26) in the English Language Teaching (ELT) department of a Turkish university's education faculty for the main study. All of the students have a listening education background during their preparatory class and first year in the department.

Preparation Process of the Scale

During the development process of the scale, an extensive literature review has been performed (Flowerdew & Miller, 1992; Hamouda, 2013; Hasan, 2000; Kalivoda, 1981; Kurita, 2012; Rubin, 1994). In this regard, the factors and conditions affecting L2 listening were examined. In addition, L2 WTC scales developed by several scholars and WTL scales (Richmond & Hickson, 2001; Roberts & Vinson, 1998) developed for L1 communication were also investigated. An item pool was created based on the relevant literature. This preliminary item pool was examined, both in terms of the coinciding items and content validity by three faculty members, two of whom were L2 listening experts and one was an English Language Teaching (ELT) expert. A 67-item pool was formed with the information gathered from the literature and the contributions of experts in the area. Twenty eight of the items in this pool consisted of positive statements and 39 consisted of negative statements. A range of 5-point choices was placed for the items in order to specify the students' attitude levels expressed in the items. These choices were organised and graded as "(1) never", "(2) seldom", "(3) sometimes", "(4) often", and "(5) always".

The draft scale was examined by 22 students from the ELT department of a public university in order to determine whether the statements were clear enough to be understood by university level learners. The finalised scale was applied to the second, third

and fourth year students in a public university's ELT department. All participants had L2 listening background through the preparatory class of the department. They also had completed a listening course in the first year. The data collected electronically were uploaded to SPSS 22.00 in order to statistically conduct the validity and reliability analyses of the scale. The values regarding negative statements were reverse coded while being uploaded to the programs.

The Nature and Content of the Items

There are 19 items in the final version of the WTL scale (See Appendix 1). The items of the scale were designed as statements starting '*I am willing to listen … / I am unwilling to listen …*' in order to investigate to what extent the participant is willing to listen in various contexts. The items include both linguistic patterns (such as rhythm, intonation, accent, and speech rate) and language used in L2 listening. As the scale was designed to be used in an L2 context, there are several items investigating listening in the classroom context. The items were designed to expand the investigation to all core concepts related to L2 listening such as listening strategies, pair/group work in listening, and latent background knowledge of the topic in listening.

Data Analysis

Before conducting any analysis, the data was checked for missing and faulty values. Sixteen papers were removed from the study as they had missing or faulty values. The exploratory factor analysis (EFA) was carried out with 319 papers.

The Kaiser-Meyer-Olkin (KMO) measure and the Bartlett test were employed to examine the appropriateness of the data for factor analysis. When the KMO value is above .60 it indicates the adequacy of sampling. The Bartlett test result should be below .05 to indicate that the data is appropriate for Exploratory Factor Analysis (EFA) (Büyüköztürk, 2012).

The anti-image correlation matrix was investigated in order to determine whether the correlations between items are at an

adequate level. Then EFA was conducted and the scale's allocation to factors was specified through maximum likelihood analysis and the factor loads were examined using the Varimax rotation method. The items with factor loads lower than .30 and the items that do not have at least .100 difference between their loads on two factors, or in other words, the items with loads separated into two factors, should be removed (Büyüköztürk, 2012).

After conducting EFA, a Confirmatory Factor Analysis (CFA) was conducted. In the scale model obtained as a result of the confirmatory factor analysis, having the observed values between the ranges of c2 /d < 3; 0 < RMSEA < .05; 0 S-RMR .05; .97 NNFI 1; .97 CFI 1; .95 GFI 1; .95 AGFI 1 and .95 IFI 1 indicates a perfect fit, while having them between the ranges of c2 /d < 5; .06 RMSEA < .08; .06 S-RMR .08; .90 NNFI .96; .90 CFI .96; .90 GFI .96; .90 AGFI .96 and .90 IFI .96 indicates an acceptable fit (Kline, 2005; Şimşek, 2007, pp. 18–71).

Discrimination is accepted as one of the most important proofs used in determining the validity of a scale (Büyüköztürk, 2002). Another way of testing a scale's discrimination is to observe the differentiation between the lowest 27% of the groups and the highest 27% of the groups after sorting raw scores obtained from an item from the highest to the lowest. In order to determine the scale's reliability, stability tests were conducted with internal consistency coefficients. The Cronbach's alpha reliability coefficient, the correlation value between two congruent halves, the Spearman–Brown formula and Guttmann split-half reliability formula were used in order to determine the internal consistency level. A reliability coefficient that is higher than .70 is accepted as an indication of the scale's reliability (Büyüköztürk, 2002; Gorsuch, 1983).
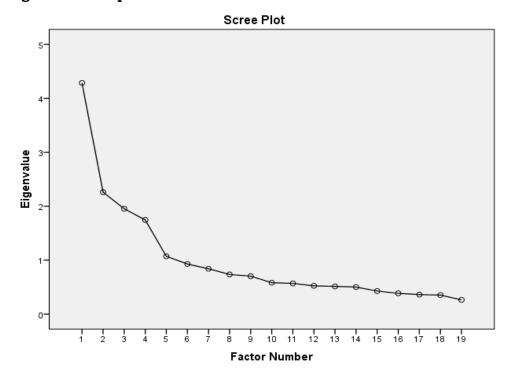
Besides, in order to analyse criterion-related validity, WTL was conducted with the Foreign Language Listening Anxiety Scale (FLLAS) developed by Kim (2000). Pearson correlation coefficients of the two scales were investigated.

Results

EFA

In order to determine the factor structure of the scale, EFA was conducted. KMO and Bartlett's test were conducted to test the adequacy of the data for factor analysis. As a result of EFA conducted for WTL scale, KMO value was calculated as .77; Bartlett's test x^2 value was found to be 2102.4 (p<.001). Data is proved to be appropriate for factor analysis as KMO is above .60 and Bartlett's test is meaningful (Büyüköztürk, 2012). Then, anti-image correlation of items was investigated and three items (#7, #4, #59) were removed as their anti-image correlations were below .50. As a result of maximum likelihood and Varimax rotation analyses, 45 items were removed as their factor loads were below .30 and factor load values between two items was less than .100. As a result of EFA, a four-factor structure, explaining 54% of total variance, was obtained. The scale's eigenvalue was found to be bigger than 1.00. Figure 1. shows the scree plot graphic of the scale.

Figure 1: Scree plot for the EFA

The first factor consisted of four items (#13, #14, #15, #16). Item loads of this factor varied between .50 and .73. This factor, explaining 10.282% of the total variance, forms the 'speaker dimension' of the scale. The second factor consisted of four items (#28, #29, #30, #34). Item loads of this factor varied between .51 and .89. This factor, explaining 11.897% of the total variance, forms the 'listener dimension' of the scale. The third factor consisted of four items (#41, #42, #43, #44). Item loads of this factor varied between .56 and .69. This factor, explaining 9.184% of the total variance, forms the 'task dimension' of the scale. The fourth factor consisted of seven items (#38, #48, #49, #50, #52, #53, #54). Item loads of this factor varied between .53 and .63. This factor, explaining 22.553% of the total variance, forms the 'topic dimension' of the scale. Factor loads of items and variances of each sub-dimension are given in Table 1.

Table 1: Item factor loads and variances of each sub-dimension for WTL

Items	Factor loads for sub-dimensions of the scale			
#15	.725			
#16	.680			
#14	.553			
#13	.502			
#29		.888		
#30		.683		
#28		.639		
#34		.510		
#43			.689	
#44			.585	
#42			.582	
#41			.555	
#52				.626
#49				.602
#48				.590
#50				.579
#54				.569
#53				.566
#38				.529
Total Variance	% 10.282	% 11.897	% 9.184	% 22.553
	Total Variance: % 53.915			

As Table 1 indicates, the scale consists of four dimensions explaining 54% of total variance. Factor loads of these dimensions are between .50 and .73 for speaker dimension, between .51 and .89 for listener dimension, between .56 and .69 for task dimension, between .53 and .63 for topic dimension. These show that item factor loads are at acceptable levels for a four-factor structure.

CFA

In order to validate the factor structure of the scale, primary and secondary level confirmatory factor analysis (CFA) has been conducted. CFA is used to investigate the balance of theoretical factors with real data. In other words, CFA is the way of confirming any pre-determined structure with collected data (Harrington, 2008; Jöreskog & Sörbom, 1996; Sümer, 2000). There are several fit indices to determine how well an a priori model fits, or reproduces the data (McDonald & Ho, 2002). Chi-square goodness, RMSEA, S-RMR, CFI, and IFI values have been investigated for the CFA conducted in this study. After a review of the prominent guidelines for interpreting fit indices, Brown (2006) has argued that RMSEA values less than 0.05 indicate good fit, those between 0.05 and 0.08 show adequate fit, the range 0.08–0.1 indicates mediocre fit and RMSEA values above 0.1 are a sign of poor model fit. In addition, SRMR values of less than 0.08 are also acceptable for model fit. Furthermore, Incremental Fit Index (IFI) and Comparative Fit Index (CFI) should be generally higher than .90 for a model to show adequate fit.

Fit indices for the scale's four-dimensional model have been investigated through CFA. The data of CFA's first step are given in Figure 2.
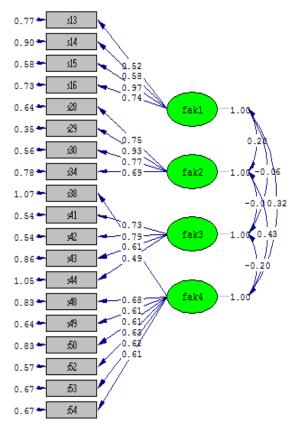
Figure 2: The results of CFA's first level for WTL

Chi-Square=489.88, df=146, P-value=0.00000, RMSEA=0.079

It can be seen from the figure that fit indices are meaningful for WTL which is a four dimensional scale with 19 items (X^2 = 489.88, sd = 146, p = .00, X^2 /sd = 3.35). The values for fit indices have been found as RMSEA = .079, S-RMR = 0.068, CFI = .90, IFI = .90. No modification is needed for the items as the fit indices of the first level CFA have been found to be good. Meydan and Şeşen (2011) emphasise the need for a second level CFA for multi-dimensional scales. The results of second level CFA are shown in Figure 3.
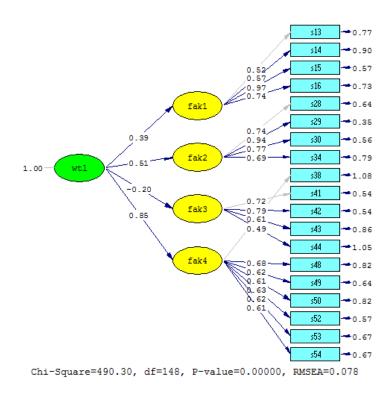
Figure 3: The results of CFA's second level for WTL

Chi-Square=490.30, df=148, P-value=0.00000, RMSEA=0.078

It can be seen from the figure that fit indices are meaningful for WTL which is a four dimensional scale with 19 items (X^2 = 490.30, sd = 148, p = .00, X^2/sd = 3.31). The values for fit indices have been found as RMSEA = .078, S-RMR = 0.069, CFI = .90, IFI = .90. No modification is needed for the items as the fit indices of the second level CFA have been found to be good.

As can be seen from Figure 3, the factor loads of the items are between .52 and .97 for speaker dimension, between .69 and .93 for listener dimension, between .49 and .73 for task dimension, between .61 and .68 for topic dimension. After standard analyses, t-value between items has been calculated and it was found to be meaningful in .05 level for all items.

Criterion-related Validity

The Participants

The study for criterion-related validity of the WTL scale has been conducted with the students of an ELT department of a Turkish public university. Ninety-seven students have been included in the study.

Procedure

In order to determine the criterion-related validity of the WTL scale, the Foreign Language Listening Anxiety Scale (FLLAS) developed by Kim (2000) has been used. The correlation between the two scales has been investigated.

Instruments

- #### *FLLAS*

The FLLAS, developed by Kim (2000), has 33 five-point scale items and it aims at investigating foreign language listening anxiety. The reason for using an anxiety scale for criterion-related validity is based on the fact that the relevant literature provides evidence for the relationship between anxiety and willingness to communicate in L2 (Baran-Lucarz, 2014; Chu, 2008; MacIntyre et al., 2003).

- #### *WTL Scale*

The current WTL scale has been used for a criterion-related validity study.

Findings

The correlation between WTL and FLLAS is shown in Table 2.

Table 2: The correlation between WTL and FLLAS

		fllastotal	wtltotal
fllastotal	Pearson Correlation	1	.599**
	Sig. (2-tailed)		.000
	N	97	97
wtltotal	Pearson Correlation	.599**	1
	Sig. (2-tailed)	.000	
	N	97	97

**p< .01

According to the analyses conducted to determine the criterion-related validity of WTL scale, a positive and meaningful relationship has been found between WTL and FLLAS (r = .599, p< .01).

Reliability

Reliability means that the scores of an instrument are stable and consistent (Creswell, 2008). In other words, a scale is accepted as reliable if it proves to have stability and consistency. In order to assess the reliability of the WTL scale, Cronbach's α, Guttman Split-Half, Spearman-Brown coefficient has been calculated. Reliability studies of the scale have been conducted with ninety-seven students from an ELT department in a public university in Turkey. Cronbach's α reliability coefficient is preferred when the items of a scale have more than three alternatives (Büyüköztürk, 2012). According to Fraenkel and Wallen (1996), the reliability of items is acceptable if the alpha value is within .70 and .99. The split-half method measures the extent to which all parts of a scale contribute to the measurement equally. Split-half reliability, conducted by splitting the scale into two equal parts and using the Spearman-Brown formula to explain the correlation of the scale, is used to prove the consistency of the scale. Reliability values of the scale are given in Table 3.

Table 3: Reliability values of the scale

Dimensions of WTL	Cronbach's α	Guttman Split-Half Coefficient	Spearman-Brown Coefficient
Speaker	.70	.61	.62
Listener	.83	.82	.82
Task	.70	.56	.56
Topic	.83	.76	.78
Total Reliability	.79	.60	.60

As can be seen from the table, Cronbach's α coefficient for WTL's total reliability has been found to be .79; Guttman split-half reliability coefficient has been found .60; and Spearman-Brown reliability coefficient has been found to be .60. Cronbach's α reliability coefficients of dimensions are: .70 for Speaker, .83 for Listener, .70 for Task, and .83 for Topic. Guttman Split-Half reliability coefficients of dimensions are: .61 for Speaker, .82 for Listener, .56 for Task, and .76 for Topic. The Spearman-Brown reliability coefficients of dimensions are: .62 for Speaker, .82 for Listener, .56 for Task, and .78 for Topic. These values prove that the scale has internal and split-half consistency. Total item correlations of the scale and t-test values for item score comparisons of 27% upper and 27% lower groups are given in Table 4.

Table 4: Total item correlations and t-tests for upper + lower 27%

Items	r	t
01	.36	5.288**
02	.39	4.967**
03	.59	6.828**
04	.46	6.773**
05	.71	7.784**
06	.72	8.586**
07	.73	6.954**

Items	r	t
08	.73	9.146**
09	.32	2.888**
10	.58	3.795**
11	.33	2.910**
12	.39	2.500*
13	.68	6.277**
14	.77	7.901**
15	.74	6.780**
16	.70	6.406**
17	.69	5.560**
18	.72	4.467**
19	.71	4.593**

* $p < .05$,   ** $p < .01$

It can be seen from Table 5. that the item-total correlation for all items is between .32 and .77, and t-values are meaningful ($p<.05$, $p<.01$). According to these values, it can be said that the items of the scale have a suitable validity and they have the proper discrimination power in terms of instrumental quality and measuring the same behaviour of the participants.

Conclusion

This paper reports the steps and findings of a set of research conducted for the development and validation of a scale which can be used to determine L2 learners' WTL. Though the concept 'willingness' exists in the relevant literature as one of the personality factors of individuals, there is no standardised scale to determine WTL structure for L2 learners. Up to now, the concept 'willingness' has been investigated in a broader sense as a 'willingness to communicate' structure which ignores listening in communication. Obviously, there is a need for a standardised scale to determine WTL. This scale is expected to fill the gap in the relevant literature and to be implemented as an instrument in determining the willingness of L2 learners.

In this paper the development and validation process of the WTL scale has three steps: a) item development; b) exploratory factor analysis; c) confirmatory factor analysis. As a result of the analyses, the scale has proved to serve the purpose of determining willingness to listen in L2 context. The statistical analyses revealed that the scale has four dimensions: speaker, listener, task and topic. These dimensions also reflect the item structure of the scale, that is, the items in the speaker dimension are related to the speaker variable in any L2 listening context; the items in the listener dimension are related to the listener variable; the items in the task dimension are related to the task variable; and the items in the topic dimension are related to the topic variable. The Cronbach's α reliability coefficient of dimensions are: .70 for Speaker, .83 for Listener, .70 for Task, and .83 for Topic and .79 for the scale's total reliability. In the light of reliability and validity analyses, it can be said that the WTL in theL2 scale is a valid and reliable scale which can be used in determining the willingness of listening in second language contexts.

This scale is the first example of its kind. Thus, it needs to be re-tested in various L2 settings. Undoubtedly, as a newly developed scale, WTL will be revised and developed according to future instrumentations. As a theoretical implication, it is suggested that the correlations between WTL and other constructs of L2 listening should be explored through quantitative methods. Also, there is a need for mixed-method studies to have a better understanding and conceptualising of the issues related to WTL. The scale is expected to be more than a theoretical structure. It should be used in the classroom by practitioners to explore learner profiles as listeners in L2.

As a limitation of the study, it should be noted that all the development and validation procedures of the current scale were conducted in a cross-sectional research design. Further studies might focus on the examination of WTL in a longitudinal research design with the participation of learners from various backgrounds. The effect of personal and individual differences might have an effect on WTL. Thus, future research trends should

be devoted to further investigations of the scale to have accurate scores in various settings.

Acknowledgement:

The Author

Dr Ahmet Selçuk Akdemir is an assistant professor at Ağrı İbrahim Çeçen University in Turkey. He was awarded his PhD from Atatürk University in Turkey on English Language Teaching programme in 2013 under the co-supervision of Dr. Michael Purdy from Governors State University, USA.

References

Alishah, A. R. (2014). Is willingness-to-speak more of a context-sensitive nature or personality? ELT Research Journal *3*(4), 169-180.

Baghaei, P., Dourakhshan, A. & Salavati, O. (2012) The relationship between willingness to communicate and success in learning English as a foreign language *The Modern Journal of Applied Linguistics*, *4* (2), 53–67.

Baker, S. & MacIntyre, P. D. (2000). The role of gender and immersion in communication and second language orientations. *Language Learning*, 50, 311-341.

Baran-Lucarz, M. (2014). The link between pronunciation anxiety and willingness to communicate in the foreign-language classroom: The Polish EFL context. *The Canadian Modern Language Review*, *70* (4): 445-473.

Bektas-Cetinkaya, Y. (2009) College students' willingness to communicate in English: Turkish context VDM Verlag, Saarbrücken.

Brown, T. A. (2006). Confirmatory factor analysis for applied research. New York: Guilford.

Büyüköztürk, S. (2002). Faktör Analizi: Temel Kavramlar ve Ölçek Geliştirmede Kullanımı. Eğitim Yönetimi Dergisi**,** Güz, 470-433

Büyüköztürk, Ş. (2012). Sosyal bilimler için veri analizi el kitabı. Ankara: Pegem Akademi.

Cao, Y. (2011). Investigating situational willingness to communicate within second language classrooms from an ecological perspective. *System*, *39*(4), 468–479.

Cao, Y., & Philp, J. (2006). Interactional context and willingness to communicate: A comparison of behavior in whole class, group and dyadic interaction. *System, 34*(4), 480–493.

Chu, H. R. (2008). Shyness and EFL Learning in Taiwan: A Study of Shy and Non-shy College Students' Use of Strategies, Foreign Language Anxiety, Motivation, and willingness to communicate. (Doctoral Dissertation). Retrieved from www.proquest.com (accessed 29/1/2015).

Clément, R., Baker, S. C. & MacIntyre, P. D. (2003). Willingness to communicate in a second language: The effects of context, norms, and vitality. *Journal of Language and Social Psychology, 22* (2), 190–209

Creswell, J. W. (2008). Educational research: Planning, conducting, and evaluating quantitative and qualitative research (3rd ed.) Upper Saddle River, NJ: Pearson Education, Inc.

Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition.* Mahwah, NJ: Lawrence Erlbaum.

Dörnyei, Z., & Skehan, P. (2003). Individual differences in second language learning. In C. J. Doughty, & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 589-630). Oxford: Blackwell.

Dunkel, P. A. (1991). Listening in the native and second/foreign language: Toward an integration of research and practice. *TESOL Quarterly, 25*, 431-457

Ehrman, M., Leaver, B., & Oxford, R. L. (2003). A brief overview of individual differences in second language learning. *System*, 31, 313–330.

Ellis, R. (1994). The study of second language acquisition. Oxford: Oxford University Press.

Flowerdew, J., &Miller, L. (1992). Student perceptions, problems a nd strategies in secondlanguage lecture comprehension. *RELC Journal*, *23*(2), 60-80.

Fraenkel, J.R., & Wallen, N.E. (2006). How to design and evaluate research in education. New York: McGraw-Hill.

Ghonsooly, B., Khajavy, G.H. & Asadpour, S.F. (2012) Willingness to communicate in English among Iranian non-English major university students. *Journal of Language and Social Psychology, 31* (2),197–211.

Grey, S., Williams, J. N., & Rebuschat, P. (2015). Individual differences in incidental language learning: The role of learning styles, personality, and working memory. *Learning and Individual Differences*, 38, 44-53.

Hamouda, A. (2013). An Investigation of Listening Comprehension Problems Encountered by Saudi Students in the EL Listening Classroom. *International Journal of Academic Research in Progressive Education and Development*. *2*(2), 113-155.

Hasan, A. (2000). Learners' perceptions of listening comprehension problems. *Journal of Language, Culture and Curriculum,* 13, 137-152.

Harmer, J. (2001). *The Practice of English Language Teaching,* London: Pearson Education Limited.

Harrington, D. (2008) Confirmatory Factor Analysis New York, New York: Oxford University Press.

Hashimoto, Y., (2002) Motivation and willingness to communicate as predictors of reported L2 use: The Japanese ESL context. *Second Language Studies, 20* (2), 29–70.

Jöreskog, K. ve Sörbom, D. (1996). *LISREL 8: Structural equation modelling with the SIMPLIS command* language. Chicago: Scientific Software International/Erlbaum.

Kalivoda, T. B. (1981). Developing advanced listening comprehension skill in a foreign language: Problems and possibilities. *Hispania*,64, 80–85.

Khatib, M., & Nourzadeh, S. (2015). Development and validation of an instructional willingness to communicate questionnaire. *Journal of Multilingual and Multicultural Development, 36*(3), 266-283.

Kim, J.-H. (2000). Foreign language listening anxiety: A study of Korean students learning English. Unpublished doctoral dissertation, University of Texas, Austin.

Kurita, T. (2012). Issues in second language listening comprehension and the pedagogical implications. *Accents Asia, 5*(1), pp. 30-44.

Lightbown, P. M. & Spada, N. (2006) How Languages are Learned. (3rd edition) Oxford: Oxford University Press.

MacIntyre, P., Baker, S., Clément, R., & Conrod, S. (2001). Willingness to communicate, social support, and language-learning orientations of immersion students. *Studies in Second Language Acquisition*, 23, 369-388.

MacIntyre, P. D., Baker, S. C., Clément, R., & Donovan, L. A. (2002). Sex and age effects on willingness to communicate, anxiety, perceived competence, and L2 motivation among junior high school French immersion students. *Language Learning, 52*(3), 537-564.

MacIntyre, P.D., Baker, S.C., Clément, R. & Donovan, L.A. (2003)Talking in order to learn: Willingness to communicate and intensive language programs. The Canadian Modern Language Review, 59 (4) (2003), pp. 589–607.

MacIntyre, P. D., Burns, C. & Jessome, A. (2011) Ambivalence about communicating in a second language: A qualitative study of French immersion students' willingness to communicate. *The Modern Language Journal, 95* (1), 81–96.

MacIntyre, P.D. & Charos, C. (1996). Personality, attitudes, and affect as predictors of second language communication. *Journal of Language and Social Psychology*, 15, 3–26.

MacIntyre, P. D., Clément, R., Dörnyei, Z., & Noels, K. A. (1998). Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation. *Modern Language Journa*l, *82*, 545-562.

McCroskey, J. C. (1992). Reliability and validity of the willingness to communicate scale. *Communication Quarterly, 40,* 16-25.

McCroskey, J. C., & Baer, J. E. (November, 1985). Willingness to communicate: The construct and its measurement. Paper presented at the annual convention of the Speech Communication Association, Denver. CO.

McCroskey, J. C., & Richmond, V. P. (March, 1985). Willingness to communicate and interpersonal communication. Paper presented at the Symposium on Personality and Interpersonal Communication, West Virginia University, Morgantown, WV.

McCroskey, J. C., & Richmond, V. P. (1987). Willingness to communicate. In J. C. McCroskey & J. A. Daly (Eds.), *Personality and interpersonal communication* (pp. 119-131). Newbury Park, CA: Sage.

McDonald, R. P., & *Ho*, M-H. R. (*2002*). Principles and practice in reporting structural equation analyses. *Psychological Methods*, 7, 64 – 82.

Meydan, H. C. & Şeşen, H. (2011), Yapısal Eşitlik Modellemesi AMOS Uygulamaları, Detay Yayıncılık, Ankara.

Murphy, J. (1991). Oral communication in TESOL: Integrating speaking, listening, and pronunciation. *TESOL Quarterly*, 25, 51-75.

Öz H., Demirezen M., Pourfeiz J., (2015). Willingness to communicate of EFL learners in Turkish context. *Learning and Individual Differences*, 37, 269-275.

Peng J. (2007) Willingness to communicate in an L2 and integrative motivation among college students in an intensive English language program in China. *University of Sydney Papers in TESOL, 2* (1), 33–59.

Peng, J. (2013). The Challenge of Measuring Willingness to Communicate in EFL Contexts. *The Asia-Pacific Education Researcher 22* (3), 281-290.

Peng, J. & Woodrow, L. (2010), Willingness to Communicate in English: A Model in the Chinese EFL Classroom Context. *Language Learning*, 60: 834–876. doi: 10.1111/j.1467-9922.2010.00576.x

Reinders, H., & Wattana, S. (2015). The effects of digital gameplay on L2 interaction. *International Journal of Computer-Assisted Language Learning and Teaching (IJCALLT), 5*(1), 1-21.

Richards J. C. & Renandya, W. A. (2002). Methodology in language teaching: An anthology of current practice, (pp.254-266), Cambridge: Cambridge University Press

Roberts, C., & Vinson, L. (1998). Relationship among Willingness To Listen, Receiver Apprehension, Communication Apprehension, Communication Competence, and Dogmatism. *International Journal of Listening.* 12, 40-56.

Rost, M. (2001). Listening. In R. Carter & D. Nunan (Eds), The Cambridge guide to teaching English to speakers of other languages(pp. 7-13). Cambridge, UK: Cambridge University Press.

Rubin, J. (1994), A Review of Second Language Listening Comprehension Research. *The Modern Language Journal*, 78: 199–221. doi: 10.1111/j.1540-4781.1994.tb02034.x

Ryan, S. (2009). Self and Identity in L2 Motivation in Japan: The Ideal L2 Self and Japanese Learners of English, in Z. Dörnyei and E. Ushioda (eds.) Motivation Language Identity and L2 Self. Bristol: *Multilingual Matters,* 120-143

Skehan, P. (1991). Individual Differences in Second Language Learning. *Studies in Second Language Acquisition*, 13, 275-298.

Sümer, N. (2000). Yapısal eşitlik modelleri: Temel kavramlar ve örnek uygulamalar. *Türk Psikoloji Yazıları, 3* (6) 49–74.

Şener, S. (2014). Turkish ELT students' willingness to communicate in English. *ELT Research Journal, 3*(2), 169-180.

Vandergrift, L., (2007). Recent developments in second and foreign language listening comprehension research. *Language Teaching,* 40, 191-210. doi:10.1017/S0261444807004338.

Weaver, C. 2005. Using the Rasch model to develop a measure of second language learners' willingness to communicate within a language classroom. *Journal of Applied Measurement, 6*(4), 396-415.

Wen, W. P. & Clément, R. (2003). A Chinese conceptualisation of willingness to communicate in ESL. *Language, Culture and Curriculum, 16*(1), 18-38.

Yashima, T. (2002). Willingness to Communicate in a Second Language: The Japanese EFL Context. *The Modern Language Journal, 86* (1), 54-66.

Yashima, T. (2009). International posture and the ideal L2 self in the Japanese EFL Context. In Dörnyei, Z. and Ushioda, E. (Eds.). *Motivation, language identity and the L2 self.* (Pp.144-163.）Clevedon, UK: Multilingual Matters.

Yashima, T., Zenuk-Nishide, L., & Shimizu, K. (2004). The influence of attitudes and affect on willingness to communicate and second language communication. *Language Learning, 54*(1), 119- 152.

Appendix:

WTL in L2 scale items

#	# in the first draft	Item	Factor (Dimension)
1.	13	I am willing to listen to a speaker with different rhythm.	Speaker
2.	14	I am willing to listen to a speaker with a high speech rate.	Speaker
3.	15	I am willing to listen to a speaker with a different/difficult vocabulary.	Speaker
4.	16	I am willing to listen to a speaker with a different or difficult accent.	Speaker
5.	28	I am unwilling to listen when I do not have enough vocabulary.	Listener
6.	29	I am unwilling to listen when I do not have enough background knowledge.	Listener
7.	30	I am unwilling to listen when I do not have enough prior knowledge.	Listener
8.	34	I am unwilling to listen when I do not have enough knowledge and application of listening strategies.	Listener
9.	43	I am willing to listen when I do pairwork.	Task
10.	44	I am willing to listen when I do groupwork.	Task
11.	42	I am willing to listen when I can predict the next.	Task
12.	41	I am willing to listen when I can predict from the title.	Task
13.	52	I am unwilling to listen because of conceptual difficulty of the text/audio.	Topic

#	# in the first draft	Item	Factor (Dimension)
14.	49	I am unwilling to listen when I have difficulty in interpreting the meaning.	Topic
15.	48	I am unwilling to listen if there are difficult grammatical patterns.	Topic
16.	50	I am unwilling to listen if the text/audio is long and dense.	Topic
17.	54	I am unwilling to listen because of syntactic complexity of the task.	Topic
18.	53	I am unwilling to listen because of syntactic complexity of the text/audio.	Topic
19.	38	I am unwilling to listen when I have difficulty in long answers/questions.	Topic