CURRICULUM & TEACHING STUDIES | RESEARCH ARTICLE

Receptive intelligibility of Turkish to Iranian-Azerbaijani speakers

Mohammad Salehi[1]* and Aydin Neysani[1]

Abstract: Azerbaijani and Turkish are two closely-related languages from Oguz branch of Turkic languages, which are said to be mutually intelligible. Regarding this background, we designed an experiment within the framework of receptive multilingualism to investigate the role of linguistic factors in intelligibility of the Turkish language to Iranian Azerbaijanis. The study was carried out by controlling extra-linguistic factors such as language contact and negative attitude in order to make the findings attributable to only linguistic factors as the main focus of the present study. In the study, 40 native speakers of Azerbaijani in Iran were exposed to the Turkish language through short videos recorded from Turkish TV programs. The participants scored in average 56% of receptive intelligibility in spoken language of Turkish; though this score is not considered a high score of receptive intelligibility of Turkish to the Azeri speakers. It provides some pieces of evidence to show how this half understanding paves the way for Azerbaijanis to learn the Turkish language through mere exposure to Turkish TV programs. Furthermore, in the present study we found lexical distance, and misleading effect of false friends to be more effective than phonetic distance between cognates of these languages, in determining the strength of receptive intelligibility of the Turkish language to Azerbaijani native speakers.

ABOUT THE AUTHOR

Mohammad Salehi is an applied linguist working and living in Tehran, Iran. He is very much interested in language-related studies. He is a bilingual Azeri-Persian speaker. He is interested in projects that maintain and preserve languages especially Azeri. Azeri has almost died in certain cities of Iran; only elders speak it. Salehi is currently working on a project to investigate the reasons for reluctance on the part of speakers to speak Azeri in work places. He is also a language testing specialist working on national tests that have life changing implications for individuals.

PUBLIC INTEREST STATEMENT

Turkish and Azeri belong to the same family of languages. Azeri speakers living in Iran visit Turkey Frequently for business and/or leisure purposes. While shopping or doing business with Turkish speakers, Azeri speakers face the daunting task of understanding Turkish speakers. Not knowing Turkish can present challenges for Azeri speakers. No study has ever investigated the understandability of Turkish for Azeri speakers living in Iran. Azeri spoken in Iran and Azerbaijan are different in some ways. This is because Azerbaijan and Turkey are closely related in terms of cultural issues. This study investigates the degree of receptive understandability of Turkish for Azeri speakers living in Iran. The participates were chosen from among Azeri speakers who were in not that much of a contact with Turkish. The results of the study show that despite similarities between the two varieties, there are also differences which can impede communication. It can be said that movies and TV channels are very effective tools to improve the understandability of Turkish.

1. Introduction

"Receptive multilingualism" (Haugen, 1966) or "semi communication" (Braunmüller & Zeevaert, 2001), has been defined as a degree of understanding between speakers of closely-related languages or different varieties of the same language in which interlocutors use their respective mother language when speaking to each other without the help of any lingua franca or use of any other language (ten Thije & Zeevaert, 2007). The notion of receptive multilingualism has gained momentum due to the fact that it has the potential to be implemented as a channel of communication instead of English as the most widely-used lingua franca in the world (Kachru, 1997). In other words, mutual intelligibility in communication between neighboring languages acts as a third language or lingua franca and consequently, this type of communication can be considered as an alternative to lingua franca in those particular situations.

Intelligibility level among neighboring languages is influenced by different factors, both linguistic and non-linguistic. In this respect Gooskens (2007) has introduced three factors which have an impact on determining the intelligibility level between closely-related languages, namely:

(1) The listener's attitude toward the language.
(2) The listener's previous contact with the language.
(3) Linguistic distance to the listener's language.

Of the mentioned factors, attitude toward the other language and amount of listener's contact with other languages are considered as non-linguistic factors; on the other hand, lexical and phonological distance between the two languages are considered linguistic factors which determine the strength of mutual intelligibility. One of the systematic ways of measuring linguistic distance among closely-related languages or different dialects of a language is the so-called "Levenshtein distance" (Gooskens & Heeringa, 2004), which, for the first time was used to measure the distance between Dutch, Frisian, and Afrikaans languages (Gooskens & van Bezooijen, 2006). This formula assesses the linguistic distance through assigning grades (from 0 to 1) to phonetic variation between cognates and lexical variation between aligned sentences of the two closely-related languages or different dialects of the same language. Gooskens (2007) used this criterion to analyze the mutual intelligibility of two groups of closely-related languages: namely, Scandinavian and West Germanic languages and finally, came to the conclusion that the phonetic distance rather than lexical distance more strongly predicts the degree of mutual intelligibility among closely-related languages.

Symmetric or asymmetric condition in mutual intelligibility is another aspect of receptive multilingualism which has been addressed in several studies (e.g. Impe, Geeraerts, & Speelman, 2008). This notion shows whether speakers of two closely-related languages understand each other to the same degree or not. It is worth mentioning that the understanding level of speakers of closely-related languages is not always equal, due to both linguistic and non-linguistic factors. Thus, if they have the same understanding level of each other's language the intelligibility is called symmetric; otherwise, it would be termed asymmetric. For instance, Gooskens, van Bezooijen, and van Heuven (2015), investigated the level of symmetry in mutual intelligibility between Dutch and German languages, and the results revealed the superiority of Dutch subjects in understanding the German cognates, which implies that these languages are not symmetrically understandable to the speakers of the other group.

Although the research in receptive multilingualism was initiated since 1952 by the study of Pierce on Indian languages in the United States, there are many closely related or similar languages which have rarely been studied within this framework. Among them are Turkish and Azerbaijani, the closely-related languages, with more than 100 million native speakers which make them suitable candidates to be investigated in order to bridge some gaps in the concept of receptive multilingualism. Although the research on intelligibility level among some of the closely-related languages, has received more attention than others, namely, Scandinavian languages (Gooskens & van Bezooijen, 2006; Haugen, 1966; Maurud, 1976; Zeevaert, 2004), receptive intelligibility among Turkic languages has not been paid much attention to. In this case, the experimental investigation of receptive intelligibility of Azerbaijani to Turkish speakers as carried out by Sağın-Şimşek and König (2011) is the only sample of literature which experimentally addressed the two languages. Thus, receptive intelligibility of Turkish to Azerbaijani also warrants investigation in order to get a deeper insight into the relationship of these languages. This is especially important because Azerbaijanis living in Iran speak differently from those in Azerbaijan who were the subjects of the above-mentioned study.

Azerbaijani and Turkish are typologically and geographically approximate languages within the Oguz branch of the Turkic languages. These languages due to many factors, both linguistic and non-linguistic, have been expected to enjoy a high degree of mutual intelligibility. In this regard, Öztopçu (1993) by comparing the most prominent features of Turkish and Azerbaijani including basic linguistic features like: "orthography, phonology, morphology, vocabulary and syntax," has introduced differences and similarities between these languages which all lead to an expectation of a strong intelligibility between these languages. Öztopçu (1993) also concludes that differences between the two languages are not that numerous. From a strictly linguistic point of view, it is difficult to call them separate languages; they look like dialects of each other.

In addition to similar linguistic features as a cause for raising the potential level of intelligibility level, there are also some extra-linguistic reasons which might lead to strengthening this mutual understanding. The most important source of exposure to the Turkish language are the Turkish TV programs in Azerbaijan. Turkish satellite TV programs are very popular among Azerbaijanis whether in republic of Azerbaijan or Azerbaijanis living in northwestern of Iran. This phenomenon implies the closeness of the Turkish language to the Azerbaijani language, which enables Azerbaijanis to understand Turkish TV programs, mostly without any prior education in Turkish in a way that most Azerbaijanis who know Turkish, consider Turkish TV programs either as the main or even only way for learning that language.

It needs to be mentioned that, the participants of the study are Azerbaijanis living in northwestern of Iran whose native language, contrary to Northern Azerbaijani does not have any official status. Northern and Southern Azerbaijani are considered distinct languages by the International Organization for Standardization (ISO), but in fact there is a high degree of mutual intelligibility between these varieties. As far as it is related to proximity of Turkish and Azerbaijani the similar cultural background of speakers and geographical proximity of the places where the languages are spoken are among the non-linguistic factors which strengthen the hypothesis of a great proximity between these languages. Official meetings between the political heads of the two countries of Turkey and Republic of Azerbaijan are a clear instance of how receptive multilingualism can work as a channel of effective communication even in a highly-sensitive political context. As an example the presidents of the mentioned countries speak in their own languages; these meetings always have been carried out without any translations, in a way that each interlocutor speaks in his native language. Implementation of receptive multilingualism is not restricted to political contexts. A case in point is that these two countries produce joint TV programs in some special occasions in which each channels' TV presenter speaks in his or her native language. Furthermore, Turkish TV series broadcast in Azerbaijan are in original languages and without any subtitle. All these are some good examples from the real world, which clearly illustrate a high level of mutual understanding among the speakers of these languages.

Sağın-Şimşek and König (2011), have taken the initial step in investigating the degree of intelligibility between Turkish and Azerbaijani. They analyzed receptive intelligibility of Azerbaijani to Turkish speakers, by giving listening and reading comprehension tests (in Azerbaijani) to 30 Turkish students with no prior contact to the Azerbaijani language. This investigation which by controlling the nonlinguistic elements, merely focused on linguistic dimension of receptive multilingualism, revealed that the expectation of a high degree of mutual intelligibility between Azerbaijani and Turkish which was raised by Öztopçu (1993) was not fulfilled at least on the part of the Turkish speakers.

Regarding what was mentioned, we found the receptive intelligibility of Turkish to Azeri speakers as an area warranting investigation in receptive multilingualism which can be probed to find the degree of proximity between these languages.

2. Methodology

2.1. Participants
The participants of the present study were 40 native speakers of Azerbaijani, from different Azeri speaking cities located in Northwestern Iran, with the age range from 22 to 38. As it was mentioned before, receptive multilingualism has proved to be affected by both linguistic and non-linguistic factors. Regarding the aim of this study which has been set only on the impact of linguistic factors on receptive intelligibility of Turkish to Azeri speakers, an attempt was made to control the non-linguistic factors by screening the participants for previous language contact with Turkish and attitude toward the Turkish language. Therefore, participants with either previous language contact with Turkish, or negative attitudes toward this language were excluded from the experiment through implementing an ethnographic questionnaire.

2.2. Instruments

2.2.1. Ethnographic questionnaire
As it was mentioned, in order to control the role of non-linguistic factors such as language contact and negative attitude toward the Turkish language an ethnographic questionnaire was employed to select the appropriate participants. The questionnaire contained questions about the age and place of birth of the participants and the language that they speak at home, and some complementary questions to check the language contact and language attitude, such as if the participant had been to Turkey, if they listen to Turkish music, if they have taken any Turkish language course, if they watch or listen to Turkish TV and radio programs, and finally, what their attitudes toward the Turkish language and its learning are. Thus, through discarding the participants with prior language contact and negative attitude to the Turkish language (via the mentioned questionnaire), we controlled the non-linguistic factors which could have impacted the performance of the participants in the receptive intelligibility test. The questionnaire appears in Appendix 2.

2.2.2. Stimuli and task
As it was mentioned before, the main source of exposure to the Turkish language among Iranian Azerbaijanis is TV programs in Turkish, and consequently acquisition of Turkish language among Azerbaijanis seems to be achieved almost only by watching these Turkish TV programs. Therefore, we recorded some short videos from recently broadcast Turkish TV programs as the test stimuli. Using this kind of authentic material as data elicitation material in the process of the study, not only enhances the authenticity of the test in the present study, but also is considered a simulation of the real situation in which an Azerbaijani native speaker hears and reacts to the Turkish language. Therefore, the participants were asked to watch three short video clips recorded from Turkish TV programs twice; the first time they got a general understanding and the second time the video was played with pauses in the termination points of the sentences or phrases allowing them to translate what they had heard in Azeri. After translation they were asked to indicate the words which they found problematic, or words which impeded their comprehension. These words were identified as critical (unknown) words in this study.

In order to establish systematic and rigorous criteria for assessing the intelligibility level, the produced Azeri translations by the subjects were scaled based on the degree of correctness (Sağın-Şimşek & König, 2011); so that, the correct translations were coded "2," partially correct translations were coded "1," and incorrect translations or no translation at all were coded "0."

In the second step, an attempt was made to get a deeper insight into either the sources of difficulty or the facilitating factors influencing the receptive intelligibility of Turkish to Azeri participants. The second part of the assessment was akin to "think aloud" process for identifying critical words. In this section, the subjects were asked to indicate the words which they found unknown and problematic. In this part, the proportion of known to unknown words to the majority of Azeri participants, or in other words, the percentage of critical words were used as the another factor influencing the degree of receptive intelligibility.

3. Results

In the study, the participants after taking the listening test were asked to report on their assumption of the degree of proximity of Turkish and Azerbaijani by giving a score from 0 to 100%, indicating how much they assumed to understand the videos. The results showed that they assumed Turkish to be about 50% intelligible which is so close to the real finding of the study. They were also asked to give a "yes" or "no" answer to the question of: "whether you believe the Turkish language to be an easier foreign language to learn in comparison to English or not?" In response, 38 out of 40 participants answered "yes." The results of the findings showed that the vast majority of them were of the idea that the Turkish language is similar to Azerbaijani and was considered an easy language. It is worth noting that the performance of some participants at distinguishing some particular keywords which were cognates with Azerbaijani was not that high as we had estimated, but they could get an overall understanding of the spoken language in the videos, similar to what was revealed by the study of Sağın-Şimşek and König (2011). This partial understanding and the ability in getting into the essence of the spoken Turkish language can be attributed to the similarities in the lexicon and syntax of Turkish and Azerbaijani languages. The reported comments of the participants during the study showed their attitudes toward the test and the Turkish language learning, so that almost all the participants believed that watching the TV programs in Turkish is sufficient to learn the Turkish language. They also believed that, in case of ample exposure to the Turkish TV programs, they would have performed far better on the test.

3.1. The role of critical words

As it was expected, in the translation part of the video clips, students generally reaped the highest scores in understanding the sentences with the maximal number of similar words to Azerbaijani, which are those sentences with the lowest lexical distance. Likewise, the lowest scores were achieved in understanding the sentences with the minimum number of similar words or with the highest lexical distance from their Azerbaijani equivalents. It is worth noting that in some sentences just a single unfamiliar word (critical word), in a grammatical position with strong semantic load resulted in subjects' failure to understand the relevant sentences, and sometimes they were able to guess the meaning of unknown words correctly by the use of contextual cues. The relevant statistical results of the test indicate that the more the number of critical words was in a sentence, the lower was the performance of Azerbaijani speakers in understanding the Turkish language. In a way that the Pearson correlation of (–0.61) which is significant at (0.01) level, shows significant negative relationship between the number of unfamiliar words within a sentence and understanding level of the participants in the present study.

Azerbaijani equivalents are also provided under the Turkish sentences of the videos in Appendix 1. In front of each sentence, the lexical and phonetic distance as well as the average scores which the participants achieved in the translation test are provided. It needs to be mentioned that even a short glance at the equivalent or aligned sentences reveals how lexically and syntactically similar these languages are.

3.2. The role of linguistic factors in receptive intelligibility

3.2.1. Lexical distance

Lexical distance in this study was calculated manually for each sentence, using the formula designed by Heeringa (2004) and called "Levenshtein Distance." According to the formula, in order to measure lexical distance the word pairs in aligned sentences were given points in accordance with their similarity: one point for non-cognates, half a point for a compound that is partly cognate, and zero points for cognates. Then, we calculated the correlation between lexical distances (in percent) for each sentence with the mean translation score (in percent) achieved for that sentence by the participants.

The total lexical distance (calculated through the mean of lexical distance for all sentences), was 12.45%. On the other side, in the whole text 19% of all the words were critical words which refer to the words that were not understandable to at least half of the participants. Words in this study were identified as critical for two linguistic reasons: first, because they were not part of the Azerbaijani lexicon or they did not have cognates in Azerbaijani and, second they were used in Azerbaijani in the form of an identical word or a cognate, but with a different meaning (i.e. false friends). In this regard, we had expected these critical words, at least partially, to block the flow of receptive communication. It is worth noting that we found the false-friends to have stronger negative effect on the flow of receptive intelligibility than unknown words; due to the fact that the false-friends were leading to a kind of misunderstanding which resulted from the participants' false assurance of the word meanings. In addition, compared to unknown words, in the case of a false-friend, a person is not expected to use contextual cues to guess the meaning of the word, which in turn leads to the misinterpretation of the actual meaning and consequently, the phrase or sentence in which the false-friend is located. The correlation of (−0.63) which is significant at 0.01 level, shows significant effect of critical words, including non-cognates and false friends and those cognates that were not recognizable to the majority of the participants, in decreasing the level of intelligibility.

During the test, the participants were trying to find a similar word or cognate in their native language for the words they were hearing, based on the assumption that the both languages are similar with a difference in pronunciation. This effort can be beneficial or harmful depending on how it guides the person to find the correct corresponding or cognate in his/her lexicon, or misguides the person toward a similar word in his or her native lexicon with a different meaning. In this regard, the effect of false friends as a negative factor in receptive intelligibility was significant. As an instance, take into the account the sentences: *bir köpük tabakası oluşturuyor*, (it forms a layer of foam) and *bir sümük tabakası oluşturuyor* (it forms a layer of mucus). These sentences are identical except for the words *köpük* (foam), and *sümük* (mucus). Both words, *köpük* and *sümük* are also used in Azerbaijani, but while *köpük* means "foam" in both languages, *sümük* is a false-friend between these languages; *sümük* in Turkish means "mucus," but it means "bone" in Azerbaijani. Thus, admittedly, the high difference in the mean translation score for the mentioned first and second sentences, can be attributed to the deleterious effect of *sümük* as a false-friend in the second sentence, which led all Azerbaijani participants to misinterpret or misunderstand the actual meaning. Similarly, in sentence (43) the critical Turkish word *kend* meaning "self," was mistaken for the Azerbaijani phonologically identical word of *kənd* which means "village" by most of the participants.

In conclusion, it seems that Turkish and Azerbaijani share a significant number of false-friends, but the extent or degree to which these types of words can decrease the level of receptive intelligibility depends on many factors, among are which syntactic category or the part of speech and semantic role that the word plays in a sentence. This finding has implications for those Azerbaijanis and Turks who tend to learn Turkish and Azeri by the help of their own language.

As it was mentioned before, the grammatical category of a critical word in a sentence is another factor appearing to have influenced the degree of intelligibility. In a way, critical words in the verb,

adjective, and noun position carrying a higher semantic load had the greatest negative impact on the receptive intelligibility, but those critical words with functional category or grammatical words had the minor or, sometimes even no role in deteriorating the level of understanding. According to what was observed in the present study, Turkish nouns and verbs, were the most critical syntactic categories either in facilitating or obstructing the flow of receptive intelligibility, while adjectives and adverbs should be put in the second place because the participant could guess the meaning of the these types of words. This factor can be attributed to the higher semantic load of the nouns and verbs compared to adjectives and adverbs which are used to modify the nouns and verbs. As an example in the sentence: "ekip ikinci dalış yerine gitmek için harekete geçtı," (the crew started to go to the second diving place), the word "dalış" (diving) is a noun which is not used in Iranian-Azerbaijani and due to its significant semantic load proved problematic for most of the participants. While, the unknown adjective of "ufak" (small) in the utterence "ufak balıklar falan varmış orda" (there are small fishes over there), was not that problematic in terms of overall understanding. Furthermore, as mentioned by Öztopçu (1993) "Turkish and Azeri due to a high typological approximity use mostly similar gramatical features such as agglutination, inflection of nouns for case, formation of passive, causative, reflexive, reciprocal, and negative verbal structures by suffixes, using postpositions in-stead of prepositions, having a subject–object–verb word order, modifiers preceding the modified head nouns, and absence of gender, definite article, duality, certain initial and final consonant clus-ters," these features can be considered as one of the other facilitating factors in receptive multilin-gulism between these languages. The transcripts in Appendix 1. reveal how similar word order these languages use. Threfore, due to the fact that there is no need to learn a different gramatical struc-ture, Azerbaijanis can focus mostly on learning the unknown words of Turkish to have a better un-derstanding of that language.

Another point that is worth mentioning is that the same or similar words in these languages have a broader meaning in one language and a narrower meaning in the other. These types of words also have to be categorized as "false friends." For instance, the widely used word of *yapmak* in Turkish, which means "to do, to make, to create, to practice, to build," in Azerbaijani is exclusively being used in the meaning of "baking," specially baking bread. The mentioned word which is frequently used in Turkish, as a false friend, confused most of the participants.

Although Turkish language is one of the languages which many Azerbaijanis in Iran are familiar with, the unofficial status of Azerbaijani language in Iran is one of the reasons for language attrition and slight deviation of it from its northern version which is spoken in Republic of Azerbaijan. Azerbaijani in Iran is influenced by Persian as the only official language in this country, so some Turkic words which are used in republic of Azerbaijan in northern Azerbaijani, in southern version in Iran have been replaced by Persian words. An example from the test is the word *dalış* (diving), which is used in the same form and meaning in the republic of Azerbaijan but almost not in southern Azerbaijani. Therefore, the expectation of a higher receptive intelligibility of Turkish to Azerbaijani speakers living in republic of Azerbaijan, where Azerbaijani is the official language of the country and the language of education, seems completely logical, due to relatively smaller lexical distance of that language from Turkish.

Turkish seems to have more English loanwords than Azeri, something which appeared to play a facilitating role or a shortcut for Azerbaijanis with knowledge of English in understanding the Turkish language. For example, some of the critical Turkish words which were loan words from English like *komple* (completely) and *fotoraf* (photograph), were identified by a large number of the participants who had at least a slight knowledge of English. So it seems that the participants in their first confrontation with Turkish language, applied all their linguistic knowledge derived from all other languages they knew. In addition, there were a significant number of cognate words in the test that were easily recognized by most of the participants, such as *dişler, ağız, keyfı, ikinci, açık, soğuk, balık, korku,* which some have either no or very low phonological distance or deviation from their Azeri counterparts. The identification of these keywords helped students to guess the meaning of the phrases and sentences.

It is clear that most of the critical words were problematic to the participants, because they were not a part of the Azerbaijani lexicon. However, it needs to be mentioned that synonyms of those critical words would not be necessarily unknown to Azerbaijanis. For example, the synonym of the critical Arabic-origin word of *sene* (year) in the study, is *yil* which is a cognate to Azerbaijani word of *il*. This shows even if some Turkish words are considered critical their synonyms are not necessarily unknown words for Azerbaijanis.

3.2.2. Phonetic distance

In another step an attempt was made to gage the role of phonetic distance as a factor which proved to influence the level of receptive intelligibility. In order to measure the phonetic distance of cognates of Azerbaijani and Turkish, we implemented the so-called "Levenshtein Distance" (Heeringa, 2004). The "Levenshtein" distance is measured based on the aligned words in equivalent sentences. The mentioned distance according to the formula was calculated between cognates in the test.

The "Levenstein" distance is measured based on the number of the symbols inserted, substituted or deleted, so the more the operations the more the phonetic distance will be. The following items are possible operations and their relevant distance:

- Insertions and deletions 1 point,
- identical symbols 0 points,
- substitutions of a consonant by a consonant and a vowel by vowel 0.5 point,
- substitutions of a vowel by a consonant or of a consonant by a vowel 1 point,
- Diacritics were joined with the preceding symbol, adding an extra 0.25 point.

Diacritics were joined with the preceding symbol, adding an extra 0.25 point. So, for example, the distance between [a] and [a:] was 0.25, that between[a] and [o] 0.5, and that between [o] and [a:] 0.75. The unwanted effect of word length was compensated for by dividing the total sum of costs by the number of symbols aligned. In Heeringa (2004) a more extensive explanation of the procedure is given.

Table 1 shows two cognates of Turkish and Azerbaijani and how phonetic distance is calculated according to phonetic variation.

Thus, phonetic distance for the mentioned cognates is the sum of the phonetic variation calculated for each alignment divided by the number of operations, which in the mentioned example the calculated distance is: (0.5 + 0.5)/6 = 0.16.

Azeri and Turkish use similar alphabetic sound letters or alphabetic signs which could add to the reading and writing intelligibility of these languages. Turkish and Azerbaijani alphabets are nearly identical. Turkish alphabet has 29 and Azerbaijani 32 characters. Thus, Azerbaijani has three more characters of *x*, *ə*, and *q* which do not exist in Turkish. So, while, Azerbaijanis are familiar with all the signs of Turkish, Turks need to learn these three more characters in order to be able to read Azerbaijani texts. These additional three characters correspond to particular Turkish sounds or signs, *x* for *h*, *ə* for *e*, and *q* for *k* in cognates or historically related words and also the common loanwords (Özel, 2008).

Table 1. An example of phonetic distance between two cognates.						
Alignments	**1**	**2**	**3**	**4**	**5**	**6**
TR: fırsat	f	ı	r	s	a	t
AZ: fürsət	f	ü	r	s	ə	t
Cost	0	0.5	0	0	0.5	0

In the present study, the effect of phonetic distance on the receptive intelligibility was investigated through running a correlation between the calculated phonetic distance of each Turkish and Azeri sentence and the achieved translation scores by the participants for the relevant sentence. The result of correlation was (−0.215) which is not statistically significant. Thus, regarding the low-phonetic variation between cognates in the present study which in average was (4.7%), and the low correlation between phonetic distance and translation scores, we came to this conclusion, that the role of phonetic distance between Turkish and Azeri is not as significant as that of Lexical distance in determining the degree of receptive intelligibility. In this case, it needs to be mentioned that the result of this study is not in consistency with those carried out by Gooskens (2006, 2007), in which phonetic distance rather than lexical distance was found to be the stronger predictor of receptive intelligibility between Scandinavian languages. Table 2 gives a deeper insight into the sound and phonetic variation between cognates of Turkish and Azeri.

Table 2 shows some examples of sound changes between cognates of Turkish and Azeribaijani. Turkish and Azerbaijani share a good number of cognates which differ in terms of pronunciation in one or more sounds and this variation, to a large extent is systematic and predictable. Some cognates differ in just one letter, for example, Turkish "bak"/bɑːk/ for Azeri "bax"/bɑːx/ in which velar/k/ changes to fricative/x/. However, in this study, we faced some words which despite their relatively great similarity to their Azerbaijani counterparts, were not successfully identified by most participants. Some examples from the transcripts are Turkish "defa" for "dəfə" and "aslinda" for "əslində" which implies that identification of that group of words within which Turkish vowel ɑ/a/ changes to Azeri vowel ə/æ/ is one of the most difficult categories for identification. Thus, the similarities are not all the time a guarantee for word identification.

The point which is worth mentioning is that Turkish and Azeri share a significant number of Arabic and Persian origin loanwords, but mostly with a different pronunciation. These loanwords in Azerbaijani are more similar to their original forms than in Turkish. In other words, common loan words in Turkish have gone through a more drastic deformation, maybe, due to the fact that phonological system of Turkish has fewer phonetic features compared to Azerbaijani, and consequently less compatible with Farsi and Arabic phonetic systems.

Furthermore, the stronger deformation that common loanwords of these languages in Turkish have gone through was not expected by Azerbaijani participants. This kind of variation in pronunciation is based on a rigorous criterion which makes it easier to adapt to by Azerbaijanis after a time of exposure, when they are learning or trying to understand the Turkish language.

Table 2. Types of systematic phonetic variation between cognates of Turkish and Azeri.			
Sound and letter correspondence	Azerbaijani	Turkish	English
Azerbaijani e,ə to Turkish e	Əl	El	Hand
Azerbaijani x to Turkish h	Xoş	Hoş	Nice
Azerbaijani ə to Turkish a	Təbəqə	Tabaka	Level
Azerbaijani q to Turkish k	Qal	Kal	Stay
Azerbaijani i to Turkishü	Kiçik	Küçük	Small
Azerbaijani o to Turkish a	Hovuz	Havuz	Pool
Azerbaijani b to Turkish p	Corab	Çorap	Socks
Azerbaijani c to Turkish ç	Borc	Borç	Debt
Azerbaijani o to Turkish a	Güc	Güç	Power
Azerbaijani g to Turkish k	Rəng	Renk	Color
Azerbaijani RP to Turkish PR	Körpu	Köpru	Bridge

Source: Özel (2008).

4. Discussion and conclusion

This study found that the spoken Turkish is approximately 56% intelligible to Iranian Azerbaijanis. This finding is similar to those of the study by Sağın-Şimşek and König (2011), who tried to measure the receptive intelligibility of spoken and written Azerbaijani to Turkish students. This level of intelligibility which was obtained by controlling the non-linguistic factors, is due to lexical and structural similarities of Turkish and Azerbaijani. Meanwhile, although the non-linguistic factors such as language contact and attitude play an important role in determining the intelligibility of a particular language to other one, the present study has focused only on the linguistic factors such as lexical and phonetic similarities between Turkish and Azerbaijani as the main factors which determine intelligibility level.

As a result this study could not find any significant relationship between phonetic distances of cognates and translation scores. This finding is not in conformity with that of studies carried out by Gooskens (2006, 2007) which found the phonetic distance between cognates as the main predictor of mutual intelligibility among Scandinavian languages. The low effect of phonetic distance in the present study can be explained by low-phonetic distance between cognates of Turkish and Azerbaijani in comparison to Scandinavian languages. Furthermore, this type of deviation in pronunciation between Turkish and Azerbaijani is highly rule-governed, and is based on some specific principles. In other words, pronunciation variation between cognates of Turkish and Azerbaijani (as was shown in the result section) can be categorized according to specific formulas. Thus, phonetic distance between cognates of Turkish and Azerbaijani is highly predictable and generalizable, that lend itself to a quick adaptation and acquisition by Azerbaijanis. Therefore, according to this study the main justification behind the small role of phonetic distance between cognates of Turkish and Azeri can be the easy adaptability and predictability of the phonetic variations by Azerbaijani participants.

Therefore, based on what we observed in this study, adapting to a different phonetic system of Turkish by Azerbaijanis wouldn't be a daunting task, and consequently this factor should be regarded as a facilitating factor for both understanding and learning the Turkish language by Azerbaijanis.

Turkish and Azerbaijani besides having a similar grammatical structure, have a significant lexical overlap (Bozkurt, 1992; Menges, 1995). The mentioned factors can be considered the main source of intelligibility between these languages. Based on what we found in this study the main source of unintelligibility should be attributed to lexical distance which was calculated based on the proportion of unknown words to all the words in a text. The study found a significant negative correlation between the number of unknown words and the translation scores for the sentences in the videos. On the other hand, Turkish and Azeri share a significant number of false-friends. These words even in this study were really confusing for the Azerbaijani participants. In conclusion, although these neighboring languages have a lot in common, an assumption of a high degree of linguistic intelligibility between these languages especially between Turkish and the southern version of Azerbaijani was not fulfilled in this study. Although both languages share a significant proportion of Persian and Arabic origin loanwords, Azerbaijani has been more influenced by Arabic and Persian origin loanwords (Sağın-Şimşek & König, 2011). The impact of Persian on the lexicon of southern version of the Azerbaijani which is spoken in Iran is even more significant which is the best explanation for the strong negative effect of lexical distance on the intelligibility level. The expectation of a high mutual intelligibility of Turkish and Azerbaijani is a false assumption and assuming these languages as two different dialects of a same language rather than two distinct languages seems unwarranted. However, regarding the fact that the similarities of these languages are more palpable than the differences, a study which engages both Azeri and Turkish participants in measuring the mutual intelligibility between these languages can be carried out to complete and confirm the finding of previous studies.

Author details
Mohammad Salehi[1]
E-mail: m_salehi@sharif.ir
Aydin Neysani[1]
E-mail: aydin_neysani@yahoo.com
[1] Languages and Linguistics Center, Sharif University of Technology, Tehran, Islamic Republic of Iran.

References
Bozkurt, F. (1992). Karaylar. *Türklerin Dili,* 335–362.
Braunmüller, K., & Zeevaert, L. (2001). *Semikommunikation, rezeptive Mehrsprachigkeit und verwandte Phänomene: eine bibliographische Bestandsaufnahme.* Hamburg: Sonderforschungsbereich 538.
Gooskens, C. (2007). The contribution of linguistic factors to the intelligibility of closely related languages. *Journal of Multilingual and Multicultural Development, 28,* 445–467. https://doi.org/10.2167/jmmd511.0
Gooskens, C., & Heeringa, W. (2004). Perceptive evaluation of Levenshtein dialect distance measurements using Norwegian dialect data. *Language Variation and Change, 16,* 189–207.
Gooskens, C., & van Bezooijen, R. (2006). Mutual comprehensibility of written Afrikaans and Dutch: Symmetrical or asymmetrical? *Literary and Linguistic Computing, 21,* 543–557. https://doi.org/10.1093/llc/fql036
Gooskens, C., van Bezooijen, R., & van Heuven, V. J. (2015). Mutual intelligibility of Dutch-German cognates by children: The devil is in the detail. *Linguistics, 53,* 255–283.
Haugen, E. (1966). Semicommunication: The language gap in Scandinavia. *Sociological Inquiry, 36,* 280–297. https://doi.org/10.1111/soin.1966.36.issue-2
Heeringa, W. J. (2004). *Measuring dialect pronunciation differences using levenshtein distance,* Groningen: University of Groningen.
Impe, L., Geeraerts, D., & Speelman, D. (2008). Mutual intelligibility of standard and regional Dutch language varieties. *International Journal of Humanities and Arts Computing, 2,* 101–117. https://doi.org/10.3366/E1753854809000330
Kachru, B. B. (1997). World Englishes and English-Using Communities. *Annual Review of Applied Linguistics, 17,* 66–87. https://doi.org/10.1017/S0267190500003287
Maurud, Ø. (1976). Reciprocal comprehension of neighbour languages in Scandinavia: An investigation of how well people in Denmark, Norway and Sweden understand each other's written and spoken languages. *Scandinavian Journal of Educational Research, 20,* 49–72. https://doi.org/10.1080/0031383760200104
Menges, K. H. (1995). *The Turkic languages and peoples: An introduction to Turkic studies* (Vol. 42). Wiesbaden: Otto Harrassowitz Verlag.
Özel, S. (2008). *Körpü-Köprü: Reading in Azerbaijani Using Turkish as a Bridge.* Retrieved March 2016 from: http://www.indiana.edu/~iaunrc/sites/default/files/bridge/bridge.pdf
Öztopçu, K. (1993). A comparison of modern Azeri with modern Turkish. *Azerbaijan International, 1*(3).
Sağın-Şimşek, Ç., & König, W. (2011). Receptive multilingualism and language understanding: intelligibility of Azerbaijani to Turkish speakers. *International Journal of Bilingualism.* 1367006911426449.
ten Thije, J. D., & Zeevaert, L. (2007). *Receptive Bilingualism: Linguistic analyses, language policies and didactic concepts.* Amsterdam: John Benjamin Publishing Company. https://doi.org/10.1075/hsm
Zeevaert, L. (2004). *Interskandinavische Kommunikation: Strategien zur Etablierung von Verständigung zwischen Skandinaviern im Diskurs.* Hamburg: Kovač.

Appendix 1

	Turkish sentences and their Azeri equivalents in videos	Critical words	Percentage of critical words	Lexical distance	Phonetic distance	Mean score of translation
1	TR: Ekip ikinci dalış yerine gitmek için hareket egeçtı	Dalış (diving)	12.5	0	0.082	47.5
	AZ: Qurup ikinci dalış yerinə getmək üçün hərəkətə keçdi					
	EN: The crew started to go to the second diving place					
2	TR: Cenkın keyfı yerınde gibi	Gibi (like)	33	33	.05	100
	AZ: Cenkin keyfı yerində kimi					
	EN: Cenk looks like he is feeling good					
3	TR: Balıkların bololduğu bir yeregidiyoruz		0	0	0.055	70
	AZ: Balıqların bololduğu bir yerəgedirik					
	EN: We are going to a place which is full of fish					
4	TR: Ufak balıklar falan varmış orda	Ufak (small)	20	20	0.040	39.5
	AZ: Kiçik balıqlar filan varmış orda					
	EN: There are small fishes living there					

	Turkish sentences and their Azeri equivalents in videos	Critical words	Percentage of critical words	Lexical distance	Phonetic distance	Mean score of translation
5	TR: Birbakıcaz güzellikler neler		0	0	0.011	81.25
	AZ: Bir baxarıq gözəlliklər nələr					
	EN: We will look for beautiful things					
6	TR: Ilk defa mide bulantisi çekmedim	Ilk (first)	40	0	0.072	42.5
	AZ: Ilk dəfə mədə bulanması çəkmədim	Defa (time)				
	EN: For the first time I did not get a nausea					
7	TR: Her hangi bir dalga yoktu	Hangi (which)	40	0	0.028	17.5
	AZ: Hər hansı bir dalğa yoxidi	Dalga (wave)				
	EN: There was no wave					
8	TR: Rahat rahat dalış yaptim	Dalış (diving)	50	25	0	0
	AZ: Rahat rahat dalış etdim					
	EN: I did a comfortable diving	Yap(do)				
9	TR: Ama bu sefer üşüdüm		0	0	0	98.75
	AZ: Amma bu səfər üşüdüm					
	EN: But I was cold this time					
10	TR: Su biraz soğuk ve serin idi burda		0	0	0.023	82.5
	AZ: Su biraz soyuq və sərin idi burda					
	EN: The water was a bit cold and cool here					
11	TR: Şimdi elbiseyle dalacam	Şimdi (now)	66.6	33.3	0.135	10
	AZ: Indi paltarla dalacaq	Dalacam				
	EN: I am diving with my clothes now					
12	TR: On sene sonar demek ki burada bunun yarisida kalmaya-cak.	Sene (year)	11.1	11.1	0.006	92.5
	AZ: On il sonar demək ki burada bunun yarisida qalmayacaq.					
	EN: We can say that ten years later we won't see any of these here.					
13	TR: Yirmi sene sonar bizim çocuklarımızın cocukları bunların hiç birini gormiyecekler	Sene (year)	20	30	0	90
	AZ: Yirmi il sonar bizim uşaqlarımızın uşaqları Bunların heç birini gormiyecekler	Çocuklar (children)				
	EN: Twenty years later, the children of our children will see none of them					
14	TR: Bu çok acı bir şey		0	0	0.033	100
	AZ: Bu çox acı bir şey					
	EN: This is a very painful thing					
15	TR: Cenk su altında aradığı keşfi yaşaya bilmediği için üzgündü	Aramakre (looking for)	25	38	0.035	2.62
	AZ: Cenk sualtında axtardığı kəşfi yaşaya bilmədiyi üçün kədərli idi					
	EN: Cenk regrets that he did not find the discovery he was looking for under water	Üzgün (regret)				
16	TR: Iş kitaba kaldi		0	0	0.033	100
	AZ: Iş kitaba qaldı					
	EN: He had to see the book					
17	TR: Bu bir gece fotosu olduğunu söyleyeyim	Foto (picture)	33.3	0	0.05	53.88
	AZ: Bu bir gecə fotosu olduğunu söyləyim					
	EN: I have to say that this is a night picture	Söylemek (to say)				

	Turkish sentences and their Azeri equivalents in videos	Critical words	Percentage of critical words	Lexical distance	Phonetic distance	Mean score of translation
18	TR: Bu balik şu anda uyuyor	Şu anda (at the momement)	60	20	0.087	22.5
	AZ: Bu baliq bu anda yatır					
	EN: This fish is sleeping right now	Uyumak (to sleep)				
19	TR: Gözleri açik uyuyor baliklar biliyorsunuz	Uyumak (to sleep)	20	20	0.070	80
	AZ: Gözləri açıq yatır baliqlar bilirsiniz					
	EN: You know fish that sleep with open eyes					
20	TR: Ama tükürüğünden bir mukoza oluşturuyor	Tükürük (spit)	40	20	0.031	6.25
	AZ: Amma tüpürmeğinden birmukoza təşkil edir	Mukoza (mucosa)				
	EN: But it builds a mucosa from its spit					
21	TR: Bir köpük tabakasi oluşturuyor.	Oluşturuyor (creates)	25	25	0.16	90
	AZ: Bir köpük təbəqəsi təşkil edir					
	EN: It creates a layer of foam					
22	TR: Bir sümük tabakasi oluşturuyor	Sümük (mucus)	50	50	0.16	0
	AZ: Bir sümük təbəqəsi təşkil edir					
	EN: It creates a layer of mucus	Oluşturuyor (creates)				
23	TR: Ve sadece ağzı ve dişleri açıktı dikkat edersiniz	Sadece (only, just)	11.1	12.5	0.066	82.5
	AZ: Və yalnız ağzı və dişləri açıq idi diqqət edərsiniz	etmek (to do)				
	EN: If you pay attention, just its mouth and teeth are open					
24	TR: Komple bütün vücudunu kapliyor	Komple (completely)	50	75	0	56.25
	AZ: Tam bütün bədənini örtür					
	EN: It covers all of its body	Kaplamak (to cover)				
25	TR: Çok şeker bu balik, dudaklara bakar mısınız?	Şeker (sweet)	16.6	0	0.093	72.5
	AZ: Çox şəkər bu balıq, dodaqlara baxarsınız mı?					
	EN: This fish is so sweet, do you look at the lips?					
26	TR: Balon balığının bu hale gelmesi onun korkması demektir		0	0	0.031	76.25
	AZ: Balon balığının bu hala gəlməsi onun qorxması deməkdir					
	EN: The balloon fish becomes this way when it is scared					
27	TR: Korktuğu zaman balon balığı iki derısının arasına su alıyor		0	11.1	0.044	93.75
	AZ: Qorxduğu zaman balon balığı iki dərisinin arasına su alır					
	EN: When it is afraid, the bubble fish gets water between the two skins					
28	TR: Ve şişiyor, şişince dikenleri ortaya çikiyor		0	0	0.085	63.75
	AZ: Və şişir, şişincə tikanları ortaya çıxır					
	EN: And when it swells, the thorns appear					
29	TR: Ve diğer baliklar onu ısıra bilmiyor	Isırmak (to bite)	16.6	16.6	0.047	39
	AZ: Və digər baliqlar onu dişləyə bilmir					
	EN: And other fishes can't bite it					
30	TR: Bu Alebot balığı planktonla yetişiyor		0	0	0.027	22.5
	AZ: Bu Alebot balığı planktonla yetişir					
	EN: This Alebot fish grows with plankton					
31	TR: Plangton yiyerek böyüyor		0	0	0	93.75
	AZ: Plankton yeyərək böyüyür					
	EN: It grows with plankton					

	Turkish sentences and their Azeri equivalents in videos	Critical words	Percentage of critical words	Lexical distance	Phonetic distance	Mean score of translation
32	TR: Ve dünyanın en büyükbalığı, ve en buyuk canlısı		0	0	0	92.5
	AZ: Və dünyanın ən böyük balığı, və ən böyük canlısı					
	EN: And the world's greatest fish, and the greatest living thing					
33	TR: Ama Ber yanındakı ağaca geçe bilirse		0	0	0	66.25
	AZ: Amma Ber yanındakı ağaca geçe bilirse					
	EN: If Ber can reach the next tree					
34	TR: Kalın sarmaşıkları kullanıp aşağıya ine bilir	Kullanmak (to use)	0	20	0.08	36.5
	AZ: Qalın sarmaşıqların istifadə edip aşağıya enə bilər					
	EN: He can go down using thick vines					
35	TR: İki ağaçı birleştiren kısa ve kalın sarmaşıklar geçiş yolu olacak		0	0	0.0358	70
	AZ: İki ağaçı birləşdirən qısa və qalın sarmaşıklar keçiş yolu olacaq					
	EN: The short and thick vines that join the two trees will be the passage way					
36	TR: Sarmaşıklar kolayca kopacakmış görünen bilsede aslında dayanıklı ve sağlam	Kolay (easy)	37.5	22	0.080	56.25
	AZ: Sarmaşıklar asanca qopacaqmış görünən bilsədə əslində dayanıqlı və möhkəm	Aslında (actually)				
	EN: Although the vines seem easy to break, they actually are durable and robust	Sağlam (Robust)				
37	TR: Burasi sıcak ve nemli bu nedenle su içmek fırsatı bulduğu zamanı sakın kaçırmayın	Sıcak (hot)	30.7	11	0.021	28.75
		Neden (reason)				
	AZ: Burasi isti və nəmli bu səbəblə su içmə fürsəti tapdığız zamanı əsla qaçırmayın.	Bulmak (to find)				
	EN: Never miss it when you find the opportunity to drink water because it is hot and humid here	Sakın (never)				
38	TR: Sanırım beni ısırdı, neyse ki zehirli bir yılan değıl	Sanmak (to think)	33.3	22	0.07	39
	AZ: Hesab edirəmki, məni dişlədi, nəysə ki zəhərli bir ilan deyil	Isırmak (to bite)				
	EN: I think it bit me, luckily it was not a poisonous snake	Neyse (luckily)				
39	TR: Yinede dikkatlı olmak lazım	Yine (again)	25	0	0.10	72.5
	AZ: Yenə də diqqətli olmaq lazım (dır)					
	EN: You need to be careful again					
40	TR: çünkü malum yılanlar diş fırçalamaz	Malum (clear, known)	20	0	0.052	63.75
	AZ: çünki məlum ilanlar diş fırçalamaz					
	EN: Because you know snakes don't brush their teeth					
41	TR: çinesi ile kendi vucudunu kavradı	Kendi (itsef)	50	0	0.017	0
	AZ: çinesi ilə öz vücudunu qavradı					
	EN: With his chin, he grasped his own body	Kavramak (to grip)				
42	TR: Bu yılanlar genelde ağaclarda yaşar	Genelde (usually)	25	20	0.01	76.25
	AZ: Bu ilanlar ümumiyyətlə ağaclarda yaşayır(lar)					
	EN: These snakes usually live on trees					
43	TR: Ve sincap, kuş ve yumurta ile beslenirler	Beslenmek (to eat)	33.3	0	0.035	82.5
	AZ: Və sincab, quş və yumurta ilə bəslənirlər					
	EN: And they eat squirrels, birds and eggs					
	Total mean		20.07	12.45	0.047	56.81

*ff: false friend.

Appendix 2

Ethnographic, language contact and attitude questionnaire

a. Age : ...

b. Gender: Male() Female()

c. Education: ...

d. In which city do you live? ....

e. Is your mother language Azerbaijani? YES() NO()

f. What language(s) do you speak at home? ......

g. Do you listen to Turkish music? YES() NO()

h. Have you ever been to Turkey? YES() NO()

i. Do you watch or have you watched any Turkish TV program? YES() NO()

j. Have you taken any Turkish language course? YES() NO()

k. Have you ever studied any book, or used any computer software or the internet in order to learn the Turkish language? YES() NO()

l. Do you like the Turkish language? YES, I DO() NO, I DON'T()

m. Do you like to learn the Turkish language? YES, I DO() NO, I DON'T