

ENGLISH–TURKISH COGNATES AND FALSE COGNATES: COMPILING A CORPUS AND TESTING HOW THEY ARE TRANSLATED BY COMPUTER PROGRAMS

LEVENT UZUN¹ AND Umut M. SALIHOGLU
Uludag University, Bursa
¹*ulevent@uludag.edu.tr*

ABSTRACT

Cognate status is one of the most complicated issues for those who deal with or are interested in linguistics. In the present study, we have provided a general overview related to this specific matter, and compiled a list of English–Turkish cognates and false cognates. According to the derived list, we determined that 2411 of English words, examined from among approximately 80,000 words, are either cognates or false cognates in Turkish. After determining the number of cognate and false cognate words, we tested and evaluated the correctness of the translations of three software programs and five websites that provide translation services using some of the cognates and false cognates from the derived list. Results suggest that cognate words are translated correctly in most sentences at lexical level, while false cognates and especially partial false cognates are mostly translated wrongly. Nevertheless, at sentential level, it is revealed that almost all sentences translated by computer are unsatisfactory, and need human correction.

KEYWORDS: Cognate; false cognate; bilingualism; machine translation; English; Turkish.

0. Introduction

What does *red* mean? There is certainly not only one answer to this. Actually, the correct answer depends either on the language of the interlocutors, the language(s) implied, or the meaning(s) attached to the word “red” apart from its meaning in the dictionary. In the following, the situation is exemplified:

A: What does *red* mean?

B: It is a colour.

C: It means ‘refusal/denial’.

D: It means ‘alright’.

E: It means ‘communist’.

F: It means all of the above and might have many more meanings.

In this example, the last definition is the most appropriate one, since the last response considers the word in its widest aspect. While the word *red* is a colour in English, it means ‘refusal/denial’ in Turkish, ‘alright’ in Bulgarian, and ‘communist/a person on the far left of the political spectrum’, which is a metaphoric meaning attached to the word apart from its literal meaning(s). Therefore, correct and full comprehension of a dialogue is a matter of what speakers mean or intend to mean, and what the listeners understand or perceive.

Those dealing with or interested in linguistics very often encounter lexical problems. Interference at the lexical level is probably the most problematic category for the linguist to account for (Hoffmann 1991: 99). Language learners, teachers, translators, and interpreters might in some cases feel confused while trying to convey messages from one language to another, which is a process that is realised through words.

The first aim of the present study is to explore cognate status in language processing, to offer comprehensible overview related to *cognate* and *false cognate* words, and to explain what kinds of problems they cause and why it is necessary to examine them. Our second aim is to derive a list of English–Turkish cognates and false cognates, which might be useful material for language learners and teachers, and which could be consulted by software programmers and website designers, specifically those dealing with machine and/or online translation. It might also attract the interest of researchers who investigate bilingual semantic and orthographic representations in the human mind, and how they are processed. The third goal of the study is to test and evaluate how cognates and false cognates from the created list are translated by some software programs and websites that provide translation services to determine the correctness level of the provided output.

1. Cognates and False Cognates

The degree of semantic and/or orthographic overlap between words in different languages is assumed to facilitate or interfere with the transmission of the intended messages. In cases where facilitation usually, but not necessarily always, occurs at the lexical level, researchers very frequently mention *cognates*, defined as words that possess the same or a similar form and meaning in two or more natural languages (e.g. English *butter* and German *Butter*; Russian *mope* [mɔre] ‘sea’ and Bulgarian *mope*; Turkish *asma* [asma] ‘grapevine’ and Bulgarian *асма*; Turkish *jelatin* [ʒelatin] and English *gelatine*). These kinds of words are reported to be quite common, especially when the two languages are from the same language family, or somehow related (Friel and Kenison 2001; Laufer in Schmitt and McCarthy 2001: 163; Chamizo Domínguez and Nerlich 2002). Cognates can be either *homographic* (orthographically identical) as in English–German *butter*–*Butter*, Russian–Bulgarian *mope*–*mope*, or *non-homographic* (spelled differently) as in Turkish–Bulgarian *asma*–*асма*.

On the other hand, when two words from different languages have the same or a similar form but do not share the same meaning, there occurs the case of false cognates, also recognised as *interlingual homographs*, *false friends*, *homographic non-cognates*, *pseudocognates*, *deceptive cognates*, *misleading cognates*, or *form-identical interlingual homographs* in literature (e.g. Turkish *moral* ‘spiritual state’ and English *moral*; Italian *casa* ‘house’ and Turkish *kasa* ‘safe/cash box’; Russian *стол* [stɔl] ‘table’ and Bulgarian *чол* ‘chair’; Russian *гост* [gɔst] ‘guest’ and English *ghost*; English *red* and Turkish *red* ‘refusal/denial’, and Bulgarian *red* ‘alright’).

Meara (1993) pointed out that false cognates are of interest to educators since they can cause problems for second language learners. Supportively, Friel and Kennison (2001) commented that once an incorrect association is learned, it might become harder for the learner to form the appropriate association than it would be with translations that are different in sound and appearance.

Chamizo Domínguez and Nerlich (2002) divided false cognates into two groups, namely *chance false friends* (words that are similar or equivalent in two or more languages, but without any semantic or etymological overlap) and *semantic false friends* (words that are graphically and/or phonetically similar in various languages and having the same etymological origin, but the meanings of which have diverged). They also divided semantic false cognates as *full false friends* (words the meaning of which diverge widely in various languages) and *partial false friends* (words that have several senses, some of which coincide in both languages while others do not).

All languages borrow lexical items from other codes, and have always done so (Hoffmann 1991: 101). Recently and rapidly, new terms created by modern technology are often adopted in similar form across the world languages, even though these languages might be historically or etymologically unrelated. Furthermore, the terms do not need to be always used exactly in the same way in all languages (e.g. in Turkish *kamera* does not mean ‘camera’ as in English, but ‘video camera’). Therefore, especially lately, it would not be naïve to conclude that *borrowings* (or *loan words*) constitute the main factor of similarity and/or difference among languages, which is maximised by the development of technology.

2. Studies on cognates and false cognates

Cognates and false cognates have caught the attention of researchers since the subject has significant implications for translation, interpretation, and foreign language learning and teaching. Awareness of false cognates should help to avoid misunderstandings or mistranslations. It can also help individuals to acquire a foreign language by making them more conscious linguistically. Another reason for investigating these words is that it is believed that they can help to reveal how the bilingual lexicon is organised and accessed during the processing of multiple languages (e.g. Dijkstra et al. 1998; French and Ohnesorge 1995, 1996; van Heuven et al. 1998; Dijkstra and van Heuven 2002).

Research conducted particularly on cognates and false cognates can be grouped mainly as studies that focus on linguistic issues such as *foreign language education* (e.g. Banta 1981; Johnston 1941; Scatori 1932; Talamas et al. 1999; Zamarin 1965; Malabonga et al. 2008; Hall et al. 2009; de Groot and Keijzer 2000; Lotto and de Groot 1998; Beltrán 2004–2005; Frunza and Inkpen 2007), *bilingual “machine” translation* (e.g. Inkpen et al. 2005; Mitkov et al. 2007; Chamizo Domínguez and Nerlich 2002; Ruiz et al. 2008; Nakov et al. 2007; Lalor and Kirsner 2000), and *cognition and bilingualism* (e.g. Dijkstra et al. 1999; Friel and Kennison 2001; Dijkstra et al. 2000; Dijkstra et al. 1998; French and Ohnesorge 1995; Lemhöfer et al. 2008; Beauvillain and Grainger 1987; Lemhöfer and Dijkstra 2004; Kerkhofs et al. 2005; Elston-Güttler et al. 2005; De Groot and Nas 1991; Sunderman and Schwartz 2008).

2.1. Educational studies

Educational studies concentrate on the facilitating or difficulty-inducing effects that cognates and false cognates might bring to the second or foreign language learning process. Talamas et al. (1999) noted that less fluent language learners suffered more from “form” related matters while more fluent learners suffered more from “meaning” related factors. Thus, it is possible to suggest that training learners specifically in cognates and false cognates might enhance the desired results. In fact, there is a great amount of research suggesting that subjects recall a higher percentage of cognates than non-cognates. Participants acquire cognate words in fewer sessions, and give faster responses in translating cognates than they do for other words (e.g. De Groot and Keijzer 2000; Ellis and Beaton 1993; Lotto and De Groot 1998; Tonzar et al. 2009). Friel and Kennison (2001) reported that for language learners it is easier to acquire cognates compared to other words. Likewise, Tonzar et al. (2009) concluded that the acquisition of cognates is less demanding compared to noncognate ones. In another study, Banta (1981) suggested some implications for the use of English cognates and loan words while teaching German vocabulary. He proposed five ways of organising learning material where cognate words would be included, and stressed that learners have to be encouraged towards intelligent guessing. He claimed that all vocabulary is initially passive, and it becomes active by practice. He proposed that ears and eyes trained to recognise cognates and common loan words will help brains to build new passive vocabulary more rapidly in the target language. This is important especially when considering the high number of Turkish people who live in Europe. Backus (2006) reported that there were more than 2.5 million Turks in various European countries, and that they often have problems related to language. More recently, Hall et al. (2009) demonstrated that similarity in form between a new word in a new language and a previously known language(s) is a significant contributory factor in the integration of initial memory traces into the lexical network. Therefore, it is clear that awareness of cognate status might enable learners to increase their level of readiness in foreign/

second language acquisition (FLA/SLA). Malabonga (2008) measured cognate awareness in Spanish speaking English language learners, and demonstrated that their scores were affected by first and second vocabulary knowledge. Their study provided support for cross-linguistic transfer suggesting that L1 and L2 are closely interrelated, and that this should be taken into consideration in order to gain the optimum advantage from this fact. Nevertheless, transfer from L1 to L2 may not be always advantageous. There is evidence in the literature that shows that cognate status (specifically false cognates and partial false cognates) may cause inhibitory effect on the acquisition and perception processes as many words have related meanings, but not exactly the same that can slow down the learners, mislead them and/or lead them to use those words in inappropriate contexts etc. (e.g. Meara 1993; Escribano 2004; Lerchundi and Moreno 1999; Tonzar et al. 2009). Lerchundi and Moreno (1999) found that most errors made by their students were caused by wrong interpretation of false cognates. Relatedly, Escribano (2004) indicated that if a new word in L2 is homographic to a word in L1, but with a different meaning, a misleading visual stimulus reaches the brain that afterwards results in wrong interpretation. Again, Tonzar et al. (2009) have suggested that when similarity in form is coupled with difference in meaning (the case of *false cognates*), it may not always be helpful for learners. Cognate status, at this point, plays an important role at lexical level.

2.2. Computational studies

Whether accomplished manually or with the help of a computer, translation services have been in demand from almost every field of interest in the world due to increased access to the Internet, and globalisation. The main tendency of studies that focus on bilingual (machine) translation is to put forward methods for cognate and false cognate extraction and/or classification. A great amount of work has contributed to automatic or semi-automatic cognate detection (e.g. Kondrak 2001; Kondrak et al. 2003; Kondrak 2004; Bergsma and Kondrak 2007). Mitkov et al. (2007) divided previous work into three groups as follows: *orthographic approaches* (studies that used approximate string matching to detect cognates), *phonetic approaches* (work aiming at recognising cognates based on similarity in the phonetic form of words rather than in their orthography), and *semantic approaches* (research examining similarity of meaning between phonetically and/or orthographically similar words). They have formulated new methods for automatic identification of cognates and false cognates from corpora, asserting that unlike previous work, which was based on translating co-occurrence data into a different language, their methodology required the translation of a smaller set of words to establish equivalence in a pair. In another study, Inkpen et al. (2005) proposed a method for disambiguating partial cognates between two languages with the claim that detecting the actual meaning of a partial cognate in context could be useful for machine translation and CALL tools.

2.2.1. How machine translation systems work

Machine translation (MT) models of today and the past are diverse. The working systems are very complicated, and there is need to know advanced engineering and mathematics to comprehend the underlying algorithms and formulas, which is not in the scope of the present study, and thus, will not be discussed. However, basic explanation related to the approaches that are used, and the philosophies lying under those models will be presented briefly in the following.

MT models, indeed, can be generalised in two groups: earlier systems, and modern systems. Earlier systems, also *traditional rule-based approaches*, use grammatical rules and dictionaries of the languages. In other words, they tend to match or convert the structure (e.g. subject+verb+object or other order) and vocabularies of the given languages to one another. *Direct*, *interlingua*, and *transfer* approaches are the three models of the earlier systems (see further Jurafsky and Martin 2000). *SYSTRAN* is one example of this type with a success rate of 60–70% (Wilks 2009: 65), which has served for a quite long period of time (over forty years). To sum up, earlier systems focus on the sentence structure and word meanings of the languages in question. On the other hand, modern systems, also *example-based*, *analogy-based*, *memory-based*, and *case-based approaches* (Mitkov 2003: 513) use a corpus or database of previously translated texts that are later matched or combined to determine the correct translation. Another, more recent, approach to MT, not the most favourite approach however, is *statistical approach* that again depends on a bilingual corpus, but the translation procedure depends on statistical modelling of the word order of the word equivalences of the two languages. So, what happens in this approach is mathematical estimation of statistical parameters related to the language data. In brief, modern systems tend to prefer to use corpora (written texts or word equivalences as chunks) belonging to the languages to be translated from or into. So, the philosophy here is rather deductive. One serious limitation of this approach is that it needs a very large and rich database in which it will be able to find and match the correct correspondence of the written texts in both languages. Regardless of the adopted approach, nevertheless, there are not considerable differences in the output quality of MT programs.

Kumano et al. (2002) stated that machine translation technology is currently incapable of producing translations of high quality. Although the existing software and websites with translational purposes seem successful in the translation of single words, translation of sentences and longer texts still seems to require skilled human translators. One important reason for this might be the specific context and genre of a written piece. In other words, online websites and software usually do translation without taking contextual meaning into consideration. Another reason for the lack of quality of text translation might be the existence of false cognates that mislead not only humans but also the computer. That is to say, while cognates cause hardly any problem, false cognates or partial cognates might trigger serious errors. For this reason, identifying false cognate

words of every natural language that translation is to be done from or to might help in minimizing translational errors at least at lexical level.

2.3. Studies on cognition and bilingualism

Works that study cognition and bilingualism aim at shedding light on how the human brain works when processing two or more languages. There is a large pool of research that provides evidence that linguistic transfer exists or occurs very frequently, either at semantic or orthographic level, among all languages that a person might know (e.g. Lemhöfer et al. 2008; Elston-Güttler et al. 2005a, 2005b; Schneider et al. 2002; Paribakht 2005; Gass and Selinker 1983; Tanaka and Abe 1985; Laufer 1990; Van Heuven et al. 1998; Dijkstra et al. 2000). Most studies that try to illuminate bilingual lexical representations and build a model for word recognition either provide evidence for the *language non-selective access hypothesis*, which assumes that a bilingual's lexicon for the two languages is integrated; or the *language selective access hypothesis*, which holds that only words of the targeted language are considered during communication. Measuring the reaction/response time (RT) that subjects give to cognates and false cognates is a popular method in the investigation of bilingual mapping. Another approach is to analyse subjects' translations provided for selected cognates and false cognates. Van Heuven et al. (1998) reported that bilingual word recognition is language non-selective since the activation of a word in the target language is initially affected by competing words in both languages. That is, words a person knows in L1 and in other languages influence L2 vocabulary acquisition by facilitating or interfering with it depending on similarity or difference (Laufer 1990). In addition, Dijkstra et al. (2000) reported that response times and language choice for false cognates were found to depend on their frequency in the L2 and L1. They also observed that L1 equivalent of a false cognate interferes with the simultaneous or temporal recognition of the L2 word. They reported that effects occurred bidirectionally between L1 and L2, and concluded that lexical competition between form identical false cognates is strongest since the orthographic overlap is maximal. Similarly, van Heuven et al. (1998) demonstrated that cross-language interference effects existed both within and between languages when words with the same or a similar orthography but with different meanings were taken into consideration.

To this end, it seems important that cognates and false cognates be investigated more deeply because the results might attract the interest not only of parties working in the field of FLA/SLA, but also of those who are interested in psycholinguistics, sociolinguistics, and bilingualism, as well as of those designing translational software and websites. In this sense, creating cognate and false cognate lists for more language pairs might provide researchers with richer and more varied data to work on, whether from a linguistic, social, or psychological perspective, and assist language teachers in helping their students to acquire awareness of form and meaning related issues more quickly.

Recognition and extraction of cognates and false cognates might be the initial step, after which the extracted words can be used in studies of whatever purpose they might have. The computational studies have contributed to this work by providing programs for easy detection and extraction of the words with similarity. The similarity-rating technique of De Groot and Nas (1991), and the translation-elicitation of Kroll and Stewart (1994) are two studies that have contributed to the field although they have not provided any lists of words that other researchers might desire to use and test.

3. Numbers and lists related to cognates and false cognates

Lobo (1966, as cited by Meara 1993) proposed that there were approximately 3000 English–Spanish cognates, which shows that these kinds of words are very frequent across languages. Seguin and Treville (1992, as cited by Meara, 1993) likewise, estimated that there were approximately 6500 English–French homographic cognates and 17,000 non-homographic cognates. Johnston (1941) provided a list of Spanish–English high frequency cognates, and concluded that learners whose native language is one of those, starting to learn the other language, would have an advantage of over 1000 known words. She suggested that if learners begin to read texts with easy comprehension, that is, texts with a high number of cognate words, their acquisition of structural forms would be more rapid. Another study was presented by Scatori (1932), in which he listed a number of false cognates, and postulated that nothing is more treacherous than the deceptive similarities of cognates. Welna (1977) also stated that there were false cognates in the lexicons of Polish and English, and provided a list of full false cognates.

Nevertheless, corpus linguistics is still lacking in that only a few lists of cognates and false cognates have been derived, and in only a few languages. In the present study, we aim at contributing to the field by adding a raw resource from the Turkish language that can be used in the comparison of vocabulary from other languages, but initially from English.

4. Method

Unlike the methods used in computational studies, in the present study we organised a procedure by which people are to decide whether a word sounds or seems the same, similar, or familiar to them, rather than having digital techniques or systems decide this. Our approach also diverges from the one of De Groot and Nas (1991), and of Kroll and Stewart (1994) mainly because we do not give subjects words to rate or translate, but rather ask them to extract those words themselves, from among a number of words. Friel and Kennison (2001) stated that words are generally encountered in contexts that allow people to infer their respective meanings, which cannot be denied. Nevertheless, there are a large number of people who use different kinds of “software” and “websites”

for translation, in which the effect of context is weakened, since that kind of human guessing might be absent. The words compiled in the present study can serve as a plain resource for the subsequent studies. In other words, our study aims at providing words that the studies like the ones mentioned above might wish to use.

The present research was carried out in four stages. The first stage involved teaching students how to read dictionaries, particularly the phonetic transcription of words, and informing them about cognates and false cognates. The second stage involved students' scanning of dictionaries as a project work for data collection. The third stage involved the analysis of the collected data, and formation of the list of English–Turkish cognates and false cognates. The last stage involved testing how a selection of words from the created list was translated by some software and online translation websites.

4.1. Participants

Seven hundred first grade university students, who were enrolled in the various departments of the Faculty of Education at Uludag University, took part in the study. They were native speakers of Turkish and had had formal experience in English for approximately three to five years. The proficiency level of the subjects varied between Beginner and Pre-Intermediate. They were taking English classes as part of their educational programme in return for three credits per semester. All of them participated for 10 marks out of 100 that would affect their general grade for the course. There were generally two groups, each of which had its sub-groups, classes from different departments (seven in one group and eight in the other), and these classes were also divided into nine groups according to the number of students in each class (between three and five students in each group). The number of students in classes varied between 35 and 65. The total number of student groups was 135 (15–AB, 15–C, 15–DE, 15–FGH, 15–IJKL, 15–MNO, 15–PQUVWXYZ, 15–RT, and 15–S). These groups were arranged according to the approximate number of pages that these letters of the alphabet comprised in the dictionaries, for an equal distribution of number of words to check.

4.2. Materials

Four dictionaries (*Longman Dictionary of Contemporary English*, *Oxford Advanced Learner's Dictionary*, *Cambridge International Dictionary of English*, and *Collins Cobuild English Dictionary*) were selected to be used in the project for the data collection phase. These dictionaries were examined for the number of pages they had, and were divided into nine parts where each part contained words that started with the specified letters. As a result it was determined to roughly divide the pages according to the following letters that formed a group: AB, C, DE, FGH, IJKL, MNO, PQU-

VWXYZ, RT, and S. For the analysis of the translation of words from the list, five websites and three translation programs (see Appendix 3) were used in the study.

The five websites and three software programs that we used in the study make use of the rule based approach. Nonetheless, Franz Och, a research scientist of Google, noted as follows:

Most state-of-the-art commercial machine translation systems in use today have been developed using a rules-based approach and require a lot of work by linguists to define vocabularies and grammars. Several research systems, including ours, take a different approach: we feed the computer with billions of words of text, both monolingual text in the target language, and aligned text consisting of examples of human translations between the languages. We then apply statistical learning techniques to build a translation model.

(Och 2006)

Nevertheless, at the time that the present study was conducted, the above-mentioned statistical machine translation approach of Google was available online just for Arabic–English and English–Arabic. More detailed information about the websites and software can be found following the web addresses in the chart below:

Website name	Information about the website
Google Translate	http://googlesystem.blogspot.com/2007/10/google-translate-switches-to-googles.html
Babylon	http://www.babylon.com/help/#WhatIsBabylon
Omniglot Translation	http://www.omniglot.com/links/translation.htm
Imtranslator	http://imtranslator.net/translator.asp
WebTrance	<a href="http://webtrance.skycode.com/default.asp?current=0<=en">http://webtrance.skycode.com/default.asp?current=0<=en
Babylon7	http://www.babylon.com/about/
ProÇeviri 2.0	http://www.proceviri.com/featurese.htm
Sametran-Same 1.0	http://www.sametran.com/index.php?content=ozellikler

4.3. Procedure

Initially, there were 15 classes in different departments at the Faculty of Education at Uludag University who were taking English as a required course in their educational programme. Seven of these classes were taught by one lecturer, and eight by another. The book, materials, and curriculum that they followed were the same.

(1) The following procedure was used in the study for data collection, analyses, and preparation of the lists of English-Turkish cognates and false cognates:

- During the normal hours of the course, 45 minutes were allocated to train students on the phonetic alphabet and to do exercises to ensure that students could read the phonetic transcription of words given in a dictionary.
- Both lecturers asked their students to form groups (three to five persons in each group) according to the number of students in the class, and each group was given a group of letters to examine, from the dictionaries mentioned above, from the nine groups previously determined by the researchers.
- Dictionaries were distributed in such a way that they were used more or less equally by each group of students and for each group of letters.
- The subjects were instructed to investigate the words and to note down the words that were written and/or pronounced the same as or similarly to any words in Turkish. They were also asked to write the Turkish word that they thought the English word evoked as pairs (eg. *bell–bel* ‘waist’ in Turkish; *image–imaj* [ɪmɑʒ]).
- Students examined the words in the pages of the dictionaries given to them and prepared the lists in a period of four weeks out of the classroom.
- Researchers collected, examined, analysed, and combined the lists that the subjects had prepared. The work of all groups of students who had examined the same groups of letters was gathered in a single list and compared. As there were 15 groups from each letter group, it was decided to detect first the shared words that the subjects had extracted. Therefore, two lists were to be derived: one list of higher frequency (for some examples see Appendix 1), in which words were noted by at least two thirds of the subjects (by 10 groups or more out of 15); and another list (for some examples see Appendix 2), in which there were words shared by fewer than one third of the groups (by 5 groups or less from the total of 15). All inflectional forms of a word were counted as one even though the meanings of the words might vary (e.g. *adapt/adapted/adapting* all related to the word *adapte* and its derivatives in Turkish). We assumed that the root of a word would give some clues towards inferring or guessing the meaning of the word. Nevertheless, words that were formed of two separate forms or were perceived as separate individual words in Turkish were counted separately (e.g. *thermal/thermometer/thermos/thermostat*). All proper nouns of people, places, and things were excluded from the list.

To sum up, English–Turkish cognates and false cognates were extracted from the specified four dictionaries, and a list, which was formed of two parts, was prepared. The first part of the list comprised the words that were more frequently regarded as cognates or false cognates by the subjects, and the second part contained less frequently noted words.

(2) The following procedure was used to check how a selection of words from the derived list was translated by some software and websites providing translation services:

- Words, cognates and false cognates from the first part of the list were randomly selected (10 words, 5 cognates and 5 false cognates) and sentences containing those words were written by researchers or taken from dictionaries and the Internet.
- Selected sentences were written in the provided space in the software or website and the available translate button was clicked.
- The Internet was scanned for online translation websites that served English–Turkish translation, as a result of which five websites were determined to use in the study. Additional to online translation programs, researchers searched for translation software, and three programs were detected to provide English–Turkish translation, and these were used in the study.
- Translation outputs of the eight computer programs were examined considering lexical accuracy and sentence structure in order to determine the quality of machine translation.

5. Results and discussion

After careful investigation of the vocabulary lists that students provided, two lists were derived (Appendix 1 and Appendix 2), and each list comprised separately-arranged groups of English–Turkish cognates and false cognates according to alphabetical order together with the frequency of each word. As stated before, the words in Appendix 1 were more frequently marked by students (at a rate of at least two in three), that is to say, those words were shared in at least 10 lists out of 15. Words in Appendix 2 were shared in less than 5 lists out of the total 15 (at a rate of less than one in three). Both lists were presented together with the frequency number of each word.

5.1. The list

According to the results, the first list is composed of 2018 words, 1259 of which are identified as cognates, and 759 of which are determined to be false cognates, 92 of which are partial false cognates. The second list is composed of 393 words, 28 of which are cognates, and 365 of which are false cognates. From among 365 false cognates, it was determined that 4 words were partial false cognates. These findings suggest that 2411 of English words, examined from among approximately 80000 words (approximate number of words in each dictionary used in the study), are either cognates or false cognates in Turkish. And it should be noted that all proper nouns of people, places, and things were excluded from the list, which indicates that the number would be higher indeed.

The total number of English–Turkish cognates and false cognates suggests that cognate status between the two languages is around 3%. For cognates, the rate is 1.6%, and for false cognates the rate is 1.2%. This might be interpreted that generally, in a

written text, every 3 words out of a hundred hold the potential to facilitate or interfere with lexical processing, depending on the subject, context, and lexical density.

The rate stated above can be evaluated as not high enough to cause huge conflicts or to serve important facilitation. However, if we consider a written paragraph of a hundred words, and if we assume that each sentence contains about twenty-five words, this would mean that two or three sentences might hold either facilitative or distractive potential, and it might be beneficial to take this into consideration. In cases of false cognates, a single word might lead to global misunderstanding, whereas in cases of cognates there might be important clues for guessing or inferencing. The examples below exemplify the situation:

*This **fabric** is very expensive.* – in Turkish *fabrika* means ‘factory’.

*What do you think about the **morals** of politicians?* – in Turkish *moral* means ‘morale’.

*The **bay** was looking really nice.* – in Turkish *bay* means ‘gentleman/Mr’.

*The **detective** solved the **problem**.* – in Turkish both ‘detective’ and ‘problem’ are cognates.

The words in bold are false cognates from the first list that might be associated with the words in single quotes for their meaning. Nevertheless, if the words in bold are replaced with the words in single quotes, communication might be corrupted. Although the sentences presented above were invented by the researchers, and thus might sound artificial, there are numerous similar examples in natural and authentic essays that mislead both people and translation machines. On the other hand, in the last sentence, both *detective* (*dedektif* in Turkish) and *problem* are cognates in Turkish, which might be taken as clues to guess the meaning of *solve*.

5.2. Machine translation outputs

In the following the quality and ability of machine translation will be discussed and evaluated. The translation quality of five websites and three software programs that provide translation services was tested. Ten sentences that contained cognate and/or false cognate words were randomly selected and entered in the programs to be translated. The output was compared with the correct sentences at lexical and sentential (structural and grammatical) level. The sentences and the translation done by computer are presented in Appendix 4. The overall rating of translation quality of each program is presented in Table 1 (overleaf).

In the evaluation of English–Turkish cognate status related to machine translation, we observed three main conditions: conditions where cognate words were involved and conditions where false cognate words were involved, which we will talk about separately as partial false cognates, and full false cognates. These factors will be discussed regarding both lexical and sentential aspects.

Table 1. The overall rating of translation programs regarding the lexical and sentential translation acceptability.

	Lexical Correctness										Sentential Correctness										Total Score	
	bank	camera	match	short	trash	dance	emblem	melody	park	style	bank	camera	match	short	trash	dance	emblem	melody	park	style		
Websites	1			√	√	√	√	√	√	√											7	
	2			√	√	√	√	√	√	a												7
	3			√	√	√	√	√	√	√												7
	4			√	√	√	√	√	√	√												7
	5	√				√	√	√		√												5
Software	6			√	√	√	√	√	√	a												7
	7			√	√	√	√	√	√	a					√	√	√					10
	8				√	√	√	√	√	√					√	√	√		√			10

√: correct meaning; a: acceptable (synonym used)

From the lexical correctness aspect, it was observed that from the 5 false cognate words, only 2 (*short* and *trash*) were translated correctly into Turkish, and the meaning of the other 3 (*bank*, *camera*, and *match*) words were translated wrongly. When words were examined more closely, it was revealed that *short* and *trash* were full false cognates in Turkish, each of which had one major meaning in English, and thus, one corresponding meaning in Turkish. The word *short* is the false cognate of Turkish *şort* [ʃort] ‘shorts’, and *trash* is the false cognate of *traş* [traʃ] ‘shave’. However, *bank* and *match* have different characteristics. These words have more than one dominant meaning, and only one of the senses is shared in Turkish. For instance, *bank* ‘ground near a river’ is the false cognate of *banka* ‘bank as an establishment’. Likewise, while *match* shared a sense with Turkish word *maç* [matʃ] ‘game’, the other meanings were absent, and when a partial false cognate is used in a sentence, that word should be processed together with its collocations, and considering all meanings of the word, for correct translation. Otherwise, leaving a sense of a partial false cognate out of the database of a translation program might cause ambiguities and incorrect outputs. Another important point to regard is when the meaning of a borrowed word has changed its sense in the opposite language, such as the word *camera*. In Turkish *kamera* means ‘video camera’, and the correct equivalent of *camera* in Turkish is *fotoğraf makinesi* ‘photograph machine’.

On the other hand, when cognate words were translated, it was observed that they were translated correctly or in an acceptable way, that is, by replacing the word with a

synonym. Problems related to cognate words were seen only in two conditions of Program 5, which also obtained the lowest score among the eight translation programs.

To sum up, when translation output of sentences is investigated, it seems necessary for translation program developers to study false cognates carefully, and more specifically the partial false cognates in a language.

When the translation outputs were considered at sentential level, it was observed that from the 10 English sentences that were translated into Turkish all, except sentences (6), (7), and (8) of Programs 7 and 8, and sentence (10) of Program 8 needed the correction of a human. This suggests that machine translation research has a significant way to go. Sentences processed in the programs were observed to lose their structural and grammatical consistency. The best translation program scored just 10 out of a possible 20.

In the present study, translation was tested only in one direction, that is from English to Turkish, and with short and simple sentences. Nevertheless, it would be interesting to investigate translation in both directions and with more and longer sentences in order to see what kind of syntactical and lexical error potential is hosted by translation programs. Considering the output provided by machine translation, it is possible to postulate that sentence translation requires human revision and correction in almost every condition, especially at sentential level. We observed that while there were only structural or some grammatical errors in the translation of sentences that contained cognate words, in the translation of sentences that contained false cognates the situation was more complicated. Words such as *bank*, *stress*, *match*, and *train* that were regarded as false cognates in Turkish and had more than one meaning were translated ridiculously, or processed with just one of their meanings. For instance, in *The bank of the river* the word *bank* was translated with the inappropriate meaning of the word. It was observed that the modified word *river* was overlooked in the translation. The corresponding output *Nehrin banka* would be translated as ‘Your river bank’ from Turkish to English, where *bank/banka* means an establishment performing services connected with money.

Full false cognates were revealed to be translated more correctly compared to partial false cognates. Nevertheless, they should be tested in translation in the reverse direction, that is to say, from Turkish to English. Since the aim of the present study was to reveal English–Turkish cognate status, and not Turkish–English cognate status, we have not checked the opposite conditions. However, for those who are interested in teaching Turkish as a foreign language, it might be interesting to investigate Turkish vocabulary that seems to be cognate or false cognate, and check whether machine translation operates in the same way while processing Turkish sentences.

Looking at these findings, one might conclude that partial false cognates are quite complicated, and potentially more problematic than full false cognates (words that look or sound alike in two or more natural languages, and have only one meaning in the respective languages). According to the translation output:

- (1) In the translation of sentences that contained cognates or full false cognates, there were structural and grammatical errors.
- (2) In the translation of sentences that contained partial false cognates, there were significant errors at lexical level.
- (3) Almost all sentences translated by computer needed to be corrected by a human grammatically, structurally, and lexically.
- (4) It seems that machine translation programs need to comprise all meanings of a given word, and if possible they should do so by processing not single words, but words together with their modifiers or modified vocabulary.

6. Conclusion

To sum up, unlike other cognate identifying methods described in the literature, we took advantage of subjects who were learning English as a foreign language to extract identically or similarly written and pronounced words from dictionaries. So, our list was formed of words that were evaluated by people rather than computers, and the words were considered both orthographically and phonetically. According to the present list, there are 1287 English–Turkish cognates and 1124 false cognates, 96 of which share at least one sense of meaning in each language, and thus are partial false cognates. Our resulting database of English–Turkish cognates and false cognates can be used as a resource for researchers investigating English–Turkish bilinguals, Turkish learners who study English as a second or foreign language, and machine translation development, and related issues. Based on our study, researchers may create new lists of cognates and false cognates from different languages, and findings can be compared interlingually. In other words, data coming from richer sources of languages might be applied in studies conducted with subjects of various linguistic backgrounds, which might contribute to a wide area of research, from pedagogy, linguistics, and pragmatics to sociology, psychology, and philosophy. Moreover, the approach that we used in the present study might be compared with the approach of other studies, that is to say, data coming from human reaction can be compared to data coming from formulated automatic identification processing. Additionally, textbook writers and text-translation software producers and developers might desire to take our list and databases into consideration when producing written or digital environment products.

7. Suggestions for further research

This study proposed a list of English-Turkish cognate and false cognate words. It was also examined how sentences that contained some of these words were translated from English to Turkish by computer programs. In future studies it might be interesting to ask people to translate sentences that contain cognate and false cognate words from one

language to another to see to what extent a human brain and computer translation program processing matches. It might also be interesting to test how sentences that contain cognates and false cognates are translated in the opposite direction that is to English from Turkish. Also, it is possible to prepare reading texts that contain words from the list presented here, and to investigate how they affect reading comprehension, guessing from context, and language learning or processing of a language issues.

REFERENCES

- Backus, A. 2006. "Turkish as an immigrant language in Europe". In: Bhatia, T.K. and W.C. Ritchie (eds.), *The handbook of bilingualism*. Oxford: Blackwell. 689–724.
- Banta, F.G. 1981. "Teaching German vocabulary: The use of English cognates and common loan words". *Modern Language Journal* 65. 129–136.
- Beauvillain, C. and J. Grainger. 1987. "Accessing interlexical homographs: Some limitations of a language-selective access". *Journal of Memory and Language* 26. 658–672.
- Beltrán, R.C. 2004–2005. "The effects of focus on form in the teaching of Spanish–English false friends". *RESLA* 17–18. 65–79.
- Bergsma, S. and G. Kondrak. 2007. "Multilingual cognate identification using integer linear programming". *Proceedings of the International Workshop on Acquisition and Management of Multilingual Lexicons*, Borovets, Bulgaria. 11–18.
- Cambridge International Dictionary of English*. 1995. Cambridge: Cambridge University Press.
- Chamizo Domínguez, P.J. and B. Nerlich. 2002. "False friends: their origin and semantics in some selected languages". *Journal of Pragmatics* 34. 1833–1849.
- Collins Cobuild English Dictionary*. (2nd ed.) 1995. London: Harper Collins.
- de Groot, A.M.B. and G.L. Nas. 1991. "Lexical representation of cognates and noncognates in compound bilinguals". *Journal of Memory and Language* 30. 90–123.
- de Groot, A.M.B. and R. Keijzer. 2000. "What is hard to learn is easy to forget: The roles of word concreteness, cognate status, and word frequency in foreign-language vocabulary learning and forgetting". *Language Learning* 50(1). 1–56.
- Dijkstra, T., H. van Jaarsveld and S. ten Brinke. 1998. "Interlingual homograph recognition: Effects of task demands and language intermixing". *Bilingualism, Language and Cognition* 1. 51–66.
- Dijkstra, T., J. Grainger and W.J.B. van Heuven. 1999. "Recognition of cognates and interlingual homographs: The neglected role of phonology". *Journal of Memory and Language* 41. 496–518.
- Dijkstra, T., M. Timmermans and H. Schriefers. 2000. "On being blinded by your other language: Effect of task demands on interlingual homograph recognition". *Journal of Memory and Language* 42. 445–464.
- Dijkstra, T. and W.J.B. van Heuven. 2002. "The architecture of the bilingual word recognition system: From identification to decision". *Bilingualism: Language and Cognition* 5(3). 175–197.
- Ellis, N.C. and A. Beaton. 1993. "Psycholinguistic determinants of foreign language vocabulary learning". *Language Learning* 43. 559–617.
- Elston-Güttler, K.E., T.C. Gunther and S. Kotz. 2005a. "Zooming L2: Global language context and adjustment affect processing of interlingual homographs in sentences". *Cognitive Brain Research* 25. 57–70.

- Elston-Güttler, K.E., S. Paulmann and S.A. Kotz. 2005b. "Who's in control? Proficiency and L1 influence on L2 processing". *Journal of Cognitive Neuroscience* 17(10). 1593–1610.
- Escribano, P.D. 2004. "Exploring cognition processes in second language acquisition: The case of cognates and false-friends in EST". *Ibérica* 7. 87–106.
- French, R.M. and C. Ohnesorge. 1995. "Using non-cognate interlexical homographs to study bilingual memory organization". *Proceedings of the 17th Annual Conference of the Cognitive Science Society*. 31–36.
- French, R.M. and C. Ohnesorge, C. 1996. "Using orthographic neighborhoods of interlexical nonwords to support an interactive-activation model of bilingual memory". *Proceedings of the 18th Annual Cognitive Science Society Conference*. 318–323.
- Friel, B.M. and S.M. Kennison. 2001. "Identifying German–English cognates, false cognates, and non-cognates: Methodological issues and descriptive norms". *Bilingualism: Language and Cognition* 4(3). 249–274.
- Frunza, O. and D. Inkpen. 2007. "A tool for detecting French-English cognates and false friends". Paper presented at *TALN 2007*, Toulouse.
- Gass, S. and L. Selinker. 1983. *Language transfer in language learning*. Cambridge, MA: Newbury House.
- Hall, C.J., D. Newbrand, P. Ecke, U. Sperr V. Marchand and L. Hayes. 2009. "Learners' implicit assumptions about syntactic frames in new L3 words: The role of cognates, typological proximity, and L2 status". *Language Learning* 59(1). 153–202.
- Hoffmann, C. 1991. *An introduction to bilingualism*. New York: Longman.
- Inkpen, D., O. Frunza and G. Kondrak. 2005. "Automatic identification of cognates and false friends in French and English". *Proceedings of RANLP 2005*, Bulgaria. 251–257.
- Johnston, M.C. 1941. "Spanish–English cognates of high frequency". *The Modern Language Journal* 25(5). 405–417.
- Jurafsky, D. and J.H. Martin. 2000. *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition*. Upper Saddle River, NJ: Prentice-Hall.
- Kerkhofs, R., T. Dijkstra, D.J. Chwilla and E.R. de Bruijn. 2005. "Testing a model for bilingual semantic priming with interlingual homographs: RT and N400 effects". *Brain Research* 1068(1). 170–183.
- Kondrak, G. 2001. "Identifying cognates by phonetic and semantic similarity". *Proceedings of the Second Meeting of the North American Chapter of the Association for Computational Linguistics*, Pittsburgh, USA. 103–110.
- Kondrak, G., D. Marcu and K. Knight. 2003. "Cognates can improve statistical translation models". *Proceedings of the Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics*, Edmonton, Alberta, Canada. 46–48.
- Kondrak, G. 2004. "Combining evidence in cognate identification". *Proceedings of the Seventh Canadian Conference on Artificial Intelligence*, London. 44–59.
- Kroll, J.F. and E. Stewart. 1994. "Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations". *Journal of Memory and Language* 33. 149–174.
- Kumano, T., I. Goto, H. Tanaka, N. Uratani and T. Ehara. 2002. "A translation aid system by retrieving bilingual news database". *Systems and Computers in Japan* 33(8). 19–29.
- Lalor, E. and K. Kirsner. 2000. "Cross-lingual transfer effects between English and Italian cognates and noncognates". *International Journal of Bilingualism* 4(3). 385–398.
- Laufer, B. 1990. "Words you know: How they affect the words you learn". In Fisiak, J. (ed.), *Further insights into contrastive linguistics*. Amsterdam: John Benjamins. 573–593.

- Lemhöfer, K., T. Dijkstra, H. Schriefers, R.H. Baayen, J. Grainger and P. Zwitserlood. 2008. "Native language influences on word recognition in a second language: A megastudy". *Journal of Experimental Psychology: Learning, Memory, and Cognition* 34(1). 12–31.
- Lemhöfer, K. and T. Dijkstra. 2004. "Recognizing cognates and interlingual homographs: Effects of code similarity in language-specific and generalized lexical decision". *Memory and Cognition* 32(4). 533–550.
- Lerchundi, M.A. and P. Moreno. 1999. "Los terminus latinos en los textos técnicos en inglés". In: Bocanegra Valle, A., M. Lario de Oñate and P. López Zurita (eds.), *Enfoques teóricos y prácticos de las lenguas aplicadas a las ciencias y a las tecnologías*. Cádiz: Universidad de Cádiz (CD-ROM). 309–312.
- Lobo, F. 1966. A 10,000 word Spanish vocabulary expanded from 3,000 English cognates. (Unpublished PhD dissertation, Georgetown University, Washington.)
- Longman Dictionary of Contemporary English*. (3rd ed.) 1995. Harlow: Longman.
- Lotto, L. and A.M.B. de Groot. 1998. "Effects of learning method and word type on acquiring vocabulary in an unfamiliar language". *Language Learning* 48(1). 31–69.
- Malabonga, V., D.M. Kenyon, M. Carla, D. August and M. Louquit. 2008. "Development of a cognate awareness measure for Spanish-speaking English language learners". *Language Testing* 25(4). 495–519.
- Meara, P. 1993. "The bilingual lexicon and the teaching of vocabulary". In: Schreuder, R. and B. Weltens (eds.), *The bilingual lexicon*. Philadelphia: John Benjamins. 279–297.
- Mitkov, R. 2003. *The Oxford handbook of computational linguistics*. New York: Oxford University Press.
- Mitkov, R., V. Pekar, D. Blagoev and A. Mulloni. 2007. "Methods for extracting and classifying pairs of cognates and false friends". *Machine Translation* 21(1). 29–53.
- Nakov, S., P. Nakov and E. Paskaleva. 2007. "Cognate or false friend? Ask the Web". Paper presented at *RANLP 2007*, Borovets, Bulgaria.
- Och, F. 2006. Statistical machine translation live.
<<http://googleresearch.blogspot.com/2006/04/statistical-machine-translation-live.html>>
- Oxford Advanced Learner's Dictionary*. (4th ed.). 1994. Oxford: Oxford University Press.
- Paribakht, S.P. 2005. "The influence of first language lexicalization on second language inferring: A study of Farsi-speaking learners of English as a Foreign Language". *Language Learning* 55(4). 701–748.
- Ruiz, C., N. Paredes, P. Macizo and M.T. Bajo. 2008. "Activation of lexical and syntactic target language properties in translation". *Acta Psychologica* 128. 490–500.
- Scatori, S. 1932. "Deceptive cognates in Spanish". *The Modern Language Journal* 16(5). 396–401.
- Schneider, V.I., A.F. Healy and L.E. Bourne, Jr. 2002. "What is learned under difficult conditions is hard to forget: Contextual interference effects in foreign vocabulary acquisition, retention, and transfer". *Journal of Memory and Language* 46. 419–440.
- Seguin, H. and M.C. Treville. 1992. "Les congeners interlinguaux: Un atout pour accélérer l'acquisition du vocabulaire et faciliter la compréhension des textes?" In: Courchene, R., J.I. Glidden, J. St. John and C. Therien (eds.), *Comprehension based second language teaching*. Ottawa: University of Ottawa Press. 477–487.
- Sunderman, G. and A.I. Schwartz. 2008. "Using cognates to investigate cross-language competition in second language processing". *TESOL Quarterly* 42. 19–33.
- Talamas, A., J.F. Kroll and R. Dufour. 1999. "From form to meaning: Stages in the acquisition of second-language vocabulary". *Bilingualism: Language and Cognition* 2(1). 45–58.
- Tanaka, S. and H. Abe. 1985. "Conditions on interlingual semantic transfer". In: Larson, P., E. Judd and D. Messerschmitt (eds.), *TESOL 84: A brave new world*. Washington: TESOL. 101–120.

- Tonzar, C., L. Lotto and R. Job. 2009. "L2 vocabulary acquisition in children: Effects of learning method and cognate status". *Language Learning* 59(3). 623–646.
- van Heuven, W.J.B., T. Dijkstra and J. Grainger. 1998. "Orthographic neighborhood effects in bilingual word recognition". *Journal of Memory and Language* 39. 458–483.
- Węlna, J. 1977. "Deceptive words: A study in the contrastive lexicon of Polish and English. *Papers and Studies in Contrastive Linguistics* 7. 73–84.
- Wilks, Y. 2009. *Machine translation: Its scope and limits*. New York: Springer.
- Zamarin, L.M. 1965. "Deceptive cognates in Portuguese". *Hispania* 48(4). 887–889.

Translation web sites and software:

<<http://translation.babylon.com/>>

<<http://translate.google.com/>>

<<http://www.omniglot.com/links/translation.htm>>

<<http://free-translation.imtranslator.net/>>

<<http://webtrance.skycode.com/gettr.asp?reqid=3989159837,962964846769&lang=en>>

Translation Software, Babylon 7.5.

Translation Software, ProÇeviri 2.0.

Translation Software, Sametran-Same 1.0 English Sentence Translator.

APPENDIX 1

Cognates	False Cognates	
	Partial FC	Full FC
Abacus (12)		Abbess (13) Abbey (13) Abide (15) Burro (12) Bus (13) Bush (13)
Cable (15)		Caboose (15)
Cacao (15)		Cache (15)
Dance (15)		Dank (15)
Danseuse (12)		Embassy (10)
Daphne (15)		
Emblem (15)		Fail (15)
Embryo (15)		Fall (15) Gel (15)
Fakir (15)		Gem (15)
Gelatine (15)		
Gendarme (15)		

APPENDIX 2

Cognates	False Cognates	
	Partial FC	Full FC
	Manual (9)	Mantle (5) Notice (9) Nought (5) Observe (2)
Octave (4)		Octopus (6) Peace (6) Peal (5) Quill (3) Quip (5) Quirk (5)
Sorbet (8)		

		Sore (4)
		Soul (4)
Tab (5)		Tabby (6)
	Table (7)	Utter (6)
		Uvula (9)
		Vandal (9)
		Vane (6)
		Wares (9)
		Warily (9)
		Warlike (5)

APPENDIX 3

Text Translation Programs

Web Sites	http://translate.google.com/	1
	http://translation.babylon.com/	2
	http://www.omniglot.com/links/translation.htm	3
	http://free-translation.imtranslator.net/	4
	http://webtrance.skycode.com/gettr.asp?reqid=3989159837,962964846769&lang=en	5
Soft-ware	Babylon 7 + Oxford Dictionary & Thesaurus	6
	ProÇeviri 2.0	7
	Sametran-Same 1.0 English Sentence Translator	8

APPENDIX 4

Sentences containing cognate and/or false cognate words and computer translation outputs.

* C: cognate; FC: false cognate; * italic sentences are computer outputs.

Word: Bank (FC)

Input: The bank of the river.

Should be: Nehrin kıyısı.

Output:

Translation Program 1	<i>Nehrin banka.</i>
Translation Program 2	<i>Banka nehri.</i>
Translation Program 3	<i>Nehrin banka.</i>
Translation Program 4	<i>Nehrin banka.</i>
Translation Program 5	<i>Banka ırmak.</i>
Translation Program 6	<i>Banka nehri.</i>
Translation Program 7	<i>Nehrin bankası.</i>
Translation Program 8	<i>Irmakın bankası.</i>

Word: Camera (FC)

Input: My camera is new.

Should be: Fotoğraf makinem yeni.

Output:

Translation Program 1	<i>Benim kamera yeni.</i>
Translation Program 2	<i>Benim kamera yeni.</i>
Translation Program 3	<i>Benim kamera yeni.</i>
Translation Program 4	<i>Benim kamera yeni.</i>
Translation Program 5	<i>Fotoğraf makinem yeni oluyor.</i>
Translation Program 6	<i>Benim kamera yeni.</i>
Translation Program 7	<i>Benim kamera yenidir.</i>
Translation Program 8	<i>Kameram yenidir.</i>

Word: Match (FC)

Input: How much is a box of matches?

Should be: Bir kutu kibrit ne kadar?

Output:

Translation Program 1	<i>Ne kadar maç bir kutu nedir?</i>
Translation Program 2	<i>Çok nasıl bir kasa bir maç?</i>
Translation Program 3	<i>Ne kadar maç bir kutu nedir?</i>
Translation Program 4	<i>Ne kadar maç bir kutu nedir?</i>
Translation Program 5	<i>Kutu emsallar how much?</i>
Translation Program 6	<i>Çok nasıl bir kasa bir maç?</i>
Translation Program 7	<i>Ne kadar çok maçların bir kutusudur?</i>
Translation Program 8	<i>Maçların bir kutusu ne kadardır?</i>

Word: Short (FC)

Input: He cut my hair very short.

Should be: O saçımı çok kısa kesti.

Output:

Translation Program 1	<i>O benim saçlar kısa kesilmiş.</i>
Translation Program 2	<i>Kesti benim saç çok kısa bir sürede.</i>
Translation Program 3	<i>O benim saçlar kısa kesilmiş.</i>
Translation Program 4	<i>O benim saçlar kısa kesilmiş.</i>
Translation Program 5	<i>O sac cutyor tam.</i>
Translation Program 6	<i>Kesti benim saç çok kısa bir sürede.</i>
Translation Program 7	<i>O benim çok saç kısa kesti.</i>
Translation Program 8	<i>O, çok şort saçımı kesti.</i>

Word: Trash (FC)

Input: Can you put out the trash?

Should be: Çöpü dışarı koyabilir misin?

Output:

Translation Program 1	<i>Bir çöp dışarı koyabilir miyim?</i>
Translation Program 2	<i>Siz söndürüldü çöp.</i>
Translation Program 3	<i>Bir çöp dışarı koyabilir miyim?</i>
Translation Program 4	<i>Bir çöp dışarı koyabilir miyim?</i>
Translation Program 5	<i>Bilyebiliyorsun sen söndürüyorsun süprüntü?</i>
Translation Program 6	<i>Siz söndürüldü çöp?</i>
Translation Program 7	<i>Sen çıkar çöpü?</i>
Translation Program 8	<i>Sen, çöpü söndürebilir misin?</i>

Word: Dance (C)

Input: Would you dance with me?

Should be: Benimle dans eder misin?

Output:

Translation Program 1	<i>Benimle dans mısınız?</i>
Translation Program 2	<i>Sen dans bana?</i>
Translation Program 3	<i>Benimle dans mısın?</i>
Translation Program 4	<i>Benimle dans misiniz?</i>
Translation Program 5	<i>-yeceksin sen benli dans edyorsun?</i>
Translation Program 6	<i>Sen dans bana?</i>
Translation Program 7	<i>Sen benle dans edecek miydin?</i>
Translation Program 8	<i>Sen, benimle dans edecek miydin?</i>

Word: Emblem (C)

Input: White dove is an emblem of peace.

Should be: Beyaz güvercin barışın amblemidir.

Output:

Translation Program 1	<i>Beyaz güvercin barışın bir amblemi olduğunu.</i>
Translation Program 2	<i>Güvercin beyaz bir amblemi barış.</i>
Translation Program 3	<i>Beyaz güvercin barışın bir amblemi olduğunu.</i>
Translation Program 4	<i>Beyaz güvercin barışın bir amblemi olduğunu.</i>
Translation Program 5	<i>White güvercin amblem barış.</i>
Translation Program 6	<i>Güvercin beyaz bir amblemi barış.</i>
Translation Program 7	<i>Beyaz güvercin barışın bir amblemidir.</i>
Translation Program 8	<i>Beyaz güvercin, barışın bir amblemidir.</i>

Word: Melody (C)

Input: This is a beautiful melody.

Should be: Bu güzel bir melodi.

Output:

Translation Program 1	<i>Bu güzel bir melodi olduğunu.</i>
Translation Program 2	<i>Bu güzel bir melodiden.</i>
Translation Program 3	<i>Bu güzel bir melodi olduğunu.</i>
Translation Program 4	<i>Bu güzel bir melodi olduğunu.</i>
Translation Program 5	<i>Bu güzel akort.</i>
Translation Program 6	<i>Bu güzel bir melodiden.</i>
Translation Program 7	<i>Bu güzel bir melodidir.</i>
Translation Program 8	<i>Bu, güzel bir melodidir.</i>

Word: Park (C)

Input: I go to run in the park.

Should be: Parkta koşmaya giderim.

Output:

Translation Program 1	<i>Ben parkta çalıştırmak gidin.</i>
Translation Program 2	<i>'e gidiyorum aday olmak park.</i>
Translation Program 3	<i>Ben parkta çalıştırmak gidin.</i>
Translation Program 4	<i>Ben parkta çalıştırmak gidin.</i>
Translation Program 5	<i>Ben park koşmak gidiyorum.</i>
Translation Program 6	<i>'e gidiyorum aday olmak park.</i>
Translation Program 7	<i>Ben parkta git koştum.</i>
Translation Program 8	<i>Ben, parkta çalıştırmaya giderim.</i>

Word: Style (C)

Input: I will try a new style.

Should be: Yeni bir stil deneyeceğim.

Output:

Translation Program 1	<i>Yeni bir stil çalışacağız.</i>
Translation Program 2	<i>Çalışacağım yeni bir tarz.</i>
Translation Program 3	<i>Yeni bir stil çalışacağız.</i>
Translation Program 4	<i>Yeni bir stil çalışacağız.</i>
Translation Program 5	<i>Ben yeni şıklık deniyeceğim.</i>
Translation Program 6	<i>Çalışacağım yeni bir tarz.</i>
Translation Program 7	<i>Ben dene yeni bir tarz yapacağım.</i>
Translation Program 8	<i>Ben, bir yeni stili deneyeceğim.</i>

Address correspondence to:

Levent Uzun

Uludag University

Faculty of Education

English Language Teaching Department

Bursa

Turkey

ulevent@uludag.edu.tr