

An English Word List for Turkish Academics

Türk Akademisyenler İçin İngilizce Kelime Listesi

Yrd. Doç. Dr. Meral Öztürk*

Abstract

The present study describes the compilation of reliable and pedagogically viable word lists from a corpus of 19 English proficiency tests officially administered by OSYM over a period of 10 years. The lists were intended for current and prospective academics in Turkish universities. The corpus consisted of previous ÜDS tests for social sciences and the word lists were generated using the Range Program (Heatley & Nation, 1994). The analysis revealed 4,452 different words which were later divided into 43 smaller-size lists. The lists contained an average of 99 words (between 185 and 41 words). They were graded in usefulness according to three criteria: ÜDS range, BNC frequency and ÜDS frequency. Function words and cognates were put into separate lists. Although ÜDS has been abolished and replaced by YDS in 2013, the words lists based on ÜDS are still relevant for YDS as the two tests are very similar in content measuring similar language skills. On the other hand, past YDS tests are not released in full to the public, and the published sections are not

Öz

Bu çalışmada, ÖSYM tarafından son 10 yıl içinde akademisyenlere yönelik olarak uygulanmış olan İngilizce Yabancı Dil Sınavlarında yer alan kelimelerin güvenilir ve pedagojik açıdan geçerli bir listesini oluşturmak hedeflenmiştir. Bu amaçla, 2000-2009 yılları arasında Sosyal Bilimler alanında yapılmış olan 19 adet ÜDS sınavını kapsayan bir metin bankası oluşturulmuş, ve bu metinlerin Heatley & Nation (1994) tarafından geliştirilmiş olan Range yazılımı kullanılarak taranmasıyla kelime listeleri elde edilmiştir. Söz konusu kelime listeleri toplam 43 tane olup 4,452 kelime içermektedir. Listeler kullanıcılar için daha yararlı olmalarını sağlamak amacıyla üç kriter gere göre sıralanmıştır: sınav sayısı, İngilizce'deki kullanım sıklığı ve ÜDS'deki kullanım sıklığı. İngilizce'nin gramerine ait olan kelimelerle, İngilizce ve Türkçe'de benzer olan kelimeler ayrı olarak listelenmiştir. ÜDS sınavları 2013 yılında sonlandırılmış olmakla birlikte, bugün uygulamada olan YDS sınavına içerik olarak büyük benzerlik gösterdiğinden

* Uludağ Üniversitesi, Eğitim Fakültesi, Yabancı Diller Eğitimi Bölümü, İngiliz Dili Eğitimi Anabilim Dalı.

sufficient for making reliable word lists. ÜDS lists were, therefore, suggested as usable for study towards YDS. It was cautioned, however, that the lists were mere word lists and not study lists.

Anahtar kelimeler: English Word List, ÜDS, YDS, Test Corpus, Proficiency Testing.

YDS sınavına hazırlık bakımından ÖSYM tarafından daha önce yayınlanmış olan ÜDS testleri hala geçerliliğini korumaktadır. ÖSYM'nin YDS testlerinin tamamını yayınlamama politikası (sadece %10) da geçmiş ÜDS testlerinin önemini arttırmaktadır. Bu çalışmada oluşturulan listeler YDS'ye hazırlık için geçerli olmakla birlikte, çalışma listesi olarak hazırlanmadığından doğrudan bu amaçla kullanımı uygun değildir.

Keywords: İngilizce Kelime Listesi, ÜDS, YDS, Sınav Metin Bankası, Yabancı Dil Sınavı.

Introduction

1.1. Background

The purpose of this study is to compile a word list for Turkish academics who wish to pass the official English proficiency test administered by ÖSYM. While the word list is based on a corpus of past ÜDS (Üniversitelerarası Dil Sınavı) tests, for two reasons, the list remains relevant for the YDS (the Foreign Language Test) in English which has recently been introduced to replace ÜDS as well as its close companion KPDS (Kamu Personeli Dil Sınavı). First, in spite of a reduction in the number of questions from 100 to 80, YDS remains similar in content to ÜDS and KPDS. All three tests measure similar language skills: vocabulary and grammar knowledge in English, ability to translate sentences to and from English and reading comprehension. An ÜDS-based list, therefore, is likely to serve just as well for preparing for the YDS. Second, it is not currently possible to make a satisfactory word list based on YDS due to a lack of test samples of sufficient length. Presently, only a small proportion (10%) of the content of previous YDS tests is published by the ÖSYM on their official website (www.osym.gov.tr). Therefore, the YDS candidates will have to continue using ÜDS or KPDS tests to prepare for the YDS.

YDS, like its two predecessors, is an important milestone in the careers of Turkish academics as passing the test is a prerequisite in career advancement. Those who aim to be promoted as assistant professors and associate professors have to score at least 65 % in this test. It is also a prerequisite in applications to postgraduate programs and to research assistant and lecturer positions. Although the minimum score required for application may be low (50% for postgraduate applications) much higher scores often need to be obtained to gain admission to these programs as well as to positions which are already highly competitive. Not surprisingly, thousands of people take the test each year and the number has steadily increased over the years. Table 1 shows that the number of people taking ÜDS increased by more than four times and KPDS by about five times in 6 years.

Table 1. Number of test takers in ÜDS and KPDS exams

Year	Number of test takers	
	ÜDS	KPDS
2005	44,138	58,891
2006	64,521	79,910
2007	73,170	90,763
2008	88,728	102,204
2009	116,532	141,426
2011	172,813	255,273

Data source: <http://www.osym.gov.tr/belge/1-12668/gecmis-yillardaki-sinavlara-ait-sayisal-bilgiler.html>

While these formal language tests have been introduced with good intentions for their assumed positive washback on the foreign language skills of test takers, they have also become a major obstacle in the career development of current and prospective academics in Turkish universities. The statistics provided by the ÖSYM (see table 2 below) indicate that many people fail to obtain the required scores. The mean scores obtained in the 41 administrations of the two tests between the years 2005 and 2010 do not even come close to the 65 points required for academic promotion, and only 10 of the means (25%) exceed the 55 points required for postgraduate study or for obtaining a research or teaching position. Having failed to obtain the desired scores, many academics re-sit these tests. A study by Yavuzer&Göver (2012) involving the academia at Nevşehir University indicated that 54% of the academics who participated in the survey took one or the other of these tests 4 times or more. For many academics, these tests are a major source of distress and frustration. In a study with academics in one faculty at Uludağ University, Küçüksüleymanoğlu (2007) found that those academics who scored below 50 and those who scored below 65 (i.e. between 50-64) experienced profound emotional burn-out.

Table 2. Mean scores for English ÜDS and KPDS

	ÜDS-Science		ÜDS-Health		ÜDS-Social Science		KPDS	
	Spring	Autumn	Spring	Autumn	Spring	Autumn	Spring	Autumn
2005	50.42	47.62	48.91	58.94	42.02	46.28	57.58	55.47
2006	42.86	51.26	52.01	55.58	47.65	50.44	53.84	59.39
2007	42.86	46.36	58.85	54.70	48.29	46.08	52.90	50.79
2008	48.73	49.03	54.30	55.88	45.43	46.09	56.24	57.81
2009	50.96	50.67	52.37	49.46	47.28	44.02	54.85	56.3
2010	51.23	*	48.92	*	45.87	*	54.28	*
Grand Mean	48.36		53.62		46.31		55.40	

*Data not available

One of the reasons why people fail to pass these tests appears to be a lack of adequate vocabulary knowledge in the foreign language. The participants in Yavuzer&Göver (2012) identified insufficient knowledge of vocabulary as the first and most important reason for their failure in the tests. The need for improving vocabulary knowledge for better test performance was felt and addressed by the many word lists that appeared on the internet as well as software that can be used in cell phones to practice a word list for ÜDS and KPDS (Çevik&Koçer, 2012). However, these lists suffer from a lack of adequate validation. Criteria for selecting the words are often not clear, and therefore it is not possible to judge if a given list is representative of the vocabulary of the respective test and whether it is useful for preparing for the test. Additionally, the available lists are often presented in a format that makes them quite unsuitable for vocabulary learning. Most are provided as a long unbroken list of 1000 or so words ordered alphabetically. Learning from these lists starting with the letter A and going over hundreds of words until you reach the letter Z is a long and tedious process, and learners are likely to become bored and demotivated after a while and quit study. The present study aims to overcome these shortcomings by compiling a principled and research-based word list which is graded in order of usefulness. The words were derived directly from a corpus of past ÜDS tests using special software that analyse the vocabulary of texts. Then, they were ordered in sublists of about a hundred

words each according to their usefulness for test preparation. Three criteria were used to evaluate usefulness: ÜDS corpus range, BNC frequency, and to a lesser extent, ÜDS corpus frequency.

2. Method

2.1. The Corpus

The corpus consists of 19 English ÜDS tests for social sciences administered between the years 2000 and 2009. There are two tests in the corpus for each year except the year 2000 for which there is only one test. The whole corpus consists of 126,117 running words. The tests were downloaded from the internet as PDF files and were then converted to word documents using Solid Converter PDF v6. The tests were then converted into plain text files and manually edited to enable better recognition of words by the text-analysis software. Manual editing involved the following modifications to the texts:

- a. All Turkish text was deleted.
- b. Typological errors were corrected.
- c. The Roman numeral “I” and the choice label “a” were deleted to prevent their recognition by the software as the first person pronoun and the indefinite article respectively.
- d. Hyphenated compounds were separated since the software recognizes only single forms, e.g. *camp-fires* was changed to *camp fires*, *old-age* to *old age*, *antique-looking* to *antique looking*, *sandstorm-scoured* to *sandstorm scoured*.
- e. American spellings were changed to British spellings as the software uses word lists derived from a British corpus, e.g. *installments* (Ame) was changed to *instalments* (Br), *lightcolored* to *light coloured*.
- f. Derived words that were not recognised by the software were reduced to base forms for better recognition. For this, prefixes were separated from their base forms by inserting a one-digit space, e.g. *nonacademic* corrected as *non- academic*, *overprotective* as *over protective*, *unprescribed* as *un- prescribed*. Transparent suffixes were deleted altogether when not recognised by the program, e.g. *symbolists* became *symbol*, *touristic* became

tourist, emotionality became emotional. Although the derived forms do not mean the same as the base forms, their meanings can be predicted by anyone who knows the meanings of the base form and the affix and should not be a problem in a receptive test of this kind. Thus, if one knows the meaning of *emotional*, they will be able to understand *emotionality* in the test. Therefore, these reductions do not cause valuable vocabulary to be left out (see the discussion on word family below).

2.2. Word Lists

A word is defined as a word family in the present study and includes the base form (e.g. *consider*) as well as any inflected and derived forms of the word (e.g. *considers, considering, consideration, considerations*). In the word lists, only the base form *consider* is listed as the other forms associated with this word are highly transparent in meaning when *consider* is known, and therefore they do not warrant separate study. However, in cases where a word appears in a derived rather than a base form in the corpus, the derived form was listed (e.g. *wedding* for *wed*).

Text-based Range Program (Heatley& Nation, 1994) adapted for web by Tom Cobb (http://http://www.lex tutor.ca/range/range_text/) was used to generate the word lists. This program counts how many times a word occurs in a number of texts (frequency) as well as in how many of the texts it occurs (range). In the present study, *range* was used as the major criterion in building the lists. Words that occurred in a greater number of tests were considered more useful for test takers to learn. Therefore, words were first grouped according to range and then frequency was used to order words internally in a given range group.

The highest possible range in the corpus was 19 which is the number of the ÜDS tests included in the corpus. Thus, a word with a range of 19 appeared in all of the tests in the corpus. A range value of 1 meant that a word appeared in only one of the exams. There was one list each for the ranges from 19 to 5 with the number of words in a given list ranging between 48 and 153. As there were too many words with a range of 4- 1, the words in these low range lists were further divided into smaller and manageable lists around one hundred words each and graded according to frequency in the BNC (the British National Corpus)

which is a corpus of general English. BNC frequency was taken as a measure of usefulness in grading as words that occur frequently in a general English corpus are more likely to occur elsewhere than less frequent words and would be more useful to learn. As such, although these low-range words are equally uncommon in the ÜDS corpus occurring only in a few of the past tests, those which are frequent in English have a higher likelihood of occurrence in future tests. Grading with BNC was performed using the BNC word frequency lists generated by Paul Nation (2004) and in-built in the Range program used here to compile the lists. There are 14 such lists which are ordered in frequency groups of 1000 words each. The first list (1K) contains the most frequent 1,000 words in the BNC, the second list (2K) contains the second most frequent 1,000 words, and so on.

The Range program could not recognize part of speech and did not count words of different parts of speech separately when they shared the same form, e.g. *state* as a verb and a noun. The program was not able to count multi-word units, either, and therefore counted each word in a multi-word unit such as a compound separately. For instance, the two words in *United States* were counted separately and were mixed with other uses of these words. Some words were not recognized by the program probably because they were low frequency such as *kinetoscope*, *monolingual*, *perimetre*, and had to be put into a separate list (i.e. Off-list words). Proper names except names of countries and affixes were not counted. There were separate lists for function words and cognates.

2.3. Function words

Function words were deleted from the range lists (except range 1 list) and put into a separate list. These words have little meaning and are mainly used to indicate grammatical relations in sentences. They belong to categories like articles, demonstratives, auxiliaries, prepositions, conjunctions, etc. As they are often dealt with in grammar books, it was considered not appropriate to include them in a word list designed for vocabulary study as they require a different kind of attention than other words. It should be noted, however, that function words are important to learn as they are very frequent in the language. They were also very frequent in the present corpus. They were used 324 times each on average and this is far more frequent than any other content word in the corpus (the three exceptions are

people (451), *make* (392), *country* (328). Most function words are far more frequent than the average and *the*, which is the most frequent function word in the corpus, occurs 7985 times. For this reason, it was thought useful to present function words in a separate list so that they could be paid a more focused attention.

Two lists of English function words were consulted for identifying function words (Higgins & Higgins, 2012; Gilner & Morales, 2005). Neither list is complete in the sense that they do not contain all of the function words in English. For this reason, a number of words that felt like function words could not be located in the above function word lists. These were included as function words if they were a part of speech other than noun, verb, adjective, or manner adverb which are often regarded as content word categories (e.g. *lest* (conjunction), *whereby* (relative pronoun), *prior* (preposition)). Others were included on the basis of their formal and functional similarity to those already in the lists (e.g. *upwards* (similar to *up*), *just* (to *already*)). A few other words were also included if the dictionary definition (OALD) involved a description of the word's function rather than meaning saying "indicating..." or "used to..." (e.g. *otherwise*, *indeed*, *perhaps*). Some function words could not be included even though they were cited in the English function word lists. These function words were homonymous and also used as content words. The Range program cannot count the range of these two uses separately. Therefore, they were left in the range lists so as not to eliminate important content words from study lists. These included both single words (*please*, *own*) and multi-word function words: (e.g. *of course*, *in case*, *as regards*).

2.4. Cognates

Words which are cognates in Turkish were also omitted from the range lists (with the exception of range 1 list) and put into a separate file. Words in the lists were identified as cognates if they were orthographically similar to their translation equivalents in Turkish. Thus, the words *police*, *traffic*, *telephone*, *terror* were identified as cognates as their equivalents in Turkish were spelled similarly, i.e. *polis*, *trafik*, *telefon*, *terör*. The meanings of these words are highly predictable and do not require previous learning for comprehension. When test takers come across with them in a foreign language test albeit for the first time, they will be able to successfully guess their meaning. Therefore, they were taken out to prevent their cluttering the lists. The list of cognates is for reference only and is not meant for study.

There were a number of cognates which could not be included in the list of cognates. Some of these were low in orthographic similarity such as *clay* (*kil*), *luxury* (*lüks*), *budget* (*bütçe*). EFL learners often fail to recognise such words as cognates to the effect that they remain as unpredictable as any other non-cognate word (Nagy et.al., 1993). Orthographic similarity was checked when in doubt with the BI-SIM string comparison method (Kondrak, 2005) using the web interface designed by Bhargava at <http://www.cs.toronto.edu/~aditya/strcmp2/>. This method involves a comparison of all pairs of adjacent letters (bi-gram comparisons) in two orthographic strings (an English word and its Turkish equivalent in this case). It yields an index of similarity between 0 and 1. Values below 0.50 meant less than 50% similarity and were not considered acceptable in the present study. The examples cited above had the following BI-SIM values after normalising for word length: *clay-kil* (0.13), *luxury-lüks* (0.25), *budget-bütçe* (0.29).

Many of the cognates had to be excluded from the cognates list for failing to display sufficient semantic similarity. The word *office*, for example, corresponds to the Turkish *ofis* in the sense of a “place of work” and is a close cognate in this sense. However, the word is also used in the ÜDS tests in the sense “a political position”, which is not shared by the word *ofis*. Similarly, *show* and *şov* were not included as cognates because *show* is used both as a verb and a noun but the Turkish word is similar in meaning only to the noun, but not to the verb. Other examples are *announce-anons* (*anons* has the meaning “calling out one’s name”, but not “declare”), *major-majör* (*majör* is used in music only), *address-adres* (*adres* does not have the sense “to address an issue” or “a talk”), *goal* (*gol* is restricted to football, and does not have the meaning “purpose”), *record-rekor* (*rekor* does not have the “album” sense), *firm* (*firma* does not have the meaning “strong”). Semantic similarity was checked using concordances that displayed words ordered alphabetically in contexts derived from the ÜDS corpus. A word was excluded from the cognates list if concordances yielded uses which were not covered by the Turkish cognate. The concordances were created using the text concordancer at http://www.lex Tutor.ca/concordancers/text_concord/. As the corpus was too large for the program it was divided into three parts and submitted separately.

Cognates were also excluded if it had more than one Turkish equivalent of which only one was cognate and if the cognate was less frequent than the alternative Turkish equivalent, e.i. *data* vs *veri* for *data*, *relaks* vs *rahatlama* for *relax*, *skor* vs *puan* for *score*, *otorite* vs *yetki* for

authority, legal vs yasal for legal. The few false cognates with similar form but quite different, often misleading meanings were also excluded, e.g. *sympathy* (“*understanding*”)-*sempati* (“*cute*”).

3. Results and Discussion

There were 4,452 words in total divided more or less evenly across 43 lists, which can be accessed at <http://uludag.academia.edu/MeralOzturk/Papers>. In most lists (31) there were less than a hundred words, while the longest list (R1-3K list, a sublist of Range 1 list) contained 185 words and the shortest (R1-K1 list) contained 41 words. These list sizes were considered reasonable in terms of study time and, with regular study, words in a given list can be mastered over a couple of weeks, which is likely to create a sense of achievement and motivate further study of words in other lists.

Table 3 presents the distribution of words across range categories. These results indicated that only a small proportion of the words (2%) occurred in all 19 tests. However, these words seem to be repeated many times and account for 10% of all occurrences in the corpus. On the other hand, nearly half of the words (47%) appeared only in one or two of the tests with very few repetitions as they accounted for only 3% of all occurrences. In other range categories, both the number and frequency of words are small and stable across increasing range categories.

Although there are around 300 function words in English, only 188 function words occurred in the ÜDS corpus. These words were used widely in the ÜDS tests. 87% (=161) of the function words occurred in 10 or more tests. They are also very common in general English. 84% (157) of them are among the most frequent 1,000 words of English according to word lists based on the British National Corpus (Nation, 2004), and only 5 words are used less frequently than the most frequent 3,000 words (i.e. *billion, prior, hence, moreover, lest*).

All in all, there were 301 words in the cognates list. These words had a total frequency of 6308 and made up 6% of all tokens (i.e. running words) in the corpus. As these can be assumed already known to the test takers, they will provide considerable advantage with the vocabulary of ÜDS. More advanced test takers may also take advantage of the less transparent cognates as increased proficiency tends to be associated with better recognition

of cognates (Molnar, 2010). Cognates appeared in all ranges, but 31 % (71 cognates) appeared in 10 or more of the ÜDS tests in the corpus. This suggests that their usefulness is not confined to a couple of tests, but most will be useful in a number of different tests. 40 % of them are among the most frequent 2,000 words of English according to the word lists based on the British National Corpus, and therefore their usefulness goes beyond the ÜDS tests.

Table 3. Number and percentage of words in range categories

Range	Number of Words	% of Words	Frequency	% Frequency
19	89	0.02	11,354	0.10
18	55	0.01	3,923	0.03
17	49	0.01	2,892	0.03
16	60	0.01	2,858	0.03
15	60	0.01	2,393	0.02
14	48	0.01	1,692	0.02
13	65	0.01	1,972	0.02
12	55	0.01	1,507	0.01
11	71	0.02	1,635	0.01
10	80	0.02	1,666	0.01
9	105	0.02	1,854	0.02
8	95	0.02	1,297	0.01
7	131	0.03	1,596	0.01
6	110	0.02	1,074	0.01
5	153	0.03	1,183	0.01
4	236	0.05	1,436	0.01
3	341	0.08	1,450	0.01
2	610	0.14	1,640	0.01
1	1,486	0.33	2,159	0.02
Function Words	188	0.04	60,845	0.54
Cognates	301	0.07	6,308	0.06
Off-list	65	0.01	-	-
Total	4,452		112,734	

In the low range categories, the division of words into BNC frequency levels revealed a concentration on the high frequency levels (cf. table 4). It turned out that 26% of words in Range 1 were in the first three frequency levels of the BNC, and the percentage increased with range (i.e. 38% in Range 2, 54% in Range 3, and 63% in Range 4). High frequency words are important to learn as they occur in a wide variety of texts and make up about 80% of words in a typical text (Nation, 2001, p.17). Direct study of these words from a list will, therefore, be beneficial. The best strategy with less frequent words, however, will be contextual learning in combination with a dictionary. Unknown words encountered when answering previous tests or reading similar academic texts can be looked up in a dictionary and checked against the context and / or noted down for future practice.

Table 4. BNC frequency levels of range 1 and range 2 words

BNC Frequency Level	Range 1	%	Range 2	%	Range 3	%	Range 4	%
1K	41	0.03	37	0.06	37	0.11	28	0.12
2K	123	0.08	98	0.16	79	0.23	74	0.31
3K	185	0.12	135	0.22	70	0.21	46	0.19
4K	154	0.10	86	0.14	63	0.18	35	0.15
5K	174	0.12	48	0.08	32	0.09	23	0.10
6K	130	0.09	55	0.09	20	0.06	13	0.06
7K	129	0.09	32	0.05	12	0.04	9	0.04
8K	108	0.07	40	0.07	7	0.02	3	0.01
9K	101	0.07	23	0.04	3	0.01	1	0.00
10K	98	0.07	8	0.01	5	0.01	-	-
11K	70	0.05	20	0.03	5	0.01	-	-
12K	50	0.03	8	0.01	1	0.00	-	-
13K	52	0.03	9	0.01	1	0.00	2	0.00
14K+	71	0.05	11	0.02	6	0.02	2	0.00
Total	1,486		610		341		236	

Conclusion

The present study used corpus analysis techniques to compile reliable and representative word lists based on an extensive ÜDS corpus and generated lists which were more

comprehensive and more pedagogically oriented than previous lists of its kind. The present list contains 4,452 words in total which is, to the author's knowledge, more extensive than any other list compiled before. While previous lists present words in a single unbroken list of words ordered alphabetically, the present study presents them in smaller size sublists, which lends them more convenient for study purposes. The sublists are graded according to several usefulness criteria (i.e. range, BNC frequency and ÜDS frequency) to allow candidates to start with the most useful words and progress from more to less useful words. The 4,452 words of the ÜDS were divided into 43 lists ordered both externally and internally in usefulness. Externally, the lists were first ordered according to the spread of words over the 19 ÜDS tests (range) from words which occurred in all of the tests to those which occurred only in one, and then lower range lists were divided into sublists which were ordered according to frequency in the BNC corpus. Internally, each list was ordered according to frequency in the ÜDS corpus from the most to the least frequent.

It must be noted, however, that the present lists are only word lists and not study lists. The words cannot be directly studied from the lists as no further information about the words such as their meaning or parts of speech have been provided. It must also be noted that mastery of the words in the present lists do not automatically lead to success in the test since the test measures language proficiency which involves other knowledge types than vocabulary knowledge alone. However, improving vocabulary knowledge is an important step towards gaining higher proficiency in the language, and therefore should contribute towards higher scores in the test.

References

- Cobb, T. *Text-Based Range*. Accessed on 27 March 2013 at http://www.lexutor.ca/range/range_text/, an adaptation of Heatley & Nation's (1994) *Range*.
- Çevik, K.K. & Koçer, H.E. (2012). Mobil cihaz tabanlı yabancı dilde kelime öğrenme uygulaması. *Journal of Technical- OnLine*, 11(2), 60-70.
- Gilner, L. & Morales, F. (2005). *The English function words set*. Accessed on 13 January, 2013 at <http://www.sequencepublishing.com>
- Heatley, A. and Nation, P. (1994). *Range*. Victoria University of Wellington, NZ. Accessed on 7 February, 2012 at <http://www.vuw.ac.nz/lals/>.

- Higgins, J. & Higgins, M. (updated 2012). *Function words*. Accessed on 13 January, 2013 at <http://myweb.tiscali.co.uk/wordscape/museum/funcword.html>
- Kondrak, G. (2005). N-gram similarity and distance. Proceedings of the Twelfth International Conference on String Processing and Information Retrieval (*SPIRE 2005*), Buenos Aires, Argentina. pp. 115-126. Accessed on 29 March, 2013 at <http://webdocs.cs.ualberta.ca/~kondrak/papers/spire05.pdf>
- Küçüksüleymanoğlu, R. (2007). Eğitim Fakültesi öğretim elemanlarının tükenmişlik düzeyleri. *Eurasian Journal of Educational Research*, 28, 101-112.
- Molnar, T. (2010). Cognate recognition and L3 vocabulary acquisition. *Acta Universitatis Sapientiae, Philologica*, 2 (2), 337-349.
- Nation, I. S. P. (2004). A study of the most frequent word families in the British National Corpus. In Bogaards, P. & Laufer, B. (Eds.), *Vocabulary in a Second Language: Selection, Acquisition, and Testing*. Amsterdam: John Benjamins. pp. 3-13.
- Nagy, W. E., Garcia, G. E., Durgunogulu, A. Y., & Hancin-Bhatt, B. (1993). Spanish-English bilingual students' use of cognates in English reading. *Journal of Reading Behavior*, 25, 241-258.
- Nation, I.S.P. (2001). *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.
- Yavuzer, H. & Göver, İ. H. (2012) Akademik personelin yabancı dil durumu ve yabancı dil sınavlarına bakışı: Nevşehir örneği. *NEÜ Sosyal Bilimler Enstitüsü Dergisi*, 1, 136-158.

