

Sebastian Moya

Thesis Advisor: Ted Fernald

12–9–08

## The History of Arabic Loanwords in Turkish

### 0. Introduction

The Turkish language provides a unique insight into the outcomes of language contact and language reform. Beginning in the 11<sup>th</sup> century and continuing through the 19<sup>th</sup> century, Turkish, specifically as used by the elites of the reigning Ottoman court, underwent a long, intense period of contact with Arabic and Persian which, as the languages of the Islamic religion, were established by the Ottomans as the model on which to base their own emerging official language. However, when the political situation changed, most radically in the 1923 establishment of the Turkish Republic, the language attitudes changed as well, and the once revered Arabic and Persian linguistic influence was quickly transformed into a common enemy for the newly united Turkish peoples to rally against. Because of the early start to both the initial period of contact and the subsequent period of reform, which definitively ended in the 1970s, the case of Turkish also gives us a certain distance from which we can more effectively evaluate the outcomes of the prolonged periods of borrowing and purification.

Taking advantage of the long history of this specific case, this thesis will survey the history specifically of Arabic loanwords in Turkish. Questions from along every step of this historical process will be addressed. Beginning in the Ottoman period, we will explore how the borrowing among Ottoman elites and the Turkish masses compared to gain insight into the mechanisms by which lexical borrowing happened in both cases. We will also discuss the interface of linguistic features specific to Turkish and Arabic, vowel harmony and templatic morphology respectively, and how these features influenced the results of borrowing. Moving on to the period of language reform, we will explore the policies through which the newly formed Turkish Republic sought to purify its language. Ultimately,

we will address the consequences of these reform efforts, placing them into the larger context of contact linguistics. The structure of this paper is thus as follows:

1. Ottoman Turkish Period
2. Borrowings
3. Language Reform
4. Results of Language Reform

Ultimately then, this thesis seeks to answer the questions of how Arabic loanwords behaved upon entering the Turkish language and how they have responded to the official efforts to rid the language of them.

An important note throughout this paper is the use of Ottoman and Turkish to refer to two distinct languages or linguistic registers. In this paper, Ottoman, or Ottoman Turkish, will be used to refer to the official language of the Ottoman court with its extensive lexical and structural borrowings from Arabic and Persian, whereas Turkish will be used to refer to the contemporaneously existent language of the masses, which was to a much larger degree free of foreign influence.

## 1. Ottoman Turkish Period

The ancestors of today's Turks populated present-day Turkey during the 10<sup>th</sup> century and had overwhelmingly converted to Islam by the early 11<sup>th</sup> century (Cücelglu, Slobin 1980). The Turks' first exposure to Islam likely came through the Persians, as evidenced by the Persian-language origins of the most basic religious terms in Turkish. Nonetheless, Arabic, as the language of the Koran, also exerted a powerful influence on Turkish and Ottoman both (Lewis 1999).

The Ottoman principality, which was founded in 1300 amid other largely independent principalities in the lands that would later become part of the Ottoman empire, only began to reach its

eventual imperial levels of power starting in the early 13<sup>th</sup> century (Kerslake 1998). All through this time, as Lewis puts it:

The perception that they were Turks was supplanted by an awareness that they were members of the *"Ummet-i Muhammed*, the Community of Believers, so the tide of Arabic and Persian flowed. (Lewis 1999: 5)

As members of this overarching Islamic community, the Ottoman ruling classes modeled their nascent courtly life on Islamic high culture (Kerslake 1998). This meant that both the Arabic and Persian languages were repeatedly referenced during the development of the official Ottoman forms of religion, government, and the arts (Kerslake 1998). This resulted in cultural adoptions such as the Persian model for Ottoman poetry as well as the use of the Arabic script in the Ottoman court (Lewis 1999). However, members of the Ottoman court also borrowed grammatical features directly from the languages themselves.

Some degree of borrowing from Arabic and Persian existed among all levels of the Turkish population, even in the poorest, least educated communities. However, in these cases Arabic and Persian borrowings were limited to relatively small amounts of lexical items, primarily new religious terms to supplement their adoption of Islam (Kerslake 1998). This kind of borrowing is representative of most cases of lexical borrowing, with the receiving speech community not achieving fluency in the language being borrowed from (Winford 2003). Among the learned Ottoman class however, fluency and literacy in Arabic and Persian were common (Kerslake 1998). Thus, Ottoman elites not only borrowed many more lexical items from Arabic and Persian than Turkish commoners, they also incorporated morphological features, such as plural and gender-markers (Kerslake 1998), and syntactic structures, including prepositional phrases (Kerslake 1998) and Indo-European-style relative clauses

(Appel, Muysken 1987) into the language of the court. All this layering of Arabic and Persian linguistic features and vocabulary over the Ottoman-era variety of Turkish resulted in a variety of linguistic registers which, at their least transparent, required fluency in Arabic and Persian in addition to Turkish to be understood (Kerslake 1998). These latter registers were of course exclusive to the Ottoman elites who had the ability to study Arabic and Persian.

The borrowing from Arabic and Persian was overwhelmingly driven not necessarily out of a need for the borrowed structures or vocabulary, though some of the borrowings did in fact fill gaps in the Turkish vocabulary or grammatical repertoire; instead they were borrowed because of the prestige associated with the Arabic and Persian languages because of their status as the languages of the Turks' new religion, Islam. As evidence of the prestige-driven nature of Ottoman borrowing from Arabic and Persian, there were several instances of borrowing of lexical items where commonly used equivalents already existed in Turkish (Lewis 1999). In fact, the massive adoption of Arabic and Persian lexical items often resulted in the decreased use and eventual disappearance of Old Turkish vocabulary (Kerslake 1998). These Arabic and Persian dominated registers, known as *inʃa*, continued to flower throughout the 16<sup>th</sup> and 18<sup>th</sup> centuries. Additionally, towards the end of the Ottoman period in the 19<sup>th</sup> century, new Ottoman vocabulary for political and scientific terminology was being created using Arabic vocabulary in conjunction with Persian syntactic rules (Kerslake 1998).

The fact that members of the Ottoman ruling class were frequently trilingual in Arabic, Persian and Turkish casts some questions onto whether their use of Arabic and Persian elements in their speech constitutes a case of fully-incorporated borrowing or, alternatively, a case of code-mixing. The difference between the two, as laid out by Appel and Muysken (1987), lies in the integration of the foreign lexical items into the recipient language's lexicon, a defining characteristic of borrowing.

Code-mixing, on the other hand, is described as the use of two separate and unassimilated systems in the same utterance. Though the difference between these phenomena is clear enough in theory, they are hard to differentiate in practice because similar processes of adaptation of foreign words are used in both cases (Winford 2003). Additionally, since the full incorporation of lexical items into the receiving lexicon is a gradual process, the line between stable situations of separate codes being mixed and cases of ongoing but still incomplete lexical borrowing is easily blurred (Appel, Muysken 1987).

The small portion of Arabic and Persian loanwords which were shared across all socio-cultural levels of the Turkish-speaking population are prime candidates for fully borrowed lexical items. This can be argued because of the fact that these were introduced in large part to provide words for new concepts, primarily those associated with the newly adopted Islamic religion. The fact that these more widely used loanwords were among the first to be borrowed also makes it more likely that they have been incorporated into Turkish more fully, as opposed to merely coexisting in a separate system alongside Turkish.

Another excellent heuristic in determining the degree of incorporation of loanwords into Turkish is whether these loans were used along with Turkish suffixes. In Turkish, which is an agglutinative language, suffixes are the primary markers of morphological information, concatenated onto the root stem with each morpheme still remaining easily distinguishable in the final product (Lehmann 1973). As will be discussed in Section (2), Zimmer (1985) provides an example of an Arabic-origin loan occurring with a native Turkish suffix. This is a good indication that the Arabic and Persian loans in use in modern Turkish have by this stage been fully incorporated into the Turkish lexicon. However, whether these same loans were used with native Turkish suffixes during the Ottoman period is another question altogether. Again, the kind of speech community using the

borrowed words is key in answering this question. The loans in use throughout the entire Turkish speech community are more likely to have been used with Turkish suffixes because, as Winford (2003) claims, it is in cases like these of less intense language contact that borrowed words are most fully adapted into the receiving language's phonology and morphology. Additionally, since most of the people who used these lexical items were not fluent in either Arabic or Persian, they had no pretensions of maintaining the words in their original form.

Among the Ottoman-era commoners, then, lexical items of foreign provenance were more likely to have been incorporated into their native Turkish lexicon. On the other hand, the Ottoman elite who were fluent in Arabic and Persian and considered these the prestigious languages surely combined both full lexical borrowing and code-switching in their speech. Nonetheless, even if certain foreign words were not fully borrowed and incorporated into the Ottoman lexicon, that does not undo the influence of their continued use in the official language and writings of the Ottoman court.

Though this thesis will be focused primarily on the lexical borrowings from Arabic to Turkish, Arabic and Persian both influenced more than just the Turkish lexicon during the Ottoman period. Indeed, the adoption of Arabic and Persian lexical, morphological, and syntactic elements into a complex, high-register Ottoman Turkish was so vast that at least certain linguistic styles of the court would have been unintelligible to the monolingual Turkish commoners (Cücelglu, Slobin 1980; Kerslake 1998). Additionally, this outcome of the protracted language contact between the three languages was also part of the impetus for the Turkish Republic's language reform.

## 2. Borrowing

Borrowing is a common outcome of prolonged language contact in which elements of one language are incorporated into another language; these elements can range from the level of individual

sounds to larger morphological or syntactic structures. Although, as discussed earlier, Ottoman Turkish incorporated lexical units as well as grammatical practices from both Arabic and Persian, the primary focus in this thesis will be on lexical borrowing from Arabic. Within this subcategory, the influence of Turkish vowel and consonant harmony and Arabic templatic morphology on the borrowing and incorporation of Arabic words into Turkish will be of specific interest.

Appel and Muysken (1987) list the possible causes of lexical borrowing as follows:

- 1) [...] Cultural influence;
- 2) Rare native words are lost and replaced by foreign words;
- 3) Two native words sound so much alike that replacing one by a foreign word resolves potential ambiguities;
- 4) There is a constant need for synonyms of affective words that have lost their expressive force;
- 5) Through borrowing, new semantic distinctions may become possible;
- 6) A word may be taken from a low-status language and used pejoratively;
- 7) A word may be introduced almost unconsciously, through intensive bilingualism.

(Appel, Muysken 1987: 165–166)

A couple of these causes can be immediately ruled out in the case of Ottoman borrowing, namely numbers (6) and (7). Firstly because the languages being borrowed from in this case were high-status; secondly because, though trilingualism was common among learned Ottomans, the borrowing happening in the Ottoman courts was an extremely conscious, researched act. Among the rest, though numbers (2) through (5) are also possible motives for borrowing in this situation, generally and

overwhelmingly the protracted borrowing in the case of Ottoman Turkish was caused by the influence and prestige of Islamic and, by extension, Arabic and Persian culture and language. As will be discussed at more length in Section 4 however, it is also true that one consequence of the protracted borrowing and the subsequent creation of replacements for these loans has been the creation of semantic and sociolinguistic distinctions between formerly-synonymous word pairs (Çelik 2008; Cücelglu, Slobin 1980).

Additionally, Appel and Muysken (1987) distinguish between different types of lexical borrowing based on whether or not parts of borrowed items are replaced according to patterns in the receiving language. Among the three kinds they identify, namely loanwords, loan blends, and loan shifts, the Turkish borrowings from Arabic considered in this thesis consist solely of loanwords. This category is defined as imported words with substitution of original features by native ones purely on the phonemic level. An alternative to the substitution of foreign phonemes by native phonemes during the process of lexical borrowing is the deletion or insertion of phonemes into the borrowed word to create a word structure in accordance with the receiving language's syllable structure. Lexical borrowing is usually followed by either of these processes, also known as phonological adaptation and accommodation, respectively (Winford 2003).

Zimmer (1985) provides an example of the gradual phonological adaptation of an Arabic loanword into Turkish with respect to the Turkish linguistic features of vowel and consonant harmony. Vowel harmony is a process in Turkish by which all of a words' vowels, even across morpheme boundaries, assimilate to the word's first vowel with respect to the phonological feature of front/backness (Underhill 1980). The complete set of rules, as laid out by Underhill (1980), is as follows:

- 8) Vowels assimilate to the immediately preceding vowel in front/backness;
- 9) A high vowel assimilates to the immediately preceding vowel in rounding;
- 10) Low vowels must be unrounded (except in first syllable).

Adapted from Underhill 1980: 25

The result of these rules is that suffixes' vowels are specified only for height, meaning they vary in terms of front/backness and rounding based on the immediately preceding vowel from the root word (Underhill 1980). The following examples show the variation in the plural and possessive suffixes' vowels based on the preceding root vowel;

11a) diʃ	11b) diʃ - ler	11c) diʃ - ler - im
<i>tooth</i>	<i>teeth</i>	<i>my teeth</i>
12a) kuʃ	12b) kuʃ - lar	12c) kuʃ - lar - ωm
<i>bird</i>	<i>birds</i>	<i>my birds</i>

Adapted from Zimmer 1985: 623

Though the suffix vowels are specified for height, with an underlyingly low vowel – [e] or [a] – in the plural suffix and an underlyingly high vowel – [i] or [u] – in the possessive suffix, their surface forms' front/backness varies based on the frontness, in the case of *diʃ*, or backness, in the case of *kuʃ*, of the root vowel.

In addition to vowel harmony, there is also the phenomenon of consonant harmony, by which consonants' place of articulation is influenced by the front/backness of the vowels in their syllables; front syllables elicit palatal consonants and a standard [l], whereas back vowels elicit velar consonants and the velarized [t̪] (Zimmer 1985). For example:

- 13) cel
- 14) qot̪

*bald*

*arm*

(Zimmer 1985: 623)

The front vowel [e] in (13) elicited the palatal stop [c] whereas the back vowel [o] in (14) elicited a velar stop [q] and the velarization of the final [l]. The backing or fronting of phonemes because of the place of articulation of a neighboring phoneme is a common phonological phenomenon, based on the actual physics of sound articulation.

Zimmer (1985) discusses the case of the Arabic loanword [idrac] which in its originally adopted form constituted a violation of consonant harmony rules, due to the palatal stop [c] occurring in the same syllable as the back-patterning vowel [a], as well as of vowel harmony rules because of the application of the possessive suffix with a front vowel, expressed as [-im]. Gradually, these consonant and vowel harmony violations were eliminated, with the original palatal [c] changing to the velar [q] and the vowel of the possessive suffix [-im] changing to its high back equivalent [ɯ], all in response to the backness of the stem's final vowel, [a] (Zimmer 1985). This change provides a gradual example of adaptation, wherein the original sounds from Arabic were, over time and in stages, replaced with the Turkish phonemes appropriate to the native phonological rules of vowel and consonant harmony.

Despite this instance of adaptation, the borrowed stem itself still constitutes a vowel harmony violation given the word-initial front-patterning [i]'s coexistence with the back-patterning [a] (Zimmer 1985). Similarly, other words of foreign origin, including those from the Ottoman-era mass borrowing of Arabic and Persian lexical items, are to this day recognizably foreign because of the unharmonized vowels in their stems:

15) *memur*

16) *insan*

*official*

*person*

(Underhill 1980: 26)

Another linguistic feature with important implications for the results of Turkish-Arabic

language contact is the templatic morphology of Arabic. This refers to the existence of strings of three or more consonants, carrying with them some abstract meaning, which serve as the basis for predictable word formation. This word formation process is predictable because the same vowel scheme, applied to different underlying roots, will result in the same corresponding grammatical forms (Watson 2007). For example, from Lewis (1999), the root consonant clusters KTB and JBR represent the concepts of writing and compulsion, respectively. Therefore, with the appropriate vowel templates inserted:

17) kataba	20) jabara
<i>he wrote</i>	<i>he compelled</i>
18) kaatib	21) jaabir
<i>writer</i>	<i>compelling</i>
19) maktuub	22) majbuur
<i>written</i>	<i>compelled</i>
	(Lewis 1999: 6)

According to Lewis (1999), the predictable nature of word-formation through templatic morphology allowed for exponential lexical borrowing from Arabic into Turkish; whenever one Arabic word was borrowed, all the other possible expressions of the underlying root would be borrowed as well. A claim like this would likely seem unusual in most situations of natural language contact, but it is important to remember here the specifics of the speech community responsible for the most heavy borrowing of Arabic and Persian linguistic features. The scholars of Ottoman Turkey were not accidentally introducing Arabic and Persian elements into their speech as a side-effect of their trilingualism, but were purposefully importing as many aspects of those two languages as they could to increase the prestige of their new official language (Kerslake 1998). These speakers' familiarity with Arabic, and

their relentless mining of the language for new prestigious linguistic forms, explains how and why all the possible expressions of a templatic root would be borrowed into Turkish. Additionally, the ease with which templatic morphology allowed every possible grammatical expression of an underlying root to be predictably formed is the primary reason why borrowed lexical items of Arabic origin would eventually surpass those of Persian origin in quantity (Lewis 1999).

### 3. Language Reform

The beginning of Ottoman Turkish language reform was strongly linked to the first stirrings of nationalist sentiment in the Ottoman Empire, embodied by the Propitious Regulations of 1839 and other mid-19<sup>th</sup> century legislation (Lewis 1999). The era between 1839 and 1876 would come to be known as the *Tanzimat* Period, after the group of reformist writers who sought to create an Ottoman national identity in large part through the study and reform of the Ottoman Turkish language (Kerslake 1998). The victories achieved by these *Tanzimat* writers included orthographical reforms towards a more phonetic use of the Arabic script, which reflected the shift of even elite Ottoman poets away from the more visual, pattern-based tradition of Persian poetry towards a form of poetry based on capturing the sounds of the Ottoman's spoken language (Gallagher 1971). Additionally, the Ottoman Turkish language began to be studied in more depth than ever before during the Ottoman Empire, both in the official curriculum of the educational system as well as by the government, with the goal of producing grammars and dictionaries which would set the stage for the standardization of the Turkish national language (Kerslake 1998). The ultimate *Tanzimat* victory for the more purely Turkish language of the masses came in the form of Article 18 of the Constitution of 1876 which designated Turkish, as opposed to Ottoman, the official language of the state (Lewis 1999). This was a blow against the far-reaching influence of Arabic and Persian on the official Ottoman language.

A crucial outcome of the mid-19<sup>th</sup> century *Tanzimat* legislation was the expansion it created of the country's journalistic infrastructure (Lewis 1999) which would become more and more influential in the field of language reform, up through the early 20<sup>th</sup> century (Cücelglu, Slobin 1980). This link between journalism and language reform is exemplified by the journalist Ali Suavi, who Lewis (1999) identifies as one of the earliest, in the 1860's, to express a position correlating language reform with nationalistic attitudes. Suavi went further than his contemporaries in advocating the use of existing Turkish words in place of the more prestigious Arabic and Persian loans that had been brought in to replace them. In an editorial for the newspaper *Muhbir*, he promised that the paper would write all its stories in a language that could be understood by all readers (Lewis 1999), which chiefly meant it would be free of the heavy Arabic and Persian influences that had come to characterize most stylistic registers of the official Ottoman language.

Writers other than journalists were also on the forefront of the debate on language reform by virtue of their profession, as in the case of the aforementioned poets whose efforts gave way to some of the first Turkish spelling reforms (Gallagher 1971). While some writers were opposed to the notion of language reform due to their love of the Ottoman language, and its use as a marker of their elite status, writers were gradually weaning themselves off of the borrowed Arabic and Persian grammatical constructions; by WWI, a majority of writers were making an effort to avoid these foreign syntactic and morphological structures (Lewis 1999). It was the period after WWI, however, that would see the language reform efforts more formalized and embraced by the government.

In 1923 the Turkish Republic was founded, with Mustafa Kemal Atatürk, a former military leader, as its first president. In the wake of WWI there was a desire to create a new sense of Turkish nationalism, and much of this was to be achieved through the reevaluation and reform of the language

(Lewis 1999, Aytürk 2004). Much scholarship was devoted to the study of common Turkish, in an effort to build up the theretofore devalued language. This reevaluation of the language was also accomplished in part by efforts to link other, more prestigious languages to Turkish; for example, Atatürk and his followers, known as Kemalists, would frequently propose Turkish etymologies for Western European words, and some reformers went so far as to claim that Turkish had a unique place in linguistic philology as the ancestor of all other European languages (Lewis 1999).

A more important component in linking language to nationalist sentiment were the language reform efforts, which had the ultimate goal of purging the Turkish language of its pervasive Arabic and Persian influences, ranging from the lexical and grammatical borrowings to the use of the Arabic script. In addition to serving the desire to foster a sense of nationalistic pride, ridding the language of its Arabic and Persian influence was a way for the new Turkish republic to distance itself from its Islamic past, and emphasize its secular, westernized future (Lewis 1999, Gallagher 1971). As part of the effort to distance the nation from its Islamic past, the switch was made, in 1928, from the Arabic script to the Roman script and in 1932 the Turkish Language Institute (TDK) was founded by Atatürk (Aytürk 2004). This organization was created to carry out research on the Turkish language, though always with the ultimate goal of language reform and purification in mind.

A crucial step taken by Atatürk and the TDK to ensure the eradication of foreign loans in Turkish was their outlawing of foreign loans from both spoken and written Turkish (Aytürk 2008). While the TDK could not feasibly control how people spoke, they could control more easily the kind of language used by the country's journalists in their publications. Thus, the primary outcome of this official position was the eradication of Arabic and Persian loanwords from Turkish newspapers. Forced upon journalists as this act was, its importance in the success of the nation's acceptance of the

reform goes to show the role the press played in validating and demonstrating the changes being made to Turkish at the time, as well as the government's commitment to the reform effort.

However, if Turkish was to rid itself of its many foreign loans, it would need to create substitute words that could be presented as legitimately Turkish. This task took many different forms and continued throughout the 1920s and 30s (Lewis 1999). Brendemoen (1998) identifies the range of techniques used by the TDK to find and create Turkish substitutes for the outbound Arabic and Persian loanwords.

- 23) Reviving words found in older, particularly Old Ottoman and Pre-Ottoman texts;

*görenek*        replacing        *an'anə*        (tradition)

- 24) Suggesting the use of words found in Anatolian dialects;

*gözgü*        replacing        *ayna*        (mirror)

- 25) Introducing words from other Turkic and Altaic languages;

*ulus*        (nation)        from Mongolian

- 26) Enlarging the semantic sphere of already existing Turkish words;

*çevir*        (to turn around)        added meaning: (to translate)

*neden*        (why)        added meaning: (reason, cause (n.))

- 27) Introducing calques and literal Turkish translations of foreign terms;

*bakan*        (minister)        from: *bak-*, (to look)

- 28) [...] Formation of compounds with prefixes, a morphological class not existing in Turkic languages;

*öngör-*        replacing        *derpiʃ et-*        (foresee)

- 29) [...] Derivation of new words by means of productive or unproductive suffixes of

Turkic or other origin

-sal/sel      ex: *dilsel*      replacing      *lisani* (linguistic)

Adapted from Brendemoen 1998: 243

These word-creation techniques ran the gamut from passive searches for preexisting words, as in techniques (23) and (24), to more creative developments of new words or word-formation processes.

One large-scale program organized by the Ministry of Education, and not exactly covered in the above list, consisted of a word-collection effort from people's everyday speech throughout the 1920s. A complex network was set up to allow for word-collection from all parts of the new republic, with committees and sub-committees set up in each province and sub-province to gather the words collected by the government employees in those areas and then ship them back to the central body of the TDK. After just a year, almost 126,000 submissions had been collected from a wide range of government employees from all parts of the country, ultimately accounting for about 35,000 distinct words (Lewis 1999). At much the same time, the Turkish Society for the Study of History compiled lists of technical terms from Ottoman Turkish with no readily available Turkish equivalents and sent them to universities and scholars around the country asking them for suggested replacements. Words suggested in this format were then approved by the Ministry of Education for use in textbooks, though they often included the very loans the TDK was trying to abolish (Lewis 1999). The intense dislike of foreign loans was not shared by all Turks, however, as evidenced by the 1933 proclamation by the newspaper *Hakimiyet-i Milliye* that whatever words were being used by the people should be considered Turkish, regardless of their etymology (Lewis 1999). This would not, however, put an end to the search for Turkish substitutes for Arabic and Persian loans, no matter how long-standing the lexical borrowings in question.

The other broad category of techniques used by the Turkish government to create Turkish equivalents for Ottoman-era foreign loanwords was the outright invention of new words. It is important here to consider the TDK's self-professed goals relating to word retrieval and creation:

- (1) Collecting and publishing the treasures of the Turkish language existing in the popular language and old books;
- (2) clarifying the methods of word-creation in Turkish and employing them to extract various words from Turkish roots;
- (3) uncovering and publicizing pure Turkish words which may be substituted for words of foreign roots widely used in Turkish, especially in the written language.

(Lewis 1999: 49)

The first and third goals fall under the umbrella of word-collecting discussed earlier, which it should be noted included words from old Turkish texts and Turkic language dictionaries alongside words found in the contemporary vernacular of Turkish speakers (Lewis 1999; Brendemoen 1998). The second by-law, however, represents the generative goal of applying existing native Turkish word-formation processes to new roots and in new ways to expand the lexicon. There was also notable potential for overlap between these two general activities since the word-formation processes collected and set out by the TDK were frequently applied to roots recently rediscovered during the process of word-collection, be it from old texts or from contemporary dialects from across the nation (Lewis 1999). Eventually, these replacements for Arabic and Persian loans were listed in regular newspaper installments or presented independently in TDK-sanctioned and published books (Lewis 1999).

Lewis (1999) provides a sample list of suffixes that were revived during the TDK's search for replacement word-formation processes. Of the twelve he discusses, seven originate in some form of

Turkish, be it contemporary spoken Turkish or the long dead Old Turkish; three of these are revivals of formerly out-of-use forms (Lewis 1999). Though one of the remaining suffixes derives from another member of the Turkic language family, the other four were extracted from pre-existing Turkish loans from less-related languages, ranging from French to Mongolian (Lewis 1999):

- 30) -çe/ça      Derived from three old Serbo-Croat loans to provide a feminine suffix
- 31) -sel/sal      Derived from the French suffix in *culturel* (cultural) and *principal* (principal)
- 32) -men/man      Derived from the Turkish loan *vatman* from French *wattman* (itself from English), referring to a tram driver; similar to English *-er* (as in *teach-er*)
- 33) -tay      Extracted from the 13<sup>th</sup> or 14<sup>th</sup> century Turkish loan *kurultay* from Mongolian; used to denote a group of people.

Adapted from Lewis 1999: 95–106

These last examples are representative of the gradual acknowledgment by language reformers, including Atatürk himself, that ridding Turkish of all foreign influence was neither practical nor truly desirable. Part of the reason for this difficulty was that the most fervent reformers were going above and beyond ridding Turkish of the excesses of the Ottoman court's borrowings from Arabic and Persian; such reformers were now arguing against the retention of any and all words of foreign origin, no matter how long-standing the borrowing or how wide its currency among the contemporary Turkish speech community (Lewis 1999). Largely because of the persistence of reformers like these, Atatürk insisted that any borrowed words remaining in Turkish would have to be presented to the public as actually Turkish in origin (Lewis 1999). Some people even suggest that Atatürk's embrace of the “Sun-Language Theory”, proposing Turkish as the global proto-language, was merely an attempt to convince

language reformers to allow known foreign words to remain in Turkish, by positing an ultimately Turkish origin for all words and languages (Brendemoen 1998).

One especially pertinent field in the debate between retaining foreign loans and developing Turkish replacements concerns technical terminology. The development of entirely new fields of terminology in response to lexical borrowing is a common phenomenon (Winford 2003), as whole semantic fields not yet established in a receiving language are introduced from external languages with the necessary vocabulary. In Turkish, the history of medical vocabulary proves especially relevant as a testament to the influence of Arabic on the language. During the Ottoman period Arabic was the source of all medical vocabulary and when the language of instruction finally changed, in the early 19<sup>th</sup> century, it was to French. Nonetheless, under pressure from students, an Ottoman-language medical dictionary was created and, more than a hundred years later in 1978, the TDK produced a similar publication in contemporary Turkish (Lewis 1999). As was so common with TDK publications however, many of the neologisms' meanings were not immediately evident to speakers because of their reliance on rare and even long out-of-use root words or affixes. This opacity of meaning, coupled with the existence of an already widely agreed-upon, albeit foreign, medical vocabulary left the new official terms without much support. It should also be noted, that by this point in time the foreign loanwords being combatted by the TDK consisted primarily of more recently introduced French and English medical terms.

Language activists continue to argue that suitable alternatives to European technical terms can be found in the Turkish vernacular. Though either item from any given foreign-native word pair can be used to refer to the same highly technical medical term, they argue, the foreign terms are the only ones that have taken on the necessary aura of scientific legitimacy. For example, these activists present

word pairs such as the following:

Ailment	Doctor's Term	Popular Term
anemia	<i>anemi</i>	<i>kansužluk</i> (bloodlessness)
appendicitis	<i>apandisit</i>	<i>apandis yangušuš</i> (appendix inflammation)
cataract	<i>katarakt</i>	<i>perde</i> (curtain)(Persian loan)
dysentery	<i>dizanteri</i>	<i>kanluš basur</i> (bloody hemorrhoids) (Arabic loan)

Adapted from Lewis 1999:126

The association of foreign terms only with scientific legitimacy, influenced by the frequent studies abroad of medical professionals, is preventing Turkish from developing a nationally agreed-upon and professionally used vocabulary for science (Lewis 1999). Ironically, some of the alternatives presented as preferable by the language activists are themselves Arabic or Persian loans. However, this reversal demonstrates the now more common contemporary pattern of English as the global dominant and prestigious language (Lewis 1999).

#### 4. Results of Language Reform

The Turkish language reform has been generally successful in creating a single, standardized language (Gallagher 1971; Brendemoen 1998), even though it was the most controversial of the

nascent Turkish republic's reform efforts (Brendemoen 1998). Gallagher (1971) cites modern mass media, standardized education, and compulsory military service as conduits for the dispersal of the new national standard language. The amount of class-based linguistic differences have decreased with the reform, especially when compared to the once exclusively elitist and administrative Ottoman language. This outcome carries to fruition the notions of egalitarianism proposed by the Turkish Republic (Gallagher 1971). As concerns the Arabic influence, Arabic words are in fact still used. Those that remain are by now so thoroughly integrated into the Turkish lexicon, though, that they are used in conjunction with native Turkish suffixes, as well as those suffixes collected and introduced into Turkish by the TDK (Gallagher 1971). Additionally, Gallagher (1971) mentions contemporary borrowings from other Turkic languages which are in some cases supplanting older terms of Arabic origin. This goes to show that the processes of language change and lexical borrowing are, as always, ongoing.

The TDK itself has fallen from its former position of influence. In the 1950s, with the rise to power of the Democratic party, the TDK lost its semi-official status and its government subsidy and began to conduct less sensationalist work than in the height of the language reform and word-creation movements (Gallagher 1971). Nonetheless, the influence of the TDK's past policies persist to this day.

Lewis (1999) identifies three possible outcomes for the neologisms created as part of the TDK's official language reform effort: they could be wholly integrated into Turkish and be unmarked as recently fabricated words, or they could fail to gain any acceptance at all. The third possibility is that the new term becomes widely used, but with a meaning other than that which was originally intended. In this kind of situation, it is furthermore possible that the recently introduced neologisms could coexist with the loanwords they were meant to replace (Çelik 2008). Whereas these loan-neologism pairs have traditionally been thought to have remained purely synonymous, Çelik (2008) argues that with time

they have undergone semantic change, resulting in differentiation between words that were originally meant to be interchangeable.

Çelik (2008) tested this hypothesis with a questionnaire asking subjects to rank a set of given definitions for each item of each loanword-neologism pair in terms of the frequency of their use. These data were then used to calculate whether the original loan and the intended replacement had differentiated in terms of their primary and/or secondary meanings. Of the 11 lexical pairs Çelik (2008) investigated, a total of nine were found to have differentiated significantly, and six of those were differentiated with respect to more than one given definition. There was no overarching pattern in the direction of semantic differentiation, nor were any of the socio-demographic factors included in the study found to predict the primary meaning assigned to the neologisms (Çelik 2008).

In such situations of lexical borrowing where semantically related or even synonymous native and borrowed words end up coexisting, it is indeed expected for the pairs of lexical items in question to undergo some sort of semantic change as they vie for the semantic space that was formerly occupied by a single lexical item (Winford 2003). Keenan (2003) explains the motivation for this semantic differentiation as a psycholinguistic disposition on the part of language learners to prioritize learning new words, rather than synonyms of already existing words. There are three different kinds of semantic change possible: broadening, narrowing, and meaning shift (Fromkin 1978). The first two refer to the growth or shrinkage of the semantic space occupied by a lexical item, respectively. While for both of these the lexical item retains some of its original meaning, all the same becoming more vague or more specific, words undergoing meaning shift take on a new, unrelated meaning and lose completely the meaning originally associated with them (Fromkin 1978). In the case of borrowing-induced semantic change, narrowing is the most common of these, as the supposedly synonymous word

pairs eventually split the original meaning between the two of them.

Operating under the assumption that speech communities will as a whole tend towards decreasing or preventing synonymy, Keenan (2003) proposes another way in which language learners may respond to synonymy. His First Exposure Effect principle proposes that when presented with two lexical items competing for the same meaning, speakers will come to prefer and thus use more frequently the one they were exposed to first. This principle would likely create a bias in favor of older lexical items because of the potential for speakers to be exposed to them earlier as well as the possibility that they occur more frequently than newer words. Given the general success of the Turkish language reform (Lewis 1999; Gallagher 1971), however, this principle cannot have held true in all instances. To get a sense of the success of some of the TDK's neologisms in modern Turkey, I collected data using the Google search engine to compare the number of tokens found for the traditional loanwords and their intended replacements in each such pair. The results of this search, using the 11 word pairs investigated by Çelik (2008), are as follows:

	Number of Tokens		
	Traditional Word	Neologism	Difference in Tokens (Traditional – Neologism)
ANSWER	68,300,000	15,100,00	53,200,000
CITY	40,800,000	3,780,000	37,020,000
DREAM	13,300,000	1,050,000	12,250,000
STORY	13,200,000	7,060,000	6,140,000
JUDGE	3,800,000	604,000	3,196,000
EVIDENCE	1,720,000	968,000	752,000
TO BE ACQUITTED	821,000	21,500	799,500
WITNESS	2,400,000	2,520,000	- 120,000
EXAM	749,000	10,300,000	- 9,551,000
STUDENT	977,000	23,100,00	- 22,123,000

RESULT	1,310,000	27,000,000	- 25,690,000
--------	-----------	------------	--------------

The results are presented here in descending order based on the difference in hits between the original borrowed word and the TDK's intended replacement word; thus, only the last four word pairs constitute pairs where the replacement words were found to be more common. On a whole then, it would seem that there has indeed been a tendency among Turkish speakers to give preference to the older items of these synonymous pairs.

There are of course limitations to this medium of data retrieval. Most obviously, the brunt of what is retrieved through the Google search engine reflects how Turkish speakers write, as opposed to how they speak in naturally-occurring conversation. Additionally, it is unclear what kinds of stylistic or sociolinguistic registers are included even with as far-reaching a corpus as the internet. Are most of the texts being surveyed representative of more formal writing, such as on newspaper or government websites, or is there also a sizable presence of amateur writing, as on blogs or message boards, in Turkey? Even if the latter material does exist it is presumably emerging from a small wealthy, technologically literate, and young subset of the larger Turkish population. Additionally, if this internet-using segment of the population is concentrated in certain parts of the country, such as Istanbul for example, then this may leave entire geographical regions of the country underrepresented. Thus, it is unclear precisely what kinds of speech communities are being represented in this kind of data collection process.

Thus, though there have been cases of successful substitution of borrowed words by the TDK's creations, the introduction into the lexicon of their neologisms has also more often than not merely added to the lexical variety available to speakers of Turkish (Çelik 2008; Cücelglu, Slobin 1980). In cases with sufficient retention of both borrowed and native forms, the choice between these distinct sets

of vocabularies can in and of itself come to represent different stylistic registers (Winford 2003). Given that the neologisms of Turkish were created as symbols of a very specific political agenda, namely Turkish nationalism, the lexical variants that came to coexist in Turkish have with time become sociolinguistic markers. That is, use of the more traditional forms, specifically Arabic and Persian loans meant to be replaced, has come to represent cultural and political conservatism whereas the words created to counteract foreign influence have become representative of the left as well as of secular and modernist leanings (Cücelglu, Slobin 1980). In fact, Cücelglu and Slobin (1980) found that subjects attributed different political attitudes and values to speakers based on identical paragraphs differing only in whether the vocabulary was pre- or post-reform (Cücelglu, Slobin 1980).

## 5. Conclusion

The history of Arabic loanwords in Turkish provides an interesting case in the study of language contact. The status of Arabic and Persian as prestige languages for the Ottoman Turks was responsible for their immense influence on and incorporation into the official Ottoman language, but when the political and cultural situation changed, these formerly accepted borrowings suddenly became a scapegoat for the Turkish language reformers. The commitment displayed by Atatürk and the TDK in ridding their language of Arabic and Persian influence during the post-WWI period of nationalization is a testament to the perceived symbolic power of language and is also a textbook example of language planning.

This protracted language reform effort also provides clear parallels to the work of language academies in other nations, charged to this day with the task of maintaining the integrity of their languages in the face of invasive vocabulary from other languages. In large part, however, the governing bodies charged with responsibilities comparable to those of the TDK have been most active

in their opposition to the encroaching influence of English. This generally means that these bodies' efforts have begun only in more recent history, as is the case with the infamous French language academy which enacted legislation to ban words of foreign origin in advertising only in the 1970's (Fromkin 1978). Comparing these reformers to the mid-to-late 19<sup>th</sup> century genesis of the Turkish language reform, it is clear that these academies are generally facing ongoing, contemporary cases of borrowing, as opposed to the TDK's original charge of retroactive combat against a historical case of borrowing. The history behind the Turkish language reform also allows us the benefit of seeing its outcomes with some distance, and could thus perhaps provide insight into how successful other institutes' specific policies or reform attempts more broadly might be in the longer run.

#### Bibliography

Appel, René, and Pieter Muysken. 1987. *Language contact and bilingualism*. 1<sup>st</sup> edn. Baltimore: Edward Arnold.

Aytürk, Iker. 2004. Turkish linguists against the West: The origins of linguistic nationalism in Atatürk's Turkey. *Middle Eastern studies* 40.6.1–25.

Aytürk, İlker. 2008. The first episode of language reform in Republican Turkey: the Language Council from 1926 to 1931. *Journal of the Royal Asiatic Society* 18.3.275–293.

Bredemoen, Brent. 1998. The Turkish Language Reform. *The Turkic Languages*, ed. by Lars Johanson and 'Eva 'Agnes Csató, 242–247. New York: Routledge.

Çelik, Mehmet. n.d. *Neologisms in contemporary Turkish*. Accessed: 9/29/08. Online: <https://www.latrobe.edu.au/linguistics/LaTrobePapersinLinguistics/Vol%202008/4Celik.pdf>

Cüceloglu, Dogan, and Dan I. Slobin. 1980. Effects of Turkish language reform on person perception. *Journal of cross-cultural psychology* 11.3.297–326.

Fromkin, Victoria. 1978. *An introduction to language*. 2<sup>nd</sup> edn. New York: Holt, Rinehart and Winston.

Gallagher, Charles F. 1971. Language reform and social modernization in Turkey. *Can language be planned? Sociolinguistic theory and practice for developing nations*, ed. by Joan Rubin and Björn H. Jernudd, 159–178. Honolulu: University Press of Hawaii.

Keenan, Edward L. 2003. An historical explanation of some binding theoretic facts in English. *The nature of explanation in linguistic theory*, ed. by John Moore and Maria Polinsky, 153–189.

Stanford: CSLI Publications.

Kerslake, Celia. 1998. Ottoman Turkish. *The Turkic Languages*, ed. by Lars Johanson and 'Eva 'Agnes Csató, 179–202. New York: Routledge.

Lehmann, W.P. 1973. A structural principle of language and its implications. *Language* 49.1.47–66.

Lewis, Geoffrey. 1999. *The Turkish language reform: A catastrophic success*. 1<sup>st</sup> edn. Oxford: Oxford University Press.

Underhill, Robert. 1980. *Turkish Grammar*. 3<sup>rd</sup> edn. Cambridge: Massachusetts Institute of Technology Press.

Watson, Janet C.E. 2007. *The phonology and morphology of Arabic*. 1<sup>st</sup> edn. Oxford: Oxford University Press.

Winford, Donald. 2003. *An introduction to contact linguistics*. 1<sup>st</sup> edn. Oxford: Blackwell Publishing.

Zimmer, Karl. 1985. Arabic loanwords and Turkish phonological structure. *International journal of American linguistics* 51.4.623–625.