21. Geminated Consonant Words in Dīwān Lughāt Al-Turk and Their Reflections in Anatolian Folk Dialects1

Mehmet YEŞİLKAYA²

APA: Yeşilkaya, M. (2025). Geminated Consonant Words in Dīwān Lughāt Al-Turk and Their Reflections in Anatolian Folk Dialects. RumeliDE Dil ve Edebiyat Araştırmaları Dergisi, (47), 344-360. DOI: https://doi.org/10.5281/zenodo.16882817

Abstract

Dīwān Lughāt al-Turk, one of the earliest and most comprehensive dictionaries of the Turkish language, provides valuable data not only in terms of vocabulary but also in the historical development of phonetic structures. This study identifies words written with geminate consonants (e.g., kk, tt, ss) in the Dīwān and compares them with their counterparts in Standard Turkish (ST) and Anatolian dialects (AD). These words are analyzed with respect to their phonetic continuity by comparing the geminated forms observed in Anatolian dialects. The research also reveals that some words were previously misclassified as geminated in earlier studies; such cases have been reevaluated and corrected statistically. It was found that geminate consonants in numeral expressions and certain lexical items still persist in many regional varieties of Turkish. The analysis is based on historical forms in the Dīwān, modern Turkish dictionaries, and regional dialect data, offering a threefold comparison for each word. This study presents significant insights into the phenomenon of consonant gemination, tracing its continuity from early written texts to contemporary spoken varieties.

Keywords: Dīwān Lughāt al-Turk, Anatolian folk dialects, consonant gemination, phonetic change, phonetic comparison

Funding: No external funding was used to support this research.

Similarity Report: Received - Turnitin / Rate: %4

Ethics Complaint: editor@rumelide.com

Peer Review: Two External Referees / Double Blind

Statement (Thesis / Paper): It is declared that scientific and ethical principles were followed during the preparation process of this study and all the studies utilised are indicated in the bibliography. **Conflict of Interest:** No conflict of interest is declared.

Copyright & Licence: The authors own the copyright of their work published in the journal and their work is published under the CC BY-NC 4.0 licence.

Source: It is declared that scientific and ethical principles were followed during the preparation of this study and all the studies used are stated in the bibliography.

Article Type: Research article, Article Registration Date: 09.07.2025-Acceptance Date: 15.08.2025-Publication Date: 16.08.2025; DOI: https://doi.org/10.5281/zenodo.16882817

Dr., MEB, Gaziosmanpaşa Bilim ve Sanat Merkezi / Dr., MNE, Gaziosmanpaşa Science and Art Center (İstanbul, Türkiye) memyesilkaya@gmail.com ORCID ID: https://orcid.org/0000-0002-8905-3306 https://ror.org/00jga9g46 ISNI: 0000 0001 2179 4856 Crossref Funder ID 501100013898

Dîvânu Lügâti't-Türk'te Geçen İkiz Ünsüzlü Kelimeler ve Bu Kelimelerin Anadolu Halk Ağızlarına Yansımaları³

Öz

Türk dilinin en eski ve kapsamlı sözlüklerinden biri olan Dîvânu Lügâti't-Türk, sadece söz varlığı açısından değil, fonetik yapıların tarihî gelişimi bakımından da önemli veriler sunmaktadır. Bu çalışmada, Dîvânu Lügâti't-Türk'te ikiz ünsüzle (ör. kk, tt, ss) yazılmış olan kelimeler tespit edilmiş ve bu kelimelerin ölçünlü Türkiye Türkçesindeki (ÖTT) karşılıkları ile Anadolu halk ağızlarındaki (AHA) biçimleri karşılaştırmalı olarak incelenmiştir. Bu kelimeler, Anadolu halk ağızlarında gözlenen ikiz ünsüzlü biçimleriyle karşılaştırılmış ve tarihsel süreklilikleri incelenmiştir. Araştırma kapsamında, bazı kelimelerin önceki çalışmalarda hatalı biçimde ikiz ünsüzlü olarak değerlendirildiği saptanmış ve bu veriler yeniden düzenlenmiştir. Rakamlarda ve bazı kelimelerde tespit edilen ikizleşmelerin Anadolu ağızlarında hâlen varlığını sürdürdüğü görülmüştür. İnceleme; Dîvân'daki tarihî biçimler, Türkiye Türkçesi sözlükleri ve bölgesel ağız verileri üzerinden yürütülmüş; her kelime için üç yönlü bir karşılaştırma yapılmıştır. Çalışma, ünsüz ikizleşmesi olgusunun tarihî metinlerden halk diline uzanan izlerini ortaya koymak açısından Türk filolojisine önemli bir katkı sunmaktadır.

Anahtar kelimeler: Dîvânu Lügâti't-Türk, Anadolu ağızları, ünsüz ikizleşmesi, fonetik değişim, fonetik karşılaştırma

Makale Türü: Araştırma makalesi, Makale Kayıt Tarihi: 09.07.2025-Kabul Tarihi: 15.08.2025-Yayın Tarihi: 16.08.2025; DOI: https://doi.org/10.5281/zenodo.16882817

Hakem Değerlendirmesi: İki Dış Hakem / Çift Taraflı Körleme

Beyan (Tez/ Bildiri): Bu çalışmanın hazırlanma sürecinde bilimsel ve etik ilkelere uyulduğu ve yararlanılan tüm çalışmaların kaynakçada belirtildiği beyan olunur.

Finansman: Bu araştırmayı desteklemek için dış fon kullanılmamıştır.

Telif Hakkı & Lisans: Yazarlar dergide yayınlanan çalışmalarının telif hakkına sahiptirler ve çalışmaları CC BY-NC 4.0 lisansı altında yayımlanmaktadır.

Kaynak: Bu çalışmanın hazırlanma sürecinde bilimsel ve etik ilkelere uyulduğu ve yararlanılan tüm çalışmaların kaynakçada belirtildiği beyan olunur.

Benzerlik Raporu: Alındı – Turnitin / Oran: %4

Etik Şikayeti: editor@rumelide.com

1. Introduction

Written by Kāshgharī Mahmūd in the 11th century, the Dīwān Lughāt al-Turk is recognized as one of the earliest known dictionaries of the Turkish language. Composed between January 1072 and February 1074, the work was presented to the Abbasid Caliph al-Muqtadī in 1077 (Tas, 2014, p. 203). Today, a manuscript copy of the $D\bar{\imath}w\bar{a}n$ is preserved in the Fatih Millet Library in Istanbul, and the text has been translated multiple times and served as the subject of various academic studies. As one of the most significant works of Karakhanid Turkish, reflecting a vast portion of the vocabulary of its time (Erdem & Coşun, 2013, p. 52; Aktan, 2010), the Dīwān Lughāt al-Turk continues to be a fundamental reference in the field of Turkology (Sakaoğlu, 1991, p. 111). Beyond its value as a lexical and grammatical resource, the work also offers detailed documentation of dialects and accents of the period, thus contributing significantly to the preservation and understanding of historical varieties of Turkish. It is considered a primary source for resolving problems encountered in the study of Turkic dialects (Gülensoy, 2014, p. 1). The $D\bar{\imath}w\bar{a}n$ can be described as a dialectal dictionary compiled from the vernaculars of various Turkic tribes (Kacalin, 2008, pp. 446–449) and is regarded as the first written source in the history of Turkish dialectology (Ulusoy, 2024, p. 842). Beyond its value as a lexical and grammatical resource, the work also offers detailed documentation of dialects and accents of the period, thus contributing significantly to the preservation and understanding of historical varieties of Turkish. The discovery of a manuscript of Dīwān Lughāt al-Turk, a huge encyclopedic monument of the 11th century, is of great importance for the literature, language, culture, history and ethnography of the eastern peoples, including the Turkic peoples (Sattarova, 2021, p. 92). According to the Islamic Encyclopedia, the work is not merely a dictionary; it also functions as a grammar book shedding light on phonology and morphology, a source of personal, tribal, and place names, and an encyclopedic text offering extensive information on Turkish history, geography, folk literature, as well as medical knowledge and treatment methods of the time (IE, 2008, p. 447). The monument also provides valuable insight into the customs, beliefs, personal names, nicknames, and titles of the people depicted, as well as the unique ways in which these elements were represented (Guliyev, 2024, p. 53).

Information about the life of Kāshgharī Maḥmūd is limited. It is estimated that he lived during the Karakhanid period and was born sometime between 1029 and 1038 (Doğan, 2016, p. 4). According to the Turkish Language Association, his year of birth is given as 1008 (TDK, 2024). While it is known that he received a solid education and authored another grammar book, no detailed information about this second work has been uncovered. His date of death is also uncertain; some sources mention the 1090s, while others refer to the 1120s. It is estimated that the author lived for approximately 85–90 years. Kāshgharī provides some personal clues in his works, which have been compared with historical sources to form various conclusions. Moreover, detailed information is included in the *Islamic Encyclopedia* (*IE*) stating that he was born in Barsghān, descended from Hārūn Bughra Khān, who conquered Bukhara in 992, and was 49 years old in the year 1057 (2008, pp. 9–16). It would be appropriate to describe him as a scholar with a command of classical Arabic and an exceptional mastery of all Turkish dialects and vernaculars (Günşen, 2011, p. 236).

Dīwān Lughāt al-Turk, authored by Kāshgharī Maḥmūd, is one of the most significant works in the history of the Turkish language. Not only does it stand as the first known dictionary of Turkish, but it also represents a unique linguistic treasure that reflects the socio-cultural structure of its period. The work has drawn scholarly attention for both its linguistic and literary aspects, and it has been the subject of numerous studies in the fields of phonetics, morphology, semantics, and lexicology. As an unparalleled source for tracing the historical development of Turkish, the Dīwān also provides a solid

foundation for comparative studies with Anatolian dialects. Many words in Standard Turkish appear in the $D\bar{\imath}w\bar{a}n$ in different forms and variants. One of the most striking examples of these differences is the phenomenon of "consonant gemination" observed in certain words.

This phonetic feature, referred to as "consonant gemination," can be explained by the use of double consonants in certain words that appear with single consonants in modern forms. In the historical development of Turkish, when a word that originally had a geminate consonant in Old Turkish (OT) evolves into a form with a single consonant in Standard Turkish (ST), the process may be described as "consonant simplification" from the perspective of OT > ST. Conversely, when the evaluation proceeds from Standard Turkish back to Old Turkish (ST > OT), the same phenomenon is referred to as "consonant gemination." Since this study examines words with geminate consonants as recorded in $D\bar{t}w\bar{d}n\ Lugh\bar{d}t\ al\ Turk$ through the comparisons of ST > DLT, ST > AD (Anatolian dialects), and DLT > AD, the term "consonant gemination" has been preferred throughout the analysis.

Consonant gemination is a phenomenon frequently observed in historical Turkish texts and has attracted the attention of many scholars. Annemarie von Gabain defines gemination as the orthographic representation of prolonged consonants such as n, r, s, s, s, and t (Gabain, 1959). Talat Tekin notes that this phenomenon generally arises under the influence of primary or secondary long vowels and provides examples such as sadda > sade, tazza > taze, pattir > bahadir, çakkal/şakkal > çakal, and kessi > kesi. He explains the word kes-si as consisting of the third person possessive suffix +si, and—referring to Doerfer—states that the /s/ consonant gradually dropped, resulting in the form kesi (2023, p. 121). Tekin also compares the geminated forms in other Turkic dialects with their equivalents in $D\bar{i}w\bar{a}n$ $Lugh\bar{a}t$ al-Turk, offering forms like $accepta k > \bar{a}ccepta k > \bar{a}c$

It has been determined that some words which appear with single consonants in Standard Turkish are found with geminated consonants in $D\bar{t}w\bar{a}n$ $Lugh\bar{a}t$ al-Turk and that this feature also persists in Anatolian dialects. This finding is significant both for tracing historical phonetic developments and for assessing the interaction between standard Turkish and dialectal data. In this context, words recorded with geminate consonants in the $D\bar{t}w\bar{a}n$ were identified and analyzed in comparison with their counterparts in Standard Turkish and Anatolian dialects. In doing so, the study aimed to reveal the historical continuity of consonant gemination between early Turkish and contemporary dialects, as well as tendencies of divergence from the standard language. Based on the data obtained, each word was evaluated in detail.

2. Methodology

In this study, words written with geminate consonants in $D\bar{i}w\bar{a}n$ $Lugh\bar{a}t$ al-Turk (DLT) were identified and comparatively analyzed in relation to their counterparts in Standard Turkish (ST) and Anatolian dialects (AD). The research was based on two main approaches: literature review and textual analysis. In the first phase, both digital and printed versions of the $D\bar{i}w\bar{a}n$ were scanned to detect words containing geminate consonants (e.g., kk, tt, ss, mm, nn, ll, rr, ss, etc.). Subsequently, the corresponding forms of these words in ST were determined using contemporary Turkish dictionaries and the Turkish Language Association's online resources. The variants of these words in Anatolian dialects were

identified through regional dialect dictionaries and the Anatolian dialect atlas.

Each word was examined in three dimensions: its form in the $D\bar{i}w\bar{a}n$, its equivalent in Standard Turkish, and its usage in Anatolian dialects. The analyses presented in this section aim to reveal the status of consonant gemination in historical texts and to trace its reflections in both the standard language and spoken vernaculars.

3. Analysis

Although the concept of "consonant gemination" was discussed in detail in the previous section, it may be briefly summarized as follows: In some words that are written with a single consonant in Standard Turkish (ST), the same consonant is pronounced more forcefully and with a longer duration in Anatolian dialects (AD), resulting in a geminated (double) consonant pronunciation. This feature primarily arises as a phonetic articulation phenomenon associated with increased emphasis or lengthening in speech.

It is well known that certain words, such as *epey*, *eşek*, *iki*, *rakam*, and *tepe*, which are written with single consonants in ST, appear in geminated forms in various Anatolian dialects. Similarly, historical sources such as *Dīwān Lughāt al-Turk* also exhibit comparable phonetic patterns. This suggests that consonant gemination is rooted in historical linguistic layers and has survived in contemporary dialects.

In Turkish texts written in Arabic script, the presence of a geminated consonant is often indicated by the use of a *shadda* (diacritic sign). In modern transliterations of such texts, these words are explicitly marked with double consonants. Although this feature is observed in Old Turkic and other Turkic dialects, it is widely accepted among scholars that geminated consonants do not typically appear in root words in the written form of modern Standard Turkish. However, some Turkish-origin words in Anatolian dialects are used in geminated forms. It is believed that this gemination arises due to phonetic factors such as consonant duration, stress, and articulatory emphasis.

Geminated consonant forms are attested in Old Turkic and several other historical Turkic dialects; however, such forms are no longer found in the contemporary written language (Karaağaç, 2010, p. 80). Despite this, this phonetic feature continues to survive with various examples in Anatolian dialects (Korkmaz, 2017). Thus, regardless of their original etymology, geminated consonant forms are clearly observable in Anatolian dialects.

Consonant gemination can be observed throughout nearly every region of Turkey, across different words and consonant types. In the Eastern dialect group, it is more common for plosive consonants to be geminated (e.g., rakkam, ikki, dokkuz, sippa, eppey), while in the Western dialect group, the phenomenon occurs with other types of consonants. Karahan (1996, p. 32) employed this phonetic feature as a criterion for dialect classification, analyzing it under the title "gemination of plosive consonants" and using it as a parameter in the distinction between dialect groups.

Kāshgharī Mahmud divided his work into eight sections; in the third section, he referred to geminated consonantal forms of verbs in $D\bar{\imath}w\bar{\imath}an\ Lugh\bar{\imath}at\ al ext{-}Turk$, noting that when a verb ends with /t/ and is followed by the past tense suffix /d/, the /d/ consonant changes to /t/ and is written with a shadda (Mahmud, 2023, pp. 34–314). However, this definition does not align with the context of the present study, as the phonetic phenomenon described here corresponds to what is now commonly referred to as "consonant assimilation" ($\ddot{u}ns\ddot{u}z\ benzesmesi$) in contemporary Turkish linguistics.

Consonant gemination observed in historical Turkic texts has been considered a noteworthy phonological phenomenon by many Turkologists. One of the most comprehensive studies on this subject was conducted by Talat Tekin. In his 1959 article titled "Ünsüz İkizleşmesi in Turkish Language and Dialects" (Consonant Gemination in Turkish Language and Dialects), Tekin presented examples from various Turkic dialects, identifying such forms as arrıg, ikki, ikkiz, and yetti in Dīwān Lughāt al-Turk, and explained these structures in terms of the progressive influence of original long vowels (Tekin, 2023, p. 202). Although Tekin provided many examples of consonant gemination across historical Turkic dialects, he cited only four examples from Dīwān Lughāt al-Turk. In another study, Tekin explored the causes of gemination and attributed the geminate structure in the word ikki (or ėkki) to the presence of the close front vowel [e] (Tekin, 2023, p. 289). In a more recent work, Ince (2017, pp. 219–229) listed a broader array of geminated forms in the Dīwān, including ottuz (thirty), issiz (rude, ungrateful person), ikki (two), ikkiz (twin), arrıg (pure, clean), ellig (fifty), bakku (hill, elevated place), basık/bassık (raided by enemies), bellüg/belgülüg (clear, evident), ira/irra (shy), kakkuk (dried fruit), sekkiz (eight), sokku (masher, pestle), tikkü (bite-sized piece), and yetti (seven).

In dialectological studies conducted across different regions of Turkey, various phonetic and morphological data have been analyzed, among which the occurrence of geminated consonants occupies a significant position. In these studies, examples of consonant gemination have typically been examined in the direction from Standard Turkish (ST) to Anatolian dialects (AD) (i.e., ST > AD) and have been categorized under the phenomenon of consonant insertion or duplication. For instance, in a study focusing on consonant insertions in Anatolian dialects, geminated forms were recorded in various regions such as Antalya, Karabük, and Erzincan. In the dialects of Karabük, Erzincan, and Antalya, forms such as bakkur, sakkm, takki, makkas, vakkit, and $\ddot{o}kk\ddot{u}z$ (gemination of /k/ > /kk/); in Antalya, dollu, cille(/l/ > /ll/); in Karabük, cille(/l/ > /ls/); and in Keban, Baskil, Malatya, Kahramanmaraş, Antalya, and Southwestern Anatolia, forms such as cille(/l/ > /ls/), cille(/l/ > /ls/), were identified. Additionally, in Karabük and Antalya, forms like cille(/l/ > /ls/),

3.1. Analysis of the Data

Dîvânu Dīwān Lughāt al-Turk has been studied, transliterated, and translated into modern Turkish by numerous scholars to date. In this study, the editions prepared by Besim Atalay (1985), Ahmet B. Ercilasun (Ercilasun & Akkoyunlu, 2014), and Mustafa S. Kaçalin (2008) have been taken as the primary sources. Among them, the most recent edition is the one published by Mustafa S. Kaçalin in 2023, with its index section compiled by Mehmet Ölmez.

Due to its digital accessibility and searchability, this study has primarily utilized the 1120-numbered edition of $D\bar{\imath}w\bar{a}n$ $Lugh\bar{a}t$ al-Turk published by the Turkish Language Association (TDK) in 2014, edited by Ahmet B. Ercilasun. In the examples provided in the "Analysis" section, if the name of the edition is not specified in parentheses, it should be understood that the citation refers to this digital version. All geminated consonant forms detected in this digital edition have also been cross-checked using Kaçalin's printed version (2023); for each identified word, an example sentence has been selected from this printed edition and cited with the abbreviation DLT-K.

This study includes both (1) words that had been previously classified under consonant gemination in earlier research but, upon review, were determined not to meet the criteria for such classification, and (2) all valid examples of geminated consonants identified by the authors through direct examination of $D\bar{i}w\bar{a}n$ Lughāt al-Turk. Previous limited studies on the $D\bar{i}w\bar{a}n$ had proposed that words such as arrug, bakkū, bassıg, bellüg/belgülüg, ellig, ırra, ıssız, ikkī, ikkiz, kakkuk, ottuz, sekkiz, sokku, tikkü, and yetti featured consonant gemination. However, it has been determined that some of these items were misclassified; these evaluations are explained below, and such examples have not been included in the statistical data set.

arriğ (**arig**): The word appears in *Dīwān Lughāt al-Turk* as *arriğ*. In contemporary usage, it means "clean" or "free from foreign substances." In Old Turkic, *arig/ariğ* was used not only in the sense of physical cleanliness but also in an abstract sense, indicating the high value placed on moral purity, suggesting that the term dates back to even earlier periods (Bozok, 2021, p. 38). In modern Turkish, the final /g/ consonant has dropped, resulting in the form *ari*. In some historical examples, the word appears with a single consonant, while in others it is written with a geminate consonant.

arrığ: arrıg neng, "very clean thing." Here, the shadda (diacritical mark) indicates emphasis or exaggeration. (BA, Vol. I, p. 143)

arıg: clean object. (DLT-K: 508)

arrig: very clean. arrig nen: very clean thing. (DLT-K: 508-510)

bakku/baku: The word appears as baku/bakku in the $D\bar{i}w\bar{a}n$. In the original manuscript, it is written without a shadda, but some sources have transliterated it as bakku. Furthermore, a footnote in the manuscript suggests that the original form may have included a geminate consonant.

baku: hill, slightly elevated place (BA, Vol. III, p. 226)

bakku: hill, elevated place. kaya körüp baku agdı ("he looked and descended from the rock") (DLT-K:525).

When analyzing the morphological structure of the word, it becomes clear that it is formed by attaching the nominalizing suffix -gu/-ku to the verb root bak- ("to look"). Eraslan has noted in two separate studies that the suffix -gu in Old Turkic served as a nominalizer (2012, pp. 378–379; 2004, pp. 117–120). Tekin similarly emphasizes that this suffix derives nouns from verbs, giving korgu ("guard") as an example (2017, p. 87). In light of this information, it can be concluded that bakku is derived from the verb bak- by means of the suffix -gu/-ku, transforming it into a noun. Therefore, this word falls outside the scope of the present study on consonant gemination and has been excluded from the analysis.

basig/bassik-: In the $D\bar{\imath}w\bar{a}n$ $Lugh\bar{a}t$ al-Turk, the forms basig, basik-, and bassik- are recorded. basig functions as a noun and refers to a place subjected to a nighttime raid. The forms basik- and bassik- written with the /k/ consonant—are verbs meaning "to be raided at night." In contemporary Turkish, this word and its associated meaning are no longer in common usage (cf. TDK Dictionary, 2025).

basig: a place to be raided at night, where the enemy is caught by surprise. Ol ani basginda tutti ("He captured him during the raid") (DLT-K:530).

basıķ-/bassıķ-: to be raided at night. er yagıka bassıktı, kişi yagıka basıktı ("the man was raided by the enemy," "someone raided the enemy") (DLT-K:531).

bellüg: In the sources we examined, no entry for the word *bellüg* was found in the *Dīwān Lughāt al-Turk*. However, the form *belgülüg* appears in both Kaçalin's and Ercilasun's editions. Although İnce (2017, p. 219) claims to have identified the form *bellüg*, this may have been either an unintentional reference from another work or a misrecording of *belgülüg*, which is attested in the *Dīwān*. Therefore, since *bellüg* is not found in the aforementioned sources, and even if it were, it would not qualify for inclusion under the category of consonant gemination.

ellig: This word, sometimes listed among geminated consonant examples, is recorded in the $D\bar{\imath}w\bar{a}n$ $Lugh\bar{a}t$ al-Turk as meaning "fifty".

ellig yarmak: elli para (fifty coins) (DLT-K:593)

The word meaning "owner of a hand" or simply "hand" appears in the $D\bar{\imath}w\bar{a}n$ as eliglig. Similarly, the expression "to hold (someone's) hand" is rendered as $elig\ tut$ -. The distinction between $elig\ (hand)$ and $ellig\ (fifty)$ is clearly demonstrated in the examples. Therefore, since $ellig\ (fifty)$ already appears in modern Turkish as elli with two consonants, this case does not constitute consonant gemination. The inclusion of this word in previous studies under the gemination category likely stems from a misclassification, which is clarified here.

ėkki/ėki (iki): The word for "two" appears 58 times in the $D\bar{\imath}w\bar{a}n$ as $\dot{e}kki$. A selection of these occurrences is presented here. In most cases, it is written with a shadda (indicating gemination), while in a few instances, it appears without it. Some sources indicate that the first vowel /i/ is rendered as a close-mid front unrounded vowel [\dot{e}], equivalent to IPA [e].

olar ėkki tütüşdiler: onlar ikisi çekiştiler "the two of them quarreled" (DLT-K:256)

ikki neŋ birle ilişdi: iki şey birbirine ilişti. "two things got entangled with one another" (p.94)

olar ikki tawarın üleşdi: o ikisi mallarını ayırıp hisselerini aldılar. "those two divided up their possessions" (p.95)

boy ikki bile alkıştı: kavimler birbirlerini yok ettiler. "the tribes destroyed each other" (p.110)

ol ikki kişi otra araladı: o, iki adamı barıştırdı. "he reconciled the two men" (p.135)

kagun karma bolsa idisi ikki elgin tegir: Kavun yağmalanırsa sahibi de iki eliyle yağmalamaya başlar. "If the melon is plundered, its owner will begin to plunder it with both hands" (p.177)

olar ikki tawar satışgan alışganlar ol: o ikisi devamlı olarak malları aralarında alıp satanlardır. "those two constantly traded goods between each other" (p.226)

ikki begler karıştılar: iki bey ihtilafa düşüp dövüştüler. "two beys got into conflict and fought" (p.260)

olar ikki yiŋ salışdı: onların ikisi birbirine yenlerini salladılar. "the two of them waved their sleeves at each other" (p.263)

olar ikki kayıtışdı: o ikisinin her biri diğerinin ardından gitti. "each of the two followed the other" (p.431)

ikki er birle süŋüşdi: savaşta iki adam birbirleriyle mızraklaştı. "in battle, two men speared each other" (p.511)

ėkkiz/ėkiz (ikiz): The word appears in the *Dīwān* in the context of "twin boys" or "twins."

ikkiz oglan: ikizler "twins" (p.73, 663)

ėkkiz oglan: ikiz çocuklar "twin boys" (DLT-K:615)

ėssiz/ėsiz/1ss1z: This entry was included based on a study that classified it under geminated consonant words in the *Dīwān Lughāt al-Turk*. In various sources, the initial vowel appears either as /1/ or the close-mid front unrounded vowel [e]. The word is used in the form *1ss1z acun* meaning "wicked world." In this context, it connotes "evil" or "wicked."

In Kaçalin's edition, under the entry *esiz* it is glossed as "wicked, evil," while *essiz* is defined as "shameless, hard-hearted, malicious" (see DLT-K: 617). Additionally, the entries *esiz* and *esizlik* are associated with meanings such as "pity" and "evil." Kaçalin clearly distinguishes the semantic differences between forms with single and double /s/. Therefore, since there is no phonetic transformation from DLT to Standard Turkish (ST) or Anatolian Dialects (AD), this cannot be considered a case of consonant gemination.

Irra: This is another word listed among geminated consonant examples in some studies. In the $D\bar{\imath}w\bar{a}n$, its verbal form appears as v bol- ("to be ashamed"). Although the modern reflex of this root is the Turkish v (to feel shame), no variant forms such as v or v have been found in modern Turkish or Anatolian dialects. Therefore, there is no evidence of consonant gemination from DLT > ST or DLT > AD.

er ır boldı: adam utandı "the man was ashamed"

urra: utanma "shame, feeling of shame" (DLT-K:623)

kakkuk: Although claimed to appear in the $D\bar{\imath}w\bar{a}n$, our investigation could not locate such a word. Instead, kak ("dry, dried") is found (DLT-K: 638). When used in the expression kak kok, it means "dried fruit" (DLT-K: 690). However, there is no evidence that these two words ever fused into a compound form such as kakkuk. Furthermore, no entries resembling kakuk/kakok or kakkuk/kakkok exist in dialect dictionaries or the contemporary TDK Dictionary (2025). Thus, this word cannot be considered an example of consonant gemination.

kurrıh/kurıh: In Dīwān Lughāt al-Turk, the expression "kurrıh kurrıh" is recorded as a term used to call a foal (DLT-K:36). In Anatolian dialects, its equivalent is *kurık*, meaning a donkey's offspring or *sıpa* (foal) (see: Yeşilkaya, 2007 for further details).

kurth: a word used to call a foal left behind its mother

kurrıh: kurrıh kurrıh tayı çağırmak için kullanılan söz "a reduplicated vocative form used to call a foal"

(DLT-K:713)

The same page also includes the expression kurr kurr with an equivalent meaning.

ottuz (thirty): This word appears as ottuz in the DLT.

ottuz: ottuz içip kıkralım "let us drink in threes and shout" (BA, vol. I, p. 142; DLT-K:769).

ottuz yarmak: thirty coins. ottuz içelim: "let us drink three times" (DLT-K:769).

sekkiz (sekiz): Found in the DLT as *sekkiz* and *sekiz* (DLT-K:796).

sekkiz: eight, cited on p. 813 of BA; indexed on p. 513 of the same edition.

sokku/soku: The word is recorded in the DLT as *sokku*, meaning *soku* or mortar (as in a mortar and pestle).

soķ-: tuzu döverek inceltmek "to pound, especially salt, to make it finer" (BA, vol. III, p. 226; DLT-K:808).

The standard Turkish form is soku. According to dialect and historical dictionaries, it appears as soku or sokku in folk usage, and sokku in Old Turkic. Due to the presence of derivational suffixes (as in -gu/-ku), this word and others formed similarly (e.g., bakku) are excluded from the scope of this study, which focuses on root-based gemination.

tikkü/tikü: In the DLT, the forms *tikü* and *tikkü* are used to denote a piece of meat or a morsel (DLT-K:861). As no contemporary equivalents could be found in dialect or historical dictionaries, and no direct correspondence in modern dialects exists, no conclusive phonological comparison could be made regarding gemination.

yetti/yeti: The word *yetti* is attested in the DLT, particularly in the expression *yetti kat kök* ("seven-layered sky"). A footnote explains that the original shadda (gemination mark) was later erased, and the word now appears as *yeti* (see: BA, vol. III, p. 27).

yetti: yedi "seven" (BA, vol. III, p. 27).

yeti: found under the lemma yeti, in the phrase yetti kat kök – "seven-layered sky" (DLT-K:945–946).

3.2. Reflections and Distribution of Geminated Consonant Words in Dialects

In light of the data examined, the geminated consonant words that have either direct counterparts in Anatolian vernaculars or exhibit phonological resemblance are evaluated below. In this context, particular attention is given to the words arrıg, bakku/baku, basıg/bassık, bellüg/belgülüg, ėkki/ėki, ėkkiz/ėkiz, ellig, ėssiz/ėsiz/ıssız, ırra, kakkuk, kurrıh/kurıh, ottuz, sekkiz, sokku/soku, tikkü/tikü, and yetti/yeti, which were identified by İnce (2017) and other researchers as examples of geminated consonants. Some of these items were excluded from the analysis due to derivational suffixes or orthographic factors. Based on the detailed evaluations above, the words bakku/baku, basıg/bassık, bellüg/belgülüg, ellig, ėssiz/ėsiz/ıssız, ırra, kakkuk, sokku/soku, and tikkü/tikü were deemed unsuitable for categorization under geminate consonant formation. Therefore, within the scope of this

study, the words arrıg, $\dot{e}kki/\dot{e}ki$, $\dot{e}kkiz/\dot{e}kiz$, $\dot{k}urrh/\dot{k}urrh$, ottuz, sekkiz, and yetti/yeti — all of which demonstrate both historical attestation and contemporary spoken usage in Anatolian dialects — have been included in the final evaluation.

Table 1. Frequency of Usage of Geminated Consonant Words in Dīwān Lughāt al-Turk

Word	Number of Occurrences
arrıg	2
ėkki	58
ėkkiz	2
kurnh/kunh	3
ottuz	2
sekkiz	2
yetti/yeti	3

3.3. Correspondences of the Identified Words in Anatolian Dialects

All of the lexical items analyzed above are attested in various regions of Anatolia. These structures, which exemplify phonological developments along the axes from Standard Turkish (ST) to Anatolian Dialects (AD) and from ST to $D\bar{\imath}w\bar{a}n$ $Lugh\bar{a}t$ al-Turk (DLT), should not be interpreted solely as phonetic phenomena. Rather, they also function as markers of semantic intensification, emphasis, or dialect-specific phonetic realization. Expressions such as arrr su ("pure water"), ekki yumurta ("two eggs"), ikkiz $do\check{g}du$ ("twins were born"), kurlk ahrrdadir ("the foal is in the stable"), kurlk kurlk gel ("come along, foal"), ottuz koyun ("thirty sheep"), sekkiz kilo ("eight kilograms"), and yeddi kat $go\check{g}$ ("seven layers of sky") are still in active use within a variety of Anatolian dialects.

Among these, the form $\dot{e}kki$ ("two") stands out as one of the most frequently attested examples of geminate consonant usage in $D\bar{i}w\bar{a}n$ $Lugh\bar{a}t$ al-Turk, with 58 documented instances. In Anatolian dialects, this form is typically employed with emphatic function, especially in contexts referring to children or animals. Similarly, the word $\dot{e}kkiz$ ("twins") appears in a limited number of examples, while numerals such as ottuz ("thirty") and sekkiz ("eight") are directly attested in geminate consonant forms. The variant yeddi/yetti/yeti ("seven") retains its current usage with the consonant /d/ in most Anatolian dialects. These numerals, as recorded in DLT, are reflected in dialectal forms such as dokkuz ("nine"), ikki ("two"), sekkiz ("eight"), yeddi ("seven"), and ottuz ("thirty"). Such forms have been documented particularly in provinces including Adıyaman, Artvin, Elâzığ, Diyarbakır, Gaziantep, Hatay, Kars, Kilis, Malatya, Mardin, and Urfa (see: Turkish Language Association [TDK] Dialect Dictionary, 2024). Moreover, it has been observed that the widespread use of geminate consonant numerals is a prominent feature in the Southeastern Anatolia region (Yeşilkaya, 2023). Karahan (1996) also confirms the occurrence of consonant gemination in numerals within Eastern Anatolian dialects.

The word arrig/arr ("pure, clean") rarely appears in a geminated form in vernaculars. This instance reflects a gemination process along the ST > DLT axis (i.e., arr > arrig).

When considered along the DLT > AD direction, the transformation from *kurnh/kurnh* to *kurnk* can also be identified in limited cases. In regions such as Gümüşhane, Erzurum, Kars, Erzincan, and Adıyaman, the form *kurnk* continues to be used with meanings such as "colt," "foal," or "donkey's offspring" (see: TDK Dialect Dictionary, 2024; Yeşilkaya, 2007).

In general, these lexical items demonstrate phonetic continuity across both historical texts and modern dialects. While geminate consonants have gradually been simplified in the written form of the language, they have largely been preserved in vernacular speech. This simplification in Standard Turkish (ST) likely stems from a linguistic tendency toward greater fluency and ease of articulation. This phenomenon can be explained in terms of the natural evolution of language and articulatory economy. Furthermore, the absence of a standardized written language during the time $D\bar{u}w\bar{u}n$ $Lugh\bar{u}d$ $ulgh\bar{u}d$

Nevertheless, factors such as historical background, emphatic expression, and phonetic variation have enabled the persistence of geminate consonant forms. The geminated pronunciations observed in Anatolian dialects are directly associated with historical Turkic varieties. In this regard, $D\bar{\imath}w\bar{a}n$ $Lugh\bar{a}t$ al-Turk serves as a unique corpus, offering invaluable insights into the phonological and lexical characteristics of different Turkic tribes and dialects.

Geminate consonants still found in some dialects and Anatolian vernaculars reflect the adaptability of the language to regional variation. For instance, in certain Northeastern Anatolian dialects, the word *çocuk* ("child") is pronounced as *çoccuk*, while in Eastern dialects, numerals are still pronounced as *ikki* ("two"), *dokkuz* ("nine"), and *sekkiz* ("eight"). Similar geminate forms are also attested in Chagatai, Modern Uyghur, and Uzbek (Çakmak, 2017, p. 96). Examples such as *sabbah* (for *sabah*, "morning"), *recca* (*rica*, "request"), *commart* (*cömert*, "generous"), *hamman* (*hemen*, "immediately"), and *eppey* (*epey*, "quite a bit") in Anatolian dialects (Yeşilkaya, 2007, pp. 25–37) further indicate the continuity of gemination. Additionally, moving westward, common geminated forms such as *amma* ("but") and *eşşek* ("donkey") are frequently encountered. In Iraqi Turkmen Turkish, geminate forms such as *yeddi* ("seven"), *kaşşık* ("spoon"), *dokkuz* ("nine"), and *ottuz* ("thirty") are also attested (Bayatlı, 1996).

A closer examination of these items reveals that geminate consonants are particularly concentrated in numerals. The attestation of forms such as *ikki*, *yeddi*, *sekkiz*, and *ottuz* in *Dīwān Lughāt al-Turk* provides strong historical evidence for the phonetic basis of gemination in this lexical category. Although gemination is also observed in other types of words, its prevalence in numerals is particularly striking, indicating a more marked continuity in phonological patterns within this subset.

3.4. Phonetic Features and Transformations in Words with Geminate Consonants

A phonetic-level analysis of lexical items containing geminate consonants reveals certain distinctive sound changes when compared with Standard Turkish (ST). Notably, the preservation of the closed front vowel $/\dot{e}/$ in the words ekki ("two") and ekkiz ("twin") is particularly striking. This phenomenon indicates the prevalence and historical continuity of the closed $/\dot{e}/$ vowel in Anatolian dialects.

In the case of the word yedi ("seven"), recorded as yetti/yeti, a phonological shift from /d/ to /t/ is observed. The fact that the word is written with the Arabic letter $\dot{}$ (indicating /t/) in $D\bar{t}w\bar{a}n$ $Lugh\bar{a}t$ al-Turk further supports this phonetic transformation. Such examples demonstrate that other phonological shifts often accompany gemination, both at the orthographic and phonetic levels.

Additionally, some words exhibit geminate consonants in writing without the use of the traditional

shadda diacritic, although this is clarified in footnotes. For instance, the pronunciation of the word dokuz ("nine") is noted as [dokuz]. Sağır (1995, p. 396) characterizes such geminate consonants as uvular or laryngeal consonants pronounced with emphatic voicing.

In the case of the word $kurr\iota h$, alongside gemination, a sound change from /h/ to /k/ is also noteworthy. These examples indicate that the process of gemination is not isolated but is frequently accompanied by other intralexical phonemic transformations.

4. Findings and Discussion

This study examines a total of sixteen lexical items found in $D\bar{t}w\bar{a}n$ $Lugh\bar{a}t$ al-Turk (DLT) that have been previously classified under the phenomenon of "consonant gemination" in various scholarly works. The phonetic correspondences of these items were analyzed in relation to their forms in Anatolian dialects. As a result of this analysis, it was determined that only seven of these words are both attested with geminate consonants in DLT and continue to exist in Anatolian dialects today. The remaining examples were excluded from the scope due to factors such as affixation, orthographic variation, semantic shift, or phonetic properties that do not constitute true gemination.

Among the seven retained words, numeral terms are particularly prominent. Forms such as *ikki* ("two"), *ottuz* ("thirty"), *sekkiz* ("eight"), and *yetti* ("seven") are attested with geminate consonants in both DLT and contemporary dialects. This pattern observed in numerals appears to be more consistent in terms of phonetic continuity compared to other word classes. This may be attributed to the fundamental and frequently used nature of numerals in language, rendering their phonetic forms more resistant to change.

Considering the diversity and historical layers of Anatolian dialects, it becomes evident that the phenomenon of consonant gemination is not merely the result of individual or regional articulatory differences but is also inherently linked to the phonetic and phonological structure of Turkish. Moreover, this study revisits a number of words previously misclassified as having geminate consonants; formations derived through derivational suffixes or misinterpreted shadda annotations have been critically re-evaluated and excluded. This underscores the importance of accounting for affixes, morphological analysis, and historical orthographic conventions in phonetic studies.

An examination of the DLT corpus reveals that the word *ikki* ("two") appears 58 times. Other numeral terms such as *ottuz*, *sekkiz*, and *yetti* are also directly attested in geminate form. Words like *arrıg* ("pure"), *kurrıh* ("colt call"), and *ikkiz* ("twin") appear less frequently, yet still retain geminate consonant forms in their recorded instances.

Overall, this study presents a comparative overview of consonant gemination as it manifests in both historical sources (DLT) and contemporary Anatolian dialects. It demonstrates that the geminate forms seen in numerals and certain lexemes contribute to a phonetic continuity that reflects a broader historical coherence within the Turkish language.

5. Conclusion and Evaluation

In this study, lexical items featuring consonant gemination attested in *Dīwān Lughāt al-Turk* (DLT) were identified and comparatively analyzed in relation to their equivalents in Standard Turkish (ST) and Anatolian Dialects (AD). The analysis revealed that seven words exhibit geminated consonants both in

Dîvânu Lügâti't-Türk'te Geçen İkiz Ünsüzlü Kelimeler ve Bu Kelimelerin Anadolu Halk Ağızlarına Yansımaları / Yeşilkaya, M.

DLT and in Anatolian dialects with comparable phonetic structures. These include: arrng, ėkki/ėki, ėkkiz/ėkiz, kurrıh/kurıh, ottuz, sekkiz, and yetti/yeti.

The systematic gemination observed particularly in numerals suggests that this phonological process is not incidental but rather constitutes a structural and formulaic feature of Turkish. The continued presence of these forms in Anatolian dialects further illustrates the persistence and phonetic tradition of such sound changes in spoken varieties of the language.

In addition, several phonetic shifts observed in specific lexemes—such as h > k (as in kurr h > kur k) and t > d (as in yetti > yeti)—indicate that this phenomenon persists with regional variation.

During the study, it was also determined that nine items previously classified as containing geminate consonants do not meet the phonetic or morphological criteria for true gemination. This finding demonstrates that phonetic analysis must not rely solely on surface forms but must be supported by indepth morphological analysis. Overlooking the morphological status of affixed forms may lead to erroneous interpretations.

In conclusion, consonant gemination is not merely a phonetic phenomenon, but also a reflection of the historical and sociolinguistic dimensions of Turkish. Future comparative studies investigating the presence of gemination in other historical texts and Turkic dialects will further enhance our understanding of this linguistic phenomenon.

Abbreviations and Other Symbols

AD : Anatolian Dialects

BA: Besim Atalay

cf. : confer (see, compare)

vol. : volume

DLT : Dīwān Lughāt al-Turk

DLT-K: Dīwān Lughāt al-Turk, edition prepared by Mustafa S. Kaçalin

IPA : International Phonetic Alphabet

ST: Standard Turkish

p. : page

TFA : Turkish Phonetic Alphabet

TTD : Dialects of Turkish in Turkey

: backward reference

> : forward reference

ė : close-mid front unrounded vowel [ė]

ķ : uvular /k/ consonant [k]

: nasal /n/ consonant [ŋ] ŋ

References

- Akün, Ö. F. (2008). In *Encyclopaedia of Islam by the Turkish Religious Foundation* (Vol. 25, pp. 9-15). Istanbul: Turkey Diyanet Foundation Publications.
- Aktan, B. (2010). Dîvânu Lügâti't-Türk'ün Söz Varlığında Yer Alan İkilemeler. Selçuk University Journal of Studies in Turcology, 1-12.
- Atalay, B. (1985). Divanü Lûgat-it Türk Tercümesi. Ankara: Türk Tarih Kurumu.
- Bayatlı, H. K. (1996). Irak Türkmen Türkçesi. Ankara: TDK Publications.
- Bozok, E. (2021). Temizliğin Türkçesi Tarihî ve Çağdaş Türk Lehçelerinde Temizlik Kavram Alanı. Istanbul: Hiperyayın.
- Cakmak, S. (2017). Uygurca Yazılmış Meslek Risalelerinin Ses Özellikleri. *International Journal of Uyghur Studies*, 81-103.
- Doğan, Y. (2016). Kâşgarlı Mahmûd'un Dîvânu Lugâti't-Türk Adlı Eserinin Dilbilim Açısından İncelenmesi. *Gaziosmanpaşa University Journal of Faculty of Theology*, 1-20.
- Eraslan, K. (2004). Tarihi yazı dillerinde -ğu/-gü, -ku/-kü isim-fiil eki, yapısı, işlevi, kullanılışı ve genişlemiş şekilleri. In Collective, *Zeynep Korkmaz Armağanı* (pp. 117-146). Ankara: TDK Publications.
- Eraslan, K. (2012). Eski Uygur Türkçesi Grameri. Ankara: TDK Publications.
- Ercilasun, A. B., & Akkoyunlu, Z. (2014). *Kâşgarlı Mahmud Dîvânu Lügâti't Türk*. Ankara: TDK Publications.
- Erdem, M., & Cosun, G. (2013). Divânü Lügâti't-Türk'te Ettirgen Yapılar. *Kahramanmaraş Sütçü İmam University Journal of Social Sciences*, 49-58.
- Ergin, M. (1980). Türk Dilbilgisi. Istanbul: Boğaziçi Publications.
- Guliyev, A. (2024). "Divani-Lüğət-it-Türk" abidəsindəki antroponimlərin leksik-semantik xüsusiyyətləri. Türkologiya, 1(42), 42–53. https://doi.org/10.59849/2313-5204.2024.1.42
- Gülensoy, T. (2014). Dîvânu Lügâti't-Türk'te Adı Geçen "Kırpılmış Türkçe"nin Anadolu Ağızlarındaki İzleri. *Diyalektolog*, 1-5.
- Günşen, A. (2011). Dîvânu Lugati't-Türk'te "Yaklaşma" Bildiren Fiiller. *Turkish Studies International Periodical For the Languages, Literature and History of Turkish or Turkic*, 235-248.
- Gabain, A. V. (1959). Die alttürkische. In Fundamenta I (p. 26). Wiesbaden.
- İnce, Y. (2017). Tarihî Türk Lehçelerinde Ünsüzler. *Unpublished Phd dissertation*. Elazığ: Fırat University.
- Kaçalin, M. S. (2008). Dîvânu Lügâti't-Türk. In *Turkish Diyanet Foundation Encyclopedia of Islam* (Vol. 9, pp. 446-449). Istanbul: Turkish Diyanet Foundation.
- Karaağac, G. (2010). Türkcenin Ses Bilgisi. Istanbul: Kesit Publications.
- Karahan, L. (1996). Anadolu Ağızlarının Sınıflandırılması. Ankara: TDK Publications.
- Korkmaz, Z. (2017). Türkiye Türkçesi Grameri (Şekil Bilgisi). Ankara: TDK Publications.
- Mahmud, K. (2023). *Dîvânu Lügâti't-Türk*. (M. S. Kaçalin, Ed.) Istanbul: Vakıfbank Culture Publications (The original work was written in 1074).
- Mangır, M. (2021). Türkiye Türkçesi Ağız Çalışmalarındaki Ünsüz Türemesi Tespitleri Üzerine. *Kesit Akademi*, 281-312.
- Sakaoğlu, S. (1991). Dîvânu Lügâti't-Türk ve Türk Halk Şiiri. *Yearbook of Turkish Language Studies Belleten*, 97-111.
- Sattorova, I. Z. (2021). About Uzbek editions of $D\bar{\imath}w\bar{a}n$ Lughāt al-Turk. The American Journal of Social Science and Education Innovations, 3(01), 92–99.

- https://doi.org/10.37547/tajssei/Volume03Issue01-18
- Sağır, M. (1995). Anadolu Ağızlarında Ünsüzler. *Yearbook of Turkish Language Studies Belleten*, 391-409.
- Taş, İ. (2014). Divanü Lügati't-Türk'ün Çevirilerinden Hareketle Fonolojik Bazı Belirlemeler (1). *Yearbook of Turkish Language Studies Belleten*, 203-220.
- TDK. (2024, 06 20). Retrieved from TDK Kâşgarlı Mahmud: https://tdk.gov.tr/divanu-lugatit-turk/kasgarli-mahmud/
- TDK Derleme Sözlüğü. (2024, 10 12). Retrieved from https://sozluk.gov.tr/
- TDK Sözlük. (2025, 06 20). Retrieved from sozluk.gov.tr
- Tekin, T. (2017). Orhon Türkçesi Grameri. Ankara: TDK Publications.
- Tekin, T. (2023). (Tanıtma) Manuscrits ouïgours du IX^e-X^e siecle de Touen-houang. In T. Tekin, *Makaleleler 2: Tarihî Türk Yazı Dilleri* (pp. 287-295). Ankara: TDK Publications.
- Tekin, T. (2023). Üçüncü Kişi İyelik Ekleri Üzerine. In T. Tekin, *Altayistik Makaleler 1* (pp. 117-122). Ankara: TDK Publications.
- Tekin, T. (2023). Türk Dil ve Lehçelerinde Ünsüz İkizleşmesi. In T. Tekin, *Makaleler 3 Çağdaş Türk Dilleri* (pp. 201-208). Ankara: TDK Publications.
- Ulusoy, Y. (2024). Çok Dillilik ve Dîvânu Lügâti't-Türk. *International Journal of Turkish Literature Culture Education*, 13(3), 838-848.
- Yeşilkaya, M. (2007). Gerger Merkez ve Köyleri Ağızları. Van: Yüzüncü Yıl University Unpublished Master's Thesis.
- Yeşilkaya, M. (2023). Güneydoğu Anadolu Bölgesi Ağız Atlası. Istanbul: Istanbul Medeniyet University Unpublished PhD Dissertation.