**UNIVERSITY OF PANNONIA**
**MULTILINGUALISM DOCTORAL SCHOOL**

Ali Can Yalçınkaya

# TURKIC LOANWORDS IN HUNGARIAN: A STUDY CONCERNING LOANWORD ADAPTATION

DOI:10.18136/PE.2024.903

Phd Dissertation

Research Supervisors: Prof. Parapatics Andrea, Prof. Szentgyörgyi Szilárd

VESZPRÉM

2024

I would like to express my heartfelt gratitude to the following individuals who have played a significant role in the completion of my Ph.D. dissertation at the Doctoral School of Multilingualism at the University of Pannonia:

First and foremost, I would like to thank my mother, Deniz Yalçınkaya, who has been a constant source of love, support, and encouragement throughout my academic journey. As an accomplished academic herself, her unwavering belief in my abilities has been a driving force behind my achievements. I am deeply grateful to my supervisors, Prof. Andrea Parapatics and Prof. Szilárd Szentgyörgyi, for their invaluable guidance, expertise, and continuous support. Prof. Parapatics, renowned for her research in multilingualism, has provided me with valuable insights and direction. Prof. Szentgyörgyi, an esteemed phonetician, has offered his expertise in navigating the intricacies of my research. I extend my heartfelt appreciation to Prof. Emeritus András Róna-Tas, a distinguished scholar whose inspirations and kind opinions have significantly influenced my work, and Prof. Péter Siptár, an eminent figure in Hungarian phonology whose vast knowledge and wisdom have provided invaluable guidance and inspiration. I would also like to express my sincere gratitude to Prof. Judit Navracsics, who accepted me into the Doctoral School of Multilingualism. Her distinguished academic career and commitment to fostering excellence have been truly inspiring. I am deeply indebted to Prof. Vincent van Heuven, an eminent figure in comparative linguistics, whose valuable lectures have inspired and shaped my research. His methodological ideas have broadened my understanding and contributed significantly to the quality of my work. To my dear girlfriend, Petra Juszkan, I am immensely grateful for her unwavering support and understanding. Despite enduring my not-so-fluent-at-all Hungarian when seeking her assistance, she has been a pillar of strength throughout this journey, providing much-needed moral support and laughter. I would also like to express my gratitude to my professors from Ankara University, both from my MA and BA programs. Their dedication to teaching, rigorous approach to linguistics, and unwavering support have equipped me with a solid foundation and the confidence to pursue my research interests. Lastly, I extend my thanks to all my friends and colleagues who have contributed to my research, either through insightful discussions or participation in data collection. Your input and collaboration have been invaluable. It is with immense gratitude that I acknowledge the support and guidance of these remarkable individuals.

## Table of Contents

**1. Introduction ..... 4**

**1.1 Area of research ..... 5**

1.1.1 Main concepts ..... 9

1.1.2 Phonological stance model (Nativization through production) ..... 10

1.1.3 Perceptual stance model & Phonetic approximation (Nativization through perception) ..... 15

1.1.4 Combined accounts ..... 23

**1.2 Objectives and importance of research ..... 28**

**1.3 Research questions ..... 30**

**1.4 Scope and limitations ..... 31**

**1.5 The exposition of the Turkic material in Hungarian ..... 35**

**1.6 Data ..... 47**

**2. Methods ..... 50**

**3. Results ..... 66**

**3.1 Borrowing process ..... 68**

**3.2 Adaptation process ..... 92**

**3.3 Patterns ..... 120**

3.3.1 Heavy syllables ..... 121

3.3.2 Nasal palatalization ..... 122

3.3.3 High back unrounded vowel ..... 123

3.3.4 Voiced velar fricative ..... 123

3.3.5 Labialization of vowels ..... 124

3.3.6 [±low] quality of the front unrounded vowels /a/ and /æ/ ..... 124

**4. Discussion ..... 125**

**5. Conclusions ..... 130**

**6. References ..... 131**

# List of Tables

Table 1. Rhotacism in Chuvash and Hungarian (Marácz, 2016) ..... 39
Table 2. Number three in Finno-Ugric languages (Laki, 1960, p. 4) ..... 46
Table 3. An example of the classical phonetic distance calculation ..... 50
Table 4. Refined distance ..... 56
Table 5. Refinement for deletions ..... 56
Table 6. Further refinement for super-sub features ..... 58
Table 7. Refinement for gemination ..... 58
Table 8. Refinement for metathesis ..... 59
Table 9. Refinement for the voiced velar fricative ..... 60
Table 10. Distinctive features matrix of Turco-Hungarian Phonemes (1) ..... 61
Table 11. Distinctive features matrix of Turco-Hungarian Phonemes (2) ..... 61
Table 12. Initial adaptation rates of WOT copies in Hungarian ..... 68
Table 13. Adaptation rates of reconstructed Old Hungarian copies in Modern Hungarian . 94
Table 14. Borrowing patterns ..... 121
Table 15. Adaptation patterns ..... 121

# CHAPTER ONE

# Introduction

## 1.1 Area of research

Loanword adaptation is a phonetic, phonological, and etymological phenomenon regarding the words that are derived from foreign languages and regarding the phonological or phonetic assimilations that occur to adapt to the features of the target language. There are diverse ideas in the literature concerning the background of these changes. Some of them consider the adaptations as minor phonetic transformations that are the results of perceptual processes (Peperkamp, 2005, p. 341) and some of them identify them as phonological equivalences: "Our results support the view of the loanword input to the phonology the borrowing language is immediately interpreted as a phonological representation by it and handled by its constraint set" (Paradis & LaCharité, 1997, p. 380). There are also some theories of interaction such as (Boersma & Hamann, 2009) and (Kenstowicz & Suchato, 2006) claiming that both viewpoints apply to some degree in their interaction theories. The phonological versus phonetic nature of loanwords has been in debate for many years now and it is a curious case whether languages follow similar patterns in nativizing the shape of loanwords or not. Nevertheless, as more languages are examined, fewer regularities seem to remain (Kertész, 2006, p. 12).

Loanwords that were derived from Turkic languages in Hungarian – being rather numerous – have been attracting various etymologists, scholars, and researchers on the topic for centuries, making them study the processes and historical periods of the borrowings to explore the actual relationship between the two languages to find out whether they are genetically related or had just shared a 'Sprachbund'. There are fairly extensive etymological dictionaries and papers on Turkic loanwords in Hungarian thus there is a strong consensus about the origins and historic journeys of many of these words. Although the phonological behavior of Turkic loanwords in Hungarian was studied in the literature (Bárcsi 1972, Ligeti 1986, Róna-Tas & Berta 2011, Honti 2017), the main aim of those studies was primarily etymological, to reveal the relatedness between the two by language contact. The regularities,

patterns, and typology ascertainment of loanword phonology in the concept of the linguistic subfield 'loanword adaptation' is still a puzzle to work on.

New vocabulary items are acquired by languages continuously, all the time. These new words are not always produced within those languages but are often borrowed from other languages. There is no doubt that although the reasons vary, borrowings occur out of necessity. One objection to this may be the fact that some societies refuse the loans in particular situations and instead produce their own words due to various cultural, pragmatic, and political reasons. It is known that some Amerindian languages did not acknowledge the lexical acculturation for some words and manufactured their own definitions such as "little maggots" for "rice" (Brown et al., 1994, p. 95). Of course, this cannot be seen as a general stance against foreign words in native American languages since they are known to have borrowed many words as well, especially from Spanish. A similar attitude on a much broader and more conscious level can be seen in Turkish in the early to mid-20$^{th}$ century that occurred as a government policy with a nationwide campaign to purify the language from heavy foreign influence that was mostly Arabic, Persian, and French. For many loanwords some of which were used frequently in daily speech, Turkic versions were proposed and when there was no such option, new words were produced from existing word roots. Some of these attempts have been successful as in *bilgisayar* (literally meaning 'information counter') instead of *computer* while some attempts were unsuccessful and forgotten as in *çekimci* 'taker' instead of *kameraman*. Technological, cultural, or conceptual innovations, more prestigious reflections of a foreign language/culture, intensive language contact, revealing unknown geographical or cultural terms of foreign societies, or even official language policies can cause borrowings of foreign words. It can be observed clearly that borrowings are mostly asymmetric so that more words are adopted from one side to the other. The reason for this is similar to the above-mentioned ones. It is known for instance that the Germanic tribes had much interaction with the Roman Empire for centuries and borrowed many Latin words asymmetrically.

"The 'early' loanwords were borrowed in the Roman and sub-Roman period. They reflect the superiority of Roman civilization and mainly denote concrete things of everyday life adopted from the Romans. A frequently practiced semantic classification of the early

loanwords shows that they can be attributed mainly to the fields of agriculture, building, food, household, and commerce. Not surprisingly these loanwords linguistically, or more strictly speaking, phonologically, exhibit a high degree of integration into the native languages." (Wollman, 1993, p. 3)

The excessive Arabic influence in Ottoman Turkish – of which the traces are still clear in modern Turkish today – is mainly a result of Arabic being the language of Islam and for a period of time, a predominant language of science, thus more legitimate than Turkish. After acquiring the title of *caliph*, Ottomans started using Arabic as the *langue officielle* in bureaucracy (Tülücü, 1997, p. 42) and also it was more common than Turkish in literature as well. This resulted in thousands of borrowings from Arabic naturally. The borrowings between Arabic and Turkish were/are eminently asymmetric.

The type of word categories that are prone to be borrowed are not very explicit. It is apparent that grammatical factors play a fundamental role in this issue. Verbs appear to be harder to borrow compared to nouns since they need more grammatical adaptation than nouns. What is called core vocabulary likewise, is rarely borrowed (e.g., numbers, organs, kinship names). Some adaptations are inevitable so that they can be understood and used in L1. Languages with gender and inflection classes regulate the loanwords morphologically so that the borrowings will fit in the syntactic system. The amount of time a particular loanword survived in L1, how familiar the speakers of L1 are with L2, and their point of view for L2 are the main criteria that define how strong the adaptation will be.

"If the donor language is well known and/or the loanword is recent, recipient language speakers may choose not to adapt the word in pronunciation, and they may borrow certain inflected forms from the donor language. In this way, English borrowed plural forms of words from Greek and Latin (phenomenon/phenomena, fungus/fungi, crisis/crises), and German even borrowed a few case forms (e.g. the genitive in das leben Jesu 'the life of Jesus')" (Haspelmath, 2009, p. 42).

A great interest in loanword adaptations has taken place in the last decades. The constraint and repair model of sound changes instead of a rule-based one is seemingly the main

motivation for the increasing studies on loanword adaptation (Kenstowicz & Suchato, 2006, p. 2).

"The interest of phonologists in loanword adaptation has mushroomed since the emergence of constraint-based theories because the patterns of sound changes in loanwords offer abundant positive evidence for the phonological constraints that a language places on output"(Paradis & LaCharité, 2011).

Eventually, two main models have become predominant, one being the phonological stance model which is based on the nativization through production view, and the other being the perceptual stance model (phonetic model) which is based on the nativization through perception view (Paradis & Tremblay, 2009, p. 212). The Perceptual/Phonetic Model suggests that loanword adaptation occurs entirely in perception. According to this model, borrowers, who may be monolingual or bilingual in monolingual L1 mode, lack access to the phonology of the source language. Adaptations happen as a result of the borrower's misinterpretation of the foreign source term. Supporters of this model argue that loanwords lack phonological representation, and the input to loanword phonology is a superficial non-linguistic acoustic signal. Borrowers perceive the incoming raw phonetic signal and create native phonological representations to fit the surface input into their native phonological system. In contrast, the Phonological Model assumes that competent bilingual speakers functioning in bilingual mode borrow words from their second language to fill gaps in their first language (L1). These speakers have a comprehensive understanding of both phonological systems and can access the categories and structures of the second language (L2). Borrowers receive the underlying representation of the borrowed word from their mental dictionary for L2 and construct its surface representation while speaking L1. Loanword adaptation in this model is based on phonological equivalency rather than perceptual similarity. When the phonological structures of the two languages do not match, the loan form is changed to the next-closest substitution based on the recipient language's phonological properties. Phonetic forms of source items and surface variations of phonemes in L2 have less importance in loanword adaptation within this model.

A third alternative model which is grounded on the Optimality Theory proposes that the input is a surface representation of L2, but indicates that also the phonological features of L1

play a role in the adaptation (Boersma & Hamann, 2009). When it comes to loanword adaptation, this theory considers both faithfulness constraints and markedness constraints. Faithfulness constraints aim to preserve the phonological structure of the source language word in the borrowed word. They penalize any changes that deviate from the original form. Markedness constraints, on the other hand, reflect the phonological patterns and constraints of the borrowing language. They favor the adaptation of the borrowed word to conform to the phonological regularities of the borrowing language. The interaction of these constraints results in the adaptation of loanwords. The specific patterns of adaptation depend on the ranking of constraints and the relative importance of faithfulness and markedness. For example, if markedness constraints in the borrowing language prioritize the deletion of certain sounds, they may override faithfulness constraints and lead to the omission or modification of those sounds in the borrowed word. Optimality Theory provides a formal framework for analyzing the complex interplay of constraints that shape the adaptation of loanwords, allowing linguists to account for both the similarities and divergences between the source and borrowing languages in the phonological structure of loanwords.

### 1.1.1 Main concepts

This section provides an overview of the main concepts used in the dissertation.

- Adaptation
- Adaptation patterns / systematic adaptation patterns
- Allophone / allophonic variation
- Alveolarization
- Assimilation
- Asymmetric borrowing
- Borrowing
- Category preservation principle
- Category proximity principle
- Chuvash
- Conservative strategy
- Constraint
- Cuman
- Deep structure
- Deletion
- Distinctive feature
- Edit distance
- Faithfulness
- Gemination
- Heavy syllable
- Importation

- Insertion
- Isomorphism Hypothesis
- L1 (recipient language)
- L2 (donor language)
- Labialization
- Levenshtein distance
- Loanword
- Loan phonology
- Mapping
- Markedness
- Middle layer words
- Nativization
- Nativization through production & perception
- Oghur
- Operation cost
- OT(optimality theory)
- Palatalization
- Parsing
- Perceptual mapping
- Perceptual stance model
- Phonemic alignment
- Phonemic distance
- Phonetic approximation
- Phonetic and phonological segments
- Phonological stance model
- Pronunciation distance
- r-Turkic
- Rule-based approach
- Salience
- Substitution
- Surface level
- Underlying structure
- Vowel epenthesis
- Vowel harmony
- Quantification
- Weighted operation
- WOT (West Old Turkic)

Below, three main models in the area of loanword adaptation are presented, each accompanied by detailed explanations. This is undertaken to demonstrate comprehension of the developmental trajectory of literature in the relevant field, showcasing significant studies by notable researchers.

### 1.1.2 Phonological stance model (Nativization through production)

In his 1954 article, the structuralist Zellig Harris makes a remark on the phonological nature of non-native sound processing by stating the following:

"Clearly, certain behaviors of the speakers indicate perception along the lines of the distributional structure: for example, the fact that while people imitate non-linguistic or foreign-language sounds, they «repeat» utterances of their own language (i.e. they reproduce the utterance by substituting, for the sounds they heard, the particular corresponding variants which they habitually pronounce; hence the heard sounds are processed as members of correspondence sets) There is also evidence of processing of sounds in terms of their morphophonemic memberships" (Harris, 1954, p. 149).

Hyman (1970) argues against the *phonetic approximation theory,* according to which, the dynamics of the adaptations are explained with physical phonetics only. He indicates that adaptations cannot be explained in mere phonetic terms. In order to advocate this, Hyman attracts attention to the fact that the realization of L1 sounds in L2 is different than in L3. He proposes the notion of *Sprachgefühl,* intuitive feeling for the nature of a language, or in his case, substituting correct sounds for adaptations intuitively. He concludes that morpheme structure conditions and phonological rules apply to adaptations.

Lovins (1975) discusses the phonological adaptations of Western loanwords in Japanese and presents the systematic processes (manually) of phonological response with thorough analyses and several phonological rules. She adopts a *rule-based approach* to investigate the loanword adaptation mechanisms, using an aural perception method by identifying a pronounceability or distinguishability hierarchy of foreign syllables that relates directly to their degree of nativization.

Danesi (1985) works on 233 loanwords in Canadian English. These are Italo-Canadian vocabulary, originating from Italian native speakers who migrated to Canada. He puts a strong emphasis on the importance of data-oriented approaches when it comes to dealing with the nativization questions, instead of theory-based ones. The essence of this study is that the author detects several phonological processes in nativization, such as *vowel suffixation* of words with consonantal endings or gender assignment to nouns. Danesi also refers to the fact that orthography does not play a significant role in nativization.

Jacobs & Gussenhoven (2000) ground on a "universal phonological vocabulary" which theoretically comprises all the processable sounds of human language and it is thus much

greater than the acoustic set of a particular native language. They argue that the freshly encountered loanwords are not mere acoustic signals but are processed phonological data, analyzed by the universal phonological vocabulary. They also argue that loan phonology has similar aspects to native language acquisition.

"If foreign inputs are faithfully parsed, are they also faithfully stored in the lexicon, in the way that first language learners store new native inputs? When speakers are aware of the fact that they are dealing with a loanword, this may indeed well be the case, but of course, loanwords often are restructured, and stored in an adapted shape in the lexicon[…] In loan phonology, it is the production parse that wins and creates changes in the lexicon, whereas in child phonology it is the comprehension/perception parse that wins and creates changes in the constraint hierarchy" (Jacobs & Gussenhoven, 2000, p. 203-204).

LaCharité & Paradis (2005) examine French and English loanwords in different languages. They strongly argue that the borrowings are not processed as surface-level phonetic data and they demonstrate the phonological nature of the nativization of loanwords. In addition, just like Danesi (1985), the authors in this study also stress the insignificance of orthography in nativization. They advocate that phonological categories are preserved in L1 if the same categories exist in both languages. This is defined as the *category preservation principle*. If a category is present in L2 and not in L1, then the adaptation will occur according to the closest category that exists in L1. This is referred to as the *category proximity principle*. Adaptations according to these principles occur regardless of the phonetic proximity. The main aim of their study is to show that the loanword adaptation is categorical. The authors defend the claim that there is a highly systemic process in nativization that works on the phonological ground.

"Since it is phonological, it follows that loanword adaptation provides one of the richest and most readily available sources of information regarding the functioning of phonology. Many languages do not have the complex morphological systems or phonological processes that phonologists heavily depend upon for information about a language; but most, if not all, languages borrow words from other languages. Through loanword adaptations we gain insight into, among other things, the constraints of the borrowing language, segment and

syllable structure, and ultimately, the workings of the human phonological capacity." (LaCharité & Paradis, 2005, p. 254)

Uffmann (2006) discusses vowel epenthesis in loanwords from an empirical and formal linguistic perspective. It challenges the assumption that epenthesis patterns are simple and argues that they are more complex, involving the interaction of three distinct processes: vowel harmony, local assimilation to the preceding consonant, and default insertion. The paper presents a statistical analysis of loanword corpora in Shona, Sranan, Samoan, and Kinyarwanda, revealing that the quality of the epenthetic vowel is determined by the interaction of these processes. It challenges the notion that a default vowel is usually inserted, showing that the choice of the epenthetic vowel is restricted to specific environments. The findings are formalized within an Optimality-Theoretic framework, considering autosegmental representations. The ranking of constraints against feature insertion and spreading determines the preferred strategy in a language. These constraints are scalar and grounded in universal markedness and prominence hierarchies, explaining the preference for certain spreading strategies in different environments. The paper argues that perception plays a minor role in vowel epenthesis, as the high amount of variation across and within languages cannot be sufficiently explained by notions of salience or perceptibility alone. The paper concludes that epenthesis is a more complex process than commonly assumed, influenced by multiple principles and constraints. The data suggest that the probability of feature spreading depends on its markedness, and the paper argues against a purely perceptual account of vowel epenthesis. Instead, it proposes a phonological model based on feature spreading and insertion to explain the observed patterns.

An overhaul of the phonological vs phonetic nativization debate is made (Paradis & LaCharité, 2011) by breaking down the issues and controversies in loanword adaptation. The authors first exhibit the implications that they made out of the debate so far. Some of these are the necessity of a large enough corpus (several hundred borrowings) and precise quantifications. Another one is the importance of the consideration of direct borrowings instead of indirect ones that were borrowed into L1 from L2 via L3. The advantage of a good insight and familiarity of L1 and L2 alongside social and historical contexts is another acknowledgment that they draw attention to. It is argued that if these factors are neglected,

this can mislead the assessment that the phonetic approximation is more reasonable in many cases. Finally, they bring out the fact that whether the borrowers are bilinguals or monolinguals is of crucial importance. About the latter problem, the authors claim that the phonological stance is much more fitting with the indication that borrowers are bilinguals in most cases. 12,452 English and French borrowings in several different languages are examined in the study. A mere 1.8% of the malformations appear as affected by orthography. This shows an analogy with several other studies that are biased on the phonological stance and suggests that the role of orthography is highly insignificant. The study reveals that while 92% of the malformations show phonological cases, 8% are non-phonological. Adaptations, importations, and deletions are addressed as phonological cases. Adaptations are simply replacements of the sounds. Deletions are omitted sounds and thus in a way, a variety of adaptations. Importations, on the other hand, are instances of on-purpose embracement of L2 sounds or phonological features. This is of course seen as a natural result of bilingualism. As the rate of bilingualism increases among L1 and L2 speakers, importations (non-adaptations or deletions) increase as well. This is referred to as *the preservation principle*.

"Crucially, intentional phonetic approximation (importation) is an attempt to have the L1 phonological system accommodate characteristics of L2. Therefore, it can introduce L2 sounds and structures into L1" (Paradis & LaCharité, 2011, p. 764).

As-Sammer (2015) demonstrates in his article 150 English loanwords display a constraint-based processing in Iraqi Arabic, and thus the adaptations strongly accommodate the grammar rules of L1. "All loanword adaptations are in conformity with IA phonology" (As-Sammer, 2015, p. 36). Discussing his findings, he then points out the argument that the main reason for adaptations is a psycholinguistic one and L2 sound structures transform into equivalents that are suitable to L1 phonology. "Few of the English loans preserve their source pronunciation because their input is in full coincidence to IA phonology" (As-Sammer, 2015, p. 37).

### 1.1.3 Perceptual stance model and Phonetic approximation (Nativization through perception)

English loanwords in Cantonese are studied in (Silverman, 1992). He argues that L1 speakers conceive foreign linguistic input according to their native phonology only. In order to do that, speakers create a surface-level phonetic representation of this new input first, only then do they adapt it into their native phonology. Silverman therefore, proposes a two-layer adaptation process; perceptual level and operative level. The former requires a phonetic resolution of the acoustic info into chunks (parsing). There is no prosody at this stage and the main point is the phonetic identification of the input. In this stage, the Cantonese speakers detect some constraints of the English input. They detect the ones that are complementary in both L1 and L2 such as *aspiration* contrasts, and they cannot detect the aspects such as English *voicing* contrasts that are lacking in L1. The second level involves prosodic processing which triggers several phonological operations possibly supplied by *universal grammar*. It applies to the output of the first level. Nonetheless, L2 phonology has no significance at either level. "For any phonetic string, it is only native speakers for whom a fully articulated phonological structure is present" (Silverman, 1992, p. 61). This study, just like several others mentioned below, is an attempt to break down the issue of loanword nativization in terms of physical phonetics.

The study of Yip (1993) is a continuation of Silverman (1992) but it focuses more on the operative (phonological) level. The essence of the idea here is that the operative level occurs on a conservative ground to merely "mimic" the perceived input. Where this is not possible, adaptations occur. The phonological output is as parallel as possible to the perceived input so that insertions, deletions, and substitutions can be avoided. The author refers to this as the *conservative strategy*, which is a linguistic approach guided by several principles: *faithfulness,* which emphasizes maintaining a close resemblance between the phonological output and the perceived input to minimize insertions, deletions, and substitutions; *parse*, where linguistic input is broken down into smaller units to align with the perceived input; and *fill*, the process of filling in gaps in linguistic input as naturally and accurately as possible, ensuring that insertions, substitutions, and deletions are minimized. According to the author, these principles are integral to the conservative strategy and are in accordance with universal

grammar. Insertions, substitutions, and deletions featurally occur however, as most of the English words cannot be perceived unadapted. The role of salience is emphasized in the case of deletions, such as when stops are deleted from word-final consonant clusters as in cases like *cast* or *band*. The author argues that salience also plays a major role in substitutions. Explicitly salient segments such as the coronal fricatives and affricates will create pressure to make sure that they are parsed.

Kenstowicz (2001) argues that perception plays a significant role in the adaptation of loanwords into a new language. He proposes a two-stage model of loanword adaptation, in which the loanword is first perceived by the speaker of the borrowing language, and then produced by the speaker. In the perception stage, the loanword is analyzed by the speaker's phonological grammar, and any segments or features that are not part of the speaker's native language are adapted or deleted. In the production stage, the adapted loanword is produced by the speaker's motor system. Kenstowicz argues that the perception stage is particularly important in loanword adaptation because it is at this stage that the speaker must decide how to interpret the foreign sounds. For example, he discusses the borrowing of the French word *gare* which means 'railway station'. In the Fon language, this word is borrowed as /ga/, omitting the final /R/ sound. The adaptation of this word is intricately analyzed in the context of the Fon phonological system.

How loanwords are adapted and pronounced by speakers of the borrowing language is explored in (Peperkamp & Dupoux, 2003), with a specific focus on the role of perception in this process. The authors propose a new model for understanding loanword adaptations, challenging the traditional view that such adaptations are solely driven by phonological and phonetic constraints. Instead, they argue that perceptual processes play a crucial role in shaping the pronunciation of loanwords. According to the authors, loanword adaptations are influenced by the listeners' perception of the borrowed sounds. They suggest that listeners' perceptual mechanisms actively reshape the phonetic representations of loanwords, making them more aligned with the phonological patterns of the borrowing language. The paper presents empirical evidence from experiments conducted with French listeners who were exposed to loanwords from English. The results demonstrate that the listeners' perception of the English sounds influenced the adaptation process. The listeners tended to assimilate the

borrowed sounds to the closest equivalent sounds in their native French, even if those sounds were not phonetically identical. The authors argue that this perceptual reinterpreting of loanword sounds allows for better integration and processing of borrowed words within the phonological system of the borrowing language. They propose a model that combines both phonological and perceptual factors to explain the loanword adaptation process. The paper highlights the importance of perceptual processes in loanword adaptations and challenges the view that phonological constraints alone drive these adaptations. By considering the role of perception, the authors provide a new perspective on how loanwords are integrated into the phonological systems of borrowing languages.

A model that combines perceptual assimilation and production processes to explain loanword adaptation is proposed in (Peperkamp, 2005). The model suggests that when speakers encounter a foreign word, they perceptually assimilate its non-native sounds to the closest phonetically similar sounds in their native language. This assimilation process allows the foreign word to become perceptually accessible to the speaker. Subsequently, during production, the adapted loanword is produced using the phonological rules and constraints of the native language, which may involve further adaptation to meet the native language's phonotactic patterns. The model is supported by experimental studies, such as an example that involves Japanese speakers perceiving English loanword stops as voiced. This finding suggests that the Japanese speakers perceptually assimilated the English stops to the closest native sounds. Additionally, the model can account for patterns of loanword adaptation observed in different languages. For instance, in Korean, loanwords from English ending in voiceless stops often undergo adaptation with an aspirated stop followed by an epenthetic vowel, particularly when the preceding vowel is tense. Peperkamp attributes this pattern to a phonological rule in Korean that prohibits voiceless stops at the end of words, and the epenthetic vowel is inserted to break up the sequence of voiceless stop consonants.

Fenyvesi & Zsigri (2006) suggest that the adaptation of loanwords with unstressed initial syllables in the source language is influenced by perception and that a perceptual approach provides more insights into the adaptation process compared to a strictly structural approach. The analysis focuses on two immigrant varieties, namely American Hungarians and American Finns, as well as their Old World counterparts. It argues that the adaptation

strategies of the immigrant varieties, which involve both stress-shifting and deletion, differ from those of the Old World varieties, which only involve stress-shifting. This difference is attributed to the relative dominance of auditory input over visual input in the immigrant varieties, while the Old World varieties receive more visual input. The explanation for this difference lies in the fact that English loanwords in today's European societies enter the recipient language partly through writing, making the written form of the loanword as salient as its phonetic form. However, for working-class and peasant-origin Hungarian and Finnish immigrants in the US, who had limited literacy in English, the phonetic form of loanwords would have been more salient due to their reliance on the oral channel. The analysis also highlights the difference between written and oral language. In written language, words are perceived as clear-cut units separated by spaces, while in oral language, stresses might be interpreted as word boundaries, particularly by listeners whose first language follows an initial stress pattern. The analysis argues that perception plays a crucial role in the adaptation of English loanwords in Finnish and Hungarian, particularly in the context of immigrant varieties compared to Old World varieties. The different adaptation strategies are explained through different rankings of constraints in Optimality Theory, with integrity constraints being more dominant when visual input is more prominent, and stress-faithfulness constraints being strengthened when auditory input is more dominant. Perceptually motivated constraints, hierarchies, and interpreted inputs are also highlighted in the analysis to emphasize the role of perception in shaping the loanword adaptation process.

The role of perception in loanword adaptation is discussed in Davidson (2007), specifically focusing on segmental and phonotactic phenomena. The paper mentions previous studies that demonstrate how perception influences the adaptation of tense and lax vowels in Japanese and the assignment of English voiceless stops to the aspirated category in Thai. It aims to extend this research by examining whether perceptual effects continue to influence loanword adaptation as words are passed from one speaker of the borrowing language to another. The specific case under examination involves the acoustic properties of transitional schwa, which English speakers produce between non-native consonant sequences. These transitional schwas have been found to have different acoustic properties than lexical schwas. The study hypothesizes that these acoustic differences occur because speakers do not sufficiently overlap the consonant sequences, resulting in the production of

a transitional vocoid between the constrictions. The study conducted experiments to investigate how English listeners categorize transitional schwas and whether they interpret them as lexical schwas. The results suggest that even with minimal exposure, English listeners do not categorically interpret transitional schwas as if they were lexical schwas. This finding indicates that recipients may infer that the presence of a transitional vowel indicates the production of a non-native consonant sequence. The study also discusses the potential impact of this perceptual information on the recipients' assignment of underlying representations to loanwords. It suggests that sensitivity to fine acoustic details in a borrowing situation, where listeners need to provide underlying representations of words becoming part of the language, may be accentuated. It further discusses the implications of the study's findings on subsequent generations of speakers and potential phonological changes in the language due to borrowing.

The factors influencing the adaptation of loanwords in Japanese are investigated in (Peperkamp et al., 2008). The authors argue that loanword adaptations are influenced not only by phonological rules but also by how native Japanese speakers perceive loanwords. To support their argument, the authors conducted experiments involving Japanese speakers who were presented with various loanwords, some adapted and some not. The results revealed that the adaptation of loanwords was influenced by the phonetic properties of the source language. Loanwords from English, lacking word-final vowels, were more likely to be adapted with a final vowel compared to loanwords from French, which already possess word-final vowels. Furthermore, the authors found that the perception of loanwords by native Japanese speakers played a role in their adaptation. Loanwords perceived as having a strong vocalic release were more likely to be adapted with a final vowel. This suggests that loanword adaptation in Japanese follows a two-step process: perception by native speakers, followed by adaptation according to the phonological constraints of Japanese. These findings have broader implications for understanding loanword adaptation in other languages. They indicate that the perception of loanwords by native speakers may significantly influence the adaptation process. This finding highlights that loanword adaptation is not solely governed by phonological rules but is also shaped by cognitive processes involved in speech perception. Practically, the authors' study has implications for teaching Japanese speakers how to pronounce loanwords from other languages more effectively. By understanding the

perceptual factors influencing loanword adaptation, educators can develop improved methods for instructing pronunciation and help learners accurately adapt loanwords into the Japanese phonological system.

In Broselow (2009), the problem addressed is the learnability of interlanguage rankings in the context of language contact situations. The author focuses on cases of loanword adaptation where systematic adaptation patterns are observed across speakers of the same native language, but the rankings of phonological constraints in the production grammar cannot be easily motivated by the data of either the native or the foreign language. The paper argues that these apparently unmotivated rankings may reflect the influence of the perception of grammar rather than the production grammar. The author presents two cases of loanword adaptation involving the preservation of foreign word stress. In Huave words borrowed from Spanish and Fijian words borrowed from English, the source stress is maintained despite the differences in the adaptation patterns. These patterns have been analyzed in terms of different rankings of a production grammar constraint related to stress preservation. However, the paper argues that the maintenance of source stress is not a result of the production grammar but rather a consequence of the native language perception/decoding principles. The paper proposes a bipartite model in which the perception grammar determines which aspects of the acoustic signal are linguistically significant, leading to the misinterpretation of certain aspects of the contact structures. The perception grammar maps the input onto a more restricted set of phonological representations, providing an impoverished base for the adaptation process. The adaptation patterns observed in loanword adaptation can be explained by the interference from the native language perception grammar, rather than unlearnable rankings in the production grammar. The paper discusses various examples and compares production-oriented and perception-oriented accounts of loanword adaptation. It concludes that the rankings of production grammar constraints in cases of loanword adaptation can be better explained by the influence of the perception grammar, which determines how foreign language stress is interpreted. The perception grammar plays a crucial role in mapping the acoustic signal onto phonological representations and affects the adaptation patterns observed in language contact situations. Overall, the paper argues that interlanguage production grammar rankings that cannot be attributed to universal grammar or input data can be understood as effects of the perception of grammar. The findings suggest

that the perception of grammar influences the adaptation process in language contact situations, shaping the rankings of phonological constraints in the production of grammar.

The phenomenon of vowel epenthesis in the speech of Korean learners of English is investigated in de Jong & Park (2012). In Korean, there are no word-final consonants, so Korean learners of English are expected to insert a vowel after a word-final consonant in order to create a syllable. The results of this study suggest that vowel epenthesis in Korean learners of English is a perceptual phenomenon. Korean learners of English are more likely to insert a vowel after a word-final consonant if they perceive the consonant as being voiceless or salient. This suggests that Korean learners of English are using their perceptual abilities to compensate for the lack of word-final consonants in Korean.

Peperkamp (2015) provides valuable insights into the role of perceived phonetic proximity in loanword adaptation and cross-linguistic speech perception. Unlike previous studies, this research focuses on on-line adaptation of auditory stimuli and directly compares it with the perception of the same stimuli. The results indicate that the way in which American English vowels are adapted into French reflects the way they are perceived, particularly when considering the consonantal context. The findings highlight the importance of controlling for coarticulation in cross-linguistic vowel comparisons. The perception of spliced vowels differs from both unspliced vowel perception and on-line adaptation, underscoring the impact of consonantal context on perception. Moreover, the study demonstrates that phonetic proximity plays a significant role in loanword adaptation, challenging the notion that adaptation is solely determined by phonological similarity. Interestingly, the research reveals that vowels considered phonologically identical in English and French can have distinct phonetic realizations, leading to unexpected adaptations. This suggests that phonetic variability due to coarticulation influences perception and emphasizes the need to account for consonantal context in cross-linguistic vowel comparisons. The author puts emphasis on the necessity that acoustic tokens should be gathered from vowels produced in utterances rather than isolated words to capture their full phonetic extension. Additionally, systemic factors should be considered in models of non-native speech perception. The study highlights the surprising finding that non-native listeners, such as Danish listeners, can outperform native listeners in discriminating certain contrasts. This suggests that broad systemic factors,

such as distinctive features in the borrowing language, may influence non-native speech perception. The research supports the hypothesis that on-line adaptation of English vowels into French is driven by phonetic proximity rather than strict phonological similarity. It sheds light on the complex relationship between perception and adaptation and calls for further exploration of perceptual and acoustic measurements, as well as the influence of systemic factors, to deepen our understanding of loanword adaptation and non-native speech perception.

The relationship between speech perception and loanword adaptation in Korean is investigated by Daland and his colleagues (2018). The authors argue that the phonological structure of the native language has a profound and lasting effect on the speech production and perception systems. Decades of research have established that monolinguals and low-proficiency bilinguals produce and perceive speech sounds in ways that are consistent with the phonotactic constraints of their native language. The authors focus on the phenomenon of vowel epenthesis in Korean loanword adaptation. Korean does not allow consonant clusters in the word-medial position. When a loanword contains a consonant cluster in the word-medial position, Korean speakers typically epenthesize a vowel between the consonants. For example, the English word *street* is adapted as *sseuteu* in Korean. The authors argue that vowel epenthesis in Korean loanword adaptation is driven by speech perception. They conducted a cross-linguistic discrimination experiment in which Korean and English speakers were asked to discriminate between Korean words that contained either a real vowel or an illusory vowel. The illusory vowel was created by adding a vowel sound to a Korean word that did not originally contain a vowel. For example, the Korean word *pak* was presented as either *pak* or *paki*. The results of the experiment showed that Korean speakers were significantly better at discriminating between real and illusory vowels in words that contained consonant clusters than in words that did not contain consonant clusters. This suggests that Korean speakers are more likely to perceive a vowel in a word that contains a consonant cluster than in a word that does not contain a consonant cluster. The authors argue that this finding can be explained by the fact that Korean speakers are less likely to encounter consonant clusters in their native language. As a result, they are less likely to be familiar with the acoustic cues that signal the presence of a consonant cluster. When they encounter a consonant cluster in a loanword, they are more likely to perceive the cluster as a vowel. The

findings of this study suggest that speech perception plays a significant role in loanword adaptation. Korean speakers are more likely to epenthesize a vowel in a loanword that contains a consonant cluster because they are less likely to perceive the cluster as a consonant cluster. This finding also has implications for our understanding of the relationship between speech perception and language acquisition. It suggests that the phonological structure of the native language can influence the way that children learn to perceive and produce speech sounds.

### 1.1.4 Combined accounts

Although there is a consensus that both phonetics and phonology may play important roles, and neither approach completely denies the significance of either aspect, all the above-mentioned research strictly lean on one side. In this section, however, some prominent intermediate positions are reviewed.

Dohlus (2005) discusses the asymmetry in the adaptation patterns of German and French mid-front rounded vowels in Japanese and explores whether the adaptation strategies are phonetically or phonologically driven. In the case of German mid-front rounded vowels /œ/ and /ø/ being adapted as Japanese /e/, it is argued to be a phonologically driven adaptation. This means that the phoneme categories of the source language (German) are maintained to the greatest extent possible in Japanese, regardless of their phonetic characteristics. On the other hand, the adaptation of French mid-front rounded vowels /œ/ and /ø/ as Japanese /u/ is considered an example of phonetic approximation. In this case, the phonological features of the sound in the source language (French) are irrelevant, and the adaptation is based on acoustic features and perceptual similarity. Japanese /u/ is perceived by Japanese listeners when hearing mid-front rounded vowels. The paper also provides further examples of phonetic approximation in Japanese loanword adaptation, such as the adaptation of word-final /n/ in English and French, and the adaptation of the English low front vowel /æ/ in different contexts. The role of orthography is discussed as well, highlighting how written forms influence loanword adaptation. The presence of written forms can provide hints about the phoneme categories of the source language and trigger phonological approximation. The author suggests that loanword adaptation is generally phonetically driven, based on

perception, and influenced by conventions, knowledge of the source language, and the role of orthography. However, there can be cases of phonological approximation, especially when direct oral input is lacking and written forms play a significant role. The asymmetry in the adaptation strategies for German and French mid-front rounded vowels in Japanese is attributed to different factors, such as the roles of German and French loans in Japanese society and the differences in spelling systems between German and French.

According to Kochetov (2008), the adaptation of Russian vowels and consonants in Japanese and Korean demonstrates a clear reference to phonemic categories. Russian vowels /i/, /e/, /a/, /o/, /u/, despite having multiple contextual realizations, are mapped onto phonemes in Japanese and Korean that share similar features. The adaptations ignore allophonic variation, focusing solely on the phonemic categories. The adaptation of Russian laryngeal and secondary articulation contrasts, as well as coda stops, is also predominantly phonological, involving a direct mapping of phonemes between the source and borrowing languages. However, the paper points out that some mappings of phonemes do not strictly follow the expected features of the interacting segments. Some adaptations involve the representation of single segments as sequences of segments, such as the representation of /ɨ/ as /ɯ/ + /i/ in Japanese and the representation of prevocalic palatalized liquids in Korean. These deviations from the expected patterns of adaptation challenge the Isomorphism Hypothesis, which predicts one-to-one mappings between segments in the source and borrowing languages. The data suggests that the phonetic properties of the source segments and their similarity to phonemic and allophonic categories in the borrowing languages play a role in loanword adaptation. Furthermore, the paper highlights cases where direct mappings occur between phonemes and allophones, as well as mappings between allophones/surface realizations and phonemes. This suggests that both phonological and phonetic information are involved in the loanword adaptation process. The study indicates that neither the strict phonological hypothesis nor the strict phonetic hypothesis can fully account for the observed patterns of loanword adaptation. The process of adaptation is influenced by native and non-native phonological categories, including both phonemic and allophonic aspects. While some attention is given to low-level phonetic information in the source language, its role is relatively limited and variable. The author supports an integrated view of loanword

adaptation that combines higher-level phonological knowledge with lower-level phonetic knowledge from both the source and borrowing languages.

An existing bidirectional model of L1 phonology and phonetics to analyze loanword adaptation in Korean is applied by Boersma & Hamann (2009). The model treats perception and production as equally important aspects of phonology and successfully explains loanword adaptation without the need for additional loanword-specific rankings, constraints, or mechanisms. The authors argue that perception should be considered as part of phonology, thus reconciling the ongoing phonology-versus-perception debate in loanword adaptation research. They propose that the behavior of listeners in their native language accounts for loanword adaptation phenomena. The model's assumptions include the differentiation between phonological and phonetic representations, bidirectional cue and faithfulness constraints, and the utilization of structural constraints in both perception and production. These assumptions have been found crucial not only for loanword adaptation but also for L1 phonology. By eliminating the need for loanword-specific phonology, the authors aim to demystify the process of loanword adaptation. They demonstrate the effectiveness of their model by successfully handling various loanword adaptation cases in Korean. They conclude by expressing their curiosity regarding the performance of their model on languages that might exhibit loanword adaptations different from those discussed in the paper.

The implications of Lin (2009) highlight the perceptual nature of loanword adaptation and the selective use of specific properties of foreign inputs for processing and adaptation. The findings suggest that some features are more salient in perception and primary in phonology, while others may be underspecified or selectively processed. This raises questions about how to predict the degree of underspecification, sources of variation in loanword adaptation, and how variation should be modeled in theoretical linguistics and psycholinguistics. The paper also has implications for feature theory, as certain features are better retained than others during the adaptation process. Manner features for consonants and backness features for vowels are shown to be more faithfully retained. This raises questions about the asymmetrical behavior of different features, the factors influencing this behavior, and how feature theory can account for these unequal relationships. The author argues that the data and processes of loanword adaptation contribute to the understanding of the interplay between phonetics and

phonology. The relative saliency and asymmetrical relationships between features observed in loanword adaptation have implications for feature theory and phonological theory as a whole.

Several case studies on loanword adaptation are reviewed by Kenstowicz (2010), and evaluated in the context of three models of loanword adaptation found in recent literature. The findings of the reviewed studies indicate that the specific phonetic details of a sound in the donor language have a crucial role in the process of loanword adaptation. These details go beyond what can be accounted for by the Phonological stance. Furthermore, the phonotactic constraints and the relationship between redundant and distinctive features are also found to be significant in shaping the loanword. Therefore, the results are inconsistent with a model that relies solely on extra-grammatical acoustic matching, such as the Perceptual stance. The Optimality Theory, with its emphasis on faithfulness constraints, is argued to provide the most comprehensive and insightful approach to loanword adaptation. This model allows for the interaction of both the phonetic details of the donor language and the phonotactic constraints of the native grammar in shaping the loanword. However, to move beyond a descriptive level and conduct a more thorough investigation of this complex phenomenon, it is essential to identify general patterns in the adaptation strategies employed in different situations.

Jaggers (2015) affirms the notion that loanword adaptation is not a simple, one-step procedure, aligning with the findings of previous research in this field. The study demonstrates that when participants created new loanwords, these words assumed an acoustic form that fell somewhere between the original form of the word and the pronunciation of established loanwords in the borrowing language. This suggests that loanword adaptation involves a gradual and nuanced transformation. Interestingly, the study found that the participants' level of experience with L2 did not significantly impact their production of loanwords. Both groups, despite being in the early stages of learning Japanese, deviated from producing the expected loanword form. This indicates that factors other than L2 experience may influence loanword production. Furthermore, the produced loanword form differed from the typical phonotactics of American English, even though it contained the necessary phonetic features. This implies a divergence from the grammar of native English words,

showcasing a distinction in loanword adaptation. The study also explored two approaches to loanword adaptation: a constraint-shifting approach and loan-specific faithfulness constraints. By examining the surrounding speech, the researchers observed that a native English function word displayed a more faithful pronunciation when preceding nonce loans compared to native English words. This finding supports the idea of upward shifting of faithfulness constraints due to the activation of the preservational style. To enhance the authenticity of loanword dissemination studies, the researchers emphasize the importance of replicating real-world scenarios. They argue that disseminators likely possess some degree of familiarity with the source language, making it crucial to consider this aspect in experimental methods. Additionally, the study advocates for conducting loanword dissemination research within a communicative task in the borrowing language. This approach better reflects the actual usage of loanwords, minimizes concerns about pronunciation accuracy, and enables analysis of surrounding native words, thereby providing a more comprehensive understanding of loanword adaptation. While the proposed theoretical account raises further questions, such as the role of constraint sensitivity and cross-linguistic loanword trends, the study recognizes the need for more extensive data analysis. Moreover, it underscores the significance of sociological aspects in loanword adaptation studies, including factors like proficiency, confidence, attitude toward the source language, and engagement outside the classroom. Methodologically, the paper suggests increasing the token count in future studies and prioritizing communication over pronunciation. It recommends considering the representativeness of loanword dissemination when eliciting responses from participants who may not naturally serve as disseminators. The author provides valuable insights into the gradual nature of loanword integration, the influence of L2 experience, the divergence from native phonotactics, and the impact of surrounding speech. The paper also emphasizes the need for further research and consideration of sociological factors, while proposing methodological improvements to enhance the study of loanword adaptation.

## 1.2 Objectives and significance of the research

Within their etymological journeys, borrowings quite often go through phonetic assimilations. The longer the time they exist in the target language, the more significant and defining the adaptations apparently are. WOT (Oghur) material in Hungarian is no exception. Given the fact that almost all the known WOT words that appear in Hungarian today are pronounced differently compared to their source forms, the question of "At what rate did that happen?" concerns this study primarily. At this point, it is essential to emphasize upfront that when loanwords integrate into the target language, they will inevitably experience the same phonological transformations as any other native words, regardless of their origin (Wieland, 2012, p. 364). This highlights that phonological changes in a language typically impact all words, including loanwords, and are not solely driven by an attempt to assimilate loanwords. This perspective is crucial to dispel any misconception that all loanwords naturally align over time with the phonological characteristics of the receiving language; at times, they may actually undergo changes in the opposite direction.

In order to quantify the rate of the phonetic changes that have occurred, a refined Levenshtein distance (LD) formula that is weighted by distinctive features is used. Other than the issue of how much the WOT words are adapted to the Hungarian language, the way they did is another ax to grind. Was it completely by perception or were there certain patterns of rule-based sound changes? The intensive Hungarian-Turkic cohabitation coincides with the period between the 6$^{th}$ and the 10$^{th}$ centuries when the late ancient Hungarian was spoken (Róna-Tas & Berta, 2011, p. 1072). This time period concurs with the post-Hunnic era in the Eurasian steppes. In the history of the Hungarian language, this period represents the late Proto-Hungarian and early Old Hungarian periods (Kiss, 2017). "During this period, the Hungarian language existed as the spoken language of tribal dialects of the changing settlement areas" (Kiss, 2018a, p. 45; translated by Andrea Parapatics, Szilárd Szentgyörgyi). As the WOT material in Hungarian is one of the most valuable sources in terms of the linguistic properties of WOT speakers, adaptational features of these words can be useful for reconstruction studies and for a better comprehension of the linguistic properties of those languages that existed in that era – including the Hunnic – and extinct with little

trace. Today the classification of some of these languages is still problematic and the acknowledgments of their phonology are limited.

"Often two languages or families show striking lexical similarities that unambiguously prove a historical relationship, but whether these lexical similarities are due to common inheritance or to borrowing is a matter of dispute. In such disputes, more systematic knowledge of the general patterns of loanword distribution will hopefully be helpful in the future" (Haspelmath, 2009, p. 36).

The objectives of the study, therefore, are as follows:

- To quantify the adaptation rates of the WOT loanwords in Hungarian.
- To detect the patterns and regularities in the nativization of WOT words in Hungarian.
- To adopt measurement methods of phonetic distance into loan phonology.
- To attempt to refine the edit distance algorithm into a more sensitive and coherent form.

The study can also contribute to the research area of the Old Hungarian language. Jenő Kiss states in his review of Honti's work (2017):

"Among the derived words that came into Hungarian before the *honfoglalás*, the overwhelming majority are those of Old Turkic origin, compared to the rest. This means that Old Turkic words play a key role in Proto-Hungarian language research" (Kiss, 2018b, p. 226; translated by Andrea Parapatics, Szilárd Szentgyörgyi).

The overall purpose of this research is as follows. It aims to develop a coherent methodology by utilizing a refined Levenshtein Distance formula weighted by distinctive features for sensitizing the analysis of loanword adaptations. For this task, firstly the initial phonetic changes that occurred during the borrowing (copying) process of Turkic words into the Hungarian language are evaluated. At this point, the comparisons are between the WOT sources and the reconstructed initial old Hungarian copies. In this part, the research aims to provide evidence for the minimality of the initial copying alterations, emphasizing that these changes were minimal and driven by phonetic necessity. Secondly, the focus is on the

adaptation rates of the modern Hungarian versions and the detection of patterns and regularities in the nativization of WOT loanwords in Hungarian.

The main aims of Turkic-Hungarian linguistic studies have always been etymological. Regularities, patterns, and a typological corroboration of loanword phonology in the concept of 'loanword adaptation' are still puzzles to work on. This study intends to contribute to the settlement of this issue.

In short, this study seeks to quantify adaptation rates and detect the rule-based systematic operations in the adaptations. By achieving these objectives, this research aims to contribute to the understanding of loanword phonology and the broader field of linguistic analysis.

## 1.3 Research questions

The following problems constitute the main objectives of the study:

- **RQ1:** What are the regular (systematic) aspects regarding phonetic adaptations of Turkic loanwords in Hungarian? Are there regularities in the first place? If the changes occur on a regular basis what are the systematic patterns behind the changes? This question addresses the adaptation patterns.
- **RQ2:** What statistical analysis can be given of loanword adaptations in Turkic-Hungarian words in terms of frequency and regularity? What is the average rate of assimilation for all the Turkic material in Hungarian? This question addresses the quantification of frequency.
- **RQ3:** What are some advantages and troublesome issues about weighted operations regarding phonetic distance calculation? How does a phonetic distance calculation based on distinctive features provide a more coherent approach compared to evaluating operation costs based on more general properties?
- **RQ4:** How adaptive are initial copies compared to long-term nativization of borrowings in the target language?

## 1.4 Scope and limitations

Weighted operations allow researchers to assign different costs to specific phonetic or phonological changes, such as substitutions, insertions, or deletions. This enables the measurement of the similarity between words while considering the specific phonetic properties of the languages involved. Loanword adaptation studies often involve comparing loanwords from one language to another. Weighted operations allow researchers to model the adaptation process by assigning higher costs to certain operations that are more or less likely to occur based on language-specific phonological constraints. Researchers can adjust the weights of different operations according to the specific linguistic characteristics they want to capture. This flexibility enables the fine-tuning of the distance metric to better reflect the patterns observed in loanword adaptation.

The limitations of the Levenshtein distance become evident when examining the fine-grained pronunciation differences between words. If a standard penalty point were to be applied for every substitution, the Levenshtein distance between /kuːl/ and /kʊl/ would be the same as the distance between /kuːl/ and /kæl/. However, in people's perception, the pronunciation of /kuːl/ is more similar to /kʊl/ than to /kæl/. Intuitively, substituting /uː/ with /ʊ/ should incur a smaller cost than substituting it with /æ/. Therefore, it becomes necessary to modify the operation cost in Levenshtein algorithms to account for these subtle distinctions in pronunciation. The refined algorithms employed in this project adjust the costs of specific operations, allowing for a more sensitive measurement of pronunciation distance and capturing the fine details of sound differences between words.

Weighted operations in the Levenshtein distance provide a valuable tool for loanword adaptation studies by allowing researchers to account for phonetic and phonological differences. However, the subjective nature of weight assignment, the lack of consensus, and the potential for overlooking subtle linguistic patterns should be considered when utilizing this approach. Assigning appropriate weights to different operations can be challenging and subjective. The weights should ideally reflect the linguistic properties being studied, but determining the precise values can be a matter of interpretation. Different weight assignments may lead to different results and potentially affect the validity of the analysis. There is no universally agreed-upon set of weights for loanword adaptation studies. The choice of

weights may vary depending on the specific research goals, languages involved, and linguistic theories employed. This lack of consensus can limit the comparability and generalizability of findings across studies. In some cases, heavily weighting specific operations may lead to the overlooking of subtle linguistic patterns or phonetic transformations that are crucial for loanword adaptation analysis. Researchers need to strike a balance between capturing important phonological differences and preserving the integrity of the data.

Evaluating the distinctive features of diphthongs or dealing with the deletion of transitional sounds like glides (also known as semivowels) can present challenges in the calculation of operation costs. There are efforts in this study to standardize these special calculations.

The western branch of Turkic languages has left a limited written record compared to Common Turkic languages. Consequently, the study of West Old Turkic requires relying on indirect sources, such as loanwords found in other languages or historical references. In the specific context of researching phonetic adaptations of West Old Turkic loanwords in Hungarian, the lack of sources on the source language may hinder the ability to compare and analyze phonetic changes that occurred during the borrowing process. To overcome this limitation, researchers often employ comparative linguistic methods, drawing on available sources and evidence from related languages to make informed hypotheses. Thus, the data that is available today are the reconstructed WOT words.

The Levenshtein distance formula offers several advantages for linguistic distance calculation:

- Flexibility

The formula can be applied to different linguistic levels, such as phonemes, morphemes, or whole words, allowing researchers to tailor its use to specific research questions. For example, in a study comparing worldwide phonemic and genetic variation in human populations, a significant association was found between phonemic distance and genetic distance, as well as between phonemic distance and geographic distance (Creanza et al., 2015). The Levenshtein distance method has also been applied successfully to evaluate

dialect distances such as Irish Gaelic and Dutch dialects (Gooskens & Heeringa, 2004, p. 197). Furthermore, the Levenshtein distance formula has been used to measure the orthographic and phonetic distances between dialects and language varieties (Mosbach et al., 2019, p. 811). It has been shown to be a versatile tool in quantifying linguistic differences and similarities.

- Language Independence

Unlike some distance metrics that rely on predefined language-specific features, the Levenshtein distance formula is applicable to any language or dialect, making it versatile for cross-linguistic comparisons (Dijkstra et al., 2018, p. 665).

- Non-parametric Nature

The algorithm does not impose assumptions about the structure of languages or the types of linguistic changes. It measures the actual number of edits required, providing a more robust and objective measure of linguistic distance (Downey et al., 2017, p. 138).

Despite its widespread use, the Levenshtein distance formula has certain limitations as mentioned below:

- Lack of Linguistic Context

The formula focuses solely on the minimum edit operations without considering linguistic patterns or rules. It does not capture the semantic or syntactic aspects of languages, limiting its ability to measure deep linguistic divergence accurately. This 'linguistically naive' (Greenhill, 2011, p. 695) characteristic may be considered a limitation in some contexts.

- Unequal Weighting

The formula treats all edit operations equally, assuming that insertions, deletions, and substitutions have the same impact on the overall linguistic distance. In reality, some operations may have different linguistic implications, and assigning equal weight to all edits can oversimplify the complexity of language variation. Researchers have explored the use of weighted or unweighted edit distance measures, where different edit operations are assigned different weights based on their linguistic implications (Cotterell et al., 2014).

- Sensitivity to Word Length

The Levenshtein distance formula tends to be more sensitive to longer words or strings, as they have a higher likelihood of containing more edit operations. This can introduce a bias in linguistic distance calculations, particularly when comparing languages with different average word lengths. To address this issue, researchers have proposed various approaches. One such approach is the use of a normalized edit distance, which takes into account the lengths of the strings being compared. By incorporating the 'Generalized Levenshtein Distance' and the lengths of the strings, this normalized edit distance provides a more balanced measure of similarity (Li & Liu, 2007; Setiabudi et al. 2021, p. 1235).

- Lack of Phonetic Consideration

While the Levenshtein distance formula captures some phonetic variation, it does not account for phonetic similarity or dissimilarity beyond single-character edits. This can be problematic when comparing languages with significantly different phonetic inventories or phonological structures. Additionally, the Levenshtein distance is based on phonetic transcriptions of isolated words, which means that intonation and tonemes are not taken into consideration (Gooskens, 2005). Therefore, when comparing languages with different phonetic characteristics, alternative approaches that go beyond the Levenshtein distance formula may be necessary to capture a more accurate measure of phonetic similarity or dissimilarity.

- Historical Considerations

Loanword adaptation is a dynamic process that can unfold over extended periods of time. The Levenshtein distance formula, as a snapshot measurement, may not account for historical or diachronic factors that influence the adaptation process. It is crucial to incorporate historical data and examine loanword adaptation in a broader temporal context. A diachronic study of loanwords in Quebec French found that the rate of non-phonological adaptations decreased over time (Paradis & LaCharité, 2008). This highlights the importance of considering historical factors in understanding loanword adaptation.

To overcome the limitations of the Levenshtein distance formula, researchers have developed various extensions and improvements:

- Phonetically Weighted Levenshtein Distance

By incorporating phonetic information or phonetic distance measures into the algorithm, the resulting metric can better account for phonetic similarities or dissimilarities between languages or dialects (Fontan et al., 2016; McCoy & Frank, 2018).

- Contextualized Linguistic Distance Metrics

Integrating linguistic context, such as syntactic or semantic information, into distance calculations can provide a more comprehensive measure of linguistic dissimilarity. For example, Herdağdelen and colleagues (2010) propose a novel approach to query reformulation that combines syntactic and semantic information using generalized Levenshtein distance algorithms. They suggest that incorporating probabilistic term rewrite functions as substitution costs improves the accuracy of the reformulated queries. This approach may require sophisticated language models or the use of machine learning techniques.

- Hybrid Approaches

Combining multiple distance metrics, such as Levenshtein distance with other algorithms like cosine similarity or string alignment methods, can yield more accurate linguistic distance measurements. These hybrid approaches may take advantage of the strengths of different methods and mitigate their weaknesses (Mosbach et al., 2019).

## 1.5 The exposition of the Turkic material in Hungarian

The world of science as we know it today is dominated by Indo-European fundamentals, particularly English being the de facto global language of science (Drubin & Kellogg, 2012, p. 1). This has been the same for a while now, considering the fact that the Age of Enlightenment, the Printing press, the Renaissance, and what is called Modern science were all born in European societies. Going further past, it is clear that the majority of written

heritages that shape our understanding of the past today and that have become a headstone for humankind to build on – in terms of history, science, and whatnot – are products of the Greco-Roman world, hence in Latin or Greek, Indo-European. Most of the scientific terminology was indeed derived from either Latin or Greek (Steffanides, 1965, pp. 785-789). For the science of Linguistics as a rather recent field, the case is not very different. This, together with a number of other factors, has led to a relative neglect of the analysis of non-Indo-European languages. However, there is another major reason, more significant than the former and that is the lack of data. As mentioned above, the majority of the written sources are from Indo-European languages, and as languages do not fossilize, it is extremely difficult to analyze their roots without written data. This has been a problem for scholars studying the ties between the Hungarian language and Turkic languages, especially since the Turkic languages – including WOT – that are subject to comparison particularly for this case are extinct with little trace (Novgorodov et al., 2015, p. 118).

Loanwords of Turkic origin in the Hungarian language are rather numerous due to extensive contact that took place in several different stages of history between the Hungarians and different Turkic-speaking groups. These age-long and complicated interactions occurred in different parts of Eurasia all the way from Magna Hungaria – as it is referred to in medieval sources – to the Pontic steppe and the Carpathian Basin. The association continued after the medieval era. Loanwords being numerous, and prototypical features of Turkic languages relatively similar to Hungarian, some scholars associated the parties with each other, especially in the late 19$^{th}$ and early 20$^{th}$ centuries. The Ural-Altaic hypothesis which was proposed by Matthias Castrén (Stammler-Gossmann, 2009, p. 202) groups Turkic, Mongolic, and Tungusic as *Altaic,* whilst Finno-Ugric and Samoyedic as *Uralic*. German philologist Friedrich Max Müller labels in his work (Müller, 1855, p. 86) non-Aryan and non-Semitic Asian languages as *Turanian* with the exception of Chinese.

Today the Altaic hypothesis is not widely credited anymore (Vovin, 2005, p. 73) as Turkic, Mongolian, and Tungusic are not seen as direct descendants of a common family. Uralic and Turkic languages on the other hand are seen as two groups of neighbors that belong to a Sprachbund (Helimski, 2003, p. 159) rather than being genetically related. Hungarian Turkologist Ármin Vámbéry tried to prove in the late 19$^{th}$ century in his work

Hungarian and Turco-Tatar word cognates (Vámbéry, 1870) by making comparisons between words, that Hungarian is a dual language with both Uralic and Turkic features and that's the result of intense early Hungarian encounters with Turkic groups (Marácz, 2012). This view started a two-decade period of academic debate called *The Ugric-Turkic (linguistic) war*. The head of the Ugric defender of the debate was József Budenz. He retaliated to Vámbéry one year after in 1871 (Budenz, 1871), arguing that all the Turkic words in Hungarian are loanwords and not genetically inherited. He indicated that most of the words that had been discussed by Vámbéry (some 740 words) were loanwords also in Turkic and were not of Turkic origin. He reduced the number of the words to 146 (Róna-Tas & Berta, 2002, p. 46). Evidently, the correspondences that Budenz acknowledged were partly correct Turkic loanwords but there were many mistakes as well. The problem is that he did not present the phonological criteria that he used for his comparative method:

"In his original works, the comparative method was either not applied correctly or was not applied at all, for two basic reasons. Firstly, the comparative method was in its infancy at the time, which means that it could not have been applied systematically and scientifically in any case" (Marcantonio et al., 2001, p. 82).

The debate also gained a political base since the Ottoman Empire of the time was neighboring the traditional rival Austro-Hungarian Empire, and was supporting the independence movements in Hungary already for more than a century.

"Vámbéry was well aware of the fact that the theory initiated by the German school of Schlözer to demonstrate the Finnish ancestry of the Hungarians was not an objective, scientific theory but was motivated at large by ideological and geopolitical reasons" (Marácz, 2012, p. 18).

The debate ended with the Ugric party getting more credit, and to this day, Hungarian is acknowledged as a Finno-Ugric language and the numerous Turkic words are all loanwords mostly borrowed from Oghur. As of today, in light of the new information in modern linguistics, the Hungarian language is accepted widely as a member of the Finno-Ugric branch of the Uralic family, and Turkic influence is acknowledged as a result of early – and latter – language contacts.

The actors on the Turkic side of these historical relationships were multifarious from different branches of the Turkic language family. Pechenegs and Ottomans from the Oghuz branch, Cumans from the Kipchak branch, Bulgar-Turkic, and Khazar (its classification is disputed) (Róna-Tas, 2007, p. 270) from the Oghur branch are some prominent ones. The group that left the most predominant impact on the Hungarian vocabulary was a member of the West Old Turkic (WOT) branch, the Oghur-speaking people. This contact happened long before Hungarians migrated into the Carpathian basin (honfoglalás). The word *Ungar* (Hungarian) itself is considered to be derived from the name *Onoğur,* meaning ten tribes or ten arrows in Turkic (Benkő & Imre, 1972, p. 31).

> "The Hunnic migrations brought Oghur Turkic nomads westwards into contact with the ancestors of the Hungarians. This contact, the details and precise chronology of which are unknown to us must have been culturally decisive for the ancestors of the Hungarians were transformed into an equestrian, pastoral-nomadic steppe people and came fully into the steppe in what is today Baskiria, the «Magna Hungaria» of the medieval sources" (Golden, 1992, p. 261).

As the Turkic language family fundamentally diverges into two main branches (Johanson, 2016) – common Turkic and Oghur – all the Turkic languages which are grouped under the common Turkic branch, differ substantially from Oghur. The only extant member of the Oghur branch today is *Chuvash* (Savelyev, 2020, p. 446), spoken primarily in the Chuvash Republic by some one million speakers (Alós i Font, 2014, p. 53). The majority of the Turkic words in Hungarian show featural analogy with Chuvash (Róna-Tas & Berta, 2002, p. 47). The distinction between Chuvash and the rest of the Turkic languages is so sharp that some ancient Turkic morphological features which have been preserved by Chuvash have been lost in all other Turkic languages (Róna-Tas, 1999). Nevertheless, Turkic languages exhibit fundamental structural and phonetic commonalities.

> "Throughout their history and in spite of their huge area of distribution, Turkic languages share essential structural features. Many of them are common to Eurasian languages of the Altaic and Uralic types. While often dealt within typologically oriented linguistic work, most aspects of the Turkic structure still call for more unbiased and differentiated description" (Johanson, 1998, p. 30).

Although there is an almost total consensus today that Turkic languages and Hungarian belong to different language families, morpho-syntactic and phonological resemblances between them such as agglutinative structures, vowel harmony, lack of grammatical gender, pro-drop features, and rich vowel inventories clearly prove the existence of a close link, setting aside the excessive number of Turkic loanwords in Hungarian. Some of the other parallel aspects are consonant assimilation or consonant harmony and affricate sounds and consonant deletions when they are accumulated (Bekar, 2013, p. 80). This has led many scholars throughout centuries to investigate the ties between Hungarian and Turkic languages to this day. Even as early as the years 1200-1230, in the medieval history book *Gesta Hungarorum – The Deeds of the Hungarians* – which was written by an unidentified author, some Turkic words such as *atil/itil* ('river'), *oba* (thought to be representing the Turkic word for 'father'; *apa/ata/opa*), *keanus* (Turko-Mongolic title *khan*) can be observed (GH, p. 65). "The comparison of Hungarian words to similar Turkic ones is as old as the Hungarian records, chronicles, and diplomas" (Róna-Tas & Berta, 2002, p. 42).

Oghur is also called *r-Turkic* since many words that end with /z/ in common Turkic languages, end with /r/ in Oghur. Therefore the term *oğur* (oghur) appears as an Oghuric reflex of the term *oğuz* (Golden, 2011, p. 30). Oğuz sub-branch comprises Turkic languages such as Turkish and Azerbaijani. The term reflects kinship as the other words in Turkic such as *oğul, oğlak, oğlan*, etc. It is thought that the term, *oğur/oğuz* may have served as a tribal union name (Golden, 1992). Among all the Turkic loanwords in Hungarian, a significant number is evidently from Chuvash, or at least they are lexically parallel to Chuvash alterations rather than the counterparts in other Turkic languages as rhotacism can clearly be observed in the Hungarian versions as shown in the diagram below.

**Table 1. Rhotacism in Chuvash and Hungarian** (Marácz, 2016)

| Meaning | Hungarian | Chuvash | Comm. Turkic | Mongolian |
|---|---|---|---|---|
| 'calf' | *borjú* | *păru* | *buzāγu* | *biraγu* |
| 'twin' | *iker* | *yĕker* | *ekiz* | *ikire* |
| 'ox' | *ökör* | *văkăr* | *öküz* | *üker* |
| 'mud' | *sár* | *šur* | *sāz* | *siroi* |
| 'to filter' | *szűr* | *sĕr-, sör-* | *süz 'to clean'* | - |
| 'ring' | *gyűrű* | *śĕrĕ* | *yüzük* | - |
| 'knee' | *térd* | *čĕr* | *tiz / tizik / tīz* | - |

The vocabulary of Turkic loanwords in Hungarian consists of mostly agricultural words. It's a problematic issue to track data for nomadic history, culture, or languages since traditionally the people that belonged to nomadic groups were used to transmigrate constantly as a way of life and did not have a settled life like for instance, Indo-European farmer societies had. This prevented them from being occupied with literal traditions or social sciences considerably. The science world did not even have full insight into the old Turkic Runic scripts – from which Hungarian Runic scripts were derived (Lisztóczki, et al. 2019, p. 422) – until Russian Turkologist Nikolai Yadrintsev discovered in 1889 the Orkhon inscriptions; two memorial installations erected by the Göktürks, written in Old Turkic alphabet in the early 8th century in the Orkhon Valley in Mongolia. The lack of historiography left many blank points for researchers today to investigate nomadic societies and their languages. That is the reason why today linguists are able to construct the Proto-Indo-European language to some extent but unable to link some of the Eurasian nomadic languages even among each other completely yet.

About the abovementioned Turkic and Hungarian Runic scripts; although it is widely accepted that the Hungarian Runic script (also called székely runes, *rovásírás* in Hungarian) is derived from Turkic scripts. There are also views suggesting that it is unlikely that the former originated from the latter. Varga (1999, p. 71) suggests that the sound representation system in the old Turkic script is radically different from its Hungarian counterpart. Turkologist and sociolinguist Klára Sándor of Szeged University who has prominent works on the Székely script, indicates that eastern Turkic runes (referring to Orkhon scripts) and the Székely scripts are related but indirectly (Sándor, 2014, p. 353).

Cultural relations of the Turkic and Hungarian people date back – as far as we know – to the time they lived together in the Ural region. Hungarians, who settled in the southwest of the Ural region, maintained a neighborhood with the Bulgarian Turks – Oğuric speakers – for centuries and finally followed them during the Bulgarian Turks' migration to the west. In the 5th century, Hungarians advanced to the Kuban River in the north of the Caucasus. Up to the year 800, they were found in the areas between the Don and Kuban rivers. Hungarians, who came together with the people under the rule of the Khazar Emperor, lived a nomadic

life in the border regions. Here they learned agriculture, fruit growing, and animal husbandry techniques from their neighbors (Düzgün, 2014, p. 2).

Language contact between Hungarians and Turkic-speaking people did not come to an end after the Hungarian conquest of the Carpathian Basin. The Turkic loanwords that were borrowed after Hungarians settled *(honfoglalás)* in where they are today in the year 895 are regarded as *middle-layer* words. Some other nomadic groups also moved into the Carpathian basin because of eastern threats gradually. Cumans and Pechenegs were two of them and they spoke Kipchak and Oğuz Turkic respectively. The Byzantines placed the first arising of the Cumans or Kipchaks in the years 894–899 when they expelled Pechenegs from the region between Ural and Volga. When the early struggles of the Cumans with the Pechenegs are examined, they are seen as allies with the Khazars. First, the Khazars and the Cumans drove the Pechenegs beyond the Don River. However, some of the second group continued to live in the deserts between Ural and Volga; the remainder gradually piled west into Hungary and the territory of the vulnerable Byzantine Empire (Howorth, 2020, p. 293). Cumans followed in the early 13$^{th}$ century, they settled in Hungary, running from Mongols. In the voluminous work of András Róna-Tas and Árpád Berta (Róna-Tas & Berta, 2011), the most comprehensive etymological dictionary of Turkic loanwords in Hungarian, 35 loanwords from Cuman are mentioned. Although Cumans/Kumans did not have a written language, it is possible to examine their language and legacy through some words. Some surnames and place names appear as examples among those words such as *agbúra* [ak buğra] 'white camel', *alacs* [alaç] 'speckled', *karakas* [kara kaş] 'black eyebrowed', *karcag* [karsak] 'steppe fox' (Yılmaz, 2015, p. 291).

Interestingly, Pecheneg words are not listed in the book although there are some words in Hungarian that are claimed to be from Pecheneg. As for the relations of the Pechenegs with Hungary, although the Pecheneg raids caused difficulties for St. Stephen and his successors, it is clear that there were peaceful relations as well. The Pechenegs, who migrated not only as prisoners of war but also voluntarily, were settled in Hungary, especially in the Tiszaroff region in Pest county, in Sárrét (in Fejér county) and in the Fertő region (Rásonyi, 1964, p. 65). "No written testimony by the Pechenegs themselves has been preserved" (Kincses-Nagy, 2013, p. 172).

Although features of the early Turkic languages are typically not well-established due to the lack of data, the Cuman language is well-documented and studied. The medieval Latin dictionary, *Codex Cumanicus* appears as a comprehensive linguistic manual that was written to ease communication of Catholic missionaries with the Cumans. The manual is well examined by some scholars including Ligeti (Ligeti, 1981). Cumans migrated to Hungary in the 13th century, primarily to Kunság, a central region in Hungary associated with a former political entity created by or for the Cumans. They were assimilated into Hungarian culture and the language was extinct by the 18th century. Tradition holds that the last speaker of the Cuman language died in 1770 in Kunság was István Varró (Salaville, 1914).

"The Cuman Lord's Prayer (Pater Noster) was written down in 1744 by Ádám Kollár court counselor from a man called István Varró, who went to Vienna as an emissary in the delegation of the Cuman Provinces. The seriously corrupted form of the prayer shows that neither the informant [nor the scriptor] knew the language anymore" (Kincses-Nagy, 2013, p. 177).

**Map 1. Cuman migration in Hungary** (Pálóczi-Horváth, 1989, p. 56).


The above-mentioned text is given below with the reconstruction by (Mándoky Kongur, 2012) and the English translation.

Bezén attamaz kenze kikte *(Biziŋ atamïz kim-siŋ kökte)*

**Our father, who are in heaven,**

szénlészen, szenádon *(sentlensin seniŋ adïŋ)*

**hallowed be your name,**

dösön szenküklön *(düššün seniŋ köŋlüŋ)*

**let your will come down,**

nitziégen, gérde, ali kékte *(nečik kim ǰerde alay kökte)*

**on earth as in heaven**

bezén akó mozne, okné mezne, bergézge *(biziŋ ekmegimizni ber bizge)*

**give our bread to us each-and-every day,**

il bézen méne mesne *(ilt biziŋ minimizni)*

**forgive us our trespasses**

neszembezde, jermezbezge, utrogergenge *(nečik kim biz de iyermiz)*

**as we also forgive it to those who come against us**

iltme bezne, olgya nanga *(iltme bizni ol ǰamanga)*

**do not lead us to the evil,**

kútkor bezne, algya manna *(qutqar bizni ol ǰamannan)*

**deliver us from the evil,**

szen borszony *(sen barsïŋ)*

**you exist**

boka csalli, bótson igyi tengere. *(bu kučli bu čïn iygi teŋri)*

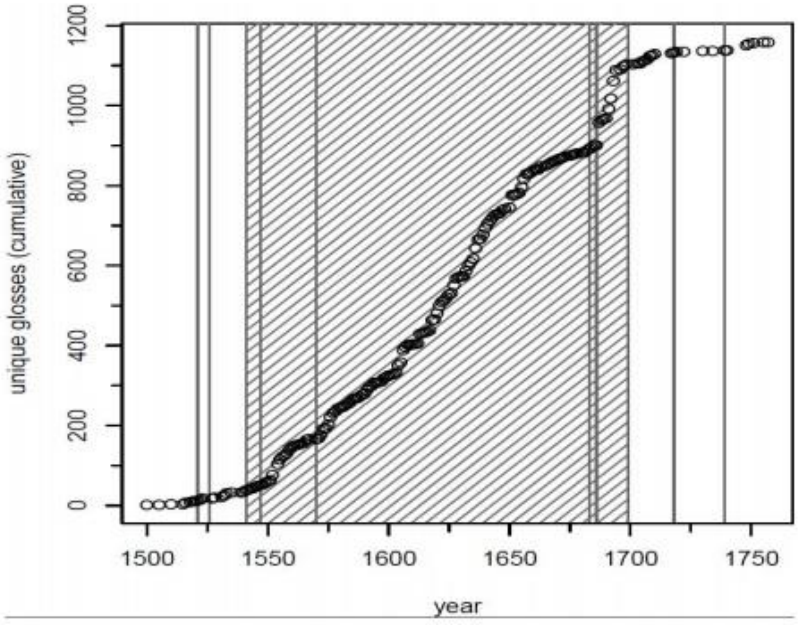
**you mighty, you righteous good lord.**

Ammen. *(Amen)*

**Amen**

Kamil Stachowski in his paper (Stachowski, 2013, pp. 100-116) examines the post-medieval texts from Hungarian that coincide with the Ottoman Era and makes remarks, indicating the correlation of them with historical events. The graph below from his paper presents a cumulative sum of unique Ottoman glosses in Hungarian texts in the years 1500–1757. Evidently, the use of Turkish material started upon the Ottoman influx and increased gradually in Hungarian.

**Figure 1. Cumulative sum of unique Ottoman glosses in Hungarian texts between 1500–1757** (Stachowski, 2013, p. 105)




"The famous Hungarian poet Bálint Balassi (1554-1594) wrote poems both in Turkish and Hungarian. Hundreds of Turkish words were used by Hungarians in their daily speech or in their written documents. The comparison of Hungarian and Turkic words was nothing special, Ottoman words were written in Latin script but with Hungarian orthography […] The Hungarian Scholar Ádám Ferenc Kollár (1718-1783) editing in 1763 the work Hungarian et Atila of Nicolaus Oláh (1493-1568), remarked in one of his notes that though

the Hungarian language seems to be related to the oriental languages and among them to Turkic, their structure is more similar than their lexical stock. Kollár had a good knowledge of Ottoman Turkish" (Róna-Tas & Berta, 2002, p. 44).

The influx affected the Hungarian language not only in terms of vocabulary but phonologically as well. The affricate [ǰ] was added among the consonants because of Turkish loanwords such as *findzsa* (< *fincan* 'teacup'), *handzsár* (< *hançer* 'dagger'), *dzsida* (< *cıda* 'lance') (Kiefer, 2010, p. 723).

The lack of data about nomadic languages prevents collecting material for reconstructions. While there is even no clear consensus on the origin of the language of Western Huns – although they had a big impact on the Western world – it is difficult to expect written sources about less significant Turkic tribes like Kumans and Pechenegs, or extinct Oghuric relatives of Chuvash. Leaving the hypothetical Uralic-Altaic macro family aside, even gathering Turkic and other central Asian languages like Mongolian and Tungusic under the roof of Altaic is problematic for that matter.

Turkic loanwords in Hungarian are indeed rather numerous, however, loanwords are not indicators of a genetic connection. The number of French loanwords in English is more than the size of its Germanic vocabulary although French is a Romance language and English is a Germanic one. That is to say, loanwords are products of a Sprachbund, and borrowing is not a selective process. Language is a living organism and depending on culture, geography, lifestyle, and many other reasons, each has its own unique vocabulary. Hungarians, after they adopted the nomadic lifestyle, borrowed many words about animal husbandry and agriculture from Turkic as a result of necessities. Other Finno-Ugric languages have some Turkic loanwords too. Rather than loanwords, one of the best ways to compare the genetic relatedness of two languages is to analyze certain areas of vocabulary that are unlikely to be replaced by loanwords in any language, such as numbers, names of body parts, or kinship terms. That is what researchers use as data when they reconstruct proto-languages. Correspondences can clearly be observed among Finno-Ugric languages in terms of those categories that are resistant to change as shown below with the number *three* in some Finno-Ugric languages.

**Table 2. Number *three* in Finno-Ugric languages** (Laki, 1960, p. 4)

| | |
|---|---|
| Hungarian | *három – harma-* |
| Finnish | *kolme* |
| Lappish | *galbma – golma* |
| Cheremis | *kom – kum* |
| Votyak | *kïjn* |
| Zyrien | *kujim* |
| Vogul | *qorem – qurem* |
| Ostyak | *kolom* |
| Mordvin | *kolmo* |

The similarities between Turkic and Hungarian are indeed striking, many shared parameter values – in the sense of Chomskyan Universal Grammar (Chomsky, 1981) – can be studied. The first-person possessive suffix is one example: Hu. *anyá-m* 'my mother' vs. Tr. *anne-m* 'my mother'. However agglutinative features, in other words, suffixal morphologies are not unique to Uralic and Altaic languages. Indigenous American languages, Austronesian languages, Niger-Congo languages, Berber languages, Dravidian languages, Eskimo-Aleut languages, Japanese and Korean (sometimes included in Altaic), Caucasian languages, and Basque are also agglutinative. Nevertheless, there is of course no reason for some of those languages not to be descendants of an older macro proto family although it is not possible with the methods of current historical-comparative linguistics to establish such a hypothesis. There are hypotheses for instance about American natives originally being Siberians and having moved to America in the ice age by walking the frozen Bering Strait (McGhee, 1989, p. 14).

In brief, we see 3 stages of (asymmetric) contact between Hungarian and Turkic languages. The first and the most important of these is the relationship that occurred in the 5$^{th}$ and the 6$^{th}$ centuries between Hungarian and a Chuvash-like language or an extinct *Proto-Bulgar – West Old Turkic* – language which was very close to Chuvash. This subject is the most important field of research in Hungarian Turcology. The second stage is the result of the migrations after the Mongol Invasion and the third one is the Ottoman influx.

After the discovery of *Orkhon Scripts* and *Uigur Manuscripts*, data on East Old Turkic has become extensive while for West Old Turkic there is still very limited material available. None of the WOT languages that were spoken in Eastern Europe were written and barely some Runic relics were left of them. In that regard, the only surviving extant member; the Chuvash language, and related material in Hungarian hold major importance for the literature in that field. These are not only important for WOT studies but also for the diachronic studies of the Hungarian language since there are no other first-hand sources to shed light on that period of the Hungarian language (Yılmaz, 2015, p. 297). The etymological dictionary (Róna-Tas & Berta, 2011) is the latest major check-over of the field where they eminently display almost the entire etymology of all the given loanwords with their phonetic backgrounds.

Ligeti (1986) mentions that there is little room for doubt that the Hungarian language once bore a more significant influence from Turkic elements than it does today.

"However, it's not merely the quantity of Turkic linguistic contributions that holds significance. At some uncertain juncture in the not-so-distant past, there appears to have been a profound transformation in the social and economic fabric of the Hungarian people, as evidenced by the Turkic vocabulary, prompting the originally Finnish-speaking Hungarian people to transition from a sedentary lifestyle to a nomadic one, characterized by horseback riding. It's crucial to stress that this shift goes beyond just the incorporation of Turkish words into the language; it represents a fundamental societal change" (Ligeti, 1986, p. 300).

## 1.6 Data

384 WOT etymologies are given in the voluminous etymological dictionary of (Róna-Tas & Berta, 2011). 377 of these are taken into account in this study. The seven entries that are excluded are *bicsak* 'pocket knife', *bicske*, 'pocket knife', *bocsájt* 'to let, to forgive', *koldus* 'beggar', *gyarat* 'to produce', *sár* 'yellow', and *seper* 'to sweep, to broom' as they are either variants or dialectal corresponding forms, but sometimes the result of a word or semantic split (Róna-Tas & Berta, 2011, p. 1143). The headwords that are taken into account instead are *bicska* 'pocket knife', *bocsát* 'to let, to forgive', *koldul* 'to beg', *gyárt* 'to produce', *sárga*

'yellow', and *söpör* 'to sweep, to broom'. Apart from that, 35 Cuman words are presented in the book. The size of this data is so far the largest ever among all relevant studies. The book has been dealt with by many Hungarian and international researchers since its publication as they are listed in (Honti, 2017, p. 7). The authors state that all the data were taken from the cited texts or dictionaries. Some of the lexemes are archaic or dialectal, some of them are not in use anymore, and most of them are from the present Hungarian literary language. While a considerable amount of these heads in the lexicon are of dubious (55 words) or possible (38 words) Turkic origin, a handful of them are of non-Turkic origin but were loaned from WOT into Hungarian. Turkic etymology of the disputed words is most likely (Róna-Tas & Berta, 2011, p. 1489). The questionable lexemes of the dictionary were analyzed by Honti: "In this study, I am dealing with all words whose origins have been criticized by the authors of the publication so far, and in connection with them I sometimes also have something to say about other etymologies" (Honti, 2017, p. 7; translated by Andrea Parapatics, Szilárd Szentgyörgyi). It is understood that no decision should be made regarding these doubtful cases as they are not the main focus of the research.

For the comparison of the 377 word pairs, the latest reconstructed WOT forms, the initial reconstructed Old Hungarian forms at the time of the borrowing process, and the modern Hungarian forms are taken into account. All the data is taken from (Róna-Tas & Berta, 2011). For the orthographic display of the WOT words, Turkish orthography, and for the display of the reconstructed and modern Hungarian words, Hungarian orthography is used. Nevertheless, orthography is only for display and thus irrelevant since it does not play a role in the calculations. Reconstructed Hungarian words are referred to as Old Hungarian (O..H..). Phonetic transcriptions of the lexemes are presented in brackets. English translations are given on top of the tables. WOT words that have a debated Turkic origin are marked with asterisks. Double asterisks indicate that the origin is non-Turkic but the word was loaned from WOT.

In the data section of this dissertation, all information pertaining to West Old Turkic material in Hungarian has been sourced exclusively from the comprehensive study conducted by Róna-Tas and Berta in 2011. This particular work, spanning a decade and executed by acknowledged experts in the field, serves as a significant reference for the subject matter.

However, it is essential to acknowledge that, despite its depth and scrutiny, this source may not be considered the ultimate and flawless output in the etymology of West Old Turkic material in Hungarian. Certain sections may possess inherent flaws, and specific etymologies remain subjects of ongoing academic discussion. Nevertheless, this source has been selected for its up-to-date nature and unparalleled comprehensiveness within the current literature. By utilizing it as the exclusive reference, the aim is to maintain the integrity of the data presented in this dissertation.

# CHAPTER TWO

## Methods

The Levenshtein distance formula measures the minimum number of single-character edits (insertions, deletions, or substitutions) required to transform one string into another. It operates under the assumption that the closer two strings are in terms of these edits, the more similar they are. This method is particularly suitable for linguistic distance calculation due to its ability to capture both phonetic and orthographic variations between languages or dialects. Comparative linguistics aims to reconstruct the evolutionary relationships between languages and uncover their common ancestral roots. The Levenshtein distance formula provides an effective means to compare languages based on their lexical, morphological, or phonetic features. By measuring the distance between words or sounds in different languages, researchers can infer patterns of language change and divergence, helping to develop language family trees and uncover the historical development of languages. In sociolinguistics and dialectology, the Levenshtein distance formula is employed to assess the linguistic variation within a speech community or region. By comparing speech samples from different speakers or geographical locations, researchers can quantify the phonetic or lexical divergence between dialects. This information aids in understanding language change over time, identifying linguistic boundaries, and investigating language contact phenomena.

Although LD is applied mainly to related languages in order to measure mutual intelligibility or to help with reconstruction studies, the formula has proven to be efficient also in loan phonology even for genetically and typologically distant languages such as Russian-Dolgan (Stachowski K., 2010) or Kipchak Turkic-Iranian (Van Der Ark et al., 2007).

“It seems that LD is much more often applied to phonologically quite similar languages such as Dutch, English, German, or Norwegian dialects. Moreover, most of these languages are phonotactically relatively rich and therefore lenient, which appears to be the key here” (Stachowski K., 2011, p. 155).

The classical application of the LD requires *phonemic alignment*, and also the "3 algorithm formulation" that presents itself as *insertion*, *deletion*, and *substitution.* The phonemic alignment is a meticulous process in which the phonemes of the pair are matched with each other. Extra or missing phonemes appear as insertions and deletions. The phonemes that are different in each alignment appear as substitutions. For each insertion, deletion, and substitution a designated penalty point (operation cost) is applied in order to reveal the necessary number of edits that are required to change one unit into another (edit distance).

In the example presented in Table 3, seven alignments can be observed between the Old Hungarian copy and its WOT original. Three of these appear as penalties, one being a deletion and two being substitutions. The high back unrounded vowel that exists in the source word is missing in the target word with no substitution counterpart. Comparisons that cost penalty points will be shown in **bold red** throughout the study. Three penalty points out of seven alignments conclude a 42.86% distance between the pair.

**Table 3. An example of the classical phonetic distance calculation**

| **gyilkos** 'murderer, killer' | |
|---|---|
| O.H. *jilkus* | [ j **i** l k u **ʃ** - ] |
| WOT. *yulkuçı** | [ j **u** l k u **t͡ʃ ï** ] |
| Penalties | 3 |
| N-align | 7 |
| LD% | 42.86 |

The problem here is that the LD supposedly fails to reflect the phonetic distance between two words in fine detail due to the reason that the phoneme pairs are rarely in equal distance to each other. It is necessary to modify the operation cost in Levenshtein algorithms so that a more sensitive and perceptually acceptable edit distance can be achieved. Several refined algorithm methods have been proposed in the literature in order to achieve the above-mentioned. One example is the pointwise mutual information (PMI) based Levenshtein algorithm, according to which the distances are measured by the PMI values, calculated by the frequency of collocation of two units in a designated corpus. Another method is

computerizing the acoustic features of sound frequencies by visualizing them on a spectrogram. This approach is referred to as the spectrogram-based Levenshtein algorithm (Zhang, 2018, p. 11). The failure of the classical LD approach and the need for a phonologically sensitive method can be seen as an example in the citation below.

"Using the Levenshtein algorithm to calculate phonological distances between lexical units does not account for the fact that two phonemes may be more or less close depending on the number of distinctive features they share together. For example, the distance between the French words [bo] *(beau)* and [pɔ̃] *(pont)* is the same as between [bo] *(beau)* and [ʁi] *(riz)*, even though the first pair of words shares more phonological features than the second pair and may thus be thought as much closer perceptively" (Fontan, et al., 2016, p. 650).

The approach which is applied in this study is a comprehensive version of the "sound class-based Levenshtein algorithm" (Zhang, 2018). This method assesses the phonemes with respect to their distinctive features. This puts the sounds in groups according to their features, making them less or more distant from each other in compliance with the number of mutual features that they share, therefore being subject to fewer or more penalty points in the operation cost accordingly. One example of this approach is given by (McCoy & Frank, 2018). They would apply one penalty point for the change of /d/ into /t/ as only one distinctive feature is changed that being the *voicing.* For /b/ into /t/ however, they would apply two penalty points since this time not only the voicing but the *place of articulation* (placing) is different as well. Approaches considering several different phonetic and phonological theories and rules can be seen in the literature such as feature theory (Jakobson et al., 1951) or probabilistic methods (Sanders & Chin, 2009, pp. 96-114) but the issue of how many penalty points are to be applied for a simple one-character change seemingly remains as a determining problem as the refinement methods differ substantially.

Key Components of the Distinctive Feature Weighted Levenshtein Algorithm (Babych, 2016; Fontan, et al., 2016; McCoy & Frank, 2018; Ghio et al., 2020; Vakulenko, 2021):

- Feature Selection: The identification of distinctive features relevant to the application is crucial for the accuracy of the Distinctive Feature Weighted Levenshtein Algorithm. These features may encompass phonetic properties, semantic

associations, contextual information, or any other attributes that provide meaningful differentiation between characters.

- Feature Weight Assignment: The assignment of appropriate weights to each distinctive feature allows for algorithm customization to specific contexts. Weights can be determined through statistical analysis, expert knowledge, or machine learning techniques, based on the available resources and application requirements.
- Weighted Edit Distance Computation: The Distinctive Feature Weighted Levenshtein Algorithm calculates the weighted edit distance by considering the distinctive feature weights during the dynamic programming phase. The weight assigned to each operation (insertion, deletion, substitution) is determined by the corresponding feature weights of the characters involved.

The Distinctive Feature Weighted Levenshtein Algorithm offers several advantages over traditional string comparison methods. Firstly, it provides increased precision by considering distinctive features and assigning weights to capture the nuances of different applications. This customization allows the algorithm to adapt to specific domains and yield more accurate results. Additionally, the algorithm is compatible with existing string comparison methods, making it easy to integrate into existing systems without extensive modifications. However, there are some limitations and areas for future improvement. The selection of distinctive features plays a crucial role in the algorithm's effectiveness, requiring further research and refinement. Additionally, determining optimal weights for these features can be challenging and subjective, warranting the development of automated approaches for weight assignment. Furthermore, efforts should focus on optimizing the computational complexity of the algorithm without compromising accuracy. The establishment of standardized evaluation metrics and benchmarks would also enhance the comparison and validation of the Distinctive Feature Weighted Levenshtein Algorithm against other string comparison techniques, facilitating a clearer understanding of its strengths and weaknesses. Expanding the algorithm to multilingual settings would further broaden its applicability in diverse global contexts.

Since the operation cost is refined according to the differences in the distinctive features of the adaptations, for each difference in features, a designated penalty point is applied. The idea behind this formulation is that a single feature change is hypothetically not enough to

match a full perceived distance on a pair of phonemes but it may be rather conceived as a modification (allophone) of the same phoneme which is indeed the case for some adaptations. Some examples are tense and lax counterparts of the vowels, velarization, palatalization, etc. There are 20 distinctive features addressed for this approach, covering Hungarian and Turkic phonemes, and they fall under 5 base categories as shown below. The effort here is to include any feature that makes a difference even in one pair among all the combinations. The *constricted glottis* feature is therefore not included as it does not apply to any phonemes in the languages that are subject to comparison in this study. In the same way, the *radical* feature is not included either since there is no epiglottal consonant in the set. The *strident* feature which is mainly used to distinguish the interdental fricative from the alveolar fricative, as well as some other types of fricatives, is neglected also since it appears as a sub-feature of a sub-feature and supposedly would need extra refining.

- Major class features (syllabic, consonantal, sonorant)
- Laryngeal features (voice, spread glottis)
- Manner features (continuant, nasal, lateral, delayed-release)
- Place features (labial (round), coronal (anterior, distributed), dorsal
- Vowel space (high/low/mid, back, tense)

Major class features are used to classify speech sounds into three broad categories: syllabic, consonantal, and sonorant. It's important to note that these categories are not mutually exclusive, and many speech segments can exhibit multiple characteristics simultaneously. Syllabic sounds form the core of a syllable and carry the syllabic nucleus, such as vowels. Consonantal sounds are characterized by some form of vocal tract obstruction or constriction and encompass a variety of segments, including plosives, fricatives, and affricates. It's worth noting that these segments, plosives, fricatives, and affricates, belong to the broader category of obstruents, which are consonantal and non-sonorant in nature. While the subsequent content in this work correctly employs these distinctions, it's essential to clarify that plosives, fricatives, and affricates are subclasses of obstruents, rather than a distinct category known as consonantal segments. Sonorant sounds are characterized by relatively open vocal tract configurations and include nasal sounds, liquids, glides, and vowels.

Laryngeal features describe the state of the vocal folds during the production of speech sounds. Two common laryngeal features are voice and spread glottis. Voice refers to the vibration of the vocal folds when producing a sound. Vowels and voiced consonants are typically produced with vocal fold vibration, while voiceless consonants are produced without vocal fold vibration. Spread glottis, on the other hand, involves a wider opening between the vocal folds, which results in voiceless sounds like the initial sound in the word *hat*.

Manner features describe the way airflow is modified or obstructed during the production of speech sounds. Continuant sounds allow a relatively unrestricted airflow and include fricatives, approximants, and vowels. Nasal sounds are produced by lowering the velum, allowing air to escape through the nose. Lateral sounds are produced by obstructing the central part of the vocal tract and allowing the airflow to escape along the sides, such as the /l/ sound in *lip*. Delayed-release sounds involve a delay in the release of the constriction, as in the affricate sound in *church*.

Place features refer to the location of the main constriction or articulation in the vocal tract during sound production. Labial sounds involve the lips and include bilabial sounds like /p/ and /b/ as well as labiodental sounds like /f/ and /v/. Coronal sounds involve the front part of the tongue and include sounds like dental, alveolar, and postalveolar. Dorsal sounds involve the back part of the tongue and include sounds like velar and uvular.

Vowel space refers to the arrangement of vowel sounds in a two-dimensional acoustic space based on their height (high, mid, low) and backness (front, central, back). Additionally, vowels can exhibit various distinctions, including tense and lax qualities. Tense vowels often involve a greater articulatory effort and can be characterized by a slightly longer duration in some languages, while lax vowels may have a relatively shorter duration. It's essential to note that the relationship between vowel tension and stress varies across languages. While in some languages, tense vowels are more likely to occur in stressed syllables, this pattern does not hold universally. For example, Hungarian features both short and long vowels, but these distinctions do not necessarily correspond to lax versus tense qualities, and they can be present in both stressed and unstressed syllables.

The combination of these features allows linguists to describe and analyze the sounds of human language in a systematic way. By identifying and categorizing these features, we can gain insights into the phonetic patterns and structures of different languages. Moreover, these features play a crucial role in phonological processes, such as assimilation, dissimilation, and vowel harmony, which shape the sound systems of languages.

For an adaptation to have 1 full penalty point from a particular base category, all the sub-features of that category should be in contrast in the pair. For example, the sound /a/ is [+low], [+back], and [-tense] while /y:/ is [+high (-low)], [-back], and [+tense]. Therefore this pair shows full contrast in terms of the "vowel space" base category. The reason that a full penalty should be equal to 1, is that every phoneme in a lexeme (or every digit in a unit) matches with only one other and takes the space of 1 alignment. Consequently, for a hypothetical one-phoneme word to be 100% distant from another one-phoneme word, there needs to be 1 penalty point between the sides. That being the case, in this approach the features that belong to a particular base category should equally participate in the measurement of the distance. In the case of vowel space, 0.33 penalty points would be applied for contrast in [±high], 0.33 for [±back], and again 0.33 for [±tense], since there are three features in this category. Considering the fact that this pair in the example above also shares a difference in placing, /a/ being [–labial] and /y:/ being [+labial (+round)], and knowing that the place category also has three main features (labial, coronal, dorsal), 0.33 additional points would be added. Therefore the distance between the pair would cost 1.33 points in total. That being the case, very distinct pairs, and vowel-to-consonant and vice versa adaptations would generally cost more than a full penalty point, further sensitizing the operation cost. A standard 2 penalty points approach for the vowel-to-consonant and vice versa adaptations is problematic since these types of assimilations are not always equally distant either. The refinement criteria are applied to the above-given example, in **Hiba! A hivatkozási forrás nem található.** below.

The explanation for the application of the refined LD algorithm in Table 4 is as follows: There are 7 phonemic matches between the words. This gives us the phonemic alignment (N-align) value. The Hungarian word has 1 less phoneme than the source word. That indicates a deletion. If an extra phoneme was added to the Hungarian word, that would mean an insertion. Deletions and insertions cost 1 penalty point. There are also 2 substitutions. "The

substitution cost of **/u/➔/i/** is 0.66. The former is [+labial] and [+back] while the latter is [-labial] and [-back]. These features cost 0.33 points each since the categories they belong to (place and vowel space categories) have three features each. Therefore 1/3 penalty points would be applied for each of these features. Finally, the cost of **/t͡ʃ/➔/ʃ/** is 0.50. The former is [-continuant] and [+delayed release] while the latter is [+continuant] and [-delayed release]. Both these features are manner features and cost 0.25 each since there are four features in that category. The sum of these penalties is equal to 2.16 which is 30.86% of the alignment value; 7." That gives us the phonemic distance between the two lexemes.

**Table 4. Refined distance**

| **gyilkos** 'murderer, killer' | |
|---|---|
| O.H. *jilkus* | [ j **i** l k u **ʃ** - ] |
| WOT. *yulkuçı** | [ j **u** l k u **t͡ʃ** **ï** ] |
| Penalties | 2.16 |
| N-align | 7 |
| LD% | 30.86 |

For insertions and deletions, 1 penalty point is applied. An example can be observed in **Hiba! A hivatkozási forrás nem található.**.

**Table 5. Refinement for deletions**

| **gyapjú** 'wool' | |
|---|---|
| O.H. *japgu* | [ j a p **-** ɣ u ] |
| WOT. *yapağu* | [ j a p **a** ɣ u ] |
| Penalties | 1 |
| N-align | 6 |
| LD% | 16.67 |

Some sub-features have their own further sub-features (Clements & Keyser, 1983). These are labial (round), coronal (anterior, distributed), and dorsal (high, back). In the case of labial, coronal, or dorsal features being parallel but the super-sub-features differing in a pair,

0.33/2=0.16 penalty points are to be applied for each of them. In **Hiba! A hivatkozási forrás nem található.**, the adaptation **/ð/➔/z/** costs 0.16 penalty points since the former is [+distributed] and the latter is [-distributed] despite both sounds being [+coronal].

**Table 6. Further refinement for super-sub features**

| **gaz** 'weed' | |
|---|---|
| O.H. *kagzu* | [ k a ɣ **z** u] |
| WOT. *kağdu** | [ k a ɣ **ð** u ] |
| Penalties | 0.16 |
| N-align | 5 |
| LD% | 3.2 |

[±tense] feature costs 0.33. It is argued that consequently gemination or consonant lengthening and degemination should cost the same penalty points as vowel lengthening and shortenings. So ideally the approach to vowel and consonant lengthenings would be coherent this way. An example is presented in Table 7.

**Table 7. Refinement for gemination**

| **csákány** 'pick-axe, war hammer' | |
|---|---|
| O.H. *csakan* | [ t͡ʃ a **k** a n ] |
| WOT. *çakkan* | [ t͡ʃ a **k:** a n ] |
| Penalties | 0.33 |
| N-align | 5 |
| LD% | 6.6 |

Metathesis is observed in some adaptations and if the sounds in transposition are the same, 1 penalty point is applied, if they are also substituted, the extra cost is additionally applied. In Table 8, [t] sound and the neighboring vowels are subject to transposition; therefore 1 penalty point is applied while another 0.33 is applied for the substitution from [ɨ] to [i].

**Table 8. Refinement for metathesis**

| **csalit** ‘thicket, brushwood’ | |
|---|---|
| O.H. *csalit* | [ t͡ʃ a l **i t** ] |
| WOT. *çaltı** | [ t͡ʃ a l **t ï** ] |
| Penalties | 1.33 |
| N-align | 5 |
| LD% | 26.6 |

The open central unrounded vowel [a] acts as a back vowel in the harmonical system of Hungarian and Turkish. Vowel harmony is very typical in all Turkic languages just as it is typical in Hungarian, and this originally central sound apparently fits in the harmonical system and displays the [+back] feature. Therefore, this sound is evaluated as [+back] in the matrix since [+central] is not included in this matrix. It is lacking in the other vowels of the set in the first place.

The voiced velar fricative [ɣ] in modern Turkish, orthographically shown with the letter <ğ> has become silent and has no sound of its own (Selen, 1979; Ergenç & Uzun, 2020; Ünal-Logacev et al., 2019; Uzun, 2021), with its effect varying depending on its location in a word and the surrounding vowels. This is a recent condition as in Ottoman Turkish the non-silent variant was in use (Ünal et al., 2019, p. 187) as well as in some dialects and other Turkic languages today. Therefore, there is no reason to think that in the Old Turkic, it was silent.

However, the voiced velar fricative /ɣ/ commonly occurring in intervocalic position raises the possibility of its transitional or sonorant nature. Nevertheless, it should be noted that the interpretation of a sound as “transitional” relies on phonological analysis, which can vary across linguistic theories and frameworks. Several justifications support considering /ɣ/ as a transitional sound. Firstly, its distribution in numerous languages, including Turkish, predominantly places it between vowels, indicating a sonority or transitional pattern as it links vowel sounds. Secondly, the Sonority Sequencing Principle, which governs sound organization within syllables, suggests that sounds with higher sonority, like vowels, tend to be closer to the nucleus of a syllable. As consonants, including fricatives such as /ɣ/, typically

exhibit lower sonority, the presence of /ɣ/ between two vowels aligns with this principle by maintaining a sonority pattern within the syllable structure. Lastly, the articulation of /ɣ/ involves a constriction in the velar region, specifically the back of the mouth, indicating an intermediate position between a complete closure (as in a plosive) and a fully open vowel. It should be emphasized that the label "transitional" is not universally applied to /ɣ/ in all linguistic frameworks. Different theories and analyses may employ alternative terms or descriptions. Nevertheless, the outlined patterns of distribution, sonority, and articulation provide support for considering /ɣ/ as exhibiting characteristics of a transitional or sonorant sound. For these reasons, the insertions and deletions of this sound cost a semi (0.5) penalty point. In the same way, insertions and deletions of [j] which is also a transitional semivowel, and [w] cost semi-penalty points as well. A demonstration of this refinement is in Table 9.

**Table 9. Refinement for the voiced velar fricative**

| **csipa** 'secretion, mucus discharged from the eyes' | |
|---|---|
| O.H. *csapa* | [ t͡ʃ a p a - ] |
| WOT. *çapağ* | [ t͡ʃ a p a ɣ ] |
| Penalties | 0.5 |
| N-align | 5 |
| LD% | 10 |

[e] and [ɛ] sounds share all the same features except [±low], although they are both officially mid-sounds. The former being closer to the high category (high-mid) and the latter being closer to the low category (low-mid) the appropriate penalty point for the [±low] feature is applied (0.33). The distinctive features matrix of Turco-Hungarian phonemes is displayed in Table 10 and Table 11.

**Table 10. Distinctive feature matrix of Turco-Hungarian Phonemes (1)**

| | MAJOR CLASS FEATURES | | | LARYNGEAL FEATURES | | MANNER FEATURES | | | |
|---|---|---|---|---|---|---|---|---|---|
| IPA | syllabic | consonantal | sonorant | voice | spread glottis | continuant | nasal | lateral | delayed release |
| [ɒ] | + | - | + | + | - | + | - | - | - |
| [a] | + | - | + | + | - | + | - | - | - |
| [aː] | + | - | + | + | - | + | - | - | - |
| [æ] | + | - | + | + | - | + | - | - | - |
| [ɛ] | + | - | + | + | - | + | - | - | - |
| [e] | + | - | + | + | - | + | - | - | - |
| [eː] | + | - | + | + | - | + | - | - | - |
| [i] | + | - | + | + | - | + | - | - | - |
| [iː] | + | - | + | + | - | + | - | - | - |
| [ɯ] | + | - | + | + | - | + | - | - | - |
| [o] | + | - | + | + | - | + | - | - | - |
| [oː] | + | - | + | + | - | + | - | - | - |
| [ø] | + | - | + | + | - | + | - | - | - |
| [øː] | + | - | + | + | - | + | - | - | - |
| [u] | + | - | + | + | - | + | - | - | - |
| [uː] | + | - | + | + | - | + | - | - | - |
| [y] | + | - | + | + | - | + | - | - | - |
| [yː] | + | - | + | + | - | + | - | - | - |
| [b] | - | + | - | + | - | - | - | - | - |
| [c] | - | + | - | - | - | - | - | - | - |
| [d] | - | + | - | + | - | - | - | - | - |
| [d͡z] | - | + | - | + | - | - | - | - | + |
| [d͡ʒ] | - | + | - | + | - | - | - | - | + |
| [f] | - | + | - | - | - | + | - | - | - |
| [g] | - | + | - | + | - | - | - | - | - |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| [ɣ] | - | + | - | + | - | + | - | - | - |
| [h] | - | - | - | - | + | + | - | - | - |
| [χ] | - | + | - | - | + | + | - | - | - |
| [j] | - | - | + | + | - | + | - | - | - |
| [ɟ] | - | + | - | + | - | - | - | - | - |
| [k] | - | + | - | - | - | - | - | - | - |
| [l] | - | + | + | + | - | + | - | + | - |
| [ɬ] | - | + | + | + | - | + | - | + | - |
| [m] | - | + | + | + | - | - | + | - | - |
| [n] | - | + | + | + | - | - | + | - | - |
| [ŋ] | - | + | + | + | - | - | + | - | - |
| [ɲ] | - | + | + | + | - | - | + | - | - |
| [p] | - | + | - | - | - | - | - | - | - |
| [r] | - | + | + | + | - | + | - | - | - |
| [s] | - | + | - | - | - | + | - | - | - |
| [ʃ] | - | + | - | - | - | + | - | - | - |
| [t] | - | + | - | - | - | - | - | - | - |
| [t͡s] | - | + | - | - | - | - | - | - | + |
| [t͡ʃ] | - | + | - | - | - | - | - | - | + |
| [v] | - | + | - | + | - | + | - | - | - |
| [z] | - | + | - | + | - | + | - | - | - |
| [ð] | - | + | - | + | - | + | - | - | - |
| [ʒ] | - | + | - | + | - | + | - | - | - |
| [w] | - | - | + | + | - | + | - | - | - |

**Table 11. Distinctive feature matrix of Turco-Hungarian Phonemes (2)**

| | PLACE FEATURES | | | | | | VOWEL SPACE | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| IPA | labial | round | coronal | anterior | distributed | dorsal | high | back | low | tense |
| [ɒ] | + | + | - | n/a | n/a | + | - | + | + | - |
| [a] | - | n/a | - | n/a | n/a | + | - | + | + | - |
| [aː] | - | n/a | - | n/a | n/a | + | - | + | + | + |
| [æ] | - | n/a | - | n/a | n/a | + | - | - | + | - |
| [ɛ] | - | n/a | - | n/a | n/a | + | - | - | - | - |
| [e] | - | n/a | - | n/a | n/a | + | - | - | - | - |
| [eː] | - | n/a | - | n/a | n/a | + | - | - | - | + |
| [i] | - | n/a | - | n/a | n/a | + | + | - | - | - |
| [iː] | - | n/a | - | n/a | n/a | + | + | - | - | + |
| [ɯ] | - | n/a | - | n/a | n/a | + | + | + | - | - |
| [o] | + | + | - | n/a | n/a | + | - | + | - | - |
| [oː] | + | + | - | n/a | n/a | + | - | + | - | + |
| [ø] | + | + | - | n/a | n/a | + | - | - | - | - |
| [øː] | + | + | - | n/a | n/a | + | - | - | - | + |
| [u] | + | + | - | n/a | n/a | + | + | + | - | - |
| [uː] | + | + | - | n/a | n/a | + | + | + | - | + |
| [y] | + | + | - | n/a | n/a | + | + | - | - | - |
| [yː] | + | + | - | n/a | n/a | + | + | - | - | + |
| [b] | + | - | - | n/a | n/a | - | n/a | n/a | | |
| [c] | - | n/a | + | - | + | + | + | - | | |
| [d] | - | n/a | + | + | - | - | n/a | n/a | | |
| [d͡z] | - | n/a | + | - | + | - | n/a | n/a | | |
| [d͡ʒ] | - | n/a | + | - | + | - | n/a | n/a | | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| [f] | + | - | - | n/a | n/a | - | n/a | n/a |
| [g] | - | n/a | - | n/a | n/a | + | + | - |
| [ɣ] | - | n/a | - | n/a | n/a | + | + | + |
| [h] | - | n/a | - | n/a | n/a | - | n/a | n/a |
| [χ] | - | n/a | - | n/a | n/a | + | + | + |
| [j] | - | n/a | - | n/a | n/a | + | + | - |
| [ɟ] | - | n/a | + | - | + | + | + | - |
| [k] | - | n/a | - | n/a | n/a | + | + | + |
| [l] | - | n/a | + | + | - | - | n/a | n/a |
| [ɬ] | - | n/a | + | + | - | + | + | + |
| [m] | + | - | - | n/a | n/a | - | n/a | n/a |
| [n] | - | n/a | + | + | - | - | n/a | n/a |
| [ŋ] | - | n/a | - | n/a | n/a | + | + | + |
| [ɲ] | - | n/a | + | - | + | + | + | - |
| [p] | + | - | - | n/a | n/a | - | n/a | n/a |
| [r] | - | n/a | + | + | - | - | n/a | n/a |
| [s] | - | n/a | + | + | - | - | n/a | n/a |
| [ʃ] | - | n/a | + | - | + | - | n/a | n/a |
| [t] | - | n/a | + | + | - | - | n/a | n/a |
| [t͡s] | - | n/a | + | + | - | - | n/a | n/a |
| [t͡ʃ] | - | n/a | + | - | + | - | n/a | n/a |
| [v] | + | - | - | n/a | n/a | - | n/a | n/a |
| [z] | - | n/a | + | + | - | - | n/a | n/a |
| [ð] | - | n/a | + | + | + | - | n/a | n/a |
| [ʒ] | - | n/a | + | - | + | - | n/a | n/a |
| [w] | + | + | - | n/a | n/a | + | + | + |

*n/a: not available.

# CHAPTER THREE

## Results

This dissertation has undergone numerous revisions and refinements to reach its current state. Initially, the adaptation process was treated as a singular entity. However, following Professor András Róna-Tas's review, the processes of borrowing and adaptation were distinguished, as the accommodation of a newly borrowed word relies on distinct factors. Primarily, it hinges on the congruence or proximity between the constituents and structures of the word in both the donor and recipient languages. Only the borrowed word itself undergoes further adaptation, a process influenced by time and frequency. Words encountered frequently undergo complete adaptation more rapidly. Meanwhile, as adaptation progresses, the recipient language undergoes its own evolution, with the borrowed word becoming integrated into this linguistic change. Furthermore, numerous morpho-phonological challenges were addressed through extensive revisions, involving the comprehensive reworking of the entire dataset several times, prompted by Professor Péter Siptár's feedback. Additionally, the calculation methodology and literature framework underwent multiple reassessments with guidance from the supervisors of this study, Andrea Parapatics and Szilárd Szentgyörgyi.

The adaptation of loanwords in language contact situations involves intricate processes that shape the incorporation of foreign words into the target language. This study focuses on the alteration rates of 377 West Old Turkic loanwords in Hungarian, aiming to shed light on the dynamics of borrowing and adaptation in this linguistic context. The investigation is carried out by examining two key processes: the initial copying process (borrowing) and the subsequent adaptation of these loanwords within the acoustic system of the target language. The first process, borrowing, involves the direct transfer of words from the source language West Old Turkic, into the target language Hungarian. During this stage, minimal changes are detected as the target language copies the words from the source language. The alteration rate for this process is determined by comparing the reconstructed West Old Turkic words with the reconstructed Old Hungarian copies. Findings indicate an initial borrowing alteration rate of 7.21% for the West Old Turkic copies in Hungarian. The second process,

adaptation, encompasses the rule-based changes that occur within the loanwords to align with the evolving phonetic and phonological system of the target language.

It is important to reiterate from the outset that when loanwords become part of a new language, they will undergo the same phonological alterations as any other native words, irrespective of their source. This underscores the fact that phonological shifts in a language usually affect all words, not just loanwords, and aren't solely influenced by the desire to make loanwords fit in. This aspect is crucial in debunking any misunderstanding that all loanwords naturally conform to the phonological features of the host language over time; sometimes, they may, in fact, go through changes in the opposite direction. It’s worth noting that the changes that have unfolded throughout the lengthy period from the times of borrowing to the present day are predominantly linked to the natural evolution of Hungarian phonology, as opposed to the process of loanword assimilation. Nevertheless, it is acceptable to use the term 'adaptations' to refer to these changes, whether these adaptations are specific to loanword integration or not. While the term is often associated with the specific modifications made to accommodate loanwords, it can also encompass broader phonological shifts that affect the language as a whole. Language adapts to the needs and preferences of its speakers, and this adaptation can be seen in various aspects, including pronunciation, vocabulary, and grammar.

In this study, the adaptation of the loanwords is evaluated by comparing the reconstructed Old Hungarian copies with modern Hungarian equivalents. This assessment reveals the extent of rule-based modifications required for the loanwords to take their current form in modern Hungarian. The results demonstrate an adaptation rate of 28.67% for the loanwords in modern Hungarian.

The comparison tables presented in this section offer a comprehensive analysis of the alteration rates observed in both borrowing and adaptation processes. These tables provide a valuable overview of the changes undergone by the loanwords during their integration into the Hungarian language, revealing the phonetic and phonological adjustments that occurred over time. The findings presented in this study contribute to our understanding of loanword adaptation in language contact situations, specifically highlighting the complex interplay between borrowing and adaptation. The adaptation rates revealed through the comparison

tables offer insights into the linguistic dynamics between West Old Turkic and Hungarian, providing a foundation for further investigations into language contact phenomena.

## 3.1 Borrowing process

By comparing the reconstructed West Old Turkic words with the reconstructed Old Hungarian copies, the analysis reveals that the borrowing alteration rate for these loanwords is measured at 7.21% (92.79% homophones). This finding highlights the minimal changes that occur during the borrowing phase, providing insights into the direct transfer of lexical material from West Old Turkic to Old Hungarian. Notably, among the 377 pairs analyzed, an intriguing finding emerges: 178 pairs, accounting for approximately 47% of the sample, show 0% distance.

**Table 12. Initial alteration rates of WOT copies in Hungarian**

| **ács** ‘carpenter’ | | | **ágyú** ‘cannon, catapult’ | | |
|---|---|---|---|---|---|
| O.H. *agacsi* | [ a ɣ a **t͡ʃ i** ] | | O.H. *aldag* | [ a l d a ɣ] | |
| WOT. *ağaççı* | [ a ɣ a **t͡ʃ: ï** ] | | WOT. *aldağ* | [ a l d a ɣ] | |
| Cost 0.66 | N-align 5 | LD% 13.2 | Cost 0 | N-align 5 | LD% 0 |
| **ajn(ároz)** ‘to fondle, pet, caress’ | | | **ál** ‘false, imitation’ | | |
| O.H. *ajan* | [ a j a n ] | | O.H. *ál* | [ **a:** l ] | |
| WOT. *ayan** | [ a j a n ] | | WOT. *al* | [ **a** l ] | |
| Cost 0 | N-align 4 | LD% 0 | Cost 0.33 | N-align 2 | LD% 16.5 |
| **alacs** ‘pied (of an animals’ coat)’ | | | **alma** ‘apple’ | | |
| O.H. *alacs* | [ a l a t͡ʃ **-** ] | | O.H. *alma* | [ a l m a ] | |
| WOT. *alaçı** | [ a l a t͡ʃ **ï** ] | | WOT. *alma* | [ a l m a ] | |
| Cost 1 | N-align 5 | LD% 20 | Cost 0 | N-align 4 | LD% 0 |
| **általag** ‘barrel (wooden container and standard quantity)’ | | | **áp(ol)** ‘to nurse, take care of’ | | |
| O.H. *altalag* | [ a l t **a** l **a** g ] | | O.H. *op* | [ o p ] | |
| WOT. *altılıg** | [ a l t **ï** l **ï** g ] | | WOT. *op** | [ o p ] | |
| Cost 0.66 | N-align 7 | LD% 9.42 | Cost 0 | N-align 2 | LD% 0 |
| **áporod(ik)** ‘to decay, putrify, turn stale’ | | | **apró** ‘small, tiny’ | | |

| | | | | | |
|---|---|---|---|---|---|
| O.H. *oporo* | [ o p **o** r **o** ] | | O.H. *opurug* | [ o p u r u ɣ ] | |
| WOT. *opura** | [ o p **u** r **a** ] | | WOT. *opuruğ* | [ o p u r u ɣ ] | |
| Cost 1 | N-align 5 | LD% 20 | Cost 0 | N-align 6 | LD% 0 |
| **arat** ‘to mow’ | | | **ár(ik)** ‘to decay, go bad (of food, water)’ | | |
| O.H. *orat* | [ o r a t ] | | O.H. *ár* | [ **a:** r ] | |
| WOT. *orat* | [ o r a t **]** | | WOT. *ar** | [ **a** r ] | |
| Cost 0 | N-align 4 | LD% 0 | Cost 0.33 | N-align 2 | LD% 16.5 |
| **árok** ‘ditch, canal’ | | | **árpa** ‘barley’ | | |
| O.H. *áruk* | [ **a:** r u k ] | | O.H. *arpa* | [ a r p a ] | |
| WOT. *aruk* | [ **a** r u k ] | | WOT. *arpa* | [ a r p a ] | |
| Cost 0.33 | N-align 4 | LD% 8.25 | Cost 0 | N-align 4 | LD% 0 |
| **ártány** ‘barrow’ | | | **ászok** ‘gauntry, gantry (for supporting barrels)’ | | |
| O.H. *aritan* | [ a r **i** t a n ] | | O.H. *aski* | [ a s k **i** ] | |
| WOT. *arıtan* | [ a r **ï** t a n ] | | WOT. *askı* | [ a s k **ï ]** | |
| Cost 0.33 | N-align 6 | LD% 5.5 | Cost 0.33 | N-align 4 | LD% 8.25 |
| **báj** ‘charm’ | | | **balkány** ‘soggy place, moor, swamp, marsh’ | | |
| O.H. *baj* | [ b a j ] | | O.H. *balkan* | [ b a l k a n ] | |
| WOT. *bay** | [ b a j ] | | WOT. *balkan* | [ b a l k a n ] | |
| Cost 0 | N-align 3 | LD% 0 | Cost 0 | N-align 6 | LD% 0 |
| **balta** ‘hatchet, axe’ | | | **bálvány** ‘pagan idol, stone or wooden pillar’ | | |
| O.H. *balta* | [ b a l t a ] | | O.H. *balvan* | [ b a l v a n ] | |
| WOT. *balta*** | [ b a l t a ] | | WOT. *balvan** | [ b a l v a n ] | |
| Cost 0 | N-align 5 | LD% 0 | Cost 0 | N-align 6 | LD% 0 |
| **bán** ‘governor of Croatia (of the southern marches)’ | | | **barom** ‘cattle, livestock’ | | |
| O.H. *bán* | [ b **a: - -** n ] | | O.H. *barum* | [ b a r u m ] | |
| WOT. *bayan* | [ b **a j a** n ] | | WOT. *barum** | [ b a r u m ] | |
| Cost 1.83 | N-align 5 | LD% 36.6 | Cost 0 | N-align 5 | LD% 0 |
| **bársony** ‘velvet’ | | | **basz(ik)** ‘vulgar expression for having an intercourse’ | | |
| O.H. *barcsun* | [ b a r t͡ʃ u n ] | | O.H. *basz* | [ b a s ] | |

| | | | | | |
|---|---|---|---|---|---|
| WOT. *barçun*** | [ b a r t͡ʃ u n ] | | WOT. *bas* | [ b a s ] | |
| Cost 0 | N-align 6 | LD% 0 | Cost 0 | N-align 3 | LD% 0 |
| **bátor** ‘courageous, brave, valiant’ | | | **becsül** ‘to estimate, to esteem, to appreciate’ | | |
| O.H. *bagatur* | [ b a ɣ a t u r ] | | O.H. *becsel* | [ b **ɛ** t͡ʃ **ɛ** l ] | |
| WOT. *bağatur* | [ b a ɣ a t u r ] | | WOT. *biçil** | [ b **i** t͡ʃ **i** l ] | |
| Cost 0 | N-align 7 | LD% 0 | Cost 0.66 | N-align 5 | LD% 13.2 |
| **béka** ‘frog’ | | | **béklyó** ‘hobble, shackle, leg-iron’ | | |
| O.H. *baka* | [ b a k a ] | | O.H. *bikagu* | [ b **i** k a ɣ u ] | |
| WOT. *baka* | [ b a k a ] | | WOT. *bıkağu* | [ b **ï** k a ɣ u ] | |
| Cost 0 | N-align 4 | LD% 0 | Cost 0.33 | N-align 6 | LD% 5,5 |
| **bélyeg** ‘stamp, bond, mark’ | | | **bér** ‘wage, rent’ | | |
| O.H. *beleg* | [ b **ɛ** l **ɛ** g ] | | O.H. *bérü* | [ b **e:** r y ] | |
| WOT. *bäläg** | [ b **æ** l **æ** g ] | | WOT. *berü* | [ b **e** r y ] | |
| Cost 0.66 | N-align 5 | LD% 13.2 | Cost 0.33 | N-align 4 | LD% 8.25 |
| **bercel** ‘name of a people’ | | | **bertű** ‘grain or small globule of precious metal’ | | |
| O.H. *berszil* | [ b **ɛ** r s i l ] | | O.H. *bürtig* | [ b y r t i ɣ ] | |
| WOT. *bärsil** | [ b **æ** r s i l ] | | WOT. *bürtiğ* | [ b y r t i ɣ ] | |
| Cost 0.33 | N-align 6 | LD% 5.5 | Cost 0 | N-align 6 | LD% 0 |
| **bese** ‘a bird of prey’ | | | **besenyő** ‘Pecheneg, name of a tribe’ | | |
| O.H. *bese* | [ b **ɛ** ʃ **ɛ** ] | | O.H. *beseneg* | [ b **ɛ** **ʃ** **ɛ** n **ɛ** **ɣ** ] | |
| WOT. *bäşä** | [ b **æ** ʃ **æ** ] | | WOT. *bäçänäg* | [ b **æ** **t͡ʃ** **æ** n **æ** **g** ] | |
| Cost 0.66 | N-align 4 | LD% 16.5 | Cost 1.75 | N-align 7 | LD% 25 |
| **betű** ‘letter (of the alphabet)’ | | | **bicska** ‘pocket knife’ | | |
| O.H. *bitig* | [ b i t i **ɣ** ] | | O.H. *bicsak* | [ b **i** t͡ʃ a k ] | |
| WOT. *bitig* | [ b i t i **g** ] | | WOT *bıçak* | [ b **ï** t͡ʃ a k ] | |
| Cost 0.25 | N-align 5 | LD% 5 | Cost 0.33 | N-align 5 | LD% 6.6 |
| **bika** ‘bull’ | | | **bilincs** ‘shackles’ | | |
| O.H. *bika* | [ b **i** k a ] | | O.H. *bilecsek* | [ b i l **ɛ** t͡ʃ **ɛ** k ] | |
| WOT. *bıka* | [ b **ï** k a ] | | WOT. *biläçäk* | [ b i l **æ** t͡ʃ **æ** k ] | |
| Cost 0.33 | N-align 4 | LD% 8.25 | Cost 0.66 | N-align 7 | LD% 9.43 |

| **bíró** ‘judge’ | | | **bocsán(ik)** ‘to be forgiven’ | | |
|---|---|---|---|---|---|
| O.H. *birug* | [ b **i** r u **ɣ** ] | | O.H. *bolcsan* | [ b o l t͡ʃ a n ] | |
| WOT. *bırug** | [ b **ï** r u **g** ] | | WOT. *bolçan* | [ b o l t͡ʃ a n ] | |
| Cost 0.58 | N-align 5 | LD% 11.6 | Cost 0 | N-align 6 | LD% 0 |
| **bocsát** ‘to forgive’ | | | **bog** ‘knot, bend’ | | |
| O.H. *bulcsat* | [ b **u** l t͡ʃ a t ] | | O.H. *bog* | [ b o g ] | |
| WOT. *bolçat* | [ b **o** l t͡ʃ a t ] | | WOT. *bog** | [ b o g ] | |
| Cost 0.33 | N-align 6 | LD% 5.5 | Cost 0 | N-align 3 | LD% 0 |
| **boglya** ‘stack of hay’ | | | **bojt** ‘tassel, fringe, tuft, node’ | | |
| O.H. *bogul* | [ b o g u l ] | | O.H. *bugut* | [ b **u** **ɣ** u t ] | |
| WOT. *bogul** | [ b o g u l ] | | WOT. *bogut** | [ b **o** **g** u t ] | |
| Cost 0 | N-align 5 | LD% 0 | Cost | N-align 5 | LD% 11.6 |
| **bojtorján** ‘burdock’ | | | **boka** ‘ankle’ | | |
| O.H. *balturgan* | [ b a l t u r **ɣ** a n ] | | O.H. *baka* | [ b a k a ] | |
| WOT. *balturgan* | [ b a l t u r **g** a n ] | | WOT. *baka* | [ b a k a ] | |
| Cost 0.25 | N-align 9 | LD% 2.77 | Cost 0 | N-align 4 | LD% 0 |
| **boly** ‘anthill’ | | | **bor** ‘wine’ | | |
| O.H. *bogol* | [ b o **ɣ** **o** l ] | | O.H. *bor* | [ b o r ] | |
| WOT. *boğul** | [ b o **ɣ** **u** l ] | | WOT. *bor* | [ b o r ] | |
| Cost 0.33 | N-align 5 | LD% 6,6 | Cost 0 | N-align 3 | LD% 0 |
| **bor(ít)** ‘to cover, to overturn’ | | | **borjú** ‘calf’ | | |
| O.H. *bur* | [ b u r ] | | O.H. *buragu* | [ b u r a **ɣ** u ] | |
| WOT. *bur* | [ b u r ] | | WOT. *burağu* | [ b u r a **ɣ** u ] | |
| Cost 0 | N-align 5 | LD% 0 | Cost 0 | N-align 6 | LD% 0 |
| **bors** ‘pepper’ | | | **borsó** ‘pea, bean’ | | |
| O.H. *burcs* | [ b u r t͡ʃ ] | | O.H. *burcsag* | [ b u r t͡ʃ a **ɣ** ] | |
| WOT. *burç* | [ b u r t͡ʃ ] | | WOT. *burçağ* | [ b u r t͡ʃ a **ɣ** ] | |
| Cost 0 | N-align 4 | LD% 0 | Cost 0 | N-align 6 | LD% 0 |
| **borz** ‘badger’ | | | **boszorkány** ‘witch, sorceress’ | | |
| O.H. *borszi* | [ b o r s **i** ] | | O.H. *baszarkan* | [ b a s a r k a n ] | |

| | | | | | |
|---|---|---|---|---|---|
| WOT. *borsı* | [ b o r s **ï** ] | | WOT. *basarkan* | [ b a s a r k a n ] | |
| Cost 0.33 | N-align 5 | LD% 6,6 | Cost 0 | N-align 8 | LD% 0 |
| **bosszant** ‘to annoy’ | | | **bot** ‘stick, cane’ | | |
| O.H. *buszan* | [ b u s a n ] | | O.H. *butig* | [ b u t **i** ɣ ] | |
| WOT. *busan* | [ b u s a n ] | | WOT. *butığ* | [ b u t **ï** ɣ ] | |
| Cost 0 | N-align 6 | LD% 0 | Cost 0.33 | N-align 5 | LD% 6,6 |
| **bő** ‘rich, abundant, roomy’ | | | **bögöly** ‘horsefly, gadfly’ | | |
| O.H. *beg* | [ b **ɛ** ɣ ] | | O.H. *bögölök* | [ b ø g ø l **ø** k ] | |
| WOT. *bäğ* | [ b **æ** ɣ ] | | WOT. *bögöläk** | [ b ø g ø l **æ** k ] | |
| Cost 0.33 | N-align 3 | LD% 11 | Cost 0.66 | N-align 7 | LD% 9.43 |
| **böjt** ‘fast(ing)’ | | | **bölcs** ‘wise’ | | |
| O.H. *bügte* | [ b y ɣ t **ɛ** ] | | O.H. *bügücsi* | [ b y ɣ y t͡ʃ i ] | |
| WOT. *büğtä** | [ b y ɣ t **æ** ] | | WOT. *büğüçi* | [ b y ɣ y t͡ʃ i ] | |
| Cost 0.33 | N-align 5 | LD% 6,6 | Cost 0 | N-align 6 | LD% 0 |
| **bölcső** ‘cradle’ | | | **bölény** ‘bison’ | | |
| O.H. *belcseg* | [ b e l t͡ʃ **ɛ** ɣ ] | | O.H. *belen* | [ b **ɛ** l **ɛ** n ] | |
| WOT. *belçiğ* | [ b e l t͡ʃ **i** ɣ ] | | WOT. *bülän* | [ b **y** l **æ** n ] | |
| Cost 0.33 | N-align 6 | LD% 5.5 | Cost 1 | N-align 5 | LD% 20 |
| **börtön** ‘prison, jail’ | | | **bú** ‘sorrow’ | | |
| O.H. *berten* | [ b **ɛ** r t **ɛ** n ] | | O.H. *bug* | [ b u ɣ ] | |
| WOT. *bärtän* | [ b **æ** r t **æ** n ] | | WOT. *buğ* | [ b u ɣ ] | |
| Cost 0.66 | N-align 6 | LD% 11 | Cost 0 | N-align 3 | LD% 0 |
| **búcsú** ‘farewell, indulgence, dedication festival’ | | | **bular** ‘Volga Bulgar’ | | |
| O.H. *bulcsug* | [ b **u** l t͡ʃ u ɣ ] | | O.H. *bular* | [ b u l a r ] | |
| WOT. *bolçuğ* | [ b **o** l t͡ʃ u ɣ ] | | WOT. *bular* | [ b u l a r ] | |
| Cost 0.33 | N-align 6 | LD% 5,5 | Cost 0 | N-align 5 | LD% 0 |
| **búsz** ‘steam, vapour, fumes’ | | | **búza** ‘wheat’ | | |
| O.H. *búsz* | [ b u: s ] | | O.H. *bugzai* | [ b u ɣ **z** a **i** ] | |
| WOT. *bús* | [ b u: s ] | | WOT. *buğday* | [ b u ɣ **ð** a **j** ] | |
| Cost 0 | N-align 3 | LD% 0 | Cost 0.49 | N-align 6 | LD% 8.17 |

| **bükk** ‘beech’ | | | **bűn** ‘sin’ | | |
|---|---|---|---|---|---|
| O.H. *bik* | [ b i k ] | | O.H. *bűn* | [ b y: n ] | |
| WOT. *bik* | [ b i k ] | | WOT. *bűn* | [ b y: n ] | |
| Cost 0 | N-align 3 | LD% 0 | Cost 0 | N-align 3 | LD% 0 |
| **bütü** ‘end of sg, a rod with an end cut’ | | | **bűv(ös)** ‘magical’ | | |
| O.H. *bütüg* | [ b y t y ɣ ] | | O.H. *bügü* | [ b y ɣ y ] | |
| WOT. *bütüğ* | [ b y t y ɣ ] | | WOT. *büğü* | [ b y ɣ y ] | |
| Cost 0 | N-align 5 | LD% 0 | Cost 0 | N-align 4 | LD% 0 |
| **bán(ik)** ‘to regret, to be sorry’ | | | **csök** ‘feast, christening feast (at baptism)’ | | |
| O.H. *bagin* | [ b a **ɣ i** n ] | | O.H. *csük* | [ t͡ʃ **y** k ] | |
| WOT. *bagın* | [ b a **g ï** n ] | | WOT. *çök** | [ t͡ʃ **ø** k ] | |
| Cost 0.58 | N-align 5 | LD% 11.6 | Cost 0.33 | N-align 3 | LD% 11 |
| **cigle** ‘a kind of willow’ | | | **csak** ‘only, just’ | | |
| O.H. *csigle* | [ t͡ʃ i g l **ɛ** ] | | O.H. *csak* | [ t͡ʃ a k ] | |
| WOT. *çiglä** | [ t͡ʃ i g l **æ** ] | | WOT. *çak* | [ t͡ʃ a k ] | |
| Cost 0.33 | N-align 5 | LD% 6.6 | Cost 0 | N-align 3 | LD% 0 |
| **csákány** ‘pick-axe, war hammer’ | | | **csalán** ‘nettle’ | | |
| O.H. *csakan* | [ t͡ʃ a **k** a n ] | | O.H. *csalagan* | [ t͡ʃ a l **a** ɣ a n ] | |
| WOT. *çakkan* | [ t͡ʃ a **k:** a n ] | | WOT. *çalığan* | [ t͡ʃ a l **ï** ɣ a n ] | |
| Cost 0.33 | N-align 5 | LD% 6.6 | Cost 0.33 | N-align 7 | LD% 4.71 |
| **csalit** ‘thicket, brushwood’ | | | **csanak** ‘(silver) cup, bowl’ | | |
| O.H. *csalit* | [ t͡ʃ a l **i t** ] | | O.H. *csanak* | [ t͡ʃ a n a k ] | |
| WOT. *çaltı** | [ t͡ʃ a l **t ï** ] | | WOT. *çanak* | [ t͡ʃ a n a k ] | |
| Cost 1.33 | N-align 5 | LD% 26.6 | Cost 0 | N-align 5 | LD% 0 |
| **csata** ‘fighting, troop, battle, quarrel’ | | | **csat(ol)** ‘to add, to join, to buckle up’ | | |
| O.H. *csata* | [ t͡ʃ a t a ] | | O.H. *csati* | [ t͡ʃ a t **i** ] | |
| WOT. *çata** | [ t͡ʃ a t a ] | | WOT. *çatı* | [ t͡ʃ a t **ï** ] | |
| Cost 0 | N-align 4 | LD% 0 | Cost 0.33 | N-align 4 | LD% 8.25 |
| **csécs** ‘measle, smallpox, pocks’ | | | **csek(él)** ‘to bind (something)’ | | |
| O.H. *csecsek* | [ t͡ʃ e t͡ʃ **ɛ** k ] | | O.H. *csik* | [ t͡ʃ i k ] | |

| | | | | | |
|---|---|---|---|---|---|
| WOT. *çeçäk* | [ t͡ʃ e t͡ʃ **æ** k ] | | WOT. *çik** | [ t͡ʃ i k ] | |
| Cost 0.33 | N-align 5 | LD% 6.6 | Cost 0 | N-align 5 | LD% 0 |
| **csöpű** ‘impurities, tow, chaff’ | | | **csiga** ‘snail’ | | |
| O.H. *csöpüg* | [ t͡ʃ ø p y ɣ ] | | O.H. *csiga* | [ t͡ʃ ï g a ] | |
| WOT. *çöpüğ* | [ t͡ʃ ø p y ɣ ] | | WOT. *çıga** | [ t͡ʃ ï g a ] | |
| Cost 0 | N-align 5 | LD% 0 | Cost 0 | N-align 4 | LD% 0 |
| **csiger** ‘a wine of low quality, a fruit one’ | | | **csipa** ‘secretion, mucus discharged from the eyes’ | | |
| O.H. *csigir* | [ t͡ʃ **i** g **i** r ] | | O.H. *csapa* | [ t͡ʃ a p a **-** ] | |
| WOT. *çıgır** | [ t͡ʃ **ï** g **ï** r ] | | WOT. *çapağ* | [ t͡ʃ a p a **ɣ** ] | |
| Cost 0.66 | N-align 5 | LD% 13.2 | Cost 0.5 | N-align 5 | LD% 10 |
| **csomak** ‘a carpenter’s axe with a long helve’ | | | **csök** ‘sexual organ of large male animals’ | | |
| O.H. *csomak* | [ t͡ʃ o m a k ] | | O.H. *csük* | [ t͡ʃ y k ] | |
| WOT. *çomak* | [ t͡ʃ o m a k ] | | WOT. *çük* | [ t͡ʃ y k ] | |
| Cost 0 | N-align 5 | LD% 0 | Cost 0 | N-align 3 | LD% 0 |
| **csökik** ‘to become smaller, remain small’ | | | **csökönyös)** ‘stubborn’ | | |
| O.H. *csük* | [ t͡ʃ **y** k ] | | O.H. *csükün* | [ t͡ʃ **y** k **y** n ] | |
| WOT. *çök* | [ t͡ʃ **ø** k ] | | WOT. *çikin** | [ t͡ʃ **i** k **i** n ] | |
| Cost 0.33 | N-align 3 | LD% 11 | Cost 0.66 | N-align 5 | LD% 13.2 |
| **csősz** ‘field-guard’ | | | **csúnya** ‘rough, coarse, ugly’ | | |
| O.H. *csaüs* | [ t͡ʃ æ **-** **y** ʃ ] | | O.H. *csunag* | [ t͡ʃ u n a ɣ] | |
| WOT. *çäwiş* | [ t͡ʃ æ **w** **i** ʃ ] | | WOT. *çunağ* | [ t͡ʃ u n a ɣ ] | |
| Cost 0.83 | N-align 5 | LD% 16.6 | Cost 0 | N-align 5 | LD% 0 |
| **csűr** ‘to wind, to turn, to distrort, misinterpret’ | | | **dara** ‘grist, groats, soft hail’ | | |
| O.H. *csevür* | [ t͡ʃ **ɛ** **v** y r ] | | O.H. *dara* | [ **d** a r **a** ] | |
| WOT. *çäwür* | [ t͡ʃ **æ** **w** y r ] | | WOT. *tarı* | [ **t** a r **ï** ] | |
| Cost 1.83 | N-align 5 | LD% 36.6 | Cost 0.83 | N-align 4 | LD% 20.75 |
| **dél** ‘noon, south’ | | | **dió** ‘nut, walnut’ | | |
| O.H. *töli* | [ t **ø** l i ] | | O.H. *jigag* | [ j **i** ɣ a ɣ ] | |
| WOT. *tüli* | [ t **y** l i ] | | WOT. *yığağ* | [ j **ï** ɣ a ɣ ] | |

| | | | | | |
|---|---|---|---|---|---|
| Cost 0.33 | N-align 4 | LD% 8.25 | Cost 0.33 | N-align 5 | LD% 6.6 |
| **disznó** ‘pig’ | | | **dől** ‘to lean, topple over, stream down’ | | |
| O.H. *jisznag* | [ j **i** s n a ɣ ] | | O.H. *dül* | [ **d** y l ] | |
| WOT. *yasnağ* | [ j **a** s n a ɣ ] | | WOT. *tül* | [ **t** y l ] | |
| Cost 0.66 | N-align 6 | LD% 11 | Cost 0.5 | N-align 3 | LD% 16.67 |
| **dug** ‘to squeeze, to thrust into’ | | | **egy(ház)** ‘church (building and organization)’ | | |
| O.H. *dig* | [ d ï g ] | | O.H. *ed* | [ **e** d **-** ] | |
| WOT. *dıg** | [ d ï g ] | | WOT. *edü* | [ **ɛ** d **y** ] | |
| Cost 0 | N-align 3 | LD% 0 | Cost 1.33 | N-align 3 | LD% 44.33 |
| **ék** ‘wedge’ | | | **eke** ‘plough’ | | |
| O.H. *ék* | [ e: k ] | | O.H. *ekeg* | [ **ɛ** k **ɛ** ɣ ] | |
| WOT. *ék** | [ e: k ] | | WOT. *äkäğ* | [ **æ** k **æ** ɣ ] | |
| Cost 0 | N-align 2 | LD% 0 | Cost 0.66 | N-align 4 | LD% 16,5 |
| **eng(ed)** ‘to allow, permit, concede, yield, give way’ | | | **enő** ‘community work in agriculture’ | | |
| O.H. *eng* | [ **ɛ** ŋ **g** ] | | O.H. *ineg* | [ i n **ɛ** ɣ] | |
| WOT. *än* | [ **æ** ŋ **-** ] | | WOT. *inäğ* | [ i n **æ** ɣ ] | |
| Cost 1.33 | N-align 3 | LD% 44.33 | Cost 0.33 | N-align 4 | LD% 8,25 |
| **ér** ‘to arrive, to reach, get to’ | | | **érdem** ‘merit’ | | |
| O.H. *er* | [ ɛ r ] | | O.H. *erdem* | [ **ɛ** r d **ɛ** m ] | |
| WOT. *er*** | [ ɛ r ] | | WOT. *ärdäm* | [ **æ** r d **æ** m ] | |
| Cost 0 | N-align 2 | LD% 0 | Cost 0.66 | N-align 5 | LD% 13.2 |
| **erkölcs** ‘morals, morality’ | | | **erny(ed)** ‘to tire, lose vitality, slacken, relax’ | | |
| O.H. *erkelcs* | [ **ɛ** r k **ɛ** l t͡ʃ ] | | O.H. *erin* | [ **ɛ** r i n ] | |
| WOT. *ärkilç** | [ **æ** r k **i** l t͡ʃ ] | | WOT. *ärin* | [ **æ** r i n ] | |
| Cost 0.66 | N-align 6 | LD% 11 | Cost 0.33 | N-align 4 | LD% 8.25 |
| **erő** ‘strength, power’ | | | **eskü** ‘oath’ | | |
| O.H. *erig* | [ **ɛ** r i ɣ ] | | O.H. *icskü* | [ i t͡ʃ k y] | |
| WOT. *äriğ* | [ **æ** r i ɣ ] | | WOT. *içkü** | [ i t͡ʃ k y ] | |
| Cost 0.33 | N-align 4 | LD% 8,25 | Cost 0 | N-align 4 | LD% 0 |

| **ész** ‘reason, mind’ | | | **etel** ‘river, name of a river’ | | |
|---|---|---|---|---|---|
| O.H. *esz* | [ ɛ s ] | | O.H. *etil* | [ **ɛ** t i l ] | |
| WOT. *es* | [ ɛ s ] | | WOT. *ätil* | [ **æ** t i l ] | |
| Cost 0 | N-align 2 | LD% 0 | Cost 0.33 | N-align 4 | LD% 8.25 |
| **gaz** ‘weed’ | | | **görény** ‘polecat’ | | |
| O.H. *kagzu* | [ k a ɣ **z** u] | | O.H. *güren* | [ g y r **ɛ** n ] | |
| WOT. *kağdu** | [ k a ɣ **ð** u ] | | WOT. *gürän* | [ g y r **æ** n ] | |
| Cost 0.16 | N-align 5 | LD% 3.2 | Cost 0.33 | N-align 5 | LD% 6.6 |
| **gödény** ‘pelican’ | | | **güzü** ‘gleaner mouse’ | | |
| O.H. *güden* | [ g y d **ɛ** n ] | | O.H. *küsegü* | [ k y s **ɛ** ɣ y ] | |
| WOT. *güdän** | [ g y d **æ** n ] | | WOT. *küsäğü** | [ k y s **æ** ɣ y ] | |
| Cost 0.33 | N-align 5 | LD% 6.6 | Cost 0.33 | N-align 6 | LD% 5.5 |
| **gyaláz** ‘to abuse, calumniate’ | | | **gyalom** ‘drag-net, a kind of fishing net’ | | |
| O.H. *jala* | [ j a l a ] | | O.H. *jalum* | [ j a l **u** m ] | |
| WOT. *yala* | [ j a l a ] | | WOT. *yalım* | [ j a l **ï** m ] | |
| Cost 0 | N-align 4 | LD% 0 | Cost 0.33 | N-align 5 | LD% 6.6 |
| **gyanú** ‘suspicion, mistrust’ | | | **gyapjú** ‘wool’ | | |
| O.H. *jonag* | [ j o n a ɣ ] | | O.H. *japgu* | [ j a p **-** ɣ u ] | |
| WOT. *yonağ* | [ j o n a ɣ ] | | WOT. *yapağu* | [ j a p **a** ɣ u ] | |
| Cost 0 | N-align 5 | LD% 0 | Cost 1 | N-align 6 | LD% 16.67 |
| **gyapot** ‘cotton’ | | | **gyarap(od)(ik)** ‘to increase, to put on weight’ | | |
| O.H. *japut* | [ j a p u t ] | | O.H. *jarpa* | [ j a r p a ] | |
| WOT. *yaput* | [ j a p u t ] | | WOT. *yarpa* | [ j a r p a ] | |
| Cost 0 | N-align 5 | LD% 0 | Cost 0 | N-align 5 | LD% 0 |
| **gyarló** ‘poor, frail, feeble’ | | | **gyárt** ‘to produce, build, fabricate’ | | |
| O.H. *jarlig* | [ j a r l ï ɣ ] | | O.H. jarat | [ j a r a t ] | |
| WOT. *yarlığ* | [ j a r l ï ɣ ] | | WOT. *yarat* | [ j a r a t ] | |
| Cost 0 | N-align 6 | LD% 0 | Cost 0 | N-align 5 | LD% 0 |
| **gyász** ‘mourning, bereavement’ | | | **gyékény** ‘bulrush’ | | |
| O.H. *jasz* | [ j a s ] | | O.H. *jeken* | [ j ɛ k **ɛ** n ] | |

| | | | | | |
|---|---|---|---|---|---|
| WOT. *yas* | [ j a s ] | | WOT. *yekän* | [ j ɛ k **æ** n ] | |
| Cost 0 | N-align 3 | LD% 0 | Cost 0.33 | N-align 5 | LD% 6.6 |
| **gyenge** ‘weak, feeble’ | | | **gyeplő** ‘rein’ | | |
| O.H. *jenge* | [ j ɛ ŋ **g ɛ** ] | | O.H. *jiplig* | [ j i p l i ɣ ] | |
| WOT. *yeni** | [ j ɛ ŋ **-** **i** ] | | WOT. *yipliğ* | [ j i p l i ɣ ] | |
| Cost 1.33 | N-align 5 | LD% 26.6 | Cost 0 | N-align 6 | LD% 0 |
| **gyepű** ‘borderland, hedge’ | | | **gyermek** ‘child’ | | |
| O.H. *jepig* | [ j **ɛ** p i ɣ ] | | O.H. *järmäk* | [ j æ r m æ k ] | |
| WOT. *yäpiğ** | [ j **æ** p i ɣ ] | | WOT. *yärmäk** | [ j æ r m æ k ] | |
| Cost 0.33 | N-align 5 | LD% 6.6 | Cost 0 | N-align 6 | LD% 0 |
| **gyertya** ‘candle’ | | | **gyertyán** ‘hornbeam’ | | |
| O.H. *jarta* | [ j a r t a ] | | O.H. *jartagan* | [ j a r t a ɣ a n ] | |
| WOT. *yarta* | [ j a r t a ] | | WOT. *yartağan* | [ j a r t a ɣ a n ] | |
| Cost 0 | N-align 5 | LD% 0 | Cost 0 | N-align 8 | LD% 0 |
| **gyilkos** ‘murderer, killer’ | | | **gyom** ‘weed’ | | |
| O.H. *jilkus* | [ j **i** l k u **ʃ -** ] | | O.H. *jom* | [ j o m ] | |
| WOT. *yulkuçı** | [ j **u** l k u **t͡ʃ ï** ] | | WOT. *yom* | [ j o m ] | |
| Cost 2.16 | N-align 7 | LD% 30.86 | Cost 0 | N-align 3 | LD% 0 |
| **gyomor** ‘stomach’ | | | **gyón(ik)** ‘to confess (sins)’ | | |
| O.H. *jumur* | [ j u m u r ] | | O.H. *juvun* | [ j **u** **v u** n ] | |
| WOT. *yumur* | [ j u m u r ] | | WOT. *yun* | [ j **u:** **-** **-** n ] | |
| Cost 0 | N-align 5 | LD% 0 | Cost 2.33 | N-align 5 | LD% 46.6 |
| **gyopár** ‘cudweed’ | | | **gyöngy** ‘pearl’ | | |
| O.H. *jipar* | [ j ï p a r ] | | O.H. *jinjü* | [ j i n j y ] | |
| WOT. *yıpar* | [ j ï p a r ] | | WOT. *yinyü* | [ j i n j y ] | |
| Cost 0 | N-align 5 | LD% 0 | Cost 0 | N-align 5 | LD% 0 |
| **gyötör** ‘to torture, to make suffer’ | | | **győz** ‘to conquer, triumph, gain victory’ | | |
| O.H. *jitür* | [ j i t y r ] | | O.H. *jegüz* | [ j **ɛ** ɣ y z ] | |
| WOT. *yitür* | [ j i t y r ] | | WOT. *yäğüz** | [ j **æ** ɣ y z ] | |
| Cost 0 | N-align 5 | LD% 0 | Cost 0.33 | N-align 5 | LD% 6.6 |

| **gyúl** ‘to catch fire, be kindled, to be ignited’ | | | **gyúr** ‘to knead, pug’ | | |
|---|---|---|---|---|---|
| O.H. *jul* | [ j u l ] | | O.H. *jugur* | [ j u ɣ u r ] | |
| WOT. *yul** | [ j u l ] | | WOT. *yuğur* | [ j u ɣ u r ] | |
| Cost 0 | N-align 3 | LD% 0 | Cost 0 | N-align 5 | LD% 0 |
| **gyűl(ik)** ‘assemble, gather’ | | | **gyűlöl** ‘to hate’ | | |
| O.H. *jigil* | [ j i **ɣ** i l ] | | O.H. *jegile* | [ j **ɛ** **ɣ** i l **ɛ** ] | |
| WOT. *yigil* | [ j i **g** i l ] | | WOT. *yägilä* | [ j **æ** **g** i l **æ** ] | |
| Cost 0.25 | N-align 5 | LD% 5 | Cost 0.91 | N-align 6 | LD% 15.17 |
| **gyümölcs** ‘fruit’ | | | **gyümölcsény** ‘a plant similar to the whitethorn’ | | |
| O.H. *jemilcs* | [ j e m i l t͡ʃ ] | | O.H. *jemilcsen* | [ j e m i l t͡ʃ e n ] | |
| WOT. *yemilç* | [ j e m i l t͡ʃ ] | | WOT. *yemilçen* | [ j e m i l t͡ʃ e n ] | |
| Cost 0 | N-align 6 | LD% 0 | Cost 0 | N-align 8 | LD% 0 |
| **gyűrű** ‘a kind of tree similar to the maple or cornel’ | | | **gyűrű** ‘ring’ | | |
| O.H. *jereg* | [ j e r e ɣ ] | | O.H. *jürüg* | [ j y r y ɣ ] | |
| WOT. *yereğ* | [ j e r e ɣ ] | | WOT. *yürüğ* | [ j y r y ɣ ] | |
| Cost 0 | N-align 5 | LD% 0 | Cost 0 | N-align 5 | LD% 0 |
| **gyűszű** ‘thimble’ | | | **hajó** ‘boat, ship’ | | |
| O.H. *jügszüg* | [ j y ɣ s y ɣ] | | O.H. *hajig* | [ **h** a j **i** ɣ ] | |
| WOT. *yüğsüğ* | [ j y ɣ s y ɣ ] | | WOT. *hayığ* | [ **χ** a j **ï** ɣ ] | |
| Cost 0 | N-align 6 | LD% 0 | Cost 1 | N-align 5 | LD% 20 |
| **harang** ‘bell’ | | | **herjó** ‘hawk’ | | |
| O.H. *hangar* | [ **h** **a** ŋ **g** a r ] | | O.H. *hirguj* | [ χ **i** r ɣ u j ] | |
| WOT. *honar* | [ **χ** **o** ŋ **-** a r ] | | WOT. *hırğuy* | [ χ **ï** r ɣ u j ] | |
| Cost 2.33 | N-align 6 | LD% 38.83 | Cost 0.33 | N-align 6 | LD% 5.5 |
| **homok** ‘sand’ | | | **hurok** ‘noose, loop, snare’ | | |
| O.H. *humoki* | [ **h** u m **o** k **i** ] | | O.H. *urok* | [ u r o k ] | |
| WOT. *humakı* | [ **χ** u m **a** k **ï** ] | | WOT. *urok* | [ u r o k ] | |
| Cost 1.33 | N-align 6 | LD% 22.17 | Cost 0 | N-align 4 | LD% 0 |
| **idő** ‘time, weather’ | | | **igen** ‘yes, affirmative particle, very’ | | |

| | | | | | |
|---|---|---|---|---|---|
| O.H. *üdeg* | [ y d **ɛ** ɣ ] | | O.H. *egen* | [ **ɛ - g ɛ** n ] | |
| WOT. *üdäğ* | [ y d **æ** ɣ ] | | WOT. *ärkän** | [**æ r k æ** n ] | |
| Cost 0.33 | N-align 4 | LD% 8.25 | Cost 2.33 | N-align 5 | LD% 46.6 |
| **ije(szt)** ‘to frighten’ | | | **iker** ‘twin’ | | |
| O.H. *äji* | [ æ j i ] | | O.H. *ikir* | [ i k i r ] | |
| WOT. *äyi** | [ æ j i ] | | WOT. *ikir* | [ i k i r ] | |
| Cost 0 | N-align 3 | LD% 0 | Cost 0 | N-align 4 | LD% 0 |
| **ildom** ‘proper behaviour’ | | | **ill(ik)** ‘to suit something, to be proper, to fit into’ | | |
| O.H. *ildam* | [ **i** l d a m ] | | O.H. *il* | [ i l ] | |
| WOT. *ıldam* | [ **ï** l d a m ] | | WOT. *il* | [ i l ] | |
| Cost 0.33 | N-align 5 | LD% 6.6 | Cost 0 | N-align 2 | LD% 0 |
| **imá(d)** ‘to adore, worship’ | | | **ing** ‘shirt’ | | |
| O.H. *vim* | [ v **i** m ] | | O.H. *ümmeg* | [ **y m** m **ɛ g**] | |
| WOT. *vım** | [ v **ï** m ] | | WOT. *önmäk** | [ **ø ŋ** m **æ k** ] | |
| Cost 0.33 | N-align 3 | LD% 11 | Cost 2 | N-align 5 | LD% 40 |
| **ír** ‘to write’ | | | **író** ‘buttermilk’ | | |
| O.H. *ir* | [ **i** r ] | | O.H. *irag* | [ **i** r a ɣ ] | |
| WOT. *ır* | [ **ï** r ] | | WOT. *ırağ* | [ **ï** r a ɣ ] | |
| Cost 0.33 | N-align 2 | LD% 16.5 | Cost 0.33 | N-align 4 | LD% 8.25 |
| **izz(ik)** ‘to glow, be hot’ | | | **jász** ‘name of an ethnic group with Iranian origins in Hungary’ | | |
| O.H. *iszi* | [ **i** s **i** ] | | O.H. *jász* | [ j a: s ] | |
| WOT. *ısı** | [ **ï** s **ï** ] | | WOT. *yas* | [ j a: s ] | |
| Cost 0.66 | N-align 3 | LD% 22 | Cost 0 | N-align 3 | LD% 0 |
| **káka** ‘bulrush, club-rush’ | | | **káliz** ‘muslim ethnic group in medieval Hungary’ | | |
| O.H. *káka* | [ k **a:** k **a** ] | | O.H. *kaliz* | [ k **–** a l i z ] | |
| WOT. *kakı* | [ k **a** k **ï** ] | | WOT. *kvaliz* | [ k **v** a l i z ] | |
| Cost 0.66 | N-align 4 | LD% 16.5 | Cost 1 | N-align 6 | LD% 16.67 |
| **kalokány** ‘a water plant’ | | | **kantár** ‘bridle, reins’ | | |
| O.H. *karlukan* | [ k a r l u k a n ] | | O.H. *kantár* | [ k **ɒ** n t **a:** r ] | |

| | | | | | |
|---|---|---|---|---|---|
| WOT. *karlukan* | [ k a r l u k a n ] | | WOT. *kantar* | [ k **a** n t **a** r ] | |
| Cost 0 | N-align 8 | LD% 0 | Cost 0.66 | N-align 6 | LD% 11 |
| **kanyaró** ‘measles’ | | | **kaptány** ‘trap, snare’ | | |
| O.H. *karamug* | [ k a r a m u ɣ ] | | O.H. *kapkan* | [ k a p k a n ] | |
| WOT. *karamuğ* | [ k a r a m u ɣ ] | | WOT. *kapkan* | [ k a p k a n ] | |
| Cost 0 | N-align 7 | LD% 0 | Cost 0 | N-align 6 | LD% 0 |
| **kapu** ‘gate’ | | | **kar** ‘arm’ | | |
| O.H. *kapug* | [ k a p u ɣ ] | | O.H. *kar* | [ k a r ] | |
| WOT. *kapuğ* | [ k a p u ɣ ] | | WOT. *kar* | [ k a r ] | |
| Cost 0 | N-align 5 | LD% 0 | Cost 0 | N-align 3 | LD% 0 |
| **kár** ‘damage, loss’ | | | **karakán** ‘man of grit, stout fellow’ | | |
| O.H. *kor* | [ k **o** r ] | | O.H. *karakan* | [ k a r a k a n ] | |
| WOT. *kor* | [ k **o:** r ] | | WOT. *karakan* | [ k a r a k a n ] | |
| Cost 0.33 | N-align 3 | LD% 11 | Cost 0 | N-align 7 | LD% 0 |
| **karám** ‘sheepfold, cattlegrid, stockyard’ | | | **karó** ‘stake, pale, stick’ | | |
| O.H. *korám* | [ k o r **a:** m ] | | O.H. *karog* | [ k a r o ɣ ] | |
| WOT. *koram* | [ k o r **a** m ] | | WOT. *karoğ* | [ k a r o ɣ ] | |
| Cost 0.33 | N-align 5 | LD% 6.6 | Cost 0 | N-align 5 | LD% 0 |
| **kárókatona** ‘a fishing water bird with black feathers’ | | | **karvaly** ‘sparrow hawk’ | | |
| O.H. *kara kotan* | [ k a r a k o t a n ] | | O.H. *kargaj* | [ k a r **ɣ** a j ] | |
| WOT. *kara* kotan* | [ k a r a k o t a n ] | | WOT. *kargay** | [ k a r **g** a j ] | |
| Cost 0 | N-align 9 | LD% 0 | Cost 0.25 | N-align 6 | LD% 4,17 |
| **katák** ‘chicory’ | | | **kazár** ‘an ethnic name, Kazar’ | | |
| O.H. *katák* | [ k a t **a:** ŋ **g** ] | | O.H. *kazar* | [ k a z a r ] | |
| WOT. *katan* | [ k a t **a** ŋ **-** ] | | WOT. *kazar* | [ k a z a r ] | |
| Cost 1.33 | N-align 6 | LD% 22.17 | Cost 0 | N-align 5 | LD% 0 |
| **kebel** ‘bosom, breast’ | | | **kecske** ‘goat’ | | |
| O.H. *kebel* | [ k **ɛ** b **ɛ** l ] | | O.H. *kecseke* | [ k **ɛ** t͡ʃ **ɛ** k **ɛ** ] | |
| WOT. *käbäl** | [ k **æ** b **æ** l ] | | WOT. *käçäkä* | [ k **æ** t͡ʃ **æ** k **æ** ] | |
| Cost 0.66 | N-align 5 | LD% 13.2 | Cost 1 | N-align 6 | LD% 16.67 |

| **kék** ‘blue’ | | | **kelengye** ‘trousseau, dowry, gift given with the bride’ | | |
|---|---|---|---|---|---|
| O.H. *keyk* | [ k **ɛ** **j** k ] | | O.H. *kälinçäğ* | [ k æ l i n t͡ʃ æ ɣ ] | |
| WOT. *kök* | [ k **ø:** **-** k ] | | WOT. *kälinçäğ* | [ k æ l i n t͡ʃ æ ɣ ] | |
| Cost 1.5 | N-align 4 | LD% 37.5 | Cost 0 | N-align 8 | LD% 0 |
| **kender** ‘hemp’ | | | **kéneső** ‘mercury’ | | |
| O.H. *kendir* | [ k **ɛ** n d i r ] | | O.H. *kenesű* | [ k ɛ n **ɛ** ʃ **y:** ] | |
| WOT. *kändir* | [ k **æ** n d i r ] | | WOT. *kenä şu* | [ k ɛ n **æ** ʃ **u:** ] | |
| Cost 0.33 | N-align 6 | LD% 5.5 | Cost 0.66 | N-align 6 | LD% 11 |
| **kép** ‘picture, shape, form’ | | | **kepe** ‘shock, shook, sheaves placed crosswise’ | | |
| O.H. *kép* | [ k **e:** p ] | | O.H. *kepeg* | [ k **ɛ** p **ɛ** ɣ ] | |
| WOT. *käp* | [ k **æ:** p ] | | WOT. *käpäğ** | [ k **æ** p **æ** ɣ ] | |
| Cost 0.33 | N-align 3 | LD% 11 | Cost 0.66 | N-align 5 | LD% 13.2 |
| **kerep** ‘a kind of ship’ | | | **kér(ődz)(ik)** ‘to ruminate (of bovines)’ | | |
| O.H. *kerep* | [ k **ɛ** r **ɛ** p ] | | O.H. *kér* | [ k **e:** **-** **-** r ] | |
| WOT. *käräp* | [ k **æ** r **æ** p ] | | WOT. *käwir* | [ k **æ** **w** **i** r ] | |
| Cost 0.66 | N-align 5 | LD% 13.2 | Cost 2.16 | N-align 5 | LD% 43.2 |
| **kert** ‘garden, place for animals’ | | | **keselyű** ‘vulture’ | | |
| O.H. *kerte* | [ k **ɛ** r t **ɛ** ] | | O.H. *kücseleg* | [ k y t͡ʃ **ɛ** l **ɛ** ɣ ] | |
| WOT. *kärtä** | [ k **æ** r t **æ** ] | | WOT. *küçäläğ** | [ k y t͡ʃ **æ** l **æ** ɣ ] | |
| Cost 0.66 | N-align 5 | LD% 13.2 | Cost 0.66 | N-align 7 | LD% 9.43 |
| **kés(ik)** ‘to be, become late’ | | | **kicsiny** ‘small, little’ | | |
| O.H. *kécs* | [ k e: t͡ʃ ] | | O.H. *kücsün* | [ k **y** t͡ʃ **y** n ] | |
| WOT. *keç* | [ k e: t͡ʃ ] | | WOT. *kiçin* | [ k **i** t͡ʃ **i** n ] | |
| Cost 0 | N-align 3 | LD% 0 | Cost 0.66 | N-align 5 | LD% 13.2 |
| **kikerics** ‘wild/meadow saffron, lion’s tooth’ | | | **kiköcsén** ‘a bird similar to a falcon’ | | |
| O.H. *kükerilcs* | [ k **y** k **ɛ** r i l t͡ʃ ] | | O.H. *kükercsen* | [ k **y** k **ɛ** r t͡ʃ **ɛ** n ] | |
| WOT. *kökärilç* | [ k **ø** k **æ** r i l t͡ʃ ] | | WOT. *kökärçän* | [ k **ø** k **æ** r t͡ʃ **æ** n ] | |
| Cost 0.66 | N-align 8 | LD% 8,25 | Cost 1 | N-align 8 | LD% 12.5 |
| **kín** ‘pain’ | | | **koldul** ‘to beg, to mendicate’ | | |

| | | | | | |
|---|---|---|---|---|---|
| O.H. *kín* | [ k i: n ] | | O.H. *koldu* | [ k o l d u ] | |
| WOT. *kin* | [ k i: n ] | | WOT. *koldu* | [ k o l d u ] | |
| Cost 0 | N-align 3 | LD% 0 | Cost 0 | N-align 5 | LD% 0 |
| **komló** ‘hops, humulus lupulus’ | | | **komócsin** ‘timothy-grass, phleum’ | | |
| O.H. *kumlag* | [ k u m l a ɣ ] | | O.H. *kamicsin* | [ k **ɒ** m **i** t͡ʃ **i** n ] | |
| WOT. *kumlağ* | [ k u m l a ɣ ] | | WOT. *kamıçın* | [ k **a** m **ï** t͡ʃ **ï** n ] | |
| Cost 0 | N-align 6 | LD% 0 | Cost 1 | N-align 7 | LD% 14.29 |
| **komor** ‘gloomy, grave, somber, dull-coloured’ | | | **koporsó** ‘coffin’ | | |
| O.H. *komur* | [ k o m u r ] | | O.H. *kopurcsag* | [ k **o** p u r t͡ʃ a ɣ ] | |
| WOT. *komur* | [ k o m u r ] | | WOT. *kapurçağ* | [ k **a** p u r t͡ʃ a ɣ ] | |
| Cost 0 | N-align 5 | LD% 0 | Cost 0.66 | N-align 8 | LD% 8.25 |
| **kor** ‘age, period’ | | | **korhány** ‘tomb, small hill, dry place in a marsh’ | | |
| O.H. *kur* | [ k u r ] | | O.H. *korgan* | [ k o r ɣ a n ] | |
| WOT. *kur* | [ k u r ] | | WOT. *korğan* | [ k o r ɣ a n ] | |
| Cost 0 | N-align 3 | LD% 0 | Cost 0 | N-align 6 | LD% 0 |
| **kóró** ‘dry stalk of weed’ | | | **korom** ‘soot’ | | |
| O.H. *kuroğ* | [ k **u -** r o **ɣ** ] | | O.H. *kurum* | [ k u r u m ] | |
| WOT. *kowrog** | [ k **o w** r o **g** ] | | WOT. *kurum* | [ k u r u m ] | |
| Cost 1.08 | N-align 6 | LD% 18 | Cost 0 | N-align 5 | LD% 0 |
| **kos** ‘ram’ | | | **kozma** ‘burn’ | | |
| O.H. *kocs* | [ k o t͡ʃ ] | | O.H. *kaszmag* | [ k a s m a ɣ ] | |
| WOT. *koç* | [ k o t͡ʃ ] | | WOT. *kasmağ* | [ k a s m a ɣ ] | |
| Cost 0 | N-align 3 | LD% 0 | Cost 0 | N-align 6 | LD% 0 |
| **ködmön** ‘frock, sheepskin waistcoat’ | | | **kökény** ‘blackthorn’ | | |
| O.H. *kedmen* | [ k **ɛ** d m **ɛ** n ] | | O.H. *köken* | [ k ø k **ɛ** n ] | |
| WOT. *kädmän* | [ k **æ** d m **æ** n ] | | WOT. *kökän* | [ k ø k **æ** n ] | |
| Cost 0.66 | N-align 6 | LD% 11 | Cost 0.33 | N-align 5 | LD% 6.6 |
| **kölcsön** ‘loan’ | | | **köldök** ‘navel, parts of agricultural instruments’ | | |
| O.H. *külcsen* | [ k **y** l t͡ʃ ɛ n ] | | O.H. *kindik* | [ k i n d i k ] | |

| | | | | | |
|---|---|---|---|---|---|
| WOT. *kölçen* | [ k **ø** l t͡ʃ ɛ n ] | | WOT. *kindik* | [ k i n d i k ] | |
| Cost 0.33 | N-align 6 | LD% 5.5 | Cost 0 | N-align 6 | LD% 0 |
| **kölyök** ‘young of an animal, kid, puppy, lad’ | | | **kölyű** ‘pounder, beater, small mortar’ | | |
| O.H. *kylek* | [ k **y** l **ɛ** k ] | | O.H. *kelig* | [ k ɛ l i ɣ ] | |
| WOT. *köläk* | [ k **ø** l **æ** k ] | | WOT. *keliğ* | [ k ɛ l i ɣ ] | |
| Cost 0.66 | N-align 5 | LD% 13.2 | Cost 0 | N-align 5 | LD% 0 |
| **könvy** ‘book’ | | | **köpcös** ‘stumpy, a small and thick man’ | | |
| O.H. *künüg* | [ k y ɲ y ɣ ] | | O.H. *köpcseg* | [ k ø p t͡ʃ **ɛ** ɣ ] | |
| WOT. *künüğ** | [ k y ɲ y ɣ ] | | WOT. *köpçäğ** | [ k ø p t͡ʃ **æ** ɣ ] | |
| Cost 0 | N-align 5 | LD% 0 | Cost 0.33 | N-align 6 | LD% 5.5 |
| **köpec** ‘a kind of Hungarian sheepdog’ | | | **köpönyeg** ‘cloak, gown, overcoat’ | | |
| O.H. *köpek* | [ k ø p **ɛ** k ] | | O.H. *kepenek* | [ k ɛ p ɛ n ɛ k ] | |
| WOT. *köpäk** | [ k ø p **æ** k ] | | WOT. *kepenek* | [ k ɛ p ɛ n ɛ k ] | |
| Cost 0.33 | N-align 5 | LD% 6.6 | Cost 0 | N-align 7 | LD% 0 |
| **köpű** ‘churn, beehive’ | | | **kőris** ‘ash tree’ | | |
| O.H. *küpüg* | [ k y p y ɣ ] | | O.H. *kewrics* | [ k ɛ **w** r i t͡ʃ ] | |
| WOT. *küpüğ* | [ k y p y ɣ ] | | WOT. *kevriç* | [ k ɛ **v** r i t͡ʃ ] | |
| Cost 0 | N-align 5 | LD% 0 | Cost 1.5 | N-align 6 | LD% 25 |
| **kőrő** ‘mellow, powdery, crumbly, rotten, mouldy’ | | | **körtvély** ‘pear’ | | |
| O.H. *kevreg* | [ k **ɛ** v r **ɛ** ɣ ] | | O.H. *kertmelig* | [ k ɛ r t m ɛ l i ɣ ] | |
| WOT. *kävräğ* | [ k **æ** v r **æ** ɣ ] | | WOT. *kertmeliğ* | [ k ɛ r t m ɛ l i ɣ ] | |
| Cost 0.66 | N-align 6 | LD% 11 | Cost 0 | N-align 9 | LD% 0 |
| **köszön** ‘to greet, to thank’ | | | **kun** ‘Cuman ethnic group in Hungary’ | | |
| O.H. *küszen* | [ k y s **ɛ** n ] | | O.H. *kuwan* | [ k u w a n ] | |
| WOT. *küsän* | [ k y s **æ** n ] | | WOT. *kuwan* | [ k u w a n ] | |
| Cost 0.33 | N-align 5 | LD% 6.6 | Cost 0 | N-align 5 | LD% 0 |
| **küllő** ‘laughing bird’ | | | **küllő** ‘a fish living at the bottom of the river’ | | |
| O.H. *külleg* | [ k y l **-** l **ɛ** ɣ ] | | O.H. *köligelig* | [ k ø l i g **ɛ** l i ɣ ] | |
| WOT. *külülägğ** | [ k y l **y** l **æ** ɣ ] | | WOT. *köligäliğ** | [ k ø l i g **æ** l i ɣ ] | |
| Cost 1.33 | N-align 7 | LD% 19 | Cost 0.33 | N-align 9 | LD% 3.67 |

| **küllő** ‘spoke (of a wheel)’ | | | **ocsú** ‘chaff, husks, winnowing, tailings’ | | |
|---|---|---|---|---|---|
| O.H. *küvey* | [ k y v **ɛ** j ] | | O.H. *ucsog* | [ u t͡ʃ o ɣ ] | |
| WOT. *küväy** | [ k y v **æ** j ] | | WOT. *uçoğ* | [ u t͡ʃ o ɣ ] | |
| Cost 0.33 | N-align 5 | LD% 6.6 | Cost 0 | N-align 4 | LD% 0 |
| **ok** ‘cause, reason’ | | | **ól** ‘sty, cattle pen, sheepfold’ | | |
| O.H. *uk* | [ u k ] | | O.H. *agul* | [ a ɣ u l ] | |
| WOT. *uk* | [ u k ] | | WOT. *ağul* | [ a ɣ u l ] | |
| Cost 0 | N-align 2 | LD% 0 | Cost 0 | N-align 4 | LD% 0 |
| **olcsó** ‘cheap’ | | | **olló** ‘kid (of a goat)’ | | |
| O.H. *ucsag* | [ u t͡ʃ a ɣ ] | | O.H. *oglag* | [ o ɣ l a ɣ ] | |
| WOT. *uçağ* | [ u t͡ʃ a ɣ ] | | WOT. *oğlağ* | [ o ɣ l a ɣ ] | |
| Cost 0 | N-align 4 | LD% 0 | Cost 0 | N-align 5 | LD% 0 |
| **ondó** ‘sperm, seminal fluid’ | | | **Orosz** ‘Russian’ | | |
| O.H. *undag* | [ u n d a ɣ ] | | O.H. *urusz* | [ u r u s ] | |
| WOT. *undağ* | [ u n d a ɣ ] | | WOT. *urus* | [ u r u s ] | |
| Cost 0 | N-align 5 | LD% 0 | Cost 0 | N-align 4 | LD% 0 |
| **oroszlán** ‘lion’ | | | **orsó** ‘spindle, whorl’ | | |
| O.H. *uruszlán* | [ **u** r u s l **a:** n ] | | O.H. *urcsug* | [ u r t͡ʃ u ɣ ] | |
| WOT. *aruslan* | [ **a** r u s l **a** n ] | | WOT. *urçuğ* | [ u r t͡ʃ u ɣ ] | |
| Cost 1 | N-align 7 | LD% 14.29 | Cost 0 | N-align 5 | LD% 0 |
| **orv** ‘thief’ | | | **orvos** ‘physician’ | | |
| O.H. *ugru* | [ **u** **ɣ** r u ] | | O.H. *urvocsi* | [ **u** r **v** **o** t͡ʃ ï ] | |
| WOT. *ogru* | [ **o** **g** r u ] | | WOT. *orwuçı** | [ **o** r **w** **u** t͡ʃ ï ] | |
| Cost 0.58 | N-align 4 | LD% 14.5 | Cost 3.33 | N-align 6 | LD% 36 |
| **os** ‘backgammon (ostábla) dice/cube’ | | | **ökör** ‘ox’ | | |
| O.H. *alcsuk* | [ a l t͡ʃ u k ] | | O.H. *ükür* | [ **y** k y r ] | |
| WOT. *alçuk* | [ a l t͡ʃ u k ] | | WOT. *ökür* | [ **ø** k y r ] | |
| Cost 0 | N-align 5 | LD% 0 | Cost 0.33 | N-align 4 | LD% 8.25 |
| **ölt** ‘to stitch, to put on a dress’ | | | **ön(ik)** ‘to elect, to select’ | | |
| O.H. *ilt* | [ i l t ] | | O.H. *ün* | [ **y** n ] | |

| | | | | | |
|---|---|---|---|---|---|
| WOT. *ilt* | [ i l t ] | | WOT. *ün* | [ **y:** n ] | |
| Cost 0 | N-align 3 | LD% 0 | Cost 0.33 | N-align 2 | LD% 16.5 |
| **őrül** ‘to become mad’ | | | **ölyv** ‘hawk, buzzard’ | | |
| O.H. *ävril* | [ æ v r i l ] | | O.H. elig | [ ɛ l i ɣ ] | |
| WOT. *ävril* | [ æ v r i l ] | | WOT. eliğ | [ ɛ l i ɣ ] | |
| Cost 0 | N-align 5 | LD% 0 | Cost 0 | N-align 4 | LD% 0 |
| **őr(öl)** ‘to grind’ | | | **örmény**: ‘Armenian’ | | |
| O.H. *evir* | [ **ɛ** v i r ] | | O.H. *ermeni* | [ **ɛ** r m **ɛ** n i ] | |
| WOT. *ävir* | [ **æ** v i r ] | | WOT. *ärmäni* | [ **æ** r m **æ** n i ] | |
| Cost 0.33 | N-align 4 | LD% 8.25 | Cost 0.66 | N-align 6 | LD% 11 |
| **örök** ‘eternal’ | | | **örül** ‘to rejoice, to be glad’ | | |
| O.H. *ürük* | [ y r y k ] | | O.H. *ögir* | [ ø ɣ i r ] | |
| WOT. *ürük* | [ y r y k ] | | WOT. *öğir* | [ ø ɣ i r ] | |
| Cost 0 | N-align 4 | LD% 0 | Cost 0 | N-align 4 | LD% 0 |
| **ör(vény)** ‘whirlpool’ | | | **sajt** ‘cheese’ | | |
| O.H. *egir* | [ **ɛ** ɣ i r ] | | O.H. *csagt* | [ t͡ʃ **a** ɣ t ] | |
| WOT. *äğir* | [ **æ** ɣ i r ] | | WOT. *çığt*** | [ t͡ʃ **ï** ɣ t ] | |
| Cost 0.33 | N-align 4 | LD% 8.25 | Cost 0.33 | N-align 4 | LD% 8.25 |
| **sár** ‘mud, marsh’ | | | **sárga** ‘yellow’ | | |
| O.H. *sar* | [ ʃ a r ] | | O.H. *sarug* | [ ʃ a r u g ] | |
| WOT. *şar* | [ ʃ a r ] | | WOT. *şarug* | [ ʃ a r u g ] | |
| Cost 0 | N-align 3 | LD% 0 | Cost 0 | N-align 5 | LD% 0 |
| **sárkány** ‘dragon’ | | | **sarló** ‘sickle’ | | |
| O.H. *sárkán* | [ ʃ **a:** r **-** k **a:** n ] | | O.H. *csarlag* | [ t͡ʃ a r l a ɣ ] | |
| WOT. *şarakan* | [ ʃ **a** r **a** k **a** n ] | | WOT. *çarlağ* | [ t͡ʃ a r l a ɣ ] | |
| Cost 1.66 | N-align 7 | LD% 23.71 | Cost 0 | N-align 6 | LD% 0 |
| **sárma** ‘a type of wild onion’ | | | **saru** ‘footwear, sandals’ | | |
| O.H. *sarmag* | [ ʃ a r m a ɣ ] | | O.H. *csarug* | [ t͡ʃ a r u ɣ ] | |
| WOT. *şarmağ** | [ ʃ a r m a ɣ ] | | WOT. *çarığ* | [ t͡ʃ a r u ɣ ] | |
| Cost 0 | N-align 6 | LD% 0 | Cost 0 | N-align 5 | LD% 0 |

| **sátor** ‘tent’ | | | **seb(es)** ‘fast, quick’ | | |
|---|---|---|---|---|---|
| O.H. *csatur* | [ t͡ʃ a t u r ] | | O.H. *seb* | [ ʃ **ɛ** **b** ] | |
| WOT. *çatur* | [ t͡ʃ a t u r ] | | WOT. *şip* | [ ʃ **i** **p** ] | |
| Cost 0 | N-align 5 | LD% 0 | Cost 0.83 | N-align 3 | LD% 27.67 |
| | | | **söpör** ‘to sweep, to broom’ | | |
| | | | O.H. *sipir* | [ ʃ i p i r ] | |
| | | | WOT. *şipir* | [ ʃ i p i r ] | |
| | | | Cost 0 | N-align 5 | LD% 0 |
| **seprő** ‘broom’ | | | **seprő** ‘draff, lees, dregs of wine’ | | |
| O.H. *sipirig* | [ ʃ i p i r i ɣ ] | | O.H. *csöpreg* | [ t͡ʃ ø p r **ɛ** ɣ ] | |
| WOT. *şipiriğ** | [ ʃ i p i r i ɣ ] | | WOT. *çöpräğ* | [ t͡ʃ ø p r **æ** ɣ ] | |
| Cost 0 | N-align 7 | LD% 0 | Cost 0.33 | N-align 6 | LD% 5.5 |
| **sér(t)** ‘to injure, to hurt’ | | | **sereg** ‘army, troops, crowd’ | | |
| O.H. *sir* | [ ʃ i r ] | | O.H. *csarig* | [ t͡ʃ æ r i g ] | |
| WOT. *şir** | [ ʃ i r ] | | WOT. *çärig* | [ t͡ʃ æ r i g ] | |
| Cost 0 | N-align 3 | LD% 0 | Cost 0 | N-align 5 | LD% 0 |
| **serke** ‘a nit’ | | | **serleg** ‘cup, goblet’ | | |
| O.H. *sirke* | [ ʃ i r k **ɛ** ] | | O.H. *serleg* | [ ʃ **ɛ** r l **ɛ** g ] | |
| WOT. *şirkä* | [ ʃ i r k **æ** ] | | WOT. *şirlig* | [ ʃ **i** r l **i** g ] | |
| Cost 0.33 | N-align 5 | LD% 6.6 | Cost 0.66 | N-align 6 | LD% 11 |
| **serte** ‘bristle’ | | | **sík** ‘even, flat, open (water)’ | | |
| O.H. *sirt* | [ ʃ i r t ] | | O.H. *sik* | [ ʃ **i** k ] | |
| WOT. *şirt* | [ ʃ i r t ] | | WOT. *şık* | [ ʃ **ï** k ] | |
| Cost 0 | N-align 4 | LD% 0 | Cost 0.33 | N-align 3 | LD% 11 |
| **sima** ‘smooth’ | | | **sió** ‘running water, running channel under a mill’ | | |
| O.H. *sima* | [ ʃ **i** m a ] | | O.H. *siu* | [ ʃ i u ] | |
| WOT. *şıma* | [ ʃ **ï** m a ] | | WOT. *şiu* | [ ʃ i u ] | |
| Cost 0.33 | N-align 4 | LD% 8.25 | Cost 0 | N-align 3 | LD% 0 |
| **sok** ‘much, many’ | | | **sólyom** ‘falcon’ | | |
| O.H. *sok* | [ **ʃ** o k ] | | O.H. *csavli* | [ t͡ʃ a v l **i** ] | |

| | | | | | |
|---|---|---|---|---|---|
| WOT. *çok** | [ **t͡ʃ** o k ] | | WOT. *çavlı** | [ t͡ʃ a v l **ï** ] | |
| Cost 0.5 | N-align 3 | LD% 16.67 | Cost 0.33 | N-align 5 | LD% 6.6 |
| **som** ‘dogwood (tree), cornel’ | | | **sör** ‘beer’ | | |
| O.H. *csum* | [ t͡ʃ u m ] | | O.H. *sere* | [ ʃ **ɛ** r ɛ ] | |
| WOT. *çum* | [ t͡ʃ u m ] | | WOT. *şire* | [ ʃ **i** r ɛ ] | |
| Cost 0 | N-align 3 | LD% 0 | Cost 0.33 | N-align 4 | LD% 8.25 |
| **sőreg** ‘starlet, sturgeon’ | | | **süllő** ‘pike perch, zander’ | | |
| O.H. *sevrege* | [ ʃ **ɛ** v r **ɛ** g ɛ **-** ] | | O.H. *silleg* | [ ʃ i l: **ɛ** ɣ ] | |
| WOT. *şivrigey* | [ ʃ **i** v r **i** g ɛ **j** ] | | WOT. *şilliğ* | [ ʃ i l: **i** ɣ ] | |
| Cost 1.16 | N-align 8 | LD% 14.5 | Cost 0.33 | N-align 5 | LD% 6.6 |
| **süly** ‘scurvy’ | | | **sűrű** ‘dense, thick (of woods or soup)’ | | |
| O.H. *sigül* | [ ʃ i ɣ y l ] | | O.H. *csireg* | [ t͡ʃ i r **ɛ** ɣ ] | |
| WOT. *şiğül* | [ ʃ i ɣ y l ] | | WOT. *çirăğ** | [ t͡ʃ i r **æ** ɣ ] | |
| Cost 0 | N-align 5 | LD% 0 | Cost 0.33 | N-align 5 | LD% 6.6 |
| **süv** ‘uncle, nephew, brother/sister in law’ | | | **szaka** ‘uvula, the recurved point of a fishing hook’ | | |
| O.H. *sig* | [ ʃ i ɣ ] | | O.H. *szakag* | [ s a k a ɣ ] | |
| WOT. *şiğ** | [ ʃ i ɣ ] | | WOT. *sakağ* | [ s a k a ɣ ] | |
| Cost 0 | N-align 3 | LD% 0 | Cost 0 | N-align 5 | LD% 0 |
| **szakáll** ‘beard’ | | | **szál** ‘raft’ | | |
| O.H. *szakal* | [ s a k a l ] | | O.H. *szal* | [ s **a** l ] | |
| WOT. *sakal* | [ s a k a l ] | | WOT. *sal* | [ s **a:** l ] | |
| Cost 0 | N-align 5 | LD% 0 | Cost 0.33 | N-align 3 | LD% 11 |
| **szám** ‘number’ | | | **szán** ‘to have pity for, to regret’ | | |
| O.H. *szam* | [ s **a** m ] | | O.H. *szagin* | [ s a ɣ **i** n ] | |
| WOT. *sam* | [ s **a:** m ] | | WOT. *sağın* | [ s a ɣ **ï** n ] | |
| Cost 0.33 | N-align 3 | LD% 11 | Cost 0.33 | N-align 5 | LD% 6.6 |
| **szán** ‘to wish, to intend something for somebody’ | | | **szapu** ‘bucket, wooden pail’ | | |
| O.H. *szán* | [ s a: n ] | | O.H. *szapag* | [ s a p a ɣ ] | |
| WOT. *san* | [ s a: n ] | | WOT. *sapağ** | [ s a p a ɣ ] | |

| | | | | | |
|---|---|---|---|---|---|
| Cost 0 | N-align 3 | LD% 0 | Cost 0 | N-align 5 | LD% 0 |
| **szár** ‘light coloured, yellowish, pale’ | | | **szárny** ‘wing’ | | |
| O.H. *szar* | [ s a r ] | | O.H. *sárn* | [ **ʃ a: -** r **- n** ] | |
| WOT. *sar** | [ s a r ] | | WOT. *çağrın* | [ **t͡ʃ a ɣ** r **ï n** ] | |
| Cost 0 | N-align 3 | LD% 0 | Cost 2.33 | N-align 6 | LD% 38.83 |
| **szatócs** ‘grocer, grand handler’ | | | **szék** ‘chair, seat, bench, throne’ | | |
| O.H. *szatigcsi* | [ s a t **ï** ɣ t͡ʃ **ï** ] | | O.H. *szekü* | [ s **ɛ** k y ] | |
| WOT. *satığçı* | [ s a t **i** ɣ t͡ʃ **i** ] | | WOT. *säkü* | [ s **æ** k y ] | |
| Cost 0.66 | N-align 7 | LD% 9.43 | Cost 0.33 | N-align 4 | LD% 8.25 |
| **szél** ‘wind’ | | | **szender(ed)(ik)** ‘to slumber, to doze’ | | |
| O.H. *szél* | [ **s** e: l ] | | O.H. *szöndür* | [ s ø n **d** y r ] | |
| WOT. *şel* | [ **ʃ** e: l ] | | WOT. *söntür** | [ s ø n **t** y r ] | |
| Cost 0.33 | N-align 3 | LD% 11 | Cost 0.5 | N-align 6 | LD% 8.33 |
| **szép** ‘beautiful’ | | | **szeplő** ‘freckle, sun spot, stain’ | | |
| O.H. *szip* | [ s i p ] | | O.H. *szepleg* | [ s **ɛ** p l **ɛ** ɣ ] | |
| WOT. *sip** | [ s i p ] | | WOT. *säpliğ* | [ s **æ** p l **i** ɣ ] | |
| Cost 0 | N-align 3 | LD% 0 | Cost 0.66 | N-align 6 | LD% 11 |
| **szer** ‘part of a village, street, group of people’ | | | **szérű** ‘treshing floor (round)’ | | |
| O.H. *ser* | [ **ʃ** ɛ r ] | | O.H. *szürüg* | [ **s** y r y ɣ ] | |
| WOT. *çer* | [ **t͡ʃ** ɛ r ] | | WOT. *çürüğ** | [ **t͡ʃ** y r y ɣ ] | |
| Cost 0.5 | N-align 3 | LD% 16.67 | Cost 0.83 | N-align 5 | LD% 16.6 |
| **szesz** ‘fumes, vapour, spirits, alcohol, humours’ | | | **szín** ‘colour, face, external appearance’ | | |
| O.H. *szisz* | [ s i s ] | | O.H. *szin* | [ s i n ] | |
| WOT. *sis* | [ s i s ] | | WOT. *sin* | [ s i n ] | |
| Cost 0 | N-align 3 | LD% 0 | Cost 0 | N-align 3 | LD% 0 |
| **szirony** ‘thin hide rope, strap’ | | | **szirt** ‘rock, cliff, occiput, crest (of mountain)’ | | |
| O.H. *sziyrum* | [ s **i** j r u m ] | | O.H. *szirt* | [ s **i** r t ] | |
| WOT. *sıyrum* | [ s **ï** j r u m ] | | WOT. *sırt* | [ s **ï** r t ] | |
| Cost 0.33 | N-align 6 | LD% 5.5 | Cost 0.33 | N-align 4 | LD% 8.25 |

| **szó** ‘word’ | | | **szór** ‘to spread, to scatter, to winnow’ | | |
|---|---|---|---|---|---|
| O.H. *szaw* | [ s a **w** ] | | O.H. *szawur* | [ s a w u r ] | |
| WOT. *sav* | [ s a **v** ] | | WOT. *sawur* | [ s a w u r ] | |
| Cost 1.5 | N-align 3 | LD% 50 | Cost 0 | N-align 5 | LD% 0 |
| **szöcske** ‘grasshoper’ | | | **szök(ik)** ‘to leap, to jump, to escape, to flee’ | | |
| O.H. *szök* | [ s **ø** k ] | | O.H. *szek* | [ s ɛ k ] | |
| WOT. *sek* | [ s **ɛ** k ] | | WOT. *sek* | [ s ɛ k ] | |
| Cost 0.66 | N-align 3 | LD% 22 | Cost 0 | N-align 3 | LD% 0 |
| **szőlő** ‘grape, wine grape’ | | | **szúnyog** ‘mosquito, gnat’ | | |
| O.H. *seleg* | [ **ʃ** ɛ l ɛ ɣ ] | | O.H. *szinuk* | [ s i **n** u k ] | |
| WOT. *çeleğ* | [ **t͡ʃ** ɛ l ɛ ɣ ] | | WOT. *sinuk* | [ s i **ŋ** u k ] | |
| Cost 0.5 | N-align 5 | LD% 10 | Cost 0.66 | N-align 5 | LD% 13.2 |
| **szűcs** ‘furrier’ | | | **szűn(ik)** ‘to cease, to stop’ | | |
| O.H. *sigcsi* | [ **ʃ** i ɣ t͡ʃ i ] | | O.H. *szön* | [ s ø n ] | |
| WOT. *çiğçi* | [ **t͡ʃ** i ɣ t͡ʃ i ] | | WOT. *sön* | [ s ø n ] | |
| Cost 0.5 | N-align 5 | LD% 10 | Cost 0 | N-align 3 | LD% 0 |
| **szűr** ‘to strain, to filter’ | | | **szűz** ‘virgin, pure’ | | |
| O.H. *szür* | [ s y r ] | | O.H. *szüzök* | [ s y z ø k ] | |
| WOT. *sür* | [ s y r ] | | WOT. *süzök** | [ s y z ø k ] | |
| Cost 0 | N-align 3 | LD% 0 | Cost 0 | N-align 5 | LD% 0 |
| **táltos** ‘sorcerer, medicine man’ | | | **tanács** ‘advice, council’ | | |
| O.H. *taltucsi* | [ t a l t u t͡ʃ **i** ] | | O.H. *tanilcs* | [ t a n **i** l t͡ʃ ] | |
| WOT. *taltuçı** | [ t a l t u t͡ʃ **ï** ] | | WOT. *tanılç* | [ t a n **ï** l t͡ʃ ] | |
| Cost 0.33 | N-align 7 | LD% 4.71 | Cost 0.33 | N-align 6 | LD% 5.5 |
| **tanú** ‘witness’ | | | **táplá(l)** ‘to feed, to nourish’ | | |
| O.H. *tanu* | [ t a n u **-** ] | | O.H. *taplag* | [ t a p l a ɣ ] | |
| WOT. *tanuğ* | [ t a n u **ɣ** ] | | WOT. *taplağ* | [ t a p l a ɣ ] | |
| Cost 0.5 | N-align 5 | LD% 10 | Cost 0 | N-align 6 | LD% 0 |
| **tapló** ‘tinder, tinder fungus’ | | | **tar** ‘bald’ | | |
| O.H. *toplug* | [ t o p l u ɣ ] | | O.H. *tar* | [ t a r ] | |

| | | | | | |
|---|---|---|---|---|---|
| WOT. *topluğ* | [ t o p l u ɣ ] | | WOT. *tar* | [ t a r ] | |
| Cost 0 | N-align 6 | LD% 0 | Cost 0 | N-align 3 | LD% 0 |
| **tár** ‘depot’ | | | **tarló** ‘plough field, arable field, stubble field’ | | |
| O.H. *tavar* | [ t a v a r ] | | O.H. *tarlag* | [ t a r - l a ɣ ] | |
| WOT. *tavar* | [ t a v a r ] | | WOT. *tarılağ* | [ t a r ï l a ɣ ] | |
| Cost 0 | N-align 5 | LD% 0 | Cost 1 | N-align 7 | LD% 14.29 |
| **tart** ‘to hold, carry, to last’ | | | **tatár** ‘tatar’ | | |
| O.H. *tart* | [ t a r t ] | | O.H. *tatar* | [ t a t a r ] | |
| WOT. *tart* | [ t a r t ] | | WOT. *tatar* | [ t a t a r ] | |
| Cost 0 | N-align 4 | LD% 0 | Cost 0 | N-align 5 | LD% 0 |
| **tátorján** ‘name of several kinds of plants’ | | | **teker** ‘to wind something round, to twist’ | | |
| O.H. *tatrang* | [ t a t r a ŋ ] | | O.H. *teker* | [ t ɛ k ɛ r ] | |
| WOT. *tatran* | [ t a t r a ŋ ] | | WOT. *täkir* | [ t æ k i r ] | |
| Cost 0 | N-align 6 | LD% 0 | Cost 0.66 | N-align 5 | LD% 13.2 |
| **teknő** ‘through, wash tub, hutch’ | | | **telek** ‘a piece of land, parcel, patch’ | | |
| O.H. *tekeneg* | [ t ɛ k ɛ n ɛ ɣ ] | | O.H. *telük* | [ t ɛ l y k ] | |
| WOT. *täkänäğ* | [ t æ k æ n æ ɣ ] | | WOT. *tilök* | [ t i l ø k ] | |
| Cost 1 | N-align 7 | LD% 14.29 | Cost 0.66 | N-align 5 | LD% 13.2 |
| **telek** ‘strap (on a whip, or on a kind of sandal)’ | | | **teng** ‘to vegetate, to linger in misery’ | | |
| O.H. *telük* | [ t ɛ l y k ] | | O.H. *teng* | [ t ɛ ŋ ] | |
| WOT. *tilök* | [ t i l ø k ] | | WOT. *tän** | [ t æ ŋ ] | |
| Cost 0.66 | N-align 5 | LD% 13.2 | Cost 0.33 | N-align 3 | LD% 11 |
| **tengely** ‘axle’ | | | **tenger** ‘sea’ | | |
| O.H. *tengel* | [ t ɛ ŋ g ɛ l ] | | O.H. *tenir* | [ t ɛ ŋ i r ] | |
| WOT. *tenäl* | [ t ɛ ŋ - æ l ] | | WOT. *tänir* | [ t æ ŋ i r ] | |
| Cost 1.33 | N-align 6 | LD% 22.17 | Cost 0.33 | N-align 5 | LD% 6.6 |
| **tér** ‘to turn, to change the original direction’ | | | **tér(d)** ‘knee’ | | |
| O.H. *tevir* | [ t ɛ v i r ] | | O.H. *tír* | [ t i: r ] | |
| WOT. *tävir* | [ t æ v i r ] | | WOT. *tir* | [ t i: r ] | |
| Cost 0.33 | N-align 5 | LD% 6.6 | Cost 0 | N-align 3 | LD% 0 |

| **terem** ‘hall, chamber, great room’ | | | **teve** ‘camel’ | | |
|---|---|---|---|---|---|
| O.H. *terem* | [ t ɛ r ɛ m ] | | O.H. *teve* | [ t ɛ v ɛ ] | |
| WOT. *terem* | [ t ɛ r ɛ m ] | | WOT. *teve* | [ t ɛ v ɛ ] | |
| Cost 0 | N-align 5 | LD% 0 | Cost 0 | N-align 4 | LD% 0 |
| **tiló** ‘hemp breaker, swingle, scutch’ | | | **tinó** ‘steer, young bullock, ox’ | | |
| O.H. *talkig* | [ t a l k **i** ɣ ] | | O.H. *tana* | [ t a n a ] | |
| WOT. *talkığ* | [ t a l k **ï** ɣ ] | | WOT. *tana* | [ t a n a ] | |
| Cost 0.33 | N-align 6 | LD% 5.5 | Cost 0 | N-align 4 | LD% 0 |
| **toj(ik)** ‘to lay eggs’ | | | **tok** ‘sturgeon’ | | |
| O.H. *tug* | [ t u ɣ ] | | O.H. *toku* | [ t o k u ] | |
| WOT. *tuğ* | [ t u ɣ ] | | WOT. *toku* | [ t o k u ] | |
| Cost 0 | N-align 3 | LD% 0 | Cost 0 | N-align 4 | LD% 0 |
| **tolmács** ‘interpreter’ | | | **torma** ‘horseradish’ | | |
| O.H. *tolmacs* | [ t o l m a t͡ʃ ] | | O.H. *turma* | [ t u r m a ] | |
| WOT. *tolmaç* | [ t o l m a t͡ʃ ] | | WOT. *turma* | [ t u r m a ] | |
| Cost 0 | N-align 6 | LD% 0 | Cost 0 | N-align 5 | LD% 0 |
| **toro(l)** ‘to pile up’ | | | **torontál** ‘a kind of small falcon’ | | |
| O.H. *toro* | [ t **o** r o ] | | O.H. *turuntay* | [ t u r u n t a j ] | |
| WOT. *turo** | [ t **u** r o ] | | WOT. *turuntay* | [ t u r u n t a j ] | |
| Cost 0.33 | N-align 4 | LD% 8.25 | Cost 0 | N-align 8 | LD% 0 |
| **Tót** ‘Slovak’ | | | **tökél** ‘to perform’ | | |
| O.H. *tát* | [ t a: t ] | | O.H. *tüke* | [ t y k **ɛ** ] | |
| WOT. *tat* | [ t a: t ] | | WOT. *tükä** | [ t y k **æ** ] | |
| Cost 0 | N-align 3 | LD% 0 | Cost 0.33 | N-align 4 | LD% 8.25 |
| **tömény** ‘concentrated, ten thousand, a military unit’ | | | **tör** ‘to break, to separate into pieces’ | | |
| O.H. *tümen* | [ t y m **ɛ** n ] | | O.H. *tügür* | [ t **y** ɣ y r ] | |
| WOT. *tümän* | [ t y m **æ** n ] | | WOT. *töğür** | [ t **ø** ɣ y r ] | |
| Cost 0.33 | N-align 5 | LD% 6.6 | Cost 0.33 | N-align 5 | LD% 6.6 |
| **tőr** ‘snare, trap’ | | | **Török** ‘Turk, Turkish, Turkic, Ottoman’ | | |
| O.H. *tőr* | [ t ø: r ] | | O.H. *türkü* | [ t y r k y ] | |

| | | | | | |
|---|---|---|---|---|---|
| WOT. *tör** | [ t ø: r ] | | WOT. *türkü* | [ t y r k y ] | |
| Cost 0 | N-align 3 | LD% 0 | Cost 0 | N-align 5 | LD% 0 |
| **történ(ik)** ‘to happen, to occur’ | | | **törvény** ‘law’ | | |
| O.H. *törtün* | [ t ø r t y n ] | | O.H. *türügen* | [ t **y** r y ɣ **ɛ** n ] | |
| WOT. *törtün* | [ t ø r t y n ] | | WOT. *törüğän* | [ t **ø** r y ɣ **æ** n ] | |
| Cost 0 | N-align 6 | LD% 0 | Cost 0.66 | N-align 7 | LD% 9.43 |
| **tulok** ‘steer, young bullock, ox or cow’ | | | **túró** ‘cottage cheese, cheese curds’ | | |
| O.H. *tukol* | [ t **u** k o l ] | | O.H. *torag* | [ t o r a ɣ ] | |
| WOT. *tokol* | [ t **o** k o l ] | | WOT. *torağ* | [ t o r a ɣ ] | |
| Cost 0.33 | N-align 5 | LD% 6.6 | Cost 0 | N-align 5 | LD% 0 |
| **turul** ‘the totem bird of the Árpád dynasty’ | | | **túzok** ‘bustard’ | | |
| O.H. *tugril* | [ t **u** ɣ r ï l ] | | O.H. *togzak* | [ t o **ɣ** **z** a k ] | |
| WOT. *toğrıl* | [ t **o** ɣ r ï l ] | | WOT. *togzak* | [ t o **g** **ð** a k ] | |
| Cost 0.33 | N-align 6 | LD% 5.5 | Cost 0.41 | N-align 6 | LD% 6.83 |
| **tükör** ‘mirror’ | | | **tűr** ‘to endure, suffer, bear, stand’ | | |
| O.H. *tüker* | [ t **y** k **ɛ** r ] | | O.H. *tür* | [ t **y** r ] | |
| WOT. *tikär* | [ t **i** k **æ** r ] | | WOT. *tör* | [ t **ø** r ] | |
| Cost 0.66 | N-align 5 | LD% 13.2 | Cost 0.33 | N-align 3 | LD% 11 |
| **tűr** ‘to roll up (a scroll or one’s sleeves)’ | | | **tyúk** ‘hen’ | | |
| O.H. *tür* | [ t y r ] | | O.H. *tiguk* | [ t i ɣ u k ] | |
| WOT. *tür* | [ t y r ] | | WOT. *tiğuk* | [ t i ɣ u k ] | |
| Cost 0 | N-align 3 | LD% 0 | Cost 0 | N-align 5 | LD% 0 |
| **ugu** ‘owl’ | | | **úr** ‘gentleman, lord, sir, high dignity’ | | |
| O.H. *ugu* | [ u **g** u ] | | O.H. *uri* | [ u r **i** ] | |
| WOT. *uğu** | [ u **ɣ** u ] | | WOT. *urı** | [ u r **ï** ] | |
| Cost 0.25 | N-align 3 | LD% 8.33 | Cost 0.33 | N-align 3 | LD% 11 |
| **üdü(l)** ‘to refresh oneself, to rest, to become cured’ | | | **ün(nep)** ‘feast’ | | |
| O.H. *edü* | [ ɛ d y ] | | O.H. *edü* | [ e d y ] | |
| WOT. *edü** | [ ɛ d y ] | | WOT. *edü* | [ e d y ] | |
| Cost 0 | N-align 3 | LD% 0 | Cost 0 | N-align 3 | LD% 0 |

| **ünő** ‘heifer, roe doe’ | | | **ürge** ‘ground squirrel’ | | |
|---|---|---|---|---|---|
| O.H. *ineg* | [ i n **ɛ** ɣ ] | | O.H. *ürge* | [ **y** r g **ɛ** ] | |
| WOT. *inäğ* | [ i n **æ** ɣ ] | | WOT. *örgä* | [ **ø** r g **æ** ] | |
| Cost 0.33 | N-align 4 | LD% 8.25 | Cost 0.66 | N-align 4 | LD% 16.5 |
| **üröm** ‘wormwood’ | | | **ürü** ‘wether, sheep’ | | |
| O.H. *erim* | [ ɛ r i m ] | | O.H. *irig* | [ i r i ɣ ] | |
| WOT. *erim* | [ ɛ r i m ] | | WOT. *irığ* | [ i r i ɣ ] | |
| Cost 0 | N-align 4 | LD% 0 | Cost 0 | N-align 4 | LD% 0 |
| **üszök** ‘hot embers, extinguished embers’ | | | **üvecs** ‘one-year-old female sheep’ | | |
| O.H. *üszüg* | [ i s **y** g ] | | O.H. *övecs* | [ ø v **ɛ** t͡ʃ ] | |
| WOT. *isig* | [ i s **i** g ] | | WOT. *öväç* | [ ø v **æ** t͡ʃ ] | |
| Cost 0.33 | N-align 4 | LD% 8.25 | Cost 0.33 | N-align 4 | LD% 8.25 |
| **váj** ‘to hollow out, to scoop out’ | | | **vályú** ‘trough, tray’ | | |
| O.H. *vay* | [ v a j ] | | O.H. *valag* | [ v a l a ɣ ] | |
| WOT. *vay* | [ v a j ] | | WOT. *valağ* | [ v a l a ɣ ] | |
| Cost 0 | N-align 3 | LD% 0 | Cost 0 | N-align 5 | LD% 0 |
| **vejsze** ‘fishweir’ | | | **vék** ‘a hole in the ice’ | | |
| O.H. *vejseg* | [ v **ɛ** j ʃ **ɛ** ɣ ] | | O.H. *vekü* | [ v **ɛ** k y ] | |
| WOT. *väyşäğ** | [ v **æ** j ʃ **æ** ɣ ] | | WOT. *väkü* | [ v **æ** k y ] | |
| Cost 0.66 | N-align 6 | LD% 11 | Cost 0.33 | N-align 4 | LD% 8.25 |
| **vendég** ‘guest’ | | | **ver** ‘to plait, to lay the rope’ | | |
| O.H. *vendeg* | [ v **ɛ** n d **ɛ** g ] | | O.H. *ver* | [ v ɛ r ] | |
| WOT. *vändäg** | [ v **æ** n d **æ** g ] | | WOT. *ver** | [ v ɛ r ] | |
| Cost 0.66 | N-align 6 | LD% 11 | Cost 0 | N-align 3 | LD% 0 |
| **vértelek** ‘gable, frontispiece’ | | | **zerge** ‘chamois’ | | |
| O.H. *vertelük* | [ v ɛ r t **ɛ** l y k ] | | O.H. *szerge* | [ s **ɛ** r g **ɛ** ] | |
| WOT. *vertülük* | [ v ɛ r t **y** l y k ] | | WOT. *särgä** | [ s **æ** r g **æ** ] | |
| Cost 0.33 | N-align 8 | LD% 4.12 | Cost 0.66 | N-align 5 | LD% 13.2 |

## 3.2Adaptation process

The adaptation of West Old Turkic loanwords in Hungarian has taken place over so many centuries. The time frame for adaptation spans from the initial borrowing of these loanwords to their subsequent assimilation and integration into the evolving phonetic and phonological framework of Hungarian. Throughout this historical process, the loanwords –just like the native words– have undergone various phonetic and phonological changes to align with the evolving sound patterns and structures of the Hungarian language. These changes reflect –to a large extent– the ongoing adaptation and assimilation of foreign linguistic elements into the native system, shaped by linguistic conventions, phonotactic constraints, and the influence of other factors such as language contact and language change.

The pronunciation distance of the 377 word pairs is on average 28.67%. Therefore, the modern Hungarian adaptations and the initial copies of this set of words are 71.33% identical (homophones). Turkic material in Hungarian is in general not very recognizable by Turkic native speakers (Navracsics, 2016, p. 15). For that reason, this distance may seem short at first glance. However, it should be noted that the common Turkic languages – all Turkic languages today, except Chuvash – are highly distant from the Oghur branch and mutual intelligibility is quite limited. One consideration that is important at this stage is the degree of relevance the source languages, from which Hungarian borrowed the words, had with modern-day Chuvash. Three dialects of WOT from the 10$^{th}$ to 13$^{th}$ centuries were reconstructed, two being extinct and one being Middle Chuvash, but what concerns the Hungarian interaction of the process more was apparently before the 10$^{th}$ century. For what is known before that period about the Chuvash-type languages such as Khazar, Saragur, Onogur, and Volga-Bulgar, the strongest source already is the WOT loanwords in Hungarian (Agyagási, 2019).

**Table 13. Adaptation rates of reconstructed Old Hungarian copies in Modern Hungarian**

| carpenter | | cannon, catapult | |
|---|---|---|---|
| H. *ács* | [ **aː - - t͡ʃ -** ] | H. *ágyú* | [ **aː - ɟ uː -** ] |
| O.H. *agacsi* | [ **a ɣ a t͡ʃ i** ] | O.H. *aldag* | [ **a l d a ɣ** ] |

| Cost 2.83 | N-align 5 | LD% 56.6 | Cost 3.16 | N-align 5 | LD% 63.2 |
|---|---|---|---|---|---|
| to fondle, pet, caress | | | false, imitation | | |
| H. *ajn(ároz)* | [ ɒ j - n ] | | H. *ál* | [ a: l ] | |
| O.H. *ajan* | [ a j a n ] | | O.H. *ál* | [ a: l ] | |
| Cost 1.33 | N-align 4 | LD% 33.25 | Cost 0 | N-align 2 | LD% 0 |
| pied (of an animals' coat) | | | apple | | |
| H. *alacs* | [ ɒ l ɒ t͡ʃ ] | | H. *alma* | [ ɒ l m ɒ ] | |
| O.H. *alacs* | [ a l a t͡ʃ ] | | O.H. *alma* | [ a l m a ] | |
| Cost 0.66 | N-align 4 | LD% 16.5 | Cost 0.66 | N-align 4 | LD% 16.5 |
| barre (wooden container and standard quantity) | | | to nurse, take care of | | |
| H. *általag* | [ a: l t ɒ l ɒ g ] | | H. *áp(ol)* | [ a: p ] | |
| O.H. *altalag* | [ a l t a l a g ] | | O.H. *op* | [ o p ] | |
| Cost 1 | N-align 7 | LD% 14.29 | Cost 1 | N-align 2 | LD% 50 |
| to decay, putrify, turn stale | | | small, tiny | | |
| H. *áporo(d)(ik)* | [ a: p o r o ] | | H. *apró* | [ ɒ p - r o: - ] | |
| O.H. *oporo* | [ o p o r o ] | | O.H. *opurug* | [ o p u r u ɣ ] | |
| Cost 1 | N-align 5 | LD% 20 | Cost 2.50 | N-align 6 | LD% 41.67 |
| to mow | | | to decay, go bad (of food, water) | | |
| H. *arat* | [ ɒ r ɒ t ] | | H. *ár(ik)* | [ a: r ] | |
| O.H. *orat* | [ o r a t ] | | O.H. *ár* | [ a: r ] | |
| Cost 0.66 | N-align 4 | LD% 16.5 | Cost 0 | N-align 2 | LD% 0 |
| ditch, canal | | | barley | | |
| H. *árok* | [ a: r o k ] | | H. *árpa* | [ a: r p ɒ ] | |
| O.H. *áruk* | [ a: r u k ] | | O.H. *arpa* | [ a r p a ] | |
| Cost 0.33 | N-align 4 | LD% 8.25 | Cost 0.66 | N-align 4 | LD% 16.5 |
| barrow | | | gauntry, gantry (for supporting barrels) | | |
| H. *ártány* | [ a: r - t a: ɲ ] | | H. *ászok* | [ a: s o k ] | |
| O.H. *aritan* | [ a r i t a n ] | | O.H. *aski* | [ a s k i ] | |
| Cost 2.33 | N-align 6 | LD% 38.83 | Cost 2.33 | N-align 4 | LD% 58.25 |
| charm | | | soggy place, moor, swamp, marsh | | |

| | | | | | |
|---|---|---|---|---|---|
| H. *báj* | [ b **aː** j ] | | H. *balkány* | [ b **ɒ** l k **aː ŋ** ] | |
| O.H. *baj* | [ b **a** j ] | | O.H. *balkan* | [ b **a** l k **a n** ] | |
| Cost 0.33 | N-align 3 | LD% 11 | Cost 1.33 | N-align 6 | LD% 22.17 |
| hatchet, axe | | | pagan idol, stone or wooden pillar, strong(smb) | | |
| H. *balta* | [ b **ɒ** l t **ɒ** ] | | H. *bálvány* | [ b **aː** l v **aː ŋ** ] | |
| O.H. *balta* | [ b **a** l t **a** ] | | O.H. *balvan* | [ b **a** l v **a n** ] | |
| Cost 0.66 | N-align 5 | LD% 13.2 | Cost 1.33 | N-align 6 | LD% 22.17 |
| governor of Croatia (of the southern marches) | | | cattle, livestock | | |
| H. *bán* | [ b a: n ] | | H. *barom* | [ b **ɒ** r **o** m ] | |
| O.H. *bán* | [ b a: n ] | | O.H. *barum* | [ b **a** r **u** m ] | |
| Cost 0 | N-align 3 | LD% 0 | Cost 0.66 | N-align 5 | LD% 13.2 |
| velvet | | | to fuck | | |
| H. *bársony* | [ b **aː** r **ʃ o ŋ** ] | | H. *basz(ik)* | [ b **ɒ** s ] | |
| O.H. *barcsun* | [ b **a** r **t͡ʃ u n** ] | | O.H. *basz* | [ b **a** s ] | |
| Cost 1.83 | N-align 6 | LD% 30.5 | Cost 0.33 | N-align 3 | LD% 11 |
| courageous, brave, valiant | | | to estimate, to esteem, to appreciate | | |
| H. *bátor* | [ b **aː** - - t **o** r ] | | H. *becsül* | [ b ɛ **t͡ʃ y** l ] | |
| O.H. *bagatur* | [ b **a ɣ a** t **u** r ] | | O.H. *becsel* | [ b ɛ **t͡ʃ ɛ** l ] | |
| Cost 2.16 | N-align 7 | LD% 30.86 | Cost 0.66 | N-align 5 | LD% 13.2 |
| frog | | | hobble, shackle, leg-iron | | |
| H. *béka* | [ b **eː** k **ɒ** ] | | H. *béklyó* | [ b **eː** k **- j oː** ] | |
| O.H. *baka* | [ b **a** k **a** ] | | O.H. *bikagu* | [ b **i** k **a ɣ u** ] | |
| Cost 1.33 | N-align 4 | LD% 33.25 | Cost 3.15 | N-align 6 | LD% 52.5 |
| stamp, bond, mark | | | wage, rent | | |
| H. *bélyeg* | [ b **eː j** ɛ g ] | | H. *bér* | [ b e: r **-** ] | |
| O.H. *beleg* | [ b **ɛ l** ɛ g ] | | O.H. *bérü* | [ b e: r **y** ] | |
| Cost 1.91 | N-align 5 | LD% 38.2 | Cost 1 | N-align 4 | LD% 25 |
| name of a people | | | grain or small globule of precious metal | | |
| H. *bercel* | [ b ɛ r **t͡s ɛ** l ] | | H. *bertű* | [ b **ɛ** r t **yː -** ] | |
| O.H. *berszil* | [ b ɛ r **s i** l ] | | O.H. *bürtig* | [ b **y** r t **i ɣ** ] | |

| Cost 1.16 | N-align 6 | LD% 19.33 | Cost 1.82 | N-align 6 | LD% 30.33 |
|---|---|---|---|---|---|
| a bird of prey | | | Pecheneg, name of a tribe | | |
| H. *bese* | [ b ɛ ʃ ɛ ] | | H. *besenyő* | [ b ɛ ʃ ɛ **ɲ ø: -** ] | |
| O.H. *bese* | [ b ɛ ʃ ɛ ] | | O.H. *beseneg* | [ b ɛ ʃ ɛ **n ɛ ɣ** ] | |
| Cost 0 | N-align 4 | LD% 0 | Cost 2.16 | N-align 7 | LD% 30.86 |
| letter (of the alphabet) | | | pocket knife | | |
| H. *betű* | [ b **ɛ** t **y: -** ] | | H. *bicska* | [ b i t͡ʃ k **ɒ** ] | |
| O.H. *bitig* | [ b **i** t **i ɣ** ] | | O.H. *bicsak* | [ b i t͡ʃ **a** k ] | |
| Cost 1.5 | N-align 5 | LD% 30 | Cost 1.33 | N-align 5 | LD% 26,6 |
| bull | | | shackles | | |
| H. *bika* | [ b i k **ɒ** ] | | H. *bilincs* | [ b i l **i n** t͡ʃ ] | |
| O.H. *bika* | [ b i k **a** ] | | O.H. *bilecs(ek)* | [ b i l **ɛ -** t͡ʃ ] | |
| Cost 0.33 | N-align 4 | LD% 8.25 | Cost 1.33 | N-align 6 | LD% 22.17 |
| judge | | | to be forgiven | | |
| H. *bíró* | [ b **i:** r **o: -** ] | | H. *bocsán*(ik) | [ b o **-** t͡ʃ **a:** n ] | |
| O.H. *birug* | [ b **i** r **u ɣ** ] | | O.H. *bolcsan* | [ b o **l** t͡ʃ **a** n ] | |
| Cost 1.5 | N-align 5 | LD% 30 | Cost 1.33 | N-align 6 | LD% 22.17 |
| to forgive | | | knot, bend | | |
| H. *bocsát* | [ b **o -** t͡ʃ **a:** t ] | | H. *bog* | [ b o g ] | |
| O.H. *bulcsat* | [ b **u l** t͡ʃ **a** t ] | | O.H. *bog* | [ b o g ] | |
| Cost 1.66 | N-align 6 | LD% 27.67 | Cost 0 | N-align 3 | LD% 0 |
| stack of hay | | | tassel, fringe, tuft, node | | |
| H. *boglya* | [ b o g **j ɒ** ] | | H. *bojt* | [ b **o j -** t ] | |
| O.H. *bogul* | [ b o g **u l** ] | | O.H. *bugut* | [ b **u ɣ u** t ] | |
| Cost 2.48 | N-align 5 | LD% 49.6 | Cost 2.15 | N-align 5 | LD% 43 |
| burdock | | | ankle | | |
| H. *bojtorján* | [ b **o j** t **o** r **j a:** n ] | | H. *boka* | [ b **o** k **ɒ** ] | |
| O.H. *balturgan* | [ b **a l** t **u** r **ɣ a** n ] | | O.H. *baka* | [ b **a** k **a** ] | |
| Cost 3.4 | N-align 9 | LD% 37.78 | Cost 1 | N-align 4 | LD% 25 |
| anthill | | | wine | | |

| | | | | | |
|---|---|---|---|---|---|
| H. *boly* | [ b o - - j ] | | H. *bor* | [ b o r ] | |
| O.H. *bogol* | [ b o ɣ o l ] | | O.H. *bor* | [ b o r ] | |
| Cost 2.75 | N-align 5 | LD% 55 | Cost 0 | N-align 3 | LD% 0 |
| to cover, to overturn | | | calf | | |
| H. *bor(ít)* | [ b o r ] | | H. *borjú* | [ b o r - j u: ] | |
| O.H. *bur* | [ b u r ] | | O.H. *buragu* | [ b u r a ɣ u ] | |
| Cost 0.33 | N-align 3 | LD% 11 | Cost 2.48 | N-align 6 | LD% 41.33 |
| pepper | | | pea, bean | | |
| H. *bors* | [ b o r ʃ ] | | H. *borsó* | [ b o r ʃ o - ] | |
| O.H. *burcs* | [ b u r t͡ʃ ] | | O.H. *burcsag* | [ b u r t͡ʃ a ɣ ] | |
| Cost 0.83 | N-align 4 | LD% 20.75 | Cost 2 | N-align 6 | LD% 33.33 |
| badger | | | witch, sorceress | | |
| H. *borz* | [ b o r z - ] | | H. *boszorkány* | [ b o s o r k a: ɲ ] | |
| O.H. *borszi* | [ b o r s i ] | | O.H. *baszarkan* | [ b a s a r k a n ] | |
| Cost 1.5 | N-align 5 | LD% 30 | Cost 2.33 | N-align 8 | LD% 29.12 |
| to annoy | | | stick, cane | | |
| H. *bosszant* | [ b o s: ɒ n t ] | | H. *bot* | [ b o t - - ] | |
| O.H. *buszan* | [ b u s a n - ] | | O.H. *butig* | [ b u t i ɣ ] | |
| Cost 2 | N-align 6 | LD% 33.33 | Cost 1.83 | N-align 5 | LD% 36.6 |
| rich, abundant, roomy | | | horsefly, gadfly | | |
| H. *bő* | [ b ø: - ] | | H. *bögöly* | [ b ø g ø j ] | |
| O.H. *beg* | [ b ɛ ɣ ] | | O.H. *bögöl(ök)* | [ b ø g ø l ] | |
| Cost 1.5 | N-align 3 | LD% 50 | Cost 1.25 | N-align 5 | LD% 25 |
| fast(ing) | | | wise | | |
| H. *böjt* | [ b ø j t - ] | | H. *bölcs* | [ b ø l - t͡ʃ - ] | |
| O.H. *bügte* | [ b y ɣ t ɛ ] | | O.H. *bügücsi* | [ b y ɣ y t͡ʃ i ] | |
| Cost 2.15 | N-align 5 | LD% 43 | Cost 3.58 | N-align 6 | LD% 59.67 |
| cradle | | | bison | | |
| H. *bölcső* | [ b ø l t͡ʃ ø: - ] | | H. *bölény* | [ b ø l e: ɲ ] | |
| O.H. *belcseg* | [ b e l t͡ʃ ɛ ɣ ] | | O.H. *belen* | [ b ɛ l ɛ n ] | |

| Cost 1.83 | N-align 6 | LD% 30.5 | Cost 2 | N-align 5 | LD% 40 |
|---|---|---|---|---|---|
| prison, jail | | | sorrow | | |
| H. *börtön* | [ b **ø** r t **ø** n ] | | H. *bú* | [ b **u: -** ] | |
| O.H. *berten* | [ b **ɛ** r t **ɛ** n ] | | O.H. *bug* | [ b **u ɣ** ] | |
| Cost 1.32 | N-align 6 | LD% 22 | Cost 0.83 | N-align 3 | LD% 27.67 |
| farewell, indulgence, dedication festival | | | Volga Bulgar | | |
| H. *búcsú* | [ b **u: -** t͡ʃ **u: -** ] | | H. *bular* | [ b u l **ɒ** r ] | |
| O.H. *bulcsug* | [ b **u l** t͡ʃ **u ɣ** ] | | O.H. *bular* | [ b u l **a** r ] | |
| Cost 2.16 | N-align 6 | LD% 36 | Cost 0.33 | N-align 5 | LD% 6.6 |
| steam, vapour, fumes | | | wheat | | |
| H. *búsz* | [ b u: s ] | | H. *búza* | [ b **u: -** z **ɒ -** ] | |
| O.H. *búsz* | [ b u: s ] | | O.H. *bugzai* | [ b **u ɣ** z **a i** ] | |
| Cost 0 | N-align 3 | LD% 0 | Cost 2.16 | N-align 6 | LD% 36 |
| beech | | | sin | | |
| H. *bükk* | [ b **y k:** ] | | H. *bűn* | [ b y: n ] | |
| O.H. *bik* | [ b **i k** ] | | O.H. *bűn* | [ b y: n ] | |
| Cost 0.66 | N-align 3 | LD% 22 | Cost 0 | N-align 3 | LD% 0 |
| end of sg, a rod with an end cut | | | magical | | |
| H. *bütü* | [ b y t y **-** ] | | H. *bűv(ös)* | [ b **y: v -** ] | |
| O.H. *bütüg* | [ b y t y **ɣ** ] | | O.H. *bügü* | [ b **y ɣ y** ] | |
| Cost 0.50 | N-align 5 | LD% 10 | Cost 2 | N-align 4 | LD% 50 |
| to regret, to be sorry | | | feast, christening feast (at baptism) | | |
| H. *bán(ik)* | [ b **a: - -** n ] | | H. *csök* | [ t͡ʃ **ø** k ] | |
| O.H. *bagin* | [ b **a ɣ i** n ] | | O.H. *csük* | [ t͡ʃ **y** k ] | |
| Cost 1.83 | N-align 5 | LD% 36.6 | Cost 0.33 | N-align 3 | LD% 11 |
| a kind of willow | | | only, just | | |
| H. *cigle* | [ **t͡s** i g l ɛ ] | | H. *csak* | [ t͡ʃ **ɒ** k ] | |
| O.H. *csigle* | [ **t͡ʃ** i g l ɛ ] | | O.H. *csak* | [ t͡ʃ **a** k ] | |
| Cost 0.33 | N-align 5 | LD% 6.6 | Cost 0.33 | N-align 3 | LD% 11 |
| pick-axe, war hammer | | | nettle | | |

| | | | | | |
|---|---|---|---|---|---|
| H. *csákány* | [ t͡ʃ **aː** k **aː ŋ** ] | | H. *csalán* | [ t͡ʃ **ɒ** l **-** **- aː** n ] | |
| O.H. *csakan* | [ t͡ʃ **a** k **a n** ] | | O.H. *csalagan* | [ t͡ʃ **a** l **a ɣ a** n ] | |
| Cost 1.33 | N-align 5 | LD% 26.6 | Cost 2.16 | N-align 7 | LD% 30.86 |
| thicket, brushwood | | | (silver) cup, bowl | | |
| H. *csalit* | [ t͡ʃ **ɒ** l i t ] | | H. *csanak* | [ t͡ʃ **ɒ** n **ɒ** k ] | |
| O.H. *csalit* | [ t͡ʃ **a** l i t ] | | O.H. *csanak* | [ t͡ʃ **a** n **a** k ] | |
| Cost 0.33 | N-align 5 | LD% 6.6 | Cost 0.66 | N-align 5 | LD% 13.2 |
| fighting, troop, battle, quarrel | | | to add, to join, to buckle up | | |
| H. *csata* | [ t͡ʃ **ɒ** t **ɒ** ] | | H. *csat(ol)* | [ t͡ʃ **ɒ** t **-** ] | |
| O.H. *csata* | [ t͡ʃ **a** t **a** ] | | O.H. *csati* | [ t͡ʃ **a** t **i** ] | |
| Cost 0.66 | N-align 4 | LD% 16.5 | Cost 1.33 | N-align 4 | LD% 33.25 |
| measle, smallpox, pocks | | | to bind (something) | | |
| H. *csécs* | [ t͡ʃ **eː** t͡ʃ ] | | H. *csek(él)* | [ t͡ʃ **ɛ** k ] | |
| O.H. *csecs(ek)* | [ t͡ʃ **e** t͡ʃ ] | | O.H. *csik* | [ t͡ʃ **i** k ] | |
| Cost 0.33 | N-align 3 | LD% 11 | Cost 0.33 | N-align 3 | LD% 11 |
| impurities, tow, chaff | | | snail | | |
| H. *csöpű* | [ t͡ʃ ø p **yː -** ] | | H. *csiga* | [ t͡ʃ **i** g **ɒ** ] | |
| O.H. *csöpüg* | [ t͡ʃ ø p **y ɣ** ] | | O.H. *csiga* | [ t͡ʃ **ï** g **a** ] | |
| Cost 0.83 | N-align 5 | LD% 16.6 | Cost 0.66 | N-align 4 | LD% 16.5 |
| a wine of low quality, a fruit one | | | secretion, mucus discharged from the eyes | | |
| H. *csiger* | [ t͡ʃ i g **ɛ** r ] | | H. *csipa* | [ t͡ʃ **i** p **ɒ** ] | |
| O.H. *csigir* | [ t͡ʃ i g **i** r ] | | O.H. *csapa* | [ t͡ʃ **a** p **a** ] | |
| Cost 0.33 | N-align 5 | LD% 6,6 | Cost 1 | N-align 5 | LD% 20 |
| a carpenter's axe with a long helve | | | sexual organ of large male animals | | |
| H. *csomak* | [ t͡ʃ o m **ɒ** k ] | | H. *csök* | [ t͡ʃ **ø** k ] | |
| O.H. *csomak* | [ t͡ʃ o m **a** k ] | | O.H. *csük* | [ t͡ʃ **y** k ] | |
| Cost 0.33 | N-align 5 | LD% 6.6 | Cost 0.33 | N-align 3 | LD% 11 |
| to become smaller, remain small | | | stubborn | | |
| H. *csök(ik)* | [ t͡ʃ **ø** k ] | | H. *csökön(ös)* | [ t͡ʃ **ø** k **ø ŋ** ] | |
| O.H. *csük* | [ t͡ʃ **y** k ] | | O.H. *csükün* | [ t͡ʃ **y** k **y n** ] | |

| | | | | | |
|---|---|---|---|---|---|
| Cost 0.33 | N-align 3 | LD% 11 | Cost 1.33 | N-align 5 | LD% 26.6 |
| field-guard | | | rough, coarse, ugly | | |
| H. *csősz* | [ t͡ʃ **ø: - s** ] | | H. *csúnya* | [ t͡ʃ **u: ɲ ɒ -** ] | |
| O.H. *csaüs* | [ t͡ʃ **æ y ʃ** ] | | O.H. *csunag* | [ t͡ʃ **u n a ɣ**] | |
| Cost 2.33 | N-align 4 | LD% 58.25 | Cost 1.83 | N-align 5 | LD% 36.6 |
| to wind, to turn, to distrort, misinterpret | | | grist, groats, soft hail | | |
| H. *csűr* | [ t͡ʃ **- - y:** r ] | | H. *dara* | [ d **ɒ** r **ɒ** ] | |
| O.H. *csevür* | [ t͡ʃ **ɛ v y** r ] | | O.H. *dara* | [ d **a** r **a** ] | |
| Cost 2.33 | N-align 5 | LD% 46.6 | Cost 0.66 | N-align 4 | LD% 16.5 |
| noon, south | | | nut, walnut | | |
| H. *dél* | [ d **e:** l **-** ] | | H. *dió* | [ **d** i **-** **o: -** ] | |
| O.H. *töli* | [ t **ø** l **i** ] | | O.H. *jigag* | [ **j** i **ɣ a ɣ** ] | |
| Cost 1.66 | N-align 4 | LD% 41.5 | Cost 3.57 | N-align 5 | LD% 71.4 |
| pig | | | to lean, topple over, stream down | | |
| H. *disznó* | [ **d** i s n **o: -** ] | | H. *dől* | [ d **ø** l ] | |
| O.H. *jisznag* | [ **j** i s n **a ɣ** ] | | O.H. *dül* | [ d **y** l ] | |
| Cost 3.07 | N-align 6 | LD% 51.17 | Cost 0.33 | N-align 3 | LD% 11 |
| to squeeze, to thrust into | | | church (building and organization) | | |
| H. *dug* | [ d **u** g ] | | H. *egy(ház)* | [ **ɛ ɟ** ] | |
| O.H. *dig* | [ d **ï** g ] | | O.H. *ed* | [ **e d** ] | |
| Cost 0.33 | N-align 3 | LD% 11 | Cost 0.66 | N-align 2 | LD% 33 |
| wedge | | | plough | | |
| H. *ék* | [ e: k ] | | H. *eke* | [ ɛ k ɛ **-** ] | |
| O.H. *ék* | [ e: k ] | | O.H. *ekeg* | [ ɛ k ɛ **ɣ** ] | |
| Cost 0 | N-align 2 | LD% 0 | Cost 0.5 | N-align 4 | LD% 12.5 |
| to allow, permit, concede, yield, give way | | | community work in agriculture | | |
| H. *eng(ed)* | [ ɛ ŋ g ] | | H. *enő* | [ **ɛ** n **ø: -** ] | |
| O.H. *eng* | [ ɛ ŋ g ] | | O.H. *ineg* | [ **i** n **ɛ ɣ**] | |
| Cost 0 | N-align 3 | LD% 0 | Cost 1.83 | N-align 4 | LD% 45.75 |
| to arrive, to reach, get to | | | merit | | |

| | | | | | |
|---|---|---|---|---|---|
| H. *ér* | [ **eː** r ] | | H. *érdem* | [ **eː** r d ɛ m ] | |
| O.H. *er* | [ **ɛ** r ] | | O.H. *erdem* | [ **ɛ** r d ɛ m ] | |
| Cost 0.66 | N-align 2 | LD% 33 | Cost 0.66 | N-align 5 | LD% 13.2 |
| morals, morality | | | to tire, lose vitality, slacken, relax | | |
| H. *erkölcs* | [ ɛ r k **ø** l t͡ʃ ] | | H. *erny(ed)* | [ ɛ r **-** **ɲ** ] | |
| O.H. *erkelcs* | [ ɛ r k **ɛ** l t͡ʃ ] | | O.H. *erin* | [ ɛ r **i** **n** ] | |
| Cost 0.66 | N-align 6 | LD% 11 | Cost 1.66 | N-align 4 | LD% 41.5 |
| strength, power | | | oath | | |
| H. *erő* | [ ɛ r **øː** **-** ] | | H. *eskü* | [ **ɛ** **ʃ** k y ] | |
| O.H. *erig* | [ ɛ r **i** **ɣ** ] | | O.H. *icskü* | [ **i** **t͡ʃ** k y ] | |
| Cost 1.5 | N-align 4 | LD% 37.5 | Cost 0.83 | N-align 4 | LD% 20.75 |
| reason, mind | | | river, name of a river | | |
| H. *ész* | [ **eː** s ] | | H. *etel* | [ ɛ t **ɛ** l ] | |
| O.H. *esz* | [ **ɛ** s ] | | O.H. *etil* | [ ɛ t **i** l ] | |
| Cost 0.66 | N-align 2 | LD% 33 | Cost 0.33 | N-align 4 | LD% 8.25 |
| weed | | | polecat | | |
| H. *gaz* | [ **g** **ɒ** **-** z **-** ] | | H. *görény* | [ g **ø** r **eː** **ɲ** ] | |
| O.H. *kagzu* | [ **k** **a** **ɣ** z **u**] | | O.H. *güren* | [ g **y** r **ɛ** **n** ] | |
| Cost 2.50 | N-align 5 | LD% 50 | Cost 1.66 | N-align 5 | LD% 33.2 |
| pelican | | | gleaner mouse | | |
| H. *gödény* | [ g **ø** d **eː** **ɲ** ] | | H. *güzü* | [ **g** y **z** **-** **-** y ] | |
| O.H. *güden* | [ g **y** d **ɛ** **n** ] | | O.H. *küsegü* | [ **k** y **s** **ɛ** **ɣ** y ] | |
| Cost 1.66 | N-align 5 | LD% 33.2 | Cost 2.66 | N-align 6 | LD% 44.33 |
| to abuse, calumniate | | | drag-net, a kind of fishing net | | |
| H. *gyalá(z)* | [ **ɟ** **ɒ** l **aː** ] | | H. *gyalom* | [ **ɟ** **ɒ** l **o** m ] | |
| O.H. *jala* | [ **j** **a** l **a** ] | | O.H. *jalum* | [ **j** **a** l **u** m ] | |
| Cost 1.91 | N-align 4 | LD% 47,75 | Cost 1.91 | N-align 5 | LD% 38.2 |
| suspicion, mistrust | | | wool | | |
| H. *gyanú* | [ **ɟ** **ɒ** n **uː** **-** ] | | H. *gyapjú* | [ **ɟ** **ɒ** p **j** **uː** ] | |
| O.H. *jonag* | [ **j** **o** n **a** **ɣ** ] | | O.H. *japgu* | [ **j** **a** p **ɣ** **u** ] | |

| Cost 3.08 | N-align 5 | LD% 61.6 | Cost 2.73 | N-align 5 | LD% 54.6 |
|---|---|---|---|---|---|
| cotton | | | to increase, to put on weight | | |
| H. *gyapot* | [ ɟ ɒ p **o** t ] | | H. *gyarap(od)(ik)* | [ ɟ ɒ r ɒ p ] | |
| O.H. *japut* | [ **j a** p **u** t ] | | O.H. *jarpa* | [ **j a** r p **a** ] | |
| Cost 1.91 | N-align 5 | LD% 38.2 | Cost 2.91 | N-align 5 | LD% 58.2 |
| poor, frail, feeble | | | to produce, build, fabricate | | |
| H. *gyarló* | [ ɟ ɒ r l **o: -** ] | | H. *gyárt* | [ ɟ **a:** r **-** t ] | |
| O.H. *jarlig* | [ **j a** r l **ï ɣ** ] | | O.H. jarat | [ **j a** r **a** t ] | |
| Cost 3.08 | N-align 6 | LD% 51.33 | Cost 2.58 | N-align 5 | LD% 51.6 |
| mourning, bereavement | | | bulrush | | |
| H. *gyász* | [ ɟ **a:** s ] | | H. *gyékény* | [ ɟ **e:** k **e: ɲ** ] | |
| O.H. *jasz* | [ **j a** s ] | | O.H. *gyékény* | [ **j ɛ** k **ɛ n** ] | |
| Cost 1.57 | N-align 3 | LD% 52.33 | Cost 3.25 | N-align 5 | LD% 65 |
| weak, feeble | | | rein | | |
| H. *gyenge* | [ ɟ ɛ ŋ g ɛ ] | | H. *gyeplő* | [ ɟ **ɛ** p l **ø: -** ] | |
| O.H. *jenge* | [ **j** ɛ ŋ g ɛ ] | | O.H. *jiplig* | [ **j i** p l **i ɣ** ] | |
| Cost 1.25 | N-align 5 | LD% 25 | Cost 3.08 | N-align 6 | LD% 51.33 |
| borderland, hedge | | | child | | |
| H. *gyepű* | [ ɟ ɛ p **y: -** ] | | H. *gyermek* | [ ɟ **ɛ** r m **ɛ** k ] | |
| O.H. *jepig* | [ **j** ɛ p **i ɣ** ] | | O.H. *järmäk* | [ **j æ** r m **æ** k ] | |
| Cost 2.41 | N-align 5 | LD% 48.2 | Cost 1.91 | N-align 6 | LD% 31.83 |
| candle | | | hornbeam | | |
| H. *gyertya* | [ ɟ **ɛ** r **c** ɒ ] | | H. *gyertyán* | [ ɟ **ɛ** r t **- j a:** n ] | |
| O.H. *jarta* | [ **j a** r **t a** ] | | O.H. *jartagan* | [ **j a** r t **a ɣ a** n ] | |
| Cost 2.9 | N-align 5 | LD% 58 | Cost 4.07 | N-align 8 | LD% 50.87 |
| murderer, killer | | | weed | | |
| H. *gyilkos* | [ ɟ i l k **o** ʃ ] | | H. *gyom* | [ ɟ o m ] | |
| O.H. *jilkus* | [ **j** i l k **u** ʃ ] | | O.H. *jom* | [ **j** o m ] | |
| Cost 4.07 | N-align 6 | LD% 26.33 | Cost 1.25 | N-align 3 | LD% 41.67 |
| stomach | | | to confess (sins) | | |

| | | | | | |
|---|---|---|---|---|---|
| H. *gyomor* | [ ɟ o m o r ] | | H. *gyón(ik)* | [ ɟ o: - - n ] | |
| O.H. *jumur* | [ j u m u r ] | | O.H. *juvun* | [ j u v u n ] | |
| Cost 1.91 | N-align 5 | LD% 38.2 | Cost 3.91 | N-align 5 | LD% 78.2 |
| cudweed | | | pearl | | |
| H. *gyopár* | [ ɟ o p a: r ] | | H. *gyöngy* | [ ɟ ø n ɟ - ] | |
| O.H. *jipar* | [ j ï p a r ] | | O.H. *jinjü* | [ j i n j y ] | |
| Cost 2.25 | N-align 5 | LD% 45 | Cost 4.16 | N-align 5 | LD% 83.2 |
| to torture, to make suffer | | | to conquer, triumph, gain victory | | |
| H. *gyötör* | [ ɟ ø t ø r ] | | H. *győz* | [ ɟ ø: - - z ] | |
| O.H. *jitür* | [ j i t y r ] | | O.H. *jegüz* | [ j ɛ ɣ y z ] | |
| Cost 2.25 | N-align 5 | LD% 45 | Cost 3.75 | N-align 5 | LD% 75 |
| to catch fire, be kindled, to be ignited | | | to knead, pug | | |
| H. *gyúl* | [ ɟ u: l ] | | H. *gyúr* | [ ɟ u: - - r ] | |
| O.H. *jul* | [ j u l ] | | O.H. *jugur* | [ j u ɣ u r ] | |
| Cost 1.58 | N-align 3 | LD% 52.67 | Cost 3.08 | N-align 5 | LD% 61.6 |
| assemble, gather | | | to hate | | |
| H. *gyűl(ik)* | [ ɟ y: - - l ] | | H. *gyűlö(l)* | [ ɟ y: - - l ø ] | |
| O.H. *jigil* | [ j i ɣ i l ] | | O.H. *jegile* | [ j ɛ ɣ i l ɛ ] | |
| Cost 3.41 | N-align 5 | LD% 68.2 | Cost 4.41 | N-align 6 | LD% 73.5 |
| fruit | | | hawthorn, a plant similar to the whitethorn | | |
| H. *gyümölcs* | [ ɟ y m ø l t͡ʃ ] | | H. *gyümölcsény* | [ ɟ y m ø l t͡ʃ e: ɲ ] | |
| O.H. *jemilcs* | [ j e m i l t͡ʃ ] | | O.H. *jemilcsen* | [ j e m i l t͡ʃ e n ] | |
| Cost 2.57 | N-align 6 | LD% 42.83 | Cost 3.58 | N-align 8 | LD% 44.75 |
| a kind of tree similar to the maple or cornel | | | ring | | |
| H. *gyűrű* | [ ɟ y: r y: - ] | | H. *gyűrű* | [ ɟ y: r y: - ] | |
| O.H. *jereg* | [ j e r e ɣ ] | | O.H. *jürüg* | [ j y r y ɣ ] | |
| Cost 3.75 | N-align 5 | LD% 75 | Cost 2.41 | N-align 5 | LD% 48.2 |
| thimble | | | boat, ship | | |
| H. *gyűszű* | [ ɟ y: - s y: - ] | | H. *hajó* | [ h ɒ j o: - ] | |
| O.H. *jügszüg* | [ j y ɣ s y ɣ ] | | O.H. *hajig* | [ h a j i ɣ ] | |

| Cost 2.91 | N-align 6 | LD% 48.5 | Cost 2.16 | N-align 5 | LD% 50 |
|---|---|---|---|---|---|
| bell | | | hawk | | |
| H. *harang* | [ h **ɒ** r **ɒ** ŋ g ] | | H. *herjó* | [ **h ɛ** r **j o: -** ] | |
| O.H. *hangar* | [ h **a** ŋ g **a** r ] | | O.H. *hirguj* | [ **χ i** r **ɣ u j** ] | |
| Cost 2.66 | N-align 6 | LD% 44,33 | Cost 2.64 | N-align 6 | LD% 44 |
| sand | | | noose, loop, snare | | |
| H. *homok* | [ h **o** m o k **-** ] | | H. *hurok* | [ **h** u r o k ] | |
| O.H. *humoki* | [ h **u** m o k **i** ] | | O.H. *urok* | [ **-** u r o k ] | |
| Cost 1.33 | N-align 6 | LD% 22.17 | Cost 1 | N-align 5 | LD% 20 |
| time, weather | | | yes, affirmative particle, very | | |
| H. *idő* | [ **i** d **ø: -** ] | | H. *igen* | [ **i** g ɛ n ] | |
| O.H. *üdeg* | [ **y** d **ɛ ɣ** ] | | O.H. *egen* | [ **ɛ** g ɛ n ] | |
| Cost 1.83 | N-align 4 | LD% 45.75 | Cost 0.33 | N-align 4 | LD% 8.25 |
| to frighten | | | twin | | |
| H. *ije(szt)* | [ **i** j **ɛ** ] | | H. *iker* | [ i k **ɛ** r ] | |
| O.H. *äji* | [ **æ** j **i** ] | | O.H. *ikir* | [ i k **i** r ] | |
| Cost 0.66 | N-align 3 | LD% 22 | Cost 0.33 | N-align 4 | LD% 8.25 |
| proper behaviour | | | to suit something, to be proper, to fit into | | |
| H. *ildom* | [ i l d **o** m ] | | H. *ill(ik)* | [ i **l:** ] | |
| O.H. *ildam* | [ i l d **a** m ] | | O.H. *il* | [ i **l** ] | |
| Cost 0.66 | N-align 5 | LD% 13.2 | Cost 0.33 | N-align 2 | LD% 16.5 |
| to adore, worship | | | shirt | | |
| H. *imá(d)* | [ **-** i m **a:** ] | | H. *ing* | [ **i n -** g ] | |
| O.H. *vim* | [ **v** i m **-** ] | | O.H. *ümmeg* | [ **y m: ɛ** g ] | |
| Cost 2 | N-align 4 | LD% 50 | Cost 2.33 | N-align 4 | LD% 58,25 |
| to write | | | buttermilk | | |
| H. *ír* | [ **i:** r ] | | H. *író* | [ **i:** r **o: -** ] | |
| O.H. *ir* | [ **i** r ] | | O.H. *irag* | [ **i** r **a ɣ** ] | |
| Cost 0.33 | N-align 2 | LD% 16.5 | Cost 1.83 | N-align 4 | LD% 45.75 |

| to glow, be hot | | | name of an ethnic group with Iranian origins in Hungary | | |
|---|---|---|---|---|---|
| H. *izz(ik)* | [ i **z:** i ] | | H. *jász* | [ j a: s ] | |
| O.H. *iszi* | [ i **s** i ] | | O.H. *jász* | [ j a: s ] | |
| Cost 0.83 | N-align 3 | LD% 27.67 | Cost 0 | N-align 3 | LD% 0 |
| bulrush, club-rush | | | a muslim ethnic group in medieval Hungary | | |
| H. *káka* | [ k a: k **ɒ** ] | | H. *káliz* | [ k **a:** l i z ] | |
| O.H. *káka* | [ k a: k **a** ] | | O.H. *kaliz* | [ k **a** l i z ] | |
| Cost 0.33 | N-align 4 | LD% 8.25 | Cost 0.33 | N-align 5 | LD% 6.6 |
| a water plant | | | bridle, reins | | |
| H. *kalokány* | [ k **ɒ -** l **o** k **a: ɲ** ] | | H. *kantár* | [ k ɒ n t a: r ] | |
| O.H. *karlukan* | [ k **a r** l **u** k **a n** ] | | O.H. *kantár* | [ k ɒ n t a: r ] | |
| Cost 2.66 | N-align 8 | LD% 33.25 | Cost 0 | N-align 6 | LD% 0 |
| measles | | | trap, snare | | |
| H. *kanyaró* | [ k **ɒ ɲ ɒ r o: -** ] | | H. *kaptány* | [ k **ɒ** p k **a: ɲ** ] | |
| O.H. *karamug* | [ k **a r a m u ɣ** ] | | O.H. *kapkan* | [ k **a** p k **a n** ] | |
| Cost 4.14 | N-align 7 | LD% 59.14 | Cost 1.33 | N-align 6 | LD% 22.17 |
| gate | | | arm | | |
| H. *kapu* | [ k **ɒ** p u **-** ] | | H. *kar* | [ k **ɒ** r ] | |
| O.H. *kapug* | [ k **a** p u **ɣ** ] | | O.H. *kar* | [ k **a** r ] | |
| Cost 0.83 | N-align 5 | LD% 16.6 | Cost 0.33 | N-align 3 | LD% 11 |
| damage, loss | | | man of grit, stout fellow | | |
| H. *kár* | [ k **a:** r ] | | H. *karakán* | [ k **ɒ** r **ɒ** k **a:** n ] | |
| O.H. *kor* | [ k **o** r ] | | O.H. *karakan* | [ k **a** r **a** k **a** n ] | |
| Cost 1 | N-align 3 | LD% 33.33 | Cost 1 | N-align 7 | LD% 14.29 |
| sheepfold, cattlegrid, stockyard | | | stake, pale, stick | | |
| H. *karám* | [ k **ɒ** r a: m ] | | H. *karó* | [ k **ɒ** r **o: -** ] | |
| O.H. *korám* | [ k **o** r a: m ] | | O.H. *karog* | [ k **a** r **o ɣ** ] | |
| Cost 0.33 | N-align 5 | LD% 6.6 | Cost 1.16 | N-align 5 | LD% 23.2 |
| a fishing water bird with black feathers | | | sparrow hawk | | |

| | | | | | |
|---|---|---|---|---|---|
| H. *kárókatona* | [ k **aː** r **oː** k **ɒ** t **o** n **ɒ** ] | | H. *karvaly* | [ k **ɒ** r **v** **ɒ** j ] | |
| O.H. *kara kotan* | [ k **a** r **a** k **o** t **a** n **-** ] | | O.H. *kargaj* | [ k **a** r **ɣ** **a** j ] | |
| Cost 3.33 | N-align 10 | LD% 33.33 | Cost 1.33 | N-align 6 | LD% 22.17 |
| chicory | | | an ethnic name, Kazar | | |
| H. *katáng* | [ k **ɒ** t aː ŋ g ] | | H. *kazár* | [ k **ɒ** z **aː** r ] | |
| O.H. *katáng* | [ k **a** t aː ŋ g ] | | O.H. *kazar* | [ k **a** z **a** r ] | |
| Cost 0.33 | N-align 6 | LD% 5.5 | Cost 0.66 | N-align 5 | LD% 13.2 |
| bosom, breast | | | goat | | |
| H. *kebel* | [ k ɛ b ɛ l ] | | H. *kecske* | [ k ɛ t͡ʃ **-** k ɛ ] | |
| O.H. *kebel* | [ k ɛ b ɛ l ] | | O.H. *kecseke* | [ k ɛ t͡ʃ **ɛ** k ɛ ] | |
| Cost 0 | N-align 5 | LD% 0 | Cost 1 | N-align 6 | LD% 16.67 |
| blue | | | trousseau, dowry, gift given with the bride | | |
| H. *kék* | [ k **eː** **-** k ] | | H. *kelengye* | [ k **ɛ** l **ɛ** n **ɟ** **ɛ** **-** ] | |
| O.H. *keyk* | [ k **ɛ** **j** k ] | | O.H. *kälinçäğ* | [ k **æ** l **i** n **t͡ʃ** **æ** **ɣ** ] | |
| Cost 1.16 | N-align 4 | LD% 29 | Cost 2.58 | N-align 8 | LD% 32.25 |
| hemp | | | mercury | | |
| H. *kender* | [ k ɛ n d **ɛ** r ] | | H. *kéneső* | [ k **eː** n ɛ ʃ **øː** ] | |
| O.H. *kendir* | [ k ɛ n d **i** r ] | | O.H. *kenesű* | [ k **ɛ** n ɛ ʃ **yː** ] | |
| Cost 0.33 | N-align 6 | LD% 5.5 | Cost 1 | N-align 6 | LD% 16.67 |
| picture, shape, form | | | shock, shook, sheaves placed crosswise | | |
| H. *kép* | [ k eː p ] | | H. *kepe* | [ k ɛ p ɛ **-** ] | |
| O.H. *kép* | [ k eː p ] | | O.H. *kepeg* | [ k ɛ p ɛ **ɣ** ] | |
| Cost 0 | N-align 3 | LD% 0 | Cost 0.5 | N-align 5 | LD% 10 |
| a kind of ship | | | to ruminate (of bovines) | | |
| H. *kerep* | [ k ɛ r ɛ p ] | | H. *kér(ődz)*(ik) | [ k eː r ] | |
| O.H. *kerep* | [ k ɛ r ɛ p ] | | O.H. *kér* | [ k eː r ] | |
| Cost 0 | N-align 5 | LD% 0 | Cost 0 | N-align 3 | LD% 0 |
| garden, place for animals | | | vulture | | |
| H. *kert* | [ k ɛ r t **-** ] | | H. *keselyű* | [ k **ɛ** **ʃ** ɛ **j** **yː** **-** ] | |
| O.H. *kerte* | [ k ɛ r t **ɛ** ] | | O.H. *kücseleg* | [ k **y** **t͡ʃ** ɛ **l** **ɛ** **ɣ** ] | |

| | | | | | |
|---|---|---|---|---|---|
| Cost 0.33 | N-align 5 | LD% 6.6 | Cost 3.91 | N-align 7 | LD% 55.86 |
| to be, become late | | | small, little | | |
| H. *kés(ik)* | [ k e: **ʃ** ] | | H. *kicsiny* | [ k **i** t͡ʃ **i** **ɲ** ] | |
| O.H. *kécs* | [ k e: **t͡ʃ** ] | | O.H. *kücsün* | [ k **y** t͡ʃ **y** **n** ] | |
| Cost 0.5 | N-align 3 | LD% 16.67 | Cost 1.33 | N-align 5 | LD% 26.6 |
| wild/meadow saffron, lion's tooth | | | a bird similar to a falcon | | |
| H. *kikerics* | [ k **i** k ɛ r i **-** t͡ʃ ] | | H. *kiköcsén* | [ k **i** k **ø** **-** t͡ʃ **e:** n ] | |
| O.H. *kükerilcs* | [ k **y** k ɛ r i **l** t͡ʃ ] | | O.H. *kükercsen* | [ k **y** k **ɛ** **r** t͡ʃ **ɛ** n ] | |
| Cost 1.33 | N-align 8 | LD% 16.62 | Cost 2.66 | N-align 8 | LD% 33.25 |
| pain | | | to beg, to mendicate | | |
| H. *kín* | [ k i: n ] | | H. *koldul* | [ k o l d u **l** ] | |
| O.H. *kín* | [ k i: n ] | | O.H. *koldu* | [ k o l d u **-** ] | |
| Cost 0 | N-align 3 | LD% 0 | Cost 1 | N-align 6 | LD% 16.67 |
| hops, *humulus lupulus* | | | timothy-grass, phleum | | |
| H. *komló* | [ k **o** m l **o:** **-** ] | | H. *komócsin* | [ k **o** m **o:** t͡ʃ i n ] | |
| O.H. *kumlag* | [ k **u** m l **a** **ɣ** ] | | O.H. *kamicsin* | [ k **ɒ** m **i** t͡ʃ i n ] | |
| Cost 1.83 | N-align 6 | LD% 30.5 | Cost 1.33 | N-align 7 | LD% 19 |
| gloomy, grave, morose, somber, dull-coloured | | | coffin | | |
| H. *komor* | [ k o m **o** r ] | | H. *koporsó* | [ k o p **o** r **ʃ** **o:** **-** ] | |
| O.H. *komur* | [ k o m **u** r ] | | O.H. *kopurcsag* | [ k o p **u** r **t͡ʃ** **a** **ɣ** ] | |
| Cost 0.33 | N-align 5 | LD% 6.6 | Cost 2.33 | N-align 8 | LD% 29.12 |
| age, period | | | tomb, small hill, elevated, dry place in a marsh | | |
| H. *kor* | [ k **o** r ] | | H. *korhány* | [ k o r **h** **a:** **ɲ** ] | |
| O.H. *kur* | [ k **u** r ] | | O.H. *korgan* | [ k o r **ɣ** **a** **n** ] | |
| Cost 0.33 | N-align 3 | LD% 11 | Cost 2.66 | N-align 6 | LD% 44.33 |
| dry stalk of weed | | | soot | | |
| H. *kóró* | [ k **o:** r **o:** **-** ] | | H. *korom* | [ k **o** r **o** m ] | |
| O.H. *kuroğ* | [ k **u** r **o** **ɣ** ] | | O.H. *kurum* | [ k **u** r **u** m ] | |
| Cost 1.5 | N-align 5 | LD% 30 | Cost 0.66 | N-align 5 | LD% 13.2 |

| ram | | | burn | | |
|---|---|---|---|---|---|
| H. *kos* | [ k o **ʃ** ] | | H. *kozma* | [ k **o z** m **ɒ -** ] | |
| O.H. *kocs* | [ k o **t͡ʃ** ] | | O.H. *kaszmag* | [ k **a s** m **a ɣ** ] | |
| Cost 0.5 | N-align 3 | LD% 16.67 | Cost 2 | N-align 6 | LD% 33.33 |
| frock, sheepskin waistcoat | | | blackthorn | | |
| H. *ködmön* | [ k **ø** d m **ø** n ] | | H. *kökény* | [ k ø k **e: ɲ** ] | |
| O.H. *kedmen* | [ k **ɛ** d m **ɛ** n ] | | O.H. *köken* | [ k ø k **ɛ n** ] | |
| Cost 1.33 | N-align 6 | LD% 22.17 | Cost 1.33 | N-align 5 | LD% 26.6 |
| loan | | | navel, various parts of agricultural instruments | | |
| H. *kölcsön* | [ k **ø** l **t͡ʃ ø** n ] | | H. *köldök* | [ k **ø l** d **ø** k ] | |
| O.H. *külcsen* | [ k **y** l **t͡ʃ ɛ** n ] | | O.H. *kindik* | [ k **i n** d **i** k ] | |
| Cost 1 | N-align 6 | LD% 16.67 | Cost 2.07 | N-align 6 | LD% 34.5 |
| young of an animal, kid, puppy, lad | | | pounder, beater, small mortar | | |
| H. *kölyök* | [ k **ø j ø** k ] | | H. *kölyű* | [ k **ø j y: -** ] | |
| O.H. *kylek* | [ k **y l ɛ** k ] | | O.H. *kelig* | [ k **ɛ l i ɣ** ] | |
| Cost 2.25 | N-align 5 | LD% 45 | Cost 3.08 | N-align 5 | LD% 61.6 |
| book | | | stumpy, a small and thick man | | |
| H. *könyv* | [ k **ø** ɲ **- v** ] | | H. *köpcös* | [ k ø p **t͡s ø ʃ** ] | |
| O.H. *künüg* | [ k **y** ɲ **y ɣ** ] | | O.H. *köpcseg* | [ k ø p **t͡ʃ ɛ ɣ** ] | |
| Cost 2 | N-align 5 | LD% 40 | Cost 2.15 | N-align 6 | LD% 35.83 |
| a kind of Hungarian sheepdog | | | cloak, gown, overcoat | | |
| H. *köpec* | [ k ø p ɛ **t͡ʃ** ] | | H. *köpönyeg* | [ k **ø** p **ø ɲ** ɛ **g** ] | |
| O.H. *köpek* | [ k ø p ɛ **k** ] | | O.H. *kepenek* | [ k **ɛ** p **ɛ n** ɛ **k** ] | |
| Cost 0.91 | N-align 5 | LD% 18.2 | Cost 2.66 | N-align 7 | LD% 38 |
| churn, beehive | | | ash tree | | |
| H. *köpű* | [ k **ø** p **y: -** ] | | H. *kőris* | [ k **ø: -** r i **ʃ** ] | |
| O.H. *küpüg* | [ k **y** p **y ɣ** ] | | O.H. *kewrics* | [ k **ɛ w** r i **t͡ʃ** ] | |
| Cost 1.16 | N-align 5 | LD% 23.2 | Cost 2 | N-align 6 | LD% 33.33 |
| mellow, powdery, crumbly, rotten, mouldy | | | pear | | |
| H. *kőrő* | [ k **ø: -** r **ø: -** ] | | H. *körtvély* | [ k **ø** r t **v e: - - j** ] | |

| | | | | | |
|---|---|---|---|---|---|
| O.H. *kevreg* | [ k **ɛ v** r **ɛ ɣ** ] | | O.H. *kertmelig* | [ k **ɛ** r t **m ɛ l i ɣ** ] | |
| Cost 3.5 | N-align 6 | LD% 58.33 | Cost 4.97 | N-align 9 | LD% 55.22 |
| to greet, to thank | | | Cuman ethnic group in Hungary | | |
| H. *köszön* | [ k **ø** s **ø** n ] | | H. *kun* | [ k u **- -** n ] | |
| O.H. *küszen* | [ k **y** s **ɛ** n ] | | O.H. *kuwan* | [ k u **w a** n ] | |
| Cost 1 | N-align 5 | LD% 20 | Cost 1.5 | N-align 5 | LD% 30 |
| laughing bird | | | a fish living at the bottom of the river | | |
| H. *küllő* | [ k y l: **ø: -** ] | | H. *küllő* | [ k **y l: ø: - - - - -** ] | |
| O.H. *külleg* | [ k y l: **ɛ ɣ** ] | | O.H. *köligelig* | [ k **ø l i g ɛ l i ɣ** ] | |
| Cost 1.5 | N-align 5 | LD% 30 | Cost 6.16 | N-align 9 | LD% 68.44 |
| spoke (of a wheel) | | | cheap | | |
| H. *küllő* | [ k y **l: ø: -** ] | | H. *olcsó* | [ **o l** t͡ʃ **o: -** ] | |
| O.H. *küvey* | [ k y **v ɛ j** ] | | O.H. *ucsag* | [ **u -** t͡ʃ **a ɣ** ] | |
| Cost 3 | N-align 5 | LD% 60 | Cost 2.83 | N-align 5 | LD% 56.6 |
| kid (of a goat) | | | sty, cattle pen, sheepfold | | |
| H. *olló* | [ o **- l: o: -** ] | | H. *ól* | [ **o: - -** l ] | |
| O.H. *oglag* | [ o **ɣ l a ɣ** ] | | O.H. *agul* | [ **a ɣ u** l ] | |
| Cost 2.33 | N-align 5 | LD% 46.6 | Cost 2.5 | N-align 4 | LD% 62.5 |
| cheap | | | kid (of a goat) | | |
| H. *olcsó* | [ **o l** t͡ʃ **o: -** ] | | H. *olló* | [ o **- l: o: -** ] | |
| O.H. *ucsag* | [ **u -** t͡ʃ **a ɣ** ] | | O.H. *oglag* | [ o **ɣ l a ɣ** ] | |
| Cost 2.83 | N-align 5 | LD% 56.6 | Cost 2.33 | N-align 5 | LD% 46.6 |
| sperm, seminal fluid | | | Russian | | |
| H. *ondó* | [ **o** n d **o: -** ] | | H. *orosz* | [ **o** r **o** s ] | |
| O.H. *undag* | [ **u** n d **a ɣ** ] | | O.H. *urusz* | [ **u** r **u** s ] | |
| Cost 1.83 | N-align 5 | LD% 36.6 | Cost 0.66 | N-align 4 | LD% 16.5 |
| lion | | | spindle, whorl | | |
| H. *oroszlán* | [ **o** r **o** s l a: n ] | | H. *orsó* | [ **o** r **ʃ o: -** ] | |
| O.H. *uruszlán* | [ **u** r **u** s l a: n ] | | O.H. *urcsug* | [ **u** r **t͡ʃ u ɣ** ] | |
| Cost 0.66 | N-align 7 | LD% 9.43 | Cost 2 | N-align 5 | LD% 40 |

| thief | | | physician | | |
|---|---|---|---|---|---|
| H. *orv* | [ **o** - r **v** ] | | H. *orvos* | [ **o** r v o **ʃ** - ] | |
| O.H. *ugru* | [ **u ɣ** r **u** ] | | O.H. *urvocsi* | [ **u** r v o **t͡ʃ ï** ] | |
| Cost 2.66 | N-align 4 | LD% 66.5 | Cost 1.83 | N-align 6 | LD% 30.5 |
| ‘cune in play’ in ostábla | | | ox | | |
| H. *os* | [ **o** - **ʃ** - - ] | | H. *ökör* | [ **ø** k **ø** r ] | |
| O.H. *alcsuk* | [ **a l t͡ʃ u k** ] | | O.H. *ükür* | [ **y** k **y** r ] | |
| Cost 4.16 | N-align 5 | LD% 83.2 | Cost 0.66 | N-align 4 | LD% 16.5 |
| to stitch, to put on a dress | | | to elect, to select | | |
| H. *ölt* | [ **ø** l t ] | | H. *ön(ik)* | [ **ø** n ] | |
| O.H. *ilt* | [ **i** l t ] | | O.H. *ün* | [ **y** n ] | |
| Cost 0.66 | N-align 3 | LD% 22 | Cost 0.33 | N-align 2 | LD% 16.5 |
| to become mad | | | hawk, buzzard | | |
| H. *őrül* | [ **ø:** - r **y** l ] | | H. *ölyv* | [ **ø** - **y v** ] | |
| O.H. *ävril* | [ **æ v** r **i** l ] | | O.H. *elig* | [ **ɛ l i ɣ** ] | |
| Cost 2.66 | N-align 4 | LD% 66.5 | Cost 2.33 | N-align 5 | LD% 46.6 |
| to grind | | | Armenian | | |
| H. *őr(öl)* | [ **ø:** - - r ] | | H. *örmény* | [ **ø** r m **e: ɲ** - ] | |
| O.H. *evir* | [ **ɛ v i** r ] | | O.H. *ermeni* | [ **ɛ** r m **ɛ n i** ] | |
| Cost 3 | N-align 4 | LD% 75 | Cost 3 | N-align 6 | LD% 50 |
| eternal | | | to rejoice, to be glad | | |
| H. *örök* | [ **ø** r **ø** k ] | | H. *örül* | [ ø **r y l** ] | |
| O.H. *ürük* | [ **y** r **y** k ] | | O.H. *ögir* | [ ø **ɣ i r** ] | |
| Cost 0.66 | N-align 4 | LD% 16.5 | Cost 1.58 | N-align 4 | LD% 39.5 |
| whirlpool | | | cheese | | |
| H. *ör(vény)* | [ **ø** - - r ] | | H. *sajt* | [ **ʃ ɒ j** t ] | |
| O.H. *egir* | [ **ɛ ɣ i** r ] | | O.H. *csagt* | [ **t͡ʃ a ɣ** t ] | |
| Cost 1.83 | N-align 4 | LD% 45.75 | Cost 1.65 | N-align 4 | LD% 41.25 |
| mud, marsh | | | yellow | | |
| H. *sár* | [ ʃ **a:** r ] | | H. *sárga* | [ ʃ **a:** r g **ɒ** ] | |

| | | | | | |
|---|---|---|---|---|---|
| O.H. *sar* | [ ʃ **a** r ] | | O.H. *sarug* | [ ʃ **a** r **u** g ] | |
| Cost 0.33 | N-align 3 | LD% 11 | Cost 1.66 | N-align 5 | LD% 33.2 |
| dragon | | | sickle | | |
| H. *sárkány* | [ ʃ a: r k a: **ɲ** ] | | H. *sarló* | [ **ʃ ɒ** r l **o: -** ] | |
| O.H. *sárkán* | [ ʃ a: r k a: **n** ] | | O.H. *csarlag* | [ **t͡ʃ a** r l **a ɣ** ] | |
| Cost 0.66 | N-align 7 | LD% 11 | Cost 2.33 | N-align 6 | LD% 38.83 |
| a type of wild onion | | | footwear, sandals | | |
| H. *sárma* | [ ʃ **a:** r m **ɒ -** ] | | H. *saru* | [ **ʃ ɒ** r u **-** ] | |
| O.H. *sarmag* | [ ʃ **a** r m **a ɣ** ] | | O.H. *csarug* | [ **t͡ʃ a** r u **ɣ** ] | |
| Cost 1.16 | N-align 6 | LD% 19.33 | Cost 1.33 | N-align 5 | LD% 26.6 |
| tent | | | fast, quick | | |
| H. *sátor* | [ **ʃ a:** t **o** r ] | | H. *seb(es)* | [ ʃ ɛ b ] | |
| O.H. *csatur* | [ **t͡ʃ a** t **u** r ] | | O.H. *seb* | [ ʃ ɛ b ] | |
| Cost 1.16 | N-align 5 | LD% 23.2 | Cost 0 | N-align 3 | LD% 0 |
| | | | to sweep, to broom | | |
| | | | H. *söpör* | [ ʃ **ø** p **ø** r ] | |
| | | | O.H. *sipir* | [ ʃ **i** p **i** r ] | |
| | | | Cost 1.33 | N-align 5 | LD% 26.6 |
| broom | | | draff, lees, dregs of wine | | |
| H. *seprő* | [ ʃ **ɛ** p **-** r **ø: -** ] | | H. *seprő* | [ **ʃ ɛ** p r **ø: -** ] | |
| O.H. *sipirig* | [ ʃ **i** p **i** r **i ɣ** ] | | O.H. *csöpreg* | [ **t͡ʃ ø** p r **ɛ ɣ** ] | |
| Cost 2.83 | N-align 7 | LD% 40.43 | Cost 2.66 | N-align 6 | LD% 44.33 |
| to injure, to hurt | | | army, troops, crowd | | |
| H. *sér(t)* | [ ʃ **e:** r ] | | H. *sereg* | [ **ʃ ɛ** r **ɛ** g ] | |
| O.H. *sir* | [ ʃ **i** r ] | | O.H. *csarig* | [ **t͡ʃ æ** r **i** g ] | |
| Cost 0.66 | N-align 3 | LD% 22 | Cost 1.16 | N-align 5 | LD% 23.2 |
| a nit | | | cup, goblet | | |
| H. *serke* | [ ʃ **ɛ** r k ɛ ] | | H. *serleg* | [ ʃ ɛ r l ɛ g ] | |
| O.H. *sirke* | [ ʃ **i** r k ɛ ] | | O.H. *serleg* | [ ʃ ɛ r l ɛ g ] | |
| Cost 0.33 | N-align 5 | LD% 6.6 | Cost 0 | N-align 6 | LD% 0 |

| bristle | | | even, flat, open (water) | | |
|---|---|---|---|---|---|
| H. *serte* | [ ʃ **ɛ** r t **ɛ** ] | | H. *sík* | [ ʃ **i**: k ] | |
| O.H. *sirt* | [ ʃ **i** r t **-** ] | | O.H. *sik* | [ ʃ **i** k ] | |
| Cost 1.33 | N-align 5 | LD% 26.6 | Cost 0.33 | N-align 3 | LD% 11 |
| smooth | | | running water, running channel under a mill | | |
| H. *sima* | [ ʃ i m **ɒ** ] | | H. *sió* | [ ʃ i **o:** ] | |
| O.H. *sima* | [ ʃ i m **a** ] | | O.H. *siu* | [ ʃ i **u** ] | |
| Cost 0.33 | N-align 4 | LD% 8.25 | Cost 0.66 | N-align 3 | LD% 22 |
| much, many | | | falcon | | |
| H. *sok* | [ ʃ o k ] | | H. *sólyo(m)* | [ **ʃ o: j - o** ] | |
| O.H. *sok* | [ ʃ o k ] | | O.H. *csavli* | [ **t͡ʃ a v l i** ] | |
| Cost 0 | N-align 3 | LD% 0 | Cost 5.57 | N-align 5 | LD% 111.4 |
| dogwood (tree), cornel | | | beer | | |
| H. *som* | [ **ʃ o** m ] | | H. *sör* | [ ʃ **ø** r **-** ] | |
| O.H. *csum* | [ **t͡ʃ u** m ] | | O.H. *sere* | [ ʃ **ɛ** r **ɛ** ] | |
| Cost 0.83 | N-align 3 | LD% 27.67 | Cost 1.66 | N-align 4 | LD% 41.5 |
| starlet, sturgeon | | | pike perch, zander | | |
| H. *sőreg* | [ ʃ **ø: -** r ɛ g **-** ] | | H. *süllő* | [ ʃ **y** l: **ø: -** ] | |
| O.H. *sevrege* | [ ʃ **ɛ** v r ɛ g **ɛ** ] | | O.H. *silleg* | [ ʃ **i** l: **ɛ ɣ** ] | |
| Cost 3 | N-align 7 | LD% 42.86 | Cost 1.83 | N-align 5 | LD% 36.6 |
| scurvy | | | dense, thick (of woods or soup) | | |
| H. *süly* | [ ʃ **y j - -** ] | | H. *sűrű* | [ **ʃ y:** r **y: -** ] | |
| O.H. *sigül* | [ ʃ **i ɣ y l** ] | | O.H. *csireg* | [ **t͡ʃ i** r **ɛ ɣ** ] | |
| Cost 3.15 | N-align 5 | LD% 63 | Cost 2.66 | N-align 5 | LD% 53.2 |
| uncle, nephew, brother/sister in law | | | uvula, the recurved point of a fishing hook | | |
| H. *süv* | [ ʃ **y v** ] | | H. *szaka* | [ s **ɒ** k **ɒ -** ] | |
| O.H. *sig* | [ ʃ **i ɣ** ] | | O.H. *szakag* | [ s **a** k **a ɣ** ] | |
| Cost 1 | N-align 3 | LD% 33.33 | Cost 1.16 | N-align 5 | LD% 23.2 |
| beard | | | raft | | |
| H. *szakáll* | [ s **ɒ** k **a: l:** ] | | H. *szál* | [ s **a:** l ] | |

| | | | | | |
|---|---|---|---|---|---|
| O.H. *szakal* | [ s **a** k **a** **l** ] | | O.H. *szal* | [ s **a** l ] | |
| Cost 1 | N-align 5 | LD% 20 | Cost 0.33 | N-align 3 | LD% 11 |
| number | | | to have pity for, to regret | | |
| H. *szám* | [ s **a:** m ] | | H. *szán* | [ s **a:** **-** **-** n ] | |
| O.H. *szam* | [ s **a** m ] | | O.H. *szagin* | [ s **a** **ɣ** **i** n ] | |
| Cost 0.33 | N-align 3 | LD% 11 | Cost 1.83 | N-align 5 | LD% 36.6 |
| to wish, to intend something for somebody | | | bucket, wooden pail | | |
| H. *szán* | [ s a: n ] | | H. *szapu* | [ s **ɒ** p **u** **-** ] | |
| O.H. *szán* | [ s a: n ] | | O.H. *szapag* | [ s **a** p **a** **ɣ** ] | |
| Cost 0 | N-align 3 | LD% 0 | Cost 1.5 | N-align 5 | LD% 30 |
| light coloured, yellowish, pale | | | wing | | |
| H. *szár* | [ s **a:** r ] | | H. *szárny* | [ **s** a: r **ɲ** ] | |
| O.H. *szar* | [ s **a** r ] | | O.H. *sárn* | [ **ʃ** a: r **n** ] | |
| Cost 0.33 | N-align 3 | LD% 11 | Cost 1 | N-align 4 | LD% 25.5 |
| grocer, grand handler | | | chair, seat, bench, throne | | |
| H. *szatócs* | [ s **ɒ** t **o:** **-** **t͡ʃ** **-** ] | | H. *szék* | [ s **e:** k **-** ] | |
| O.H. *szatigcsi* | [ s **a** t **ï** **ɣ** **t͡ʃ** **ï** ] | | O.H. *szekü* | [ s **ɛ** k **y** ] | |
| Cost 2.83 | N-align 7 | LD% 40.43 | Cost 1.66 | N-align 4 | LD% 41.5 |
| wind | | | to slumber, to doze | | |
| H. *szél* | [ s e: l ] | | H. *szender(ed)(ik)* | [ s **ɛ** n d **ɛ** r ] | |
| O.H. *szél* | [ s e: l ] | | O.H. *szöndür* | [ s **ø** n d **y** r ] | |
| Cost 0 | N-align 3 | LD% 0 | Cost 1.33 | N-align 6 | LD% 22.17 |
| beautiful | | | freckle, sun spot, stain | | |
| H. *szép* | [ s **e:** p ] | | H. *szeplő* | [ s ɛ p l **ø** **-** ] | |
| O.H. *szip* | [ s **i** p ] | | O.H. *szepleg* | [ s ɛ p l **ɛ** **ɣ** ] | |
| Cost 0.66 | N-align 3 | LD% 22 | Cost 1.16 | N-align 6 | LD% 19.33 |
| part of a village, street, group of people | | | treshing floor (round) | | |
| H. *szer* | [ **s** ɛ r ] | | H. *szérű* | [ s **e:** r **y:** **-** ] | |
| O.H. *ser* | [ **ʃ** ɛ r ] | | O.H. *szürüg* | [ s **y** r **y** **ɣ** ] | |
| Cost 0.33 | N-align 3 | LD% 11 | Cost 1.83 | N-align 5 | LD% 36.6 |

| fumes, vapour, spirits, alcohol, humours | | | colour, face, external appearance | | |
|---|---|---|---|---|---|
| H. *szesz* | [ s **ɛ** s ] | | H. *szín* | [ s **i:** n ] | |
| O.H. *szisz* | [ s **i** s ] | | O.H. *szin* | [ s **i** n ] | |
| Cost 0.33 | N-align 3 | LD% 11 | Cost 0.33 | N-align 3 | LD% 11 |
| thin hide rope, strap | | | rock, cliff, occiput, crest (of mountain) | | |
| H. *szirony* | [ s i **-** r **o ɲ** ] | | H. *szirt* | [ s i r t ] | |
| O.H. *sziyrum* | [ s i **j** r **u m** ] | | O.H. *szirt* | [ s i r t ] | |
| Cost 1.83 | N-align 6 | LD% 36 | Cost 0 | N-align 4 | LD% 0 |
| word | | | to spread, to scatter, to winnow | | |
| H. *szó* | [ s **o: -** ] | | H. *szór* | [ s **o: - -** r ] | |
| O.H. *szaw* | [ s **a w** ] | | O.H. *szawur* | [ s **a w u** r ] | |
| Cost 1.5 | N-align 3 | LD% 50 | Cost 2.5 | N-align 5 | LD% 50 |
| grasshoper | | | to leap, to jump, to escape, to flee | | |
| H. *szöcske* | [ s ø **t͡ʃ** k **ɛ - -** ] | | H. *szök(ik)* | [ s **ø** k ] | |
| O.H. *szököső* | [ s ø **-** k **ø ʃ ø:** ] | | O.H. *szek* | [ s **ɛ** k ] | |
| Cost 3.66 | N-align 7 | LD% 52.29 | Cost 0.33 | N-align 3 | LD% 11 |
| grape, wine grape | | | mosquito, gnat | | |
| H. *szőlő* | [ **s ø** l **ø -** ] | | H. *szúnyog* | [ s **u: ɲ o g** ] | |
| O.H. *seleg* | [ **ʃ ɛ** l **ɛ ɣ** ] | | O.H. *szinuk* | [ s **i n u k** ] | |
| Cost 1.5 | N-align 5 | LD% 30 | Cost 2.5 | N-align 5 | LD% 50 |
| furrier | | | to cease, to stop | | |
| H. *szűcs* | [ **s y: -** t͡ʃ **-** ] | | H. *szűn(ik)* | [ s **y:** n ] | |
| O.H. *sigcsi* | [ **ʃ i ɣ** t͡ʃ **i** ] | | O.H. *szön* | [ s **ø** n ] | |
| Cost 2.50 | N-align 5 | LD% 50 | Cost 0.66 | N-align 3 | LD% 22 |
| to strain, to filter | | | virgin, pure | | |
| H. *szűr* | [ s **y:** r ] | | H. *szűz* | [ s **y:** z ] | |
| O.H. *szür* | [ s **y** r ] | | O.H. *szüz(ök)* | [ s **y** z ] | |
| Cost 0.33 | N-align 3 | LD% 11 | Cost 0.33 | N-align 3 | LD% 11 |
| sorcerer, medicine man | | | advice, council | | |
| H. *táltos* | [ t **a:** l t **o s -** ] | | H. *tanács* | [ t **ɒ** n **a: -** t͡ʃ ] | |

| | | | | | |
|---|---|---|---|---|---|
| O.H. *taltucsi* | [ t **a** l t **u** **t͡ʃ** **i** ] | | O.H. *tanilcs* | [ t **a** n **i** **l** **t͡ʃ** ] | |
| Cost 2.49 | N-align 7 | LD% 35.57 | Cost 2.33 | N-align 6 | LD% 33.33 |
| witness | | | to feed, to nourish | | |
| H. *tanú* | [ t **ɒ** n **u:** ] | | H. *táplá(l)* | [ t **a:** p l **a:** **-** ] | |
| O.H. *tanu* | [ t **a** n **u** ] | | O.H. *taplag* | [ t **a** p l **a** **ɣ** ] | |
| Cost 0.66 | N-align 4 | LD% 16.5 | Cost 1.16 | N-align 6 | LD% 19.33 |
| tinder, tinder fungus | | | bald | | |
| H. *tapló* | [ t **ɒ** p l **o:** **-** ] | | H. *tar* | [ t **ɒ** r ] | |
| O.H. *toplug* | [ t **o** p l **u** **ɣ** ] | | O.H. *tar* | [ t **a** r ] | |
| Cost 1.5 | N-align 6 | LD% 25 | Cost 0.33 | N-align 3 | LD% 11 |
| depot | | | plough field, arable field, stubble field | | |
| H. *tár* | [ t **a:** **-** **-** r ] | | H. *tarló* | [ t **ɒ** r l **o:** **-** ] | |
| O.H. *tavar* | [ t **a** **v** **a** r ] | | O.H. *tarlag* | [ t **a** r l **a** **ɣ** ] | |
| Cost 2.33 | N-align 5 | LD% 46.6 | Cost 1.83 | N-align 6 | LD% 30.5 |
| to hold, carry, to last | | | tatar | | |
| H. *tart* | [ t **ɒ** r t ] | | H. *tatár* | [ t **ɒ** t **a:** r ] | |
| O.H. *tart* | [ t **a** r t ] | | O.H. *tatar* | [ t **a** t **a** r ] | |
| Cost 0.33 | N-align 4 | LD% 8.25 | Cost 0.66 | N-align 5 | LD% 13.2 |
| name of several kinds of plants | | | to wind something round, to twist | | |
| H. *tátorján* | [ t **a:** t **o** r **j** **a:** **n** ] | | H. *taker* | [ t ɛ k ɛ r ] | |
| O.H. *tatrang* | [ t **a** t **-** r **-** **a** **ŋ** ] | | O.H. *teker* | [ t ɛ k ɛ r ] | |
| Cost 2.83 | N-align 8 | LD% 35.37 | Cost 0 | N-align 5 | LD% 0 |
| through, wash tub, hutch | | | a piece of land, parcel, patch | | |
| H. *teknő* | [ t ɛ k **-** n **ø** **-** ] | | H. *telek* | [ t ɛ l **ɛ** k ] | |
| O.H. *tekeneg* | [ t ɛ k **ɛ** n **ɛ** **ɣ** ] | | O.H. *telük* | [ t ɛ l **y** k ] | |
| Cost 2.16 | N-align 7 | LD% 30.86 | Cost 0.66 | N-align 5 | LD% 13.2 |
| strap (on a whip, or on a kind of sandal) | | | to vegetate, to linger in misery | | |
| H. *telek* | [ t ɛ l **ɛ** k ] | | H. *teng* | [ t ɛ ŋ ] | |
| O.H. *telük* | [ t ɛ l **y** k ] | | O.H. *teng* | [ t ɛ ŋ ] | |
| Cost 0.66 | N-align 5 | LD% 13.2 | Cost 0 | N-align 3 | LD% 0 |

| | | | | | |
|---|---|---|---|---|---|
| Axle | | | sea | | |
| H. *tengely* | [ t ɛ ŋ g ɛ **j** ] | | H. *tenger* | [ t ɛ ŋ **ɛ** r ] | |
| O.H. *tengel* | [ t ɛ ŋ g ɛ **l** ] | | O.H. *tenir* | [ t ɛ ŋ **i** r ] | |
| Cost 1.25 | N-align 6 | LD% 20.83 | Cost 0.33 | N-align 5 | LD% 6.6 |
| to turn, to change the original direction | | | knee | | |
| H. *tér* | [ t **e: - -** r ] | | H. *tér(d)* | [ t **e:** r ] | |
| O.H. *tevir* | [ t **ɛ v i** r ] | | O.H. *tír* | [ t **i:** r ] | |
| Cost 2.66 | N-align 5 | LD% 53.2 | Cost 0.66 | N-align 3 | LD% 22 |
| hall, chamber, great room | | | camel | | |
| H. *terem* | [ t ɛ r ɛ m ] | | H. *teve* | [ t ɛ v ɛ ] | |
| O.H. *terem* | [ t ɛ r ɛ m ] | | O.H. *teve* | [ t ɛ v ɛ ] | |
| Cost 0 | N-align 5 | LD% 0 | Cost 0 | N-align 4 | LD% 0 |
| hemp breaker, swingle, scutch | | | steer, young bullock, ox | | |
| H. *tiló* | [ t **i** l **- o: -** ] | | H. *tinó* | [ t **i** n **o:** ] | |
| O.H. *talkig* | [ t **a** l **k i ɣ** ] | | O.H. *tana* | [ t **a** n **a** ] | |
| Cost 3.5 | N-align 6 | LD% 58.33 | Cost 1.66 | N-align 4 | LD% 41.5 |
| to lay eggs | | | sturgeon | | |
| H. *toj(ik)* | [ t **o j** ] | | H. *tok* | [ t o k **-** ] | |
| O.H. *tug* | [ t **u ɣ** ] | | O.H. *toku* | [ t o k **u** ] | |
| Cost 1.15 | N-align 3 | LD% 38.33 | Cost 1 | N-align 4 | LD% 25 |
| interpreter | | | horseradish | | |
| H. *tolmács* | [ t o l m **a:** t͡ʃ ] | | H. *torma* | [ t **o** r m **ɒ** ] | |
| O.H. *tolmacs* | [ t o l m **a** t͡ʃ ] | | O.H. *turma* | [ t **u** r m **a** ] | |
| Cost 0.33 | N-align 6 | LD% 5.5 | Cost 0.66 | N-align 5 | LD% 13.2 |
| to pile up | | | a kind of small falcon | | |
| H. *toro(l)* | [ t o r o ] | | H. *torontál* | [ t **o** r **o** n t **a: l** ] | |
| O.H. *toro* | [ t o r o ] | | O.H. *turuntay* | [ t **u** r **u** n t **a j** ] | |
| Cost 0 | N-align 4 | LD% 0 | Cost 2.25 | N-align 8 | LD% 28.12 |
| Slovak | | | to perform | | |
| H. *tót* | [ t **o:** t ] | | H. *töké(l)* | [ t **ø** k **e:** ] | |

| | | | | | |
|---|---|---|---|---|---|
| O.H. *tát* | [ t **a:** t ] | | O.H. *tüke* | [ t **y** k **ɛ** ] | |
| Cost 0.66 | N-align 3 | LD% 22 | Cost 1 | N-align 4 | LD% 25 |
| concentrated, ten thousand, a military unit | | | to break, to separate into pieces | | |
| H. *tömény* | [ t **ø** m **e: ɲ** ] | | H. *tör* | [ t **ø - -** r ] | |
| O.H. *tümen* | [ t **y** m **ɛ n** ] | | O.H. *tügür* | [ t **y ɣ y** r ] | |
| Cost 1.66 | N-align 5 | LD% 33.2 | Cost 1.83 | N-align 5 | LD% 36.6 |
| snare, trap | | | Turk, Turkish, Turkic, Ottoman | | |
| H. *tőr* | [ t ø: r ] | | H. *török* | [ t **ø** r **ø** k ] | |
| O.H. *tőr* | [ t ø: r ] | | O.H. *türkü* | [ t **y** r k **y** ] | |
| Cost 0 | N-align 3 | LD% 0 | Cost 1.66 | N-align 5 | LD% 33.2 |
| to happen, to occur | | | law | | |
| H. *történ(ik)* | [ t ø r t **e:** n ] | | H. *törvény* | [ t **ø** r **v - e: ɲ** ] | |
| O.H. *törtün* | [ t ø r t **y** n ] | | O.H. *türügen* | [ t **y** r **y ɣ ɛ n** ] | |
| Cost 1 | N-align 6 | LD% 16.67 | Cost 3.5 | N-align 7 | LD% 50 |
| steer, young bullock, ox or cow | | | cottage cheese, cheese curds | | |
| H. *tulok* | [ t u l o k ] | | H. *túró* | [ t **u:** r **o: -** ] | |
| O.H. *tukol* | [ t u k o l ] | | O.H. *torag* | [ t **o** r **a ɣ** ] | |
| Cost 1 | N-align 5 | LD% 20 | Cost 2.16 | N-align 5 | LD% 43.2 |
| the totem bird of the Árpád dynasty | | | bustard | | |
| H. *turul* | [ t u **-** r **u** l ] | | H. *túzok* | [ t **u: -** z **o** k ] | |
| O.H. *tugril* | [ t u **ɣ** r **ï** l ] | | O.H. *togzak* | [ t **o ɣ** z **a** k ] | |
| Cost 0.83 | N-align 6 | LD% 13.83 | Cost 1.83 | N-align 6 | LD% 30.5 |
| mirror | | | to endure, suffer, bear, stand | | |
| H. *tükör* | [ t y k **ø** r ] | | H. *tűr* | [ t **y:** r ] | |
| O.H. *tüker* | [ t y k **ɛ** r ] | | O.H. *tür* | [ t **y** r ] | |
| Cost 0.66 | N-align 5 | LD% 13.2 | Cost 0.33 | N-align 3 | LD% 11 |
| to roll up (a scroll or one's sleeves) | | | hen | | |
| H. *tűr* | [ t **y:** r ] | | H. *tyúk* | [ **c - - u:** k ] | |
| O.H. *tür* | [ t **y** r ] | | O.H. *tiguk* | [ **t i ɣ u** k ] | |
| Cost 0.33 | N-align 3 | LD% 11 | Cost 2.5 | N-align 5 | LD% 50 |

| owl | | | gentleman, lord, sir, high dignity | | |
|---|---|---|---|---|---|
| H. *ugu* | [ u g u ] | | H. *úr* | [ **u:** r **-** ] | |
| O.H. *ugu* | [ u g u ] | | O.H. *uri* | [ **u** r **i** ] | |
| Cost 0 | N-align 3 | LD% 0 | Cost 1.33 | N-align 3 | LD% 44.33 |
| to refresh oneself, to rest, to become cured | | | feast | | |
| H. *üdü(l)* | [ **y** d y ] | | H. *ün(nep)* | [ **y n -** ] | |
| O.H. *edü* | [ **ɛ** d y ] | | O.H. *edü* | [ **e d y** ] | |
| Cost 0.66 | N-align 3 | LD% 22 | Cost 2.24 | N-align 3 | LD% 74.67 |
| heifer, roe doe | | | ground squirrel | | |
| H. *ünő* | [ **y** n **ø: -** ] | | H. *ürge* | [ y r g ɛ ] | |
| O.H. *ineg* | [ **i** n **ɛ ɣ** ] | | O.H. *ürge* | [ y r g ɛ ] | |
| Cost 1.83 | N-align 4 | LD% 45.75 | Cost 0 | N-align 4 | LD% 0 |
| wormwood | | | wether, sheep | | |
| H. *üröm* | [ **y** r **ø** m ] | | H. *ürü* | [ **y** r **y -** ] | |
| O.H. *erim* | [ **ɛ** r **i** m ] | | O.H. *irig* | [ **i** r **i ɣ** ] | |
| Cost 1.33 | N-align 4 | LD% 33.25 | Cost 1.16 | N-align 4 | LD% 29 |
| hot embers, extinguished embers | | | one-year-old female sheep | | |
| H. *üszök* | [ **y** s **ø k** ] | | H. *üvecs* | [ **y** v ɛ t͡ʃ ] | |
| O.H. *üszüg* | [ **i** s **y g** ] | | O.H. *övecs* | [ **ø** v ɛ t͡ʃ ] | |
| Cost 1.33 | N-align 4 | LD% 33.25 | Cost 0.33 | N-align 4 | LD% 8.25 |
| to hollow out, to scoop out | | | trough, tray | | |
| H. *váj* | [ v **a:** j ] | | H. *vályú* | [ v **a: j u: -** ] | |
| O.H. *vay* | [ v **a** j ] | | O.H. *valag* | [ v **a l a ɣ** ] | |
| Cost 0.33 | N-align 3 | LD% 11 | Cost 3.08 | N-align 5 | LD% 61.6 |
| fishweir | | | a hole in the ice | | |
| H. *vejsze* | [ v ɛ j **s** ɛ **-** ] | | H. *vék* | [ v **e:** k **-** ] | |
| O.H. *vejseg* | [ v ɛ j **ʃ** ɛ **ɣ** ] | | O.H. *vekü* | [ v **ɛ** k **y** ] | |
| Cost 0.83 | N-align 6 | LD% 13.83 | Cost 1.66 | N-align 4 | LD% 41.5 |
| guest | | | to plait, to lay the rope | | |
| H. *vendég* | [ v ɛ n d **e:** g ] | | H. *ver* | [ v ɛ r ] | |

| | | | | | |
|---|---|---|---|---|---|
| O.H. *vendeg* | [ v ɛ n d **ɛ** g ] | | O.H. *ver* | [ v ɛ r ] | |
| Cost 0.66 | N-align 6 | LD% 11 | Cost 0 | N-align 3 | LD% 0 |
| gable, frontispiece | | | chamois | | |
| H. *vértelek* | [ v **e:** r t ɛ l **ɛ** k ] | | H. *zerge* | [ **z** ɛ r g ɛ ] | |
| O.H. *vertelük* | [ v **ɛ** r t ɛ l **y** k ] | | O.H. *szerge* | [ **s** ɛ r g ɛ ] | |
| Cost 1.33 | N-align 8 | LD% 16.62 | Cost 0.5 | N-align 5 | LD% 10 |

### 3.3 Patterns

In this section, frequency analysis will be practiced to highlight key findings related to the most common phonological adaptations. The most frequent phonological tendencies in terms of adaptations mirror the phonological disposition of Hungarian towards the nativization process of loanwords. Only the most striking adaptation patterns are displayed in this section. The results demonstrate certain phonological processes that the Hungarian language applies when nativizing loanwords. The significant point here is that the adaptations display – to a large extent – systematic patterns through which we understand that phonological constraints behave in a determined unarbitrary fashion when integrating borrowings. This, of course, can be evaluated as an indication that the nativization through production stance has a valid basis. The patterns that are displayed below demonstrate that the adaptations are not the results of the random perception of mere phonetic inputs. Due to the insignificance of the alterations made, the generation of patterns for the "initial copying process" proves to be unfeasible. The only exceptions are presented in Table 14. Consequently, almost all the subsequent data and patterns presented, solely pertain to the "adaptation process". From this juncture, the focus is mostly directed towards the examination and analysis of the adaptation process. When certain patterns are mentioned though, the comparisons that are being referred to are the ones that are parallel with Table 14 and 15. For example, adaptation of [ï] is a pattern that is revealed in the borrowing process, therefore the comparisons are between the WOT and O.H. words.

**Table 14. Borrowing patterns**

| Pattern | Occurrence rate in the whole data | Occurrence rate in the relevant data | Common adapted forms |
|---|---|---|---|
| Adaptation of [ɨ] | 11.67% (377/44) | 88% (50/44) | [o], [i] |
| Adaptation of [æ] | 20.69% (377/78) | 92.86% (84/78) | [ɛ], [ø] |
| Adaptation of [t͡ʃ] to [ʃ] | 1.86% (377/7) | 8.86% (79/7) | [ʃ] |

**Table 15. Adaptation patterns**

| Pattern | Occurrence rate in the whole data | Occurrence rate in the relevant data | Common adapted forms |
|---|---|---|---|
| Heavy syllables | 52.25% (377/197) | 52.25% (377/197) | vowel lenghtening |
| Nasal palatalization | 7.43% (377/28) | 35.44% (79/28) | [ɲ] |
| Adaptation of [ɣ] | 30.77% (377/116) | 100% (116/116) | [-], [j], [v] |
| Labialization of vowels | 46.15% (377/174) | - | [a]==>[ɒ][o][u], [e][ ɛ]==>[ø][y], [i]==>[ø][y], [ɨ]==>[o][u] |
| Adaptation of [j] ➔ [ɟ] | 8.75% (377/33) | 68.75% (48/33) | [ɟ] |
| Adaptation of [t͡ʃ] ➔ [ʃ] | 5.57% (377/21) | 29.17% (72/21) | [ʃ] |

### 3.3.1 Heavy syllables

"In a number of foreign words, Hungarian has to meet a peculiar requirement, they have to contain a heavy syllable" (Kertész, 2006 p. 12). It's important to note however that heavy syllable production in Hungarian is not necessarily an attempt to make borrowed words conform to the Hungarian phonological system; sometimes, it's the opposite. Heavy syllables appear either as syllables that contain long vowels or as syllables that display the gemination of consonants, or consonant clusters. In Hungarian, syllable weight is determined primarily by the length of vowels. A syllable is considered heavy if it contains a long vowel or a long vowel in combination with a consonant. Short vowels, on the other hand, contribute to the

perception of a light syllable. Vowel length is an inherent property of the vowel itself, and it is not influenced by surrounding consonants, and it is the primary factor determining syllable weight. Although long vowels are very scarce in Turkic phonotactics today (they are almost entirely from loanwords), they are very typical in Hungarian. In Old Turkic though, long vowels did exist. This section primarily (almost entirely) explores the phenomenon of long vowels. The starting point is the recognition that a light syllable typically consists of a single mora, which can be a short vowel, a diphthong, or a single consonant. Conversely, a heavy syllable contains two moras, typically in the form of a long vowel, a vowel followed by a single consonant, or a vowel followed by a consonant cluster. Notably, nearly half of the comparisons made reveal a significant prevalence of vowel lengthening, underscoring its status as a highly recurrent pattern. It is seen in this study that 197 lexemes out of 377, exhibit this adaptation. Adaptations occur both in initial and non-initial or terminative syllables. Gemination is observed in 2 adaptations, *bükk* [byk:] 'beech' and *szakáll* [sɒka:l:] 'beard'. Degemination on the other hand can be seen in 2 words, *ács* [a:t͡ʃ] 'carpenter' and *csákány* [t͡ʃa:ka:ɲ] 'pick-axe'. Heavy syllables that display consonant clustering are *bilincs* 'shackles', *bors* 'pepper', *borz* 'badger', *böjt* 'fast', *bölcs* 'wise', *bükk* 'beech', *erkölcs* 'morals', *gyárt* 'to produce', *gyöngy* 'pearl', *gyümölcs* 'fruit', *ijeszt* 'to frighten', *könyv* 'book', *körtvély* 'pear', *orv* 'thief', *ölt* 'to stitch', *sajt* 'cheese', *szakáll* 'beard', *szárny* 'wing', *szirt* 'crest', *tart* 'to hold', *teng* 'to be in misery', *térd* 'knee'. Among these, *bors*, *erkölcs*, *gyümölcs*, *körtvély*, *ölt*, *sajt*, *szirt*, *tart* already display clustering in their source counterparts.

### 3.3.2 Nasal palatalization

28 adaptations exhibit nasal palatalization. This sound change is almost always observed in final syllables although contrarian examples such as *kanyaró* [kɒɲɒro:] do exist. Nasal palatalization is known to occur in Hungarian phonology before palatal consonants (Siptár, 1994, pp. 5-32). When the case is loanword adaptations, there seems to be no regular aspect to it, except that it mainly occurs if the nasal sound is the final phoneme in the lexeme. The process of palatalization in Hungarian, as described in historical linguistics, seems to have begun as context-induced palatalization, particularly before palatal vowels. However, it later extended to the ends of words, becoming more generalized in its application. While this

process may have been assimilatory in nature initially, it is not identical to the present-day processes discussed in the above-given study.

### 3.3.3 High back unrounded vowel

There is strong evidence to suggest that /ɯ/ existed in Hungarian at some point in its history. This evidence includes the existence of words in modern Hungarian that have /ɯ/ in some dialects, and the fact that /ɯ/ is found in many related Uralic languages and in the reconstructed Proto-Uralic language (Pystynen, 2017). However, there is some debate among linguists about when /ɯ/ disappeared from Hungarian. Some scholars believe that it disappeared by the early Old Hungarian period, others believe that it may have persisted in some dialects of Hungarian until the late Old Hungarian period, or even later (cf. Kiss & Pusztai ed., 2003; É. Kiss et al., 2013; Kiss & Pusztai eds., 2018). Some argue against the existence of the /ɯ/ sound in the history of the Hungarian language (Kis, 2005).

Today, /ɯ/ – or /ɨ/ as in scholarly writing of Turkic languages (Erdal, 2004, p. 52) – is missing from the Hungarian native sound system. 44 adaptations are discovered in the initial borrowing stage. In the adaptation stage, on the other hand, only 7 are detected. Nativization of this sound in the data commonly appears either as /o/, /ɒ/ or /i/. Since /ɨ/ is a [+high] and [+back] vowel, it is perceptible that the three predominant substitutions share at least one of these features. The salient outlook in the adaptation of this sound in Hungarian would be that it goes out of the classical front-back contrast of the harmonical boundaries (of Turkic and Hungarian both) as it assimilates into /i/ very often. In the entirety of the WOT words containing /ɨ/, there are no instances where the front-back vowel harmony is disrupted. However, in the case of Hungarian adaptations, it becomes apparent that the nativization process does not conform to the expected harmonic rules governing the front-back vowel contrast in Hungarian.

### 3.3.4 Voiced velar fricative

There is no doubt that the /ɣ/ was not silent in Old Turkic as mentioned before in this study. Borrowings such as *yogurt* that preserved this sound as a non-silent variant in foreign languages affirm this. This being the case, it can be argued that the transitional feature of this

sound was still present nevertheless. A shred of strong evidence is that out of all the occurrences, in only three of the WOT words it is not preceded by a vowel; *balturgan* [balturɣan] > *bojtorján* [bojtorja:n], *hırğuy* [χïrɣuj] > *herjó* [hɛrjo:], *korğan* [korɣan] > *korhány* [korha:ɲ]. Another indication is that out of the 116 occurrences in the adaptation stage, 95 of them turn out to be a deletion in the Hungarian adaptations. There are 11 substitutions to the semivowel /j/, and 5 substitutions to /v/. As a result /ɣ/ emerges as another pattern in which the adaptations display a similar characteristic to the "soft g" in modern Turkish. It mainly functions as a vowel lengthener in the adaptations.

### 3.3.5 Labialization of vowels

Substitution into the low back rounded vowel /ɒ/ is a common adaptation. It takes place of the /a/ sound in 66 words and thus appears as a common pattern. /a/➔/o/ pattern is observed in 24 comparisons while /ɛ/➔/y/ is observed in 6, /ɛ/➔/ø/ in 38, /e/➔/ø/ in 1, and /e/➔/y/ in 4 comparisons. /i/➔/ø/ and /i/➔/y/ patterns are displayed in 11 and 17 occasions respectively. The sound /ï/ exists in only in 7 O.H. words and 2 of them labialize as /u/, while 3 of them as /o/. Apart from these changes, 2 occurrences of /a/➔/u/ are detected. As a result, labialization of vowels takes place in 174 pairs.

### 3.3.6 [±low] quality of /æ/

Apart from /a/ and /ɒ/, another [+low] vowel /ä/ – being absent in Hungarian – is used in the lexicon in some Turkic words as the [+front] variant. However, the quality of this sound in Turkic words is multifarious. In most cases in Turkic languages, the quality is near-low, therefore /æ/. Considering the fact that in the comparisons, it is always assimilated into either /e/, /ɛ/, and /ø/ but never /a/, advocates this condition. 78 word pairs displaying substitutions from /æ/ are observed in the adaptations.

# CHAPTER FOUR

## Discussion

One important observation from this study is that the integration of loanwords when they are borrowed or copied from West Old Turkic into Hungarian, appears to be minimal during the initial stage. The average pronunciation distance of 7.21% indicates a very close phonetic resemblance between the borrowed words and their West Old Turkic originals. This finding suggests that at the time of borrowing, Hungarian did not undergo significant modifications. The borrowed words were largely reproduced or copied in their original phonetic form.

This minimal change during the initial stage can be attributed to several factors. Firstly, it could reflect the nature of borrowing between these particular languages during the studied period. Given the fact that there was frequent contact and interaction between speakers of Hungarian and West Old Turkic, it is reasonable that the borrowed words were assimilated into Hungarian with minimal phonetic alterations. This could be due to the familiarity of the borrowing community with the pronunciation and phonetic features of West Old Turkic.

Secondly, the minimal initial change observed in this study might be influenced by the phonetic similarities between Hungarian and West Old Turkic. If the two languages shared similar phonetic systems or had overlapping phonetic features, it would require fewer adjustments for borrowed words to align with the phonetic patterns of Hungarian. This would result in a smaller pronunciation distance between the borrowed words and their WOT origins. While the snapshot analysis presented in the dissertation suggests a close resemblance at the time of borrowing, it is crucial to acknowledge that this process was not a one-time event. The borrowing of Turkic elements spanned a period of bilingualism, during which significant changes occurred in the Hungarian language, such as the introduction of new sounds and phonemes. Moreover, the Ugric vowel system underwent significant alterations under the influence of Turkic (Róna-Tas & Berta, 2011, p. 1121). Exploring the changes that transpired in the Hungarian language from its separation from the Ugric community through its encounter with Turkic languages and continuing observation of these changes during the borrowing process would reveal a significantly greater degree of difference.

Additionally, the relatively minimal alterations at the initial stage could be attributed to the cultural or social factors surrounding language contact and borrowing. If there was a strong cultural or social prestige associated with West Old Turkic or a positive perception of its linguistic features, there might have been a tendency to preserve the phonetic integrity of borrowed words in Hungarian. This could be seen as a way to maintain the perceived authenticity or prestige associated with the borrowed vocabulary. However, it is important to note that the exploration of the subsequent stages of adaptation within Hungarian is also a matter of importance. The study takes up this challenge also and examines the "adaptation" of the borrowed material within the Hungarian language. It is evident that the borrowed words underwent additional modifications and phonetic adjustments over time, in line with the evolution of the Hungarian language.

The minimal change observed during the initial stage of the loan process suggests a close phonetic resemblance between Hungarian borrowings and West Old Turkic counterparts. Factors such as frequent language contact, phonetic similarities between the two languages, and cultural or social influences may have contributed to this minimal alteration rate. Further research to investigate the subsequent stages of adaptation to gain a more comprehensive understanding of the overall loanword adaptation process in Hungarian is presented separately.

This subsequent process reveals an adaptation rate of 28.67% for the loanwords in modern Hungarian. The presence of certain patterns in loanword adaptations reflects the phonological constraints and tendencies of the recipient language, Hungarian, when accommodating borrowed words. These patterns arise due to a combination of linguistic factors, including phonotactic constraints, historical phonological changes, and the interaction between the borrowing and recipient languages. Understanding these patterns sheds light on the underlying mechanisms and principles governing loanword adaptation in Hungarian.

One prominent pattern observed in the study is the preservation of heavy syllables in loanword adaptations. The starting point here is the recognition that a light syllable typically consists of a single mora, which can be a short vowel, a diphthong, or a single consonant. Conversely, a heavy syllable contains two moras, typically in the form of a long vowel, a vowel followed by a single consonant, or a vowel followed by a consonant cluster. Notably,

nearly half of the comparisons made reveal a significant prevalence of vowel lengthening, underscoring its status as a highly recurrent pattern.

The adaptation of the high back unrounded vowel (/ɯ/ or /ɨ/) in loanword nativization reveals interesting tendencies. The observed substitutions of /o/, /ɒ/, or /i/ for this vowel suggest that loanword adaptations prioritize maintaining certain features, such as [+high] and [+back], rather than strictly adhering to the front-back harmony rules of Turkic and Hungarian. This deviation from expected patterns may be influenced by factors such as ease of pronunciation or the influence of similar sounds in the recipient language. It indicates that loanword adaptations are not solely driven by phonetic similarity but also involve complex interactions between the phonological systems of the borrowing and recipient languages.

The adaptation of the voiced velar fricative (/ɣ/) in loanword nativization demonstrates patterns resembling the “soft g” phenomenon found in modern Turkish. The observed patterns, including deletion, substitution to semivowel /j/, or vowel lengthening, suggest a convergence between the historical phonological changes of Hungarian and the phonotactic preferences of loanword adaptations. This convergence reflects the dynamic nature of language contact and the influence of historical linguistic developments on the adaptation of borrowed words.

The preference for labialization can be attributed to several factors. One may be the influence of adjacent consonants. When a borrowed word contains labial or labialized consonants in proximity to a vowel, the vowel often undergoes labialization to match the articulatory characteristics of the consonants. This assimilation process ensures phonetic harmony and facilitates the smooth pronunciation of loanwords within the phonological framework as in the example of *gyümölcs* where the vowels adjacent to the labial /m/, labialize during the adaptation.

The substitution of /j/ with /ɟ/ can be attributed to the principle of phonetic similarity and ease of articulation. Both /j/ and /ɟ/ are palatal sounds, characterized by articulatory movements involving the raising of the tongue towards the hard palate. The transition from the approximant /j/ to the stop /ɟ/ involves a similar articulatory movement with the closure

of the vocal tract at the place of articulation. This shared feature in articulation makes the substitution a natural and perceptually plausible choice.

Although the principles are default, refinement approaches to the operation cost in the Levenshtein algorithms, make significant differences in terms of acquired results. Reflection rates of some adaptations in human perception may be little to none. Some adaptations, on the other hand, may sound obvious but still less distant than some others. Perhaps, to begin with, a computerized (or non-computerized) advancement of a measurement methodology that evaluates the perceived distance of a certain set of sounds particular to two phonological systems may be a good approach. A database that is big enough to consist of a sufficient number of co-occurrences of the same sound pairs is needed to achieve this. In any case, it is necessary to have big data with a large word base to apply the LD and obtain meaningful, consistent findings. Narrow data for example, such as 35 acknowledged Cuman loanwords in Hungarian, is a limitation and appears as a disadvantage of the application of the LD. This research is conducted with significantly large data on WOT loanwords in Hungarian. Another disadvantage is that the refinement strategies require a good command of the phonological knowledge of the compared languages. If any of the languages lack comprehensive phonological data as in Cuman, being an extinct language, this can also emerge as a disadvantage. Advantages on the other hand are being able to quantitate the relatedness of compared lexemes and acquiring data concerning the commonness of certain adaptations as regular patterns. It should be noted, as Honti also states, that the changes in consonants are known to show a greater degree of regularity than those of vowels (Honti, 2017, p. 189). The automatic phonetic comparisons may also contribute to the settlement of the discussion of the phonological vs perceptional nature of adaptations. Phonologically sensitized automatization methods for calculating the adaptation rates of loanwords between languages may help to attest to a correct acknowledgment of the dynamics of loanword adaptation.

One aspect to take into consideration is that in the Sprachbund of the Eurasian steppes, Turkic and Hungarian speakers were not the only ones, needless to say. Iranian speakers such as Alans, Germanic speakers such as Goths, and Slavonic speakers co-existed with the former two in different periods. Doubtlessly, the interactions with these groups must have affected the loan phonology of Hungarian borrowings. In this regard, a versatile approach to

adaptations instead of a single-ended one could bring a better understanding. Stachowski brings up this issue and suggests a comparison of phonetic adaptation processes of Turkic loanwords in Hungarian with those of Pannonian Slavonic (Stachowski, 2014 p. 221).

Another assessment for further research can be to test native speakers of the borrowing language, asking them about their perception of the source words, or vice versa with the source language speakers, and comparing the results with the automated rates. It may also be worthwhile to carry out the study described earlier in this study, taking into account the findings of Honti (2017) for the selection of lexemes to be analyzed.

# CHAPTER FIVE

# Conclusions

In conclusion, this research focused on the adaptation rates of West Old Turkic loanwords in Hungarian, examining both the borrowing and adaptation processes. The findings indicate an initial borrowing assimilation rate of 7.21% for the loanwords, highlighting the minimal changes that occur during the direct transfer of lexical material from West Old Turkic to Old Hungarian. The subsequent process reveals an adaptation rate of 28.67% for the loanwords in modern Hungarian, reflecting the rule-based changes necessary to align them – to a large extent – with the ever-evolving phonetic and phonological system of the target language.

The comparison tables presented in this study provide a comprehensive analysis of the adaptation rates observed in both borrowing and adaptation processes, shedding light on the phonetic and phonological adjustments that occurred over time. The results contribute to our understanding of loanword adaptation in language contact situations, highlighting the intricate dynamics between borrowing and adaptation. These findings offer valuable insights into the linguistic dynamics between West Old Turkic and Hungarian, serving as a foundation for further investigations into language contact phenomena.

Patterns of adaptation were also observed in this research, demonstrating the phonological tendencies and systematic figures in the nativization of loanwords in Hungarian. The patterns identified, such as heavy syllables, nasal palatalization, adaptation of specific sounds like [ï] and [ɣ], labialization of vowels, and others, reveal the consistent behavior of phonological constraints during the integration of borrowings, more often than not. These patterns support the validity of the nativization process and indicate that adaptations are mostly not arbitrary but follow systematic rules. This study provides valuable insights into the intricate processes of loanword adaptation in language contact situations. The findings contribute to our understanding of the borrowing and adaptation dynamics between West Old Turkic and Hungarian, revealing the phonetic and phonological adjustments that occur during the integration of loanwords. Further research can build upon these findings to explore additional aspects of language contact phenomena and deepen our understanding of language evolution and change.

# References


Agyagási, K. (2019). *Chuvash Historical Phonetics: An areal linguistic study with an Appendix on the Role of Proto-Mari in the History of Chuvash Vocalism*. Wiesbaden: Harrassowitz Verlag.

Alós i Font, H. (2014). Chuvash Language in Chuvashia's Instruction System: An Example of Educational Language Policies in Post-Soviet Russia. *Journal on Ethnopolitics and Minority Issues in Europe: JEMIE 13*(4), 52–84.

van der Ark, R., Mennecier, P., Nerbonne, J., & Manni, F. (2007). Preliminary identification of language groups and loan words in Central Asia. In P. Osenova et al. (Eds.), *Proceedings of the RANLP Workshop on Computational Phonology* (pp. 13–20). Borovets. Retrieved from: http://www.let.rug.nl/nerbonne/ (11.11.2022)

As-Sammer, M. A. (2015). Phonetic and Phonological Adaptations of English Loanwords into Iraqi Arabic: A Generative Study. *Journal of the College of Arts, 73*, 1–46.

Babych, B. (2016). Graphonological Levenshtein Edit Distance: Application for Automated Cognate Identification, Baltic Journal of modern computing, 4(2), (pp. 115–128).

Bárcsi, G. (1972). Quelques Conclusions Tirées De L'Étude Des Plus Anciens Mots D'Emprunt Turcs Du Hongrois. *Acta Orientalia*, 10(15), (pp. 383–390).

Bekar, İ. P. (2013). Türkçe ve Macarcanın ses yapılarının karşılaştırılması üzerine bir gözlem. *Dil Dergisi, 160*, 66–81.

Benkő, L., & Imre, S. (1972). *The Hungarian Language*. Budapest & Paris: Mouton, The Hague-Paris, & Akadémiai Kiadó.

Boersma, P., & Hamann, S. (2009). Loanword adaptation as first-language phonological perception. In A. Calabrese, & W. L. Wetzels (Eds.), *Loan Phonology* (pp. 11–58). Amsterdam/Philadephia: John Benjamins.

Broselow, E. (2009). Stress adaptation in loanword phonology: perception and learnability. In P. Boersma, & S. Hamann (Eds.), *Phonology in Perception* (pp. 191–234). Mouton de Gruyter.

Brown, C. H. et al. (1994). Lexical Acculturation in Native American Languages [and Comments and Reply]. *Current Anthropology, 35*(2), 95–117.

Budenz, J. (1871). Jelentés Vámbéry Á. magyar-török szóegyezéseiről. *Nyelvtudományi Közlemények, 10*, 67–135.

Chomsky, Noam. (1981). Lectures on Government and Binding: The Pisa Lectures. Foris Publications

Clemente, C. L., & Kingston, J. (2016). Phonological Adaptation of English Loanwords in Brazilian Portuguese: A Production Study. *Language and Speech, 59*(2), 229–252.

Clements, G. N., & Keyser, S. J. (1983). *CV Phonology: A Generative Theory of the Syllable.* Cambridge: MIT Press. Retrieved from: http://seas3.elte.hu/egg12/clements-keyser-83.pdf (17.11.2022)

Cotterell, R., Peng, N., & Eisner, J. (2014). Stochastic Contextual Edit Distance and Probabilistic FSTs. In Toutanova Kristina, Wu Hua (Eds.) *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics* (pp. 625–630). Baltimore, Maryland: Association for Computational Linguistics.

Creanza, N., Ruhlen, M., Pemberton, T., Rosenberg, N., Feldman, M., & Ramachandran, S. (2015). A comparison of worldwide phonemic and genetic variation in human populations. *PNAS. Proceedings of the National Academy of Sciences of the U.S.A.*,112(5), (pp. 1265–1272).

Daland, R., Oh, M., & Davidson, L. (2018). On the relation between speech perception and loanword adaptation: Cross-linguistic perception of Korean-illicit word-medial clusters. *Natural Language & Linguistic Theory, 37(3)*, (pp. 825–868).

Danesi, M. (1985). *Loanwords and phonological methodology*. Ottawa: Studia Phonetica.

Davidson, L. (2007). The relationship between the perception of non-native phonotactics and loanword adaptation. *Phonology, 24(2)*, 261–286. https://doi.org/10.1017/S0952675707001200

Dijkstra, T., Wahl, A., Buytenhuijs, F., Van Halem, N., Al-Jibouri, Z., De Korte, M., & Rekké, S. (2018). Multilink: A computational model for bilingual word recognition and word translation. *Bilingualism: Language and Cognition*, 22(4), (pp. 657–679). doi:10.1017/S1366728918000287

Dohlus, K. (2005). Phonetics or Phonology: Asymmetries in Loanword Adaptations - French and German Mid Front Rounded Vowels in Japanese. *ZAS Papers in Linguistics, 42*, 117–135.

Downey, S., Sun, G., & Norquest, P. (2017). alineR: an R Package for Optimizing Feature Weighted Alignments and Linguistic Distances. *R JOURNAL*, 9(1), (pp. 138–152).

Drubin, D. G., & Kellogg, D. R. (2012). English as the universal language of science: opportunities and challenges. *Mol Biol Cell, 23*(8), 1399. https://doi.org/10.1091/mbc.E12-02-0108

Düzgün, Ü. K. (2014). Türk–Macar Kültüründe Ortak Unsurlar. In Salih Ünver (Eds.), *Uluslararası Türkiye-Macaristan İlişkileri Sempozyumu Bildiri Kitabı*, 51, Budapest: Halk kültürü araştırmaları kurumu.

Erdal, M. (2004). *A Grammar of Old Turkic. Handbook of Oriental Studies. Section 8 Uralic & Central Asian Studies Volume 3.* Brill Academic Publishers.

Fenyvesi, A., & Zsigri, G. (2006). The Role of Perception in Loanword Adaptation: The Fate of Initial Unstressed Syllables in American Finnish and American Hungarian. *SKY Journal of Linguistics, 19,* 131–146.

Fontan, L., Ferrané, I., Farinas, J., Pinquier, J., & Aumont, X. (2016). Using Phonologically Weighted Levenshtein Distances for the Prediction of Microscopic Intelligibility. *Annual conference Interspeech (INTERSPEECH 2016), Sep 2016, San Francisco, CA, United States.* Retrieved from: https://hal.archives-ouvertes.fr/hal-01474904.

GH = The *Gesta Hungarorum* of Anonymous, the anonymous notary of King Béla. A translation by Martyn Rady. Retrieved from: https://discovery.ucl.ac.uk/id/eprint/18975/1/18975.pdf (13.05.2022)

Ghio, A., Lalain, M., Giusti, L., Fredouille, C., & Woisard, V. (2020). How to Compare Automatically Two Phonological Strings: Application to Intelligibility Measurement in the Case of Atypical Speech. In Nicoletta Calzolari, Frédéric Béchet, Philippe Blache, Khalid Choukri, Christopher Cieri, Thierry Declerck, Sara Goggi, Hitoshi Isahara, Bente Maegaard, Joseph Mariani, Hélène Mazo, Asuncion Moreno, Jan Odijk, Stelios Piperidis (Eds.) *Proceedings of the 12th Conference on Language Resources and Evaluation (LREC 2020)* (pp. 1689–1694). Marseille: European Language Resources Association (ELRA).

Golden, P. B. (1992). *An Introduction to the History of the Turkic Peoples.* Wiesbaden: Harrassowitz Verlag.

Golden, P. B. (2011). *Studies on the Peoples and Cultures.* Bucarest: Editura Academiei Române.

Gooskens, C. (2005). Travel time as a predictor of linguistic distance. *Dialectologia et Geolinguistica*, 13, 38–62. doi:10.1515/dig.2005.2005.13.38

Gooskens, C., & Heeringa, W. (2004). Perceptive evaluation of Levenshtein dialect distance measurements using Norwegian dialect data. *Language Variation and Change*, 16(03), 189–207. doi:10.10170S0954394504163023

Greenhill, S. (2011). Levenshtein Distances Fail to Identify Language Relationships Accurately. *Computational Linguistics*, 37(4), (pp. 689–698). doi:10.1162/COLI_a_00073

Harris, Z. S. (1954). Distributional Structure. *Word, 10*(2-3), 146–162. https://doi.org/10.1080/00437956.1954.11659520

Haspelmath, M. (2009). Lexical borrowing: Concepts and issues. In M. Haspelmath, & U. Tadmor (Eds.), *Loanwords in the World's Languages: A Comparative Handbook* (pp. 35–54). Mouton de Gruyter.

Herdağdelen, A., Ciaramita, M., Mahler, D., Holmqvist, M., Hall, K., Riezler, S., & Alfonseca, E. (2010). Generalized syntactic and semantic models of query reformulation. In Fabio Crestani, Stéphane Marchand-Maillet, Hsin-Hsi Chen, Efthimis N. Efthimiadis, Jacques (Eds.) *SIGIR '10: Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval* (pp. 283–290). New York: Association for Computing Machinery. doi:doi.org/10.1145/1835449.1835498

Helimski, E. (2003). Areal groupings (sprachbünde) within and across the borders of the Uralic language family: A survey. *Nyelvtudományi Közlemények, 100*, 156–167.

Honti, L. (2017). *A magyar és a nyugati ótörök szókészleti kapcsolatairól.* Budapest: Tinta Könyvkiadó.

Howorth, H. (2020). V. Yüzyıldan XIX. Yüzyıla Kadar Göçebelerin Batı'ya Doğru Sürüklenişi III. Bölüm Kumanlar ve Peçenekler. *Türk Tarihi Araştırmaları Dergisi, 5*(1), 284–301.

Hyman, L. M. (1970). *The role of borrowing in the justification of phonological grammar*. California: University of California.

Jacobs, H., & Gussenhoven, C. (2000). Loan phonology: Perception, salience, the lexicon and OT. In J. Dekkers, F. van der Leeuw, & J. van de Weijer (Eds.), *Optimality Theory: Phonology, syntax, and acquisition* (pp. 193–210). Oxford: Oxford University Press.

Jaggers , Z. (2015). A Constraint-Shifting Account of Loanword Adaptation: Evidence from the early stages of dissemination. *University of Pennsylvania Working Papers in Linguistics*, *21*(1), Article 33.

Jakobson, R., Gunnar, C., Fant, M., & Halle, M. (1951). *Preliminaries to speech analysis: The distinctive features and their correlates.* Cambridge: MIT Press.

Johanson, L. (1998). The structure of Turkic. In L. Johanson, & É. Csató, *The Turkic Languages* (pp. 30–66). Routledge.

Johanson, L. (2016). *Classification of Turkic languages*. Retrieved from: http://www.turkiclanguages.com/www/classification.html (09.05.2022)

de Jong, K., & Park, H. (2012). Vowel epenthesis and segment identity in Korean learners of English. In L. Plonsky (Ed.), *Studies in Second Language Acquisition* (pp. 127–155). Cambridge University Press.

Kenstowicz, M. (2010). Loanword Phonology and Enhancement. In *Proceedings of the 2010 Seoul International Conference on Linguistics, Universal Grammar and Particular Languages, June 23–25, 2010, Seoul, Korean, South.* Retrieved from https://dspace.mit.edu/handle/1721.1/71827 (15.02.2023)

Kenstowicz, M. (2001). The role of perception in loanword phonology. *Studies in African Linguistics*, 32(1), 95–112.

Kenstowicz, M., & Suchato, A. (2006). Issues in loanword adaptation: A case study from Thai. *Lingua, 116*(7), 921–949.

Kertész, Zs. (2006). Approaches to the phonological analyses of loanword adaptation. *The Even Yearbook 7, ELTE SEAS Working Papers in Linguistics*, 1–15.

Kiefer, F. (2010). Hungarian. *Revue belge de Philologie et d'Histoire, 88*(3), 715–739.

Kincses-Nagy, É. (2013). A disappeared people and a disappeared language, the Cumans and the Cuman language in Hungary. *TDD/JofEL, Summer*, 171–186. Retrieved from: https://www.acarindex.com/pdfler/acarindex-496547dc-1671.pdf (23.05.2022)

Kis, T. (2005). A veláris i a magyarban. *Magyar Nyelvjárások, 43*, 5–26.

Kiss, J. & F. Pusztai eds. (2003). *Magyar nyelvtörténet.* Budapest: Osiris Kiadó.

Kiss, J. (2017). A magyar nyelvtörténet korszakolásának kérdésköréhez. *Magyar Nyelv, 113*(2), 129–144.

Kiss, J. (2018a). Honti László, A magyar és a nyugati ótörök szókészleti kapcsolatairól. *Magyar Nyelv, 114*(2), 225–227.

Kiss, J. (2018b). A magyar nyelvtörténet korszakolása és korszakai. In J. Kiss, & F. Pusztai eds., *A magyar nyelvtörténet kézikönyve* (pp. 43–47). Budapest: Tinta Könyvkiadó.

Kiss, J. & F. Pusztai eds. (2018). *A magyar nyelvtörténet kézikönyve.* Budapest: Tinta Könyvkiadó.

É. Kiss, K., K. Gerstner & A. Hegedűs eds. (2013). *Kis magyar nyelvtörténet.* Piliscsaba: Pázmány Péter Katolikus Egyetem Bölcsészettudományi Kar.

Kochetov, A. (2008). Phonology and phonetics of loanword adaptation: Russian place names in Japanese and Korean. *Toronto Working Papers in Linguistics, 28*, 159–174.

LaCharité, D., & Paradis, C. (2005). Category preservation and proximity versus phonetic approximation in loanword adaptation. *Linguistic Inquiry, 36*(2), 223–258.

Laki, K. (1960). The number system based on six in the Proto Finno-Ugric Language. *Journal of the Washington Academy of Sciences, 50*(4), 1–11.

Li, Y., & Liu, B. (2007). A Normalized Levenshtein Distance Metric. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(6), 1091–1095. doi:10.1109/TPAMI.2007.1078

Ligeti, L. (1981). *Codex Cumanicus with the prolegomena to the Codex Cumanicus.* Budapest: Library of the Hungarian Academy of Sciences.

Ligeti, L. (1986). *A magyar nyelv török kapcsolatai a honfoglalás elott és az Árpád-korban.* Budapest: Akadémiai Kiadó.

Lin, Y.-H. (2009). Loanword Adaptation and Phonological Theory. In Y. Xiao (Ed.), *Proceedings of the 21st North American Conference on Chinese Linguistics (NACCL-21)* (pp. 1–12). Smithfield, Rhode Island: Bryant University.

Lindsay, R. (2010). *Mutual Intelligibility Among the Turkic Languages.* Fresno: California State University.

Lisztóczki, L., Pusztay, J., & Bakay, M. (2019). Decipherment Challenges Due to Tamga and Letter Mix-Ups in an Old Hungarian Runic Inscription from the Altai Mountains. Humanities, 8(9), 421–433. [DOI: 10.3390/h8090421]

Lovins, J. B. (1975). *Loanwords and the phonological structure of Japanese.* Indiana: Indiana University Linguistics Club.

Marácz, L. (2012). The "Ugric-Turkic War" and the Origin of the Hungarian Language. *International Review of Turkish Studies, 2*(4), 8–23.

Marácz, L. (2016). *Towards Eurasian Linguistic Isoglasses: The case of Turkic and Hungarian.* Astana: House Gylym.

Marácz, L. (2018). Revisiting the theory of the Hungarian. In A. Marcantonio (Ed.), *The State of the art of Uralic Studies: tradition vs innovation: Proceedings of the 'Padua Uralic seminar' University of Padua, November 11–12, 2016.* (pp. 59–86). Roma: Universitá degli Studi di Roma 'La Sapienza'.

Marcantonio, A., Nummenaho, P., & Salvagni, M. (2001). The "Ugric–Turkic Battle": A Critical Review. *Linguistica Uralica, XXXVII,* 81–102.

McCoy, T. R., & Frank, R. (2018). Phonologically Informed Edit Distance Algorithms for Word Alignment with low resource languages. *Proceedings of the Society for Computation in Linguistics (SCiL) 2018,* 102–112.

McGhee, R. (1989). Who owns prehistory? *The Bering Land Bridge Dilemma. Canadian Journal of Archaeology, 13*, 13–20.

Mosbach, M., Stenger, I., Avgustinova, T., & Klakow, D. (2019). A Toolbox for Calculating Linguistic Distances and Asymmetries between Related Languages. In Ruslan Mitkov, Galia Angelova (Eds.) *Proceedings of Recent Advances in Natural Language Processing*, (pp. 810–818). Varna. doi:10.26615/978-954-452-056-4_094

Müller, F. M. (1855). *The Languages of the Seat of War in the East. With a survey of the three families of language, Semitic, Arian and Turanian.* London: Williams and Norgate.

Myers, J. (2009). Perception, phonology, and bilingualism. J. Paradis, & S. Mcleod (Eds.), *The Handbook of Bilingualism* (pp. 267–286). Oxford: Blackwell Publishing.

Navracsics, J. (2016). Living with two languages and cultures: The complexity of self-definition for bilingual individuals. *Alkalmazott Nyelvtudomány, 16*(2), 1–23.

Novgorodov, I., Lemskaya, V., Tokmashev, D., & Aktaş, E. (2015). On a Multidisciplinary Study of South Siberian Turkic Varieties (in Comparison with Yakut). Part I. *Procedia - Social and Behavioral Sciences, 206*, 114–122.

Pálóczi-Horváth, A. (1989). *Pechenegs, Cumans, Iasians: Steppe peoples in medieval Hungary.* Budapest: Corvina Kiadó.

Paradis, C., & LaCharité, D. (1997). Preservation and minimality in loanword adaptation. *Journal of Linguistics, 33*(2), 379–430.

Paradis, C., & Lacharité, D. (2008). Apparent phonetic approximation: English loanwords in Old Quebec French. *Journal of Linguistics*, 44(1), 87–128. doi:10.1017/S0022226707004963

Paradis, C., & Lacharite, D. (2011). Loanword Adaptation: From Lessons Learned to Findings. In J. Goldsmith, J. Riggle, & A. C. Yu (Eds.), *The Handbook of Phonological Theory.* Second Edition (pp. 751–778). Oxford: Wiley Blackwell.

Paradis, C., & Tremblay, A. (2009). Nondistinctive features in loanword adaptation: The unimportance of English aspiration in Mandarin Chinese phoneme categorization. In A. Calabrese, & W. L. Wetzels (Eds.), *Loan Phonology* (pp. 211–224). Amsterdam/Philadelphia: John Benjamins.

Peperkamp, S. (2005). A psycholinguistic theory of loanword adaptations. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society, 30*(1), 341–352.

Peperkamp, S. (2015). Phonology versus phonetics in loanword adaptations: A reassessment of English vowels in French. In J. Romero, & M. Riera (Eds.), *Oxford Handbook of Language Acquisition* (pp. 349–366). Amsterdam/Philadephia: John Benjamins.

Peperkamp, S., & Dupoux, E. (2003). Reinterpreting loanword adaptations: The role of perception. In M.J. Solé, D. Recasens & J. Romero (Eds.) *Proceedings of the 15th International Conference of Phonetic Sciences,* 367–370.

Peperkamp, S., Vendelin, I., & Nakamura, K. (2008). On the perceptual origin of loanword adaptations: experimental evidence from Japanese. *Phonology, 25*(1), 129–164. https://doi.org/10.1017/S0952675708001425

Polgárdi, K., & Rebrus, P. (1996). *There is no labial harmony in Hungarian. A government phonology approach.* Budapest: Research Institute for Linguistics, Hungarian

Academy of Sciences. Retrieved from: http://real-eod.mtak.hu/8189/1/WorkingPapersInTheTheoryOfGrammar_03-3_1996.pdf

Pystynen, J. (2017). The Proto-Uralic vocalism revisited. 20th Conference of the Finno-Ugric Studies Association of Canada. Toronto.

Rásonyi, L. (1964). *Tuna Köprüleri.* Ankara: Türk kültürünü araştırma enstitüsü.

Róna-Tas, A. (1999). Chuvash and historical morphology. *Acta Orientalia Academiae Scientiarum Hungaricae, 52*(1), 1–15.

Róna-Tas, A., & Berta, Á. (2002). *Old Turkic loanwords in Hungarian.* Budapest: Akadémiai Kiadó.

Róna-Tas, A. (2007). The khazars and the magyars. In P. Golden, H. Ben-Shammai, & A. Róna-Tas (Eds.), *The world of the Khazars. New perspectives. Selected papers from the Jerusalem 1999 International Khazar Colloquium* (pp. 269–278). Leiden & Boston.

Róna-Tas, A., & Berta, Á. (2011). *West Old Turkic: Turkic Loanwords in Hungarian.* Wiesbaden: Harrassowitz Verlag.

Salaville, S. (1914). Un peuple de race turque christianisé au XIIIe siècle: les Comans. *Revue des études byzantines, 106*, 193–208.

Sanders, N. C., & Chin, S. B. (2009). Phonological Distance Measures. *Journal of Quantitative Linguistics, 16*(1), 96–114.

Sándor, K. (2014). *A székely írás nyomában.* Budapest: Typotex Kiadó.

Savelyev, A. (2020). Chuvash and the Bulgharic languages. In M. Robbeets, & A. Savelyev (Eds.), *The Oxford Guide to the Transeurasian Languages* (pp. 446–464). Oxford: Oxford University Press.

Setiabudi, R., Iswari, S., & Rusli, A. (2021). Enhancing text classification performance by preprocessing misspelled words in Indonesian language. *Telkomnika*, 19(4), 1234–1241. doi:10.12928/TELKOMNIKA.v19i4.20369

Shoibekova, G., Odanova, S., Sultanova, B., & Yermekova, T. (2016). Vowel Harmony is a Basic Phonetic Rule of the Turkic Languages. *International Journal of Environmental & Science Education, 11*(11), 4617–4630.

Silverman, D. (1992). Multiple Scansions in Loanword Phonology: Evidence from Cantonese. *Phonology, 9*(2), 289–328.

Siptár, P. (1994). Palatalization rules in Hungarian. *Acta Linguistica Hungarica, 42*(1/2), 5–32.

Stachowski K. (2010). Quantifying Phonetic Adaptations of Russian Loanwords in Dolgan. *Studia Linguistica Universitatis Iagellonicae Cracoviensis, 127*, 101–177.

Stachowski K. (2011). A note on levenshtein distance versus human analysis. *Studia Linguistica Universitatis Iagellonicae Cracoviensis, 128*, 155–160.

Stachowski, K. (2013). The influx rate of Turkic glosses in Hungarian and Polish post-mediaeval texts. In R. Köhler, & G. Altmann (Eds.), *Issues in Quantitative Linguistics 3* (pp. 100–116). Lüdenscheid: RAM Verlag.

Stachowski, M. (2014). Remarks on the investigation of the oldest layer of Turkic loanwords in Hungarian. *Studia Etymologica Cracoviesia,19,* 215–222.

Stammler-Gossmann, A. (2009). A life for an idea: Matthias Alexander Castrén. *Polar Record, 45*(3), 193–206.

Steffanides, G. F. (1965). The Role of Greek and Latin in Science. *The American Biology Teacher, 27*(10), 785–789.

Surhone, L. M., Timpledon, M. T., & Markesen, S. F. (2010). *Ural-Altaic Languages: Language Family, Uralic Languages, Matthias Castrén, Finno-Ugric Languages, Mongolic Languages, Tungusic Languages, Genetic Relationship, Taxon.* Betascript Publishing.

Tülücü, S. (1997). Arapça ve Farsçanın Türkçeye tesiri. *Atatürk Üniversitesi Ilahiyat Fakültesi Dergisi, 0*(13), 31–51. Retrieved from: https://dergipark.org.tr/tr/download/article-file/30806 (02.02.2022)

Uffmann, C. (2006). Epenthetic vowel quality in loanwords: Empirical and formal issues. *Lingua, 116*, 1079–1111. https://doi.org/10.1016/j.lingua.2005.06.009

Vakulenko, M. (2021). Calculation of Phonetic Distances between Speech Sounds. *Journal of Quantitative Linguistics, 28*(3), 223–236. doi:https://doi.org/10.1080/09296174.2019.1678709

Vámbéry, Á. (1870). Magyar és török-tatár szóegyezések. *Nyelvtudományi Közlemények, 8*.

Varga, G. (1999). *The Origins of Hunnish Runic Writing.* Budapest: Írástörténeti Kutató Intézet.

Vovin, A. (2005). *The End of the Altaic Controversy In Memory of Gerhard Doerfer. Central Asiatic Journal, 49(1)*, 71–132.

Wieland, G. (2012). Language Contact. In A. Bergs, & L. Brinton, *English Historical Linguistics: An International Handbook* (pp. 362–373). De Gruyter Mouton.

Wollman, A. (1993). Early Latin loan-words in Old English. *Anglo-Saxon Englad, 22*, 1–26.

Yılmaz, E. (2015). Altayistik bağlamında Macarca-Türkçe dil ilişkileri. *Türk dünyası sosyal bilimler dergisi, 74*, 287–308.

Yip, M. (1993). Cantonese loanword phonology and optimality theory. *Journal of East Asian Linguistics, 2*(3), 261–291.

Zhang, L. (2018). *A More Sensitive Edit-Distance for Measuring Pronunciation Distances and Detecting loanwords.* Master's Thesis. Groningen & Malta: University of Groningen & University of Malta.