Werk

Titel: Number 2

Ort: Wiesbaden

Jahr: 2015

PURL: https://resolver.sub.uni-goettingen.de/purl?666048797_0019|LOG_0021

Kontakt/Contact

Digizeitschriften e.V.
SUB Göttingen
Platz der Göttinger Sieben 1
37073 Göttingen

info@digizeitschriften.de

Contents

Turkic Languages, Volume 19, 2015, Number 2

Editorial note

Turkic Languages, Volume 19, 2015, Number 2

The present issue of TURKIC LANGUAGES begins with Heidi Stein's obituary of the Hungarian Turcologist György Hazai (1932–2016).

Öner Özçelik deals with word-level prominence in Kazakh and Modern Uyghur. It is generally assumed that most Turkic languages behave like Turkish in realizing final prominence. According to the author, the situation is more complex. Modern Uyghur and Turkish mark prominence only by a pitch rise, not by greater intensity or duration. The Kazakh distribution of stress is similar, though the nature of the prominence is different. Kazakh uses duration and/or intensity in addition to pitch. The situation may be explained through influence of Russian, but it must be confirmed that the findings also hold true for Kazakhs not influenced by Russian.

Peter Sauli Piispanen deals with loans related to reindeer terminology in several languages of north-eastern Siberia. There are numerous known (Pre-)Yakut and Tungusic borrowings in the Yukaghiric languages. The author proposes and discusses eleven new etymologies, Turkic borrowings in Yukaghiric and Uralic, and likely borrowings or cognates between Uralic and Yukaghiric.

Kentaro Suganuma proposes three phonological constraints on the distribution of consonants in Turkish mimetic words as representatives of sound symbolism. Two of them are related to the markedness of consonants, and one is related to assimilation of consonants. The author also discusses some phonological similarities and differences between Turkish mimetic and non-mimetic words.

Umarjan Kurban offers a descriptive analysis of canonical causative constructions in Modern Uyghur. It is shown that double causatives derived from intransitive and transitive bases result in different surface realizations. Unlike other Turkic languages such as Turkish, Modern Uyghur does not permit triple causative constructions.

Lars Johanson deals with four degrees of grammaticalization of nominal relators in Kazakh.

Éva Á. Csató, Hüner Kaşıkara, Beáta Megyesi, and Joakim Nivre report on an Uppsala-based project on Turkish parallel corpora and on the application of the annotation scheme Universal Dependencies to Turkic.

Henryk Jankowski reviews the volume *Kazakh in Post-Soviet Kazakhstan*, edited by Raihan Muhammedova. Finally, Ümit Deniz Turan reviews the volume *Ankara Papers in Turkish and Turkic Linguistics*, edited by Deniz Zeyrek, Çiğdem Sağın Şimşek, Ufuk Ataş, and Jochen Rehbein.

Lars Johanson

Obituary

In memoriam György Hazai (1932–2016)

Heidi Stein

Stein, Heidi 2015. György Hazai 1932–2016). *Turkic Languages* 19, 151–155.

Heidi Stein, Johannes Gutenberg University of Mainz, Hegelstraße 59, DE-55122 Mainz, Germany. E-mail: Heidi Stein <hstein@uni-mainz.de>

On January 7, 2016, Prof. Dr. Dr. h.c. mult. György Hazai, the great Hungarian Turcologist, Ottomanist and linguist, passed away in Budapest shortly before his 84th birthday. His scientific work was always internationally oriented, but thanks to his twenty-year stay at Humboldt University Berlin, he rendered great service especially to German Turcological studies.

Hazai was born on April 30, 1932, in Budapest. From 1950 until 1954 he studied Oriental Philology and Turcology at the Eötvös Loránd University in Budapest under the guidance of the eminent scholars Gyula Németh, Lajos Fekete and Lajos Ligeti, as well as Slavic studies under István Kniezsa. In 1956 he was offered a position as visiting lecturer for three terms at the University of Sofia where he also found impulses and materials for his first scientific fields of interest: Turkish dialects and Turkish transcription texts. In Sofia, he began his career as an academic teacher. His students from that relatively brief period—some of whom are today successful scholars at the Department of Turkish Studies at Sofia University—still remember with much gratitude his helpfulness and energy. Returning to Budapest in

1958, he became a Research Fellow at the Hungarian Academy of Sciences and in the same year received his Ph.D. degree for a study on *Problems of the Turkish linguistic monuments in Cyrillic script from Bulgaria*.

In 1962 he was appointed by the initiative of the Finno-Ugrist Wolfgang Steinitz of the Academy of Sciences in Berlin to rebuild the formerly famous Turcological studies program there. In this, Nâzim Hikmet, with whom Hazai was rather well acquainted, played an intermediary role, because Hazai had translated some of Hikmet's works. So in 1963 Hazai began his work, first as Visiting Associate and then Full Professor, at the Near Eastern Institute ("Vorderasiatisches Institut") of Humboldt University. In 1964 he habilitated at the Hungarian Academy of Sciences with a study on *The Turkish text of Jakab Nagy de Harsány. Linguistic problems of Ottoman Turkish in the 17th-century*. It is thanks to Hazai that a Turcological course of studies, independent from Oriental studies in general, and also from the politically determined area studies ("Regionalwissenschaften"), could be established and successfully developed. Humboldt University remained the only university in Eastern Germany to offer a Turcology program. Having been one of his first students in 1963, I will always remember his inspiring attitude, trying to fill us with some enthusiasm for Turkish philology, and his active support when we were taking our first steps in this field. Still rather young, he had already acquired broad scientific experience, but he generally treated the students and young assistants as equals and took their imperfect efforts seriously. Hazai's numerous contacts with scholars in Turkey and other countries were also a great advantage for his young collaborators when entering into the international exchange of scientific ideas. One highlight in this regard was his organization of the XIIth Meeting of the Permanent International Altaistic Conference in Berlin in 1969.[1] This was the first Altaistic conference where a large number of scholars from East and West could meet each other.

Another scientific field in Berlin is connected with Hazai's name: the intensive study of the Turfan texts. From the beginning, it was Hazai's task to revitalize the research on this eminent collection. Together with Wolfgang Steinitz, he founded in 1963 the *Turfan Research Group* at the Oriental Institute of the Berlin Academy of Sciences. He was its head until 1974, and also promoted its continuing research in many ways. Some of Hazai's first Turcological graduates took part in the early research on the Old Turkic Turfan texts, and have continued until today.

Hazai's work in Berlin, which was also characterized by his efforts to create a friendly and cooperative scientific climate between Turcological colleagues from his institute and others, ended in 1982. He returned to Budapest, working there from 1984 to 1990 as Research Professor and Publishing Director of the Hungarian Academy of Sciences. He became Honorary Professor of the University of Budapest.

1 Hazai, Georg & Zieme, Peter (eds.) *Sprache, Geschichte und Kultur der altaischen Völker. Protokollband der XII. Tagung der Permanent International Altaistic Conference 1969 in Berlin*. Berlin 1974.

During 1992–1999, he took up a new challenge: He was appointed as Head of the newly founded Department of Turkish Studies at the University of Cyprus in Nicosia. With his authority, balanced personality, and broad international contacts, he was able to do good work at this young university that was developing in a difficult political situation. Even at the age of 70 he took on another official position of responsibility, discharging his duties with his characteristic energy and diplomatic skills. From 2002 until 2003 he acted as the Founding Rector of the Andrássy University of Budapest, a German-speaking university which plays an important role in the academic cooperation between Hungary, Germany, Austria and Switzerland.

Alongside all his official work, however, he continued his research, his publishing and his scientific organizational activities. His main areas of interest within Turcology were the linguistic history of Ottoman Turkish, Turkish dialectology (especially Balkan-Turkish), Ottoman history, and Old Turkic.[2]

In the field of the linguistic history of Turkish he focused on the problems of phonology and morphophonology, and the appropriate sources for solving them. Early on he directed his attention to a special group of sources: the so-called transcription texts, or better, texts in non-Arabic scripts. His habilitation thesis on the texts of Harsány[3] became a standard methodological work on how to analyze such texts for results relevant to language history. He continued these studies with many articles about shorter texts and about methodological principles in this field.[4] By means of those texts, which mainly belong to the Middle Ottoman period from the 16th until the 18th century, he developed criteria for the periodization of the Turkish language history.

His interest in dialectology, especially the so-called Rumelian dialects, was awakened and promoted by his teacher Gyula Németh, and like Németh, Hazai approached the field from the perspective of language history. The same is true of his studies on different forms of language contact in the Balkans.[5]

2 See the overview of his research themes in Barbara Kellner-Heinkele et al. (eds.) *Monumenta et Studia Turcologica. Ausgewählte Schriften von György Hazai.* Berlin 2012.

3 *Das Osmanisch-Türkische im XVII. Jahrhundert. Untersuchungen an den Transkriptionstexten von Jakab Nagy de Harsány*. Budapest 1973.

4 For the special group of texts in Cyrillic script see: Monuments linguistiques osmanlis-turcs en caractères cyrilliques dans les receuils de Bulgarie. *Acta Orientalia Academiae Scientiarum Hungaricae* [AOH] 11 (1960), 221–233. For methodological principles see e.g.: Zum balkanischen Hintergrund der osmanisch-türkischen Transkriptionstexte von Bartholomaeus Georgievits. *Studia Slavica Academiae Scientiarum Hungaricae* 20 (1974), 71–106.

5 See e.g. Les dialectes turcs du Rhodope. *AOH* 9 (1959), 205–222; Remarques sur les rapports des langues slaves des Balkans avec le turc-osmanli. *Studia Slavica Academiae Scientiarum Hungaricae* 7 (1961), 97–138; Mes enquêtes sur les parlers turcs des Balkans. In: Van Windekens, A. J. (ed.) *Communications et rapports du Premier Congrès International de Dialectologie Générale 2.* Louvain 1964, 85–90; Probleme und Aufgaben der Balkan-Turkologie (Sprachwissenschaft). In: *Association Internationale*

From the beginning Hazai was concerned with Ottoman manuscripts. He published Ottoman documents connected with historical studies, and later a *Catalogue of the Turkish manuscripts in the Library of the Hungarian Academy of Sciences.*[6] For the solution of problems in language history, he always strived to bring together all relevant groups of sources; "synthesis" was one of his favored terms in this respect. Hence, in his later years he did a great deal of work with Ottoman texts in Arabic script, i.e. vocalized manuscripts that can also shed light on the phonology of Old Anatolian and Middle Ottoman Turkish. Four important texts which are also of literary and historical interest were made available by him in critical editions.[7] One of them, the edition of *Ferec baᶜd eş-şidde*, was realized in close collaboration with his longtime colleague and friend Andreas Tietze.

His quest for synthesis also led to two works on the state of the art in the linguistics of Ottoman and Modern Turkish: his *Kurze Einführung in das Studium der türkischen Sprache* (1978, in Turkish 2012),[8] and, the *Handbuch der türkischen Sprachwissenschaft* (1990) initiated and edited by him.[9]

As head of the *Turfan Research Group*, Hazai promoted wide-ranging publication activities. He founded the series *Berliner Turfantexte*, which regularly makes available the texts of this eminent collection, and he personally took part in producing editions of Old Turkic texts.[10]

d'Études du Sud-Est Européen. 1er Congrès International des Études Balkaniques et Sud-Est-Euro–péennes. Résumés des communications. Linguistique. Sofia 1966, 92–100; Beiträge zu einigen Problemen der Lehnwörterforschung in den osmanisch-türkischen Mundarten des Balkans. *AOH* 18 (1965), 183–190; Die Turzismen in den Sprachen Südosteuropas. Erfassung, Bedeutung, Status. In: Majer, Hans Georg (ed.) *Die Staaten Südosteuropas und die Osmanen.* München 1989. 205–214.

6 Parlatır, İsmail & Hazai, György & Kellner-Heinkele, Barbara: *Catalogue of the Turkish manuscripts in the Library of the Hungarian Academy of Sciences*. Budapest 2008.

7 Ein "İskendernāme" als politische Zweckschrift aus der Zeit von Süleyman dem Prächtigen. *Archivum Ottomanicum* [AO] 15 (1997), 221–308, *AO* 16 (1998), 125–277, *AO* 18 (2000), 251–305; Hazai, György (ed.) *Die altanatolisch-türkische Übersetzung des Taẕkaratu l-Awliyā von Farīduddīn 'Aṭṭār* 1–2. Berlin 2008; Hazai, György & Tietze, Andreas (eds.) *Ferec baᶜd eş-şidde. "Freud nach Leid". (Ein frühosmanisches Geschichtenbuch)* 1–2. Berlin 2008; Hazai, György (ed.) *Die Geschichte der Ungarn in einer osmanischen Chronik des 16. Jahrhunderts: Tercümān Mahmūds Tārīh-i Ungurus*. Berlin 2009.

8 *Kurze Einführung in das Studium der türkischen Sprache*. Budapest & Wiesbaden 1978. A revised and completed translation is: *Türkiye Türkçesinin dünü ve bugünü. Türk dili araştırmalarına kısa bir giriş.* Ankara 2012.

9 Hazai, György (ed.) *Handbuch der türkischen Sprachwissenschaft* 1. Budapest & Wiesbaden 1990.

10 E.g. the reprint of editions of the Berlin tradition in Turfan research: *Sprachwissenschaftliche Ergebnisse der deutschen Turfan-Forschung. Text-Editionen und Interpretationen von Albert August von Le Coq, Friedrich Wilhelm Karl Müller, Willy Bang, Annemarie*

Hazai's general interest in bringing together Turcological scholars from different countries, and to contribute to a better exchange of their publications, resulted in his editing of several bibliographies. He started with the bibliographical monograph *Sovietico-Turcica* (1960),[11] and continued with *Bibliographisches Handbuch der Turkologie* (1986),[12] and the vital series *Turkologischer Anzeiger—Turcology annual* (since 1975).

He was also the editor or co-editor of several journals and monograph series such as *Studien zur Sprache, Geschichte und Kultur der Turkvölker*, *Bibliotheca Orientalis*, *Archivum Ottomanicum*, *Zeitschrift für Balkanologie*, and served on the advisory boards of other prestigious series and journals. He held posts in international bodies like Vice-President of the *Comité International des Études Pré-Ottomanes et Ottomanes*, Secretary General of the *International Union of Oriental and Asian Studies*, president of international congresses, and others.

In recognition of his prolific scientific work and its extraordinarily international scope a vast number of honors and awards were bestowed on Hazai such as Doctor honoris causa of the *University of Ankara* and the *University of Istanbul*; Honorary member of the *Turkish Linguistic Society*; Ordinary member of the *Hungarian Academy of Sciences*; Honorary Member of the *Turkish Academy of Sciences*; Honorary Professor of the *Oriental Institute of the Russian Academy of Sciences*; Honorary member of *American Oriental Society*, *Deutsche Morgenländische Gesellschaft*, *Societas Uralo-Altaica*. He was awarded Orders of Merit by the Federal Republic of Germany, the Republic of France, and the Republic of Turkey, as well as other national and international honors.

A great number of colleagues and friends came together at the touching memorial ceremony in Budapest, to pay their last respects to the great scholar. In spite of the sad occasion they were happy to learn that György Hazai's legacy is in good hands. His daughters Kinga and Cecilia have decided to create an archive of his works and make his extraordinarily valuable private collection of books available to the public. By 2017 his home will be converted into a library dedicated to research and study for scholars in Turcology and Oriental Studies wherever they may come from. György Hazai, who did not spend a single day away from his beloved work, would especially appreciate this kind of honor and remembrance.

von Gabain, Gabdul Rašid Rachmati, Wilhelm Thomsen. Gesammelte Berliner Akademieschriften 1908–1938. Mit Vorwort von Georg Hazai 1–2. Leipzig, 1972.

11 *Sovietico-Turcica. Beiträge zur Bibliographie der türkischen Sprachwissenschaft in russischer Sprache in der Sowjetunion 1917–1957*. Budapest 1960.

12 Hazai, György & Kellner-Heinkele, Barbara: *Bibliographisches Handbuch der Turkologie. Eine Bibliographie der Bibliographien*. Budapest & Wiesbaden 1986.

Degrees of grammaticalization of Kazakh nominal relators

Lars Johanson

Johanson, Lars 2015. Degrees of grammaticalization of Kazakh nominal relators. *Turkic Languages* 19, 156–162.

The paper deals with degrees of grammaticalization of nominal relators in Kazakh, suggesting a new kind of systematic classification. Grammatical relators are understood as free or bound elements occurring after nominals to mark their syntactic functions. Four levels of relators are distinguished: (i) case relators, (ii) opaque postpositional relators, (iii) less grammaticalized relators, and (iv) complex postpositional relators. The four levels represent decreasing degrees of grammaticalization.

Lars Johanson, Institut für Slavistik, Turkologie und zirkumbaltische Studien, Johannes Gutenberg-Universität, Hegelstraße 59, D-55122 Mainz, Germany.
E-mail: johanson@uni-mainz.de

Introduction

The present paper deals with degrees of grammaticalization of nominal relators in Kazakh, suggesting a new kind of systematic classification.[1] Grammatical relators are understood as free or bound elements occurring after nominals to mark their syntactic functions. Nominals are parts of speech that can be inflected for number, possession, and case: nouns, nominalized adjectives, pronouns, and numerals. An extended nominal is a nominal group larger than a single nominal.

The left-branching syntax of Kazakh has a rich system of grammatical relators functioning as postposed case markers and postpositions. The relators form relator phrases with their complements, linking these to other elements in the clause, in particular to verbs.

Level 1: Case relators

At level 1, the level of maximal grammaticalization, the relators are bound case markers. They combine morphologically with their hosts, forming phonological words with them and often showing sound harmony.

Kazakh cases cover semantically broad, less differentiated relational concepts. As is well known, five oblique cases are expressed by accentable suffixes, genitive

1 The paper is based on a talk presented at the *International Symposium on Kazakh Philology*, 26–27 October 2013, Minzu University, Beijing.

{-NIŋ}, accusative {-NI}, dative {-GA}, locative {-DA}, and ablative {-DAn}. The markers are slightly different after possessive suffixes. The accusative is marked with {-n} after third-person possessive suffixes. A so-called "pronominal *n*" appears in locative {-ndA} and ablative {-nAn} < {-ndAn}. Personal and demonstrative pronouns have a different declension, displaying some irregular forms, e.g. *men* 'I', genitive *meniŋ*, accusative *meni*, dative *maɣan*, locative *mende*, and ablative *menen*. The final *-l* in the demonstratives pronouns *bul, ol, sol* is replaced by *-n* in oblique cases except in the dative and in some ablative forms such as *bu-dan.* The old compound suffix {-Iŋ-KI} exhibits a variation, e.g. *men-iŋ-ki* ~ *men-i-ki* 'mine'. The *i* of the latter form looks like an accusative marker but may be a shortened genitive or an oblique stem form.

There have been speculative attempts to trace core case suffixes back to originally independent words, particularly nouns. Gerhard Doerfer (1977: 208–214) has suggested that the Orkhon Turkic dative marker {-KA}, which exhibits adlocational (allative, directive, terminative) functions, might go back to a Proto-Turkic root **ḳa* 'near(ness)', present in words such as *ḳat* 'side', *ḳat-* 'to join (tr.)' and East Old Turkic *ḳa:* 'family' (Clauson 1972: 578a). The markers may have developed from early level 2 postpositions (see below), though this is impossible to prove.

Kazakh relators of level 1 do not govern other cases. There are no complex forms of pronominal declension of the type documented in East Old Turkic, e.g. locatives such as *biz-ịn-dä* 'in us' instead of *biz-dä.* There are no traces of composite case suffixes, so-called 'double declension' markers, which are common in Mongolic and Tungusic and also occur in Yakut and Dolgan (e.g. comitative + accusative). It is highly improbable that the dative marker {-KA} goes back to two Proto-Turkic elements, as has been suggested by some scholars. The alleged second element has been compared to the Mongolic dative-locative marker {-A} (Tekin 1968: 130). The original dative suffix is most likely {-KA}, not {-A}.

There are other level 1 relators which are not core case markers, e.g. the unstressed bound equative marker {-šA}, developed from {-čA} and mostly expressing comparison ('like', 'as'), e.g. *ḳazaḳ-ša* '(in) Kazakh'.

Level 2: Opaque postpositional relators

Kazakh has, as other Turkic languages, a rich system of grammatical relators functioning as postpositions. Though they have functions similar to those of case markers, they form a distinct grammatical category. The number of cases is very limited, whereas there are numerous simple and complex postpositions. Postpositions mostly occur as free wordforms, but may also have bound variants, sometimes showing sound harmony. Postpositions may contribute to various kinds of complement phrases. Unlike case markers, they can even be separated from their complements by other intervening words.

Postpositions may determine the case of their complements. Some may even govern more than one case, depending on their meanings. Postpositions are fre-

quently classified according to such criteria. Here we will only deal with them as representing certain degrees of grammaticalization.

Postpositions express more differentiated relational concepts than cases do. Kazakh has numerous postpositions that can be used to express a wide range of spatial, temporal, and other semantic relations. They mostly locate their complement in space or time, but also express comparison, cause, purpose, instrument, etc.

Level 2 consists of highly grammaticalized, opaque postpositions of high age, including the four primeval Turkic postpositions described by Grönbech (1936: 35): *biːrlä:* 'with', *täːg* 'like', *üčün* 'for', *üzä:* 'over'. It is unknown what lexical elements they may go back to.

In modern languages, postpositions of this type constitute a closed class. They do not occur as other parts of speech, for instance as adverbs. They are not derived productively from words of other categories. They normally do not carry case markers. They are not used attributively, for instance by taking on the adjectivizing suffix {-GI}. They cannot be immediately preceded by quantifiers.

The East Old Turkic postposition *täːg* 'like', e.g. *täŋri täːg* 'god-like', is represented by {-DAy} in Kazakh, e.g. *ḳar-day appaḳ* 'white as snow, snow-white'. In the southwestern branch of Turkic, it has been replaced by the type *kib-i,* which belongs to level 3, a noun *kiːb* 'mould, model' plus a third-person singular possessive suffix, e.g. Ottoman *ḳar gibi* 'like snow'. In nominalized use, modern Kazakh {-DAy} can be followed by case markers in forms of pronouns and proper names, e.g. *mun-day* 'like this', genitive *mun-day-dïŋ* 'of something/someone like this', *Ḳanat-tay-dïŋ* 'of those like Kanat'.

Modern Kazakh *üšin* has prodessive, causal, and purposive functions: 'for', 'for the sake of', 'because of', 'in order to'. Its Old Uyghur equivalent *üčün* 'for' can govern the genitive or nominative case of personal and demonstrative pronouns in singular, e.g. *sen-iŋ üčün ~ sän üčün* 'for you'; cf. Uzbek *sen-iŋ üčün ~ sen üčün.* Some modern languages employ the genitive, e.g. Turkmen *θen-iŋ üčüːn* ‹seniň üçin›, Turkish ‹sen-in için›. Certain languages attach an accusative-like suffix, certainly a shortened genitive marker, to the complement, e.g. Karachay-Balkar *sen-i üčün*. Kazakh or Noghay *üšin* does not use the genitive case of pronominal complements, e.g. *sen üšin* 'for you'. The equivalents in other Turkic languages mostly occur as independent words, but they may also have bound variants, e.g. {-čUn}. Chuvash grammarians reckon with a causal or purposive case in {-šӖn}.

The old postposition *biːrlä:* 'with' survives as modern Uyghur *birlä ~ bilä ~ birlän ~ bilän*, Tuvan *bilä*, Uzbek *b(i)lán,* Turkmen ‹bile› ~ ‹bilen› etc. Its shape is rather different from that of the Kazakh and Kirghiz instrumental-comitative postposition *menen* 'with', which has equivalents in several languages, e.g. Bashkir *mi̯nän* ‹менән›. Kaare Grönbech (1936: 35) analyzed *biːrlä:* as *biːr* 'one' plus an otherwise unknown "emphatic particle" *-laː*. Sir Gerald Clauson (1972: 364b) analyzes it as an abbreviated converb of *biːrläː-* 'to unite', which would mean that it was used as an adverb. The word also came to be used as an adverb meaning 'even', e.g. Ottoman *bilä*. Kazakh *menen* is, however, not used adverbially, and clearly belongs to level 2.

Kazakh *menen* and its Kirghiz counterpart tend to occur in bound forms, as the non-harmonic enclitic {-Men}, e.g. *ušaḳ-pen* 'by air'. Neither *menen* nor {-Men} are case markers, since they themselves govern a case, namely an oblique case of singular personal and demonstrative pronouns, e.g. *men-i menen* 'with me'. The equivalents of *bi:rlä:* behave in similar ways, e.g. Uzbek *meniŋ* (genitive) ~ *men* (nominative) *blän,* Turkish ‹benim ile› ~ ‹benim-le› 'with me' (Deny 1921, § 437), the latter containing the harmonic enclitic {-(y)lA}.

The marker attached to the Kazakh complement gives the impression of being an accusative suffix (Geng Shimin et al. 1999: 205). It can, however, be a shortened form of the genitive marker; cf. *biz-di-ki* 'ours' and *men-iŋ-ki* ~ *men-i-ki* 'mine' (Menges 1959: 469). Zhang Dingjing (2004: 212) states that a high unrounded vowel is inserted before {-Men}, when it is attached to pronouns such as *men*, *sen*, and *ol*. This is a correct synchronic description. Diachronically, however the element *i* attached to the complement, e.g. in *men-i menen* 'with me', may reflect an old stem-final oblique segment (Menges 1959: 467). As noted, singular personal and demonstrative pronouns exhibit oblique stems ending in a "pronominal *n*" in many Turkic languages. East Old Turkic thus shows locatives such as *biz-iṇ-dä* ← *biz* 'we' rather than *biz-dä.* Kazakh has oblique pronominal stems in locatives such as *bun-da* 'in this'. In Uzbek, which has lost the "pronominal *n*" in the nominal declination, the case markers {-gȧ} (dative), {-dȧ} (locative), {-dȧn} (ablative), {-čȧ} (equative), and a few other suffixes are attached to the oblique stem of certain pronouns, e.g. *bun-dȧ* 'in this', *bun-dȧy* 'like this'. Oblique pronominal stems of this kind also occur in Mongolic, Tungusic, and Japanese.

Level 3: Less grammaticalized relators

Level 3 comprises less grammaticalized postpositions. They are formally indistinguishable from other parts of speech such as adverbs, adjectives, or nouns.

Postpositions sharing their shape with adverbs are often petrified converbs, e.g. *ḳara-y* 'toward' ← *ḳara-* 'to look', *gör-i* 'in comparison with, rather than' ← *kör-* 'to see', *bol-a* 'for' ← *bol-* 'to (be)come', *say-ïn* 'each' (< 'counting' ← *say-* 'to count'), e.g. *kün sayïn* '(counting) each day'. Other types contain adverbializing markers, e.g. *bas-tan* 'from ... on' ← *bas* 'head'. The type *boy-ïn-ša* 'according to', provided with the bound equative marker {-šA}, rather represents level 4 (below).

Postpositions sharing their shape with adjectives include the type *basḳa* < *bašḳa* 'other (than)', 'except (for)', 'apart (from)', 'excluding', 'without' ← *baš* 'head', governing the ablative case, e.g. *bu-dan basḳa* 'other than this'. Relator phrases at this level may be formally identical to adjective phrases with a complement that modifies a noun, e.g. *üy-ge žaḳïn* 'close to the house' < *yaḳïn* 'near' ← *yaḳ-* 'to approach, to be near'. Another type, containing adjectivizing markers, is represented by *sïyaḳ-tï* 'like' and *tuwra-lï* 'about', both governing the nominative, e.g. *sen tuwra-lï* 'about you'.

A number of postpositions consist of plain nouns without any adverbial endings. For example, Kazakh *burun* ‹бұрын› ‘before’ has clearly nominal properties. It goes back to the noun *burïn* ‘nose’, ‘beak’, which was also used for other protruding objects, generalized in the sense of ‘place in front (of)’. In the temporal sense, it did not mean ‘time still to come’ (time following the moment of speaking) but, on the contrary, ‘earlier (than)’, ‘prior (to)’, e.g. Karakhanid *bu kün-dä burun* ⟨this day-ABL before⟩ ‘before today’. According to Maḥmūd al-Kāšγarī, it was used for ‘the first (*al-awwal*) of anything’ (Clauson 1972: 366b–367a); cf. postpositions of the type *äv:äl* in several Turkic languages. Kazakh *burun* is regarded as an adverb that is also used as a postposition. It basically seems to be a noun meaning ‘time or place before’. It can take on case markers and be used attributively, provided with the suffix {-GI}, *burïn-γï* ‘previous’. It can also be immediately preceded by quantifiers, e.g. *keški tamaḳtan eki saγat burun* ⟨dinner-ABL two hour before⟩ ‘two hours before dinner’. Both features are impossible in the case of postpositions of level 2.

In the same way, Kazakh *keyin* is basically a noun meaning ‘time or place after’ with the adjectival form *keyin-gi*. It goes back to *ke:-dịn*, a denominal noun ← **ke:* ‘place behind’, ‘time after’. The East Old Turkic marker {-dỊn} forms denominal place nouns, e.g. *Säläŋä ke:-dịn* ‘the place behind (west of) the Selenga’. It is more or less synonymous with the old word *ke:n* ‘after’, an allegedly instrumental form of **ke:,* used as an adverb or postposition, meaning ‘place behind’, ‘time after (later)’, e.g. *män-dä ke:n* ‘after me’. It can be declined as a noun, which, according to Clauson 1972: 724b, might mean that its origin had been forgotten. It rather seems to be a noun which, like *ke:-dịn,* could occur in both adverbial and postpositional functions.

Kazakh *soŋ* ‘end’ is a noun that can be used as a postposition without an adverbial ending. Like other nouns, it can take on the adjectivizing suffix {-GI}, i.e. *soŋ-γï.* The older *soŋï-ra* ‘afterwards’, ‘after’ is a form in {-rA} ← *soŋ* ‘end’. Though the East Old Turkic marker {-rA} is mostly said to form locational adverbs, it rather forms nouns that can be used both locationally and adlocationally, e.g. *ič-rä* ‘in, inward’ ← *ič* ‘interior’, *taš-ra* ‘out, outward’ ← *taš* ‘exterior’, *öŋ-rä* ‘in front of, forward’ ← *öŋ* ‘front’, *as-ra* ‘below, downward’. The derived forms are not primarily adverbs, but place nouns, e.g. Ottoman *taš-ra* ‘the outside’, ‘the provinces’. Similarly, *soŋï-ra* had nominal properties, forming nouns meaning ‘time or place after’. Compare East Old Turkic *üzä-rä*, one of the four primeval postpositions that was also employed as an adverb. It replaced *üzä:* ‘over’, originally an adverb ‘above’ and also was used as a postposition meaning ‘above, upon, on’, governing the nominative, occasionally locative or genitive of pronouns (Clauson 1972: 280b).

Kazakh *beri* ‘since’ is a deictic space- and time-related postposition which also occurs as an adverb ‘here’, ‘hither’, ‘on this side’. It originally contained {-rU} an alleged directive suffix that rather is a marker of place and time nouns. It was often used in antithesis to *aŋaru* ‘that place/time', ‘that side’. Like other nouns, it could take on an adjectivizing suffix, e.g. *bärü-ki* ‘(situated) towards this side’. Compare Kazakh *keri* ‘backward place’ < *ke:rü* ← **ke:* ‘place behind’, Turkish ‹geri›.

The markers of place and time nouns are comparable to space nouns such as Kazakh *žoγarï*. The old form *yoːḳaruː* 'upper place' < **yoːḳ-ɣaruː* is derived from *yoːḳ* 'high ground' by means of the suffix {-GArUː}, which is a space-noun suffix rather than a directive marker. In adverbial expressions with verbs expressing motion, such nouns may occur in the meaning 'up', or 'upwards' without an adverbial marker of direction or goal. Thus Turkish ‹yukarı› 'upper place', can, if employed adlocationally, be synonymous with ‹yukarı doğru› 'upwards'. Used together with dynamic verbs it does not need a dative marker to express direction or goal (Johanson 2012).

Level 4: Complex postpositional relators

The complex relators found at level 4 are the least grammaticalized, most transparent relators. They provide more differentiated semantic specifications than other relators do, particularly expressing various exact spatial and temporal relations. They are based on auxiliary nouns provided with possessive suffixes and case markers. The possessive suffixes agree with the preceding complement, which is in the genitive or the nominative case, e.g. *üy-diŋ ald-ïn-da* ⟨house-GEN front-POSS3SG-LOC⟩ 'in front of the house', *stol üst-in-de* ⟨table top-POSS3SG-LOC⟩ 'on the table'.

These constructions are similar to free morphological combinations, but they are petrified ("frozen") grammaticalized items; compare the status of the complex English preposition *on top of* as against the locative phrase *on the top of*. The possibility of inserting elements between the complement and the relator is also very limited. The boundary between complex postpositions and free combinations is sometimes less distinct. It may be difficult to distinguish the postpositions from the phrases they derive from, particularly as most auxiliary nouns function as regular nouns.

The complex relators generally refer to place and time. Most of them go back to nouns denoting body parts. The relevant notions are mapped onto various spaces that may be conceived of as bodies or body parts. Most postpositions express relations in space: attachment ('at'), inclusion ('in'), superposition ('on', 'over'), subposition ('under'), exteriority ('outside'), laterality ('beside'), proximity ('near'), circumposition ('around'), contraposition ('against', 'opposite to'), intermediacy ('between', 'among'), etc. Spatial relators may describe dynamic situations, implying goal, direction, delocation ('motion to', 'motion toward', 'motion away from'). Relations in time are mostly expressed by the same postpositions. Some complex postpositions express comparison, cause, purpose, instrument, etc.

Kazakh employs auxiliary nouns such as *ald* 'front', *ara* 'interval', *art* 'back', *ast* 'bottom', 'lower part', *iš* 'interior', *orta* 'middle', *sïrt* 'exterior', *žan* 'side', *üst* 'top'. Some of them, e.g. *al-d*, *ar-t*, *as-t*, *üs-t*, are petrified relators that have lost their original functions. The auxiliary nouns can take on the adjectivizing suffix {-GI}, e.g. *art-ḳï* and *sïrt-ḳï*.

Notations

The following brackets and arrows are used for notations.
Graphic forms are given in angle brackets of the type ‹ ›
Glosses are given in angle brackets of the type ⟨ ⟩
Morphophonemic transcriptions are given in curly brackets of the type { }
← means 'is derived from'
< means 'has developed from'

References

Clauson, Sir Gerald 1972. *An etymological dictionary of pre-thirteenth-century Turkish.* Oxford: Clarendon.

Deny, Jean 1921. *Grammaire de la langue turque (dialecte Osmanli).* (Bibliothèque de l'École des langues orientales vivantes 5.) Paris: Éditions Ernest Leroux.

Doerfer, Gerhard 1977. Zu türkisch *bana* 'mir', *sana* 'dir'. *Central Asiatic Journal* 21, 208–214.

Geng, Shimin et al. 1999. *Ḳazirgi ḳazaḳ tili*. [Modern Kazakh]. Beijing: Minzu Publishing House.

Grönbech, Kaare 1936. *Der türkische Sprachbau.* Kopenhagen: Munksgaard.

Johanson, Lars 2012. Pyramids of spatial relators in Northeastern Turkic and its neighbors. In: Suihkonen, Pirkko & Comrie, Bernard & Solovyev, Valery (eds.) *Argument structure and grammatical relations. A crosslinguistic typology.* Amsterdam & Philadelphia: Benjamins.191–210.

Menges Karl H. 1959. Die aralo-kaspische Gruppe. In: Deny, Jean et al. (eds.) *Philologiae turcicae fundamenta* 1. Wiesbaden: Franz Steiner. 434–488.

Tekin, Talât 1968. *A grammar of Orkhon Turkic* (Indiana University Publications. Uralic and Altaic Series 69.) Bloomington: Indiana University; The Hague: Mouton & Co.

Zhang Dingjing 2004. *Xiandai hasakeyu shiyong yufa* [A practical grammar of Modern Kazakh]. Beijing: The Publishing House of the Central University of Nationalities.

Stress or intonational prominence? Word accent in Kazakh and Uyghur

Öner Özçelik

Özçelik, Öner 2015. Stress or intonational prominence? *Turkic Languages* 19, 163–192.

This paper investigates word stress/prominence in two Turkic languages, specifically Kazakh and Uyghur, and challenges the widespread assumption that the Foot is a universal constituent of the Prosodic Hierarchy (see e.g. Selkirk 1995, Vogel 2009). Instead, as was proposed for Turkish and French by Özçelik (2014, to appear), it is argued here that some languages, such as Uyghur, are footless, and thus, that the presence vs. absence of the Foot is parametric, thereby extending Özçelik's proposal to additional Turkic languages. We present three types of evidence for this proposal: (i) phonetic, (ii) formal phonological and (iii) syntactic (from syntax-prosody interface). Phonetic evidence, for example, demonstrates, based on the results of an experiment with native Uyghur speakers, that final prominence in (at least the predominant variety of) Uyghur is accompanied only by an optional "pitch" rise, and not by greater "intensity" or "duration". This is in stark contrast with the situation in Kazakh, in which, as with true iambic languages (i.e. languages with head-final feet), final prominence is also accompanied by greater duration, as evidenced by the results of an experiment with native Kazakh speakers. Languages that mark prominence only by a pitch rise (like Uyghur and Turkish) are classified, by several researchers, as pitch-accent languages, and not as stress-accent. The latter use duration and/or intensity, in addition to pitch (see e.g. Beckman 1986, Ladd 1996, Hualde et al. 2002). That is, although, on the surface, most Turkic languages seem to be the same in that they have finally prominent words, the source and nature of this prominence are different: it is an iambic foot for Kazakh, whereas it is footless intonational prominence for Uyghur and Turkish.

Öner Özçelik, Department of Central Eurasian Studies, Indiana University, 355 N. Jordan Ave., Bloomington, IN 47405, USA. E-mail: oozcelik@indiana.edu

1. Introduction

It is commonly assumed that prosodic constituents are organized into a hierarchy, where syllables are formed into feet, feet into prosodic words (PWds), PWds into phonological phrases (PPhs), and finally, phonological phrases into intonational phrases (Is). Furthermore, although there are various different adaptations of the Prosodic Hierarchy, it is assumed that all of these main constituents are universally available in every language (see e.g. Selkirk 1995, Vogel 2009), i.e. that every language has feet, PWds, etc., just as every language is assumed to have morphological words. In this paper, I argue, contra previous approaches, that the presence/absence

of the Foot is parametric; that is, while some languages, such as English, require every prosodic word (PWd) to have at least one foot, other languages, such as Turkish and Uyghur, are footless in the default case, or at least, their grammar does not assign foot structure (although the grammar can keep a foot that was assigned lexically). That Turkish grammar does not, in the usual case, parse syllables into feet has already been proposed in previous research (Özçelik 2014, and was later extended to French (Özçelik, to appear). Whether it is true for additional languages, or whether this holds true for, for example, all Turkic languages, remains to be investigated, however.

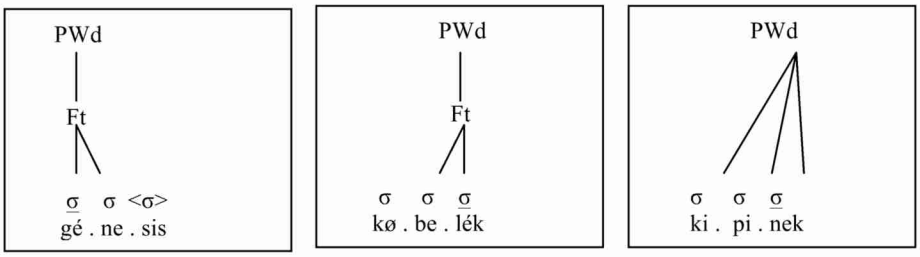
The first candidates that come to mind for a footless language are, of course, other Turkic languages, which share many linguistic characteristics with Turkish, and are generally assumed to have similar prosodic properties, e.g. in that prominence falls on the final syllable of words in most Turkic languages (see e.g. Johanson 1998, Schiering & van der Hulst 2010). Furthermore, as Johanson (1998) notes, there is a difference in many Turkic languages in the manifestation of regular final vs. (exceptional) non-final prominence; whereas final prominence is realized as pitch, non-final stress is dynamic stress-accent (which can also bear pitch), which can be taken to imply that final prominence may not involve stress to begin with. This phenomenon is, in fact, one type of evidence, among others, that Özçelik (2014, to appear) presents for the footless status of Turkish (final stress) based on experimental findings from Konrot (1981) and Levi (2005). The situation in other Turkic languages could be very similar to that in Turkish, as is predicted by Johanson's (1998) observations. However, although much research has been done on Turkish stress (see e.g. Johanson & Csató 1998, Hayes 1995, Inkelas 1999, Inkelas & Orgun 1995, 1998, Kaisse 1985, Kabak & Vogel 2001, Lees 1961, Lewis 1967, Özçelik 2014, to appear, Sezer 1983, Underhill 1976, van der Hulst & van de Weijer 1991), stress/prominence in other Turkic languages has rarely been studied. And the little research that is available does not present any experimental findings, or an in-depth discussion of cases that diverge from the norm, such as exceptional stress (see e.g. Alpysbaeva, Ismagulova & Turguzhanova 1995, Балақаев & Исқақов 1954, Баскаков et al. 1966, Batayeva 2013, Кайдаров 1997, Kirchner 1998, Krippes 1993, Muhamedowa 2016, Pratten & Omarova 1994, Somfai-Kara 2002 on Kazakh; Comrie 1997, Hahn 1991, Jiang et al. 2010, Liang 2009, Liang & Zhang 2008, 1998, Muti 2007, Yakup 2013 on Uyghur; Johanson 1998 on general Turkic). This paper aims to close this gap, and makes an initial attempt to shed light on the status of stress/prominence in two Turkic languages: Kazakh and Uyghur. In doing so, as much as it answers some previously unanswered questions, it also poses additional ones for future research to address.

The assumption that the Foot is a universal constituent of the Prosodic Hierarchy is held despite the fact that even when learning languages with foot structure, children's initial utterances do not contain any evidence of feet. In fact, children's initial outputs are not in the form of (the unmarked) binary feet; they are rather monosyllabic (Jakobson 1941/68), and critically monomoraic, utterances (see e.g. Fikkert

1994, Demuth 1995, Goad 1997). These findings do not receive a reasonable explanation under the assumption that the Foot comes as part of the Prosodic Hierarchy, and thus UG, given that children receive positive evidence containing foot structure from the beginning of the acquisition process. Thus, if the Foot is universal, it is not clear why children would not produce the unmarked form of prosodic words (PWds) from the very beginning, words composed of binary feet. If, on the other hand, the presence/absence of the Foot was parametric, and thus the Foot was available only in some languages, children could start the language acquisition process with footless utterances, and then add the Foot to their grammar based on positive evidence in learning a language that has foot structure, such as English. This would solve the problem posed by child language acquisition research, but would require the existence of footless languages, i.e. languages whose *grammar* does not assign feet, such as Turkish (Özçelik 2004, to appear).

In this paper, I argue that such languages do exist, and are not limited to Turkish. For example, having both regular and exceptional stress, Uyghur presents *formal* (as well as acoustic) evidence of lack of foot structure, both from the PWd-level and from facts at the syntax-prosody interface (PPh-level and beyond). In addition to formal evidence, I also present evidence in the form of acoustic correlates of prominence in reporting the findings of an experiment. Kazakh, on the other hand, I argue, differs from both Turkish and Uyghur in that it is a truly iambic language, one in which heads of feet are heavier in weight than dependents. Further, I also demonstrate, for Uyghur, that there are three different varieties as far as stress/prominence is concerned; although the footless Turkish-like variety is the predominant one, it is by no means the only variety, thereby partially accounting for the conflicting proposals raised for Uyghur stress in previous literature.

(1) a. English: /dʒɛnəsiz/ ‘genesis’ b. Kazakh: /købelek/ ‘көбелек’ c. Uyghur: /kipinek/ ‘كۆپەنەك’



On an account that views the presence/absence of the Foot as parametric, prosodic representations of the world’s languages can be summarized as follows: Whereas (1a) illustrates the internal structure of a PWd for a footed trochaic language like English, (1b) demonstrates the same thing for the iambic counterpart of footed languages, i.e. languages like Kazakh. Finally, (1c) presents the prosodic representation of words for Uyghur, in which words are footless in the default case. Thus, although

in both Kazakh and Uyghur (or Turkish), primary prominence falls on the final syllable of prosodic words, being iambic (and thus footed), Kazakh is more like English with respect to having feet than other Turkic languages.

(2) below presents the settings of the parameters of prosody that are assumed in order to reach these assumptions. These will be motivated in the rest of this paper.

(2)	*English*	*Kazakh*	*Uyghur/Turkish*
a. Footedness	Yes	Yes	No
b. Extrametricality	Yes	No	N/A
c. Directionality	R-L	R-L	N/A
d. Foot Binarity	Yes	Yes	N/A
e. Headedness	Left(Troc)	Right(Iamb)	N/A
f. Iterativity	Yes	Yes	N/A

2. "Stress" in Turkic languages: the case of Turkish, Uyghur and Kazakh:

2.1. Regular stress

It is commonly assumed that "stress" in Turkic languages falls on the final syllable of prosodic words (see e.g. Johanson 1998, Schiering & van der Hulst 2010).[1] (3) below illustrates this for Turkish, the Turkic language whose stress pattern has most commonly been investigated (see also Inkelas & Orgun 1998, Kabak & Vogel 2001, Özçelik 2014, Sezer 1983, van der Hulst & van de Weijer 1991).

(3)	a. *kedí*	b. *kedi-lér*	c. *kedi-ler-ím*	d. *kedi-ler-im-dé*
	cat	cat-PL	cat-PL-POSS1SG	cat-PL-POSS1SG-LOC
	'cat'	'cats'	'my cats'	'on my cats'

Uyghur works very similarly to Turkish, as is illustrated in (4) below. As with Turkish, each time a suffix is added, prominence moves to the rightmost syllable, demonstrating that, as with Turkish, prominence is word-final in Uyghur. This is the line held by most previous research on Uyghur prominence (e.g. Engesaeth et al. 2009, Jiang et al. 2010, Liang & Zhang 2008, Muti 2007, Yakup 2013).

(4)	a. *müʃúk*	b. *müʃük-lér*	c. *müʃük-lir-ím*	d. *müʃük-lir-im-dé*
	cat	cat-PL	cat-PL-POSS1SG	cat- PL-POSS1SG-lOC
	'cat'	'cats'	'my cats'	'on my cats'

1 There are Turkic languages that do not comply with this generalization, such as Chuvash, which is a Default-to-Opposite Edge stress language, in which leftmost light syllable is stressed in words that are composed only of light syllables, and when a heavy syllable is available, the rightmost heavy syllable bears stress (Krueger 1961, Dobrovolsky 1999, Gordon 2000).

One widely cited account that differs from mainstream descriptions of Uyghur, and posits Uyghur to be rather different from other Turkic languages is that of Hahn (1991) (see also Comrie 1997, among others), according to which stress is assigned to the last heavy syllable (CVC or CVV) in ultimate or penultimate position (e.g. LLH́, LHH́, LH́L), and if no such syllable exists, then primary stress is assigned to the final syllable of the word (e.g. LHLĹ, LLĹ, HLLĹ). To date, the differences between the two contrasting descriptions of Uyghur stress have not been addressed.

On the surface, Kazakh looks very similar to Turkish and Uyghur, as is illustrated in (5); the only apparent difference in Kazakh is the fact that the final open syllable is lengthened, as illustrated in (5d).

(5)	a. *almá*	b. *alma-lár*	c. *alma-lar-ím*	d. *alma-lar-im-dá:*
	apple	apple-PL	apple- PL-POSS1SG	apple- PL-POSS1SG-LOC
	'apple'	'apples'	'my apples'	'in my apples'

Closer investigation, however, reveals some interesting differences, all seemingly motivated by weight and complementing the picture of final lengthening observed in (5) (though see below). Examine the forms in (6).

(6)	a. *kép.ka*	b. *fe.gírt.ke*	c. *der.ba.zá:*	d. *sa.rim.sák*
	'hat'	'grasshopper'	'door'	'garlic'

These data indicate that, at least in some varieties of Kazakh, syllable weight is crucial in determining which syllable bears stress within a given word; that is, heavy syllables attract stress, whether they are heavy due to the length of the nucleus vowel (i.e. CVV) or by means of having a coda consonant (i.e. CVC). Although, at a first look, these forms seem to be exceptional in the same way that Turkish and Uyghur have exceptional (non-final) stress (see (7) and (8)), I will argue in Section 5.2 that these are in fact part of the norm in Kazakh, and are thus different from Turkish and Uyghur. This is because weight plays a crucial role in the assignment of stress in this language, as will be revealed by experimental findings. Before moving on to a more detailed explanation of these forms in Kazakh, we first present a description of exceptional stress in the three languages.

2.2. Exceptional stress

When stress is non-final in Kazakh, Turkish and Uyghur, it is said to be exceptional. In general, there are two types of exceptional stress in these languages. One involves roots, and the other, which has been more commonly studied, involves a small set of suffixes. These suffixes, when available, cause non-final stress. In general there are two types of non-final stress driving suffixes. One involves pre-stressing suffixes, and the other stressed suffixes.

2.2.1. Pre-stressing suffixes

(7), (8) and (9) below illustrate pre-stressing suffixes in Turkish, Uyghur and Kazakh respectively. (Pre-stressing suffixes are underlined throughout the paper.) Notice that these suffixes, when available, cause stress to fall on the immediately preceding syllable, irrespective of where in the word they are located.

(7) a. *dinle-dí*
listen-PAST
'(He) listened.'

b. *dinle-dí-mi*
listen-PAST-Q
'Did he listen?'

c. *dinlé-me-di*
listen-NEG-PAST
'He didn't listen.'

(8) a. *aŋli-dí*
listen-PAST
'(He) listened.'

b. *aŋli-dí-mu?*
listen-PAST-Q
'Did he listen?'

c. *aŋlí-mi-di*
listen-NEG-PAST
'He didn't listen.'

Note that these suffixes are (almost) always monosyllabic in Turkish and Uyghur, and monosyllabic exceptional suffixes never exceptionally bear stress themselves (i.e. are stressed despite other suffixes/syllables following); they are always pre-stressing. The two languages do indeed have some suffixes that exceptionally bear stress, but these are always bisyllabic (see section 2.2.2 below).

Kazakh is also stated to have exceptional pre-stressing suffixes, most notably pronominal agreement suffixes, examples of which are presented below in (9):

(9) a. *oqɨtɨwʃí-min*
teacher-COP1SG
'I am a teacher'

b. *oqɨtɨwʃí-ma?*
teacher-Q
'teacher?'

c. *oqɨtɨwʃɨ-lar-dán-bɨz*
teacher-PL-ABL-COP1PL
'We are some/among the teachers'

2.2.2. Stressed exceptional suffixes

Turkish and Uyghur also have a set of suffixes that are exceptionally stressed, and these are always bisyllabic. In addition, it is always the initial syllable of a bisyllabic exceptional suffix that is stressed:

(10) Turkish
gel-íyor-du
come-CONT-PAST
'(X) was coming'

(11) Uyghur
kel-íwat-idu
come-CONT-PRES
'(X) is coming.'

Kazakh also has some stressed exceptional suffixes, but unlike Turkish and Uyghur, these can also be monosyllabic, such as the present suffix, which gets stressed, whether or not it is followed by a person agreement suffix, as is exemplified in (12).

(12) Kazakh

a. *ket-é-siŋ*	b. *bar-á-dı*
leave-PRES-COP2SG	go-PRES-COP3SG
'You (will) leave.'	'He goes/will go.'

3. Current account

This paper proposes a unified account of regular and exceptional stress in the two Turkic languages of Uyghur and Kazakh. The analysis for Kazakh is the simpler one and is very straightforward from the perspective of phonological theory: It is a typical iambic language, which is weight-sensitive, as is true for iambic languages in general, which are arguably always weight-sensitive (see e.g. McCarthy & Prince 1986, 1993, Hayes 1991, 1995, but cf. Altshuler 2009). Final stress in Kazakh words would then be ensured through Right-to-Left footing, meaning that a single weight-sensitive iambic foot is built at the right edge of a word, leading to word-final stress, e.g. [LL(LH́)].

The analysis for Uyghur is very different; though with its regular word-final prominence, it looks very similar to Kazakh. In Uyghur, we have noticed three different patterns of stress, depending on the speaker, all essentially different from Kazakh. In the most predominant pattern, as opposed to Kazakh, regular word-final "stress" is best analyzed not as stress, but as intonational prominence falling on the final syllable of prosodic words; i.e. it does not involve foot structure. Exceptional stress, however, does involve foot structure in Uyghur, and is trochaic. The idea here is very similar to that of Özçelik (2014, to appear). As with Turkish, the grammar of the predominant variety of Uyghur (according to our data) does not assign foot structure at all, but if certain morphemes (i.e. exceptional stress driving suffixes) are already footed in their underlying representation/lexicon, then they will be footed in the surface representation too (i.e. the output of the grammar in Optimality Theoretic terminology). In this proposal, the whole grammar is trochaic; it is just that Trochaicity is vacuously satisfied for words that do not have any feet at all (regularly stressed words), since there are no feet in these words to begin with. In sum, Turkish/Uyghur and Kazakh are very different under this proposal, with the former two being trochaic languages and the final one, Kazakh, iambic.

The prosodic grammars of the two types of Turkic languages are presented in the form of an OT style constraint ranking in (13) and (14) below.

(13) Kazakh: iambic with weight sensitivity
IAMBIC, FT-BIN >> WSP[2] >> TROCHAIC

(14) Uyghur and Turkish: trochaic with no weight sensitivity
TROCHAIC, FT-BIN >> FINAL-PROMINENCE >> IAMBIC

That is, Kazakh is a typical iambic language on this proposal, as is also illustrated in the following OT tableau.

(15)

/alma:/		IAMBIC	FT-BIN	WSP	TROCHAIC
☞ *a*	[(almá:)]				*
b	[(álma)]	*!			
c	[(álma:)]	*!		*	*
d	[(almá)]			*!	
e	[al(má)]		*!	*	

Turkish and Uyghur, on the other hand, are both trochaic. The question that arises then is how a trochaic language receives prominence on the final syllable in the usual case (i.e. the so-called regularly stressed words). This is because regular "stress" in these two languages is not stress but is instead intonational prominence, and, as such, is not assigned through feet, but rather via an independent constraint FINAL-PROMINENCE, which places intonational prominence on the final syllable of PWds in the absence of foot structure. This is illustrated in (16) below for Uyghur (the predominant variety).

(16)

/müʃük-lér/		TROCH	FT-BIN	FINAL-PROM	PARSE-σ
a	[mü.(ʃúk.ler)]			*!	*
b	[mü.(ʃük.lér)]	*!			*
c	[mü.ʃük.(lér)]		*!		**
☞ *d*	H% | [mü.ʃük.lér]				***

2 Weight-to-Stress Principle (WSP) dictates (i) that stressed syllables are heavy, and (ii) that unstressed syllables are light, although not all linguists argue for both of the conditions, and there are variations in the literature (see e.g. Prokosch (1939), Chomsky & Halle (1968), Hayes (1981), Halle & Vergnaud (1987), Kager (1989), Prince (1990), Prince & Smolensky (1993), Burzio (1994), and Hammond (1999).

That is, although the language is trochaic, trochaicity does not become obvious unless one of the exceptional stress driving suffixes are available, as the language cannot assign foot structure to begin with.

When an exceptional suffix is available, however, things are a little different, because these suffixes already come footed in their UR, and given certain prosodic faithfulness constraints, they are footed in the SR, too. So we will need to update our constraint ranking presented above in (14) as follows.

(17) TROCHAIC, FT-BIN >> ANCHOR-RIGHT >> ANCHOR-LEFT, FINAL-PROMINENCE >> IAMBIC

The Anchor constraints here basically ensure that whatever is footed in the input/UR is footed in the output of the grammar/SR. To be more specific, ANCHOR-RIGHT ensures that a foot edge that is at the right edge of a syllable in the input should surface at the right edge in the output. ANCHOR-LEFT is its left counterpart; that is, material that is at the left edge in the input should surface at the left edge in the output. Of course, all this assumes inputs that are already footed in the lexicon, in the sense of Özçelik (2014), as illustrated below for Uyghur for monosyllabic and bisyllabic exceptionally stressed suffixes respectively.

(18) a. Inputs (URs) for pre-stressing suffixes:
$(mi)_{Ft}$
NEG

b Inputs (URs) for stressed suffixes:
$(iwat)_{Ft}$
PRES-CONT

The fact that ANCHOR-RIGHT and Ft-BIN both rank above ANCHOR-LEFT ensure further that monosyllabic exceptional suffixes, such as (18a), surface as pre-stressing, i.e. not stressed, nor post-stressing, as illustrated in (19) below.

(19) Monosyllabic exceptional suffixes: prestressing

/aŋli-(mi)-di/	TROCH	FT-BIN	ANCHOR-RIGHT	ANCHOR-LEFT
a [aŋ.li.(mí).di]		*!		
b [aŋ.(li.mí).di]	*!			*
c [aŋ.li.mi.dí]			*!	*
d [aŋ.li.(mí.di)]			*!	
e [aŋ.(lí.mi).di]				*

Here, since the suffix is monosyllabic, there is a violation of one of the Anchor constraints (i.e. ANCHOR-LEFT), so that a higher-ranking constraint, Ft-BIN, is satisfied. If both ANCHOR constraints were satisfied, as in (19a), it would lead to a violation of FT-BIN, as the foot would no longer be binary.

Indeed, for bisyllabic exceptional suffixes, we see a violation of neither constraint, as they are bisyllabic, and therefore keeping the underlying foot edges as they are will not incur a violation of FT-BIN, which is the reason why monosyllabic exceptional suffixes surface as pre-stressing (not ever stressed), whereas bisyllabic exceptional suffixes are stressed on their initial syllable. Examine (20).

(20) Bisyllabic exceptional suffixes: stressed on the first syllable

	/kel-(iwat)-idu/	TROCH	FT-BIN	ANCHOR-RIGHT	ANCHOR-LEFT
a	[kel(íwat)idu]				
b	[(kéli)watidu]			*!	*
c	[keliwatidú]			*!	*

4. Experimental and acoustic evidence

The hypothesis put forward above rests on the premise that whereas Kazakh is a truly iambic language, both Turkish and Uyghur are footless. If this proposal, for which we presented formal evidence above, is correct, one type of empirical evidence would be supplied by acoustics. While iambic languages tend to be uneven in that the head of a foot is heavier than the dependent, which usually surfaces in the form of greater duration for stressed syllables (e.g. Bolton 1894, Woodrow 1909/1951, Hayes 1991, 1995), neither duration nor intensity are correlates of prominence in languages without stress; prominence in these languages is cued only by an F0 rise at the most (see e.g. Beckman 1986, Ladd 1996, Hualde et al. 2002). Languages without stress are, in turn, good candidates for being footless (Özçelik 2014, to appear).

Findings of previous research already offer some insight into the issue, although exclusively from Turkish. It has been found that whereas the correlates of Turkish exceptional stress include both significantly greater intensity and F0 rise on the stressed syllable, final prominence is accompanied only by an optional slight rise in F0 (see e.g. Konrot 1981, 1987, Levi 2005, Pycha 2006; see also Kamali 2011, 2014, Güneş 2015, and Ipek 2015 for work that not only provides additional evidence for these findings, but also takes into account the effect of prosodic constituency on prominence as well as the effect of speech acts on intonational tunes; see also Özçelik (2014, to appear) for an interpretation of these facts (along with other evidence) as indicative of lack of foot structure for Turkish. Since the Turkish facts are better established, we will focus on a comparison between Kazakh and Uyghur in the remainder of this paper. The hypotheses are (i) that stressed syllables should be longer than unstressed syllables in Kazakh, potentially also accompanied by

greater intensity, and (ii) that no such pattern should be observed in Uyghur, i.e. that stressed syllables should be of equal duration and intensity as unstressed syllables.

In order to test these hypotheses, as well as to draw a general picture of stress/prominence in Kazakh and Uyghur, we conducted two experiments, one in Kazakh and the other in Uyghur, focusing on words with regular stress/prominence for the purpose of this study.[3]

4.1. Subjects

A semi-controlled production experiment was conducted with seven Kazakh and eight Uyghur native speakers. As was determined by a background questionnaire and a read-aloud task in Russian (for the Kazakh speakers), Chinese (for the Uyghur speakers), and English (for both, as the subjects were located in the USA), these speakers were near-monolingual; although Kazakh or Uyghur was their first language, they also had some non-native knowledge of English in addition to Russian or Chinese. The subjects ranged in age from 25 to 47. All of them lived in the USA, but regularly continued to speak their native language every day. Finally, all of them had at least college education, or, at the time of testing, were attending college or university.

4.2. Stimuli

The task involved responding to stimuli of various syllable structure profiles. There was a total of 70 words, all of which were common nouns. 20 of these were bisyllabic, comprised of all possible combinations of open vs. closed syllables (i.e. five under each condition, see below), and 40 were trisyllabic, again equally distributed among all possible combinations of open and closed syllables. The remaining 10 were four- and five-syllable words, composed of various different syllable structure profiles.

Tables 1 and 2 below present an example of each condition with bisyllabic and trisyllabic stimuli in Uyghur and Kazakh respectively. L represents a light/open syllable, whereas H represents a heavy syllable (although, strictly speaking, these were not necessarily heavy, but rather closed). These, along with the rest of the stimuli, were all regularly stressed.

3 The prediction is that exceptional stress should behave differently regarding acoustic correlates. In this paper, we focused on regular stress, but findings of research on exceptional stress in languages like Turkish (see above) confirm this prediction, as do some preliminary findings on Uyghur exceptional stress (see below).

Table 1. Uyghur stimuli

Bisyllabic words:			
LL	LH	HL	HH
paqa	supun	alma	jastuq
frog	soap	apple	pillow
پاقا	سۇپۇن	ئالما	ياستۇق

Trisyllabic words:							
LLL	LLH	LHL	LHH	HLL	HLH	HHL	HHH
tʃymyle	kipinek	tʃiketke	køzeynek	taʃpaqa	zendʒiwil	xizmettʃi	ytʃurgutʃ
ant	butterfly	grasshopher	glasses	turtle	ginger	servant	eraser
چۈمۈلە	كۆپىنەك	چېكەتكە	كۆزەينەك	تاشپاقا	زەنجۈبىل	خىزمەتچى	ئۆچۈرگۈچ

Table 2. Kazakh stimuli

Bisyllabic words:			
LL	LH	HL	HH
baqa	qalam	alma	dɛpter
бақа	қалам	алма	дәптер

Trisyllabic words:							
LLL	LLH	LHL	LHH	HLL	HLH	HHL	HHH
ʒygeri	købelek	jumirtqa	sarimsaq	tasbaqa	qolʃatir	ørmekʃi	erkekter
жүгері	көбелек	жұмыртқа	сарымсақ	тасбақа	қолшатыр	өрмекші	еркектер
corn	butterfly	egg	garlic	turtle	umbrella	spider	men

4.3. Task and procedure

Subjects were placed in an experimental condition where they saw pictures of each stimulus. They had to utter the name of each stimulus first in isolation and then once again in a carrier sentence. The carrier sentences used for Uyghur and Kazakh are presented in (21) and (22) below respectively.

(21) بۇ رەسىمدە بىر ____ بار.
'There is an__X__in this picture.'

(22) Бұл суретте _____ бар.
'There is __X__ in this picture.'

The stimuli were presented in three steps: First, the subjects had to guess the name of the pictured item, whose first letter was provided in order to make this task easier. At the second step, they had to utter the stimulus again. And at the third and final step, they finally produced the name of the item in a carrier sentence. Only the stimuli in the carrier sentence were transcribed and analyzed for acoustic measures.

The steps of the experimental procedure are exemplified below in (23) for Uyghur and Kazakh.

(23) Uyghur and Kazakh; experimental steps:
Step 1: guess the name of the pictured stimulus and utter it in isolation
Step 2: Utter the stimulus again in isolation
Step 3: Utter the stimulus in a carrier sentence ((21) for Uyghur; (22) for Kazakh)

The reason for not analyzing the words uttered in isolation was to prevent potential confounding variables such as utterance-final lengthening and phrase-level accent, both factors that tend to happen utterance-finally (Gordon 2014; Hyman 2014).

Each subject was tested individually and in a sound-attenuated booth. They were audio-recorded using Audacity (http://audacity.sourceforge.net) onto a Mac computer, with the help of an external Sony microphone, which was placed approximately 20 cm from the speakers' lips. The tasks were administered in the following order: a background questionnaire, production experiment, and the read-aloud tasks (in English and Chinese and/or Russian). The entire procedure took about one hour per subject.

4.4. Analysis

The responses were transcribed and annotated for placement of stress/prominence using Praat acoustic analysis software (Boersma & Weenink 2013). For each experimental word, vowel and syllable duration (in ms), average and peak intensity (in dB), average fundamental frequency (F0, in Hz), and time of F0 peak were measured. In addition, the first and second formant frequencies (F1 and F2, in Hz) were noted with the aim of mapping the subjects' vowel space, because certain vowels (e.g. high vowels) inherently have lower duration than others (e.g. low vowels) (Fonagy 1966, Beckman 1988). Finally, sentences in which there was a pause between the syllables of the experimental word were excluded from the analysis.

Although all the words were analyzed for stress placement in order to get a general picture of stress/prominence in both languages, the acoustic analysis here focused on bisyllabic stimuli composed of two open syllables, i.e. the LL condition. There were several reasons for focusing on the LL condition in the phonetic analysis: First, acoustic cues in longer words tend to be affected more easily by extralinguistic factors such as word-internal pauses; such effects are best controlled in bisyllabic words. Second, since these words do not involve any syllables with codas, vowel duration alone is a good predictor of syllable length. In conditions that have syllables with codas, vowel shortening before codas (a phenomenon observed in many natural languages) might, otherwise, influence duration measurements since vowels in open syllables are normally longer than those in closed syllables. At the same time, closed syllables contribute to syllable length by virtue of the presence of the additional segment in the coda position, and some consonants, like continuants,

can sometimes arbitrarily be produced longer. Words composed only of open syllables are therefore ideal to keep the effects of context on vowel/syllable length under control. Third, since both syllables in the LL condition are underlyingly short, this condition allows us to best distinguish between a trochee and an iamb as well as a footless representation, and, crucially, permits us to spot cases where syllable weight is derived; certain syllables (such as final syllables) may, after all, be lengthened for reasons of e.g. iambic lengthening, which may not as readily be observed if those syllables already bear weight by means of having a coda. In sum, words under this condition will best serve to distinguish the acoustic cues of trochaic, iambic, and footless grammars. In iambic grammars, for example, heads of feet (the final syllables of LL forms) are typically longer than non-heads (see e.g. Hayes 1995), whereas no such effect is expected in grammars that are footless. Similarly, in trochaic grammars, heads (initial syllables in LL forms) are expected to bear higher intensity than non-heads, which, again, is not expected in footless or iambic systems.

As suggested by Peterson & Lehiste (1960), segmentation criteria were based on an analysis of both spectrograms and waveforms cues. In particular, the following criteria were used in determining syllable and vowel boundaries: (i) The first upward-going zero crossing at the beginning of the waveform was used to isolate the onset of the first syllable; (ii) for the offset of the first and the onset of the second syllable, in words without a stop consonant as the onset of the second syllable, the boundary between the offset of the first and the onset of the second syllable was determined as the transition between the spectrographic pattern of the last segment of the first syllable and the first segment of the second; (iii) in words with a stop consonant as the onset of the second syllable, this boundary was defined as the starting point of the silence of stop closure; (iv) second syllable/word offset was determined as the end point of the waveform at the final downward-going zero crossing. These criteria were used, in making decisions on vowel/syllable length, in both the initial transcriptions and TextGrid annotations, though syllable and vowel boundaries were marked only in TextGrids, for statistical purposes.

5. Results and discussion

We present the results in two different ways. First, overall results are provided, with illustrations of spectrograms, followed by more specific acoustic results on the LL condition with statistical analyses.

5.1. Overall results

Before we present more specific results on the bisyllabic stimuli with two open syllables (i.e. the LL condition), we take a quick look at the visual differences between Uyghur and Kazakh, as revealed through a spectrogram, and then do the same for Uyghur regular vs. exceptional stress/prominence. The evidence offered in this section is weaker than that in Section 5.2., where statistical information is presented based on data from numerous responses generated from multiple speakers. Section

5.1. only serves to present a preliminary look at some of the patterns that hold and are crucial in understanding the arguments laid out in this paper.

5.1.1. A comparison of Kazakh vs. Uyghur: a preliminary examination

Figures 1 and 2 below illustrate duration, intensity and pitch values/tracks for the word /alma/ 'apple' for Uyghur (the most predominant variety) and Kazakh, which happens to be the same segmentally in both languages. As seen though, the two words appear quite differently at the suprasegmental level, mainly because the final open syllable is lengthened in Kazakh, unlike in Uyghur.

Figure 1: Uyghur /alma/ < ئالما > "apple"

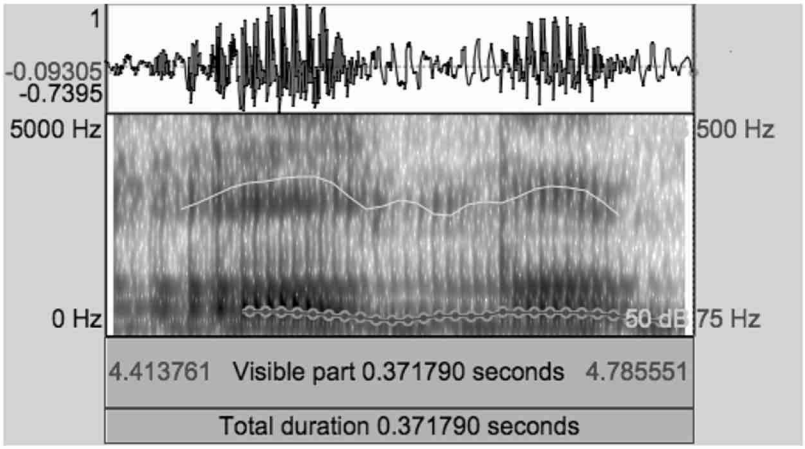
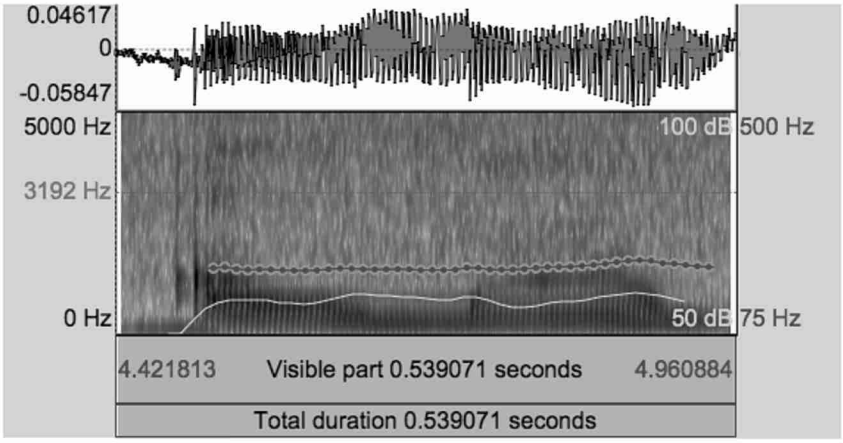


Figure 2: Kazakh /alma/ < алма > "apple"



What these spectrograms show is that in Kazakh the duration of the second vowel (and thus the second syllable) is much greater than that of the first vowel (and the first syllable). This is in stark contrast to the situation in Uyghur, where the first and the final syllables are of about the same duration. That is, duration is an important correlate of stress in Kazakh, whereas it is not in (the predominant variety of) Uyghur. Regarding intensity and pitch, on the other hand, we see that neither language uses these as correlates of prominence. Intensity (as indicated by the solid curve on the spectrogram) is about the same on both syllables for both Uyghur and Kazakh, although the first syllable seems to be slightly higher in intensity in Uyghur, and the second syllable in Kazakh. As for pitch (which is indicated by the dotted curve), it seems more or less flat for both languages, although there seems to be a slight pitch rise on the final vowel for Kazakh. All in all, given only these two words (quantitative evidence to be given in Section 5.2), which are segmentally identical in the two languages, we could reach the preliminary conclusion that none of intensity, duration and pitch are correlates of prominence in Uyghur, whereas duration is an important correlate for Kazakh (to be investigated in more detail later). If the observation made based on these two words holds true more generally (see Section 5.2.), and given the criteria presented in previous formal phonological literature (see e.g. Beckman 1986, Ladd 1996, Hualde et al. 2002, Özçelik 2014, to appear), it follows that Uyghur regular stress is not really (foot-based) stress in accordance with the observation of Johanson (1998) for Turkic languages, but is, instead, intonational prominence, much like in Turkish (Özçelik 2014, to appear). Kazakh, on the other hand, seems to be a true (foot-based) iambic language, in that duration is an important cue in marking stressed syllables.

5.1.2. A preliminary comparison of regular vs. exceptional stress/prominence in Uyghur

If regular (final) prominence in Uyghur is not foot-based stress, the question that follows is whether exceptional stress in Uyghur is really (trochaic) stress, as we argued above in Section 3. Although this paper does not present an in-depth statistical investigation of this issue, a comparison of the two spectrograms below in Figures 3 and 4 is quite informative. Whereas Figure 3 depicts the representation of a verb that only has a regular suffix attached, Figure 4 has the same word with an additional exceptional suffix. In both words, the same syllable, i.e. *-di*, is the most prominent one, although the correlates of prominence are rather different.

Figure 3: Uyghur: regular/word final prominence: /bol-dí/ "(it) happened"

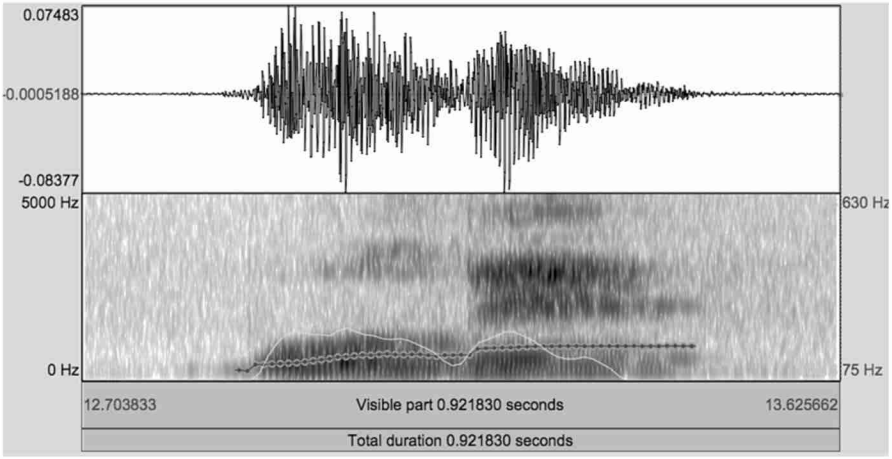
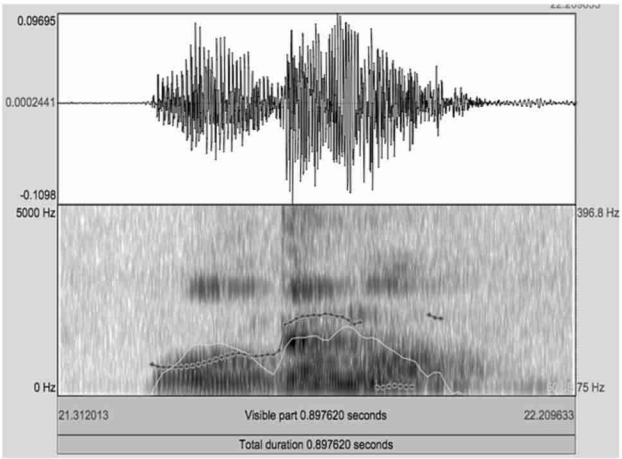


Figure 4: Uyghur: exceptional stress: /bol-dí-mu/ "Did (it) happen?"



Notice, in the regularly stressed word /boldí/ (Figure 3), that the first and second syllable have about the same level of intensity and duration, although pitch values rise slightly for the second syllable. For the exceptionally stressed word /boldí-mu/, on the other hand, the correlates of stress are manifested in strikingly different ways: The stressed second syllable has a very sharp rise in pitch, as well as much greater intensity than the first syllable. Intensity and especially pitch values then drop dramatically for the third syllable.

From these findings, it can be concluded, at least tentatively, in the absence of statistical data and data from other words/suffixes, that whereas regular final prominence in Uyghur is accompanied only by a slight pitch rise (which is optional, as will be shown later), exceptional stress is accompanied by both a sharp F0 rise, as well as greater intensity (similar to Konrot's (1981, 1987) and Levi's (1995) findings for Turkish), suggesting that, unlike regular prominence, exceptional prominence is trochaic, as is argued formally above in Section 3.[4] Once again, it should be emphasized, however, that this comparison should be taken with a grain of salt in the absence of quantitative data. More research is clearly needed, which can build on the insights provided here.

5.1.3. Three different varieties of Uyghur

The discussion above regarding word-level prominence in Uyghur is rather simplified, as our starting point was the research question of whether there are additional footless languages like Turkish, the most likely candidates for which, we hypothesized, would be other Turkic languages. The answer to this question was in the affirmative, with evidence coming from the most predominant variety of Uyghur found in our sample, but not from Kazakh.

This was, by no means, however, the only variety of Uyghur revealed in our data. Although the utterances of 5 of the 8 Uyghur subjects tested were consistent with this pattern, the remaining three behaved rather differently, illustrating two additional patterns: First, 2 subjects (interestingly comprising the only two female subjects in the group) consistently lengthened final open syllables, leading to a pattern that is in essence similar to that in Kazakh, but that also differs in that Kazakh does not *consistently* lengthen final open syllables with final stress each time; when a final open syllable is at times not lengthened, stress is nonfinal in Kazakh (see e.g. the data in (5)). For these two Uyghur subjects, however, stress was more consistently final, with final open (but not closed) syllables lengthened across the board.

For the remaining one subject, a pattern such as that described by Hahn (1991) arose. That is, this subject placed stress either on the final or the penultimate syllable depending on weight, but not on a preceding syllable even if that syllable was the only heavy one. In the grammar of this speaker, if both the penultimate and the final syllable were heavy, main stress fell on the ultimate syllable; if the penultimate syllable was heavy and the ultimate light, then, stress fell on the penultimate syllable. If neither were heavy, stress still fell on the final syllable, even if a preceding syllable was heavy, meaning that a two-syllable window existed for stress at the right edge, as in the pattern reported by Hahn (1991).

It follows from these findings that syllable weight in Uyghur was important for three of the eight subjects, whereas it played no role whatsoever for the remaining

4 Note, however, that, as one reviewer correctly points out, the sharp F0 rise here may have been further bolstered by the fact that a question particle is used, and interrogatives in Turkish may have an extra pitch accent not found in declaratives (Kamali 2014).

five. These findings help account for the conflicting views on Uyghur stress reported in the literature, with most research suggesting that it is consistently word-final (Jiang et al. 2010, Liang & Zhang 2008, 1998, Muti 2007, Yakup 2013), but some suggesting that syllable weight plays a role within a window of two syllables at the right edge on the other (e.g. Comrie 1997 and Hahn 1991). Further, it is possible that these differences point to language change in progress, either from a weight-based footed language to a footless one or from a footless language to a footed one in which syllable weight is important. It is interesting in this regard that the two subjects whose utterances consistently exemplified word-final open syllable lengthening were female, which may point either to language change led by female speakers or a register used only by females, although the small sample size employed here is not sufficient to lead us to a definitive conclusion. Either way, additional research in the issue is needed, which may lead to important conclusions about our knowledge of footless languages, and whether footed languages lead to footless languages or vice versa, historically speaking. It may also have certain implications about language change. I leave further investigation of these issues to future research.

Regardless, the three varieties of Uyghur can be summarized as follows vis-à-vis stress/prominence, with example prosodic representations.

(24) a. Turkish-like (predominant) pattern, as in Özçelik (2014, to appear):
intonational (footless) prominence on the final syllable:
Ex: H%
|
[σ σ σ]

b. Hahn (1991) pattern:
stress on the final or penultimate heavy syllable, depending on weight:
Ex: [L(LH́)], [(LH́)L], [H(LĹ)], [L(LĹ)]

c. Final stress with final open syllable lengthening (Kazakh-like):
Ex: [L(LH́)], [LH(H́)], [HL(H́)], [L(LH́)]

Now that we have outlined the general stress/prominence patterns of both Uyghur and Kazakh, the following section will focus on comparing the two languages by means of a quantitative analysis. In doing so, and in accordance with the research question posed at the beginning of this paper, the relevant sections will concentrate on the predominant (footless) variety of Uyghur found in our data (exemplified in (24a) and how it compares with Kazakh, which we have argued to be iambic.

5.2. Acoustic correlates of prominence: footed vs. footless outputs

One of the most important predictions made in this paper was that whereas (the predominant variety of) Uyghur grammar does not assign foot structure (for regularly stressed words), Kazakh does.

One main assumption behind this proposal is that for prominence to involve foot structure, it must be accompanied by intensity and/or duration at the word level. That is, neither pitch rise alone nor a flat structure (i.e. one that has no intensity, duration or pitch) involves foot structure, as in Turkish, where regular word-final "stress" is either flat, or involves a slight pitch rise on the word-final syllable (see Özçelik 2014, in press). The best way to determine whether subjects have foot structure or not is, then, to investigate the phonetic correlates of prominence to determine whether they include intensity and/or duration, in addition to pitch rise.

Kazakh speakers' stressed syllables should, then, be accompanied by greater duration and/or higher intensity, whereas this should not be the case for the Uyghur-speaking subjects. Further, if they act according to cross-linguistic tendencies argued for by Bolton (1894) and Hayes (1985), Kazakh-speaking subjects' iambs should be accompanied mostly by greater duration, more so than greater intensity (since they are iambs, rather than trochees).

In order to test these predictions, a subset of the data, the five stimuli in the LL condition, were further analyzed using TextGrids. The results are reported below for each of duration, intensity and pitch respectively. We start with duration, which is the most important correlate to be discussed here, for it reveals significant differences between the Uyghur-speaking subjects and the Kazakh-speaking subjects.

5.2.1. Duration

Table 3 below presents the durations (in seconds, s.) of the first and second vowels in the LL condition (which included five tokens) for the Uyghur- and Kazakh-speaking subjects.

Table 3: Duration measurements in the LL condition (in s.)

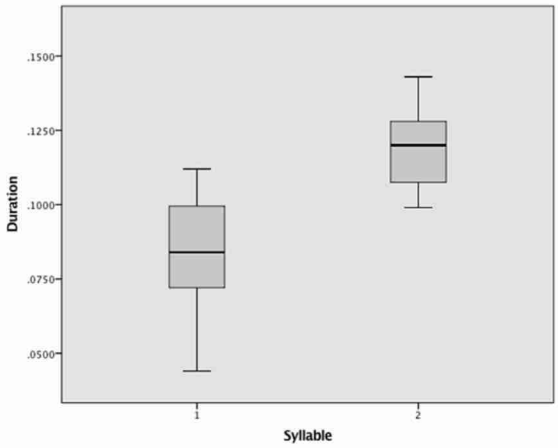
	First Vowel Duration (s.)	Second Vowel duration (s.)	Difference (s.)	p-value
Kazakh:				
(n=7)[5]	0.084614	0.118771	0.034157	$p < 0.0001$
Uyghur (predominant variety):				
(n=5)	0.119684	0.121940	0.002256	$p = 0.959$

As Table 3 demonstrates, although the vowel of the second syllable was longer than that of the first syllable for both Kazakh- and Uyghur-speaking subjects, the differ-

5 One Kazakh speaker was excluded from analysis as an outlier, taking the number of Kazakh speakers tested down from 8 to 7.

ence between the first and the second syllable was much greater for the Kazakh-speaking subjects. The first vowel was, on average, 0.085 seconds (s) for them, while the second vowel was about 0.119 s. In other words, the difference in duration between the first and the second vowel for this group was about 0.034 s. This difference is statistically significant, ($F = (1, 68) = 91.968$, $p < 0.0001$), as confirmed by the results of a one-way ANOVA. This is illustrated by the box plots in Figure 5 below. The horizontal line represents the first and second (final) vowels, indicated as 1 and 2 respectively, while the vertical line represents the duration of the corresponding vowel in s.

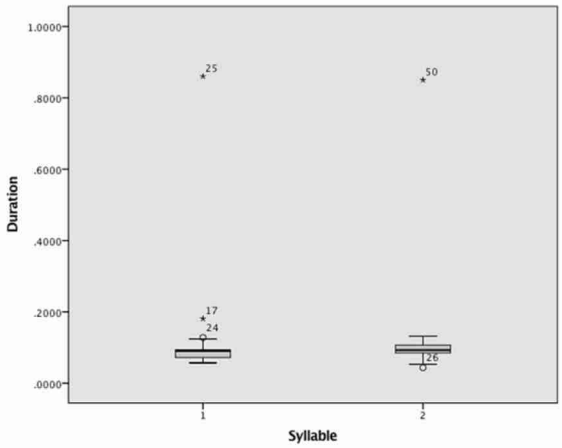
Figure 5: Durational differences between the first and second vowels for Kazakh-speaking subjects



In short, the durational differences between the first and second vowels were not a chance factor for the Kazakh-speaking subjects, who had iambic systems; these subjects significantly lengthened the final vowels of words.

For Uyghur-speaking subjects, however, a different pattern emerged. In contrast to the Kazakh-speaking subjects, differences in the mean durations of first and second syllable vowels were much smaller for the Uyghur speakers; the first vowel was about 0.120 s. while the second vowel was about 0.122 s. The difference was, in other words, only about 0.002 s., and this difference was not statistically significant, $F = (1, 48) = 0.003$, $p = 0.959$. This is also illustrated by the box plots in Figure 6 below.

Figure 6: Durational differences between the first and second vowels for Uyghur-speaking subjects



To summarize thus far, the average durations of the first and second vowels in target LL words differed significantly for the Kazakh-speaking subjects, who had an iambic grammar, whereas the means more or less coincided for the Uyghur-speaking subjects, who had footless representations (at least the five that were tested here). This presents strong evidence for the proposal made earlier in this paper; Kazakh speakers produce final prominence through foot structure, whereas Uyghur speakers produce footless utterances.

A well-known tendency in natural languages is that duration is a better cue than intensity for iambic systems, whereas the converse is true for trochaic systems (Bolton 1894, Hayes 1995). In sum, the Kazakh-speaking subjects who stressed the final syllables of LL forms did indeed have iambs; in their productions stressed syllables were accompanied by significantly greater duration than unstressed syllables. This was not the case with the Uyghur-speaking subjects, however; their final syllables were only slightly longer than nonfinal syllables. As opposed to the iambic Kazakh group, the durational differences between the two syllables for the Uyghur groups were statistically insignificant.

5.2.2. Intensity

Intensity was not found to be an important correlate in either group. Table 4 presents a summary of the findings for all Kazakh- and Uyghur-speaking subjects. The results are given in terms of maximum intensity (in dB).

Table 4: Intensity measures

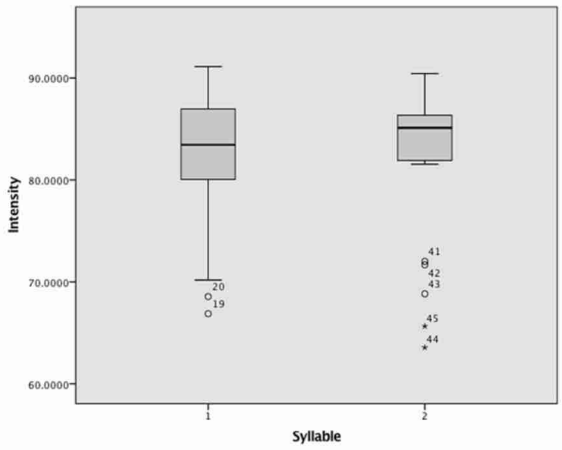
	First σ (peak) Intensity (dB)	Second σ (peak) Intensity (dB)
Kazakh:		
(n=7)	69.003	68.857
Uyghur:		
(n=5)	81.873	82.068

As predicted, intensity was not an important correlate of final prominence for the Uyghur-speaking subjects, as was the case with duration (see section 5.2.1). As indicated in the last row of Table 4, the intensity of the second vowel was slightly higher, on average, than that of the first vowel, however this was not statistically significant, as the results of a one-way ANOVA indicate, $F = (1, 48) = 0.009$, $p = 0.925$.

The difference in intensity between the first and second syllables is illustrated, for these subjects, by the box plots in Figure 7.

It is not surprising for intensity to play no role for the Uyghur-speaking subjects; as was argued above, neither intensity nor duration is expected to play an important role for word-level prominence in footless languages, and the predominant variety of Uyghur (represented by 5 out of the 8 Uyghur subjects tested here), I argue, is footless.

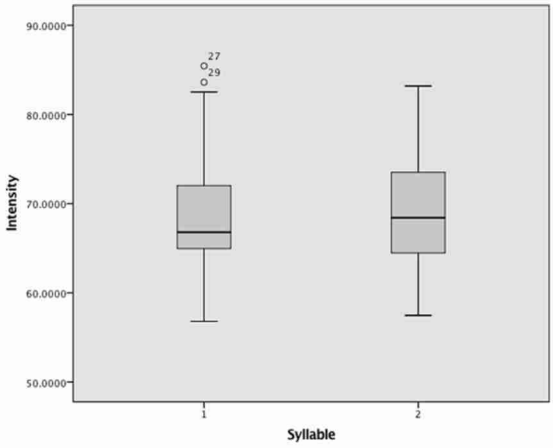
Figure 7: Maximum intensity for Uyghur-speaking subjects



Intensity was not an important correlate for the Kazakh-speaking subjects, either. As first noted by Bolton (1894) and later argued by Hayes (1991, 1995) among others, iambic languages have elements contrasting in duration, not intensity, whereas tro-

chaic languages tend to have elements contrasting in intensity but not duration. The results of the current research parallel this finding. While the Kazakh-speaking subjects with iambic grammars consistently lengthened the final syllables of Kazakh PWds (see section 5.2.1), they did not rely on intensity as a correlate of prominence. In fact, as is illustrated in Table 4 above, initial syllables were slightly higher in intensity for these subjects than final syllables, although, as with the Uyghur-speaking subjects, this was not statistically significant, $F = (1, 68) = 0.008$, $p = 0.930$. This is illustrated by the box plots below in Figure 8.

Figure 8: Maximum intensity for Kazakh-speaking subjects.



5.2.3. Pitch

For both Kazakh- and Uyghur-speaking subjects, pitch measurements were also taken for the LL condition. Table 5 below presents average maximum pitch values associated with the first and second syllables for both Kazakh- and Uyghur-speaking subjects.

Table 5: Pitch measurements

	First V Max Pitch (Hz)	Second V Max Pitch (Hz)
Kazakh:		
(n=7)	223.551	207.377
Uyghur:		
(n=5)	143.815	154.496

For the Kazakh-speaking group, maximum pitch was higher overall on the first syllable than the second, whereas for the Uyghur-speaking group, maximum pitch was higher on the second syllable. A one-way ANOVA was conducted in order to test for statistical significance. For the Kazakh group, the difference in pitch between the two syllables was far from significant: $F = (1, 68) = 0.832$, $p = 0.365$. Likewise, for the Uyghur group, there was no significant effect associated with (maximum) pitch; $F = (1, 48) = 0.701$, $p = 0.407$ for maximum pitch.

All things considered, unlike duration, and similar to intensity, pitch was not a significant correlate for stress/prominence for either group. It should be noted, however, that to a large extent this is true because, as with most similar research, overall results are examined here (i.e. individual results are all collapsed together under one language, Kazakh or Uyghur). Looking at individual subjects reveals alternate patterns; some subjects had higher pitch on the first syllable than the second, while others had higher pitch on the second syllable than the first, even within the same language. For example, as is illustrated in Table 5, although the Kazakh speakers as a group had higher pitch values overall on the first syllable than the second, three of them consistently placed higher pitch on the second syllable than the first, and this was true for these three subjects across all LL stimuli. Clearly, there is much individual variation when it comes to using pitch as a correlate of prominence in either language, and especially in Kazakh, and the issue thus requires further investigation. The presence of so much variation is of course in itself evidence that pitch values alone are not significant in determining the place of prominence in either language, although this does not exclude the possibility that pitch "shape", if not pitch values, may play a role in either language.

5.2.4. Summary and discussion of phonetic correlates

The results of acoustic measurements reported in Section 5.2 above indicate that whereas duration was a significant correlate of stress in the productions of the Kazakh-speaking subjects, none of duration, intensity or pitch was a good correlate for the Uyghur-speaking subjects (those who are speakers of the predominant variety—see Section 5.1 for other variants where weight and duration play a role). The acoustic findings with said variety of Uyghur are very similar to what was previously found to be the case with Turkish (see e.g. Konrot 1981, 1987, Levi 2005), and are in line with the observation made by Johanson (1998) regarding the status of dynamic stress vs. pitch in Turkic languages, although Kazakh diverges significantly from this pattern. One reason for Kazakh being so different could be the influence of Russian, which is a stress language (see below for more on this).

The finding that neither intensity nor duration was an important correlate of prominence in the grammars of the Uyghur-speaking subjects was predicted, since the variety of Uyghur they speak lacks foot structure, similar to Turkish, as was proposed above, and these two cues, duration and intensity, have been argued to be associated with the presence of foot structure (Özçelik 2013, 2014, to appear). Simi-

larly, the fact that Kazakh-speaking subjects relied to a great extent on duration was not surprising, since, as was argued above, Kazakh is an iambic language, and iambic languages generally rely on duration in marking stressed syllables. Seemingly more problematic is the fact that the Uyghur-speaking subjects did not rely on pitch either, raising the question of what was the cue for final prominence in their language. Although I do not have a definitive answer to this question, it may be that the specific shape of the pitch track is of some importance, or that a flat structure (regarding pitch) leads to perception as final prominence, as is argued for Turkish by Özçelik (2014), and alluded to for French by Ladd (1996) (see also Özçelik to appear on French). Either way, as with Turkish (Özçelik 2013, 2014), pitch seems relevant for some Uyghur speakers, those who consistently have higher pitch on prominent syllables, and irrelevant for others who have a flat structure regarding pitch, the combined result of which led to the lack of statistical significance vis-à-vis pitch for Uyghur speakers.

6. Conclusion

In this paper, we have examined word-level prominence in two Turkic languages, Kazakh and Uyghur. Using both formal and experimental evidence, we have argued that although Kazakh and Uyghur look alike on the surface, in that the word-final syllable is often the most prominent syllable within a word, the nature of this prominence differs. Although Kazakh is a typical iambic language that adheres to weight sensitivity, Uyghur, in its most predominant variety, is footless, unless certain morphemes come pre-footed in the lexicon, as in Turkish (Özçelik 2014, to appear). However, this was only the case for some of the subjects in our experiments, albeit the majority. Some Uyghur-speaking subjects demonstrated a pattern quite like that of Kazakh, whereas the behavior of one subject was consistent with the system described by Hahn (1991), who proposes a stress pattern for Uyghur that differs from most analyses of the language. In this regard, the current paper also offers some insight into the reasons behind the differing views on accent/stress/prominence in the Uyghur language.

Kazakh's apparent dependence on duration may be explained either by its historical development and subsequent divergence from other Turkic languages like Turkish and Uyghur, or by the greater influence of Russian, a stress language, unlike Chinese, which is tonal. After all, all the speakers in our experiments had at least some knowledge of Russian (in the case of the Kazakh) or Chinese (in the case of the Uyghur), although their native language (Kazakh or Uyghur) was clearly predominant. In other words, Russian may have influenced Kazakh speakers' productions by giving their Kazakh some characteristics of a stress language, such as stress syllables bearing greater duration, although the distribution of stress is still similar to that in other Turkic languages like Turkish and Uyghur, in that it is still mostly final. That is, it remains to be confirmed that these findings do indeed hold true for Kazakhs not influenced by Russian.

To sum up, relatively little research or scholarly analysis has been conducted on stress or prominence in Turkic languages, although Turkish, the most commonly taught Turkic language, has received a great deal of attention recently regarding its accentual pattern. It is commonly assumed that most Turkic languages behave like Turkish in demonstrating final stress. This paper has shown that the situation is much more complex in other Turkic languages, including those that are relatively better researched, such as Kazakh and Uyghur, and has made an initial attempt at a descriptively and theoretically accurate analysis of the accentual patterns of these languages. While the paper provided answers to many of the questions posed at the beginning of this paper, it has also raised additional questions. Clearly, more research needs to be done on the stress/prominence patterns of Turkic languages, including Kazakh and Uyghur.

Acknowledgments

I would like to thank Kurban Niyaz and Gulnisa Nazarova for help with the Uyghur data and recruitment of Uyghur-speaking subjects, as well as Saulet Alpysbayeva and Jonathan N. Washington for help with the Kazakh data and discussion of Kazakh results. I am also thankful to Can Ayter for help with recruiting the subjects and conducting the experiments in my lab at Indiana University, as well as all the subjects who have taken their time to take part in the experiments. Many thanks also to the audience of ICTL 2014 at the University of Rouen and WAFL 10 at MIT for their helpful comments on an earlier version of the proposal presented in this paper. Finally, I am also very grateful to the anonymous reviewers of *Turkic Languages*, as well as the editor, for helpful comments and suggestions.

References

Alpysbaeva, Bayan & Ismagulova, Bayan & Turguzhanova, Uphilmalik 1995. *Kazakh language course for peace corps volunteers in Kazakhstan.* Almaty: Peace Corps.

Altshuler, Daniel 2009. Quantity insensitive iambs in Osage. *International Journal of American Linguistics* 75, 365–398.

Балақаев, М., А. Исқақов 1954. *Қазіргі қазақ тілі (лексика, фонетика, грамматика).* Алматы: Қазақ ССР Ғылым академиясының баспасы.

Баскаков, Н. А. & Хасенова, А. К. & Исенгалиева, Б. А. & Кордабаев, Т. Р. 1966. *Сопоставительная грамматика русского и казахского языков. Морфология*. Алма-Ата: Издательство «Наука» Казахской ССР.

Batayeva, Zaure 2013. *Colloquial Kazakh*: *The complete course for beginners*. London, New York: Routledge

Beckman, Mary E. 1986. *Stress and non-stress accent*. Dordrecht: Foris.

Boersma, Paul & Weenink, David 2013. *Praat*: Doing phonetics by computer. Accessible: http://www.fon.hum.uva.nl/praat.

Bolton, Thaddeus 1894. Rhythm. *American Journal of Psychology* 6, 145–238.

Burzio, Luigi 1994. *Principles of English stress*. Cambridge, UK: Cambridge University Press.

Chomsky, Noam & Halle, Morris 1968. *The sound pattern of English*. New York: Harper and Row.

Comrie, Bernard 1997. Uyghur Phonology. In: Kaye, Alan (ed.) *Phonologies of Asia and Africa* 2. Winona Lake: Eisenbrauns. 913–925.

Csató, Éva Á. & Johanson, Lars 1998. Turkish. In: Johanson, Lars & Csató, Éva Á. (eds.) *The Turkic languages.* London & New York: Routledge. 203–235.

Demuth, Katherine 1995. Markedness and development of prosodic structure. In: Beckman, Jill (ed.) *Proceedings of the North Eastern Linguistic Society (NELS 25)*. Amherst, MA: GLSA, University of Massacchusetts. 13–25.

Dobrovolsky, Michael 1999. The phonetics of Chuvash stress: Implications for phonology. *Proceedings of the 14th International Conference of Phonetic Sciences*. 539–42.

Engesaeth, Tarjei & Yakup, Mahire & Dwyer, Arienne M. 2009. *Greetings from the Teklimakan: A handbook of Modern Uyghur*. University of Kansas Scholarworks.

Fikkert, Paula 1994. *On the acquisition of prosodic structure.* PhD dissertation, Leiden University/HIL. Published 1994, The Hague: Holland Academic Graphics.

Fónagy, Iván 1966. Electrophysiological and acoustic correlates of stress and stress perception. *Journal of Speech and Hearing Research* 9, 231–244.

Goad, Heather 1997. Codas, word minimality, and empty-headed syllables. *Child Language Research Forum* 28. Stanford: Center for the Study of Language and Information. 113–122.

Gordon, Matthew 2000. Re-examining default-to-opposite stress. In: *Proceedings of the 26th Berkeley Linguistics Society (BLS 26)*. 101–112.

Gordon, Matthew 2014. Disentangling stress and pitch-accent: a typology of prominence at different prosodic levels. In: van der Hulst, Harry. (ed.) *Word accent: Theoretical and typological issues*. Cambridge: Cambridge University Press. 83–118.

Güneş, Güliz 2015. Deriving prosodic structures. PhD dissertation, University of Groningen.

Hahn, Reinhard F. 1991. *Spoken Uyghur*. Seattle: University of Washington Press.

Halle, Morris & Vergnaud, Jean-Roger 1987. *An essay on stress*. Cambridge, MA: MIT Press.

Hammond, Michael 1999. *The phonology of English: A prosodic Optimality Theoretic approach*. Oxford: Oxford University Press.

Hayes, Bruce 1981. *A metrical theory of stress rules*. PhD dissertation, MIT. [distributed by Indiana University Linguistics Club].

Hayes, Bruce 1985. Iambic and trochaic rhythm in stress rules. In: Niepokuj, M. et al. (eds.) *Proceedings of the Eleventh Annual Meeting of the Berkeley Linguistics Society (BLS 11)*. 429–446.

Hayes, Bruce (1987). A revised parametric metrical theory. In McDonough, J. & Plunket, B. (eds.) *Proceedings of the 17th North East Linguistic Society (NELS 17)*. 274–89.

Hayes, Bruce (1995). *Metrical stress theory: Principles and case studies*. Chicago & London: The University of Chicago Press.

Hualde, José Ignacio & Elordieta, Gorka & Gamind, Iñaki & Smiljanic, Rajka 2002. From pitch-accent to stress-accent in Basque. In: Gussenhoven, C. & Warner, N. (eds.) *Laboratory Phonology* 7. Berlin: Mouton de Gruyter. 547–584.

Hulst, Harry van der & van de Weijer, Jeroen 1991. Topics in Turkish phonology. In: Boeschoten, H. & Verhoeven, L. (eds.) *Turkish linguistics today*. Leiden: Brill. 11–59.

Hyman, Larry M. 2014. Do all languages have word accent? In: van der Hulst, H. (ed.) *Word accent: Theoretical and typological issues*. Cambridge: Cambridge University Press. 56–82.

Inkelas, Sharon 1999. The exceptional stress-attracting suffixes in Turkish: Representations versus the grammar. In: Kager, R. & van der Hulst , H. & Zonneveld, W. (eds.) *The prosody-morphology interface*. Cambridge: Cambridge University Press. 134–187.

Inkelas, Sharon & Orgun, Cemil Orhan 1995. Level ordering and economy in the lexical phonology of Turkish. *Language* 71, 763–793.

Inkelas, Sharon & Orgun, Cemil Orhan 1998. Level (non)ordering in recursive morphology: evidence from Turkish. In: Lapointe, S. G. & Brentari, D. K. & Farrell, P. M. (eds.) *Morphology and its relation to phonology and syntax*. Stanford: CSLI. 360–410.

Ipek, Canan 2015. The phonology and phonetics of Turkish intonation. PhD dissertation, USC.

Jiang, Haiyan, Liu, Yan, & Lu, Li 2010. 维吾尔语词重音实验研究。《民族语文》，第三期， [An experimental study of stress in Uyghur words. *Ethnic Study* 3.], 67−71.

Jakobson, Roman 1968 [1941]. *Child language, aphasia, and phonological universals*. Mouton: The Hague.

Johanson, Lars 1998. The structure of Turkic. In: Johanson, L. & Csató, É. Á. (eds.) *The Turkic languages.* London & New York: Routledge. 30–66.

Kabak, Barış & Vogel, Irene 2001. The phonological word and stress assignment in Turkish. *Phonology* 18, 315–360.

Kager, René 1989. *A metrical theory of stress and destressing in English and Dutch*. Dordrecht: Foris Publications.

Kaisse, Ellen 1985. Some theoretical consequences of stress rules in Turkish. *Chicago Linguistic Society* 21: 1, 199–209.

Kamali, Beste 2011. Topics at the PF interface of Turkish. (PhD dissertation) Harvard University.

Kamali, Beste 2014. Beyond morphosyntax: Interrogative intonation and its role in Turkish. *Turkic Languages* 13, 189–206.

Қайдаров, А. Т. 1997. Казахский язык: In: Тенишев, Э. Р. (ed.). *Языки мира. Тюркские языки*. Бишкек: Кыргызстан. 242–254.

Kirchner, Mark 1998. Kazakh and Karakalpak. In: Johanson, L. & Csató, É. Á. (eds.) *The Turkic languages*. London & New York: Routledge. 318–332.

Konrot, Ahmet 1981. Physical correlates of linguistic stress in Turkish. *University of Sussex Language Centre Occasional Papers* 24, 26–52.

Konrot, Ahmet 1987. Stress in Turkish: Is it determined phonologically or morphologically? In: Boeschoten, H. E. & Verheoven, L. T. (eds.) *Studies on modern Turkish: Proceedings of the third conference on Turkish linguistics*. Tilburg: Tilburg University Press.

Krippes, Karl A. 1993. *Kazakh (Qazaq) Grammatical sketch with affix list.* Columbia, MD: Hierglyphic Press.

Krueger, John Richard 1961. *Chuvash manual*. [Uralic and Altaic series 7.] Bloomington: Indiana University Press.

Ladd, D. Robert 1996. *Intonational phonology*. Cambridge: Cambridge University Press.

Lees, Robert B. 1961. *The phonology of Modern Standard Turkish*. Bloomington: Indiana University Publications.

Levi, Susannah V. 2005. Acoustic correlates of lexical accent in Turkish. *Journal of the International Phonetic Association* 35: 1, 73–97.

Lewis, Geoffrey 1967. *Turkish grammar*. Oxford: Oxford University Press.

Liang, Jie 2009. An analysis of the stress features of disyllabic words in Uygur. *Journal of Minzu University of China* 2, 115–119.

Liang, Jie & Zhang, Jisheng 2008. 维吾尔语双音节词重音的产出实验研究。《第八届中国语音学术会议暨庆贺吴宗济先生百岁华诞语言科学前沿问题国际研讨会论文集》 [A Production Study on Stress Pattern on the Disyllabic Words in Uyghur Language. *8th Chinese Phonetics Conferences and Language Science International Symposium on Frontier Issues in the Honor of Celebrating Mr. Wu Zongxian's Centary Birthday.*]

McCarthy, John & Prince, Alan 1986. *Prosodic morphology*. Ms., Brandeis University and University of Massachusetts, Amherst.

McCarthy, John & Prince, Alan 1993a. Generalized alignment. In: Booij, G. E. & van Marle, J. (eds.) *Yearbook of Morphology*. Dordrecht: Kluwer. 79–153.

Muhamedowa, Raihan 2016. *Kazakh: A comprehensive grammar*. Routledge: New York.

Muti, Ibrahim 2007. Uyghur tilidiki söz urghusi we ibare intonatsiyisi. Memlikitimizdiki Turkiy tillar we bu heqtiki ilmiy tetiqiqat. *Ibrahim Muti maqaliliri*. Milletler Neshiryati, Beijing. 40–59.

Özçelik, Öner. 2013. Exceptions in stress assignment: Feet in input. In: Kan, S. et al. (eds.) *Proceedings of NELS 40.* GLSA: Amherst, MA. 107–121.

Özçelik, Öner 2014. Prosodic faithfulness to foot edges: the case of Turkish stress. *Phonology* 31, 229–269.

Özçelik, Öner (to appear). The Foot is not an obligatory constituent of the Prosodic Hierarchy: "Stress" in Turkish, French and child English. *The Linguistic Review* 34: 1 [2017].

Pratten, David & Omarova, Anargool 1994. *Kazakh notes*. Almaty: Senim.

Peterson, Gordon & Lehiste, Ilse 1960. Duration and syllable nuclei in English. *Journal of the Acoustic Society of America* 32, 693–703.

Prince, Alan 1990. Quantitative consequences of rhythmic organization. *Papers from the 26th Regional Meeting of the Chicago Linguistic Society*. 355–398.

Prince, Alan & Paul Smolensky 1993. *Optimality theory: Constraint interaction in generative grammar*. Ms., Rutgers University and University of Colorado.

Prokosch, Eduard 1939. *A comparative Germanic grammar*. Philadelphia, PA: Linguistic Society of America.

Pycha, Anne 2006. A duration-based solution to the problem of stress realization in Turkish. *UC Berkeley Phonology Lab Annual Report*. 141–151.

Schiering, René & van der Hulst, Harry 2010. Word accent systems in the languages of Asia. In: van der Hulst, H. & Goedemans, R. & van Zanten, E. (eds.) *A survey of word accentual systems in the languages of the world.* Berlin: Mouton de Gruyter. 509–614.

Selkirk, Elisabeth 1995. Sentence prosody: Intonation, stress and phrasing. In: Goldsmith, J. (ed.) *The handbook of phonological theory*. Cambridge, MA: Blackwell. 550–569.

Sezer, Engin 1983. On non-final stress in Turkish. *Journal of Turkish Studies* 5, 61–69.

Somfai-Kara, Dávid 2002. *Kazak*. Lincom Europa.

Underhill, Robert 1976. *Turkish grammar*. Cambridge, MA: MIT Press.

Vogel, Irene 2009. Universals of prosodic structure. In: Scalise, S. & Magni, E. & A. Bisetto (eds.) *Universals of language today*. Dordrecht: Springer. 59–82.

Woodrow, Herbert 1909. A comparative study of rhythm. *Archives of Psychology* 14. 1–66.

Woodrow, Herbert 1951. Time perception. In: Stevens, S. S. (ed.) *Handbook of experimental psychology.* New York: Wiley. 1224–1236.

Yakup, Mahire 2013. *Acoustic correlates of lexical stress in native speakers of Uyghur and L2 learners*. PhD Dissertation, University of Kansas.

Phonological constraints restricting the consonantal distribution in Turkish mimetic words

Kentaro Suganuma

Suganuma, Kentaro 2015. Deverbal nominals in Altaic: In search of a framework (not only) for reconstruction. *Turkic Languages* 19, 193–212.

This paper proposes the following three constraints which restrict the consonantal distribution in Turkish mimetic words: (i) Coronal constraint (domain: base): alveolars must not precede palatals; (ii) Agree-[voice]/obs (domain: base): if C_1 and C_2 are obstruents, their feature [voice] must agree; (iii) $*D]_{\sigma}$: voiced obstruents with the feature [-continuant] must not appear in coda positions. Furthermore, the paper indicates the following two things: (i) Coronal constraint and $*D]_{\sigma}$ relate to the markedness of consonants, and Agree-[voice]/obs relates to the assimilation of consonants; (ii) there are phonological similarities and differences between Turkish mimetic and non-mimetic words in terms of whether or not the constraints above hold.

Kentaro Suganuma, Kyushu University, 6-10-1 Hakozaki Higashi-ku Fukuoka, Japan.
E-Mail: baykentaro@gmail.com

1. Introduction

Generally, almost all languages have *mimetic words*. Mimetic words are words which imitate sounds, psychological states, or manners. For example in Turkish, *hav hav* 'bow wow' imitates dogs' barking, and *hüngür hüngür* 'sobbing' imitates the sound of crying.

Mimetic words have *sound symbolism* as their defining characteristic. Sound symbolism refers to a close relationship between the sound and meaning of words. De Saussure (1916 [1984]: 97–102) notes that language symbols are arbitrary and that no intrinsic association holds between the sound and meaning of words. However, the existence of sound symbolism has been demonstrated by many researchers (Sapir 1929, Ohala 1994, and Shinohara & Kawahara 2015), and it also has been suggested that sound symbolism is reflected in mimetic words (Zülfikar 1995 and Hamano 1998).

Previous studies of mimetic words (Zülfikar 1995, Hamano 1998, Ido 1999, etc.) have focused mainly on sound symbolism and/or morphology. Accordingly, the phonological system of mimetic words has received less attention. Although mimetic words are exceptions to arbitrariness, they are not outside the language system. They should have a phonological system and it ought to be investigated. In fact,

some previous studies show that mimetic words indeed follow a phonological system (see Nasu 1999, 2005, 2007 for Japanese mimetic words and Jun 1994 for Korean mimetic words). Furthermore, in Nasu (1999, 2005, 2007) it is said that there are similarities and differences between the phonological system of mimetic words and that of non-mimetic words. Understanding the phonological system of mimetic words and understanding what similarities and differences exist between mimetic and non-mimetic words will lead us to a better understanding of the whole phonological system of an individual language.

As one such attempt, this paper intends to describe the phonological system of Turkish mimetic words. In particular, the following three points will be discussed:

(1) a. There are three phonological constraints which restrict the distribution of consonants in Turkish mimetic words.
b. Two of these constraints are related to the markedness of consonants, and one to the assimilation of consonants.
c. There are phonological similarities and differences between Turkish mimetic and non-mimetic words.

In the next section, we shall discuss the characteristics of Turkish phonology and Turkish mimetic words.

2. An outline of Turkish phonology

2.1. Consonant phonemes of Turkish

The subject language of this paper is Turkish, which is spoken mainly in the Republic of Turkey. Turkish has 24 consonant phonemes.

(2)

	Labial	Coronal		Dorsal	Glottal
		Alveolar	Palatal		
Plosive	*p b*	*t d*	*(k') (g')*	*k g*	
Fricative	*f v*	*s z*	*ş (j)*	*ğ*	*h*
Affricate			*ç c*		
Nasal	*m*	*n*			
Liquid		*r l*	*(l')*		
Approximant			*y*		

Note: Although *ğ* is categorized as a velar fricative in this table, its feature value is controversial, as will be discussed below.

Parenthesized phonemes are distinctive only in loanwords, such as the loanword *k'ar* 'profit' vs. the native word *kar* 'snow'. These parenthesized phonemes do not appear in mimetic words.

As shown in (2), Turkish consonants can be classified into four groups according to their place of articulation: (i) labial, (ii) coronal, (iii) dorsal, and (iv) glottal. Coronal consonants can be further divided into alveolar and palatal phonemes. Turkish has a voiced vs. voiceless contrast in obstruents, but no contrast in sonorants, as shown in (3).

(3) Voiced obstruents ([+voiced]): *b, d, g, v, c, z*
Voiceless obstruents ([-voiced]): *p, t, k, f, s, ç, h, ş*
Sonorants ([+voiced]): *m, n, r, l, y*

Another consonant remains: *ğ* (*soft-g*). In (3), *ğ* is not included in any category because it has so many allophones that we cannot easily determine its feature values. As Göksel & Kerslake (2005: 7) say, *ğ* is pronounced as a voiced velar fricative /ɣ/ in some dialects. However, in standard Turkish, *ğ* is never pronounced as [ɣ]. Roughly speaking, *ğ* is realized as a palatal approximant between front vowels (e.g. *eğitim* [ejitim] 'education') and is deleted between back vowels (e.g. *ağır* [aɨr] 'heavy'). In coda position, it triggers the lengthening of preceding vowels (e.g. *dağ* [da:] 'mountain'). Since the feature values of *ğ* are difficult to determine, mimetic words containing *ğ* are excluded from the analysis in this paper.

2.2. Mimetic words in Turkish

2.2.1. Data

The source of mimetic word data used in this paper is the mimetic dictionary in Zülfikar (1995: 281–695). However, Zülfikar's (1995) dictionary contains many dialectal mimetic words, some of which are gathered from relatively old literature (earlier than the 20th century). To get rid of the effects of individual dialectal phonology, and to focus on more common and generally modern mimetic words, two Turkish native speakers were asked whether the mimetic word data in Zülfikar's (1995) dictionary contained mimetic words that they use or have heard in normal conversation. The mimetic words for which both responded "I have used or heard it" are treated as the data for this study. Both of the Turkish native speakers aforementioned are male; one is from Izmir and the other from Nevşehir. They went through their language acquisition stages in these areas, and they speak standard Turkish, as far as investigated. (We will briefly discuss all the data in Zülfikar (1995) in section 5, and the data obtained through the method mentioned above are shown in the appendix.)

Turkish has reduplicated mimetic words. They are generally used as adverbs or adjectives, and some are used as nouns, as shown in (4).

(4) a. Adverbial usage
Demirler çın çın çarpıştı.
iron: PL clang clash: REC: PAST3SG
'Irons clashed with each other with clanging sounds.'

b. Adjectival usage
çıtır çıtır patates cipsi
crispy potato chips
'crispy potato chips'

c. Nominal usage
cav cav
babbling
'babbling person'

These reduplicated elements (e.g. for *çın*, *çıtır* and *cav*, we call these elements *bases* in this paper) can be used without reduplication. Generally, such non-reduplicated forms are used as adverbs with *diye* following (the converbial form of the verb 'say').

(5) *Cart diye yırttım.*
tearing sound say:CON tear up:PAST1SG
'I tore up (something).'

Some mimetic words in Zülfikar (1995) have no reduplicated form, and among them some are words about which it is uncertain whether or not they can be regarded as mimetic words (e.g. *beçik* 'young goat'). So this paper deals only with the mimetic words that are always used in a reduplicated form.

2.2.2. Phonological and morphological structure of bases

When we classify bases by syllable structure, we can see that there are five types of base in Turkish.

(6) a. C_1VC_2 e.g. *çın* → *çın çın* 'clang' 89 items
b. C_1VC_2V e.g. *gulu* → *gulu gulu* 'cry of a turkey' 12 items
c. C_1VC_2VC e.g. *çıtır* → *çıtır çıtır* 'crispy' 102 items
d. C_1VC_2CVC e.g. *hüngür* → *hüngür hüngür* 'sobbing' 37 items
e. C_1VC_2C e.g. *cart* → *cart cart* 'tear up' 10 items

In this paper, we will examine the consonant distribution in Turkish mimetic words, focusing on C_1 and C_2, i.e., on the onset of the first syllable of the base and the next consonant after C_1. For this reason, bases that do not have C_1 (e.g. *efil efil* 'breezing') are excluded from the analysis.

Previous studies (Zülfikar 1995 and Ido 1999) point out that some of these bases contain one or more affixes. At first, almost all CVCVC and CVCCVC bases have *-Ir* or *-Il* as the final VC part. The large capital *I* represents the archiphoneme for high vowels. It alternates among *ı ~ i ~ u ~ ü* according to vowel harmony. Zülfikar (1995: 101) and Ido (1999: 72) analyzes these sequences as suffixes that affix to mimetic morphemes, and Ido (1999: 72) suggests that CC sequences in CVCCVC bases are infix-like morphemes, which he calls *extenders*. These CC sequences are almost always *mb*, *ld,* or *ng*, and we see an affix-like regularity in these sequences, as shown in (7).

(7)

mb	*gümbür*	'rolling'
	tımbır	'sound of a tambourine'
ld	*güldür*	'a sound of a waterfall'
	feldir	'a watchful nasty eye'
ng	*hüngür*	'sobbing'
	şıngır	'jingling'

Furthermore, almost all CVCC bases have *t* in their base-final positions. Ido (1999: 72) argues that the final *t* is also an extender. Zülfikar (1995: 98), on the other hand, suggests that CVCC bases are derived from CVC bases by liquid insertion into the C_2 position, or that CVC bases are derived from CVCC bases by liquid deletion. In the data for this paper, CVCC bases do not have corresponding CVC bases. However, if CVCC bases are derived from CVC bases by *r*-insertion, that inserted *r* can also be regarded as a kind of affix.

(8) e.g. *cart* 'tear up'

Ido (1999: 72)	*car-t* → *cart*	
Zülfikar (1995: 98)	*cat*→ Liquid insertion	*ca-r-t* → *cart*

Zülfikar (1995) and Ido (1999) do not describe the derivation or functions of these affixes in detail, and some bases lack these affixes (e.g. *fossur* and *zonk*). Thus, the status of these affixes remains controversial. This paper, however, assumes that these sequences are infixes and suffixes, as they have argued. Their arguments give us the following morphological structures as mimetic bases.

(9) a. C_1VC_2
b. C_1VC_2V
c. C_1VC_2-VC
d. C_1V-C_2C-VC
e. C_1V-C_2-C

Hyphens represent morpheme boundaries.

(9d, e) show that the C_2s of CVCCVC and CVCC bases are consonants belonging to infixes or suffixes and not consonants of mimetic roots. On the other hand, C_2s of CVC, CVCV, and CVCVC bases are true components of mimetic roots. This means that we cannot treat C_2s in (9a–c) and in (9d, e) the same way. Thus, this study examines only CVC, CVCV, and CVCVC bases, while CVCC and CVCCVC bases are excluded from the analysis and set aside for future research.

The following sections examine the consonant distribution and propose phonological constraints.

3. Analysis 1: Focusing on place of articulation

3.1. An asymmetrical coronal distribution

First, we examine the distribution of consonants in C_1 and C_2 with respect to their places of articulation. When we classify bases according to the place of articulation of C_1 and C_2, we can see a distribution of consonants shown in (10).

(10)

C_1 \ C_2	Labial	Coronal	Dorsal	Glottal
Labial	2	49	5	2
Coronal	30	37	27	0
Dorsal	7	17	2	0
Glottal	5	13	1	0

Note: This table includes CVC, CVCV, and CVCVC bases.

The vertical column represents the consonants in C_1, while the horizontal row represents the consonants in C_2. The numbers indicate the frequencies of bases with the particular consonantal sequences. For instance, the 2 in the upper-left cell of the table indicates that there are two bases with labial consonants in both C_1 and C_2.

When we look at this table giving special attention to the bases in which C_1 and C_2 have identical place of articulation (highlighted by thick frames), there are not many bases in which both C_1 and C_2 are labial, dorsal or glottal consonants. On the other hand, there are many bases where both C_1 and C_2 are coronal. This distribution

is reasonable because the number of phonemes in each place of articulation differs, and coronal consonants are the most numerous among them.

(11)

	Number	Inventory
Labial	5	*p, b, f, v, m*
Coronal	11	*t, d, s, z, n, r, l, ş, ç, c, y*
Dorsal	3	*k, g*
Glottal	1	*h*

When we look more closely at the bases in which both C_1 and C_2 are coronal, however, we can see an asymmetrical distribution of coronal consonants.

(12)

C_1 \ C_2	Alveolar	Palatal
Alveolar	16	0
Palatal	18	3

Table (12) indicates that Turkish has bases containing the sequence [C_1: Palatal C_2: Alveolar] but not the sequence [C_1: Alveolar C_2: Palatal]. These two sequences are mirror images of each other, but only the former exists.

(13)

[ok] [C_1: Palatal	C_2: Alveolar]	*[C_1: Alveolar	C_2: Palatal]
e.g. [ok]*çat*	[ok]*cızır*	e.g. **taç*	**zıcır*

In the next section, we will discuss why such an asymmetrical distribution can be observed.

3.2. Markedness of palatals

Since the aforementioned asymmetrical distribution can be seen in C_1 and C_2, differences between phonological positions (i.e., C_1 and C_2) correlate with the distribution. Beckman (1998: 1) has said that there are two types of phonological position, as shown in (14). One is the privileged position, where phonological contrasts are preserved and marked segments (complex, specialized, or phonologically active) tend to appear, and the other is the non-privileged position, where phonological contrasts are not preserved and unmarked segments tend to appear.

(14)

	Privileged positions	Non-privileged positions	
a.	Root-initial syllables	Non-initial syllables	
b.	Stressed syllables	Unstressed syllables	
c.	Syllable onsets	Syllable codas	
d.	Roots	Affixes, clitics, function words	
e.	Long vowels	Short vowels	Beckman (1998: 1, (1))

When we focus on (14a, c), we can see that the contrast between C_1 and C_2 of CVCV and CVCVC bases corresponds to (14a), and the contrast between C_1 and C_2 of the CVC base corresponds to (14c). When they are all put together, C_1 is in the privileged position and C_2 is in the non-privileged position, as shown in (15).

(15)

	Privileged positions	Non-privileged positions
CVCV, CVCVC	C_1: (Onset of) initial syllables	C_2: (Onset of) non-initial syllables
CVC	C_1: Syllable onsets	C_2: Syllable codas

This paper argues that these differences between C_1 and C_2 cause the asymmetrical distribution. Although both alveolar and palatal consonants are coronal, their properties are not equal: one is marked and the other is unmarked, and if marked and unmarked segments appear in the same base, only the marked segments can appear in privileged positions, while unmarked segments have no choice but to appear in non-privileged positions. In other words, unmarked segments cannot appear in privileged positions over marked segments.

If it is correct, palatal consonants must be considered marked and alveolar consonants must be unmarked. Since the segments appearing in privileged position (i.e., C_1) are marked and those appearing in non-privileged position (i.e., C_2) are unmarked and as we have seen in (12), the segments in C_1 are palatal and those in C_2 are alveolar. According to Lees (1966: 36) and Clements & Sezer (1982: 249–250), palatal consonants change preceding vowels into high unrounded vowels in Turkish, especially in certain varieties of the Istanbul dialect, as shown in (16).

(16) Lees (1966: 36, line 9 to 12)

3. Palatal Umlaut: A harmonic short vowel is unrounded and de-harmonified immediately before a palatal /*y, ş, j, c, ç*/ within word boundaries if morpheme-final, or if not, in the first syllable, and is also raised there if that palatal consonant is followed by a vowel.

e.g. /*güneş-ye*/ ⇒ [*güneşe*] ~ [*günişe*]
sun:DAT
/*düş-üş-ecek*/ ⇒ [*düşüşeʤek*] ~ [*düşüʃicek*]
fall: REC:FUT3SG

On the other hand, alveolars do not trigger such phonological changes. This phenomenon shows that palatal segments are phonologically *active*, in the sense that they trigger some phonological changes. Such active segments are generally marked, according to Hume (2011: 80).

In sum, markedness of consonants is related to the asymmetrical distribution of coronal consonants. This paper proposes a *Coronal constraint*, shown in (17), to capture that distribution. This constraint disallows unmarked segments (i.e., alveolars) preceding marked segments (i.e., palatals).

(17) Coronal constraint (domain: base)
Alveolars must not precede palatals.
* [C_1: alveolar C_2: palatal]

It should, however, be noted that this constraint only holds with mimetic words, as shown in (18), and there are many non-mimetic words to which it does not apply; i.e. they fit into the pattern [C_1: Alveolar, C_2: Palatal].

(18) *saç* 'hair' *taş* 'stone' *duy-* 'to hear'

In the following section, we will see the consonant distribution in terms of feature [voice].

4. Analysis 2: Focusing on the feature [voice]

4.1. Distribution of consonants

To understand the distribution of consonants, we classify C_1 and C_2 in Turkish mimetic bases into three groups: (i) voiced obstruents, (ii) voiceless obstruents, and (iii) sonorants. The results for CVCV and CVCVC bases are shown in (19) and (20), respectively.

(19) CVCV bases

C_1 \ C_2	Obstruents [+voiced]	Obstruents [-voiced]	Sonorant
Obstruents [+voiced]	6	0	2
Obstruents [-voiced]	0	4	0
Sonorant	0	0	0

(20) CVCVC bases

C_1 \ C_2	Obstruents [+voiced]	Obstruents [-voiced]	Sonorant
Obstruents [+voiced]	16	6	10
Obstruents [-voiced]	1	44	10
Sonorant	3	5	2

When we look at bases with obstruents in C_1 and C_2 (highlighted by thick frames in Tables (19) and (20)), there are more bases in which C_1 and C_2 have identical values for the feature [voice] than in which C_1 and C_2 have different values for the feature [voice].

On the other hand, the result for CVC bases is the following.

(21) CVC bases

C_1 \ C_2	Obstruents [+voiced]	Obstruents [-voiced]	Sonorant
Obstruents [+voiced]	3	9	18
Obstruents [-voiced]	1	36	16
Sonorant	1	4	1

As in (19) and (20), more bases have the sequence [C_1: voiceless obstruent C_2: voiceless obstruent]. However, bases with the form [C_1: voiced obstruent C_2: voiced obstruent] are less numerous than those with the form [C_1: voiced obstruent C_2: voiceless obstruent].

(22) C_1 [+voiced] C_2 [+voiced]: 3 items
C_1 [+voiced] C_2 [-voiced]: 9 items

Thus, we can summarize this distribution of consonants as follows.

(23) If C_1 and C_2 are obstruents:

a. Many bases have C_1 and C_2 with identical values for the feature [voice].
C_1 and C_2 are voiced obstruents:

vɪdɪ	**fɪdɪ*	'bigmouth'
cɪz	**cɪs*	'frizzing'
gɪcɪ	**gɪçɪ*	'screaming'

C_1 and C_2 are voiceless obstruents:

piti	**biti*	'toddling'
fos	**vos*	'whispering'
küp	**küb*	'thumping'

b. In CVC bases, the number of bases with [C_1: voiced obstruent C_2: voiced obstruent] are less numerous than the number of bases with [C_1: voiced obstruent C_2: voiceless obstruent].

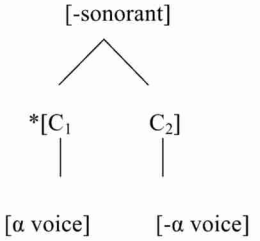
dık	**dıg*	'staring'
güp	**güb*	'thumping'
bıç	**bıc*	'squelch'

In the following sections, two phonological constraints will be proposed to account for this distribution.

4.2. Agree-[voice]/obs

First, we will focus on (23a). Since C_1 and C_2 have identical values for [voice], a phonological constraint can be formalized, as seen in (24).

(24) Agree-[voice]/obs (domain: base)
If C_1 and C_2 are obstruents, their feature [voice] must agree.



This is an assimilation-oriented constraint. This paper argues that the consonant distribution summarized in (23a) is caused by this constraint.

Just as for the Coronal constraint, this constraint does not hold in C_1 and C_2 of non-mimetic words as shown in (25).

(25) C_1 and C_2 of non-mimetic words: there is no restriction on feature [voice]

a. *kadın* 'woman', *taban* 'floor'	C_1: voiceless obstruent, C_2: voiced obstruent
b. *getir-* 'to bring', *batak* 'swamp'	C_1: voiced obstruent, C_2: voiceless obstruent

However, it could be said that this constraint holds in non-mimetic words in a different style. In Turkish, consonants in affixes assimilate to the [voice] value of pre-

ceding consonants. As a result, obstruents forming consonant clusters share a common value for the feature [voice]. Such assimilation could potentially be accounted for by the constraint Agree-[voice]/obs.

(26)		Locative /-*dA*/	Diminutive /-*cIk*/
	a. [+voiced] obstruent	*kazda* **kazta*	*kazcık* **kazçık*
		goose:LOC	goose :DIM
		evde **evte*	*evcik* **evçik*
		house:LOC	house:DIM
	b. [-voiced] obstruent	*kasta* **kasda*	*kasçık* **kascık*
		muscle:LOC	muscle:DIM
		lafta **lafda*	*lafçık* **lafcık*
		talk:LOC	talk:DIM

Note: Large capital *A* represents the archiphoneme of high vowels. It alternates between *a* ~ *e* according to vowel harmony.

In the next section we will focus on (23b).

4.3. *D]$_{\sigma}$

As we saw in (23b), some CVC bases violate Agree-[voice]/ obs, and they have [C_1: voiced obstruent C_2: voiceless obstruent]. This paper assumes that this is because C_2 of CVC bases is in coda position. It is well known that in Turkish coda position generally does not allow the emergence of voiced stops (including affricates). For example, some Turkish non-mimetic stems have voiced stops in stem-final position, but when these voiced stops are in coda position, they become voiceless as shown in (27).

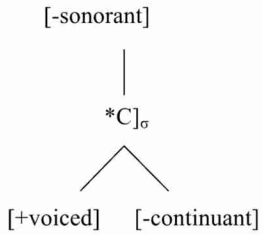
(27)				
	/dolab/	*do.lap*	*do.la.bı*	*do.lap.tan*
		cupboard	cupboard:ACC	cupboard:ABL
	/kilid/	*ki.lit*	*ki.li.di*	*ki.lit.ten*
		lock	lock:ACC	lock:ABL
	/güc/	*güç*	*gü.cü*	*güç.ten*
		power	power:ACC	power:ABL

Notes: Periods indicate syllable boundaries.

Some loanwords and some monosyllabic non-mimetic native words do not undergo this alternation. (e.g. /*dialog*/ *dialog*, **dialok* 'dialogue', /*ad*/ *ad*, **at* 'name').

The constraint which triggers this change can be formalized as shown in (28).

(28) *D]$_{\sigma}$ Voiced obstruents with the feature [-continuant] must not appear in coda positions.



The fact that mimetic words are also subject to this constraint could be said to give rise to the consonantal distribution shown in (23b). In other words, mimetic words are subject to *D]$_{\sigma}$, and since the requirement of *D]$_{\sigma}$ is stronger than that of Agree -[voice]/obs, bases with [C_1: voiced obstruent C_2: voiced obstruent] are less numerous than bases with [C_1: voiced obstruent C_2: voiceless obstruent].

It should be mentioned that *D]$_{\sigma}$ allows voiced fricatives to appear in the coda position, as shown in (29) below, since fricatives do not have [-continuant].

(29)

/*kaz*/	*kaz*	**kas*	*kaz.lar*	**kas.lar*
	goose		goose:PL	
/*ev*/	*ev*	**ef*	*ev.ler*	**ef.ler*
	house		house:PL	

The same can be said in mimetic words, where we can see that only voiced fricatives can appear in C_2 of CVC bases, as shown in (30).

(30) *cav* 'babbling,' *cız* 'sizzling or hissing sound,' *vız* 'buzzing'
Note: *v* and *z* are voiced fricative consonants.

Furthermore, *D]$_{\sigma}$ also holds for bases other than CVC. When we look at bases which have obstruents as a final consonant, we see that voiced stops do not appear, while voiced fricatives or voiceless obstruents do appear, as shown in (31).

(31) a. voiced fricatives *miyav* 'mewing'
b. voiceless obstruents *vıcık* 'gooey', *yalap* 'sparklingly'

In this section, we saw that *D]$_{\sigma}$ restricts the consonant distribution of mimetic words. *D]$_{\sigma}$ prohibits the emergence of voiced obstruents (or more accurately, voiced stops). Generally, voiced obstruents are regarded as marked elements be-

cause of their articulatory complexity. Therefore, it is clear that *D]$_{\sigma}$, as well as the Coronal constraint in (17), is related to the markedness of consonants.

5. Do the proposed constraints hold for all the mimetic words listed in Zülfikar (1995)?

This paper has proposed three constraints so far, but at the present stage these constraints hold for only limited data. In this section, we will examine whether or not the proposed constraints hold for all of the data provided by Zülfikar (1995). As we have done so until now, we will consider CVC, CVCV and CVCVC bases which have full-reduplicated form in this section, and as mentioned in 2.2.1., Zülfikar (1995) contains relatively old (pre 20th-century) mimetic words. Such bases are represented within parentheses in this section.

5.1. Coronal constraint

In Zülfikar (1995), there are only three exceptions to the Coronal constraint, as shown in (32).

(32) *dayır* 'suddenly and quickly'
tayır 'to walk around rapidly'
tıyır 'to flee secretly'

These exceptions have *y* as C_2 in common. The phoneme *y* is the only palatal sonorant in Turkish (except for *l'* which appears only in loanwords). It is conceivable that sonority and/or manner of articulation has something to do with the Coronal constraint, although it is not obvious at the present stage. In any case, since there are very few exceptions, it could be said that the Coronal constraint holds in almost all of Zülfikar's (1995) data.

5.2. Agree-[voice]/obs

Unlike the Coronal constraint, there are a considerable number of exceptions to Agree-[voice]/obs, as shown in (33). Accounting for these exceptions is a subject for future analysis. Depending on the results of future analysis, it will most likely be necessary to revise the constraint. We will return to this problem again in section 6.

(33) a. [C_1: voiced obstruent C_2: voiceless obstruent]: 18 items
beçi, bıçı, biçi
caf, cakır, cas, cepir, cıf, cisen, cisil, cokur, copur, cukur
dakır
gepir, gıçı, güpür
vaf, vakır, vıkır
zıpır

b. [C_1: voiceless obstruent C_2: voiced obstruent]: 13 items
(çaġ), (çaġıl), çigir
hıcıl
kıbıl, kıdı, kıdık, kıdım, (ḳıġış), kıvıl, kıvış
sızım
şabır
hav
Notes: Meanings of mimetic words are omitted.
ḳ and *ġ* represent *k* and *g* in a syllable which contains a back vowel, respectively.

5.3. *D]$_\sigma$

There are only three exceptions to *D]$_\sigma$ as shown in (34). They are all monosyllabic.

(34)

bıd	'to stagger'
bid	'small animals' hopping'
(çaġ)	'to cascade'

Çaġ is obtained from old literature (this is taken from *Muhammediyye*, written by Yazıcıoğlu Mehmed in 1449, and *Cinânü'l-Cenân*, written by Hacı Ivazoğlu Mehmed in the 15th century as quoted in *Tarama Sözlüğü* (1965)), so if we exclude it, the only exceptions that remain end with *d*. Actually, *d* in coda position is also allowed in several monosyllabic non-mimetic native words (e.g. *ad* 'name' *öd* 'bile'), although native words are also subject to *D]$_\sigma$. Thus, monosyllabicity and *d* have something to do with these exceptions. At the very least, it could be said that *D]$_\sigma$ holds in almost all of Zülfikar's (1995) data, since the exceptions are so few in number.

6. Concluding remarks

The main findings of this paper are the following.

(35) = (1)

a. There are three phonological constraints restricting the distribution of consonants in Turkish mimetic words.
b. Two of these constraints are related to the markedness of consonants, and one is related to the assimilation of consonants.
c. There are phonological similarities and differences between Turkish mimetic and non-mimetic words.

As for (35a), this paper proposes the Coronal constraint, Agree-[voice]/obs, and *D]$_\sigma$, discussed in section 3 and 4. Furthermore, as for (35b), this paper indicates that the Coronal constraint and *D]$_\sigma$ relate to the markedness of consonants and

Agree-[voice]/obs relates to the assimilation of consonants. Finally, as for (35c), this paper indicates that there are phonological similarities and differences between Turkish mimetic and non-mimetic words, in terms of whether these constraints hold or not. The Coronal constraint holds only for non-mimetic words. On the other hand, Agree-[voice]/obs holds for both non-mimetic and mimetic words, but the domains differ. In the former, the domains are the adjacent consonant sequences, which are in the morpheme boundary. In the latter, the domain is the base. Finally, *D]$_{\sigma}$ holds for both non-mimetic and mimetic words. We can sum up the findings as in (36).

(36)

	Non-mimetic words	Mimetic words
Coronal constraint	does not hold e.g. *saç, taş* = (18)	holds e.g. ^{ok}*cızır*, **zıcır* = (13)
Agree-[voice]/obs	holds (domain: C-C) e.g. *kazda*, **kazta* = (26)	holds (domain: base) e.g. *gıcı*, **gıçı* = (23a)
*D]$_{\sigma}$	holds e.g. *dolap*, **dolab* = (27)	holds e.g. *dık*, **dıg* = (23b)

Additionally, the findings indicate that the Coronal constraint and *D]$_{\sigma}$ hold in almost all of Zülfikar's (1995) data.

Several problems remain that need to be addressed in the future. Above all, there are exceptions to Agree-[voice]/obs (regardless of whether we consider all Zülfikar's (1995) data or not). These exceptions need to be explained, and in order to do so, we need to find what common characteristics these exceptions have. At the present stage, some features of these exceptional cases can be identified as shown in (37); however, it is unclear how these characteristics relate to Agree-[voice]/obs.

(37) (Data is the same as in (33))

a. C_1: Voiced obstruent C_2: Voiceless obstruent

i. Some of these exceptions have the same consonantal sequence.

ii. There are many bases that have *c* in C_1.

C_1: *b*, C_2 *ç* : *beçi, bıçı, biçi*

C_1: *c*, C_2 *k* : *cakır, cokur, cukur*

C_1: *c*, C_2 *s* : *cas, cisen, cisil*

C_1: *c*, C_2 *f* : *caf, cıf*

C_1: *c*, C_2 *p* : *cepir, copur*

C_1: *g*, C_2 *p* : *gepir, güpür*

C_1: *v*, C_2 *k* : *vakır, vıkır*

(Bases with unique consonantal sequences: *dakır, gıçı vaf, zıpır*)

b. [C_1: voiceless obstruent C_2: voiced obstruent]: 13 items
   i. Some of these exceptions have the same consonantal sequence.
   ii. There are many bases with *k* in C_1.
   C_1: *ç*, C_2 *ġ* or *g*:*(çaġ), (çaġıl), çigir*
   C_1: *k*, C_2 *d* : *kıdı, kıdık, kıdım*
   C_1: *k*, C_2 *ġ* : *(ḳıġış)*
   C_1: *k*, C_2 *v* : *kıvıl, kıvış*
   (Bases with unique consonantal sequences: *kıbıl, hav, hıcıl, sızım, şabır*)
   Note: If we assume that *h* is a sonorant, as did Chomsky & Halle (1968: 177), bases containing *h* (e.g., *hav*) do not need to be treated as an exception.

Problems such as the above remain. If this research is pursued further, our knowledge of Turkish phonology will become richer, and such research may also contribute to the phonological research on mimetic words in other languages.

Acknowledgements

I thank the reviewers for their cooperation and highly appreciate the comments and suggestions given. I also thank the two Turkish native speakers for providing data essential for this research.

In addition, this paper is based on my two previously published papers Suganuma (2012a and b). I thank the reviewers of those papers as well, since their comments and suggestion are also reflected in this paper. The constraints proposed in this paper are basically the same as those in the previously published papers, although the data is different. Each set of data in the two previous papers was obtained from a different native speaker. For this paper, I again asked the native speakers who had helped with the past papers about the mimetic words. Those mimetic words to which both of them responded 'I have used or heard it' are treated as the data for this paper, as explained in Section 2.2. Moreover, this paper adds a new discussion in Section 5.

Abbreviations

1	first person		
3	third person		
ABL	ablative	ACC	accusative
CON	converb	DAT	dative
DIM	diminutive	FUT	future
LOC	locative	PL	plural
REC	reciprocal	SG	singular

References

Beckman, Jill N. 1998. Positional faithfulness: An Optimality Theoretic treatment of phonological asymmetries. [Ph.D. dissertation, University of Massachusetts, Amherst].

Chomsky, Noam & Halle, Morris 1968. *The sound pattern of English*. Cambridge, MA: MIT Press.

Clements, George N. & Sezer, Engin 1982. Vowel and consonant disharmony in Turkish. In: van der Hulst, Harry & Smith, Norval (eds.) *The structure of phonological representations* 2. Dordrecht: Foris. 213–255.

De Saussure, Ferdinand 1916 [1984]. *Cours de linguistic générale*. Paris: Payot.

Hamano, Shoko 1998. *The sound-symbolic system of Japanese*. Stanford: CSLI Publications.

Hume, Elizabeth 2011. Markedness. In: van Oostendorp, Marc & Ewen, Colin J. & Hume, Elizabeth & Rice, Keren (eds.) *The Blackwell companion to phonology 1: General issues and segmental phonology*. Chichester: Wiley-Blackwell. 79–106.

Ido, Shinji Ghaeyri 1999. Turkish mimetic word formation. *Asian and African Studies* 8, 67–73.

Jun, Jongho 1994. Metrical weight consistency in Korean partial reduplication. *Phonology* 11, 69–88.

Lees, Robert B. 1966. On the interpretation of a Turkish vowel alternation. *Anthropological Linguistics* 8 (9), 32–39.

Nasu, Akio. 1999. Onomatope ni okeru yuuseika to [p] no yuuhyousei [Voicing and markedness of [p] in mimetic words]. *Onsei kenkyuu* [Journal of the Phonetic Society of Japan] 3, 52–66.

Nasu, Akio. 2005. Onomatope no on'inkoozoo ni miru hitaishoosei [Asymmetrical nature in phonological structure of mimetic words]. *Onsei kenkyuu* [Journal of the Phonetic Society of Japan] 9, 20–29.

Nasu, Akio. 2007. Onomatope no gengogakuteki tokuchoo [The linguistic characteristics of onomatopoetic words]. *Nihongo gaku* [Japanese linguistics] 26, 4–15.

Ohala, John J. 1994. The frequency codes underlies the sound symbolic use of voice pitch. In: Hinton, Leanne & Nichols, Johanna & Ohala, John J. (eds.) *Sound symbolism*. Cambridge: Cambridge University Press. 325–347.

Sapir, Edward. 1929. A study in phonetic symbolism. *Journal of Experimental Psychology* 12, 225–239.

Shinohara, Kazuko & Kawahara, Shigeto 2015. A cross-linguistic study of sound symbolism: The images of size. *Proceedings of BLS 36*. Berkeley: Berkeley Linguistic Society. 396–410.

Suganuma, Kentaro. 2012a. *Torukogo onomatope ni okeru zecchoosiin no hairetu ni kansuru seiyaku* [The phonotactic constraint of coronal consonants in Turkish mimetic words]. *Onin kenkyuu* [Phonological Studies] 15, 59–66.

Suganuma, Kentaro. 2012b. *Torukogo onomatope ni okeru siin no yuuseisei ni kansuru seiyaku* [Constraints on the [voice] of consonants in Turkish onomatopoetic words]. *Kyuushuu daigaku gengogaku ronsyuu* [Kyushu University Papers in Linguistics] 33, 41–68.

Tarama Sözlüğü. Ankara: Türk Dil Kurumu Yayınları. 1965.

Zülfikar, Hamza 1995. *Türkçede ses yansımalı kelimeler*. Ankara: Türk dil kurumu yayınları.

Appendix. The list of Turkish mimetic bases

CVC bases

bam	*çim*	*cuk*	*düm*	*hav*
caf	*çın*	*dan*	*fır*	*hır*
cak	*çip*	*dar*	*fıs*	*hış*
çak	*çıp*	*dik*	*fış*	*hop*
çan	*cır*	*dık*	*fos*	*hor*
cap	*çır*	*dım*	*gar*	*kih*
çat	*çıt*	*dın*	*gır*	*kiş*
cav	*cız*	*dır*	*güm*	*kıs*
cik	*çok*	*dom*	*gür*	*kış*
çık	*cor*	*dum*	*har*	*kös*

küp	*pıs*	*şor*	*tok*
küt	*pıt*	*tak*	*ton*
lak	*poh*	*tik*	*tos*
lap	*rap*	*tık*	*vik*
löp	*şak*	*tin*	*vın*
mır	*şap*	*tın*	*vır*
mız	*şar*	*tip*	*vız*
pat	*şik*	*tıp*	*zıp*
peh	*şık*	*tir*	*zır*
pır	*şır*	*tıs*	

CVCV bases

badi	*gıdı*
bici	*gulu*
bıcı	*pisi*
bıdı	*piti*
bili	*tepe*
çiki	*vıdı*

CVCVC bases

bayır	*cayır*	*ciyak*	*fışır*	*gurul*	*homur*
bicik	*cıbıl*	*cıyak*	*fokur*	*gurül*	*hopur*
bıcık	*cıbır*	*cızır*	*fosur*	*gürül*	*horul*
bıcır	*ciğil*	*cokur*	*foşur*	*hapır*	*hosur*
bıdık	*cığıl*	*dakır*	*gıcık*	*harıl*	*hoşur*
bücük	*çiğil*	*fakır*	*gıcır*	*haşır*	*hösür*
cağıl	*çipil*	*fasır*	*gıdık*	*hatır*	*kikir*
çağıl	*cırıl*	*fıkır*	*gımıl*	*hıkır*	*kıkır*
cakır	*çıtır*	*fısıl*	*gümül*	*hışır*	*kımıl*
çatır	*cıvıl*	*fısır*	*güpür*	*hıtır*	*kıpır*

kıpış	*miyav*	*şapır*	*tıkır*	*vizir*
kıtır	*mızık*	*şarıl*	*tıpır*	*vızır*
kütür	*mızır*	*şıpır*	*tıpış*	*zıpır*
lakır	*parıl*	*şırıl*	*tiril*	*zırıl*
lıkır	*pırıl*	*sızım*	*tısıl*	
lokur	*pısır*	*şorul*	*tosur*	
löpür	*pıtır*	*şupur*	*vazır*	
mıcık	*püfür*	*takır*	*vıcık*	
mırıl	*pütür*	*tapır*	*vıkır*	
mışıl	*şakır*	*tıkıl*	*viyak*	

Causatives in Uyghur

Umarjan Kurban

Kurban, Umarjan 2015. Causatives in Uyghur. *Turkic Languages* 19, 213–239.

The present study is intended to provide a descriptive analysis of canonical causative constructions in modern Uyghur. Lexical, morphological and analytic causative constructions are employed in modern Uyghur. Lexical causative is less productive, and its use seems to be restricted to very specific domains without a morphological marker. By contrast, morphological causatives are undoubtedly the most frequent means of expressing causatives, and they are highly productive with both transitive and intransitive verbs. These productive single causatives are mostly regular; the choice of the suffixes is phonologically determined. Double causatives derived from intransitive and transitive bases in Uyghur result in different surface realizations. No more than two different causative morphemes can be iterated in the causative construction. Therefore, Uyghur does not permit any iteration of the triple causative. In this sense, the double causative in Uyghur is different from that of genetically related languages (for example Turkish) with respect to the nature and amount of causative morpheme reduplication. This study shows that the morphological causatives in Uyghur share universal tendencies in terms of causative devices and valence increasing operations in the argument structure of causative verbs in canonical situations. Although the scope of this paper is limited to one individual language, its findings would be relevant for studying causative constructions in other languages.

Umarjan Kurban, College of Humanities, Xinjiang University, Urumqi, 830046 Xinjiang, China. E-mail: tarim333@hotmail.com

0. Introduction

Causatives have been the subject of intense study in linguistic literature and have long been discussed on the basis of cross-linguistic investigations in terms of functional typology (Comrie 1975, 1976, 1981, 1985, 1989, 1993; Haspelmath 1993; Dixon 2000), morphological and syntactic classification (Nedjalkov & Silnitsky 1973; Shibatani 1975, 1976; Aissen 1979; Givon 1980, Foley & Van Valin 1984; Haiman 1985; Song 1996), semantic types of causatives (Shibatani 1975, 1976; Talmy 1976; McCawley 1976; Dixon 2000; Pylkkänen 2000; Haspelmath 2008; Schafer 2009), grammatical and semantic relations (Dowty 1972, 1979; Shibatani 1975, 1976, 2001; Comrie 1975, 1976, 1981; Foley and Van Valin 1984; Baker 1988; Levin and Hovav 1995,2009; Chierchia 2004; Harley 2008), and argument structure and mapping (Alsina 1992; Pesetsky 1995; Baker 1997; Pinker 1989; Dowty 1991; Grimshaw 1990; Reinhart 2002; Pylkkänen 2008; Schafer 2008; Harley 2013; Croft 2012). Comrie (1981: 158–177) characterizes causative events in terms of two (or more) micro-events, perceived as making up a macro-event, that are encoded in a

single expression (of varying size and form). He makes a three-way typological distinction between lexical, morphological, and analytical causatives. This three-way categorization forms a continuum from lexical to morphological and analytical, and corresponds to the continuum from indirect to direct causation. In other studies, causative constructions are classified according to a number of semantic contexts. Shibatani (1975) presents four pairs of semantic factors that determine the semantic types of causatives in specific contexts. These pairs are referred to as: Coercive vs Noncoercive, Directive vs Manipulative, Direct vs Indirect, Ballistic vs Controlled.

This paper provides an overview of causative constructions in Uyghur analyzed on the basis of a number of syntactic and semantic aspects such as suffixation rules and their restrictions, argument structures and their syntactic realization, case marking, and animacy restrictions of obligatory arguments. The paper is divided into three parts. Part one deals with distinctive semantic and syntactic properties of lexical causatives. In part two, the semantic features and syntactic manifestations of arguments of single causatives derived from intransitive and transitive verbs in Uyghur are examined in terms of argument coding, animacy properties of core arguments in causal chains, and the case marking of obligatory arguments. In part three, distinctive semantic properties and syntactic realizations of double causatives derived from transitive and intransitive bases through iteration of causative suffixes in modern Uyghur are briefly discussed. Finally, a brief comparison is made between lexical and morphological causative alternations in Uyghur and a general conclusion is drawn concerning the correlation between productive and nonproductive causative constructions. The data used in this study were taken from the Uyghur Language Database collected in the Laboratory of Xinjiang Multilingual Information Technology, founded at Xinjiang University in 2008, and from my Ph.D. thesis (Kurban forthcoming).

1. Lexical causatives

As in many other languages, lexical causative in Uyghur is manifested through a group of morphologically irregular, stem-specific, non-productive, transitive verbs with specific idiosyncratic meaning. These are transitive verbs which do not undergo morphological change or derivational processes when they denote a causative situation. Unlike productive causatives, they are not productive in formation, and semantically bear the causative meaning inherent to them. These are considered to be the most compact causatives by Dixon (2000: 74ff). For example, Uyghur transitive verbs like *qaz-* 'to dig', and *yirt-* 'to tear' are not formed through an overt derivational process; nevertheless they indicate a causal-resultant relationship in the causative event structure.

Hopper and Thompson (1980: 251) write: "Transitivity is understood to be about an activity being carried over or transferred from an agent to a patient. Transitivity in this traditional view necessarily involves at least two participants". Based on this prototype approach, the prototypical transitive event involves two core participants

which are conceptualized as agent (causer) and patient (causee) respectively. Hence, lexical causative verbs are dyadic (bivalent) transitive verbs and thus bear relationship with dominative agent and affected patient in their argument structure. "The notions 'agent' and 'patient' can be defined in terms of the starting point and endpoint respectively, of the prototypical event; they are natural delimiters: an event with an agent and patient is maximally delimited" (Croft 1994: 39).

(1) a. *Adil bügün čirayliq bir näqiš oy-di.*
Adil today beautiful a engraving make^{CAUS3}-DI.PST.3SG
'Adil made a beautiful engraving today.'

b. *Bala-lar därizä-ni čaq-iwät-ti.*
child-PL window-ACC broke^{CAUS}-ASP-DI.PST.3SG
'The children broke the window.'

As exemplified in (1a, b), the lexical causatives in Uyghur are expressed by a bare verb (a simple verb root) without any suffixation, yet by themselves they can indicate a causal-resultant event structure: the agent's activity (causing event) and the patient's change of state (resulting state) in mono-clausal structure.

The two core participants in lexical causatives display distinctive semantic properties. The prototypical lexical causative construction in Uyghur is characterized by an eventuality in which the agentive participant is often encoded as willful and animate entity which directly manipulates the patientive participant. In such cases, the volitional agent transfers physical energy directly and intentionally onto the patient which, in turn, undergoes a change of state. In this sense, the semantic role of the agent can be encoded as an "*Actor* that expresses the participant which performs, effects, instigates, or controls the situation denoted by the predicate", while the patient can be encoded as an "*Undergoer* that expresses the participant which does not perform, initiate, or control any situation, but is affected by it in some way" (Folley & Van Valin 1984: 29). What is indicated by (1a, b) is that there is a direct causal relationship between the causing and caused event in this mono-clausal event structure; i.e., the patient is intentionally manipulated by the direct presence of the agent without intervening of a third party in the causal event denoted by the underlying clause.

The subject choice in lexical (as well as morphological and periphrastic) causative sentences is determined by the respective roles of the arguments in the causal-resultative event structure. According to Comrie's Case Hierarchy (1981: 169; 1985: 337–340), the most agent-like argument, the Actor, which is an animate, high volitional and intentional entity, maps to nominative subject position, as exemplified in (2a, b), while the most affected argument, the Undergoer or patient, which is generally an inanimate entity, is often realized as object in the accusative; see as (2a, b).

(2) a. *Kona öy-ni buz-du-q.*
old house-ACC destroy^{CAUS}-DI.PAST-1PL
'We destroyed the old house.'

b. *Iščilar qädimiy buyum-lar-ni qaz-iwal-di.*
workers ancient relics-PL-ACC excavate^{CAUS}-ASP-DI.PST.3PL
'Workers excavated the ancient relics.'

A cross-linguistic generalization in the literature devoted to causative structure is that mono-clausal (lexical) causatives tend to express direct causation, while bi-clausal (periphrastic) causatives may express indirect causation (Shibatani 1976, 2001; McCawley 1978; Pinker 1989; Levin & Hovav 1995; 2009; Wolff 2003). Direct causation involves a (relatively) direct interaction between the subject and the object, which represent the initiator and the endpoint of the causal chain respectively, thus excluding any kind of intermediary factors. The complete physical involvement of the agent with the patient is viewed as the canonical example of direct causation. Shibatani (1976: 3) claims that "in both Japanese and Korean the productive causative expresses directive causation and the lexical causative manipulative causation". However, "when a lexical causative sentence expresses a situation that is associated with a conventionalized purpose, the sentence allows directive interpretation" (Shibatani 1976: 38). As indicated by (2a, b) above, the prototypical (transmission of force) relation between the participants encoded as subject and object in Uyghur lexical causatives is also expressed by direct manipulation of the patient by the agent without the intervention of any type of mediating entity in lexical causative construction. Wolff (2003) shows that speakers are more willing to use lexical causatives to describe causal chains yielding intended results than chains yielding non-intended results.

Unlike periphrastic causatives, lexical causatives in Uyghur are not productive and occur only in a particular class of verbs, namely, those denoting externally caused change of state. The verb (root) used in causatives is intrinsically transitive and indicates an externally caused event which involves the external causer of an action (the causer), and the recipient of the action (the causee), which are represented by the subject and object in causal-resultant event structure. Levin and Rappaport (1995) state that externally caused change-of-state events imply the existence of an external causer with immediate control. For instance, what is described by (1a, b) is that the agent's action (external force) affects the state of the patient. The agents, *Adil* and *balilar* 'children', are volitional animate entities, and the action denoted by the transitive verbs *oy-* 'to make' and *čaq-* ' to break' are directly completed by the external manipulation and control of *Adil* and *balilar* 'children' respectively in the causal event. "Direct causation is at issue when the agent controls the final result, whereas indirect causation holds when the agent controls the input situation but not all intervening stages" (Wunderlich 1997: 38). The agent's volition in the event

comes both from the context and our shared knowledge. In this sense, a lexical causative transitive verb denotes an externally caused eventuality.

2. Morphological causatives

Morphological causatives, universally considered one of the most productive strategies for forming causative constructions, result in the introduction of a new causer argument that is absent from the syntax of the non-causative counterpart. A distinctive characteristics of morphological causatives lies in their productivity (Comrie 1989, Dixon 2000, Shibatani 2001). In Uyghur morphological causatives, the causing event and the caused event are encoded in a mono-clausal structure by attaching a set of causative suffixes to an intransitive or transitive verb root along with appropriate argument adjustment. The specific condition for these causative suffixes are as follows.

(1) The causative suffixes *-dur/-tur/-dür/-tür* or *-ɣuz/ -quz/-güz/-küz* are attached to monosyllabic verb stems ending in a consonant (excluding those ending in *y* or in a vowel). They yield semantic interpretations of "permission" or "coercion" when they are attached to monosyllabic verb stems. The distinctive meaning conveyed by the causatives with these suffixes depends on the specific situation as well as the speaker's perspective. Therefore, they are partially interchangeable in most cases, but in some cases they cannot substitute for one-another. The choice of these two types of suffixes is made according to semantic and phonological properties of the given verb. Some verbs with transitive and intransitive properties (for example: *sun-* 'to snap', 'to stretch') place restrictions on the selection of these suffixes.

yaz- 'to write'	+ *-dur/-ɣuz*	'to make (sb) write'
bar- 'to go'	+ *-dur/-ɣuz*	'to make (sb) go'
maŋ- 'to move, go'	+ *-dur/-ɣuz*	'to make (sb) move/go'
yat- 'to lie down'	+ *-quz*	'to make (sb) lie down'
yä- 'to eat'	+ *-güz/-dür*	'to make (sb) eat'
qil- 'to do'	+ *-dur/-ɣuz*	'to make (sb) do'
siz- 'to draw'	+ *-dur/-ɣuz*	'to make (sb) draw'
tur- 'to stand'	+ *-ɣuz*	'to make (sb) stand'
eyt- 'to say'	+ *-quz*	'to make (sb) say'
bär- 'to give'	+ *-güz/-dür*	'to make (sb) give'
tap- 'to find'	+ *-tur/-quz*	'to make (sb) find'
täp- 'to kick'	+ *-küz/-tür*	'to make (sb) kick'
köm- 'to bury'	+ *-güz/-dür*	'to make (sb) bury'
öl- 'to die'	+ *-tür*	'to kill (sb) '
äm- 'to suck'	+ *-güz/-dür*	'to make (sb) suck'
käl- 'to come'	+ *-güz/-tür*	'to make (sb) come'
kät- 'to leave'	+ *-küz/-tür*	'to make (sb) leave'
or- 'to scythe'	+ *-dur/-ɣuz*	'to make (sb) scythe'

käs- 'to cut'	+ *-güz/-tür*	'to make (sb) cut'
tik- 'to sew'	+ *-küz/-tür*	'to make (sb) sew'
sun (v.t.)- 'to snap'	+ *-dur*	'to make (sb) snap'
sun (v.i.)- 'to stretch'	+ *-γuz*	'to make (sb) stretch'
qon- 'to stay over'	+ *-dur/-γuz*	'to make (sb) stay over'
min- 'to ride'	+ *-güz/-dür*	'to make (sb) ride'
bil- 'to know'	+ *-dür/ -güz*	'to make (sb) know'

(2) The suffix *-t* is attached to polysyllabic verb stems ending in *y*, *r* or a vowel.

taray- 'to become narrow'	+ *-t*	'to make narrow'
köpäy- 'to increase'	+ *-t*	'to make increase'
aqar- 'to become white'	+ *-t*	'to make white/whiten'
sämir- 'to become fat'	+ *-t*	'to make fat'
eri- 'to melt'	+ *-t*	'to make melt'
oyna- 'to play'	+ *-t*	'to make/let play'

(3) The suffixes *-ar/-är/* or *-ur/-ür* are attached to monosyllabic verb stems ending in *č* or *š*.

köč- 'to move'	+*-ür/-är*	'to make move/copy'
ič- 'to drink'	+*-ür/-är*	'to make drink'
uč- 'to fly'	+*-ur/-ar*	'to make fly'
piš- 'to ripe'	+*-ur/-ar*	'to make ripe'
čüš- 'to descend'	+*-ür/-är*	'to make descend'
aš- 'to increase'	+*-ur*	'to make increase'

(4) The suffixes *-it/-ut/-üt* are attached to some monosyllabic stems ending in *q* or *k*.

aq- 'to flow'	+ *-it*	'to make flow'
qorq- 'to fear'	+ *-ut*	'to make sb afraid of'
ürk- 'to startle'	+ *-üt*	'to cause to be startled'

These productive forms above are entirely regular—the choice of the suffixes is phonologically determined. Thus, the relevant suffixes are chosen according to the rule noted above, and removing the causative suffixes yields well-formed non-causative expressions. There are, however, certain forms in which this regularity is obscured.

kör- 'to see'	+ *-sæt*	'to show'
qayt- 'to go back'	+ *-ar/-ur*	'to return'
čiq- 'to go up'	+ *-ar*	'to cause to go up / come out of, to extract'

The causative forms of these verbs are unique and should be individually memorized. The forms involve suffixes, which can be easily segmented, and they qualify as morphological causatives, but they are irregular and functionally more similar to unanalyzable lexical causatives than to productive morphological forms. On the basis of the Japanese data discussed by Shibatani (1976a) Comrie (1981: 170) recognizes the possibility that certain non-productive morphological causatives may align with lexical causatives in their function.

In Uyghur, all transitive and intransitive verbs can undergo causative changes. These two types of verbs, however, give rise to different semantic and syntactic structures. Chomsky (1965: 189) maintains that a causative transformation can account for the derivation of a transitive verb from its intransitive one, as in *Jane dropped the pen*, which can be derived from *The pen dropped.* Transitive and intransitive verbs differ in terms of the number of arguments they require; i.e., intransitive verbs generally take a single argument which is encoded as subject, while transitive verbs require two arguments encoded as the subject and the object respectively. When an intransitive verb is causativized, the subject of the intransitive verb gets the accusative case in the object position whereas the newly introduced causer occupies the subject position and is in the nominative (3b, 4b).

(3) a. *Mašina maŋ-di.*
car move-DI.PST3SG
'The car moved.'

b. *Tursun mašina-ni maŋ-dur-di.*
Tursun car-ACC move-CAUS-DI.PST3SG
'Tursun made the car move.'

(4) a. *Bala yiɣla-di.*
child cry-DI.PST3SG
'The child cried.'

b. *Adil bala-ni yiɣla-t-ti.*
Adil child-ACC cry-CAUS-DI.PST3SG
'Adil made the child cry.'

(5) a. *Tursun saät-ni yasa-di.*
Tursun watch-ACC repair-DI.PST3SG
'Tursun repaired the watch.'

b. *Tursun-γa saät-ni yasa-t-ti-m.*
Tursun-DAT watch-ACC repair-CAUS-DI.PST-1SG
'I had Tursun repair the watch.'

When a transitive verb is causativized, the transitive verbs becomes ditransitive in (5b), and the newly introduced causer occupies the subject position and is in the nominative. The original object is maintained as an affected patient in the accusative (5b). Here I wish to draw attention to a distinction found in Uyghur between two types of morphological causatives, namely the single causatives derived with one causative suffix, and the double causatives derived with two causative suffixes.

2.1. Single causatives

Single causatives in Uyghur can be derived from both intransitive and transitive verbs. For convenience, I will call the former Type A, and the latter Type B.

Type A

Causativization is actually a transitivization process (Comrie 1981; Shibatani 1975, 2002; Alsina 1992; Dixon 2000 and others) by which an agent argument is introduced into the underlying intransitive clause producing a transitive clause. Morphological causatives in Uyghur, as in many other Turkic languages, are characterized by distinct morphological suffixation that can give rise to increased valence in the argument structure. In Type A causatives, the causative suffix not only introduces a new argument into the underlying clause by transforming the monovalent intransitive verbs into bivalent transitive verbs (Kratzer 1996; Marantz 1997; Pylkkanen 2008), it also changes the syntactic environment by placing certain requirements on the surrounding arguments. Thus, derived transitive verbs have two core (obligatory) arguments in the mono-clausal event denoted by the underlying clause. The two obligatory arguments of bivalent transitive verbs are encoded as causer and a causee in the mono-clausal event, and they display universal tendencies as well as language specific characteristics when they are syntactically realized. The syntactic realization of the argument structure in causative constructions can be explained with Comrie's Case Hierarchy (1975, 1976: 263) schematized in (6).

(6) subject > direct object > indirect object > oblique > genitive

Based on this universal reflecting a scale of thematic prominence, the logical subject role is assigned by default to the external argument (the most prominent or highest argument). Comrie (1985) holds the view that many causativized base-transitive

constructions mark the new argument as belonging to the leftmost available slot in the hierarchy. The marking of the new argument can thus, to a certain degree, be crosslinguistically predicted by the marking of the core arguments in the non-causativized counterpart of the same clause.

(7) a. *Bala-lar* *oyna-di.*
child-PL play-DI.PST3SG
'The children played.'

b. *Biz* *bala-lar-ni* *oyna-t-tu-q.*
we child-PL-ACC play-CAUS-DI.PST-1PL
'We had the children play.'

As seen in (7), the newly introduced external argument *biz* 'we' occupies the subject slot as causer (8b) in the derived structure, while the nominative-marked original subject argument *balilar* (7a) becomes the accusative-marked direct object (7b) in the derived structure of the bivalent causative verbs, resulting in a canonical transitive clause. The original subject argument must be demoted one step to the direct object position in the Case Hierarchy, since the subject position has already been occupied by the new causer.

In Type A causative constructions in Uyghur, both animate and inanimate entities can function as causer and causee, yielding distinctive semantic readings. The causer, being an external force or trigger, brings about the change in the state of the causee through direct or indirect participation in the event as a volitional entity. "The change of state in the patient follows directly and immediately from the action carried out by the agent in canonical transitive events" (Comrie 1989[1981]: 165). Under such an interpretation, the causer should at least be an animate entity with its own volition.

When both the causer and the causee are animate entities in Type A causatives, high intentional volition on the part of the causer and submissive volition as well as resistance on the part of the causee are implied. This yields two possible semantic interpretations:

(i) The causee possesses a degree of autonomy in carrying out the caused event; thus the causer's role is limited to supplying oral directions or instructions (indirect) to the causee (8a), rather than getting physically involved in the execution of the caused event. It can also imply that the caused event *saqla-* 'to wait' may be completed by the intervention of other means, such as conveying the message either by phone or through other persons. There may also be some spatial distance between the causer and the causee when the causer is giving orders or instructions. Such cases can be ascribed to Directive, Indirect Coercive and Ballistic causation (Shibatani 1975).

(ii) When the causer's influence on the causee is not merely restricted to oral directions or instructions, but includes active physical involvement in the caused event

(8b), the causer acts volitionally, either trying to get direct control over the causee (8c) or helping the patient to complete the causative event process (8d) in spite of the fact that the causee is also a volitional entity with submissive volition as well as resistance. Wolff (2003: 5) holds that direct causation is present between the causer and the final causee in a causal chain in the following cases:

(a) If there are no intermediate entities at the same level of granularity as either the initial causer or final causee.

(b) If any intermediate entities that are present can be construed as an enabling condition rather than an intervening causer. In this sense, this type of causative denotes Ballistic, Directive, Direct and Coercive/Permissive readings.

(8) a. *Sän biz ni bäk saqla-t-ti-ŋ.*
you we-ACC very wait-CAUS-DI.PST2SG
'You have kept us waiting long.'

b. *Bala-lar-ni oyna-t-tuq.*
child-PL-ACC play-CAUS-DI.PST1PL
'We had the children play.'

c. *Bala-lar yilan+ ni öl-tür-di.*
child-PL snake-ACC die-CAUS-DI.PST3SG
'The children killed the snake. '

d. *Sestra ayal-ni tugh-dur-di.*
nurse woman-ACC bear-CAUS-DI.PST3SG
'The nurse made the woman give birth (to her baby).'

It should be noted here that Type A causatives in Uyghur are ambiguous, allowing for a permissive and a coercive reading. It is apparent that different verbs give rise to different readings, and even the same verb might present different readings in different contexts. In this sense, (8a, c) indicate coercive while (8b, d) are either coercive or permissive depending on context as well as the speaker's perspective.

In cases where the causer is inanimate, causative verbs—mostly stative—involve a mental state or condition. In such cases, the causer does not have volition and can only trigger the causee by passive participation in the event by exerting influence in a direct or indirect manner without any intention and awareness. In this sense, the change of state of the causee might be conceived as the result of the causee's affectedness by the event. In other words, psychic verbs used in morphological causatives denote a sense of obligation or an affectedness imposed on the causee (Kayne 1994, Guasti 1996, Folli & Harley 2007). As indicated in (9a, b), the inanimate causer *iš* 'matter' and *muzika* 'music' stimulate the occurrence of a change in the mental state of the causees *bala* 'child' and *Alim*. In this sense, the influence of the causers *bu iš* and *muzika* can be conceptualized as the instigator of the change of state, whereas

the causees *Alim* and *bala* can be interpreted as the experiencers of the change of condition or state in the caused event. Besides, (90a) implies that the causee undergoes the change of mental state either by directly participating in the caused event or by other indirect means. Therefore, (9a) allows for direct or indirect readings depending on the implication denoted by the event in the given context. However, (9b) implies that the causee is directly affected by the causer without the intervention of other factors in the process of changing the state. So it can be conceived as a direct causation.

(9) a. *Bu iš Alim-ni bäk terik-tür-di.*
this matter Alim-ACC very angry-CAUS-DI.PST3SG
'This matter made Alim extremely angry.'

b. *Muzika axiri bala-ni uxla-t-ti.*
music at last child-ACC sleep-CAUS-DI.PST3SG
'The music made the child sleep at last.'

On the other hand, when the causee is an inanimate entity without any volition, the execution of the caused event is wholly dependent on the causer's active participation (10a, b) while the causee is characterized only by affectedness. However, the sentences have slightly different readings. In (10a), the causee is under the control of the causer when the event is viewed as a whole. It also implies either that the causer directly participates in the event or that other factors are involved, excluding the causer's direct involvement, although the causer has continuous control over the causee. By contrast, (10b) denotes that the causee is in a state of being accompanied by, or under continuous direct control of the causer from the beginning until the end. It also indicates the causee's physical manipulation of the causee, which is an inanimate entity without volition and resistance. In such non-volitional situation, the inanimate causee is unable to begin, continue or end the action or state of affairs. "Only the accusative coding of the causee is allowed" in describing such event (Kozinsky & Polinsky 1993: 202). In other words, the coding of the inanimate causee as an accusative direct object indicates that the inanimate causee's volition is completely suppressed such that this causee has no control over the situation, but only performs the action directed by the causer. In this sense, it is the causer, rather than the causee, who controls the situation. Hence, the causative with inanimate causee, illustrated by (9a, b), can be interpreted as a controlling and manipulative causation (Shibatani 1975).

(10) a. *Qar-ni eri-t-tu-q.*
snow-ACC melt-CAUS-DI.PST-1PL
'We made the snow melt.'

b. *Tursun mašina-ni maŋ-dur-di.*
Tursun car-ACC move-CAUS-PST3SG
'Tursun made the car move.'

There are also some Type A causatives in which both causer and causee are inanimate entities. In such cases, the causer's influence on the causee is rendered directly or indirectly regardless of causee's active or passive participation in the caused event (11a, b). Hence, in Uyghur the causer does not necessarily have to be an animate entity to be able to exert influence on the patient. On the contrary, it is common for an inanimate causer to bring about changes in the state of a causee, even though it does not have volition.

(11) a. *Ički uruš dölät-ni ajizla-t-iwät-ti.*
civil war country-ACC become weak-CAUS-ASP-DI.PST3SG
'Civil war weakened the county.'

b. *Qar-šiwirɣan hawa-ni xelila sowu-t-ti.*
snow-storm weather-ACC rather become cold-CAUS-PST3SG
'The snowstorm made the weather rather cold.'

Here the correlation and distinction between the morphological causative and inchoative verbs has to be mentioned. The relationship between the causative (transitive) and inchoative (intransitive) verbs can be understood from semantic and morphological perspectives. The derivational relationship between them has been controversial but it could be generalized as follows:

(i) Some argue that causative verbs are derived from inchoative verbs and use the term Causativization (Perlmutter 1978, Dowty 1979, Burzio 1986, Parsons 1990; Harley 2008; Marantz 2009).

(ii) Others hold that inchoative verbs are derived from causatives and call this Anticausativization or Decausativization (Levin & Hovav 1995: 108, Chierchia 2004, Reinhart 2002: 241, Koontz-Garboden 2009; Horvath & Siloni 2011).

(iii) Others still argue that each member is derived independently from a common stem, which is called Equipollence (Harley 1995, 2008, Pesetsky 1995: 70).

In Uyghur, as in other Turkic languages, the inchoative variant (with intransitive verbs) of the causative alternation is basic, and morphological causative variants (with transitive verbs) are derived from an inchoative counterpart which is the root form. Erdal (1991: 709) states that in Old Turkic "causatives formatives can, in principle, be added both to intransitive and to transitive bases". This also holds true for Mongolian causative and inchoative verb alternations "in which the causative member is marked and derived from the inchoative member" (Haspelmath 1993: 89). Here it should be noted that, in Uyghur, not only the causative form, but also other voice alternations are derived from intransitive or transitive verb root like *öl-* 'to die', *yaz-* 'to write'. In this sense, the assertion that the causative alternation

(transitive) is derived from its inchoative (intransitive) counterpart seems justified for Uyghur.

(12) a. *Yilan öl-di.* (*öl-* basic root form)
snake die-DI.PST3SG
'The snake died.'

b. *Yilan-ni öl-tür-dü-m.* (*öl-tür-* causative)
snake-ACC die-CAUS-DI.PST3SG
'I killed the snake.'

c. *Maqalä tünügün yaz-il-di.* (*yaz-il-* passive)
essay yesterday write-PASS-DI.PST3SG
'The essay was written yesterday.'

d. *U-lar maqalä yaz-iš-ti.* (*yaz-iš-* reciprocal)
He/she-PL essay write-REC-DI.PST3SG
'They wrote an essay together.'

(13) a. *Kemä čök-ti.*
boat sink-DI.PST3SG
'The boat sank.'

b. *Alim kemä-ni čök-tür-di.*
Alim boat-ACC sink-CAUS-DI.PST3SG
'Alim made the boat sink.'

(14) a. *Bolka piš-ti.*
bread be baked-DI.PST3SG
'The bread is baked.'

b. *Bolka-(ni) piš-ur-du-m.*
bread-(ACC) bake-CAUS-DI.PST3SG
'I baked (the) bread.'

The syntactic distinction between the causative and inchoative verb alternations in Uyghur is that the former are labeled with a morphological marker (12b, 13b, 14b) on the causative while the latter are characterized by a verb root without a morphological marker (13a, 14a, 15a). In other words, the inchoative form of a verb is often called the basic, unmarked voice form, while the causative is conceived of as the marked, derived voice form, due to its morphological complexity and the way it affects the argument structure of the verb by increasing it with a new obligatory argument. The semantic difference between the causative and inchoative verbs is

that the former involve a syntactically expressed controlling initiator (causer) as a subject in addition to a theme object (causee), while the latter do not involve an existentially bound external argument, i.e. lack a causer in their semantic representation.

Type B

Type B causatives in Uyghur also represent a mechanism to increase valency. Bivalent transitive verbs are converted into trivalent ditransitive verbs through causativization. Thus, ditransitive verbs involve three core (obligatory) arguments in the causal event denoted by the underlying clause. This clearly constitutes the main pattern of Type B causatives. The valence-increasing operation adds a causer argument which acts upon a causee to perform an event (Dixon 2000: 30). However, languages vary significantly in terms of the semantic role of arguments and syntactic realization (Dixon 2000: 62–74). Zubizarreta (1985) claims that the argument structure of a causative verb has three slots which are satisfied by the causer, the causee, and, possibly, the internal argument of the causativized predicate. Alsina (1992) holds that the three arguments saturating a causative verb are the causer, the causee, and the caused event, which in its turn may include an internal argument (IA). According to some other scholars (Burzio 1986, Guasti 1996, Folli and Harley 2007), however, a causative verb only has two argument slots which are satisfied by the causer and by the caused event which involves the causee and, eventually, IA. As for Turkic languages, Erdal (1991: 710) asserts that "causative verbs derived from transitive bases can allot three participant tasks: the Instigator, the Subject, and the (ultimate) Object".

The three obligatory arguments of trivalent ditransitive causative verbs in Uyghur are encoded as causer, causee, and patient in the embedded clause, and are similar to other languages with respect to the readjustment of the core arguments in ditransitive causative structure, but differ from them in terms of the syntactic realization of each obligatory argument as well as the semantic role of causee in particular. According to the Thematic Hierarchy (Fillmore 1968: 33; Bresnan & Kanerva 1989: 23–24), the most prominent (highest) argument is identified as the logical subject. The newly introduced argument (causer) is regarded as the most prominent argument in the Thematic Hierarchy; thus the logical subject role is naturally assigned to this highest argument by default. As also pointed out by Comrie (1989: 191) in his Case Hierarchy, the causee is assigned the leftmost available position. Since the subject and direct-object positions are already occupied in a causative based on a transitive verb, the causee occupies the indirect object position.

(15) a. *Tursun ayaq-ni yama-di.*
Tursun shoes-ACC mend-DI.PST3SG
'Tursun mended the shoes.'

b. *Tursun-γa* *ayaq-ni* *yama-t-ti-m.*
Tursun-DAT shoes-ACC mend-CAUS-DI.PST-1SG
'I had Tursun mend the shoes.'

As mentioned above, causativization usually signals some rearrangement of argument structure. The most usual scenario is for the causee and the patient argument to retain all the semantic properties that are found in the basic non-causative construction, while the causer is understood as an external instigator or bearer of the volitional component. As seen in (15), the newly introduced external argument *män* (covert in syntactic structure) occupies the subject slot as a causer in the derived argument structure, while the accusative-marked patient *ayaq* (16a) still retains its role as the patient and is assigned to the slot of the accusative-marked direct object (15b). The nominative-marked subject argument *Tursun* in the basic construction (15a) cannot rise to subject position since the latter has been occupied by the new causer, so it has to be demoted to the dative-marked indirect object position as causee in the derived argument structure of trivalent causative verbs (15b), which produces a canonical transitive clause.

The three core participants of the ditransitive causative exhibit distinctive semantic properties. The causer, an external force or trigger rather than the real performer of the action or event, promotes or stimulates an action or event, and produces changes in the state of the causee through indirect participation rather than physical involvement. However, in some instances the causer's direct involvement in the event can be found in Uyghur ditransitive causatives as well.

(16) a. *Oquγuči-lar* *muällim-gä* *hikayä* *eyt-quz-di.*
students-PL teacher-DAT story tell-CAUS-DI.PST3SG
'The students had the teacher tell stories.'

b. *Muällim* *bala-lar-γa* *resim* *siz-γuz-di.*
teacher child-PL-DAT picture draw-CAUS-PST3SG
'The teacher made the children draw pictures.'

c. *Aygül* *qončaq-qa* *kiyim* *kiy-güz-di.*
Aygül doll-DAT clothes put on-CAUS-PST3SG
' Aygül put the clothes on the doll.'

As illustrated in (16a), the first subevent is the causer's causal event, which is carried out by the causer directly in terms of verbal interaction; and the second subevent is the causee's performance of the caused event. In such cases, the causer is an initiator or trigger rather than the actual agent of caused event. It is the causee that functions as an actual agent. This is not always the case, however. In some cases, the causer either verbally or physically participates in the caused event, as in (16b). In such situations, (16b) might have two different semantic readings:

(i) The causer does not participate physically in the execution of the caused event; his role is confined to verbal instructions.

(ii) Not only the causee but also the causer are viewed as the agent of the caused event; i.e. it is the causer himself that shows the children how to draw. In (16c), the causer, rather than the causee, carries out the caused event, and has total control over the causee from the beginning to the end, because the causee is an inanimate entity without volition. In this sense, it can account for direct, manipulative and controlling causation.

The causee is always required to be animate in Type B constructions in Uyghur. For example, with an animate causer, transitive verbs such as *sal-* 'to build', *yasa-* 'to repair', *yaz-* 'to write', *eyt-* 'to tell/recite', *teri-* 'to till the land', *üz-* 'to pick' express (after their conversion into ditransitive verbs) a situation where the causer gets something done with a tangible beneficial effect. Thus, it can be viewed as a Benefactive (Babby 1993: 344) causative, in terms of the beneficiary of the consequences generated in the caused event. In this type of causatives, two conditions must be met so as to yield a benefactive reading:

(i) A strict animacy requirement is placed by the ditransitive verbs on the causer and causee who must possess intentional volition to be able either to perform an action or receive the benefit from the caused event (18a, b).

(ii) There should be a patientive object that is transferrable, either verbally (17b) or physically (17a), to the causer. By contrast, if the causer is an inanimate entity, it does not have any intention and ability to receive the benefit from the caused event, and the verb fails to be categorized as benefactive causative (17c).

(17) a. *Tursun-γa* *resim* *siz-γuz-du-m.*
Tursun-DAT picture draw-CAUS-DI.PST-1SG
'I made Tursun draw a picture.' (coercive, benefactive, transferrable physically)

b. *Tursun-γa* *öy* *sal-dur/ghuz-du-m.*
Tursun-DAT house build-CAUS-DI.PST-1SG
'I made Tursun build a house.' (coercive, benefactive)

c. *Kälkün* *dehqanlar-γa* *qaytidin* *öy* *sal-dur/γuz-di.*
flood farmer-PL-DAT again house build-CAUS-DI.PST3SG
'The flood made the farmers build the house again.'
(Coercive, not benefactive, due to inanimate causer.)

As seen in (17), the ditransitive causative in Uyghur, irrespective of the animacy of the causer, denotes a coercive situation by nature. In (17a, b), the animate causer with intentional volition exerts influence on the causee, by verbally or physically triggering the event. However, in (18c), the inanimate causer (a natural force) does not affect the causee as directly as the animate causer does. The only difference is

that the ditransitive causative with an animate causer expresses a direct causation, while that with an inanimate causer indicates indirect causation.

On the other hand, in Turkic and Mongolic languages the causer in ditransitive causatives tends to suffer a disadvantage rather than getting a benefaction—at least in Uyghur, though it might be unintentional in other Turkic languages, see Johanson (1998: 56)—as illustrated in (18a, b). Therefore, this type of causative can be accounted for semantically as a passive causative, in striking contrast to the benefactive causatives.

(18) a. *Čiš-im-ni* *tart-quz-du-m.*
teeth-POSS1SG-ACC pull-CAUS-DI.PST-1SG
'I had my teeth pulled.'

b. *Bu* *ayal* *bala-si-ni* *al-dur-iwät-ti.*
this woman child-POSS3SG-ACC take-CAUS-ASP-DI.PST3SG
'This woman had an abortion.'

The causee in Uyghur ditransitive causatives is always an animate entity that is the actual performer of the action in the caused event, though it is influenced by the causer in the execution process of the caused event. It is always marked with dative case in the syntactic structure, but it may be covert. Many languages place restrictions on the number of overt arguments in the syntactic realization, especially languages with ditransitive causatives. In this type of causative, one of the obligatory arguments (generally causee) of ditransitive verbs may remain covert in the syntactic realization. With regard to the suppression and extraction of the causee argument, Kozinsky and Polinsky (1993: 230) propose the following two regularities:

(19) a. The causee nominal can be extracted, passivized and/or suppressed only in the presence of the overt patient nominal.

b. The patient nominal can be extracted, passivized and/or suppressed only if the causee/recipient nominal is not overtly expressed.

This is particularly true for the languages like Uyghur in which ditransitive causatives are one of the prominent categories. The three core arguments of ditransitive verbs in Uyghur can be overtly mapped in a canonical syntactic structure. The causee, however, remains covert unless it is has to be overt. Even when it is omitted, its absence from the syntactic manifestation does not necessarily alter the meaning of the causative sentence because its identity is recoverable from the context. If it is necessary to express it overtly in the syntactic structure, then it is put in the dative.

(20) a. *Saät-im-ni* *yasa-t-ti-m.*
watch-POSS1SG-ACC repair-CAUS-DI.PST-1SG
'I had my watch repaired.'

b. *Tursun-γa* *saät-im-ni* *yasa-t-ti-m.*
Tursun-DAT watch-POSS1SG-ACC repair-CAUS-DI.PST-1SG
'I had Tursun repair the watch.'

As pointed out above, according to Comrie's *Case Hierarchy*, the causee is assigned to the indirect object position on the condition that the patient of the base construction still retains its position as direct object in the derived argument structure. In many languages, the causer mainly stimulates or triggers the causee through verbal commands or instructions—although there are some instances in which the causer gets partially physically involved in the caused event—but it is the causee, rather than causer, that is the actual executor of the action (under the influence of the causer) in the ditransitive causative construction. In this sense, the semantic role of the causee is similar to that of the agent to a greater extent, at least in Uyghur. The only difference between the semantic role of the causee in Type A and Type B causatives is that in the former, the causee, being a subject, is viewed as a high volitional agent with total control over the patient without the intervention of an external force in the performance of the caused event. In the latter, the causee is also considered a high volitional agent-like entity with flexible control over the patient, even if it is more or less influenced by an external force (causer) in the execution of the caused event. This means that the causee has the ability to control the action or state on his own by resisting the causer's instructions or demands in the given situation. Thus, the causee resembles the subject, rather than the object in terms of semantic function, and can be encoded as an agentive entity (Shibatani 1976: 33).

Being one of the obligatory argument in the ditransitive causatives, the patient (direct object) is usually an inanimate entity. As mentioned above, the causee can be left out of ditransitive causative constructions in Uyghur when it is retrievable from the context. In addition, even the causer can be covert in the causative structure. It can also be inferred from the personal suffix of predicate verbs, as the verbs always preserve agreement with the subject in the syntactic structures. The direct object is a prerequisite for the existence of the causal relationship between the causer and the causee. The canonical syntactic framework of ditransitive causative verbs in Uyghur could not be established without the participation of the patient in the causal event structure (21c). Therefore, the patient is viewed as an indispensable part of the ditransitive causatives in Uyghur. For example, the patient *maqale* 'essay' and *räsim* 'picture' in (21a, b) could not be omitted in the ditransitive causatives construction.

(21) a. *Muällim Tursun-γa* *maqalä* *yaz-dur-di.*
teacher Tursun-DAT essay write-CAUS-DI.PST3SG
'The teacher had Tursun write an essay.'

b. *Rässam-γa räsim-im-ni siz-γuz-du-m.*
painter-DAT picture-POSS1SG-ACC draw-CAUS-DI.PST-1SG
'I had the painter draw my picture.'

*c. *Muällim yaz-dur-di.* (incomplete)
teacher write-CAUS-PST.3SG
'The teacher had ??? write ???.'

d. *Räsim-ni män rässam-γa siz-γuz-du-m.*
picture-ACC I painter-DAT draw-CAUS-DI.PST-1SG

*e. *Räsim män rässam-γa siz-γuz-du-m.*
picture I painter-DAT draw-CAUS-DI.PST-1SG

The core arguments of ditransitive causative verbs in Uyghur take different forms when they are syntactically realized. The causer maps to the subject position in the nominative (21a, b), while the causee is always manifested in the dative (21a), and the direct object is realized in the accusative (21b). The direct object can only be moved to the head of the sentence when the accusative case marker is assigned and it is the focus of information (21d), and it is regarded as ungrammatical without the case marker (21e).

2. 2. Double causatives

In this section, I will briefly discuss double causative constructions, which have been dealt with in the literature (Comrie 1989, Zimmer 1976, Aissen 1979, Kulikov 1993, Kural 1996), and are referred to as "Second Causative" (Kulikov 1993: 121). Double causatives in Uyghur are formed by adding the causative suffixes *-GUz* to a stem which already has a causative suffix.

(22) a. *Alim yilan-ni öl-tür-di.*
Alim snake-ACC die-CAUS-DI.PST3SG
'Alim killed the snake.'

b. *Alim-γa yilan-ni öl-tür-güz-düm.*
Alim-DAT snake-ACC die-CAUS-CAUS-DI.PST-1SG
'I had Alim kill the snake.'

(23) a. *Tursun-γa saät-im-ni yasa-t-ti-m.*
Tursun-DAT watch-POSS-1SG-ACC repair-CAUS-DI.PST-1SG
'I had Tursun repair the watch.'

b. *Adil-γa* *däp* *Tursun-γa* *saät-im-ni*
Adil-DAT say-CONV Tursun-DAT watch-POSS-1SG-ACC
yasa-t-quz-du-m.
repair-CAUS-CAUS-DI.PST-1SG

c. *Adil-ni* *čaqirip* *Tursun-γa* *saät-im-ni*
Adil-ACC CALL-CONV Tursun-DAT watch-POSS-1SG-ACC
yasa-t-quz-du-m.
repair-CAUS-CAUS-DI.PST-1SG

d. *Adil-arqiliq* *Tursun-γa* *saät-im-ni*
Adil.NOM-POST Tursun-DAT watch-POSS-1SG-ACC
yasa-t-quz-du-m.
repair-CAUS-CAUS-DI.PST-1SG
'With the help of Adil, I had Tursun repair the watch.'

e. *Tursun-γa* *saät-im-ni* *yasa-t-quz-du-m.*
Tursun-DAT watch-POSS-1SG-ACC repair-CAUS-CAUS-DI.PST-1SG
'I had Tursun repair the watch.'

*f. *Adil-γa* *Tursun-γa* *saät-im-ni* *yasa-t-quz-du-m.*
Adil-DAT Tursun-DAT watch-POSS-1SG-ACC repair-CAUS-CAUS-DI.PST-1SG
'I made Adil make Tursun repair the watch.'

As seen in (22, 23), the double causatives derived from an intransitive or transitive base in Uyghur result in different surface realizations. Subject, direct and indirect objects cannot be doubled. In (22b), there is only one causee argument marked with the dative. In contrast, in (23b, c, d), there are apparently two causees in the surface structure. Since the direct and indirect object positions are already occupied and marked with the accusative and the dative, the second causee is either demoted to oblique with a postposition (23d), is included in a construction based on a converb (23b, c) or is omitted (23e). Therefore, (23f) is not acceptable due to the occurrence of two dative causees in the same sentence.

In morphological causatives, the causative morphemes are iterative, so double (or even triple) causatives are licensed in many languages. For example, Turkish has been claimed to permit up to three iterations (Çetinoğlu et al. 2009) or to have no upper limit (Kural 1996). Kulikov (1993:124) also states that double (or triple) causatives are frequent in Turkish. In Uyghur, however, both intransitive and transitive bases can be causativized only twice; that is, the limit on iteration of causative morpheme is two (23b, 24b, c, d, e). In this respect, Uyghur is different from genetically related languages, for example Turkish.

Causative suffixes in the double causative construction in Uyghur are subject to strict ordering. The causative suffix in *-DIR* obligatorily precedes the causative suffix *-GUz;* see (24b). The reverse is not acceptable (24c).

(24) a. *Rässam-γa* *räsim* *siz-dur/-γuz-du-m.*
painter-DAT picture draw-CAUS-PST.1SG
'I had the painter draw a picture.'

b. *Rässam-γa* *räsim* *siz-dur-γuz-du-m.*
painter-DAT picture draw-CAUS-CAUS -PST.1SG
'I made (someone made) the painter draw the picture.'

*c. *Rässam-γa* *räsim* *siz-γuz-dur-du-m.*
painter-DAT picture draw-CAUS-CAUS -PST.1SG

The coercive reading is a common prototypical interpretation of double causatives in many languages. Zimmer (1976: 412) points out that Turkish double causatives denote a single act of causation with emphasis on its forcefulness. The double causative suffixes in Uyghur can optionally or obligatorily be deleted without any change in meaning when the causative is derived from a transitive base. In such cases, the single causative morpheme may be used instead of the double causative morpheme without any change in causative meaning (25b). In this sense, the double causative derived from a transitive base may yield an intensified reading. In the case of an intransitive base, however, the deletion of double causative suffixes gives rise to the omission of the causee argument in surface structure (26b). Therefore, the deletion of double causative suffixes in Uyghur results in distinctive semantic readings in the surface structure of causative constructions derived from transitive and intransitive bases respectively.

(25) a. *Alim-γa* *mašina-m-ni* *yasa-t-quz-du-m.*
Alim-DAT car-POSS-1SG-ACC repair-CAUS-CAUS-DI.PST-1SG

b. *Alim-γa* *mašina-m-ni* *yasa-t-ti-m.*
Alim-DAT car-POSS-1SG-ACC repair-CAUS-DI.PST-1SG
'I had Alim repair the watch.'

(26) a. *Alim-γa* *bolka* *piš-ur-γuz-du-m.*
Alim-DAT bread.NOM bake-CAUS-CAUS-DI.PST-1SG
'I had Alim bake the bread.'

b. *Bolka* *piš-ur-du-m*
bread.NOM bake-CAUS-DI.PST-1SG
'I baked the bread.'

On the contrary, in Uyghur causative suffixes may be iterated without adding further cause events or arguments to the single causative constructions when the double causative is derived from transitive base (25a). In such cases, an increase in morphological complexity does not necessarily mean an increase in semantic complexity. Thus, there is no one-to-one correlation between the number of suffixes and the number of events. In this sense, the double causative construction derived from a transitive base in Uyghur may produce vacuous semantic readings. However, in the case of a double causative based on an intransitive base, increasing the number of causative suffixes definitely leads to an increase in the number of cause events or arguments (26a).

Double causatives in Uyghur may also have a permissive reading in addition to the coercive and intensive readings mentioned above. The permissive reading is not as straightforward and obvious as the coercive and intensive readings indicated in the double causative construction derived from both intransitive (27a) and transitive (27b) bases. The distinction between the permissive and the intensive reading derived from the double causative construction usually depends on the specific context and the speaker's perspective.

(27) a. *Alim-γa* *bala-lar-ni* *oyna-t-quz-du-m.*
Alim-DAT child-PL-ACC play-CAUS-CAUS-DI.PST-1SG
'I caused Alim to let the children play.'

b. *Bala-lar-γa* *räsim* *siz-dur-γuz-du-m.*
Child-PLU-DAT picture draw-CAUS-DI.PST-1SG
'I caused (someone) to let the children draw the picture.'

3. Conclusion

The present article has attempted to provide a descriptive analysis of lexical and morphological causative constructions employed in modern Uyghur from semantic, syntactic and diathetic perspectives.

Lexical causatives in Uyghur are manifested through a group of morphologically irregular, stem-specific, non-productive, bivalent (dyadic) transitive verbs without particular morphological markers. Lexical causatives in Uyghur involve two core participants in their argument structures, and which are encoded as the external causer and the internal causee in the causative event. The external causer invariably occupies nominative subject position while the internal causee is realized as an accusative-marked object in canonical lexical causatives in modern Uyghur.

Morphological causatives are undoubtedly the most frequent means of expressing the causative in Uyghur, and their high productivity is characterized by almost unrestricted derivation from both transitive and intransitive verbs via distinctive morphological suffixation, which in turn leads to an increase in valence-number by converting monovalent (monadic) intransitive verbs into bivalent (dyadic) transitive verbs, and bivalent transitive verbs into trivalent (triadic) ditransitive verbs in the

argument structure. Both animate and inanimate entities can function as causer and causee in the single causative construction, but they produce distinct semantic readings. The morphological causatives derived from intransitive verbs in Uyghur are ambiguous, allowing for a permissive or a coercive reading depending on the semantic and syntactic properties of the causative verb. The striking characteristics of morphological causatives derived from transitive verbs in Uyghur lie in the covert manifestation of the causee in specific situations. Even it remains covert in a particular situation, its absence from the syntactic manifestation does not necessarily alter the meaning of a causative sentence, as its identity can be recovered from the context.

Double causatives derived from intransitive and transitive bases in Uyghur result in different surface realizations. Both intransitive and transitive base can be causativized twice at most; that is, no more than two different causative morphemes can be iterated in a causative construction. Thus Uyghur does not permit any triple causative. In this sense, Uyghur is different from that of certain genetically related languages (for example Turkish). Besides, there is strict restriction as to the suffixation ordering of double causatives in Uyghur. The deletion of double causative suffixes in Uyghur results in distinct semantic readings in the surface structure of causative constructions derived from transitive and intransitive bases respectively.

Acknowledgements

This research was supported by grants for scientific research from (1) National Social Science Foundation of China (Project No: 12CYY063); (2) Humanities and Social Sciences Foundation of the Ministry of Education (Project No: 10XJJC 740003); (3) Xinjiang University Doctoral Science Foundation (Project No: 07020 428065)

This paper was originally drafted and completed while I was a visiting scholar at the Center for Languages of the Central Asian Region (CeLCAR) at Indiana University (February 2015, February 2016). I am grateful to Öner Özçelik for facilitating my visits to CeLCAR. I wish to thank anonymous reviewers for their comments and valuable suggestions. I am also thankful to Turghun Ibrahim of Laboratory of Xinjiang Multilingual Information Technology for providing me with relevant linguistic data.

Abbreviations

AC	actor	NUM	number
ACC	accusative	OBJ	object
AD	adverb	PASS	passive
ADJ	adjective	PL	plural
CAUS	causative;	POSS	possessive
^{CAUS}	lexical causative	POST	postposition
DAT	dative	PST	past

DEF	definite	REC	reciprocal
DO	direct object	SG	singular
GEN	genitive	SUBJ	subject
IA	internal argument	UN	undergoer
NEG	negative	V	verb
NOM	nominative	VN	verbal noun

References

Aissen, J. 1979. *The syntax of causative constructions*. New York: Garland.

Alsina, A. 1992. On the argument structure of causatives. *Linguistic Inquiry* 23, 517–555.

Alsina, A. 1997. A theory of complicate predicates: Evidence from causatives in Bantu and Romance. In: Alsina, A. & Bresnan, J. & Sells, P. (eds.) *Complex predicates.* Stanford: CSLI Lecture Notes 64, 203–246.

Babby, L. H. 1993. Hybrid causative constructions: Benefactive causative and adversity passive. In: Comrie, B. & Polinsky, M. (eds.) *Causatives and transitivity.* Amsterdam: Benjamins. 343–367.

Baker, M. 1988. *Incorporation: A theory of grammatical function changing.* Chicago: Chicago University Press.

Baker, M. 1997. Thematic roles and syntactic structure. In: Haegeman, L. (ed.) *Elements of grammar: Handbook of generative syntax.* Dordrecht: Kluwer. 73–137.

Bresnan, J. & Kanerva, J. M. 1989. Locative inversion in Chichewa: A case study of factorization in grammar. *Linguistic Inquiry* 20: 1, 1–50.

Burzio, L. 1986. *Italian syntax: A government and binding approach*. Dordrecht: Reidel.

Çetinoğlu, Ö. & Butt, M. & Oflazer, K. 2009. Mono/biclausality of Turkish causatives. In: Ay, S. & Aydın, Ö. & Ergenç, İ. & Gökmen, S. & İşsever, S. & Peçenek, D. (eds.) *Essays on Turkish Linguistics.* (Turcologica 79.) Wiesbaden: Harrassowitz Verlag. 43–52.

Chomsky, N. 1965. *Aspects of the theory of syntax*. Mass: MIT Press.

Chierchia, G. 2004. A semantics for unaccusatives and its syntactic consequences. In: Alexiadou, A. & Anagnostopoulou, E. & Everaert, M. (eds.) *The unaccusativity puzzle: Explorations of the syntax-lexicon interface*. Oxford: Oxford University Press. 22–59.

Comrie, B. 1975. Causative and universal grammar. *Transactions of the Philological Society* 1974, 1–32.

Comrie, B. 1976. The syntax of causative constructions: Cross-language similarities and divergences. In: Shibatani, M. (ed.) *The grammar of causative constructions*. New York: Academic Press. 261–312.

Comrie, B. 1981. *Language universals and linguistic typology: Syntax and morphology*. Chicago: University of Chicago Press. 158–177.

Comrie, B. 1985. Causative verb formation and other verb-deriving morphology. In: Shopen, T. (ed.) *Language typology and syntactic description* 3. Cambridge: Cambridge University Press. 309–348.

Comrie, B. 1989[2]. *Language universals and linguistics typology*. Chicago: Chicago University Press.

Comrie, B. 2006. Transitivity pairs, markedness and diachronic stability. *Linguistics* 44, 303–318.

Comrie, B. & Polinsky, M. 1993. *Causatives and transitivity*. Amsterdam: John Benjamins.
Croft, W. 1994. The semantics of subjecthood. In: Yaguello, M. (ed.) *Subjecthood and subjectivity: The status of the subject in linguistic theory*. Paris: Ophrys. 29–75.
Croft, W. 2012. *Verbs: Aspect and argument structure*. Oxford: Oxford University Press.
Dixon, R. M. W. 2000. A typology of causatives: form, syntax and meaning. In: Dixon, R.M.W. & Aikhenvald, A. (eds.) *Changing valency: Case studies in transitivity.* Cambridge: Cambridge University Press. 30–83.
Dowty D. 1979. *Word Meaning and Montague Grammar: The semantics of verbs and times in generative semantics and in Montagues PTQ.* Dordrecht: Reidel.
Dowty D. 1991. Thematic proto-roles and argument selection. *Language* 67, 547–619.
Erdal, M. 1991. *Old Turkic word formation: A functional approach to lexicon.* Wiesbaden: Otto Harrassowitz.
Fillmore, C. J. 1968. The case for case. In: Bach, E. & Harms, R. T. (eds.) *Universals in linguistic theory*. New York: Holt, Rinehart and Winston. 1–88.
Fillmore, C. J. 1977. The case for case reopened. In: Peter C. & Jerrold M. S. (eds.) *Syntax and Semantics: Grammatical relations 8,* New York, San Francisco & London: Academic Press. 59–81
Foley, W.A. & Valin, V. 1984. *Functional syntax and universal grammar.* Cambridge: Cambridge University Press.
Folli, R. & Harley, H. 2007. Causation, obligation, and argument structure: On the nature of little *v*. *Linguistic Inquiry* 38, 197–238.
Givon, T. 1980. The binding hierarchy and the typology of complements. *Studies in Language* 4, 333–377.
Grimshaw, J. 1990. *Argument structure*. Cambridge. Mass.: MIT Press.
Guasti, M. 1996. Semantic restrictions in Romance causatives and the incorporation approach. *Linguistic Inquiry* 27, 294–313.
Haiman, J. 1985. *Natural syntax.* Cambridge: Cambridge University Press.
Harley, H. 1995. Subjects, events, and licensing. Ph.D thesis, MIT.
Harley, H. 2008. On the causative construction. In: Miyagawa S. & Saito M. (eds.) *Oxford handbook of Japanese linguistics*. Oxford: Oxford University Press. 20–53.
Harley, H. 2013. External arguments and the Mirror Principle: On the distinctness of voice and *v*. *Lingua* 125, 34–57.
Haspelmath, M. 1993. More on the typology of inchoative/causative verb alternations. In: Comrie, B. & Polinsky, M. (eds.) *Causatives and transitivity.* Amsterdam: Benjamins. 87–120.
Haspelmath, M. 2008. Causatives and anticausatives. Handout for the Leipzig Spring School on Linguistic Diversity.
Hopper, P. J. & Thompson S. A. 1980. Transitivity in grammar and discourse. *Language* 56, 251–299.
Horvath, J. & Siloni, T. 2011. Causatives across components. *Natural Language & Linguistic Theory* 29: 3, 657–704.
Hovav, M. R. & Levin, B. 2012. Lexicon uniformity and the causative alternation. In: Martin E. & Marelj, M. & Siloni, T. (eds.) *The theta system: Argument structure at the interface*. Oxford: Oxford University Press. 150–176.
Johanson, L. 1998. The structure of Turkic. In: Johanson, L. & Csató, É. Á. (eds.) *The Turkic languages.* London: Routledge. 30–66.

Kayne, R. 1994. *The antisymmetry of syntax.* Cambridge, Mass.: MIT Press.
Koontz-Garboden, A. 2009. Anticausativization. *Natural Language and Linguistic Theory* 27, 77–138.
Kozinsky, I. & Polinsky, M. 1993. Causee and patient in the accusative of transitive: Coding conflict or doubling of grammatical relations. In: Comrie, B. & Polinsky, M. (eds.) *Causatives and transitivity.* Amsterdam: Benjamins. 177–240.
Kratzer, A. 1996. Severing the external argument from its verb. In: Rooryck, J. & Zaring, L. (eds.) *Phrase structure and the lexicon*. Dordrecht: Kluwer Academic. 109–137.
Kural, M. 1996. *Verb incorporation and elementary predicates*. Doctoral dissertation. UCLA.
Kurban, U. (forthcoming) A contrastive study of verbal valence in Chinese and Uyghur. (PhD Dissertation) College of Humanities, Xinjiang University, Urumqi.
Levin, B. & Hovav, M. R. 1994. A preliminary analysis of causative verbs in English. *Lingua* 92, 35–77.
Levin, B. & Hovav, M. R. 1995. *Unaccusativity: At the syntax-lexical interface*. Cambridge, Mass.: MIT Press.
Levin, B. & Hovav, M. R. 2005. *Argument realization*. Cambridge: Cambridge University Press.
Levin, B. 2009. Further explorations of the landscape of causation: Comments on the paper by Alexiadou and Anagnostopoulou. *MIT Working Papers in Linguistics* 49, 239–266.
Lewis D. 2004. Causation as influence. In: Collins, J. & Hall, N. & Paul, L. (eds.) *Causation and counterfactuals*. Cambridge: Cambridge University Press. 75–106.
Marantz, A. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. In: Dimitriadis, A. et al. (eds.) *Proceedings of the 21st Annual Penn Linguistics Colloquium. Working Papers in Linguistics* 4: 2, 201–25.
Marantz, A. 2009. Root, *re-,* and affected agents: can roots pull the agent under little v? [Paper presented at Roots 2 Workshop, Stuttgart].
McCawley, J. 1976. Remarks on what can cause what. In: Shibatani, M. (ed.) *The grammar of causative constructions*, New York: Academic Press. 117–129.
McCawley, J. 1978. Conversational implicature and the lexicon. In: Cole, P. (ed.) *Pragmatics*. New York: Academic Press. 245–59.
Nedjalkov, V. P. & Silnitsky, G. G. 1973. The typology of morphological and lexical causatives. In: Kiefer, F. (ed.) *Trends in Soviet theoretical linguistics*. Dordrecht: Reidel. l–32.
Parsons, T. 1990. *Events in the semantics of English: A study in subatomic semantics*. Cambridge, Mass.: MIT Press.
Perlmutter, D. 1978. Impersonal passives and the Unaccusative Hypothesis. In: *Procedings of the Fourth Annual Meeting of the Berkeley Linguistics Society,* 157–189.
Pesetsky, D. 1995. *Zero syntax.* Cambridge, Mass.: MIT Press.
Pinker, S. 1989. *Learnability and cognition: The acquisition of argument structure.* Cambridge, Mass.: MIT Press.
Pylkkänen, L. 2000. On stativity and causation. In: Tenny, C. & Pustejovsky, J. (eds.) *Events as grammatical objects.* Stanford: CSLI Lecture Notes 100, 417–445.
Pylkkänen, L. 2008. *Introducing arguments*. Cambridge, Mass.: MIT Press.
Reinhart, T. 2002. The theta system: An overview. *Theoretical Linguistics* 28: 3, 229–290.
Schäfer, F. 2008. *The syntax of (anti-)causatives. External arguments in change-of-state contexts.* Amsterdam: Benjamins.
Schäfer, F. 2009. The causative alternation. *Language and Linguistic Compass* 3: 2, 641–681.

Shibatani, M. 1973a. A linguistics study of causative constructions. P.hD thesis, University of California, Berkeley.

Shibatani, M. 1973b. Semantics of Japanese causativization. *Foundations of Language* 9, 327–373.

Shibatani, M. 1975. *A linguistic study of causative constructions*. Bloomington*:* Indiana University Press.

Shibatani, M. 1976. The grammar of causative constructions: a conspectus. In: Shibatani M. (ed.) *The grammar of causative constructions.* New York: Academic Press. 1–42.

Shibatani, M. & Pardeshi, P. 2001. The causative continuum. *Kobe Paper in Linguistics* 3, 136–177.

Song, J. J. (ed.) 1996. *Causatives and causation: A universal-typological perspective.* London: Longman.

Talmy, L. 1976. Semantic causative types. In: Shibatani, M. (ed.) *The grammar of causative constructions*. New York: Academic Press. 43–116.

Talmy, L. 1988. Force dynamics in language and cognition. *Cognitive Science* 12, 49–100.

Talmy, L. 2000. *Toward a cognitive semantics* 1–2. Cambridge, Mass.: MIT Press.

Wolff, P. 2003. Direct causation in the linguistic coding and individuation of causal events. *Cognition* 88, 1–48.

Wunderlich, D. 1997. Cause and the structure of verbs. *Linguistic Inquiry* 28: 1, 27–68.

Zimmer, K. 1976. Some constraints on Turkish causativization. In: Shibatani, M. (ed.) *The grammar of causative constructions*. New York: Academic Press. 399–412

Zubizarreta, M. L. 1985. The relation between morphophonology and morphosyntax: The case of Romance causatives. *Linguistic Inquiry* 16, 247–289.

Extensive borrowing of reindeer terminology in north-eastern Siberia

Peter Sauli Piispanen

Piispanen, Peter Sauli 2015. Extensive borrowing of reindeer terminology in north-eastern Siberia. *Turkic Languages* 19, 240–258.

This paper presents borrowings, mostly related to reindeer husbandry in the far north-eastern Siberian area, between several non-genealogically affiliated languages. The semantics, phonology and chronology are discussed. The following are etymologized as Ewenki borrowings: Yukaghiric *moll'e* 'small wild reindeer', *oŋul* 'reindeer', *šaqala* 'fox', *ugur* 'spine' and *joɣul* 'nose'; and (Pre-)Yakut borrowings: Yukaghiric *saa-laaɣare* 'south, lit. tree left', *saaɣare* 'left side of a yurt; West', *šajɣər* 'aside'. A possible Turkic borrowing is found with (Proto-)Turkic *qan 'blood' > Proto-Samoyed *kẹm 'blood'. Lastly, Uralic cognates or borrowings in Yukaghiric consist of *kedie-* 'obstinate (of a tied reindeer)', *petčigije* 'reins', *a:čə* 'domestic reindeer' and *sierdiid-ile* 'reindeer not selected for slaughter'.

Peter Sauli Piispanen, Department of Slavic and Baltic Languages, Finnish, German and Dutch, Stockholm University, 106 91 Stockholm, Sweden.
E-mail: peter.piispanen@finska.su.se

Introduction

The languages of reindeer herders

Reindeer husbandry is of great economic and cultural importance for many of the populations residing in north-eastern Siberia. The herding groups periodically travel very long distances, at least did so historically, during which they come into contact with populations speaking other genealogically or non-genealogically related languages. As a consequence of such language contacts, many groups speak not only their own native language, and very often Russian, as well as at least one more language, easily making them bi-, tri- or multilingual. Since many of the (encountered) populations also deal with reindeer herding and the reindeer industry, extensive interlingual borrowings related to reindeer terminology are quite expected. Over the centuries several languages have served as *lingua franca* for reindeer herders—an early example being Chukchi (used in Chukotka until the beginning of the 20th century by Russians, Chukchi, Ewen, Yukaghirs, Koryak and Yup'ik; Krupnik (1993))—and the most recent probably being Yakut (used along the trail from

Dudinka to Khatanga);[1] Yakut is still used (in the Sakha Republic) by many Dolgans, Ewenkis, Ewens and Yukaghirs as a *lingua franca*.[2]

Populations involved in the reindeer economy

To the traditional populations engaged in reindeer husbandry over a very large geographical area belong the Ewen and Ewenki (Northern Tungusic languages), Yakut and Dolgan (Turkic languages), Chukchi and Koryak (Chukotko-Kamchatkan languages), Yukaghir (in particular Kolyma Yukaghir (KY) and Tundra Yuka-ghir (TY); Uralo-Yukaghiric languages),[3] Ket (a Yeniseian language), Yup'ik (Eskimo-Aleut), some Samoyed groups (Uralic languages), etc. Another such group is the Chuvantsi, originating in the easternmost Yukaghir populations, who speak both Koryak and Chukchi (Gurvič 1982); historically they also spoke the now extinct Yukaghir language of the Chuvantsi before being assimilated by the Koryak. The Koryak were very often represented as invading forces by the Yukaghirs (Bogoras 2009); this historical aspect may explain the cultural resistance among Yukaghirs to borrowing Koryak vocabulary, and indeed very few such borrowings are known. Relations with the Yakut were traditionally much better, and there are numerous Yakut as well as Tungusic borrowings in Yukaghiric.

The Yukaghirs

Taking the Yukaghirs as a further example, numerous known borrowings are at least peripherally related to reindeer terminology (summarized and referenced in Nikolaeva, I. 2006) in the Yukaghiric languages, with the donor languages being Yakut

1 Yakut also developed into the creole language of the Dolgan a few centuries ago (by mixing Yakut grammar and vocabulary with Ewenki and Russian vocabulary). Even today the Dolgans remain nomadic reindeer herders and hunters in the distant Taymyr Peninsula.

2 I am grateful to the anonymous reviewers for their valuable comments during manuscript preparation.

3 The question of Yukaghiric genealogic affiliation is a controversial and there are two prevalent schools of thought on the question. The first holds that the relationship between the Uralic and Yukaghiric languages is one of borrowings only, and the second that the two are actually genealogically affiliated language families going back to a common language referred to by different authors as Proto-Uralo-Yukaghiric, Proto-Sibero-Uralic (or Uralo-Siberian), or Pre-Proto-Uralic (Pre-PU), the last of these being my personal preference. I further assume that Pre-PU is phonologically, morphologically and lexically very close to, and immediately precedes the linguistic stage of Early Proto-Yukaghir, aka EY (> Middle Proto-Yukaghir, aka MY > Late Proto-Yukaghir, aka PY > Kolyma and Tundra Yukaghir, etc.). The question of genealogic affiliation, the background of Yukaghir studies, and previous research in the field were most recently summarized and discussed by Piispanen (2013b; 2015, 2016) and Aikio (2015), who are proponents of the two different schools of thought.

(2),[4] Ewen/Ewenki < Tungusic (19),[5] Chukchi (1),[6] Koryak (1),[7] etc. The importance of reindeer terminology among the Yukaghirs is also evident in a related suffixation system: KD *-c-*, *-rej-* 'suffix to get a reindeer-related verb' and TY *-aa* 'suffix to get a reindeer-related noun'.

This paper focuses in particular on borrowings into Yukaghiric. Many Ewen ~ Ewenki (=Tungusic) and Yakut borrowings into Yukaghiric were summarized in the *Comparative Dictionary of the Tungus-Manchu Languages* (TMS 1 & 2), and are

4 These are: 1. Yakut *sanajaq* 'fur coat', borrowed as: KY *šaja:q* 'coat with fur outside made of a large reindeer skin', and 2. Yakut *u:čaq* 'saddle-reindeer', borrowed as: KY *u:čaŋ*, *u:čaq* 'saddle-reindeer'. Clearly, the Yukaghirs regarded the Yakut mainly as reindeer herders; there is an old Kolyma Yukaghir word, KD *ilbied'i* 'Yakut', which was derived directly from KD *ilbe* 'domestic reindeer'.

5 These are: 1. Ewen *amarka:n* 'four- or five-year-old male reindeer', borrowed as: TY *amarkanel* 'five-year-old male reindeer'; 2. Ewen *a:w-a:w* 'imitation of the noise made by a reindeer calf', borrowed as: TY *awńe-* 'to make noise (of a reindeer calf)'; 3. TU ***čur-**, ***čir-** 'three-year-old wild deer', borrowed as: PY ***čurqə** > TY *čurγa* 'two-year-old female reindeer', *čurγan-purewre* 'three-year-old female reindeer, lit. over two-year-old female reindeer'; 4. Ewen *e:ni* 'one-year-old female reindeer with a calf', borrowed as: TY *eenil* 'one-year-old female reindeer'; 5. Ewen *i:te:nken* 'three-year-old wild bull reindeer', borrowed as: TY *iiteenken*, KD *itenken* 'three-year-old wild bull reindeer', TY *iiteelnel* 'three-year-old male reindeer'; 6. Ewen *ketem* 'barren (of a female reindeer)', borrowed as: TY *ketemel* 'barren reindeer heifer'; 7. TU ***lökü-** 'elk, deer', borrowed as: PY ***löγ-** > TY *lögur* 'call of reindeer male' > MC *logu* 'reindeer'; 8. Ewenki *melele*, *me:lan* 'elk calf, small tundra reindeer', borrowed as: PY ***mejl'inč'ə**; J > KY *mejl'id'ə* 'one-year-old reindeer or elk'; 9. Ewen *mo:mina* 'reindeer intestine filled with lard', borrowed as: PY ***mo:mina** > KY *mo:mina*: 'large intestine', TY *momneń* 'part of a reindeer intestine'; 10. Ewenki *na:ra:*, *nara* 'castrated reindeer', borrowed as: TY *naareŋol-* 'half-castrated'; 11. TU ***ńar-gu-** 'new, fresh', borrowed as: PY ***ńarqə** > TY *ńarqa-jewlid'e* 'new-born reindeer'; 12. Ewen *ńo:rkan*, borrowed as: TY *ńuorkanal* 'four- or five-year-old male reindeer'; 13. ?NT ***awa-la** 'reindeer', borrowed as: ?PY ***o:wə** > MU *end-schdsche-ówa* 'elk'; 14. TU ***kula** 'light brown with black mane and tail (of a horse)', borrowed as: PY ***qulu-** > TY *quluruo-* 'to have hair that has white tips with a dark tinge (of a reindeer)', TK *quluruod'ed-ile* 'white reindeer with a grayish hue'; 15. Ewen *ke:nde* 'draught reindeer of the Chukchi or Koryak breed', borrowed as: PY ***qunde**: > TY *qundietege* 'draught reindeer more than six years-old'; 16. Ewen *kuńa:-* 'to gallop (of a reindeer or a horse)', borrowed as: PY ***quńe** > TY *quńe* 'two-year-old male reindeer'; 17. Ewen *tambaka*, borrowed as: TY *tambakaa* 'Chukchi child's overall made of reindeer skin', 18. NT ***ire** 'male elk, wild reindeer, smoked reindeer skin', borrowed as: PY ***yra-** > TY *iraŋal* 'light brown reindeer', *irul* 'biggest wild reindeer', and 19. Ewen *nolima* 'sledge', borrowed as: PY ***lolimə** > TY *lalime* 'sledge', *lalimed'aa* 'people with sledges', etc. (unless borrowed into Ewen from Yukaghir).

6 This is: 1. Chukchi *sawsi*, *čawču* 'reindeer breeder', borrowed as: PY ***ča:ča:** > TY *čaačaa* 'a reindeer-breeding Yukaghir tribe'.

7 This is: 1. Koryak *ineŋ*, borrowed as TY *iniŋie* 'sledge for transporting lumber and the poles of a yurt'.

referenced in Nikolaeva's *A Historical Dictionary of Yukaghir* (2006). Old and new borrowings into Yukaghiric from (Pre-)Yakut[8] have recently been discussed elsewhere (Piispanen 2013a: 115–139). Borrowings between languages must of course be precisely described in lexical, phonological and semantic terms, and attempts at a chronology should be made.

In the case of borrowings into Yukaghiric, it is known that Tungusic *-u- and *-o- of *early* borrowings (dated to ca. 1500 BP) are both found as Late Proto-Yukaghir (PY) *-o-, while the Tungusic *-u- and *- o- of *late* borrowings (dated to ca. 1000 BP) are both found as PY *-u- (Nikolaeva 2006: 58).[9] It can be assumed that Turkic borrowings into Yukaghiric follow similar vocalic change patterns and chronology (Piispanen 2013a; reading suggested for a brief background to Yukaghir vowels, prosody and Yakut and Tungusic borrowings and their chronology). Below, a number of new borrowings and a few new suggested Uralic–Yukaghiric cognates directly related to reindeer economy are presented.

New borrowings between Tungusic and Yukaghiric

> *New borrowing*
> Ewenki *mullikan* 'reindeer which cannot be trained' (Nedjalkov 1997: 333), *mulliikaan* 'wild reindeer; reindeer strayed from the flock' (Vasilevitch 1958: 261), borrowed as: PY ***moll'ə** > TY *moll'e*, KJ *molle*, *mole*, KD *molle* 'small wild reindeer'.

This constitutes another borrowing related to reindeer economy between Yukaghir and Ewenki. The direction of borrowing is from Ewenki to Yukaghir, even though the end suffix is missing in Yukaghir, since there are also other known reindeer-related borrowings in this direction (as shown in footnote 5 earlier). Furthermore, there are Ewen *mụlqan* 'deer', and Negidal (a dialect of Eastern Ewenki) *molkān* 'deer' (TMS 1 534, 555), clearly making this a fully Tungusic word, but neither of the two can be the source of borrowing due to phonological reasons. While Yukaghir usually borrows only roots, this could possibly be a Pre-Ewenki borrowing before the suffix *-kan* was added, although the presence of the same suffix in the Ewen and Negidal words as well shows that the suffix was already in place by the time of borrowing; the Yukaghirs thus only borrowed the root, obviously knowing

8 With the term *Pre-Yakut* I indicate an earlier form of the modern Yakut language, which had not yet undergone certain sound changes, and appears to have been spoken some 1000–1500 years ago; example: Pre-Yakut ***ytymaq**~***ytymeq** 'fishing pole' > Yakut *ytymex* 'fishing pole' (Piispanen 2013a).

9 Furthermore, recent population genetic studies actually prove—to a surprisingly accurate degree—the suggested ancient tribal contacts between the populations who borrowed lexical items from each other during these *early* and *late* periods of time, suggesting that the numbers are relatively accurate. (I intend to present a paper on such details in the future.)

that the suffix was indeed a suffix. Indeed, the vowel correspondence PY *-o-* ~ Ewenki *-u-* would suggest that it is an *old* borrowing, placing it perhaps around 1500 BP (Piispanen 2013a: 120). Further, the geminate in Yukaghir, which would have been lost far earlier with universal degemination (Piispanen 2013b) also suggests that the Yukaghir words are (Pre-)Ewenki borrowings post-dating degemination (in EY or MY). The palatalization of *-l-* is of course related to the easily palatalized cluster *-li-*. Semantically, there is a clear connection between 'wild animal' and 'animal that cannot be trained', as many trained animals have been raised as domesticated animals; an *untrainable animal*, i.e. *untamable animal*, is a *wild animal*. Curiously, there is also Finnish *mulli* 'young oxen', although this must be only a chance similarity.

The PY root in itself means 'small wild reindeer'. In this case, the Ewenki item was suffixed with *-kan*, which has a specifically diminutive meaning (Nedjalkov 1997: 298). The same suffix is found with other reindeer terms in Ewenki as well, and probably constitutes an early productive suffix in several Tungusic languages. Examples (Ewenki vocabulary taken from Nedjalkov 1997: 333–334) include *avlakan* 'one-year-old reindeer', *ńogarkan~ńovarkan* 'four-year-old reindeer', *amarkan* 'five-year-old reindeer', and *bagdakatkan* 'wild calf-reindeer', and I conjecture that the suffix may have been present earlier in other reindeer terms such as *bagdaka*: (I suggest: < *bagdakan) 'wild reindeer', *engneken* (I suggest: < *engnekan) 'young calf-reindeer' and *kumaka* (I suggest: < *kumakan) 'red deer'.[10] However, terms such as *songgačan* 'new-born calf-reindeer' and *epkača:n* 'less than one-year-old female reindeer' are likely instead to bear the suffix *-čan*, with a pejorative meaning, which semantically can be traced to the fact that very young reindeer may appear not only small, but also fragile and weak. Given this, it is noteworthy that the meaning 'small' is still found in Yukaghiric, suggesting that the (Pre-)Ewenki word was already suffixed and carried this exact meaning.

> *New borrowing*
> Ewenki *aŋa* 'wild game; beast' (Vasilevitch 1958: 30), *aŋa-* 'to graze (of deer)', *aŋan* 'enclosure for deer', dial. *anaŋ* 'mountain ram', also: *ongkovor*, *ongoskocho* 'reindeer with a skewbald patch, spot on a muzzle', borrowed as: PY ***oŋ-**, SU *oŋyl*, B *onye*, ME *ongei* 'reindeer'.

This is a likely borrowing from Ewenki into Yukaghiric. The Tungusic forms appear to originate in the root *aŋ-, which generally relates to (rein)deer. This has been raised phonologically with *ong-* in compounds meaning 'reindeer with skewbald patch'. The same *ong-* (orthographically more properly *oŋ-*)—meaning 'reindeer'—

10 Actually, the words for many animals in Ewenki—big or small—also surprisingly have the suffix *-kan*, including *kulikan*, *amika:n* 'bear', *čipkan* 'sable' and *munnukan* 'hare'. The very same suffix seems to be present also in *kungakan* 'child', suggesting affective uses for it.

is a borrowing into Yukaghiric found only within a very limited geographic area, i.e. in dialects only (where different suffixation patterns have been applied). This root is suffixed in Yukaghiric (*-l is a nominal derivational suffix in Yukaghir (Nikolaeva 2006: 81)).

Assuming that the change *aŋ- > *oŋ- predates the borrowing, the vocalism of Tungusic *-o- > Yukaghiric *-o- suggests that this is an *early* borrowing, i.e. from around 1500 BP (Piispanen 2013a: 120).

The semantics do suggest '(rein)deer' for all words relating to both the Ewenki and Yukaghiric sets. In dialectal Ewenki the semantically shifted meaning 'mountain ram' is also found, while 'wild game' is a quite natural referent for 'reindeer' for populations mainly involved in reindeer husbandry. The meaning of 'reindeer with a skewbald patch' is semantically specified in Ewenki, constructed either through a complex, non-identified suffixation pattern (less likely) or through some sort of lexical compounding (more likely) to the ancient root *aŋ- (> *oŋ-).

> *New borrowing*
> Ewenki *sulaki* 'fox' (Nedjalkov 1997: 334) *sulakii* 'fox' (Boldyrev 1994: 182), borrowed as: PY ***saqoli** (*saqa- in Nikolaeva, I. 2006:396) > KY *šaqala* 'fox', *šaqalən-* 'yellow', *šaqal'əlbo*:- 'yellowish', *šaqalədaj-* 'to become yellow', *šaqal'ənilbo*:- 'pitted' > *šaqan* 'foxy, fox', *šaqad- abut* (< *saqa/on-t-aw-ut) 'a place in the upper Jasačnaja, lit. den of the fox', TY *saaqid'eń-~saaqičeń-* 'yellowish-grey', *saaqiń* 'a man in a story, lit. foxy, sly', *saaqid'aa* 'a man in a story; yellowish-grey dog with a black tinge', RS *šakoli* 'fox', *šoxolonei* 'yellow', SU *čoxóla*, *čoxolod-* 'fox', KD *caxale~cexel'e* 'fox', *čaxaladail'el~caxaladailel* 'isterus', *caxaluo* 'one-year old elk', KL *čoqolojent* 'fox', B *tshakala* 'fox', *tshakolonni* 'yellow', ME *tschokola* 'fox', *tschakolonni* 'yellow', MK *tschochála-* 'fox', *tschochólani* 'green', etc.

The direction of this borrowing which interestingly exhibits metathesis, is Ewenki > Yukaghir. First, the KY form appears to be homogenized into all-identical vowels, a form that would not have yielded the heterogeneous vowels of the Ewenki word. Second, on the Tungusic side, we also find Ewen *hụlää* 'fox' and Negidal *solaxįį* 'fox', clearly showing that the borrowing is from Ewenki to Yukaghiric.

The presence of the word-initial sibilant must be carefully accounted for. Early *s- in Yukaghir—in both inherited and borrowed vocabulary—would be subject to either retention (producing *š-* in KY (Nikolaeva 2006: 66–68)), deletion (producing ø-) or lateralization (producing *l-*), depending on the exact phonological surroundings (Piispanen 2015); the structure *sul- would have deleted the sibilant altogether, which has not occurred here, while *saq- would have retained the sibilant (in contrast to *sak-). Therefore, we must assume that this word was metathesized during

borrowing, and that the *-k-* changed into *-q-* due to the Yukaghir rules of synharmonism.[11]

The different Yukaghir languages display a large number of vowel variations. The previously suggested *saqa- takes into account neither the high-voweled forms nor the *-l- that is present in practically all of the words, and it should therefore be reconstructed as ***saqo-** or even ***saqoli** (as perhaps directly shown by the archaic RS form!). In other borrowings listed in *A Historical Dictionary of Yukaghir*, we find TU *-a- > PY *-a- or, rarely, *-o- or *-y-, which explains the various Yukaghir forms. Further, the shift from Ewenki *-u-* to Yukaghir *-o-* suggests that this is an *early* borrowing from ca. 1500 BP (Piispanen 2013a), which also accounts for its presence in practically all Yukaghir languages and dialects. Yukaghirs have also borrowed other names of fauna from surrounding populations, an example being Chukchi *milúte* 'hare', borrowed as PY ***milúte** > KY *melate* 'hare' (noted in Nikolaeva 2006, entry 269), a vocalism that would also be consistent with an *early* borrowing. Furthermore, 'fox' is also borrowed into certain other Yukaghiric dialects such as MC *jajdel'* 'fox', borrowed from Chukchi *jájcol~játjol* 'fox' (noted in Nikolaeva 2006, entry 626).

While the source has not been determined, MK *indéndsche* 'fox', RS *kinlič'a* 'fox', MO *kille* 'fox' (< *kinle) and RS *ńandimide* 'black and gray fox' are also quite likely borrowings.

> *New borrowing*
> Ewenki *ikeri* 'spine, vertebrae' (Nedjalkov 1997: 329) *iikeerii* 'spinal cord; bone' (Vasilevitch 1958: 161; Boldyrev 1994: 302), borrowed as: KY *ugur* 'spine, ridge', KJ *ugur* 'spine' (? < PY ***öγur** in Nikolaeva 2006, entry 1591).

The form has limited spread in Yukaghiric, perhaps suggesting borrowing. The vowelism of this prospective borrowing is irregular—which, however, is extremely common with Tungusic borrowings in Yukaghiric. Perusing *A Historical Dictionary of Yukaghir* one finds the following borrowing correspondences:

TU *-a- > PY *-a- or, rarely, *-o- or *-y-; TU *-o- > PY *-o- or *-u-, or rarely, *-ö-;

TU *-ö- > PY *-ö-; TU *-u- > PY *-u- or *-o- or, more rarely, *-i- or *-a- or *-ö- or *-y-;

11 Interestingly, there is also Ewenki śekalan 'lynx', which has the same consonants in the same order as Yukaghiric *saqoli. Perhaps the metathesis exhibited by the Yukaghiric borrowing (*sulaki > *saqoli) was influenced by the lexical structure of this animal name in Ewenki, even though a *lynx* and a *fox* are rather different in physical shape, color and manner.

TU *-ü- > PY *-u-; TU *-i- > PY *-e- or *-i- or, more rarely *-y-;

TU *-e- > PY *-e- or, more rarely *-y- or *-u- or *-ö-.

As mentioned, changes to *-o- and *-u- in particular show the age of the borrowing for Tungusic and probably also (Pre-)Yakut and other borrowings. However, one may assume that in root-initial positions vowels are changed extra prominently. One can therefore posit the following not entirely satisfactory development: ?*ikeri > *yker (desyllabification with the borrowing) > *ukur (progressive vowel assimilation) > KY *ugur* (voicing of intervocalic plosive). Semantically, of course, all the meanings are practically identical. The direction of borrowing (Ewenki > Yukaghiric) is indicated by the homogenization of vowels that has occurred in Yukaghiric.

> *New borrowing*
> Ewenki *oŋokto* 'nose' (Nedjalkov 1997: 329; Boldyrev 1994: 232), borrowed as: PY ***joŋq***- > KY *joγul* 'nose, cape, promontory', *joγud-aŋil'* 'nostril, lit. nose opening' (< *joŋq-u(l)-nt-aŋ-i-l), *joŋ-ža*: 'beak, spout', *joγulə-qoqšəš-* 'to snore, lit. to choke the nose', TY *joγul* 'nose', BO *júngol* 'nose', MO *niongol'* 'nose', etc.

Suggestion: *oŋok-to > *oŋk- > *joŋq- > KY *joγul* 'nose'.

The common Yukaghir word for 'nose' is an Ewenki borrowing. Also found are Negidal (a Western Ewenki dialect) *oŋokto* 'nose', Ulchi (southern Tungusic) *χoŋqo* 'front part of a boat', and Nanai (southern Tungusic) *qoŋtoro* 'nose' (<?Proto-Tungusic *xoŋo-), showing that this is indeed a Tungusic word. The *-k-* of the Ewenki word, which has a direct correspondence in PY *-q- after assimilation, clearly shows that this is an Ewenki borrowing, long after the Proto-Tungusic stage. Further, the cluster *-ŋkt- is impossible in PY, forcing cluster simplification (> *-ŋk-) after assimilation if the Ewenki third syllable was originally also borrowed. This is an *early* borrowing (likely from around 1500 BP) as shown by the correspondence Tungusic *-o- > Late Proto-Yukaghir (PY) *-o- (Piispanen 2013a). PY *j- is likely to have developed secondarily for some vowel-initial roots due to the influence of the following cluster *-ŋk- (> KY -*γ*- regularly; again, *-l is a nominal derivational suffix in Yukaghir (Nikolaeva 2006: 81)); note that there are very few PY roots of the type *oŋk-~*oŋq-, all of which could alternatively be reconstructed as *oγ-.[12] Actually, many if not all PY roots having the cluster *Vŋk/q-, with V being a front vowel like *-e-, *-i-, *-ü-, *-ö-, may have originated from forms that originally had the proto-sibilant *s- (Piispanen 2016). On the other hand, the numerous PY roots with the back vowel *-o-, like *joŋq-, *joŋč-, *joŋl-, *joŋn- and *joŋt-,

12 The only exception would appear to be PY *öŋkə- > KY *oγo*:- 'to stand'.

may suggest that the root-initial *j- is secondary for roots that originally had *o- only; such a secondary phonologic development would also be evident with this borrowing. Furthermore, many *e- initial PY roots have alternatively been constructed as *je-, again perhaps suggesting epenthetic *j- with some vowel-initial roots as a (semi-)regular change in Yukaghiric. Such epenthetic effects could in some cases be Russian influences.

New borrowings between Turkic and Yukaghiric

> *New borrowing*
> (Pre-)Yakut *soɣuruu~soŋuruu* 'south' (cf. Old Turkic *yoqarï* 'up(wards)'; *yoqar* 'to rise' (Nadeljaev et al. 1969: 273)), *saɣuruuŋu* 'southern, south' (a derived adjective), borrowed as: TY *saa-laaɣare* 'south, lit. tree left', TY *sespe-saaɣare* 'Southern part of the yurt, left of the entrance, lit. entrance left', TY *jawun-saaɣara* 'southern side of a road, lit. road left', TY *saaɣare* 'left side of a yurt; West', TK *saɣand'a* 'a little aside from smth', *saɣar* 'side', *saɣargudeŋ* 'aside', and also KY *šajɣər* 'aside' (< PY *sa:ɣər).

This constitutes a new (Pre-)Yakut borrowing in the Yukaghiric languages, as attested by the presence of Old Turkic *yoqarï*. In Yakut, all word-initial s- disappeared, and the modern s-initial words are therefore believed to be borrowings from other Turkic languages or other sources. Yakut s- regularly developed from Turkic *y- or *č-, or appeared with later Turkic borrowings (Anderson 1998). In this case, the Yakut words go back to Turkic *y-. In Yukaghiric, on the other hand, the proto-sibilant *s- would be retained, lateralized or deleted in the modern languages depending on a number of phonological factors (Piispanen 2016; this also discusses Yakut root-initial sibilants); Pre-PY *sVɣ- > *θVɣ- > PY *lVɣ-, with back vowels, is expected. This lateralization is indeed realized in TY *saa-laaɣare* 'south, lit. tree left', the only form also preceded by a low back vowel. The other (compound) forms all retain the sibilant, the reason for which being that the borrowed PY form was the long-voweled *sa:ɣər, and not the theoretical *saŋqər—as is still found with some TY forms and with the KY form (demonstrating *-a:ɣ- > *-ajɣ-*)—as this would likely have blocked the lateralization in most cases (Piispanen 2016) had this been a very early borrowing. In summary, these factors seem to show that the borrowing took place after the change *y- > Yakut s-, and also probably, but not necessarily, after the sibilant alternation sound changes had already occurred in the Yukaghiric languages.

Yakut *saɣuruuŋu* and *soɣuruu* are clearly related to Old Turkic *yoqarï* (note that *-aɣu-* is not typical in modern Yakut, which instead prefers forms like *-aɣï* or *-oɣu-*). It is known that Tungusic borrowings with *-o- were retained as *-o- in Yukaghiric, but also that they could sporadically change to *-a- in *early* borrowings. Since such vocalism may also be the case here—if the Yakut form with a first-syllable mid

vowel was indeed the original form—then one may assume that the borrowing took place after 1630 BP (as suggested by population genetics data in Piispanen 2013a: 134). If the Yakut form with a low first-syllable vowel—i.e. the derived form *saγuruuŋu*—was the original form, however, it could have occurred a bit later as a *late* borrowing from 900–1300 BP and undergone a sporadic change *saγ- > *soγ-* before the borrowing.

The Yakut phonological forms are only slightly odd, and possibly are so due to some unusual clusters. In Yakut *soγuruu* and *saγuruuŋu,* the word-initial *s-* simply goes back to Turkic *y-. Original Turkic *-oγur- and *-aγur- would regularly be found as Yakut *-uor-*, which is not found in this case, suggesting instead a probable original Turkic *-q- for these words. Further, Yakut *-uu-* often goes back to *-uγu- (Anderson 1998: 1–32), but this cannot be the case here, suggesting derivational markings in the Yakut words instead. There is no doubt, therefore, from either a semantic or phonological viewpoint, that Turkic *yoqarï, or something very close to it, is the origin of the Yakut words, which then were borrowed into the Yukaghiric languages.

As for the semantics, Yukaghir appears to be a west-oriented system in its world view, and so turning left from there would lead to the south. In Yakut, the meaning of 'south' appears to originate simply from 'to rise' and 'upwards', which may be related to the position of the sun during midday. The meaning of 'left (not west)' would then be 'south' in a Yukaghir borrowing. Comparing the orientation systems to those of Kyrgyz, another Turkic language, is interesting but reveals significant differences: in Kyrgyz directions are instead given by specifying 'side': Kyrgyz *zhak* 'side' > *tyn zhak* 'north' (< *tyn* 'true; firstborn'), *kyn zhak* 'south' (< *kyn* 'sun'), *kybla zhak* 'west' (Judaxin 1985: 217) and *kökurök zhak* 'east' (< *kökurök* 'breast, chest') (Judaxin 1985: 419). Interestingly, some of the Yukaghir words do signify '(a)side'. These phonological and semantic factors clearly show this to be a (pre-)Yakut borrowing in the Yukaghiric languages.

New borrowings between Turkic and Uralic

New borrowing

Proto-Turkic ***qan**~***kan** 'blood'; Old Turkish *qan* 'blood', *qanlïγ* 'bloody' (Nadeljaev et al. (eds.) 1969: 416–417); Turkish *kan* 'blood', *kana-* 'to bleed'; Azerbaijani *qan* 'blood'; Turkmen *gan* 'blood'; Tatar *qan* 'blood'; Bashkir *qan* 'blood'; Kazakh *qan* 'blood'; Kyrgyz *kan* 'blood' (Yudaxin 1985: 338); Uzbek *qon* 'blood'; Uyghur *qan* 'blood'; Yakut *qa:n* 'blood'; Chuvash *jun* 'blood', etc., borrowed as: PS ***kẹm*** 'blood' (SW 65): Nganasan *ka(a)m* 'blood'; Enets *kî'* (gen.sg. *kio'*) 'blood'; Yurak *χyva* 'he bleeds'; Tundra Nenets *χ3.m'* 'blood', *χ3.vas'* 'пойти (о крови)' (?<*kẹmå-); Forest Nenets *kị̀ẹ̀m* 'blood'; Selkup *kem* 'blood'; Kamassian *k^{c}ə̂`m* 'blood'; Koibal *kam* 'blood'; Mator *kem* 'blood', borrowed as: Tofalar *gomdu* 'he bleeds'.

While this case may have little to do with reindeer husbandry—unless the word for 'blood' was borrowed precisely because of reindeer husbandry—it may solve a long-standing etymological problem. The common Samoyed root for 'blood' has been reconstructed as PS ***kẹm** and a noteworthy resemblance to Turkic words meaning 'blood' can be noted. This may suggest that the Samoyed root is a Turkic borrowing. The phonology, however, is complicated, and the tentative chronology is uncertain. Old Turkish is attested in the 7th century, while Proto-Turkic (the ancestor of both the Eastern Turkic languages (Siberian, Kyrgyz-Kipchak, Arghu) and the Western Turkic languages (Oghuz, Kipchak, Karluk, Oghur)) must be quite a lot older. Clauson suggests that Proto-Turkic existed long before the Christian era and then broke up at the start of the Christian era into the Shaz-Turkic and Lir-Turkic branches (Clauson 1962). Róna-Tas estimates the beginning of Proto-Turkic at around 4000–4500 BCE (Róna-Tas 1998: 67–80). Proto-Samoyed, however, is dated at perhaps no more than 2000 BCE, which suggests that the borrowing would have to be from a specific Turkic language, and there are many from which to choose. Borrowing could even have taken place from a now extinct Turkic source.

I suggest the following scenario. The Samoyed root is a Turkic borrowing from a form like *kan or, more likely, *qan, which had the root-final sonorant altered from *-n to *-m, possibly during the borrowing itself, or under the influence of an original final vowel.[13] Such alternations between *-m and *-n are both known and are not uncommon in other Uralic languages; cf. Fin. *minä olen* 'I am' < PU **mon wolem* 'I am'.

Actually, it cannot be completely excluded that the Proto-Turkic form did not have *-m in the first place. Often, Chuvash has *-m* where Common Turkic has *-n*, a previous matter of controversy between Doerfer and Hovdhaugen, where the former suggested (Doerfer 1973) that only Chuvash retains an original *-m* from Proto-Turkic *-m* whereas it would have been changed in Common Turkic to *-n*. However, in this case, Chuvash actually has a *-n*, and not a *-m* so here that argument seems moot. Still, the history of this Turkic word may have to be rethought considering this possible Samoyed counterpart.

Here I must also mention that the Turkic-speaking Chuvash are genetically a Turkicized Uralic population (Graf et al. 2012) that had earlier spoken a Finno-Volgaic language. I therefore conjecture that some of the peculiarities of Oghur Turkic may be explained by the phonology of the early Mari or Mordvinic languages or something quite close to them (all of which still today partake in a *Sprachbund*). From both a geographic and historic point of view it seems that the most likely conquering group was either the Khazar Khanate (618–1048 AD) or Volga Bulgaria (7th century–1240 AD), which is why the language change to what

13 Most PS roots are shortened forms of PU roots; disyllabic roots have often become monosyllabic roots, which may also apply to PS *kẹm, and thus they may originally have been borrowed as *kani or *kanə (?> *kam) from a Turkic source.

is today Chuvash may occurred as early as the 7th or 8th centuries or as late as the 12th or 13th centuries.

Since Proto-Samoyed had no *q it would have been borrowed as PS *k. The Proto-Samoyed vowel, then, has been reconstructed as *-ę -, although some of the languages do have an -a- (Nganasan and Koibal), and -e-, -o-, -î-, -ə̂ and -3.(!) are even encountered. In this context, the possible change of *-a- > *ę is interesting, as it is also encountered, for example, with: PU ***apte*** 'hair' (UEW 14-15) > PS ***ęptə̂*** '(head) hair'; PU ***ajtV*** 'to get loose; to free oneself' > Kamassian *ęl'èm* 'I let loose', as well as the similar: PU ***jäsne*~*jäsen** 'joint, limb' > PS ***ęsen*~ *ęsən** 'joint, limb' (in detail also: EM *ezńe* 'knee, joint', KB M Mari *ježaŋ* 'Gelenk; Knötchen einer Pflanze', VO KZ *jez* 'Gelenk', *jęz- vij* 'Gliedgelenk', P KZ *jęzna* 'muscle; joint', Tundra Nenets *ŋeso'(n)* 'joint; section, segment'). Hence, an argument can be made for Turkic ***qan*** 'blood', borrowed as: PS ***kęm*** 'blood'.

The vowel change may actually be motivated by avoidance of homonymy with PU ***kama*** 'peel, crust' (UEW 121–122) > PS ***kamə*** 'Schuppe' (PS *kämg- 'Schuppe' in SW 63; PS *ä should rather be reconstructed as PS *a according to Helimski, E. 2005) > O. Nenets *śāw*, J. Nenets *śem* 'scale, flake'; Nganasan *kamu* 'peel, crust'; Selkup *qā m* 'fish scale'; Kamassian *kåm* 'scale'; Mator *kamengapty*. While the full details of this borrowing scenario remain to be worked out, it is at least a worthwhile and rather reasonable proposal.

Further, an alternative thesis could be entertained: Proto-Turkic *kem 'river; Yenisei' (> Old Turkic *kem (*käm) 'river; Yenisei') > Tuvan *xem* 'river', Khakas *kim* 'Yenisei' (Janhunen 2012: 70–73) borrowed as PS *kęm 'blood'. The phonology is closer than with the main suggestion, and semantically it is possible to see *blood* as the *river inside the body*. This seems to be paralleled semantically by PIE *bhlo-to- 'to swell, to gush, to spurt; that which bursts out' > Old Eng. *blod* 'blood' > Eng. *blood* (Etymon). However, the identical semantics in the main suggestion above seems much more convincing, and so the phonological similarities of this second suggestion must be ascribed to a chance occurrence only.

Likely borrowings or cognates between Uralic and Yukaghiric

> *New borrowing or cognate*
> PY ***kentə***, TY *kedie-* '(of a tied reindeer that does not want to follow a sledge)' is either a Ugric borrowing or cognate with: Khanty *kĕnt* 'wrath, anger, malice', *kănt* 'anger', Mansi *känt*, *kant*, *kantaŋ* 'angry', Hung. *kedv* '(to be in a) mood, emotional, desire', Pug ***kVntV-*** 'whim, mood, caprice' (UEW 861–862).

In Yukaghir, a valid disyllabic noun root must have a prosodic structure of (C)VCCə-. Therefore, the final vowel could have been anything else in MY (if inherited) or in an original donor language (if borrowed). Semantically, the meanings

of the TY item—'stubborn, obstinate, scornful'—are mirrored by the meanings found in PUg 'mood, caprice', as well as in Hungarian 'to be in a mood' and Khanty/Mansi 'angry'. The Hungarian form suggests an original first-syllable front vowel. The correspondences may thus suggest that an earlier, prospective Pre-PU *kentV- 'whim, mood, caprice'—as a cognate—had a front-voweled first syllable (likely *-e-), or that a Uralic donor language, such as *early* Khanty, was of the form *kent(V).

> *New borrowing or cognate*
> PY ***pet***-, TY *petčigije* 'reins' (> *petčigije-moojnijaa* 'skilled reindeer-driver'), *petče* 'light sledge for carrying people', *pettes*- 'to drive reindeer (TR)', *petčigijes*- 'to put the front part of the harness on a reindeer' <> Nenets *pod'ern* 'collar for horse, dog', Enets *fóren*, *fòden* 'pull strap', ?*poδida~poδer* 'reindeer harness', Nganasan *fúdar* 'yoke, pull strap', *huturə*- 'to harness', ?Selkup *paater* 'garter', PS ***potV*** 'harness, gear for the sleigh' (UED 794).

A rare reindeer-husbandry related borrowing between Proto-Samoyed and Yukaghir can be presented to show that borrowings are indeed possible between the geographically closest modern Uralic languages and Yukaghir. The semantics have good correlations which have been further expanded after PY through suffixation, which can be expected in a culture deeply involved in reindeer husbandry. No other earlier correspondences have shown the specific vowel correspondence as seen with PS *pot(V)- <> PY *pet- (i.e. back *contra* front vocalism), which makes this something of a special case; it could suggest palatalization in the PS branch only, which would be unusual. While one could imagine labialization in Proto-Samoyed, this could possibly constitute a rare borrowing between Samoyed and Yukaghir, both populations being far-traveling precisely because of their dependence on reindeer husbandry. If it is a borrowing, it may go back perhaps to ~1500 BP, but the direction of borrowing is not at all clear from the phonological correspondences alone.

This root pertains to reindeer technology and it seems implausible that it could constitute a common Pre-PU root going back several millennia; reindeer husbandry may simply not be old enough to have been present as common terminology that reaches back thousands of years. Still, if it were a cognate, it would accord well in accordance with the fact that PU *-o- is often found as PY *-e- with other cognate suggestions; examples include PFU ***kolV*** 'rift, cleft, interval' (UEW 174-175) <> PY ***ke:***- > KY *ke:l* 'slot', etc. (noted in Nikolaeva 2006, entry 768); PFU ***wolka*** 'shoulder' (UEW 581) <> PY ***wele***- > KY *ejeji*:- 'to carry', etc. (noted in Nikolaeva 2006, entry 2603); PU ***ńole***~***ńo:le***- 'to lick' (UEW 321) <> PY *ńel- > KY *ńel'i*:- 'to lick' (noted in Nikolaeva 2006, entry 1401), etc.

The Samoyed words have not spread to other Uralic languages either, not even the Ugric languages. For these reasons this set may best be described as a rarely

borrowed technological term instead of a cognate set. Enets *poδida~poδer* is my own prospective addition to this set.

New borrowing or cognate
PY ***a:čə~*wa:čə**, KY *a:čə* 'domestic reindeer' > KY *a:ča* 'domestic reindeer', B *aaitsha* 'domestic reindeer', ? *ishakalloo* 'young reindeer', M *áča*, *áčapul*, *aača*, *aače* 'domestic reindeer', ?ME *tschakalloa* 'young reindeer', etc. <> PFV ***wača~*waća** 'young animal, female reindeer, foal' (UEW 808–809), ?Fin. *vaadin*, dial. *vaame*, *vaami*, *vaatimo*, *vain*, *vaija* 'reindeer cow' (possibly < N. Saami *vājâ-ččâm-* 'full-grown female reindeer (which already has or has had a calf)', L. Saami *vātjav* (L) 'reindeer cow', T. Saami *vāj* 'young reindeer cow', Kld. Saami *vā^{Di}Zì*, *vaÌij(a)* 'reindeer cow' (> Russ. *важенка* > KZ & Udm. *važenka*?), Kar. *voajin* 'reindeer cow', ?EM *vašo*, ?MM *vaša* 'foal'. Also possibly: PFU ***wasa** 'calf, deer calf' (UEW 814-815), Fin. *vasa* 'Kalb, einjähriges Renkalb', *vasikka* 'Kalb', Est. *vasik*, *vasikas*, I. Saami *vyesi*, T. Saami *viisse*, Kld. *vūiss* 'kleines Rentierkalb, bis es um den Peterstag neues Haar bekommt', EM & MM *vaz*, EM *vaznìe*, MM *vaznìä* 'calf', ?Mansi *wēsəj*, *wēsəγ*, *wāsiγ* 'elk calf'.

In the UEW, Rédei considered PFU *wasa to be an independent Iranian borrowing. However, while such a hypothesis is certainly possible, if it were an original Uralic item one would be able o trace the proto-form back to at least the Proto-Finno-Ugric stage. It seems to me that the etymons *wača & *wasa may both be traceable back to some earlier common etymon related to *young elks* or *(rein)deer* even though the phonological details are not at all clear, since an earlier *-č- may regularly reflect a later *-t-, but not really an *-s-. Perhaps the two roots were independently borrowed into the two proto-languages, but actually originated from a single foreign etymon that changed over time. The Late Proto-Yukaghir form *wa:čə is particularly close to PFV *wača, both in phonological and semantic terms, meaning that the two could constitute either true cognates (< Pre-PU *wača) or ancient borrowings, with the common semantics appearing mainly to pertain to 'young reindeer'. The first-syllable long vowel and the second-syllable final vowel of the PY root are explained by Yukaghir prosody, where CV:Cə- is one of the valid disyllabic nominal root structures.

New borrowing or cognate
PY ***se:rti:**, TY *sierdiid-ile* 'reindeer not selected for slaughter' (< TY *ile* 'domestic reindeer') <> PU ***śarta** 'young deer, reindeer' (UEW 464), EM & MM *śarda*, EM *śardo* 'elk, dial. reindeer', Mari *šarδę* , *šorδo*, 'elk', Khanty *surti* 'one-year-old tame reindeer', *sŭrtĭ* 'one-year-old calf/foal of a cow, horse, elk or tame reindeer', Mansi *surti* 'year-old reindeer calf', Nenets *siraj* 'one-year-old reindeer cow' (> Khanty *sira*, *siràj* '1.5-year-old reindeer' >

Ewenki *siru* ‘male reindeer in the period of pairing’), ?Selkup *sjaera* ‘Cervus tarandus’.

This is an interesting set showing both Uralic and Yukaghiric cognancy (or borrowing) and further borrowing into the Tungusic language of the Ewenki (the latter being noted in the UEW). The UEW entry also compares the Uralic words with the Turkic Shor *sartak* ‘reindeer’ and sagaisch-koibalisch *sardak* ‘einjähriges Maral’, and with Mongol *sarluγ* and Khalkha Mongolian *sarlaγ* ‘yak’. As such, this may be an ancient *Wanderwort*, and the borrowing sequence would therefore be extremely difficult to trace. The second part of the TY compound *-ile*, is also a likely borrowing: cf. Ewenki *elkēn* ‘wild deer’; Ewen *iēlken~elken* ‘domestic deer’; Yakut *elik* ‘чубарый олень’ (cf. Old Turkic *älik* ‘roe-buck’); Middle Mongolian *ele'ut* ‘a kind of camel’ (Haenisch 1939), another possible *Wanderwort*.

Several factors also point to the PY root itself being a borrowing: first, it has a long vowel in a closed syllable, the anomalous vocalism (front *contra* back vocalism as compared to Uralic); second, the root is only found with TY and no other dialects and as no other derivatives; third, the disyllabic prosodic structure of *CV:CCV:* appears in this case to result from a contraction of a final short vowel *-i- with a word-final glide like *-j (i.e. < *se:rtij) (Nikolaeva 2006: 76)—just as happens to be found word-finally in the Nenets word (> Khanty > Ewenki)—but *-j is not a known or fitting suffix. On the other hand, Yukaghir has a nominal derivational marker PY *-i: (Nikolaeva 2006: 80), which may explain the form of the PY root. If a borrowing, it is likely from around 1500 BP, as was suggested for the case of PS *potV <> PY *pet- above. The PY form should probably be newly reconstructed as *śe:rtij.

Semantically, the entire set pertains to *young reindeers* (semantically shifted to *yak* in Mongolic). A ‘tame reindeer’ may in some sense be considered a ‘reindeer not selected for slaughter’, since pets are not usually slaughtered. Furthermore, Ewenki *siru* appears to be a direct Khanty borrowing; indeed, a male reindeer becomes independent in early fall after almost one year as a calf (being born around May–June) and can then take part in the period of pairing the following year in October–January, depending on species, at an approximate age of 1.5 years. Likewise, female reindeers can be sexually mature at 16 months of age (~1.5 years), but are more commonly so a year later (i.e. in the following pairing season).

Structured semantic fields

Most of the borrowings presented in this paper pertain to reindeer terminology. Further dividing the borrowings into cultural or technological sub-spheres of semantics (such as those presented in Rédei (1999)) yields the following categories:

a. body parts of humans and animals—KY *ugur*, TY *joγul*, PS *kẹ̄m
b. animal kingdom (i.e. fauna)—PY *moll'ə, PY *oŋ-, PY *saqoli, PY *wa:čə, PY *se:rti:

d. nature, natural phenomena and natural places—KY *šajɣər*
e. types of work and tools—PY *pet-
m. elementary phenomena, actions and perceptions—PY *kentə

Most of the borrowings pertain to fauna, in particular, semantically specified descriptions of reindeer as seen from the point of reindeer husbandry (age, level of domestication, suitability, etc.). Interestingly, a technological term, or rather, a root, also pertaining to reindeers, was found as a borrowing between Yukaghiric and Samoyedic. Perhaps a bit surprisingly, a few body parts were also found as borrowings into Yukaghiric (as well as Samoyedic), as were a geographic direction and an emotion, which is now only used about reindeers in Yukaghir. Practically all the borrowings—except perhaps KY *šajɣər*—were found to be chronologically *old* (i.e. of around 1500 BP) in Yukaghiric as evaluated by the vowelism. The Uralic–Yukaghiric correspondences could alternatively be considered cognates.

Summary and conclusions

In this paper, a total of eleven new lexical borrowings (or cognates in the case of Uralic–Yukaghiric) have been presented. These results add to the older research on rather extensive borrowings of reindeer terminology in greater north-eastern Siberia, where several languages, including Yakut (and Dolgan), Chukchi, Koryak, Ewen, Ewenki, Yup'ik, Yukaghir, Russian, etc. are extensively used.

Abbreviations

Fin. = Finnish
Hung. = Hungarian
KY = Kolyma Yukaghir
KZ = Komi-Zyrian
MM = Moksha Mordvin
MY = Middle Proto-Yukaghir
N. Saami = Northern Saami
Kld. Saami = Kildin Saami
S. Saami = South Saami
I. Saami = Inari Saami
T. Saami = Ter Saami
L. Saami = Lule Saami
NT = Northern Tungusic
PIE = Proto-Indo-European
PS = Proto-Samoyed
PU = Proto-Uralic
PUg = Proto-Ugric
PY = Late Proto-Yukaghir
TU = Proto-Tungusic

TY = Tundra Yukaghir
Udm. = Udmurt

Abbreviations of the linguistic resources

B = Materials of Billings 1787
KD = Kolyma Yukaghir from Jochelson's manuscript dictionary
KJ = Kolyma Yukaghir materials of Jochelson in 1898 and 1900
KL = Materials of Klitschka 1781
KK = Kolyma Yukaghir materials of Krejnovič in 1982
M = materials by Maydell presented by Schiefner in 1871
MC = čhuvan materials of Matjuškin (Wrangel 1841)
ME = Materials of Merk 1787
MK = Kolyma Yukaghir materials of Mueller and Lindenau in 1741
MO = Omok materials of Matjuškin (Wrangel 1841)
RS = materials of Rajskij and Stubendorf presented by Schiefner in 1871
SD = Kolyma Yukaghir materials of Spiridonov in 2003
SU = materials by Suvorov presented by Schiefner in 1871
TD = Tundra Yukaghir materials of Jochelson's (1926) manuscript dictionary
TK = Tundra Yukaghir materials of Krejnovič in 1958 and 1982
W = early materials of Witsen in 1692.

All these older materials are described and referenced in Nikolaeva (2006).

References

Aikio, A. 2015. The Uralic-Yukaghir lexical correspondences: Genetic inheritance, language contact or chance resemblance? *Finnisch-Ugrische Forschungen* 61, 7–76.

Anderson, G. D. S. 1998. Historical aspects of Yakut (Saxa) phonology. *Turkic Languages* 2, 3–32.

Bogoras, W. 2009. *Tales of Yukaghir, Lamut, and Russianized natives of Eastern Siberia*. Hanover, NH: International Polar Institute Press. A reprint of the 1918 edition.

Boldyrev, B. V. 1994. *Russko-evenkijskij slovar'*. Novosibirsk: Rossijskaja Akademija Nauk.

Clauson, G. 1972. *An etymological dictionary of pre-thirteenth-century Turkish*. Oxford: Clarendon Press.

Doerfer, G. 1973. *Tschuwaschisch -m < urtürkisch *-m (> gemeintürkisch -n)*, Ural-Altaische Jahrbücher, 45.

Etymon = Online (English) Etymology Dictionary (www.etymonline.com).

Graf, O. M. & Johnson, S. M. & Mitchell, J. & Wilcox, S. & Livshits, G. & Crawford, M. H. 2012. *Analysis of Chuvash mtDNA points to Finno-Ugric origin*. Abstract from the 81st Annual Meeting of the American Association of Physical Anthropologists in Portland, Oregon, April 11–14, 2012. Found online at: http://physanth.org/annual-meeting /2012/aapa-meeting-program-2012 [retrieved 2012-10-23].

Gurvic, Il'ja 1982. *Ėtničeskaja istorija narodov severa* [The ethnic history of the North], Moscow: Nauka.

Haenisch, E. 1939. *Wörterbuch zu Manghol-un Niuča Tobčaan (Yüan-ch'ao pi-shi), Geheime Geschichte der Mongolen*. Leipzig.
Helimski, E. 2005. The 13th Proto-Samoyedic vowel. In: Wagner-Nagy, Beáta (ed.) *Mikola-konferencia 2004*. Szeged. 27–39.
Janhunen, J. 2012. Etymological and ethnohistorical aspects of the Yenisei. *Studia Etymologica Cracoviensia* 17, 67–87.
Judaxin, K. K. 1985. *Kirgizsko-russkij slovar' 1. A–K*. Moscow: Sovetskaja ėnciklopedija.
Judaxin, K. K. 1985. *Kirgizsko-russkij slovar' 2. L–Ja*, Moscow: Sovetskaja ėnciklopedija.
Krupnik, I. 1993. *Arctic adaptations: Native whalers and reindeer herders of Northern Eurasia*, University Press of New England.
Nadeljaev, V. M., Nasilov, D.M., Tenishev, E.R. & Scherbak, A.M. 1969. *Drevnetjurkskij slovar'*, Leningrad: Izatel'stvo 'Nauka', Leningradskoje otdelenie.
Nedjalkov, I. 1997. *Evenki*. London: Routledge.
Nikolaeva, I. 2006. *A historical dictionary of Yukaghir*. (Trends in Linguistics Documentation 25.) Berlin & New York: Mouton de Gruyter.
Piispanen, P. S. 2013a. Further lexical borrowings from (Pre-)Yakut into the Yukaghiric languages. *Turkic Languages* 17, 115–139.
Piispanen, P. S. 2013b. The Uralic-Yukaghiric connection revisited: Sound correspondences of Geminate clusters. *Journal de la Société Finno-Ougrienne* 94, 165–197.
Piispanen, P. S. 2015. Evaluating the Uralic-Yukaghiric word-initial, proto-sibilant correspondence rules. *Journal de la Société Finno-Ougrienne* 95, 237–274.
Piispanen, P. S. 2016. A prosody-controlled semi-vowel alternation in Yukaghir. *Journal of Historical Linguistics*, Accepted.
Rédei, K. 1999. Zu den uralisch-jukagirischen Sprachkontakten. *Finnisch-Ugrischen Forschungen* 55, 1–58.
Róna-Tas, A. 1998. The reconstruction of Proto-Turkic and the genetic question. In: Johanson, Lars & Csató, Éva Á. (eds.) *The Turkic languages*. (Routledge Language Family Descriptions.) London: Routledge. 67–80.
Sleptsov, P. A. 1972. *Jakutsko-russkij slovar' (JRS)*. Moscow: Sovetskaja ėnciklopedija.
SW = Janhunen, J. 1977. *Samojedischer Wortschatz—gemeinsamojedische Etymologien*. (Castrenianumin toimitteita 17.) Helsinki.
TMS 1 = Cincius, V. I. 1975. *Sravnitel'nyj slovar' tunguso-man'čžurskich jazykov* [Comparative dictionary of the Tungus-Manchu languages] 1. Leningrad: Nauka.
TMS 2 = Cincius, V. I. 1977. *Sravnitel'nyj slovar' tunguso-man'čžurskich jazykov* [Comparative dictionary of the Tungus-Manchu languages] 2. Leningrad: Nauka.
UED = Uralic etymology database (found online at: http://starling.rinet.ru/cgibin/ query.cgi?basename=\data\uralic\uralet&root=config&morpho=0)
UEW = Rédei, K. 1988–1991. *Uralisches etymologisches Wörterbuch*. Budapest: Akadémiai Kiadó.
Vasilevitch, G. M. 1958. *Evenkijsko-russkij slovar'*. Moskva: Gosudarstvennoje isdatel'stvo inostvannyx i natsional'nyx slovarej.

Yukaghir-related etymological dictionaries and glossaries

The etymological reference works consulted in Nikolaeva, I. (2006), the source of Yukaghir used for this research, were, given with common short-hand:

CED (Fortescue, M. et al., 1994),
DEWOS (Steinitz, W. 1966–1993),
EDAL (Starostin, S. A. et al., 2003),
ESRD (Anikin, A. E. 2000),
ESRZ (Anikin, A. E. 2003),
HUV (Collinder, B. 1965),
JRS (Slepcov, P. A. 1972),
JU (Collinder, B. 1940—notes Yukaghir parallels and arguments),
LR (Fortescue, M. 1998) and
UJN (Collinder, B. 1957).

Further, the following glossaries were consulted and collected therein:

Angere, J. (1957, based on Kurilov, G. N. 1900),
Kurilov, G. N. (2001),
Veenker, W. (1989),
Endo, F. (1997 & 2001),
Nyikolajeva, I. (2000),
Nikolaeva, I. and Shalugin, V. (2003) and
Maslova, E. (2001 & 2003).

Report

Parallel corpora and Universal Dependencies for Turkic

Éva Á. Csató & Hüner Kaşıkara & Beáta Megyesi & Joakim Nivre

Csató, Éva Á. & Kaşıkara, Hüner & Megyesi, Beáta & Nivre, Joakim 2015. Parallel corpora and Universal Dependency for Turkic. *Turkic Languages* 19, 259–273.

The first part of this paper presents ongoing work on Turkic parallel corpora at the Department of Linguistics and Philology, Uppsala University. Moreover, examples are given of how the Swedish-Turkish-English corpus is used in teaching Turkish and in comparative linguistic studies. The second part deals with the annotation scheme Universal Dependencies (UD) used in treebanks, and its application to Turkic languages.

Éva Á. Csató, Beáta Megyesi, and Joakim Nivre, Department of Linguistics and Philology, Uppsala University, Box 635, SE-751 26 Uppsala, Sweden. E-mail: eva.csato@lingfil.uu.se, beata.megyesi@lingfil.uu.se, joakim.nivre@lingfil.uu.se
Hüner Kaşıkara, Boğaziçi University, Istanbul, Turkey. E-mail: huner.kasikara@boun.edu.tr

Turkic languages at Uppsala University

The Faculty of Languages at Uppsala University, Sweden, is recognized for maintaining a rich linguistic environment. Regular courses are offered in over forty languages including both dead languages and large, small and vanishing contemporary vernaculars. Turkic languages are well represented in the curriculum. Besides Turkish, which is the main focus in teaching, there are campus and distance courses in the classical literary language Ottoman, modern Turkic languages such as Uzbek, Uyghur, Kazakh, and Azeri, and the highly endangered Karaim language. The Department of Linguistics and Philology, where the Turkic languages program is situated, has a tradition of supporting scholarly environments for lesser taught languages. For several years the department has received financial aid to build up resources for teaching and research. Thanks to this support, Turkic linguistics at Uppsala can offer a relatively large number of language courses and conduct research in comparative Turkic linguistics.

Some projects at the department aim at giving support to lesser taught languages through developing parallel treebanks. Treebanks are parsed corpora, collections of texts with morphological and syntactic annotation. Treebank projects are carried out by experts in less resourced languages such as Hindi, Persian and Turkic together

with internationally outstanding scholars in computational linguistics working at the department. The collaboration has proved to be mutually advantageous. The focus in language technology has been on English and major Western languages. English is one of the languages that has been included in many of the existing parallel treebanks. Access to data on languages representing language types that are widespread in the world, as for instance Turkic, provides excellent opportunities for computational linguistics to develop their language technological theories and tools.

Part 1: Parallel corpora

A parallel corpus is a collection of multilingual text material containing original texts in one language and translations into one or more other languages, with texts placed alongside their corresponding translations. Parallel texts are usually aligned, meaning that corresponding units (sentences, phrases, or even words) from the different language versions are explicitly linked together. The texts can be linguistically annotated with respect to part-of-speech and morphological features, and on the syntactic level. Syntactically annotated parallel corpora are called parallel treebanks.

A number of parallel corpora are available on the Internet, one example being the Farkas Translations (http://www.farkastranslations.com/bilingual_books.php). Farkas presents out-of-copyright literary works in sentence aligned format without any further linguistic analysis. For instance Mark Twain's novel The Adventures of Tom Sawyer can be found in this corpus in the English original and in German, Hungarian, Dutch and Catalan translations. The resource is provided free of charge for the purpose of language learning, language teaching and translation research, and as a demonstration of the text alignment services offered by the website. Farkas has accessed many of the texts published in the Gutenberg Project, a digital library of free books which also can serve as an excellent resource for further projects aiming to build new parallel corpora and other language resources (https://www.gutenberg.org/wiki/Main_Page).

A computational linguist at Uppsala, Jörg Tiedemann, has built a much more advanced collection of multilingual parallel corpora, OPUS, providing various tools and interfaces and a growing collection of language samples collected from the web for building parallel corpora and related tools (http://opus.lingfil.uu.se/). OPUS provides freely available data sets in various formats together with basic annotation to make it useful for applications in computational linguistics, translation studies and cross-linguistic corpus studies. OPUS is probably the largest collection of freely available parallel corpora in the world. It covers over 90 languages and includes data from several domains. Altogether, there are over 3,800 language pairs with sentence-aligned data comprising a total of over 40 billion tokens in 2.7 billion parallel units (aligned sentences and sentence fragments), Including a parallel corpus for Romanian-Turkish. Another improvement of recent versions of OPUS is the availability of various download formats for all sub-corpora. Different tools are available on the OPUS webpage for language-specific taggers and parsers, including for Turk-

ish, as well as alignment tools to automatically link sentences, phrases and words. The texts are processed by Uplug (https://bitbucket.org/tiedemann/uplug), a collection of tools and scripts for processing text-corpora to create (annotated) parallel corpora (Tiedemann 2003, 2012).

Building a Swedish-Turkish-English parallel corpus

The Uppsala parallel corpus containing parallel texts in Turkish, Swedish and English is the first of its kind (http://www2.lingfil.uu.se/corpora/) and was created between 2006 and 2010, as previously described (Megyesi et al. 2006; Megyesi & Dahlqvist 2007; Megyesi et al. 2008; Megyesi et al. 2010). The corpus contains syntactically annotated parallel texts with various annotation layers ranging from part-of-speech tags and morphological features to dependency annotation. Each layer is automatically annotated and the sentences and words are aligned, and the results are partly manually corrected.

In order to build the treebank automatically, a basic language resource kit (BLARK) was used for the included languages. This consists of (i) tokenizers to segment words and punctuation marks and to mark sentence endings, (ii) taggers for the annotation of part-of-speech and morphological features, (iii) parsers for the annotation of syntactic structures in terms of dependency relations, and (iv) aligners to mark the related sentences and words in the translations. Also, we developed tools for the manual correction of the automatic linguistic annotation and alignment. The tools were included in a pipeline for easier processing. The annotation procedure is shown in Figure 1 below.

First, the original materials received from publishers in various formats were normalized to create a consistent, machine-readable format across the corpus data. For example, rtf, doc, and pdf documents were converted into plain text files. After cleaning up the original data, the texts were processed automatically by using tools for formatting, linguistic annotation and sentence and word alignment. During formatting, the texts were encoded using UTF-8 (Unicode) and marked up structurally using XML Treebank Encoding Standard (XCES). The text files were processed by various tools in the BLARKs developed for each language separately.

A tokenizer was used to split the text into tokens such as words and punctuation marks. Sentence segmentation was also performed to break the texts into sentences. Once the sentences and tokens were identified, the data was linguistically analyzed. The annotation was represented in several annotation layers, first on a morphological level, then on a syntactic level. For the linguistic annotation, external morphological analyzers, part-of-speech taggers and syntactic dependency parsers were used, which were trained on annotated treebanks developed for the specific languages. The annotations and labels for linguistic analysis were the de facto standards for the various languages at that time. For example, for Swedish we used the Stockholm Umeå Corpus tagset (SUC 1997) for the morpho-syntactic annotation and the functional annotation of Talbanken05 (Nivre et al. 2006b), while for the syntactic

analysis of Turkish we derived the linguistic annotation from the METU-SABANCI Turkish Treebank (Oflazer et al. 2003). For English, we used the Penn Treebank tagset.

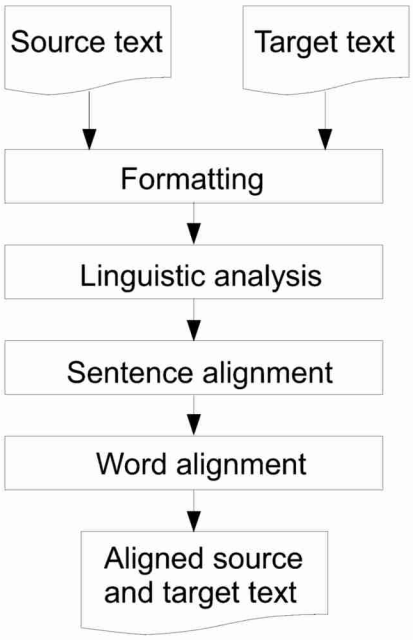


Figure 1. Annotation procedure

The Swedish and English texts were morphologically annotated with the open source HunPoS tagger (Halácsy et al. 2007). The tokens were annotated with parts of speech and morphological features and then disambiguated. The results for the morphological annotation of Swedish show an accuracy of 96.6% (Megyesi 2008). The Turkish material was morphologically analyzed and disambiguated using a Turkish analyzer (Oflazer 1994) and a disambiguator (Yuret & Türe 2006) with an accuracy of 96%. The English data contained less error, approximately 2–3%.

The other linguistic layer contains information about the syntactic analysis. We use dependency rather than constituent structures, as the former has been shown to be well suited for both morphologically rich and free-word-order languages such as Turkish, and for morphologically simpler languages like Swedish. The English, Swedish and Turkish data were annotated syntactically using MaltParser (Nivre et

al. 2006a), and trained on Penn Treebank, Talbanken05 (Nivre et al. 2006b), and the Metu-Sabancı Turkish Treebank (Oflazer et al. 2003) respectively. The annotation includes approximately 15–20% errors, depending of the language, which need to be manually corrected.

The processing tools were implemented in a framework with a graphical user interface, UplugConnector (Megyesi and Dahlqvist 2007), which is based on the modules in the Uplug toolkit (Tiedemann 2003). Our goal was to produce user-friendly tools to make annotation, alignment and correction easy for people with little computer skills.

Resources in the corpus

The corpus includes texts representing both fiction—novels and short stories—and non-fiction, such as information material for immigrants, a book on political history, and texts issued by international organizations. Most of the original texts are in either Swedish or Turkish accompanied by a translation into the other languages. English translations are in several cases missing. The corpus consists of 288 701 tokens in Swedish, 162 302 in Turkish and 140 848 in English, comprising both fiction and technical documents. In order to make the corpus useful for users, search functions are included. Users can search for complete words, beginnings of words, ends of words, and parts of words, being case-sensitive or not. The target language can be Swedish or Turkish and the corpus can be limited to one specific text or text type. Turkish letters are provided.

Another tool produces frequency lists of a freely selectable Turkish text, which has to be copied into a window. The frequency list may be displayed in different formats and contains the number of words, number of different words, average length of words, and type-token ratio in the given text.

Using the corpus in teaching and research

The aim of the Swedish-Turkish treebank is to provide Swedish speaking students and researchers with easily accessible annotated linguistic data on Turkish. The corpus is now being complemented with English texts. The web-based corpus can be used both by regular and distance students in their data-driven acquisition of new vocabulary items and usage. It also functions as a learning platform and for testing hypotheses concerning the morphological and syntactic aspects of Turkish grammar. It further helps students to practice translation between Swedish and Turkish. All this is possible because the Swedish-Turkish parallel texts are available in annotated form. On request, the annotations are visualized in pop-up windows. The morphological analyses are currently given in clumsy, parser-generated formulas but in the near future they will be replaced by more intelligible labels based on the grammatical terms employed in textbooks and in the Turkish Suffix Dictionary (Csató and Nathan 2003). The interface for displaying syntactic information is not ready yet. A search tool assists the students in creating concordance lists. The concordance lists

display whole sentences where the target item appears (and is highlighted). The selected sentences are aligned with their translational equivalents, as illustrated in Figure 2.

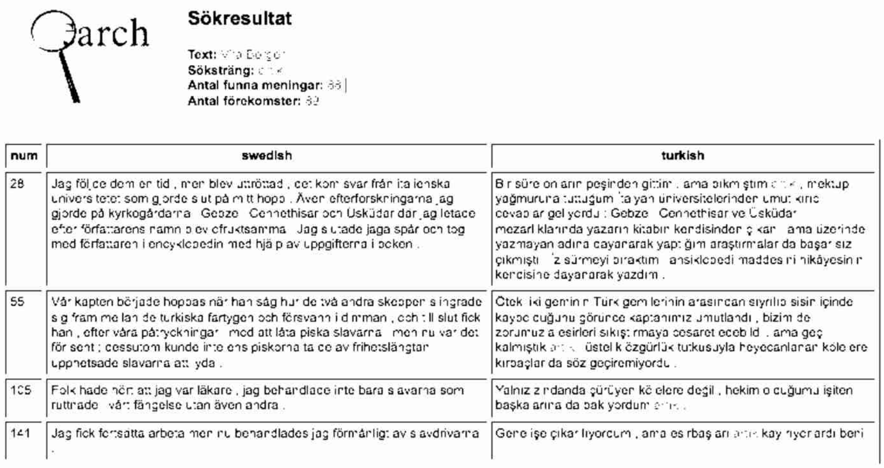


Figure 2. Concordance list

Such lists are used to find frequent patterns of usage, transformational equivalents, different meanings of polysemic words, translational equivalents of Turkish grammatical categories, etc. Different types of exercises are designed and published on the Internet. Students of Turkic languages also use the corpus while writing their theses. Bergdahl (2006) studied the meanings of the Turkish word *gölge* 'shadow' and the corresponding Swedish word *skugga*. Dadasheva (2005) investigated how the Turkish indirective category marked by *-mIş / imiş* is translated into Swedish and Russian. Hedman and Uyghur (2009) compared the meanings of the Swedish and Turkish verbs 'to give', 'to do' and 'to make'. Haktanır (2006) reviewed the ambiguous Turkish morphological forms in one of the parallel texts and described different types of morphological ambiguities.

Apart from being used in learning environments, the corpus is also used in research. One example is the article "Rendering evidential meanings in Turkish and Swedish" (Csató 2009), which examined the Turkish evidential category of indirectivity and the less grammaticalized or lexical strategies used in Swedish to render evidential nuances. The description of the strategies used in the two languages was complemented with an analysis of data from one of the parallel Turkish-Swedish texts. It was found that although Swedish has means to express evidential

nuances, they were used much less in the Swedish translations than expected. The article describes several possible reasons for this. One is that the Turkish category allows three different types of reading. This ambiguity is significant in certain texts. The Swedish devices can render a particular evidential nuance, but not the whole range of ambiguity of the Turkish forms.

Turkic-Turkic corpora

The development of parallel Turkic-Turkic corpora is a project we would like to carry out in an international collaboration. Our corpus includes a Turkmen novel, Ak Welsapar's *Kepjebaş*, in the original and in Turkish and Swedish translations. The texts have been entered into the database and the sentence alignment is currently underway. The Turkic-Turkic parallel corpora would mostly serve Turcological research interest. Turkic languages are relatively similar, and their morphologies and syntax can easily be annotated in a coordinated way. This raises the need to implement some measures of standardization with regard to transliteration/transcription, and annotation.

Part 2. Universal Dependencies

The annotation of the Uppsala Parallel Corpus as described above was based on de facto standards in 2010. Since then, new annotation standards have been developed for the included languages, especially concerning the syntactic annotation in terms of dependency structures. A recent development is our effort to contribute to the Universal Dependencies (UD) for Turkic languages project. UD aims to develop cross-linguistically consistent treebank annotation for many languages, that is, syntactically annotated corpora that can be used in natural language processing and linguistics, and that allow for meaningful comparisons across typologically diverse languages.

As far as Turkish is concerned, there are two existing treebanks: METU-SABANCI and ITU-METU-SABANCI. Both use dependency structure annotation and the latter builds on the former. Dependency parsing of Turkish has been studied by Eryigit et al. (2008); part-of-speech tagging for Turkish has been investigated by Dincer et al. (2008); and morphological processing of Turkish has been pursued in the work of Sak et al. (2011). Apart from Turkish, work on morphological tagging has been carried out for: Uyghur (Altenbek 2006) and Kazakh (Makazhanov et al. 2014). The listed works target the treebanks and the annotation methods but none have used Universal Dependencies.

The UD subproject has two goals. The first is to investigate the methods for implementing UD for Turkish and provide guidelines. The second is to help with the ongoing development of conversion methods from the ITU-METU-SABANCI treebank to the UD scheme. In order to discuss the issues of conversion and methods for implementing UD to Turkish, a workshop has been arranged involving the team(s) at ITU (Istanbul Technical University) and Francis Tyers from The Arctic Universi-

ty of Norway and interested parties at Uppsala University. We hope that in the future, the discussions will be broader and the UD scheme will be extended to Turkic languages in general.

What is Universal Dependencies?

Universal Dependencies (UD) is an annotation scheme the main goal of which is to provide guidelines for consistent annotation of similar constructions across languages while also allowing language specific extensions, where necessary. Currently 33 languages have been annotated according to the UD scheme and work is ongoing on Turkish and Kazakh. The goals and characteristics of UD have been discussed in Nivre (2015). The following brief introduction to Universal Dependencies is based on Nivre (2015) and the UD documentation.[1]

In order to achieve the goal of cross-linguistically consistent treebanks, UD works on different layers. These are tokenization, morphology and syntax. Words are the basic units of annotation in UD and they are considered as syntactic words. The goal of tokenization is to segment the words between which the dependency relations hold. Thus, syntax represents the relations between words. Each word is either a dependent of another word in the sentence or the root of the sentence. The morphological specification of a syntactic word or a unit in the UD scheme consists of lemma, part-of-speech (POS) tag, and a set of features. The lemma represents the semantic content of a unit. The POS tag represents the abstract lexical category associated with the unit. Finally, the features represent the lexical and grammatical properties that are associated with a particular word form or lemma. Every feature has the form Name=Value where Name is the feature name and Value is the value of the feature. If a feature is not mentioned in the data, this implies an empty value. When there is more than one feature for a word, they are ordered alphabetically. A fully annotated text in UD contains the right features and categories of dependency relationships between segmented words.

Application of Universal Dependency (UD) to Turkish

The first part of this section will list the problems and possible solutions related to the verbal lexical category of Turkish when we implement the UD scheme. We have chosen first to implement the scheme for one lexical category and make an exhaustive study of it rather than giving brief descriptions of each lexical category. It seems appropriate to start with verbs. In the second part we will report on the decisions made during the workshop at Uppsala University regarding a conversion method for the ITU-METU-Sabanci treebank.

1 For the documentation on Universal Dependencies together with the languages that are annotated with this scheme; see http://universaldependencies.org/

Universal Dependency annotation of Turkish verbs

We will here deal with the subcategorization of verb forms. We first deal with finite verb forms and then with non-finite verb forms without their nominal morphology.

Finite verbs

The following morphological issues, which are problematic for the annotation schemes will be presented: the segmentation of an orthographic word, multiple voice, multiple tense, multiple aspect and multiple modality.

Segmentation of an orthographic word: In UD, clitics are separated from their host and treated as separate words in most but not all cases. We will discuss whether the clitics in Turkish should be separated or not. We will look at three types of clitics, namely; (i) the clitics *idi* (denoting past tense), *imiş* (denoting indirectivity/evidentiality), *ise* (denoting hypothetical mood); (ii) the generalized modality marker *-DIr* which conveys different modal notions such as presumption; and (iii) the question particle *mI*.

The clitics in (i) and (ii) add grammaticalized meanings to the proposition. UD represents the features of words (or units) in the form of a morphological feature list. We choose not to split these types of clitics because they can be represented in the morphological features of the host as a feature value.

The third type of clitic, the question particle, is orthographically separated and the existing treebanks consider it as a separate token. We propose that it can be kept separate from the verb and have an *aux* dependency relation with the verb. Consequently, it carries the feature "question" and in cases where it bears any agreement feature, these can be represented in its feature list.

Multiple voice: Turkish allows verb to carry more than one voice morpheme. The reflexive suffix is unproductive and the reciprocal suffix combines with very few stems. Thus they can be viewed as separate lexical entries. The double passive usage is also very limited and, when present, is perceived as a single passive. The extreme cases of multiple voice are observed when there are multiple causative voices or when a causative and passive suffix are observed together, example (1). When some features are marked more than once on the same word, UD suggests the usage of *layered features.*[2] We propose to apply the layered features idea to multiple voice marking in Turkish verbs. In this case, the voice features of example (1) will be: Voice[1]=Caus|Voice[2]=Pass. Note that the numbering in the brackets indicates multiple layers of the feature. However, the numbering does not necessarily indicate an ordering between the morphemes.

2 *Layered features:* According to the UD documentation; "when some features are marked more than once on the same word, it is said that there are several layers of the feature". This type of representation has been implemented for Basque verbs in the agreement paradigm.

(1) *Giy-dir-il-di.*
wear-CAUS-PASS-PAST-3SG
'It was made put on.'

Multiple tense marker: The only tense that is marked by the morphemes is the past tense morpheme *-DI* (TAM1) or clitic *idi* (TAM2). In the absence of the past tense clitic *idi*, the TAM1 morphemes have other meanings than tense. As a result, the multiple tense issue is resolved trivially.

Multiple aspect/viewpoint: No two morphemes come together in such a way that both represent aspect/viewpoint syntactically thus cause a multiple aspect/viewpoint problem in the Turkish verbal system. Postverbs that mark actionality modification, such as *-(y)Adur-* marking durativity or repeatedness, *-(y)Iver-* marking 'to do something uncontrolled or quickly', should be treated as derivations.

Multiple modality: In the current version[3] of UD, evidentiality is analyzed under modality. However, for Turkic languages evidentiality has to be analyzed as a separate feature just like tense, aspect and modality. We propose that an evidentiality feature can/should be added to the feature list in UD to cover languages like Turkish that mark evidentiality separately.

Another issues is the marker *-(y)Abil*, originally a combination of a converb and a postverb, but which has been further grammaticalized as marker of possibility/probability/capability. Possibility can also be combined with the necessitative, presumptive, and hypothetic markers. The negated form of the possibility marker can be combined with its affirmative form; see (2).

(2) *Birinci ol-ama-yabil-ir.*
first be-NEGATED.POSSIBILITY-POSSIBILITY-AORIST3SG
"S/he might not be able to be first."

Each case can be resolved in a similar fashion as multiple voice marking, by using layered features. The feature representation of example (2) is Mood[1]=Pot|Mood[2]=Pot.

To sum up, we have gathered all of the proposals that we have made in this section in Table 1. On the left side, the issues are listed, and on the right side, under Comments, our suggested proposal within the UD scheme can be found.

3 Version 1.2.

Issues on Finite Verbs		Comments
Segmentation of clitics	Clitic Type	
	TAM2	Not segmented
	Presumptive	Not segmented
	Question Particle	Orthographically segmented. It has an aux-dependency relation with the host
Tense	Morpheme(s)	
	-mIş or *imiş*	These forms refer to evidentiality (Turkic indirectivity). A new feature that represents this is needed.
	-(y)AcAktI	The combination of the two morphemes corresponds to future past tense (English)
Actionality (lexical aspect)		Lexical aspect should be viewed as derivation. In this case there is no multiple aspect.
Modality		Layered features
Voice		Layered features

Table 1. Summary of the proposal for finite verbs

Non-finite verbs

Between verb root and subordinating suffixes the following morphology could be found: voice, negation and mood/aspect. This means that the voice problem we have observed in finite verbs will be repeated for non-finite verbs. However, the layered feature solution is still applicable.

Another issue with the non-finite verbs relates to terminology. Studies on Turkic languages group the subordinating suffixes into three categories. These are: verbal nouns where the subordinated verb functions as predicate in a nominal clause, participles where it functions as predicate in a relative clause, and converbs where it functions as an adverbial clause.

The definition of transgressive in the UD documentation corresponds to that of the converbs. Since the goal of UD is not to define similar things differently, we will be using transgressive instead of introducing a new value under verbform. The basic syntactic dependency relationship for transgressives is the *advcl*. When there is a postposition, there should be a *case* relation between the converb and the postpositional element. Some of the complex verbal forms are treated as multiword expressions in the existing treebanks.

The existing treebanks treat *-mAk, -mA, -(y)Iş* as infinitives. We would like to preserve the insight of these treebanks and keep them separate from gerunds.

The issues related to non-finite verbs are summarized in Table 2, including other types of verbforms in this category.

Issues on non-finite verbs	Comments	
VerbForm	Terminology	
	Infinitive	*-mA, -mAk, -(y)Iş*
	Gerund (verbal noun)	*-(y)AcAk, -DIK*
	Participle	*-(y)AcAk, -DIK, -(y)An*
	Transgressive	converbs such as *-(y)Ip, -(y)ArAk* etc.
Voice	Voice problem is similar to that with finite verbs and is resolved by the same method.	

Table 2. Summary of the proposals for non-finite verbs

Conversion

A workshop has been arranged to discuss the details of the conversion of the existing ITU-METU-Sabancı treebank to the Universal Dependency scheme. The workshop was organized by the Department of Linguistics and Philology, Uppsala University in 26–27 November 2015. The final version of the conversion will be contributed by the team at Istanbul Technical University, who are working on the software that converts each particular treebank scheme to the UD format. The following are the results of the discussions.

During the workshop, mainly the syntax-related issues were discussed. These involve the types of dependency relationships. Some differences can be solved by changing labels. PREDICATE, PUNCTUATION, DETERMINER, COORDINATION and, Intensifier in ITU-METU-SABANCI treebank are re-labeled as *root*, *punct*, *det*, *conj* and *advmod*, respectively.

The basic difference between the annotation scheme used in the ITU-METU-SABANCI treebank and the UD scheme is that UD makes a distinction between clausal and non-clausal dependency relations. For example, a MODIFIER in ITU-METU-SABANCI annotation may correspond to a clausal adjective or just a lexical adjective. In light of this the following conversion table, Table 3, has been agreed upon. Under the ITU-METU-SABANCI label, the names of dependency relations are listed as they are found in the treebank. Under the UD label are the corresponding relations in the UD listed. Each label in the ITU-METU-SABANCI treebank has more than one corresponding UD label. The definition of the relation determines what type of UD label will be used.

The ITU-METU-SABANCI annotation method uses umbrella terms which correspond to more detailed descriptions in UD. APPOSITION, POSSESSOR, ARGUMENT, MODIFIER and VOCATIVE in ITU-METU-SABANCI treebank correspond to a variety of relations in UD.

ITU-METU-SABANCI label	UD label	Definition of the relation
RELATIVIZER	Ccomp	Complement clause
	advcl	Adverbial clause
	acl	Adjective clause
MODIFIER	acl	Adjective clause
	amod	Adjective
	advcl	Adverbial clause
	advmod	Adverb
OBJECT	ccomp	Complement clause
	dobj	Direct object
SUBJECT	csubj	Subject clause (Any clause that functions as the subject)
	nsubj	Subject

Table 3. Conversion table

UD and Turkic

Applying the UD scheme to Turkic languages is a new movement. There are ongoing projects with Turkish and Kazakh. Makazhanov et al. (2015) is an example of applying UD to a Kazakh treebank. Çöltekin (2015) lists the issues related to Turkish and annotation schemes. Again Çöltekin provides an annotation tool for UD and Apertium[4] has a Turkic lexicon which may be useful.

Most Turkic languages do not have a treebank, but this will change in the future. In this report, we have aimed to highlight issues that may be problematic in Turkish, and suggest methods to solve them in a cross-linguistic annotation scheme.

References

Altenbek, G. 2006. Automatic morphological tagging of contemporary Uyghur corpus. *Proceedings of the 2006 IEEE International Conference on Information Reuse and Integration*. 557–560.

Apertium—Turkic lexicon. http://wiki.apertium.org/wiki/Turkic_lexicon# (accessed 19 March 2016).

Bergdahl, E. A. 2006. Shadow. From a relaxing spot to darkness and death. A semantic study of how the word shadow is used in Swedish and Turkish. [Term thesis]. Batı Dilleri ve Edebiyatı Bölümü, Bogaziçi University & Department of Linguistics and Philology, Uppsala University.

Çöltekin, Ç. 2015. A grammar-book treebank of Turkish. In: Dickinson M. et al. (eds.) *Proceedings of the 14th workshop on Treebanks and Linguistic Theories (TLT 14)*. 35–49.

Çöltekin, Ç. 2016. *Brat Annotation Tool*. http://coltekin.net/cagri/ud/#/ (accessed 19 March 2016).

4 http://wiki.apertium.org/wiki/Turkic_lexicon#

Csató, É. Á. 2009. Rendering evidential meanings in Turkish and Swedish. In: Csató, É. Á. et al. (eds.) *Turcological letters to Bernt Brendemoen*. Oslo: Novus. 77–86.

Csató, É. Á. & Nathan, D. 2003. Turkish suffix dictionary. http://www.dnathan.com/language/turkish/tsd

Dadasheva, Sabina 2005. Den turkiska indirektiva kategorin. En undersökning av återgivningen av den turkiska indirektiva kategorin i ryska och svenska autentiska översättningar. [Term thesis.] (In Swedish.) Department of Linguistics and Philology, Uppsala University.

Dincer, T. & Karaoglan, B. & Kisla, T. 2008. A suffix based part-of-speech tagger for Turkish. *Information Technology: New Generations, 2008. ITNG 2008,* 680–685.

Eryiğit, G., J. Nivre & K. Oflazer. 2008. Dependency parsing of Turkish. *Computational Linguistics* 34, 3.

Göksel, A. & Kerslake, C. 2005. *Turkish*. London: Routledge.

Haktanır, M. 2006. Orhan Pamuk'un Beyaz Kale adlı eserinde çok anlamlılık. [Term paper] (In Turkish.) Department of Linguistics and Philology, Uppsala University.

Hedman, M. 2009. Verbet 'göra' i svenska och turkiska. [Term thesis.] (In Swedish.) Department of Linguistics and Philology, Uppsala University.

Halácsy, P. & Kornai, A. & Oravecz, Cs. 2007. Hunpos—an open source trigram tagger. In: *Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics, Companion volume, proceedings of the demo and poster sessions, Prague, Czech Republic.* 209–212.

Makazhanov, A. & Yessenbayev, Z. & Sabyrgaliyev, I. & Sharafudinov, A. & Makhambetov, O. 2014. On certain aspects of Kazakh part-of-speech tagging. *Proceedings of IEEE AICT 2014*. 1–4.

Makazhanov, A. & Sultangazina, A. & Makhambetov, O. & Yessenbayev, Z. 2015. Syntactic annotation of Kazakh: Following the Universal Dependencies guidelines. A report. In: *Proceedings of TurkLang 2015*. Kazan: Academy of Sciences of the Republic of Tatarstan Press. 338–350.

Megyesi, B. 2008. The open source tagger HunPoS for Swedish. Department of Linguistics and Philology, Uppsala University.

Megyesi, B. & Sågvall Hein, A. & Csató, É. Á. 2006. Building a Swedish-Turkish parallel corpus. In: *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC'06).* Genoa, Italy.

Megyesi, B. & Dahlqvist, B. 2007. A Turkish-Swedish parallel corpus and tools for its creation. In: *Proceeding of Nordiska Datalingvistdagarna*, *NoDaLiDa* 2007.

Megyesi, B. & Dahlqvist, B. & Pettersson, E. & Nivre, J. 2008. Swedish Turkish parallel treebank. In: *Proceedings of the Sixth International Conference on Language Resources and Evaluation* (LREC'08), Marocco. 470–473.

Megyesi, B. & Dahlqvist, B. & Csató, É. Á. & Nivre, J. 2010. The English-Swedish-Turkish parallel treebank. In: *Proceedings of Language Resources and Evaluation (LREC 2010)*, May 2010.

Nivre, J. 2008. Treebanks. In: Kytö, M. & Lüdeling, A. (eds.) *Corpus linguistic: An international handbook.* Berlin: Mouton de Gruyter. 225–241.

Nivre, J. 2015. Towards a Universal Grammar for natural language processing. In: Gelbukh, A. (ed.) *Computational linguistics and intelligent text processing. 16th International Conference, Cairo, Egypt, April 14-20, 2015, Proceedings* 1. Heidelberg: Springer.

Nivre, J. & Hall, J. & Nilsson, J. 2006a. MaltParser: A data-driven parser-generator for dependency parsing. In: *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC)*. 2216–2219.

Nivre, J. & Hall, J. & Nilsson, J. 2006b. Talbanken05: A Swedish treebank with phrase structure and dependency annotation. In: *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC)*. 1392–1395.

Oflazer, K. 1994. Two-level description of Turkish morphology. *Literary and Linguistic Computing,* 9: 2.

Oflazer, K. & Say, B. & Hakkani-Tür, D. Z. & Tür, G. 2003. Building a Turkish treebank. In: Abeillé, A. (ed.) *Treebanks: Building and using parsed corpora*. Kluwer. 261–277.

Sak, H. & Güngör, T. & Saraçlar, M. 2011. Resources for Turkish morphological processing. *Language Resources and Evaluation* 45: 2, 249–261.

Sulubacak, U. 2015. *ITU Treebank annotation guide* v2.7. Istanbul Technical University.

Sulubacak, U. & Eryigit, G. (to be published). A redefined Turkish Dependency Grammar and its implementations: The revised Turkish treebank and a new Turkish web treebank.

Taylan, E. E. 2014. The structure of Modern Turkish. Morphology. [Class Notes]. Boğaziçi University.

Universal Dependencies documentation. http://universaldependencies.org/ (accessed 19 March, 2016).

Tiedemann, J. 2003. Recycling translations. Extraction of lexical data from parallel corpora and their applications in Natural Language Processing. PhD Thesis. Uppsala University.

Tiedemann, J. 2012. Parallel data, tools and interfaces in OPUS. In: *Proceedings of the 8th International Conference on Language Resources and Evaluation (LREC'2012)*.

Uplug documentation and download: https://bitbucket.org/tiedemann/uplug (accessed April 29, 2016).

Uyghur, D. 2009. Semantiken av turkiska verbet *ver-* 'att ge' och dess motsvarighet i svenska. [Term thesis.] (In Swedish.) Department of Linguistics and Philology, Uppsala University.

Yuret, D. & Türe, F. 2006. Learning morphological disambiguation rules for Turkish. In: *Proceedings of HLT NAACL'06*, New York. 328–334.

Reviews

Henryk Jankowski: Review of Raihan Muhamedowa (ed.) *Kazakh in Post-Soviet Kazakhstan. Proceedings of the International Symposium on Kazakh, November 30–December 2, 2011, Giessen.* (Turcologica 102.) Wiesbaden: Harrassowitz, 2015. 175 pp. ISBN 978-3-447-10440-1.

Henryk Jankowski, Katedra Studiów Azjatyckich, Uniwersytet im. A. Mickiewicza, al. Niepodległości 24, 61-714 Poznań, Poland. E-mail: henko@amu.edu.pl

The reviewed volume contains 12 selected papers read at the symposium *Kazakh in Post-Soviet Kazakhstan* held in 2011 at the University of Giessen, Germany. The papers were originally presented in Kazakh, Russian, German and English, but all published in English. While some papers were submitted as read at the symposium, others were supplemented with new data and references as far as 2014. The editor decided not to unify transliterations. As a result, we have various transliteration systems and Kazakh in its original Cyrillic script. The editor divided the articles into six sections: Language Planning, Language Contact, Language and Identity, Latinization of the Alphabet, Terminology, and Varieties of Kazakh. The section Language Planning contains two papers, Language Contact contains one, Language and Identity and Latinization of the Alphabet contain two each, Terminology contains four, and Varieties of Kazakh one.

The volume opens with Eleonora Suleimenova's article "The vitality of the Kazakh language and language planning" (pp. 11–25). This paper comprises seven sections, including an introduction. The first section following the introduction discusses different approaches to language vitality, one of which is hierarchy. Suleimenova places Kazakh at the low hierarchical scale and gives reasons for this. Later in the paper, the author also points to historical reasons rooted in the USSR (p. 14). The following section discusses demographic indicators of the vitality of Kazakh. In general, both the number of ethnic Kazakhs and the competence in Kazakh show a growing tendency after the dissolution of the Soviet Union. Later in the article the author evaluates proficiency in Kazakh. As Suleimenova stresses, various data are presented in various studies and these data are mostly incomparable (p. 15). She argues that low figures provided in some studies are based on beliefs and impressions rather than good analytical grounds, and according to her survey the share of Kazakhs who do not know Kazakh at all is not higher than approximately 2%. Therefore, in Suleimenova's view low proficiency in Kazakh poses less of a threat to the language than often perceived (p. 16). However, we have to add that Suleimenova's arguments are in favour of low proficiency, and the percentage of the Kazakhs with high proficiency level Kazakh is in fact low. Concerning language of instruction, Kazakh is increasingly used as the language of instruction in all types of schools. Despite this, the figures are not good for Kazakh. One third of students still

prefer instruction in Russian (p. 18). Suleimenova then examines the strength of Kazakh in various domains of communication. Also in this respect, the vitality of Kazakh is low (p. 18). In conclusion the author tries to justify the situation of Kazakh in modern Kazakhstan by various well-known factors, such as multilingualism, minorities, the power of Russian and influence of Russian-speaking population. Unfortunately, Suleimenova does not give a straightforward answer to the question whether language planning in Kazakhstan is successful or not. Despite this, the careful reader sees that the answer must be negative. At least, if we assess language policy from the viewpoint of declarations and measures taken, we see that the progress is slow. In fact, Suleimenova says implicitly what many other authors have already observed, namely that the enhancing of the role of Kazakh was more successful in the symbolic sphere than in practice. There is one observation that in my opinion is of particular importance. This concerns the growth of English and Russian in many spheres of life that "cannot be regulated by the state", e.g. science, consumption, fashion, high technology, popular culture and music, the Internet, satellite television, and mobile and network communication. In my opinion, Kazakh authorities missed the opportunity to make Kazakh a full-fledged official language in Kazakhstan before these factors had become powerful. Now it may be too late.

Some bibliographical items in Suleimenova's article are in wrong alphabetical order, e.g. Artygalin (2002) and (2003), and the items "Sbornik" from 2001 to 2008 have been placed after Spolsky (2004). A few studies referred to in the article, e.g. Suleimenova (2011a) and Suleimenova (2011b), are absent from the bibliography. Lastly, I might add that the low vitality of Kazakh in science is also evident from Suleimenova's article. Out of 49 items published in Kazakhstan shown in the bibliography only two (4%) are in Kazakh, one being Kazakh-Russian bilingual, all the others in Russian.

The second article in this section is Kussaiyn Ryssaldy's "Problems of the Kazakh language as the state language in modern Kazakhstan" (pp. 27–33). This article, as its title shows, is not exclusively devoted to language planning; it assesses the present state of the language, and the achievements and failures of Kazakh language policy. The author starts with providing a historical background (p. 27) and proceeds to present some legal acts regulating the status and the spheres of use of Kazakh (p. 28). He observes that Russian is used not only in the spheres stipulated by the constitution, but also in many other areas. Ryssaldy's article contains a few postulates about the development of Kazakh, such as the need to shift to the Latin alphabet and to improve language status. The position of Kazakh, as shown by the author is rather weak, especially in the media. For example, he says that there are only two Kazakh-language television channels out of approximately 100, and only two radio stations that broadcast exclusively in Kazakh out of 12 (pp. 29–30).[1] The situation is also

1 In addition, one of these two radio stations also broadcasts news and ads in Russian, so it is not exclusively Kazakh.

gloomy as regards the newspapers and periodicals. Although Ryssaldy emphasizes the role of two weeklies, *Qazaq Tili* 'Kazakh language' and *Qazaq Ädebiyeti* 'Kazakh literature', it should be added that they have a very limited impact on the public sphere. The former is distributed in 26,000 copies only, which means a great fall in relation to the 1990s when it was distributed in 100,000 copies (http://anatili.kazgazeta.kz/?page_id=8638). The latter has an even smaller distribution, not exceeding 15,000 copies. Ryssaldy points to some problems with language related antagonism saying that the language issue is dividing the population of Kazakhstan into "Kazakhs" and "Kazakhstanis" (p. 31).

Ryssaldy's paper is written in a patriotic tone, as is clear from the author's claim that "Kazakh is one of the richest languages in the world" (p. 32). In contrast to Suleimenova's article, the references mostly include studies and articles written in Kazakh. We can show some weaknesses of transliteration, e.g. "Qimanova" (p. 32) versus "Qiymanova" (p. 33) and "muwra" (p. 33) 'heritage', the correct form being *mura* [mụra].

Juldyz Smagulova's "Language shift and maintenance: Primary language socialization in urban Kazakh families" (pp. 37–49) is an article in the section Language Contact. It presents extracts of two talks by a grandfather to his grandson, one registered when the baby was 12 months old, the other when he was a 26-month-old toddler. While the earlier talk was in Kazakh, the later was in Russian, except for the form of address *balam* 'my child' to the baby. The author searches for the reason for the switch into Russian, which in her opinion is typical of middle class urban Kazakh families. She finds that the basic reason is that Russian is associated with "positive and socially valued aspects of the self" or with social status and being adult. At the same time Smagulova shows that the talk in Russian is more complex both lexically and grammatically than the talk in Kazakh, which is mostly composed of simple and repeated utterances (pp. 42–43). Naturally one may expect the author to supply evidence for the typicality of such situations.

I have two remarks on this otherwise interesting article. First, the word *aytšï* 'tell (me)' is not a typical form used in talking with pre-verbal children as Smagulova claims (p. 47), but a normal expression used in many situations without restrictions, except for the restriction on the politeness register. Second, the Russian form *nosok* 'sock' is unusual. To the best of my knowledge, this Russian loanword in Kazakh is only used in the plural form *noski*, mostly pronounced [næskị], even with the numeral 'one', i.e. *bir noski*.[2]

Ruth Bartholomä's article "Between 'duty' and 'prestige': The Kazakh language in the discourse of contemporary Kazakhstan" (pp. 53–67) is the first of two papers

2 Since this is a sub-standard Russian borrowing, it is not listed by dictionaries, but it is evidenced on websites, e.g. https://vk.com/public100179203, while **bir nosok* could not be found in the Internet resources. Therefore, it may be argued that *nosok* is only used by the Russian-speaking Kazakhs.

in the section Language and Identity. It evaluates the axiological force of some slogans about the Kazakh language which have figured in the public discourse in recent years. The author concentrates on official documents aiming to shape language policy as well as how they are reflected in print media. As for the latter, only the titles of some articles are analysed, but Bartholomä is right when she says that the title of an article has special informative and expressive force. The author's basic aim is to examine what functions the creators of language policy ascribe to language in terms of duty and prestige. The seven most often repeated themes are analyzed, quoted from documents and articles written in Kazakh, Russian, and both: (1) Kazakh as the language of the Kazakh ethnic group; (2) Kazakh as the state language; (3) Kazakh as a factor/guarantee of unity; (4) Kazakh as the language of all Kazakhstanis; (5) speaking Kazakh as a duty and obligation; (6) Kazakh as a language of prestige; and (7) Kazakh as a language of the future. Some other themes of slogans are also pointed out, e.g. that Kazakh is/should/must be the fate of the nation, or that without its language the country/nation has no future. Bartholomä stresses that the first theme was avoided in official documents, but often promoted by patriotic press. In the author's opinion, these ambitious aims of Kazakh language policy can hardly be achieved unless decisive measures are taken to implement them (p. 66).

Bartholomä's article is an important voice in the discussion about Kazakh language policy, since it examines a few ideas on the basis of official documents and in some influential newspapers. It cites the sources in the original Kazakh alphabet. From a few titles and names provided in transliteration we see the infeasibility of the system applied; e.g. *Eluwbay* (p. 56) should be transcribed *Elüwbay*.

The next article in this section is "Kazakhstan and Tatarstan—building identities in Russian-Turkic speech communities" by Aksana Braun and Monika Wingender (pp. 69–94). It consists of six sections including an introduction and conclusion. The aim of this article is to compare the situation in Tatarstan and Kazakhstan. Both Kazakhstan and Tatarstan share many characteristic features, but there are some remarkable differences between them. Most importantly, Tatarstan is part of the Russian Federation and has two state languages, and the size of its Russian population is higher than that in Kazakhstan. However, in both countries there is an asymmetric bilingualism, the Russians being predominantly monolingual. Even in Kazakhstan the law on languages is differently assessed by the two groups, but both the Kazakhs and the Russians regard it as unfavourable. The authors demonstrate that despite the loss of many functions by Russian, the titular languages cannot function as state languages (p. 70). They seek answers to the questions of how successful the current language policies are and what measures should be taken to improve the position of titular languages. However, they stress that Russians in the speech communities in post-Soviet states and some other Turkic-Russian areas in Russia are the minority per political definition, but the majority per language spread and prestige (p. 71). The next section overviews past and current research, while the two subsequent ones examine questions of identity in Kazakhstan and Tatarstan, respectively. The last section is devoted to the current discussion in both speech communities.

Naturally, the conclusion is clear. There are many similarities between Kazakhstan and Tatarstan, but Tatarstan is faced with more constraints due to its status of being subject to the laws and practices imposed by Russian Federation. Concerning Kazakhstan, the authors discuss Nazarbayev's idea of the Kazakhstani nation, much criticized by the Kazakhs, as well as by some Russians (pp. 76–77). The same is true of the president's doctrine of Kazakh, Russian and English trilingualism. In contrast to Kazakhstan, Tatarstan has no official concept of state identity (p. 81). Russian federal law sometimes contradicts Tatar republican law, which is unfavourable to the Tatar national culture and language.

This article was written by competent authors who are keen observers, and one of them also has a good academic background.

The section "Latinization of the Alphabet" begins with Damina Shaibakova's article "'Clothing or costume'? A further [sic] article on the graphics of the Kazakh language" (pp. 97–106). Shaibakova gives some details on the application of Cyrillic script to Kazakh and Bashkir, and characterizes the structural and practical features of all three alphabets she deals with: Arabic (also called "Arabian lettering", p. 98), Latin and Cyrillic. She reminds the reader that the adoption of Latin was a compromise between Cyrillic, regarded as technologically backward, and Arabic, considered to be a remnant of medieval Muslim obscurantism (p. 99). She presents the arguments for the Cyrillic, Latin, and even the Arabic alphabet put forward by various groups. For her part, the author advocates Cyrillic (p. 105), though one can hardly agree with her vague reasoning: "Education favours the Cyrillic literature".

The remarks to this article are the following. Firstly, "Sotsiyalistik Qazaqstan" was not published in two alphabets at the same time. In fact, this newspaper was first called *Satsïyaldï* (later *Sotsïyaldï*) *Qazaġïstan* and only in 1938 was renamed. Like all Kazakh publications at that time, it was published in Latin script, and from 1939 onwards in Cyrillic. Quite interestingly, all titles provided in the references to this article are in Russian; there is not a single paper in Kazakh or any other language, though the title of one of them must have been translated into Russian, because it was published in *Ana Tili*, a monolingual Kazakh weekly. This is also an example of the preference for Russian as a more prestigious language in the article of a Kazakh researcher who works on Kazakh.

The second paper in this section is "The issue of alphabet latinization in independent Kazakhstan" (pp. 107–118) by Barbara Kellner-Heinkele and Jacob M. Landau. In the short introduction the authors stress that few languages—the exceptions being Turkish and Hausa—have successfully carried out the change of alphabet (p. 107). After the introduction, they provide an overview of the events of alphabet change in the post-Soviet Turkic states, underlining the well-known fact that three post-Soviet Turkic states, Azerbaijan, Turkmenistan and Uzbekistan, shifted to Latin, while Kazakhstan and Kyrgyzstan did not (p. 109). The authors then proceed to discuss the pros and cons of a prospective change in Kazakhstan. It is clear that here is not sufficient political and ideological support for the change in Kazakhstan, since the other conditions are fulfilled (or rather were fulfilled at the time of writing

of the article, for the economic situation deteriorated due to falling oil prices). The authors present the arguments put foward by supporters and opponents of the change, focusing on the president and his officials who have "all the cards in their hands" (p. 114). However, some other institutions' and prominent figures' initiatives and opinions are also described. The article also discusses the prospects for the co-existence of the two alphabets, since such a strange proposal has been put forward.

This is a very good article written by authors of great erudition and experience. Kellner-Heinkele and Landau stress that "President Nazarbaev's imprint on language policies in independent Kazakhstan has been decisive throughout" (p. 109). However, they provide examples to demonstrate that the president constantly moves forward and backward on the question of alphabet change (pp. 110–111), as well as other aspects of language policy. In general, there are many pronouncements in support of the titular language of Kazakhstan, but little has been done to implement an effective language policy favouring Kazakh.

Köbey Khusaiyn's article "New words in the modern Kazakh language" (pp. 121–127) begins the largest section in the book, which is "Terminology". He demonstrates basic ways of forming new terms, such as adopting foreign words from western languages via Russian, adopting Russian words, adopting Arabic, Turkish and Persian words (though no Persian words are listed in the examples, p. 122), and forming new words and terms from the internal resources of Kazakh. Khusaiyn gives a brief outline of the trends in terminology from the 1920s onwards, showing neologisms from all periods. We can add that some of the first terms have fallen into disuse, e.g. *šaytanarba* 'bike', while others have survived (e.g. Baytursïnulï's linguistic terminology) or have recently been revived, e.g. *töraġa* 'chairman' (p. 123). Khusaiyn says that in 1935 a Soviet congress decided that new terms must be borrowed from Russian instead of other languages, and that Russian words must be written as in the original. This decision regulated the ways of creating new terminology for decades and still exerts influence (p. 124). Khusaiyn gives examples of various techniques of creating new terms, such as compounding (e.g. *beynežazba* 'video recording') or derivation (e.g. *muzdatqïš* 'refrigerator'). The word for 'vacuum cleaner', *šaŋsorġïš* (p. 125), is an example of both compounding and derivation (< *šaŋ* 'dust' + *sor-* 'to suck' + *-ġïš*).

One statement in this article should be corrected. It is not true that "the Kazakh language gained official status according to the Constitution and the 'Law on the Languages' of 1997" (p. 121). This status of Kazakh was stipulated in 1989, confirmed in 1990, to be achieved by 2000. Kazakh was declared the state language of the Republic of Kazakhstan by the first Constitution promulgated in 1993, while Russian was defined as the language of interethnic communication. This status of Kazakh was confirmed by the second Constitution of 1995, with Russian being raised to the status of the second official language. Therefore, this occurred a few years before 1997. I am sure that this mistake in Khusaiyn's article stems from an inexact translation. There is another point that also should be commented. It concerns the Kazakh word *därihana* 'pharmacy' (glossed as 'drugstore') which Khu-

saiyn claims was coined by Seifullin (p. 126). This word needs deeper analysis, for it is present in all Central Asian Turkic languages except Turkmen (more precisely, we do not have good evidence from Turkmen), cf. Kirghiz *darïqana*, Uzbek *darixona*, and Uighur *darixana* (cf. Jankowski 2010: 136).

Khusaiyn's article is followed by Sherubay Kurmanbaiuly's paper "The development of Kazakh terminology, 1990–2010" (pp. 129–136). Kurmanbaiuly traces the development of Kazakh terminology during the 20 years of independent Kazakhstan, but also provides a short historical background. He stresses that Kazakh terminology began developing in the second half of the 19th century. Therefore, the following periods are established: (1) –1910, (2) 1910–1930, (3) 1930–1990), (4) 1990–2010. Naturally, these periods can be debated, especially the choice of 1910 and 1930 as dividing points seems to be arbitrary. In the following, Kurmanbaiuly shows increas in dissertations and studies on terminology during the last period, compared with the Soviet period of 1930–1990. The final part of the article is devoted to the creation of new terms. The author illustrates his ideas with numerous examples of new terms, old terms revived, and phonetically adapted Russian loanwords—not necessarily special terms—; e.g. *mäyki* 'T-shirt' < Russian *majka*, *silesir'* 'metalworker' < Russian *slesar'*.

Some forms provided in transcription are odd, e.g. *žasauwïl* 'executor of instructions' (p. 133). *Tintuwir* 'keyboard mouse' (p. 134) should be transcribed *tintüwir*. It should be noted that, unlike the other articles in the volume, this one does not contain any references.

The third paper in this section is Gulmira Madieva's "Language policy in the regulation of proper names in the communicative space of Kazakhstan" (pp. 137–145). This article is very important, for few international papers have dealt with the changes in proper names from the viewpoint of language policy. The author provides some figures to illustrate the volume of changed proper names, basically place names (pp. 139–140), and also examines the question of the emergence of new names (p. 141). Madieva presents the principles for the conversion of Kazakh proper names into Russian ones, which were employed in Soviet times and still are in force.

The transcription used by Madieva is quite inconsistent; it is rather a simplified Romanization of the Kazakh alphabet. For instance, in most names the vowel [ä] is rendered as a, which is not correct, since /ä/ and /a/ are distinct phonemes. However, in one case we read "Žämilya" (p. 143). Another inadequacy is the rendering of Kazakh [ɪ] (written as i) with y, e.g. *Aldaberdyqïzï* (p. 142), correctly *Aldaberdiqïzï*.

The last article in this section is Saltanat Rakhimzhanova's "Terminology work in contemporary Kazakh" (pp. 147–155). The article is divided into two parts, as a study on language policy should be, i.e. first dealing with planning, then implementation, and mentions two basic stages of language planning: status planning and corpus planning. In accordance with this principle, Rakhimzhanova discusses first legal documents related to terminology (p. 148), then terminology work in practice. Of the various legal acts and bodies, she allots the most space to the State Terminological Commission, which plays a key role in terminology work in Kazakhstan.

Thus we see the basic principles of creating and changing terminology. As in the other articles, Rakhimzhanova demonstrates the techniques applied in terminology, i.e. extension of meaning of existing words, word formation, reviving archaisms or formerly existing terms, adopting dialect words, adopting words from historical contact languages such as Arabic, Persian and Turkish, and others.

The author of this article employs the best and the most consistent system for transcription of Kazakh. Despite this, there are some minor weaknesses in the article; e.g. the verbs are written in various forms, once in the form of the verbal noun, as *tirkew* (p. 150) 'joining', and once in the form of the stem, *tur-* (p. 151) 'to stand up'. Regarding a new term for 'traffic jam' to replace Russian *probka*, I would add another term, *keptelis* to the author's *tïğïn*.

The last paper in the volume, in the section Varieties of Kazakh, is Aynur Abish and Éva Á. Csató's "Language attitudes and linguistic habits of young Kazakhs in Xinjiang" (pp. 159–175). As is evident from the title, this article does not deal with the main geographic region covered by the volume, but it is a very interesting paper which complements the other papers on the problems of bilingualism and is just as relevant to the Kazakh language as the remaining articles. Abish and Csató's paper is a small-size study, as they say, of situations which lead to a shift from Kazakh to Chinese. The authors demonstrate the rapid rise of bilingualism among young Kazakhs in China after the introduction of bilingual and monolingual Chinese education. The study examines three university students, two of whom attended bilingual Kazakh-Chinese schools, and a monolingual Chinese school, the latter also coming from a mixed Kazakh-Chinese family. Their present-day university education is monolingual Chinese. Abish and Csató prepared a questionnaire for these three students and analysed the results. All their work is firmly based on existing linguistic surveys and presents interesting answers to some key questions of language contact, bilingualism, code-mixing and language shift. One of its results is the confirmation of the great importance of education (p. 164). If children receive education in their first native language and become bilingual later, the bilingual contact situations do not threaten their native language. Another point worth stressing is that bilingual speakers tend to use the dominant language in more prestigious domains such as writing and counting.

In my opinion the argument that negative attitudes towards copying prevent speakers from using Kazakh (p. 165) should be used with caution, but in any event it deserves further study. Another remark concerns the choice of informants. I believe that selecting students who all have the same or similar family background would be more reasonable, especially in such a small study. It seems probable that someone with a mixed family will prefer the second language in some situations more than someone from a monolingual Kazakh family. It is interesting to note that the authors match the Chinese word *shūdiàn* with Kazakh *kitäpḳana* (p. 168, 171, 172) which is the word for 'library' (as translated in the example on page 165), while *shūdiàn* means 'bookstore' and should be translated *kitäp dükeni*. However, the local Kazakhs really call the single Kazkah bookstore in the region *kitäpḳana*.

In conclusion I can say that the book under review contains contributions of various levels and importance. One of them is a report rather than a scholarly article. Language is also an issue in some of the papers. As was demonstrated above, it is characteristic of some Kazakh scholars that they predominantly work with Russian, and do not use studies written in other languages, including Kazakh. In this respect, there are great differences between particular contributors. The most obvious issue is the lack of some relevant studies written in Kazakh. For instance, none of the articles on Romanization cite a very important edited collection by Wäli et al. (2007), published by Baytursïnulï Institute of Linguistics, in which some important historical documents have also been published. Despite some weaknesses, the book edited by Muhammedowa should be positively evaluated, since it makes available articles in English in a field where studies normally are written in either Kazakh or Russian and that therefore are inaccessible to many researchers.

References

Jankowski, Henryk. 2010. Orta Asya Türk dillerinin bazı ortak özellikleri. *Ankara Üniversitesi Dil ve Tarih-Coğrafya Fakültesi Türkoloji Dergisi* 17: 2, 131–142.

Wäli = Уәли, Нұргелді & Күдеринова, Құралай & Фазылжанова, Анар & Жұбаева, Орынай (eds.) 2007. *Латын графикасы негізіндегі қазақ әліпбиі: тарихы, тағылымы және болашағы*. Алматы: Арыс.

Ümit Deniz Turan: Review of Zeyrek, Deniz & Sağın Şimşek, Çiğdem & Ataş, Ufuk & Rehbein, Jochen (eds.) 2015. *Ankara papers in Turkish and Turkic linguistics.* (Turcologica 103.) Wiesbaden: Harrassowitz. ISBN: 978-3-447-10523-1, 676 pages.

Ümit Deniz Turan Anadolu University, Turkey E-mail: udturan@anadolu.edu.tr

Introduction

This reviewed volume is an edited collection of papers presented at the *Sixteenth International Conference on Turkish Linguistics* (ICTL), held at Middle East Technical University, Ankara 18–21 September 2012. It reflects the nature of the ICTL conferences, which bring together researchers with diverse interests in various topics in the linguistics of Turkish and other Turkic languages. The volume contains studies carried out within different theoretical and functional frameworks, with different methodologies and written by scholars from different countries. There are six papers on phonetics and phonology, eleven on syntax and morphology, eight on acquisition of Turkish as a first or second language, ten on discourse, semantics, and pragmatics, seven on language contact and sociolinguistics, ten on other Turkic languages, and five on prospective aspect in Turkic languages. In the following, the papers will be briefly summarized in the order in which they appear in the book, after which a general evaluation of the collection will be presented.

Brief review of the papers

Phonetics and phonology

Sıla Ay and İpek Pınar Bekâr in their study "A devoicing analysis of vowel [i] in voiceless consonant surroundings" carry out an experimental study with forty speakers of Standard Turkish, who read a list of CVC combinations, where the vowel is [I]. They found that the vowel is devoiced between almost all voiceless consonants, but maximally when it precedes or follows the voiceless bilabial plosive /p/. This result is consistent with those findigs in the literature. The authors suggest that it may be due to the fact that this consonant is a bilabial plosive.

Marcel Erdal in "Buffering, linking or latent consonant deletion" shows that Turkish makes use of three types of epenthesis. The first is grammatically motivated, i.e. when the pronoun is case marked, the latent coda /n/ is used between the vowels in the demonstrative pronouns *o* 'he, she, it', *bu* 'this' as in *ona* 'to him', *buna* 'to this', and in the reflexive pronoun *kendi* 'herself /himself'; as in *kendini*. The second type, phonetic epenthesis, is commonly used in Turkish to break vowel sequences by inserting one of the three approximant glides [w], [y] or [ɯ]. The third type of epenthesis is morphologically motivated. When a group of functional morphemes, e.g. *ile* (with), *iken* (while), *idi* (copula-past tense), become clitics, then [y]

is inserted as *-yle, -yken, -ydi*, respectively. This latent onset /y/ can be explained with reference to historical developments.

Senka Ivoşeviç and İpek Pınar Bekâr investigate the "Acoustic correlates of focus in Turkish". They study recordings of 21 native speakers of Turkish to collect elicited data from three types of focus constructions: broad focus (where the domain of focus is the whole sentence), informational focus (where the constituent that provides new information in the sentence is focused), and contrastive focus (where two entities are contrasted). The results show that word duration significantly increases in informational focus but not in broad focus. There is no significant difference between information and contrastive focus in canonical word order structures.

Beste Kamali in "Information structure in Turkish yes/no questions" investigates different types of yes/no questions and focus. The particle *mI* is placed right after the interrogated constituent in yes/no questions; the subject, object, adverb can be used with the particle *mI*, which is the interrogated part of the sentence. If the whole sentence is questioned, the question particle is placed at the end of the sentence, i.e. "verum question/verum focus". Change in word order is not necessary in yes/no question formation. Kamali discusses the differences between these different types of focus and shows that Turkish yes/no questions allow for diverse information structural interpretations.

Beste Kamali and Didem İkizoğlu in their study titled "Compound stress in Turkish is phrase stress" argue that compound stress in Turkish is the same as phrase level stress, contrary to previous analyses in the literature. After critiquing the previous literature, based on both impressionistic judgments and analyses of pitch tracks, they conclude that productive noun-noun and verbal compounds display prosodic properties of syntactic compounds and, therefore, take phrase stress. On the other hand, the least productive compounds take word stress. Their results suggest that phrase-level stress and lexical stress are sufficient to account for the stress rules of compounds in Turkish, depending on their productivity.

Mehmet Akif Kılıç's "Türkçe'deki ötümsüz sürtünmeli ünsüzlerin akustik özellikleri" 'Acoustic properties of voiceless fricative consonants in Turkish' is another contribution in phonetics. The author recorded ten native speakers of Turkish, who read words and non-words that contain voiceless fricatives in initial and final position in a carrier clause, such as *Ahmet __ dedi* 'Ahmet said __'. The recordings were analyzed in Praat 5.3.1.3 after which TextGrid files were formed. For statistical analysis SPSS was used. The results show that sibilant fricatives have higher amplitude and duration in Turkish and that the acoustic parameters of fricatives in final position are longer than those in initial position. The author suggests that the acoustic properties of voiceless fricatives in Turkish appear to be similar to those in other languages.

Syntax and morphology

Sinan Çakır in "Island constraints in Turkish: a grammaticality judgment study" examines whether there is an argument/adjunct asymmetry in *wh*-questions in Turkish. The data come from grammaticality judgment tests. The results show that there is indeed an asymmetry. Secondly, sentential adjuncts such as *why* are interpreted to be less grammatical than VP-adjuncts when extracted out of islands as *wh*-expressions. These results demonstrate that island constraints indicate divergent effects and that there is a continuum rather than a clear-cut dichotomy in the interpretation of island constraints.

Mevlüt Erdem in "Transitive verbal reflexives in Turkish: Synchronic and diachronic perspective" studies reflexive verbal markers and concludes that these do not have a single syntactic function or meaning. The author presents how the valency of verbs is influenced by the reflexive suffix—also based on a hierarchy of grammatical relations: Subject > D. Object > I. Object > Oblique.An argument NP that is deleted is always at the lower end of this hierarchy. In addition, the paper also addresses the historical changes pertaining to reflexives in Turkish.

Atakan İnce, Gülşat Aygen, and Özgür Aydın in "Copular structures as (non)phases" discuss an interesting property of equative copular structures in Turkish within the Minimalist Program. In Turkish equative copular structures, the copula can either agree with the subject DP or the second DP, i.e. the one in the predicate: *Sen Ali'sin* [you Ali-COPULA 2ND PERSON] 'You are Ali' or *Ali sensin* [Ali you-COPULA-2ND PERSON SINGULAR] 'Ali is you'. The authors claim that this can only occur with first or second person *phi*-features. They propose that in Turkish, T^{o} has a tendency to agree with a [participant] DP which contains only first and second person *phi*-features if and only if no phase edge intervenes between the DPs. To support this claim evidence is provided from subordinate copular clauses that contain Negative Polarity Items in Turkish. This is an elegant analysis and has theoretical implications; however, some native speakers accept similar structures with third person singular *phi*-features such as *O Ali'dir* 'He is Ali'. *Ali o(dur)* 'Ali is he'. This intuition should also be accounted for.

Barış Kahraman in "Processing Turkish relative clauses in context" examines the validity of the recently proposed Discourse Function Hypothesis (DFH) to explain a well-known universal tendency: namely that subject relativization is easier to process than object relativization. DFH suggests that the relative difficulty of processing object relative structures is eliminated given appropriate context in discourse. The author carries out an experimental study with 24 sets of subject and object relative clauses in a self-paced reading task performed by 35 native speakers of Turkish. The results show that object relative clauses are relatively more difficult to process even in the appropriate context. This result indicates that the difficulty of object relativization cannot be explained by contextual factors in discourse; instead structural explanations may need to be taken into consideration in future studies.

Meltem Kelepir in "Nominalization morphemes are underspecified participial markers" shows that indicative nominal clauses lack evidential interpretation because there is no syntactic head that licenses it. However, nominal clauses have a tense interpretation although they are not morphologically marked with explicit tense morphemes. The author attributes the interpretation of tense to the existence of a tense operator, which contains semantic tense features, but lacks morphosyntactic features.

Greg Key and Deniz Tat in "Structural variation in Turkish predicates" analyze Turkish complex predicates with the light verb *et-* 'to make/to do' within the framework of Distributed Morphology (DM). They provide the following taxonomy of these complex predicates. Type I transitive complex predicates, the verb *et-* contains the features of CAUSE and BECOME: *Mehmet Berna'yı rahatsız etti* 'Mehmet disturbed Berna' (Mehmet CAUSEd Berna to BECOME disturbed) and *Berna rahatsız oldu* 'Berna was disturbed' (Berna BECOME-past disturbed).Type II is still transitive, but this time, the object does not undergo a change of state by using the verb *olmak*; i.e. it lacks BECOME. This can be exemplified as in *Mehmet, Berna'yı davet etti* 'Mehmet invited Berna', but **Berna davet oldu* 'Berna became invited' is not possible. Type III contains unergatives such as *dans etmek-* 'to dance'. Type IV contains also unergatives, but they are performative verbs such as *iltifat etmek* 'to compliment'.

Jaklin Kornfilt in "Two types of free relatives in Turkish in disguise: One is headed the other a correlative" analyzes two "fake" free relative clauses in Turkish. These are similar to free relative clauses in other languages, however, syntactically they are not free relative clauses. The first has a silent pronominal like *one* in English and a gap in the modifying clause. In the second, there is an additional conditional suffix *-sA* and a *wh*-element, which corresponds to the RC-head, e.g. *Hangi kitabı aldıysan onu okuyacağım* 'Whatever book you borrowed, I'll read it'. The author claims that these constructions are not real relative clauses, and also that from a typological perspective, the previous claims that Turkish does not have *Wh*-words in relative clause-like constructions cannot be maintained. This is because, as seen in the example above, *hangi kitap* 'which book' contains a *wh*-word in a relative clause-like construction.

Aysun Kunduracı in "A paradigm within process morphology" addresses noun-noun compounds which are marked by the morpheme *-(s)I* in Turkish, such as *kadın sihirbaz-ı* 'magician for women' Although these are superficially similar to N-N structures, such as *kadın sihirbaz* 'woman magician', the author argues that while the latter is an NP, the former is a compound that involves an associative attribution but not an intersective attribution. It is suggested that the morpheme *-(s)I* is a word formation suffix that marks the compounding process, and that these compounds are obtained through a word formation process at the level of morphology.

A. Sumru Özsoy in "Linearization in Turkish and minimality in binding" argues that fronting in Turkish results from a syntactic operation which involves moving constituents to a sentence-initial position outside their canonical merge positions.

The author claims that this movement is triggered by the requirements of information structure marking (topic/focus) and that it takes place at the level of syntax, not phonology, by providing evidence from the facts on binding and quantifier scope. Özsoy further argues that Turkish has two [topic/focus] positions in sentence initial position, and that locality constraints determine the nature of the constituents that can move into these positions.

Balkız Öztürk, Eser Erguvanlı Taylan, and Karl Zimmer in "Possessive-free genitives in Turkish" argue that there are two types of Possessive-free constructions (PFCs), with different syntactic derivations. PFCs lack a possessive morpheme on the head noun, e.g. *benim çanta* 'my bag'. These constructions are in contrast with the default Genitive Possessive Constructions (GPCs), which contain both a genitive suffix on the possessor and a possessive suffix on the possessed, e.g. *benim çantam*. The authors claim that the genitive in GPCs and type I PFCs are similar in that both have compositional semantics with a predicative relationship between the possessor and the possessed. On the other hand, type II PFCs do not imply any possession; the first part is used as a modifier that restricts the identification of the referent. The authors also discuss the discourse properties of PFCs, namely that they are always presuppositional and always used in casual speech.

Yağmur Sağ in "Spell-out of the combined predicate structures in Turkish: A nano-syntactic analysis" claims that the indicative perfective morpheme *-DI* contains both lexical-lexical and feature-feature pointers, while the evidential morpheme *-mIş* only has feature-feature pointers when considered within the domain of nano-syntax, a novel approach to the spell-out mechanism.

First and second language acquisition

Sıla Ay in "The vowel epenthesis and deletion errors of Japanese students of Turkish" describes Japanese learners' tendency to use an epenthetic vowel in CC clusters and to delete vowels when writing. The author discusses the patterns and potential reasons for these erroneous usages.

Özge Cengiz and Hamide Çakır in "An investigation of maternal input within the framework of Bloom's taxonomy of the cognitive domain" studied elicited and tape-recorded conversations between ten mothers and their five-year-old children. The authors found that mothers with higher social and educational backgrounds produced utterances at almost all levels of Bloom's taxonomy, while those with low social educational backgrounds did not.

Selma Elyıldırım and Güven Mengü in "Production of Turkish vowels by Swahili speakers" recorded the speech of three learners of Turkish from Africa whose native tongue was Swahili, and analyzed their use of vowels in Turkish. The results reveal that the speakers had difficulty producing vowels that do not exist in their native language. They replaced these vowels with the nearest corresponding sound.

F. Nihan Ketrez in "Acquisition of Turkish alongside a secret twin language" analyzed the speech of two identical twin boys at the age of 4:11. The speech of the

children was recorded when they were completing a puzzle with their parents. The results show that although the twins seem to be fluent in a secret language they share, their language development in Turkish is delayed at all levels of language.

F. Nihan Ketrez, Funda Kamiloğlu, Aslı Özkul, and Esra Yıldız in "Early vocabulary size in Turkish: Twins and singletons" studied the speech of 25 singletons and 28 twin children between 13 and 24 months of age. The results show that both singletons and twins have equal vocabulary sizes and that their vocabulary size grows in the second year.

Mine Nakipoğlu and Esra Yıldız in "Complementation and acquisition: The case of Turkish" investigate the productive skills of children in complementation. The data was collected using a picture-prompted elicitation task. The results show that children make errors in their choice of nominalizer suffixes, the acquisition of which presents a challenge for them.

Treysi Terziyan, Ayhan Aku Koç, and Eser Erguvanlı Taylan in "Acquisition of modality" present a comprehensive study on the acquisition of modal forms, such as modal verbal suffixes, predicates, and adverbs. The data was collected from natural spontaneous conversation of two girls and their caregivers over a four-year-period. The analysis revealed that the developmental stages in the acquisition of modality take place in the following order: First, dynamic modality is acquired followed by evidential modality; then deontic and epistemic modality emerge almost concurrently.

Burcu Ünlütabak and Ayhan Aksu Koç in "Developmental relations between reference to characters in narratives and theory of mind" investigate children's ability in elicited narratives to refer to characters in stories by considering the state of listener's knowledge. They found that age is a factor that determines the performance; 4- and 5-year-olds were better than 3-year-old children. Another factor that determined success was the complexity of narratives. Children were more successful with simple narratives with one main character, than more complex ones with two main characters.

Discourse, semantics, and pragmatics

Cengiz Acartürk in "Referring expressions in communication through line graphs" focuses on the referring expressions used to identify statistical graphics in an experiment with 125 participants in 6 groups. Materials used were cued and non-cued graphs. The most frequent verbal descriptions pertained to the domain value, reference to time, units of time, unit domain value. This experiment was conducted in order to discover typical verbal description templates for an automatic discourse generation system. Thus, it was important to determine these expressions in order to develop a multimodal corpus of referring expressions.

Mustafa Aksan and Yeşim Aksan in "Multi-word units in imaginative and informative domains" investigate the recurrent word groups, or multi-word units (MWUs) in sub-corpora of *Turkish National Corpus* in order to identify the structural types of MWUs and their functional categories. They also investigate whether

MWUs distinguish one domain from another. In terms of structure, the findings show that MWUs are mostly NPs, and that MWUs with verbal elements are rarely encountered in the data except for light verb constructions. As for the functions of MWUs, it is observed that they are mostly used as referential expressions (time, place, person reference, etc.). Other functions, such as text organizers, stance expressions, and conversational features follow referential expressions in terms of their functions. The results also show that there are domain-specific MWUs, which distinguish between imaginative and informative domains. The findings also suggest that MWUs in Turkish are far fewer in number than in English, which may be due to the fact that Turkish is an agglutinative language. The structural forms of MWUs, however, are similar in both languages. This is one of the pioneering studies within the field of MWUs in corpus studies in Turkish.

Başak Alango in "*Bu*'nun sözcenin düzenlenişindeki rolü" studies the role of the demonstrative pronoun *bu* 'this' in Turkish in the organization of the utterance. The author argues that the constituents with demonstratives in sentence-initial position are either subjects or subject-like entities, with old information status.

Engin Arık, Beril Arık, Esen Büyüksökmen, and Bade Dalahmetoğlu in "The relationship between figure/ground and frontness/backness: Evidence from Turkish" collected grammaticality judgment tests from sixty-four participants who were asked to mark these clauses on a seven-point Likert scale for grammaticality. The researchers aimed to test their hypothesis that there is a Figure—Ground and front—back asymmetry in locative clauses in Turkish. Their findings verified their hypothesis: Turkish speakers prefer F<G, e.g. *There is mouse in front of an elephant,* rather than F>G sentences, e.g. *There is an elephant in front of the mouse*.

Engin Arık, Pınar Öztop, and Esen Büyüksökmen in "An experimental approach to new and old information in Turkish locatives and existentials" investigated native speaker judgments on the appropriateness of locative and existential constructions in Turkish based on old and new information status of the NPs. Their data consisted of 12 test item sentences coded by 121 native speakers for appropriateness on a 7-point Likert scale. The results of repeated measure ANOVA showed that native speakers preferred existential sentences over locatives in presenting NPs with new information status. Given the previous findings on subjects and givenness in the literature, this is no mystery, since it is by now a well-established fact that subjects in sentence-initial position tend to mark referents with old information as in the locative sentences used in this study, such as *Kitap masada* 'The book is on the table', while existential sentences are used to introduce new information. The results of this study thus experimentally confirm what is already a consensus.

Gülsüm Atasoy in "Dictionary word definitions versus corpus-based word definitions" explores a corpus-based approach to lexicography, exemplified by the specific word *millet* 'nation' and its collocations within the context and co-text of this particular lexical item. The study suggests that dictionary entries must be based on unbiased and objective analyses of corpus studies rather than intuitions of those who prepare dictionaries.

Hamide Çakır and Özlem Fidan in "Contrastive study of the rhetorical structure of Turkish and English research article abstracts" examine the cross-linguistic and cross-disciplinary features of the move rhetorical structures observed in the abstracts of research articles in English and Turkish written by the native speakers of these two languages. The results show that three moves, Purpose, Method, and Result, are obligatory in the abstracts written in both languages and in both soft and hard sciences.

Mine Güven in "Some preliminary observations on restitution in Turkish" shows how the semantic domain of restitution is expressed and how it differs from the closely related notion of repetition of events in subtle ways. Restitution is a return to a previous state or position. It is closely linked to lexical aspect and event plurality. Restitution is obtained by various verbs and adverbs in Turkish, such as *dön-* 'to return', *geri* 'back', *tekrar/yeniden* 'again'. The author claims that the repetition is event-related; while restitution yields a result, i.e. a state. In other words, restitution involves a movement towards an earlier state in a reverse direction. This is obviously a notion different from repetition.

Tooru Hayashi and A. Sumru Özsoy in "*Şu* or *bu/o*: Turkish nominal demonstratives with concrete referents" collected data from 40 participants who collaboratively constructed three objects out of Lego blocks. Their dialogs were videotaped and transcribed. The aim was to investigate the relevant parameters in the choice of demonstratives, such as spatial proximity, etc. Their results show that the referents of *o* are more distant from the speaker, while those of *bu* and *şu* are closer to the speaker. Furthermore, the researchers found that most tokens of the demonstrative *bu* are accompanied by a gesture. *Şu* is also accompanied by a gesture though less often. The authors claim that demonstratives have a two-way distinction as parameters: proximity to and distance from the speaker. They further conclude that *şu* is not a demonstrative specified for spatial relations, but a proximity-neutral term whose function is to convey the speaker's assumption about the identifiability of the referent by the hearer.

Celile Eren Ökten and Duygu Çandarlı in "Preparing a translation corpus for raising awareness on translation errors" attempt to identify student errors in translated texts. They identified syntactic, lexical, and textual errors. This seems to be a work in progress, the ultimate aim of which is to build a parallel corpus from student translations and improve students' performance.

Language contact and sociolinguistics

Seyyare[i] Duman and Gönül Karasu in "Günlük söylemde yer alan eleştiriler" 'Criticisms in colloquial discourse' provide examples of positive and negative criticisms.

Barış Giray in "Code-switching among Bulgarian Muslim Roma in Berlin" investigates code switching within the framework of Muysken (2000). The Roma are Turkish speaking Muslim people from Bulgaria, who live in Germany. The data come from tape-recorded and transcribed informal, spontaneous interviews in Turk-

ish. The participants spoke Bulgarian, Romani, German and Turkish. Unsurprisingly, it was found that mostly content words were code-switched and that Turkish bound morphemes were attached to these words. Code-switching within this group had a social motivation: they tended to speak the language of Turks who live in Berlin in order not to be discriminated and to find jobs in that community.

Hristo Kyuchokov in "Syntactic complexity in the Turkish dialect spoken by Muslim Roma in northeast Bulgaria" explores relative clauses and word order used by Roma speakers of Turkish. The data come from video recordings and transcribed interviews. The author found that relative clauses used by these speakers follow the structure of Bulgarian. On the other hand, while simple sentences reflect the canonical SOV order of Turkish, complex sentences do not. This shows that complex sentences are influenced by Romani and Bulgarian due to language contact in the multilingual speech community.

Maria Petrou in "Lexical borrowing and code-switching in Turkish varieties of Western Thrace" describes contact induced phenomena at the level of phonology, morphology, and syntax from data collected in field work. The author suggests that phonological integration correlates with the level of bilingual proficiency of the speaker, while morphological and syntactic integration does not necessarily.

Carol W. Pfaff in "(How) will Turkish survive in Northwestern Europe? 50 years of migration, 35 years of research on sociopolitical and linguistic developments in diaspora Turkish" explores the place of Turkish in various European countries, such as Germany, France, etc., citizenship, and migration policies and their impact on Turkish. She gives some examples of research carried out on Turkish in the European context. It is a very ambitious paper for it evaluates the status of Turkish in Europe from many different perspectives: socio-cultural factors in the contexts of different countries, policies, as well as linguistic aspects of language contact phenomena. The author concludes that language attrition, incomplete language acquisition, and language change can be observed in Turkish communities in Europe. Nevertheless, at present Turkish is still strong among immigrant languages in Europe due to access to Turkish media, travel to Turkey, and spousal migration. Whether its strong status will continue in the future depends on the policies and attitudes that are adopted towards the language.

Jochen Rehbein and Annette Henkenrath in "Converbs in monolinguals' and bilinguals' Turkish" study converbial constructions used by bilingual Turkish speakers in order to understand whether language contact leads to changes in the use of converbs. They found that converb constructions remain relatively robust in language contact.

Kutlay Yağmur and Gülcan Çolak-Bostancı in "Intergenerational acculturation orientations of Turkish speakers in the USA" attempt to answer the question: Are there any differences in terms of acculturation among Turkish immigrants across generations, and to what extent is language use an indicator of acculturation? Data were collected mainly from participants who live in the New Jersey area in the United States and also online by using a questionnaire. The researchers found that there

were differences in acculturation between the first and second generation Turkish immigrants in the U.S.A., with the first generation tending to maintain the Turkish language and cultural values. The results also show that language is an important indicator for explaining acculturation differences. In the literature on Turkish immigrants, many studies are carried on immigrants in the European context, and fewer in the U.S. context. This is a study that contributes to our understanding of the attitudes towards the Turkish language and culture within the U.S.A.

Turkic languages

Christiane Bulut in "The Urum ('Pontioi') of Cyprus: A multi-lingual minority" gives a first description of the Turkic variety of the Urum. According to the author, no previous study of this language exists. Urum has predominantly Turkish lexicon with a high number of Azeri words. The author discusses phonological, morphological and syntactic aspects of the language and to what extent it has been influenced by language contact. For example, contact with Greek influenced the word order, so complex sentences follow the rules of Indo-European languages. Nevertheless, the existence of Turkish media also seems to be effective in replacing the rules of Indo-European languages with the Turkish ones.

Balázs Danka in "Hypercorrect orthographic forms in the Pagan Oguz-Nāmā" attempts to investigate the words with hypercorrect orthography and the quality of the sounds in <VqV>. The author claims that <VqV> and <VkV> sequences mark diphthongs as well as long vowels.

Fuyuki Ebata in "Valency retention in Sakha (Yakut) derivational nominalization" discusses three types of Sakha nominals: lexical derivation, syntactic derivation, and participle. The syntactic derivation is interesting because it is like a nominalization but it retains verbal properties, e.g. valency of the verb, at the same time.

Arman Eleusin in "Palatalization in the Mishar dialect of Kazan Tatar" describes the environments in which palatalization occurs in Mishar Tatar. For example, the reduction of the palatal /j/ triggers the palatalization of the following dental consonants. Additionally, palatal consonants are used with back vowels, but not with front vowels. The author states that palatalization in Mishar Tatar is different from Chuvash, Gagauz, and Karaim, but resembles palatalization in the geographically closest languages, as in Moksha and Eerzya Mordvinian.

Hossein Hashemi Zarajabad in "Copied passives in Southern Azerbaijani, Northern Azerbaijani and Turkish" investigates the role of language contact with Persian on the copied passives in the three varieties, Southern Azerbaijani (SA), Northern Azerbaijani (NA), and Turkish. In Persian, passive is obtained by a passive and inchoative auxiliary verb *ʃodaen* periphrastically. The author claims that in Southern Azerbaijani passives, the stative-inchoative verb *ol-* 'to be' is used as a result of contact with Persian. The author further argues that the changes in copied passives due to Persian influence in NA and Turkish have become stable over the years and have spread to the speech of monolinguals, but in SA, copied passives are mainly found

in the speech of bilingual speakers because of an ongoing change as a result of Persian influence.

Lars Johanson in "So close and yet so distant... On Turkic core structures, genealogical and typological grouping of varieties, and mutual intelligibility" asks how one should study distance between languages. He distinguishes between distances in terms of genealogical relations, typological similarities, areal features, lexicostatistical measurements, and degree of intelligibility. Especially interesting is the overview how Turkic languages diverged from one another historically through migrations and contacts with other languages. Johanson shows that convergence has also taken place due to intensive contact between Turkic varieties, e.g. in Central Asia and Anatolia. Turkic varieties have been in contact both with other Turkic varieties and with non-Turkic languages such as Iranian, Slavic, and others. Nevertheless, Turkic languages maintained core structures such as head final and left branching syntactic structure, agglutinative morphology, etc. A core property of Turkic, namely sound harmony, tends to exist across Turkic. Maintenance of genealogical stability is also observed in aspect and mood systems. Turkic languages influenced neighboring languages as well as being influenced by them.

Yuu Kuribayashi in "Causative/anti-causative alternations in Turkish, Old Turkic and Khalaj" investigates the derivation of transitive and intransitive verbs in Turkish in comparison with other Turkic languages by conducting a statistical analysis of verbs found in a Turkish-Japanese dictionary. Contrary to the claims presented in previous studies, this study shows that intransitive verbs are not always the basic forms from which transitive verbs are derived. In some cases transitive verbs are the basic forms and intransitives are derived from them. Still, in Old Turkic the dominating tendency is to derive transitive verbs from intransitives.

Eszter Ótott-Kovács in "On the causative-marked passive clauses in Old Turkic" shows that clauses containing verbs with causative suffixes can be interpreted as passive sentences in Old Turkic. A potential multi-functionality of a voice morpheme may not be too surprising, the author claims, when some facts on Indo-European languages are taken into account. In Old Turkic, passives marked with causative suffixes are strictly limited to certain syntactic environments. Firstly, they can only appear with transitive verbs. Secondly, they can take place in an underlying sentence in which the causer and the direct object are co-referential, as in *'We cause ourselves to suffer...'*, where the underlined constituents are co-referential. Thirdly, once the passive is formed, the underlying object is omitted as expected. The ambiguity between causative and passive clauses is resolved since the clause is only interpreted as passive when verbs denote actions that cannot be performed by the subject on himself, such as acts of violence or acts expressing respect, etc., unless the causer is inanimate. Finally, citing the literature the author claims that Old Turkic does not have morphosyntactic formal devices that mark passive sentences. This lack of grammatical devices to mark passives may be the reason why some causative clauses can get a passive interpretation in Old Turkic according to the author. If this

is correct, it shows how a lack of one formal aspect in language causes another to obtain multiple functions.

Arzhaana Syuryun in "Multi-application of grammatical suffixes in Tyvan" presents cases in Tyvan when two conditional morphemes are used together. Tyvan (Tuvan) is a language that belongs to the subgroup of the Siberian Turkic branch. One of the conditional suffixes is internal; it can precede person and tense morphemes. The other is external, i.e. follows them. These two morphemes can only be used together in the case of first and second persons. Although the internal morpheme historically comes from the past tense suffix, it does not have past reference but functions as a conditional, something which is peculiar to Tyvan.

Arzhaana Syuryun and Fuyuki Ebata in "Derivation from plural stems in Tyvan and Sakha" discuss the unexpected cases where derivational morphemes follow an inflectional one, i.e. the plural morpheme. This is unexpected, because it is a well-known language universal that if both derivational and inflectional morphemes follow or precede the root, the derivation is always between the root and inflection. Thus, in Turkic an inflectional suffix is not expected to precede a derivational one. According to the authors, the reason why the plural suffix precedes derivational morphemes may have to do with the blurred status of the plural morpheme. The authors also argue that the so-called "derivational" morphemes that follow the plural suffix not be derivational morphemes, but rather borderline cases in the inflection-derivation continuum. Otherwise, the plural suffix can be in an ongoing process of de-categorization of the plural morpheme in these languages. This study shows that in certain cases derivation and inflection are not clearly distinguishable and that the derivation inflection distinction forms a continuum rather than a strict dichotomy.

Prospective in Turkic languages

The final five papers discuss the semantic domain of the prospective, which is considered to be the immediate, imminent and near future.

Irina Nevskaya in "Prospective in Turkic languages" provides an overview of the category of the prospective aspect and also an evaluation of papers presented in this section.

Aminem Memtimin in "The prospective in Modern Uyghur" shows that prospective, near future, in Modern Uyghur, with its rich inventory, is complicated both in terms of structure and meaning as well as factors involved such as animacy, the semantics of auxiliary verbs, the choice of lexical aspect, etc. The past can metaphorically be interpreted as prospective, and the verbs *bol-* 'to be', *qal-* 'remain', *qil-* 'to do' can be used as auxiliaries that mark the prospective.

Astrid Menz in "Expressions for prospective and avertive in Turkish and Gagauz" deals with grammatical items that mark prospective ("a state [...] related to a subsequent situation") and the avertive (an expression that marks that "a situation was about to happen but did not") in two closely related languages, Turkish and Gagauz. In both languages, there is no specific marker that exclusively marks prospec-

tive; instead the future suffix *-(y)AcAk* is used. As for the avertive, Turkish has a special suffix that attaches at the verb stem: *-AyAz-*, but not Gagauz. However, lexical items, such as adverbials, such as *az kalsın* 'about to', etc. are used.

Irina Nevskaya and Saule Tazhibayeva in "The category of prospective in Modern Kazakh" show that Kazakh has a proper prospective *-GAll* + copula and also that *-MAQşI* can be used with animate subjects. Furthermore, *-(I)p qal* can also have prospective meaning.

Monika Rind-Pawlowski in "Prospective, intention and avertive in Dzungar Tuvan" states that the prospective in Dzungar Tuvan follows the pattern found in South Siberian. *-Vr-GA* +auxiliary and the length of time before the realization of the event is also marked *jit-* for short and *dur-* for longer duration. Moreover, another prospective marker *-MAKşI* is borrowed from Kazakh.

Conclusion and evaluation

As the brief review of the papers illustrates, this volume is a collection of studies with very diverse research interests, topics, and methodologies in various subfields of linguistics, within the frameworks of theoretical, empirical, functional and descriptive linguistics. Accordingly, it is a challenge to provide a unified evaluation.

As is well known, Turkish is relatively less investigated than Indo-European languages like English. Thus, it seems that we have a long way to go; that is why each contribution in the volume is a welcome addition to improving the knowledge within the field of Turkish/Turkic linguistics. The papers in the volume represent current topics, recent theoretical and empirical discussions within the field, some introducing new methods, and theoretical approaches. Some papers contribute to even less studied aspects, such as supra-segmental phonology, new aspects of morphology and syntax, recent topics in semantics, language contact, etc. All in all, the papers deepen our knowledge on Turkish and Turkic linguistics. They also have implications for the study of other languages, in terms of both their findings and methodology. Thus the papers in this volume can inspire further research in the field. Researchers and graduate students who work on Turkic as well as other languages and/or typology of languages will find the book very valuable.

TURKIC LANGUAGES

Information and Guidelines for Contributors

Contents. TURKIC LANGUAGES contains articles, review articles, reviews, contributions to discussions (remarks and replies), reports, surveys of publications, lists of publications received, and addresses of contributors.

Original contributions. As a rule, only original contributions are accepted. Submission of a manuscript implies that it has not been previously published and that it is not under review elsewhere.

Language. The preferred language of publication is English. Authors not writing in their mother-tongue are asked to have their manuscripts checked with respect to language and style.

Length of contributions. The suggested length of articles and review articles is 10–30 pages.

Submission of manuscripts. Authors are requested to send their manuscripts in electronic form, preferably as a doc-file together with a pdf version. Manuscripts should have a separate title page with the author's name and full mailing address. Please submit your contribution electronically, by e-mail or on disk. Manuscripts will not be returned.

Style sheet. Contributors are requested to comply with the style sheet which will be sent on request.

Permissions. It is the author's responsibility to request any permission required for the use of material owned by others.

Abstracts. Authors submitting articles and review articles are requested to include an English abstract of about 100 words.

Referee system. The journal uses a referee system in order to ensure a high standard. The anonymous reviewers will only read copies of the contributions without any identification of the author. While they may give suggestions for improvements, they are not obliged to specify the motivations for their decision to recommend or not to recommend publication in each case.

Proofs. Authors will receive one copy of the first proof, which, duly corrected, should be returned normally within two weeks. Changes against the original manuscript must be avoided or kept to a minimum.

Free author copies. Authors will receive a free copy of the volume in which their article appears. Additionally they will receive the print-ready pdf-file version of their article.

Prof. Dr. Dr. h.c. Lars Johanson, Turkic Languages, Institute of Oriental Studies,
University of Mainz, D-55099 Mainz, Germany
E-mail: johanson@uni-mainz.de