



Proceedings of the 1st  
Conference on Central Asian  
Languages and Linguistics

(ConCALL -1)

Volume 1

16- 17 May 2014  
Bloomington, Indiana



Edited by  
Dr. Öner Özçelik and  
Amber Kennedy Kent

(C) ConCALL-1 2014 Proceedings copyrighted by Indiana University, CeLCAR

(C) All papers copyrighted by the authors 2014

Proceedings of the 1st Conference on Central Asian Languages and Linguistics (ConCALL)

Volume 1/edited by Dr. Öner Özçelik and Amber Kennedy Kent

ISBN: 978-0-9961762-0-0

1. Central Asian Languages - Congresses

For information on ConCALL-1 (held on 16-17 May 2014, Bloomington, Indiana), visit our website at

<http://www.iub.edu/~celcar/ConCALL/>

Published by

Center for Languages of the Central Asian Region

Eigenmann Hall 702

1900 E. Tenth Street

Bloomington, IN 47406-7512

USA

<http://www.iub.edu/~celcar/>

## History of ConCALL

The Conference on Central Asian Languages and Linguistics (ConCALL) was founded in 2014 at Indiana University under the leadership of Dr. Öner Özçelik, the residing director of the Center for Languages of the Central Asian Region (CeLCAR), with grants and contributions from the U.S. Department of Education and several units at Indiana University, including the Ostrom Grant Programs, College of Arts and Humanities Center (CAHI), Inner Asian and Uralic National Resource, College of Arts and Sciences, School of Global and International Studies (SGIS), Sinor Research Institute for Inner Asian Studies (SRIFIAS), Department of Central Eurasian Studies (CEUS), and Department of Linguistics.

As the nation's sole U.S. Department of Education funded Language Resource Center focusing on the languages of the Central Asian Region, CeLCAR's main mission is to strengthen and improve the nation's capacity for teaching and learning Central Asian languages through teacher training, research, materials development projects, and dissemination. As part of this mission, CeLCAR has an ultimate goal to unify and fortify the Central Asian language community by facilitating networking between linguists and language educators working on Central Asian languages, encouraging research projects that will inform language instruction, and provide opportunities for professionals in the field to both showcase their work and receive feedback from their peers.

Thus, ConCALL was established to be the first international academic conference to bring together linguists and language educators working on the languages of the Central Asian region, focusing primarily on Altaic (Turkic, Mongolic, Tungusic) and Eastern Indo-European languages, among others, with the aim of encouraging research into how these specific languages are (i) represented formally, (ii) acquired by second/foreign language learners, and (iii) best taught given research driven teaching methods.

Languages served by ConCALL 2014 include, but are not limited to: Azerbaijani, Dari, Karakalpak, Kazakh, Kyrgyz, Mari, Mongolian, Pamiri, Pashto, Persian, Russian, Shughnani, Tajiki, Tatar, Tibetan, Tofalar, Tungusic, Turkish, Turkmen, Tuvan, Uyghur, Uzbek, Wakhi and more! In general, ConCALL represents all languages spoken in Central Asia and the surrounding areas, as well as languages that are genetically related to Central Asian languages (such as Turkish and Persian, which are directly related to Central Asian Turkic and Iranian languages respectively)!

## Conference Proceedings

Conference presenters were selected via a peer-review process. All conference invited speakers and accepted presenters were invited to publish their papers as part of the conference proceedings.



# CONTENTS OF VOLUME 1

<i>History of ConCALL</i>	iii
<i>Editor's Introduction</i>	vii
<i>Acknowledgements</i>	viii
<b>PART 1. PLENARY AND INVITED PAPERS OR ABSTRACTS</b>	
I KORNfilt, Jaklin, Syracuse University Turkish Relative Clauses: How Exceptional are they from a Central Asian Turkic Perspective?	3
II BECKWITH, Chris, Indiana University What Did the Earliest Central Asian Languages Sound Like?	25
III HANKAMER, Jorge, University of California, Santa Cruz Turkish Suspended Affixation	26
IV KARIMI, Simin, University of Arizona Contrast in Focus Persian and Turkish focus constructions	27
<b>PART 2. FORMAL LINGUISTICS AND RESEARCH</b>	
1 ABUDUREHEMAN, Kaili Binu, Northwest University for Nationalities Uyghur Consonants and Electropalatograph	31
2 BAEK, Sangyub, Hokkaido University Tungusic Converb -mi from the Perspective of Linguistic area	36
3 BOKHARI, Hassan, and WASHINGTON, Jonathan, Indiana University Stress in Saudi Diaspora Uzbek and its relation to stress in Turkish	51
4 BOLGUN, Ali, Monterey Institute of International Studies Why situation-theory cannot explain the accusative case marking in Turkish	62
5 DEMIR, Duygu, Hacettepe University Adverbial Clauses in Modern Turkish	73
6 DJURABAEVA, Sharifa, Ball State University Definition and clarification discourse markers in the speech of Uzbek speakers	79
7 HIDAKA, Shinsuke, Tokyo University Uzbek's Suffix Expressing Possession -li	86
8 ILOLOV, M., ILOLOVA, P., YUSUFBEKOV, Sh. P., Academy of sciences of the Republic of Tajikistan Lingvo-Methodical Objectives and Frequency-Specific Approaches for Kazakh Language	92

9	KURYSHZHANOVA, Aliya, Hankuk University of Foreign Studies Lingvo-Methodical Objectives and Frequency-Specific Approaches for Kazakh Language	97
10	NEMATOLLAHI, Narges, Indiana University Development of the Progressive Construction in Modern Persian	102
11	ŞAHİN, Savaş, Amasya University Metaphorical Approach to Mental Verbs in Turkmen Turkish	114
12	SUGAR, Alexander, University of Washington Implementing a Machine-Readable Grammar of Uyghur	123
13	YAMAKOSHI, Yasuhiro, Tokyo University of Foreign Studies How Can We Define The “Participles” In Mongolic Languages: Two Problems In Shinekhen Buryat	131

### **PART 3. PEDAGOGY AND TEACHING**

1	FEDERE, Marla and LEISHMAN, Clay, Naval Special Warfare Developing a Learner-Centered Pashtu Curriculum for Navy SEALs: A Case Study	143
2	KRAUSE, Johnathan, Western Carolina University Afghan Languages as a Source of Unity Rather Than Division	154
3	KUZEKOVA, Zeinekhan, Nazarbayev University Functional-Semantic Fields and Teaching of Kazakh Language	162
4	SAYDEE, Farid, San Diego State University Determinant Factors in Effective Teaching and Learning of Dari and Pashto Languages: An Action Research Paper	167
5	STERN-GOTTSCHALK, Ariann, Indiana University Trends in US-based Central Asian Language Instruction: The Case of Indiana University’s Summer Language Workshop	186

## Editors' Introduction

On May 16th and 17th, 2014, the Center for Languages of the Central Asian Region (CeLCAR) hosted the first ever Conference on Central Asian Languages and Linguistics (ConCALL) at Indiana University in Bloomington, Indiana as a biannual academic conference. This first ConCALL was established to bring together linguists and language educators specializing in the languages of the Central Asian region, including Turkic, Iranian, Mongolic, Tungusic and Tibetan languages spoken in the region, among others.

The theme of this founding conference, “**Building a Bond: Strengthening the Central Asian Language Community,**” was chosen based on our goals to create a stronger network of Central Asian language experts and to provide a unique opportunity for researchers, pedagogists, linguists, and educators in the field to present their work to an audience of their peers in the same language field.

We feel pleased at the overall success of this first time conference. We had 79 registered attendees that came from all over the globe, including not only the Central Eurasian region, including Afghanistan, Azerbaijan, Georgia, Kazakhstan, Kyrgyzstan, Russia, Tajikistan, Turkey, and Uzbekistan, but also attendees from such diverse countries as Sweden, India, China, Japan, Malaysia, Pakistan, and Australia! Additionally, we had attendees from various American universities and institutions including Yale University, University of Kentucky, San Diego State University, University of Washington, University of Wisconsin, and the Defense Language Institute in Monterey. And of course our illustrious guest speakers Jaklin Kornfilt, Jorge Hankamer, Simin Karimi, and Chris Beckwith represented Syracuse University, University of California - Santa Cruz, University of Arizona, and Indiana University, respectively.

We received 76 presentation proposals/abstracts. Although this was the first gathering of our conference, we were impressed and inspired by the high level of quality submissions, and selected 30 of these for oral presentations, i.e. with an acceptance rate of 39.47%. Out of the remaining 46, we chose 15 as poster presentations (32.61% acceptance rate). The selected presentations and posters covered an assortment of research topics related to syntax, semantics, phonetics, phonology, language acquisition, pedagogy, and more in an assortment of Turkic, Iranian, Mongolic, and even some Tungusic and (eastern) Uralic languages.

We cannot express enough our sincere pleasure at being a part of the founding of this conference which we genuinely believe is just the beginning of a movement towards strengthening and uniting the community of our field, and increasing the teaching of these less commonly taught languages in more language programs across the world.

And of course, once again, we would like to thank our main conference supporters: IU's Ostrom Program Grant, IU's College of Arts and Humanities Institute, the Inner Asian & Uralic National Resource Center, IU's College of Arts and Sciences, IU's Central Eurasian Studies Department, and IU's Department of Linguistics. And we look forward to seeing everyone again in Spring 2016, and hopefully some new faces as well.

Sincerely,

Dr. Öner Özçelik, Director  
Amber Kennedy Kent, Language Instructional Specialist

Center for Languages of the Central Asian Region  
Indiana University, Bloomington, Indiana  
May 2014



## ACKNOWLEDGEMENTS

The Conference on Central Asian Languages and Linguistics held at Indiana University on 16-17 May 1014 was made possible through the generosity of our sponsors and the help from all of the CeLCAR staff and student volunteers whose diligence and hard work contributed to the successful organization and execution of the first ever ConCALL.

Major sponsors of ConCALL include:

Center for Languages of the Central Asian Region (CeLCAR)

Ostrom Grant Programs

The College of Arts and Humanities Center (CAHI)

Inner Asian and Uralic National Resource Center (IAUNRC)

College of Arts and Sciences

School of Global and International Studies (SGIS)

Sinor Research Institute for Inner Asian Studies (SRIFIAS)

Department of Central Eurasian Studies (CEUS)

Department of Linguistics

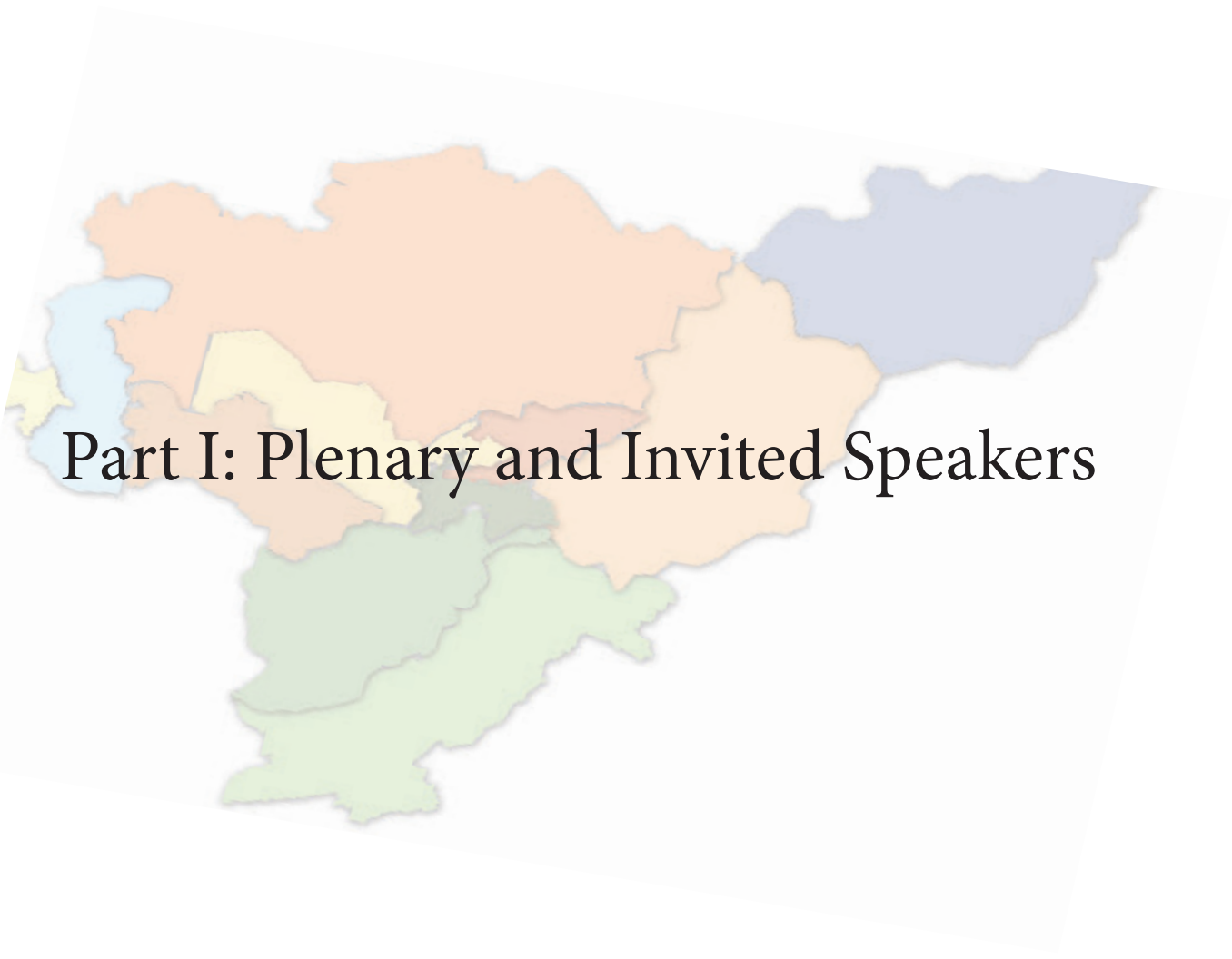
The conference would not have been as successful without their generous contributions.

Additionally, we would like to thank the anonymous reviewers for donating their time reviewing the many conference proposals.

Furthermore, we are eternally grateful for all of the conference presenters and participants who came from all over the globe to attend this first time conference and made it a huge success.

And finally, special thanks to all of the contributors of the ConCALL Proceedings for their submissions.





Part I: Plenary and Invited Speakers



# Turkish Relative Clauses: How Exceptional are they from a Central Asian Turkic Perspective?\*

Jaklin Kornfilt  
*Syracuse University*

## 1. Introduction

This paper has a historical and a synchronic dimension; the two dimensions have similarities in that they make related claims, as follows:

The historical claim is that Turkish phrase structure has developed functional projections over time; the synchronic, typological claim is that some of these functional projections are not shared by all of the cognate languages of Turkish, i.e. by some of the other Turkic languages. These projections are CP, TP, DP, and possibly Agr(S)P (if the existence of this projection is assumed); to be more specific, I claim that some of the contemporary Turkic languages, most of them being spoken in various parts of Central Asia, have also developed functional projections, but not all of those which Turkish, in its migration to the Mediterranean area, has developed. Conversely, it is possible that those languages have developed, or perhaps inherited, functional projections which Turkish has either not developed or has lost; more specifically, this might well be the case in the area of tense and aspect, given that the verbal morphology of some of those cognate languages is richer than that of Turkish in those respects. In Central Asia, the number of functional projections is probably smaller, as is their size: I shall claim that for at least those Central Asian languages this paper will look at in some detail, CP is missing—at least in relative clauses (RCs), and probably altogether.

A number of syntactic phenomena, such as certain types of movement (e.g. DP- or NP- movement in passives), certain checked structural cases (as opposed to default case), and certain operator — variable constructions (e.g. full-fledged RCs instead of reduced ones) have arisen in Turkish as a consequence, due to the additional specifier positions of these functional projections.

Especially with respect to RCs, this paper claims that the modifier clauses of Modern Standard Turkish (MST) are CPs, while they are TPs or even lower Aspect/ Mood-phrases in many other contemporary cognate languages. Historically, in Old and Middle Turkic, these clauses were Aspect/Mood-clauses, or possibly even smaller projections, i.e. bare VPs.

The paper is organized as follows: I first look at the historical dimension of the issue and present some case studies from older stages of Turkic, pointing at morphological as well as syntactic differences between them and the corresponding constructions in MST, making the claim that those disparate differences can be insightfully described in a unified fashion if they are ascribed to a difference in the functional projections available to the different stages, and if it is claimed that Turkish added functional projections in the course of its historical development, with the new additions being higher than the existing projections. I then turn to other contemporary Turkic languages, again presenting case studies (and, most importantly, studies of relative clauses or constructions that function as such) which make the same point: MST has a richer array of higher functional

---

\* This paper corresponds to a large part to the presentation I gave at the Conference on Central Asian Languages and Linguistics held at Indiana University in May 2014. Colleagues and informants gave me help of different kinds and at different times; it would have been impossible to write this paper without their contributions. I thank Öner Özçelik for inviting me to the conference, and for his and his team's generous hospitality. I also thank Philip LeSourd for his feedback on the Modern Uighur data after the conference, as well as for generously sharing his transcriptions of his own informant sessions on this language. I am grateful to Nadya Vinokurova for sharing her native intuitions in Sakha, as well as her insights into the syntax of that language, and to Abdurishid Yakup, to Kenjgül Kalieva, and to Raihan Muhamedowa for the examples they provided for Modern Uighur, Kirghiz, and Kazakh, respectively. Amber Kennedy Kent deserves particular mention for her help in formatting the paper. Any shortcomings are my responsibility.

projections than its contemporary cognate languages, especially having developed a CP which its contemporary sisters don't have (at least not in their RCs).<sup>1</sup>

A number of languages, English included, have the possibility of both pre-nominal (e.g. the recently visited island) and post-nominal (e.g. the island that was visited recently) RCs. Note that the pre-nominal RC is more reduced than the post-nominal RC. Modern Turkish pre-nominal RCs are different from their English counterparts, however, in at least two crucial respects: they can display agentive “by-phrases” in passive RCs, which is not possible in English (e.g. \*the by many tourists visited island), and they can target non-subjects—as discussed, to a large extent, in the present paper. The RC targeting a non-subject is a type of construction which is also not possible in English (e.g. \*the many tourists visited island, versus the well-formed post-nominal, non-subject RC the island that many tourists visited). This shows that Turkish pre-nominal RCs are not as reduced as their English counterparts and instead correspond to their post-nominal English counterparts (cf. also Kornfilt 2000 for more discussion).

## 2. Case studies

### 2.1. *TP versus VP*

In Old Turkic (OT)<sup>2</sup>, Agreement is severely limited, as is Tense. There is only one genuine Tense: the simple past, and it takes a genuine Agreement form, i.e. a bound morpheme:

- (1) biz az är-ti-miz  
we few be-PST-1.PL  
'We were few' (BK E 32, cited in Tekin 1967, 138)

I suggest that this single Tense was an innovation—an intimation of more tenses developing. Note also that this Tense marker is immediately followed by a subject agreement marker, as genuine Tense markers are in MST, as well. But everywhere else in OT verb forms, there is a postverbal personal pronominal form instead of a suffix. (A corresponding, mostly identical, pronominal form is only sometimes found in canonical preverbal subject position at the same time.) Thus, more typical examples are of the following sort:

- (2) biz az biz  
we few we  
'We are few' (O F 7, cited in Tekin 1967, 139)
- (3) saqin -ur män  
mourn-PRSPART I  
'I mourn' (lit.: 'I am mourning, I am the mourning one') (BK W 6; Tekin 1967, 139)

I propose the following analysis for such forms: At this stage, in this more general pattern, there is no T-head with its own projection. There either is no Tense marker at all (as in (2), where the interpretation is that of Present Tense, by default), or there is a marker, as in (3), which is a participle and has the Aspectual interpretation of habitual, thus a default temporal interpretation of present.

Consequences of this analysis: There is no Spec,TP in most instances of predicates; the subject is VP-internal in general—i.e. not only is the subject base-generated VP-internally, I claim that it also remains in that

<sup>1</sup> The relative clauses considered in this paper are limited to restrictive relative clauses. For concise explanations of the nature of restrictive versus nonrestrictive relative clauses as well as of (other) apparent relative clauses, the reader is referred to Keenan (1985), a very useful chapter on relative clauses.

<sup>2</sup> For the purposes of this study, I limit myself to the earliest body of documentation in Old Turkic (OT), i.e. to the inscription in the runiform script, dating from the 7th to 10th centuries AD, found in the area of the second Turkic empire and the Uighur steppe empire, i.e. present-day Mongolia, and the Yenisey basin. The two other corpora are documents of somewhat later stages (9th to 13th century Old Uighur manuscripts, and 11th century texts from the Karakhanid state), and they illustrate somewhat more “evolved” characteristics; I leave a study of the language in those documents to future research. For an overall description of OT in this wider sense, see Erdal (1998). (I have chosen the orthography of “Uighur” over “Uyghur” in this paper.)

position. Passive constructions should not involve moving the subject to a Spec,TP position. If so, how is nominative Case (or any other subject Case) checked? Two answers come to mind: 1. The subject Case is checked with the subject in-situ, VP-internally; 2. subject Case is a default Case.

My claim here is that the second answer is the appropriate one: In OT, subject Case is default. Where MST has a licensed structural Case on a specifier, with the licenser the functional head of the projection, OT has default Case and no functional element as the head of a potential functional projection.

I will discuss other instances of specifiers in the next (sub)sections.

## 2.2. Possessive phrases and nominal compounds

This subsection addresses a comparison of Old Turkic possessive phrases and nominal compounds with their Modern Turkish counterparts.

Certain Case and agreement markers that obligatorily co-occur in Modern Turkish do not show up in Old Turkic. Consider the following example of an Old Turkic inscription and its counterpart in Turkish, where we find possessive phrases and one nominal compound in both examples:

- (4) qayan it yıl onunç ay altı otuz-ga uça bar-dī  
 khan dog year tenth month six thirty-LOC die-PST  
 ‘The Kagan passed away on the twenty-sixth day of the tenth month of the Year of the Dog’  
 (BK S 10; Tekin 1967, 246)

MST:

- (5) kağan köpek (“it”) yıl -in -in onuncu ay -in -in  
 kagan dog year- COMPD -GEN tenth month-3.SG-GEN  
 yirmi altı-sın -da öl-dü  
 twenty six- 3.SG-LOC die-PST  
 ‘The Kagan died on the twenty sixth of the tenth month of the Year of the Dog’

Note that in (5), the nominal compound *köpek yıl-ı(n)* ‘dog year’ bears a compound marker on its head noun; the corresponding nominal compound *it yıl* ‘dog year’ in (4) has no such marker.

Likewise, in the possessive phrases of (5), we find that the “possessors” are all marked with the genitive, and the “possessee”, i.e. the heads of the possessive phrases, are marked with a third person singular agreement marker, thus agreeing with their possessors in the features of person and number; thus, we find, for example, for the meaning ‘the tenth month of the dog year’ (literally, ‘the dog year’s tenth month’), a genitive marker on *köpek yıl-ı(n)* ‘dog year – compound marker’, yielding *köpek yıl-ın-ın* ‘of the dog year; the dog year’s’; we further find the just-mentioned agreement marker on the “possessee”, i.e. on ‘tenth month’: *onuncu ay-ı(n)*.<sup>3</sup>

My account for the appearance of these markers is as follows:

Possessive phrases: In MST, the genitive has to be checked by a phi-feature agreement element (cf. Kornfilt, 2003; 2006; 2009a). In the corresponding OT example, the specifiers (“possessors”) of the possessive phrases are bare, and there is no agreement element on the heads (“possessee”), as we saw, thus raising the question of why no checking mechanism is required.

I propose the following account: The bare DPs in OT in specifier position are in the nominative Case (i.e. a Case which is represented by a morphologically null element, also in MST), and the nominative is a default case in OT. The possessive phrases in OT are bare NPs, which are not dominated by functional projections (or at least not by an *AgrP* or an *nP*—cf. Kornfilt 2006), the way they are in MST; this is why there is no

<sup>3</sup> Note that the compound marker is actually the third-person singular agreement marker, which I take to be the default agreement marker. In both of its usages, i.e. as a nominal compound marker as well as when used as a genuine agreement marker in possessive phrases, this marker has a word-final /n/, which gets deleted in word-final position, but does show up otherwise, i.e. both before a vowel or a consonant. This is why the /n/ is placed between parentheses in the text.

morpho-phonologically realized agreement morpheme, and this is why there is no overt genitive (or any other licensed, specifier-specific case, either). Given that the nominative is a default case, at least in OT, it does not need a licenser.

More detailed representations of the relevant parts in the previous examples, as discussed above, follow:

- (6) MST:  $[_{DP} [_{DP} \text{ köpek yıl} \text{ -in} \text{ -in}] [_{D'} [_{NP} \text{ onuncu ay}] [-1]]]$   
           dog year- CMPD-GEN tenth month -3.SG  
           ‘The tenth month of the Year of the Dog’
- (7) OT:  $[_{NP} [_{NP} \text{ it yıl}] [_{NP} \text{ onunç ay}]]$   
           dog year tenth month  
           ‘The tenth month of the Year of the Dog’

Returning to nominal compounds, as we saw earlier, in Modern Turkish, the nominal compound must be headed by default agreement (i.e. the third person singular agreement marker), which has become a nominal compound marker in MST; the OT nominal compounds lack this, or any other, marker; additional examples for this contrast follow:

- (8) a. OT: *türük bodun*  
           Turk people  
           ‘The Turkish people’ (KT E 6; Tekin 1967, 203)
- b. MST: *Türk millet -i*  
           Turk people -CMPD (=3.SG)  
           ‘The Turkish people’
- (9) a. OT: *şantung yazı*  
           Shantung valley  
           ‘The Shantung valley’ (BK E 15; Tekin 1967, 203)
- b. MST: *Şantung ova -sı*  
           Shantung valley -CMPD (=3.SG)  
           ‘The Shantung valley’

The default agreement as a nominal compound marker isn’t found with other kinds of compounds or phrases in MST, i.e. where the non-head is not a nominal; e.g. with an adjective:

- (10) a. MST: *yeşil ova*  
           green valley  
           ‘The green valley’
- a. MST: *açık deniz*  
           open sea  
           ‘The open sea’

I sketch the following analysis for MST: Suppose that the default agreement marks a small *n*-head which checks a special nominal case on the nominal which is in-situ in the complement position of the head noun, via AGREE. Non-nominal categories don’t need case and hence no licenser, which thus explains why no compound marker is found in examples such as (10). But in OT, this special nominal case is a default case that doesn’t need to be checked by a licenser, thus there is no default agreement on the head of the nominal compound.

### 2.3. *Passive*

In Modern Turkish, it is clear that there is syntactic passive; derived subjects can be thematically unrelated to the passive verb:

- (11) Ali sen -i<sub>i</sub> [<sub>t<sub>i</sub></sub> uyuyakal-dı] san -ıyor  
       Ali you-ACC fall asleep -PST believe -PRSPRG  
       ‘Ali believes you to have fallen asleep’

- (12) *Sen*<sub>i</sub> (*Ali* tarafından) *t*<sub>i</sub> [*t*<sub>i</sub> *uyuyakal-dı*] *san* -il -iyor -sun  
 you Ali by fall asleep-PST believe-PASS-PRSPRG-2.SG  
 ‘You are believed by Ali to have fallen asleep’

In (12), the subject of the passive main clause, *sen* ‘you’, is in the nominative and agrees with the main clause predicate in person and number; however, this subject is not related thematically to that main clause predicate. This means that at least for instances like these, there must be a genuinely syntactic (rather than lexical) passive; such constructions are due to DP-movement and cannot be ascribed to a purely lexical phenomenon based on a verb’s externalizing an internal argument; as just mentioned, in the example above, the derived matrix subject *sen* ‘you’ is not a thematically related argument of *san* ‘believe’.

There are no passives of this sort in OT at all. This may be due to an accidental gap in the documentation, but given that there is quite a wealth of documentation for the relevant period, this is unlikely. We have to note that there is no motivation in OT for DP-movement that would be case-driven; if nominative is a default rather than a licensed case, it can be assigned to a DP without that DP’s needing to move to a designated position where nominative is licensed. Instead, there is lexical passive, which externalizes an internal argument, which then receives default case. Also, if there is no TP, there is no specifier position of TP as a landing site (and a site where subject case is licensed) for DP-movement. Clearly, this account is more explanatory than the assumption of an accidental gap.

OT did have instances of passive predicates; a passive marker that is the same as that of MST does exist. However, the phenomenon of passive as such must have been different from what we find in MST. Tekin (1967) observes that a number of passive verbs are found without active counterparts, offering support for the claim advanced here that the passive in OT was a lexical rather than movement-based syntactic operation—in other words, morphologically passive forms without corresponding active verbs:

- (13) a. *tir-il* ‘to come to life’ (BK E 31); <\**tir* ‘to live’ (Tekin 19’67, 115)  
 b. *adr-İL* ‘to be separated, be disjoined’ (TI W 2); <\**adır* ‘to separate’ (Tekin 1967,115)  
 c. *ök-ül* ‘to be planned’ (TI N 8); <\**ök* ‘think’ (Tekin 1967, 115)

Data where a given morphological form (here, the passive) has no morphologically simpler (here, active) counterpart is typical of a lexical process and has been posited as a criterion (among others) to distinguish between lexical (adjectival) and syntactic (verbal) passives in English (cf. Levin and Rappaport 1986, Lightfoot 1979, Wasow 1977).

An additional observation about morphological passives in OT is that even where the “passive” verbs do have a corresponding active form, they have a reflexive or middle/inchoative reading in context, rather than a genuinely passive interpretation; for example, the passive verb in the following example, *adr-İL* ‘split from, separate’, is one of the verbs whose active counterpart is not attested. In addition, the context makes it clear that the verb bearing the “passive” morpheme does not have a passive meaning:

- (14) *türk bodun ... tabyač-da adr -il -tı, qan-lan -tı*  
 Turk people Chinese-ABL separate-PASS-PST khan-‘get’-PST  
 ‘The Turkish people parted from/left the Chinese and got themselves a khan’  
 (TI W 2, Tekin 1967, 249)

If this were an example of a true passive, the translation should be: ‘The Turkish people were separated from the Chinese (by someone else, i.e. by the agent of the separation)’, which doesn’t correspond to the historical realities. What we know about the events in question is that the Turkic people, under the leadership of Tonyukuk (the ‘Prime Minister’), decided that the time had come for their independence from the Chinese; they left for a different area and crowned a leader for themselves. (There are more examples illustrating passives of this sort in the inscriptions from which this and the other OT examples above are taken.)



2.4. *Passives and relative clauses*

If OT did not have a syntactic TP projection, it would be surprising (although not altogether impossible) if it had a CP projection, i.e. a functional projection higher than TP. There is evidence that OT did not have a CP projection, either.

In the documented OT inscriptions, there are no passive predicates in pre-nominal modification, i.e. in constructions corresponding to relative clauses. As a consequence, examples such as the following representative one, where the head corresponds to the thematic role of the predicate's patient, are systematically ambiguous between an active and a passive reading (assuming that passive interpretations are universal, and are present in a language that has no syntactic passive).

- (15) [qañimiz      äçimiz      qazyan-miš ]      bodun  
 our khan      our uncle      conquer-PRFPART      people  
 'The peoples who were conquered by our father and uncle' (= 'the peoples whom our father and uncle conquered')  
 (BK E 22, Tekin 1967, 179)

In MST, there would be two separate corresponding constructions: RCs with passive and RCs without passive. In both instances, the modern constructions bear evidence of functional syntactic structure which is lacking in the OT examples. I start by illustrating a corresponding RC with passive in the modifying clause:

- (16) a. [e<sub>i</sub> kağan-ımız ve amca-mız tarafından yen -il -miş ] millet-ler<sub>i</sub>  
 khan-1.PL and uncle-1.PL by      conquer-PASS- PRFPART      people-PL  
 'The peoples who were conquered by our khan/father and uncle'

Note that in addition to the passive marker on the verbal predicate, we also have an indication that the agents, i.e. the conjuncts within the coordination of *kağanımız* 'our khan/father' and *amcamız* 'our uncle', are not a coordinated sequence of subjects, but of non-subject agents, given that they show up as complements of the postposition *tarafından* 'by'. There is no such indicator in the OT example, thus leaving it open whether the corresponding coordination of agents is actually a subject or a non-subject agent phrase. In this MST example, the head of the RC corresponds to a derived subject, i.e. to the thematic patient of the predicate, which is the derived subject, via syntactic passive, i.e. via DP-movement.

Next, a non-subject RC without any sign of passive in MST:

- (16) b. [kağan-ımız-ın      ve amca-mız -ın e<sub>i</sub>      yen -diğ-i ] millet-ler<sub>i</sub>  
 khan-1.PL-GEN      and uncle-1.PL-GEN      conquer-FN-3.      people-PL  
 'The peoples whom our khan/father and uncle conquered'

Here, the target is not only the thematic patient, but clearly a syntactic non-subject, given that the subject is clearly marked as such via the licensed subject case of genitive (i.e. the subject case for specific subjects in nominalized clauses). The (specific, referential) subject must be in the specifier position of a nominalized TP or Aspect phrase, which is headed by agreement, in order to have its genitive case licensed; cf. Kornfilt, 2003; 2006; 2008).

In MST, we have now seen ways in which it is clear whether the target of relativization is a subject or a non-subject (i.e. independently from its thematic role); in OT, the syntactic relation of the RC-target to the clause's predicate is ambiguous. The reasons for this ambiguity are as follows: 1. The subject is not marked with any dedicated licensed case, but is in a default case. 2. There is no relativization morpheme on the predicate that would give any clues: the perfective participial marker on the predicate in our example is also found in simple clauses; furthermore, there is no subject agreement marker. Therefore, the shape of the predicate does not indicate anything with respect to the target's relation to that predicate. In contrast, in MST, the so-called factive nominalization marker *-DIK*<sup>4</sup> signals an embedded clause, and the third person agreement marker not only

<sup>4</sup> I follow general Turkological practice (and the practice of generative studies where Turkish and Turkic are concerned) of using capital letters to represent archephonemes whose full values are determined by vowel harmony for vowels, and by voicing and devoicing processes for consonants.



licenses the genitive case on the subject, as just mentioned; it also signals the fact that the target of the RC is not a subject, given that in subject RCs, there is never an agreement marker—something which can be seen easily in comparison with the preceding subject RC example under (16a) for MST.

The factive nominalization marker found in non-subject RCs in MST is an (impoverished) Tense marker, given that it can differentiate between future (in which case it is *-(y)AcAK*) and non-future (in which case it is *-DIK*). This goes along with the approach here that proposes a TP for MST clauses, both when they are fully finite and when they are nominalized (at least in this type of factive RC). In contrast, I have proposed that there is no TP in OT. TP provides a site for the moved DO; it also provides a site for the licenser of the Genitive and Nominative (=subject case in nominalized and fully tensed clauses, respectively), this licenser being the agreement under T.

### 2.5. RCs: CPs in MST, AspPs in OT

The morphology of the predicate in MST RCs differs in two distinct ways, depending on the target of the RC. The first difference concerns the presence or absence of agreement morphology on the predicate, expressing agreement with the subject in terms of person and number: in subject RCs, there can't be any overt agreement marker on the predicate; in non-subject RCs, the agreement marker is obligatory. The absence of agreement in subject RCs will be addressed in the synchronic part of the paper. The presence of agreement in non-subject RCs has been mentioned earlier: it is needed as the licenser of the subject Case, i.e. of the genitive in these nominalized RCs.

The second way in which the predicates differ with respect to the RC-target is in the shape of the nominalization marker: while in non-subject RCs, this is the general factive nominalization marker *-DIK*, found in any factive nominalized embedding, there is a special marker, *-(y)An*, in subject RCs. The following examples illustrate both differences:

A subject as the target of relativization:

- (17) a. [[<sub>i</sub>e<sub>i</sub> geçen yaz ada-da ben-i gör-en] kişi-ler<sub>i</sub>]  
 last summer island-LOC I-ACC see-(y)An person -PL  
 ‘The people who saw me on the island last summer’  
 (1: No phi-feature morphology on the predicate of the modifying clause; 2: special nominalization form on that predicate)

A non-subject as the target of relativization (traditionally so-called “object relativization”):

- (17) b. [[*pro* geçen yaz ada-da e<sub>i</sub> gör-düğ -üm ] kişi -ler<sub>i</sub>]  
 last summer island-LOC see-FN -1.SG person -PL  
 ‘The people who(m) I saw on the island last summer’  
 (1: Phi-feature morphology; 2: general indicative nominalization form on predicate)

The next example illustrates a general factive nominalized embedding, thus showing the same morphology on the predicate as what is exhibited by non-subject RCs:

- (18) [öğrenci -ler-in ben-i ada-da gör-dük -lerin]-i duy-du-m  
 student -PL-GEN I-ACC island-LOC see-FN -3.PL -ACC hear-PST-1.SG  
 ‘I heard that the students saw me on the island’

Based on these data, I make the following two central claims:

Claim 1: The difference between the predicate forms in (17a) and (17b) is best viewed as a version of the “*que-to-qui* conversion” in French, however analyzed. Almost all analyses of this phenomenon in the literature, however much different from each other, involve the CP-projection.<sup>5</sup> Suppose that the regular factive nominalization (FN) form *-DIK* corresponds to the general indicative complementizer *que*, and the special form exhibited by subject RCs (*qui*) corresponds to *-(y)An*. This means that in RCs in MST, the CP-projection is



At a stage following the stage discussed above as Old Turkic, namely the stage usually referred to as Middle Turkic, which is usually assumed to start around the 13th C, the OT construction survives in some languages.

In (Early) Chagatai, a version of Middle Turkic, it is still possible to have the OT construction for non-subject RCs, with a subject which is not morphologically marked for Genitive:

- (21) evvali [Sultan Abu Sa'id Mirza qoy-gan ] Mihr Nigar Xanım idi  
 previous/first Sultan Abu Sa'id Mirza make-P Mihr Nigar Lady was  
 'The first was Lady Mihr Nigar whom Sultan Abu Sa'id Mirza made [his fiancée]'.  
 (Baburname, as cited in Schönig 1992/93)

Note that the subject is bare, and no overt agreement with the subject is to be seen on the predicate of the modifying clause in this example. The subject case is apparently still a default case.

But in some other languages of the same time, there is subject — predicate agreement. However, the agreement is mostly on the clause-external head of the relative clause (like some of the modern Turkic languages to be discussed shortly) rather than, as expected, on the modifier domain (i.e. on the clause's head, namely the participial predicate)<sup>7</sup>. For example, Khorezmian, another Turkic language of the Middle Turkic period, exhibits some examples where construal of the agreement on the clause-external head as agreement with a possessor would be semantically strange; thus, we have to analyze the agreement marker on the clause-external head as agreeing with the clause-internal subject, rather than with a possessor:

- (22) [täfä -niŋ kel -ür ] yol -in -da  
 camel-GEN come- PRFPART road-3.SG-LOC  
 'On the way on which the camel came/comes along' [habitual]  
 (Rabyu:zi: 13; Schinkewitsch 1927, as cited in Schönig 1992/93)

The subject bears overt genitive case. Agreement starts playing a role in case assignment, i.e. instead of default case, a system with case licensing under Agree (or Spec/Head agreement) is emerging (i.e. nominal agreement on the predicate and genitive subjects start emerging in embedded clauses that are nominalized).

At this point, the following question arises: How can the head agree with a non-local subject?

Answer: The modifying nominalized clause in Khorezmian (as well as its contemporary languages with similar properties) is not a CP, but rather a reduced Aspectual projection. The higher DP, headed by the (clause-external) head noun, is a full-fledged, non-reduced maximal projection, and the agreement marker on that head marks the DP as such. Thus, the relevant local domain to license subject case (which now has evolved into a licensed, rather than default, case) is that of the (higher) DP. Within that DP, the genitive case is indeed licensed locally in Khorezmian (and its kin).

Note that Khorezmian also has instances of the older type, with unmarked subject and no agreement; e.g.

- (23) män ayt-ma-mis hadi:s  
 I say-NEG-PRFPART hadith  
 'A hadith which I have not said' (Schönig 1992/93)
- (24) [Päyya:mbar oltur-γan] yär-dä oltur -ayın  
 prophet sit -P place-LOC sit -(OPT)1.SG  
 'I wish to sit in the place where the Prophet sat' (Schönig 1992/93)

It is not unusual for innovative constructions to be used along constructions which survive a previous stage, and this is what we find in some of the modern Turkic languages, too (where, apparently, the two constructions are essentially synonymous, with some pragmatic, discourse-based differences that are hard to pinpoint at the current stage of our knowledge). At this stage, this innovation with respect to genitive subjects (rather than subjects unmarked for case) and their co-occurrence with Agreement on the RC-head seems to be limited to certain Tenses or Aspects.

<sup>7</sup> There are some examples in OT of this kind, too, but those are best analyzed as possessive constructions, with possessive agreement on the head, i.e. agreeing with a possessor, not as agreement with the subject of the modifier clause.

At a third stage (Early Ottoman=Early Anatolian Turkish, late Middle Turkic), we find the following complex situation: the nominal Agreement has become generalized to all Tenses/Aspects. I therefore claim that the Agreement morpheme now systematically marks a full-fledged syntactic domain, and thus shows up on the head of that phase. (I don't take a stand here on whether the Agreement morpheme is attached to D, or whether it projects in addition, or instead, an AgrS-phrase.) At this stage, yet another innovation has taken place: the difference in the predicate's morphology with respect to the target, i.e. a different morphology for subject RCs, and regular factive nominalization morphology for non-subject RCs. I thus claim that the modifier clause in these RCs is now a full CP. This analysis is supported by the fact that the subject Agreement morpheme is found on the predicate of the clause (as in MST, as we saw above), and not on the head of the relative clause.

- (25) [sän işlä-düg -ün] iş  
 you do -FN -2.SG deed  
 'The deed that you did' (Süheyl-ü Nevbahar [14th C]; Banguoğlu 1938, 115)

Note that the pronoun is not in the genitive; the unmarked case here is either default, or licensed by the (local) Agreement. At this stage, the genitive seems to be optional (perhaps only morphologically).

With the exception of the genitive subject (where the genitive is obligatory), the Modern Standard Turkish construction is the same: the same local Agreement on the participle is found on the nominalized predicate of the modifying clause, as we saw earlier:

- (26) [(biz-im) e<sub>i</sub> oku-duğ -umuz] kitap-lar<sub>i</sub>  
 we-GEN read- FN -1.PL book-PL  
 'The books that we read'

To recapitulate: the Agreement between subject and participial predicate is local in both constructions, i.e. the earlier one with overt Agreement on the head and the later one with overt Agreement on the participle.

A further development took place: during this third (Early Anatolian Turkish) stage, a new future Tense/potentiality Aspect morpheme was "borrowed" into this branch of Turkic from related languages. The Agreement morpheme associated with this form was the nominal form found on nominalized predicates, and, just as was *previously* the case for *all* Aspects, is found attached to the RC-head, rather than to the predicate. This means that the modifying clauses in these future tense/potential Aspect RCs were bare Aspect Phrases, and not full CPs:

- (27) [[var -acaq ] yer -ümüz] İraq-raq -dur  
 arrive -FUTPART place-1.PL far-somewhat -is  
 'The place where we shall arrive is rather far'  
 (Marzuba:n-na:me [14th C]; Kleinmichel 1974, 315)

The continuation of the sentence after the RC shows that the nominal Agreement form on the head noun cannot be interpreted as a possessive marker, but is indeed an Agreement marker with the subject of the modifier clause (with the Agree relation between the subject and the clause-external Agreement morpheme being local, i.e. within the same phase, given that the modifying clause is not a CP, but rather an AspP, and thus not a phase itself). The next example, from a different document from the same period, makes the same point:

- (28) [gäl -äcäk] yer -üm çünkü ol-a qapu-ŋ  
 come- FUTPART place -1.SG because be-shall door-2.SG  
 'Because the place to which I shall be coming is your door'  
 (Süheyl-ü Nevbahar [14th C]; Banguoğlu 1938, 126.)

These two structures (i.e. the more evolved one with the subject agreement marker on the participle in non-Future/potential Aspect relative clauses, and the more archaic one with the subject Agreement marker on the head noun in Future/potential Aspect relative clauses) co-existed during this stage but "collapsed" into the "evolved" structure in a later stage, during Ottoman Turkish, when the marker for Future/potentiality became "regularized", and with it also its associated Agreement marker, which lost its restriction for purely nominal hosts and became a suffix able to show up on nominalized verbal predicates. This is also the situation we find in contemporary MST:

- (29) var -acağ -ımız yer  
 arrive- FUTPART-1.PL place  
 ‘The place where we shall arrive’

The Central Asian RCs have now arrived in the Mediterranean world, and in the process have grown from bare ASPPs to full-fledged CPs.

In the next part of the paper, I turn to the synchronic dimension of this study, and I discuss RCs in contemporary Turkic languages which “stayed behind” in Central and Northern Asia.

#### 4. Other Turkic languages

In some contemporary Turkic languages, presence versus absence of Agreement is the only difference between the two main types of RCs (i.e. subject and non-subject RCs) and is determined similarly to Turkish (cf. Csató 1996).

##### 4.1 Sakha (Yakut) RCs: the apparently long-distance subject—Agr relation

Relative clauses in Sakha resemble the Middle Turkic situation: They do have Agreement with the subject, and thus are different from their counterparts in Old Turkic. Also, this Agreement is placed on the head of the RC, which makes them similar to the innovative RCs in Middle Turkic. At the same time, the situation is somewhat different from what we saw in Middle Turkic: Sakha does not have remnants of the Old Turkic alternative RC, i.e. the construction where there is no Agreement at all in non-subject RCs. Also, Sakha has largely lost the morphology of the genitive case (with a very limited genitive, to be discussed later in this paper); therefore, even when the Agreement is present, the subject typically shows up bare, apparently in the nominative.

- (30) a. [it e<sub>i</sub> ih -iex -teex ] üüt<sub>i</sub> -e  
 dog (NOM) drink -FUT -MOOD milk -3.SG  
 ‘the milk the dog should drink’ (Kornfilt & Vinokurova 2001)
- b. [[kini aqa -ta ] e<sub>i</sub> öl -ör -büt] oquh<sub>i</sub> -a  
 he(NOM) father -3.SG(NOM) die-CAUS -P ox -3.SG  
 ‘The ox which his father killed’ (Kornfilt & Vinokurova 2001)

1. The subject—agreement relationship, and the licensing of the case on the subject by the Agreement on the RC-head, appear to violate locality. Because such relations are assumed to be local universally, the Turkish situation with the Agreement on the predicate is expected, while the situation in Sakha is surprising.

2. If the lack of overt *Agr* in subject RCs is to be explained via Generalized Binding, then the relationship between the subject “gap” as *pro* and its licenser, i.e. overt *Agr*, seems to violate locality, as well.

Can we derive locality where the source structure doesn’t seem to offer a local relationship? Is that relation similar to proposals made earlier in this paper for similar constructions in Middle Turkic, based on an Agree relationship between the Agreement element on the clause-external head of the entire construction, i.e. on D, along with the claim that the modifying clause is smaller than CP, and therefore not a full-fledged syntactic domain, while the DP is such a domain?

There is an alternative: Raising of the subject to the specifier position of the higher DP. However, I shall claim here that this alternative is not viable for Sakha.

I shall first consider possessive DPs in Sakha. This is relevant, because in the literature that favors the raising analysis of subjects in RCs in Altaic languages (e.g. Hale & Ning 1996 and Hale 2002 for Dagur Mongolian, and Aygen 2006 for Turkic languages such as Kazakh and Kazan Tatar), it is the specifier position of the higher DPs in RC constructions, i.e. the “possessor” position, which has been taken to be the target position of the subject’s raising.



4.2 *Possessive DPs in Sakha*

Possessive phrases in Sakha are quite similar to their counterparts in Turkish: The head of the possessor phrase, i.e. the possessee, agrees in phi-features with the specifier of the possessive phrase, i.e. with the possessor. The only difference is that the possessor is not in the genitive in Sakha in general. This is expected: as mentioned earlier, Sakha has, by and large, lost the genitive.

- (31) a. *kīis*      *oquh-a*  
 girl(NOM) ox -3.SG  
 ‘The girl’s ox’ (Kornfilt & Vinokurova 2001)
- b. *kini*      *aqa -ta*  
 he (NOM) father -3.SG  
 ‘His father’ (Kornfilt & Vinokurova 2001)
- c. *min*      *oquh -um*  
 I(NOM) ox -1.SG  
 ‘My ox’ (Kornfilt & Vinokurova 2001)

Conditions under which a relic genitive does show up on DPs in Sakha are as follows: This happens when the possessor is itself complex. In other words, if the possessor is itself a possessive phrase, then the complex possessor does get marked with a morpheme which is a relic of a previously productive genitive case; the possessor within that complex possessor is, as expected, in the nominative:

- (32) a. [*kini*      *aqa -tī*] -n *oquh -a*  
 he (NOM) father -3.SG -GEN ox -3.SG  
 ‘His father’s ox’ (Kornfilt & Vinokurova 2001)
- b. [*kīis*      *oquh-u*] -n *kuturug -a*  
 girl(NOM) ox -3.SG -GEN tail -3.SG  
 ‘The girl’s ox’s tail’ (Kornfilt & Vinokurova 2001)

This relic genitive is limited to third person possessives, i.e. it does not show up after first or second person possessives.

4.3 *First empirical argument against subject raising in RCs: Sakha RCs with complex possessive subjects*

The relic genitive just mentioned does not show up on the subject of an embedded clause, even if that clause is nominalized, and even if the subject is a third person complex possessive phrase:

- (33) a. [[*Kini aqa -ta*] *üüt -ü*  
 He father -3.SG (NOM) milk-ACC  
*ih -iex -teek -i*] -n *bil -e -bin*  
 drink-FUT-MOD -3.SG-ACC know-AOR-1.SG  
 ‘I know that his father should drink the milk’ (Kornfilt & Vinokurova 2001)
- (33) b. \*[[*Kini aqa -tī -n*] *üüt -ü*  
 He father -3.SG -GEN milk-ACC  
*ih -iex -teek -i*] -n *bil -e -bin*  
 drink-FUT-MOD-3.SG -ACC know -AOR-1.SG  
 Intended reading: The same as for (33a). (Kornfilt & Vinokurova 2001)

Conclusion: the relic genitive can show up only on (complex, third person) specifiers of bona fide DPs, enforcing the interpretation of such specifiers as possessives; subjects of nominalized clauses don’t qualify.

If the raising analysis were correct for the subject of Sakha non-subject RCs, we would expect for such a complex subject, if it is a third person, to be marked with the relic genitive. While this is possible, the resulting construction has somewhat different syntactic and semantic properties than the RC with the nominative subject, where the subject has obviously the same case it has in a regular embedding such as in (33a):

- (34) a. [[kini aqa -ta ] öl -ör -büt] oquh-a  
 he (NOM) father-3.SG (NOM) die -CAUS-P ox -3.SG  
 ‘The ox which his father killed’ (Kornfilt & Vinokurova 2001)

The corresponding example with a DP that’s understood as the subject is, under this subject interpretation, less preferred (N. Vinokurova, personal communication). As alluded to above, that DP can, of course, be interpreted as the possessor and is fine under that interpretation:

- (34) b. ?[[kini aqa -ti ] -n öl -ör -büt] oquh -a  
 he(NOM) father-3.SG -GEN die-CAUS-P ox -3.SG

Reading with less preferred grammaticality judgment: ‘The ox which his father killed’

Interpretation with perfect grammaticality judgment: ‘His father’s ox which he (his father) killed’ (Kornfilt & Vinokurova 2001)

It appears to be questionable whether RCs with the genitive subjects are instances of genuine RCs at all. In addition to the fact, already mentioned, that they are less preferred (for the non-possessor, subject reading of the clause-initial DP), at least by a number of speakers, they seem to have different syntactic properties, as mentioned earlier. According to N. Vinokurova (p.c.), the traditional view of RCs with genitive subjects is to characterize them not as RCs, but as noun phrases with participial modifiers; this view, according to her, is supported by the fact that the putative genitive subject can be separated from the predicate, for example by other modifiers of the head; e.g.:

- (35) [[kini aqa -ti-n] **eder** atılas-pit ] at-a  
 he(NOM) father-3.SG-GEN young buy -P horse-3.SG  
 ‘the young horse which his father bought’ (i.e. ‘the young, bought (by his father) horse’; interestingly, this can also mean ‘his father’s young horse which he, his father, bought’; N. Vinokurova, p.c.)

In contrast, the RC version with the nominative subject, preferred for the non-possessive subject reading for that DP, does not tolerate such separation of the subject from the predicate by material which is external to the clause:

- (36) \*[[kini aqa-ta] **eder** atılas-pit] at-a  
 he(NOM) father-3.SG (NOM) young buy -P horse-3.SG

Intended reading: ‘the young horse which his father bought’<sup>8</sup>

This is not surprising, if in (35), the genitive DP is base-generated as a possessor of the head, i.e. in Spec,DP position (which is consistent with the fact that this example also has a reading under which the genitive DP is interpreted as the possessor of the head), and if the predicate is not directly predicated of that possessor, but is a participial modifier of the head. As such a modifier, the participle can be preceded by other modifiers of the head. The “possessor” DP is then either syntactically (via a silent element, i.e. PRO) or perhaps only pragmatically linked to the participle and takes on a secondary reading as its subject. In contrast, (36) is a genuinely clausal headed construction, with the subject in-situ and hence in the nominative; therefore, no clause-external material such as an additional adjectival modifier of the head can be inserted into the clause.

Therefore, derived locality between the subject of the clause and the clause-external subject Agreement element on the RC-head via a raising operation to Spec,DP is empirically problematic and unlikely.

Conclusion: Sakha RCs with apparent genitive subjects are actually base-generated with a genitive DP in specifier-of-DP position, and are not the result of subject raising. They have a different syntax from that of a genuine RC, while the construction with nominative subject is indeed a genuine RC, but does not undergo raising of the subject to Spec, DP. But if we cannot derive a local relation between the subject and the clause-external Agreement marker via subject raising, there must be another way of doing so; I claim that this is via Agree,

<sup>8</sup> These data about the contrasts between these headed constructions with nominative versus genitive subjects were originally provided by Nadya Vinokurova and were subsequently checked with Vladimir Monastirev and Svetlana Prokopiëva. I thank all three native informants. Additional discussion of the raising versus in-situ controversy concerning Sakha can be found in Baker & Vinokurova (2010) and Kornfilt (2009b).

which can take place across the clause boundary, given that the clause in Sakha RCs is not a CP, and thus not a phase. Therefore, this Agree relationship is indeed local, as it takes place within a phase: the higher DP.

#### 4.4 Second empirical argument against subject raising in RCs: Sakha RCs embedded in possessive DPs

There are instances in Sakha of RCs within higher possessive phrases, where the phi-features of the possessor and those of the subject differ:

- (37) [[aqa -n ] öl -ör -büt] (min) oquh -um  
 father-2.SG(NOM) die -CAUS -P (I[NOM]) ox -1.SG  
 ‘My ox which your father killed’ (Kornfilt & Vinokurova 2001)

Note that the RC-head bears the Agreement marker for the possessor, i.e. for the specifier of the higher DP, and not the marker for the subject, which would have been third person singular. This is not (or perhaps not only) an issue of a hierarchy between different persons, with, for example, the first person winning over the third, as the next example shows:

- (38) [min öl -ör -büt] (kini) oquh -a  
 I(NOM) die -CAUS -P he (NOM) ox -3.SG  
 ‘His ox which I killed’ (Kornfilt & Vinokurova 2001)

Here, the (syntactically) higher “possessive Agreement” wins over the (syntactically) lower “subject Agreement”, although the winner is the third person Agreement marker, usually classified as lower on person hierarchies than the first person, over which it has obviously won in this example.<sup>9</sup>

The fact that not both Agreement morphemes, i.e. one for the subject, and one for the possessor, can be displayed on the head *oquh* ‘ox’, is due to a constraint against immediate sequences of the same type of morpheme (cf. Kornfilt (1986), where the constraint is labeled “the Stuttering Prohibition”, and Göksel 1997).

The importance of these examples arises from the following reasoning: If there is a possessor for the head in the construction, and if that possessor is of a different person than the subject, it becomes difficult or even impossible to rely on an analysis involving the raising of the subject to a “possessor” (i.e. Spec,DP) po-

<sup>9</sup> A person hierarchy is somewhat involved, however. In data provided more recently with Vinokurova (p.c.), involving subjects or possessors that consist of third person possessive phrases (and which therefore can potentially be marked with the relic genitive morpheme), examples with third person subjects and first or second person possessors are judged to be fine, while examples with first or second person subjects and complex (and thus genitive) third person possessors are judged to be bad. Furthermore, examples where both the subject and the possessors are complex and in the third person are judged to be acceptable, with the possessors obligatorily in the genitive, and the subject preferably nominative. The less-preferred examples of this sort where the complex subject is also genitive would be characterized as analogically copying the possessive genitive in the construction, given that there is no licenser for it, i.e. no separate, dedicated agreement morphology; the head of the construction would be the entire possessive expression:

- (i) Lena aqa – tyn aldjap – pyt Masha aqa – tyn cahyy-ta  
 Lena father-GEN break-PSTPRT Masha father-GEN watch-3.SG  
 ‘Masha’s father’s watch which Lena’s father broke’

As with other RCs, the preferred counterpart has a nominative subject:

- (ii) Lena aqa – ta aldjap – pyt Masha aqa – tyn cahyy-ta  
 Lena father-3.SG (NOM) break-PSTPRT Masha father-GEN watch-3.SG  
 ‘Masha’s father’s watch which Lena’s father broke’

In such expressions, I claim, the entire possessive sequence *Masha aqatyn cahyyta* ‘Masha’s father’s watch’ is the head of the construction and is what’s relativized in (ii), and is the head of the construction in the “fake” RC in (i), too. The fact that this sequence cannot be interrupted, and that the possessor in particular cannot be separated from the rest of the possessive sequence by the relative clause (N. Vinokurova, p.c., and also checked with Monastirev and Prokopieva) supports this analysis.

It should be noted that not all native speakers of accept examples of possessed RCs (or, given the analysis advocated here, with the entire possessed head being the target of the RC, these should perhaps be called RCs with possession). In Göksel (1997), such examples are reported as ill-formed, based on the judgments of the author’s informants; Vinokurova herself and her informants do accept them as well-formed; Monastirev and Prokopieva seem to be somewhere between those extremes.



sition, given that the target position of this raising will be occupied by the overt possessor. Therefore, in the examples such as (37) and (38), the subject is in-situ, as maintained in this paper throughout.

Analysis: The in-situ subject and the Agreement on the head (i.e. on the D-head of the entire DP) are in a local relationship—a relation crossing the clause but nonetheless local, because the modifying clauses of RCs in languages such as Sakha are reduced with respect to their counterparts in Turkish. I propose to impute the category of TAM [Tense/Aspect/Mood]-P to them. They are not CPs. This is also why the Agreement element does not show up on the clause’s predicate: as proposed in Kornfilt (2008) and in Miyagawa (2008; 2011), Agreement marks a phase-head, and thus shows up on the CP-clause in Turkish RCs, and on the DP-head of Sakha RCs, given that in Sakha, the clause is not a CP (and in Turkish, the clause is a CP and thus the smallest relevant phase, i.e. lower and smaller than the higher DP-phase).

#### 4.5 The A'-disjointness requirement in Sakha

In further support of locality for Agr (on D) and the in-situ subject, I advance the following consideration: Csató’s generalization holds in Sakha RCs just as it holds in Turkish RCs: subject RCs cannot exhibit local subject Agreement—due to the principle in (19), i.e. the A'-binding restriction.

The Sakha facts can be explained via this principle only if the RC-head and its Agreement marker are in a local relationship with the subject of the modifying clause.

Subject RCs in Sakha, showing the lack of Agreement morphology with the subject, thus contrasting with non-subject RCs, are illustrated by the following example:

- (39) [e<sub>i</sub> üüt ih -iex -teex ] it<sub>i</sub>  
           milk drink-FUT -MOD dog  
       ‘The dog which should drink the milk’

There is no overt Agreement with the subject here—neither on the predicate of the modifier clause, nor on the RC-head. In this respect, the subject RC in Sakha is similar to its Turkish counterpart.

Note also the contrast with a Sakha non-subject RC:

- (40) [it e<sub>i</sub> ih -iex -teex ] üüt<sub>i</sub> -e  
       dog (NOM) drink-FUT -MOD milk -3.SG  
       ‘The milk which the dog should drink’

In (39), then, the variable in subject position and the RC-head are in a local relationship, because they are in the same phase, by virtue of the embedded clause not being a CP. Agreement on the RC-head would therefore have violated the principle in (19), i.e. the generalized binding principle against *locally* bound resumptive pronouns.

The predicates of the embedded clauses are identical in these last two examples, even though (39) is a subject RC, and (40) is a non-subject RC. The subject—non-subject asymmetry we have observed and discussed for Turkish RCs holds only with respect to the absence versus presence of overt Agreement, but not with respect to different predicate shapes. Given that we had treated the difference in predicate shapes as a diagnostic for CP-status of the clause, again we conclude that Sakha modifying clauses in their RC-constructions are not CPs but rather are projections smaller than CP, thus making an Agree relation between the *Agr* element on the clause-external D and the clause-internal, in-situ subject possible. The fact that the A'-disjointness Condition treats the subject position and the clause-external Agreement element as standing in a local relationship to each other offers additional support to the analysis advocated.

4.6 *Modern Uighur RCs*

In Modern Uighur non-subject RCs, *Agr* is placed on the RC-head, similar to Sakha non-subject RCs. This is illustrated by the next example:

- (41) [sin -iŋ e<sub>i</sub> izdä -ydiyan] adəmi -iŋ  
 you -GEN search for -FUT man, person-2.SG  
 köč -üp kät -ti  
 move-CONV go/leave<sup>10</sup> -PST  
 ‘The person whom you will look for has moved away/left’ (Abdurishid Yakup, p.c.)

Also as in Sakha, the predicate of the modifier clause does not change its shape according to the target of the RC, while it can change its shape with respect to Tense or Aspect:

- (42) [e<sub>i</sub> ürümçi -dɛ tur - idiyan ] siŋl<sub>i</sub> -iŋ  
 Ürümçi -LOC live-FUT/PRES younger sister-2.SG  
 ‘Your younger sister who lives in Ürümçi’ (LeSourd 1989)
- (43) [(min -iŋ ) e<sub>i</sub> tut - idiyan ] at<sub>i</sub> -im  
 I -GEN catch -FUT/PRES horse-1.SG  
 ‘The horse that I catch/will catch’ (LeSourd 1989)

The shape of the predicate can change with a change in the Tense/Aspect, even where the target remains the same:

- (44) [min -iŋ al -yan ] xotun-um dunya-da bir  
 I -GEN take -PST/PERF lady -1.SG world -LOC one  
 ‘The lady I married is unique in the world’ (Abdurishid Yakup, p.c.)

There are additional similarities to RCs in Sakha; for example, subject RCs are similar to those in Sakha, as well: Csató’s Generalization holds in Uighur, too. In other words, subject RCs cannot display Agreement with the subject on the RC-head:

- (45) [ e<sub>i</sub> kel -gen ] kiši<sub>i</sub>  
 come -PST/PERF person  
 ‘A/The person who has come’ (Johanson 1998, 61)

However, somewhat differently from Sakha, we do find overt genitive marking in a productive way; in other words, genitive morphology is not limited to (certain) third person subjects. The genitive on the subject is dependent on the Agreement morphology on the RC-head:

- (46) a. [(sin -iŋ ) e<sub>i</sub> kör -idiyan ] adəmi -iŋ  
 you -GEN see -FUT/PRES man -2.SG  
 ‘The man you will see’ (LeSourd 1989)

Uighur has a second type of non-subject RC—a type which, in some other Turkic languages, is the only type available, and which we saw in the first part of this paper to have been the only RC type documented in the Old Turkic inscriptions; that type does not have any overt subject Agreement morphology at all—neither on the RC-head, nor on the predicate; crucially, the subject is bare, i.e. in the nominative:

- (46) b. [sen e<sub>i</sub> kör -idiyan ] adəmi  
 you see -FUT/PRES man  
 ‘The man you will see’ (LeSourd 1989)

In the version with the genitive subject, the genitive is clearly licensed by the overt Agreement on the RC-head; a genitive subject leads to ill-formedness in the absence of overt Agreement:

<sup>10</sup> While the root of this verb is glossed correctly for occurrences of this verb as a main verb in general, it is used here as part of a serial verb, and is thus used as a “light verb” or auxiliary.

- (46) c. \*[sɨn -iŋ e<sub>i</sub> kōr -idiyan ] adəm<sub>i</sub>  
 you -GEN see -FUT/PRES man  
 Intended reading: Same as in (46)a. and b. (LeSourd 1989)

There are two important factors about overt Agreement in Uighur non-subject RCs that emerge from these examples and which we have to note: 1. overt Agreement expresses the phi-features of the apparently non-local subject; 2. this overt Agreement licenses the genitive case on this seemingly non-local subject. Therefore, the question arises once again whether the subject is raised to the specifier position of the higher DP, i.e. to “possessor” position.

Modern Uighur does offer support for an in-situ analysis of the embedded subject, against a raising analysis; the evidence is of a sort which is different from what we had seen in Sakha, and is based on word order. If the subject could raise, out of its clause, to a higher position within the DP, then it should also be able to raise to other positions outside of its clause; however, this is not possible, as the following examples illustrate. The constituents used to check potentially possible positions will be, in addition to the subject, two different adverbs, and their construal possibilities with the embedded versus the matrix predicate will be important.

- (47) [[ min-iŋ εtε bar -idiyan] yer -im ]  
 I -GEN tomorrow arrive -FUT place -1.SG  
 enigki nahayiti yiraq  
 obviously very far  
 ‘The place where I will arrive/go tomorrow is obviously very far’  
 (Abdurishid Yakup, p.c.)

The adverb of the modifier clause, εtε ‘tomorrow’, can freely scramble over the subject, as long as this is local scrambling, i.e. within the clause:

- (48) [[εtε min-iŋ bar -idiyan] yer -im ]  
 tomorrow I -GEN arrive -FUT place -1.SG  
 enigki nahayiti yiraq  
 obviously very far  
 Same meaning as (47) (Abdurishid Yakup, p.c.)

Similarly, the root adverb *enigki* ‘obviously’ can freely move within the root clause:

- (49) enigki [[εtε min-iŋ bar -idiyan] yer -im ]  
 obviously tomorrow I -GEN arrive -FUT place -1.SG  
 nahayiti yiraq  
 very far  
 Same meaning as (47) and (48) (Abdurishid Yakup, p.c.)

However, neither the genitive subject, nor the adverb of the embedded clause can move into the root clause:

- (50) \* min-iŋ enigki [[εtε bar -idiyan] yer -im ]  
 I -GEN obviously tomorrow arrive-FUT place -1.SG  
 nahayiti yiraq  
 very far  
 Intended reading: Same as in (47), (48), and (49) (Abdurishid Yakup, p.c.)

- (51) \*εtε enigki [[ min-iŋ bar -idiyan] yer -im ]  
 tomorrow obviously I -GEN arrive -FUT place -1.SG  
 nahayiti yiraq  
 very far  
 Intended reading: Same as in (47), (48), and (49) (Abdurishid Yakup, p.c.)<sup>11</sup>

<sup>11</sup> These examples were constructed along similar ones in LeSourd (1989). I have followed here the guidance of Dr. Abdurishid Yakup, a native Uighur speaker and Turkologist, who helped me construct examples which he found to be more colloquial and acceptable.

(Note that in (49), the immediate sequence of two adverbs is fine, as long as the root adverb is higher than the embedded adverb; the reversed order and thus hierarchy between the two adverbs leads to ill-formedness in (51).)

While the ill-formedness of these last two examples might be explained by referring to subjacency effects, i.e. to the unsuccessful attempt to move constituents out of a complex DP, this is actually a piece of evidence against a raising analysis of genitive subjects in Uighur non-subject RCs.

In Uighur, as in other Turkic languages, possessors in possessive phrases with an Agreement marker on the head scramble quite freely out of the possessive phrase. Thus, the fact that (50) with its topicalized subject is ill-formed despite the Agreement marker on *yer* ‘place’, the RC-head, shows that the topicalization originated from the subject position, rather than from the position of the specifier of the higher DP, i.e. from the possessor position—and it is the possessor position which would have been the source of the topicalization, if the subject had undergone raising to Spec, DP.

Conclusion: the genitive subject in Uighur non-subject RCs remains in-situ.

Secondly, the well-formedness of (48), in conjunction with the preceding discussion, also argues against raising of the genitive subject to a higher Spec,DP position. If the subject had risen to the specifier position of the higher DP, so as to enter a local relationship with the Agreement element on the RC-head, then the temporal adverb must also have risen to an even higher position than the subject (given that the adverb precedes the subject), presumably to some high topic position. But ill-formed examples such as (51) show that adverbs can scramble only locally; hence, we conclude that in (48), the adverb has scrambled to a clause-internal topic position; but if the adverb is still within the clause, then the genitive subject that follows that adverb is in-situ and has not undergone raising.

Kazakh and Kirghiz, two Turkic languages with RCs very similar to those in Uighur, offer similar evidence.

Kazakh:

- (52) Erteŋ men-iŋ älbette bar -atın žer -im alısta  
 Tomorrow I -GEN certainly arrive-FUT place-1.SG far  
 ‘\*The place where I will arrive tomorrow is certainly far.’  
 ‘The place where I will certainly arrive tomorrow is far.’ (Raihan Muhamedowa, p.c.)

Here, the temporal adverb of the embedded clause and the genitive subject precede the adverb which, in an “unscrambled” counterpart, is primarily construed with the matrix predicate:

- (53) Älbette erteŋ men-iŋ bar -atın žer -im alısta  
 Certainly tomorrow I -GEN arrive-FUT place-1.SG far  
 ‘The place where I will arrive tomorrow is certainly far.’  
 ‘The place where I will certainly arrive tomorrow is far.’ (Raihan Muhamedowa, p.c.)

In (53), *älbette* ‘certainly’ can be construed with the embedded predicate, but it can also be construed with the matrix predicate ‘be far’. This reading is not possible in (52), where the same adverb can only be construed with the embedded predicate ‘arrive’. This shows that neither the embedded temporal adverb, nor the genitive subject of the embedded clause, are able to move out of that clause; if they had been able to, the second adverb *älbette* ‘certainly’ should have been able to act as a matrix constituent, which it is not.

Kirghiz:

The sentence-initial adverb *sözsüz* ‘doubtlessly, certainly’ can be construed with either the embedded or the matrix predicate:

- (54) Sözsüz Ali-nin erteñ bara tur -gan žer -i uzak  
 Certainly Ali-GEN tomorrow arrive AUX-P place-3.SG far  
 ‘The place where Ali will arrive tomorrow is certainly far’  
 ‘The place where Ali will certainly arrive tomorrow is far’ (Kenjegül Kalieva, p.c.)

In the following example, where the genitive subject precedes the adverb *sözsüz* ‘certainly’, that adverb can only be construed with the embedded predicate:

- (55) Men-in sözsüz erteñ bara tur -gan žer -im uzak.  
 I -GEN certainly tomorrow arrive AUX -P place-1.SG far  
 ‘\*The place where I will arrive tomorrow is certainly far’  
 ‘The place where I will certainly arrive tomorrow is far’ (Kenjegül Kalieva, p.c.)

The genitive subject cannot move out of its clause, into the matrix clause. If that subject had risen into Spec, DP of the RC, then such movement into the matrix should not have been problematic. I conclude that the genitive subject in Kirghiz non-subject RCs, just as its counterpart in Uighur and Kazakh, remains *in-situ*.

Let us now turn back to the locality of the relationship between the genitive subject and the overt Agreement on the RC-head: Both for the sake of phi-feature Agree (cf., among others, Chomsky 1999 and 2008), and for the sake of genitive case licensing, this relationship should be local. If this can’t be achieved via the raising of the subject to the specifier of the RC’s head/D, then it must be the case that the probe-goal relationship in question is indeed local, in spite of being established with the *in-situ* subject.

Proposal: The same approach already proposed for Sakha (as well as for Old and Middle Turkic) can succeed for Uighur, as well as for Kazakh and Kirghiz: the modifier clause is not a CP, but just a bare TAM-phrase. Therefore, the subject and the RC-head/D are in the same phase (namely, the DP), and the relationship between the two is thus local. The fact that Csató’s Generalization holds, as well as the lack of any “*que-to-qui* conversion” effects involving C and the level of CP in these “Central Asian” languages, are two pieces of independent evidence that support this proposal.

## 5. Conclusions

This study proposes that non-default, *licensed* subject (and possessor) case is licensed locally, within a relevant, complete syntactic domain, via an Agree relationship between a probe (D and the associated *Agr*, or C and the associated *Agr*), and the goal, i.e. the subject. In constructions such as relative clauses, where a clause is headed by a clause-external nominal, licensed subject case is possible on an *in-situ* subject, licensed by C (and the associated *Agr*), as in Turkish, when the clause is a CP. When the clause is smaller than a CP, and thus smaller than a complete syntactic domain, the probe can be clause-external, i.e. D and the associated *Agr*; here, the smallest relevant phase is the DP; examples were Middle Turkic as well as Sakha, Uighur, Kazakh and Kirghiz.<sup>12</sup>

In its historical part, the paper has also claimed that Turkish must have developed larger functional projections in the course of its development, and that in particular the CP is a “Mediterranean” development, which sets Turkish (and its very close relative, Azerbaijani/Azeri) apart from its Central and Northern Asian sister languages. I would tentatively suggest that those languages have also expanded their inventory of functional projections, e.g. having already inherited an Aspect-Phrase projection, they have probably developed a Tense-Phrase projection, but that they have not (yet?) developed a CP (at least not in their externally headed constructions such as RCs and noun-complement constructions).

<sup>12</sup> A somewhat different approach to these facts was taken in Kornfilt (2005), based on Kayne (1994).

**Abbreviations:***A. Orkhon Inscriptions*

BK: Bilge Kagan  
 O: Ongin  
 KT: Kül Tigin  
 T I: Tonyukuk, first monument  
 E: Eastern façade of a monument  
 W: Western façade of a monument  
 F: Front façade of a monument  
 S: Southern façade of a monument  
 N: Northern façade of a monument

The numbers after the abbreviations for the monument's name and the direction of a façade refer to the line from which the example is taken.

*B. Glosses*

ABL	Ablative
ACC	Accusative
AGR	Agreement
AGR(S)P	(Subject) Agreement Phrase
AOR	Aorist
ASP	Aspect
ASPP	Aspect Phrase
AUX	Auxiliary
CAUS	Causative
CFC	Complete Functional Complex
CMPD, CMPDMRK	Compound, Compound marker
CONV	Converb
CP	Complementizer Phrase
DP	Determiner Phrase
FN	Factive Nominalization
FUT	Future
FUTPART	Future participle
GEN	Genitive
LOC	Locative
MOD	Modal
MST	Modern Standard Turkish
NEG	Negation, Negation marker
NOM	Nominative
NP	Noun Phrase
OP	Operator
OPT	Optative
OT	Old Turkic
P	Participle
PASS	Passive
PL	Plural
PRES	Present
PRF	Perfect
PRFPART	Perfect Participle
PRSPART	Present Participle
PRSPRG	Present Progressive
PST	Past
PSTPRT	Past Participle
RC	Relative Clause
SG	Singular



T	Tense
TAM	Tense/Aspect/Mood
TP	Tense Phrase
VP	Verb Phrase
1.	First person
2.	Second person
3.	Third person

## References

- Aoun, Y. 1986. *Generalized Binding, the Syntax and Logical Form of Wh-interrogatives*. Dordrecht: Foris.
- Aygen, G. 2006. Case and agreement in reduced versus full relative clauses in Turkic languages. In S. Yağcıoğlu & A. C. Değer eds. *Advances in Turkish Linguistics*. İzmir: Dokuz Eylül Yayınları; 643–653.
- Baker, M. & Vinokurova, N. 2010. Two Modalities of Case Assignment: Case in Sakha. *Natural Language and Linguistic Theory* 28:3, 593–642.
- Banguoğlu, T. 1938 *Altosmanische Sprachstudien zu Süheyl-ül Nevbahar*. Doctoral Dissertation, Breslau.
- Borer, H. 1984. Restrictive relatives in modern Hebrew. *Natural Language and Linguistic Theory* 2: 219–260.
- Chomsky, N. 1999. Derivation by phase. *MIT Occasional Papers in Linguistics* 18.
- Chomsky, N. 2008. On phases. In R. Freidin, C. P. Otero, and M. L. Zubizarreta eds. *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud*. Cambridge, Mass.: MIT Press.
- Comrie, B. & E. L. Keenan 1979. Noun Phrase Accessibility Revisited. *Language* 55:3; 649–664.
- Csató, É. 1996. A typological review of relative clause constructions in some Turkic languages. In B. Rona ed. *Current Issues in Turkish Linguistics: Proceedings of the Fifth International Conference on Turkish Linguistics*. Ankara: Hitit Yayınevi [copyright London: SOAS]; 8–32.
- Erdal, M. 1998. Old Turkic. In L. Johanson & É. Csató ed. *The Turkic Languages*. London: Routledge; 138–157.
- Göksel, A. 1997. Morphological asymmetries between Turkish and Yakut. In K. İmer & N. E. Uzun eds. *Proceedings of the VIIIth International Conference on Turkish Linguistics*. Ankara: Ankara University publications; 69–76.
- Hale, K. 2002. On the Dagur Object Relative: Some Comparative Notes. *Journal of East Asian Linguistics* 11 (2); 109–122.
- Hale, K. & C. Ning. 1996. Raising in Dagur relative clauses. In B. Agbayani, K. Takeda, and S.W. Tang eds. *UCI Working Papers in Linguistics* 1; 35–38.
- Johanson, L. 1998. The structure of Turkic. In L. Johanson & É. Csató eds. *The Turkic Languages*. London: Routledge; 30–66.
- Kayne, R. S. 1976. French relative que. In M. Luján and F. Hensey eds. *Current studies in Romance linguistics*. Washington, DC: Georgetown University Press; 131–152.
- Kayne, R. S. 1994. *The Antisymmetry of Syntax*. Linguistic Inquiry Monograph 25. Cambridge, Mass.: MIT Press.
- Keenan, E. L. 1985. Relative clauses. In T. Shopen ed. *Language typology and syntactic description II: Complex constructions*. Cambridge: Cambridge University Press; 141–170.
- Keenan, E. L. & B. Comrie 1977. Noun Phrase Accessibility and Universal Grammar. *Linguistic Inquiry* 8:1; 63–99.
- Keenan, E. L. & B. Comrie 1979. Data on the Noun Phrase Accessibility Hierarchy. *Language* 55:2; 333–351.
- Kleinmichel, S. 1974. Beitrag zur Geschichte der Nomenverbindungen im Osmanisch-Türkischen (Izafet). In *Sprache, Geschichte und Kultur der altaischen Völker: Protokollband der XII. Tagung der Permanent International Altaistic Conference, 1969*; G. Hazai & P. Zieme 9 ed. Berlin: Akademie-Verlag, 1974; 307–316.
- Koopman, H. & D. Sportiche. 2008. The *que-qui* alternation: New analytical directions. UCLA ms.
- Kornfilt, J. 1984. *Case Marking, Agreement, and Empty Categories in Turkish*. Doctoral dissertation, Harvard University.
- Kornfilt, J. 1986. The stuttering prohibition and morpheme deletion in Turkish. In A. Aksu Koç & E. Erguvanlı-Taylan eds. *Proceedings of the Second International Conference in Turkish Linguistics*. Istanbul:

- Boğaziçi University Publications #400; 59–83.
- Kornfilt, J. 1988. A typology of morphological agreement and its syntactic consequences. In D. Brentari, G. Larson and L. MacLeod (eds.) *Papers from the Parasession on Agreement in Grammatical Theory*: 24-II. Chicago: Chicago Linguistic Society; 117–134.
- Kornfilt, J. 1991. Some current issues in Turkish syntax. In H. Boeschoten & L. Verhoeven (eds.) *Turkish Linguistics Today*. Leiden: Brill; 60–92.
- Kornfilt, J. 2000. Some Syntactic and Morphological Properties of Relative Clauses in Turkish”. In A. Alexiadou, C. Wilder & P. Law eds. *The Syntax of Relative Clauses*. Amsterdam and Philadelphia: John Benjamins; 121-159.
- Kornfilt, J. 2003. Subject case in Turkish nominalized clauses. In U. Junghanns & L. Szucsich eds. *Syntactic Structures and Morphological Information*. Berlin: Mouton de Gruyter; 129–215.
- Kornfilt, J. 2005. Agreement and its Placement in Turkic Non-Subject Relative Clauses”. In G. Cinque & R. Kayne eds. *Handbook of Comparative Syntax*. Oxford: Oxford University Press; 513-541.
- Kornfilt, J. 2006. Agreement: The (unique and local) syntactic and morphological licenser of subject Case. In J. Costa & M. C. Figueiredo Silva eds. *Studies on Agreement*. Amsterdam/Philadelphia: John Benjamins; 141–171.
- Kornfilt, J. 2008. Subject case and Agr in two types of Turkic RCs. In Ulutaş, S. & C. Boeckx eds. *Proceedings of WAFL 4: MITWPL 56*. Cambridge, Mass.: MIT; 145-168.
- Kornfilt, J. 2009a. Placement of Agreement and Subject licensing in Turkish/Turkic relative clauses. In S. Tatevosov ed. *Investigations into Formal Altaic Linguistics: Proceedings of WAFL 3*. Moscow: MAKSPress; 69-87.
- Kornfilt, J. 2009b. Subject–Agreement correlations and their syntactic effects in some Turkic relative clauses. *Turkic Languages* 13:1. Wiesbaden: Harrassowitz; 70-96.
- Kornfilt, J. & N. Vinokurova. 2001. Transcriptions and notes of one field work session on Sakha. Ms., Syracuse University and Utrecht University.
- LeSourd, P. 1989. Transcriptions of four informant sessions on Modern Uighur. Ms., Indiana University.
- Levin, B. & M. Rappaport 1986. The Formation of Adjectival Passives. *Linguistic Inquiry* 17:4, 623-662.
- Lightfoot, D. W. 1979. *Principles of Diachronic Syntax*. Cambridge: Cambridge University Press.
- Meral, H. M. 2006. Resumptive pronouns in Turkish. In S. Yağcıoğlu & A. C. Değer eds. *Advances in Turkish Linguistics*. İzmir: Dokuz Eylül Yayınları; 223–233.
- Miyagawa, S. 2008. Genitive subjects in Altaic. In Ulutaş, S. & C. Boeckx eds. *Proceedings of WAFL 4, MIT-WPL 56*, Cambridge, Mass.; 181-198.
- Miyagawa, S. 2011. Genitive Subjects in Altaic and Specification of Phase. *Lingua* 121:7; special issue: *Nominizations in Linguistic Theory*; J. Kornfilt & J. Whitman (guest editors); 1265–1282.
- Ouhalla, J. 1993. Subject-extraction, negation and the anti-agreement effect. *Natural Language and Linguistic Theory* 11: 477–518.
- Pesetsky, P. 1982. Complementizer-Trace Phenomena and the Nominative Island Condition. *The Linguistic Review* 1:297-344.
- Rizzi, L. 1990. *Relativized Minimality*. Linguistic Inquiry Monograph 16. Cambridge, Mass.: MIT Press.
- Schinkewitsch, J. 1927. Rabgusis Syntax. In *MSOS XXX:II*.
- Schönig, C. 1992/93. Relativsatzbautypen in den sogenannten altaischen Sprachen. In *Acta Orientalia Academiae Scientiarum Hung.* Tomus XLVI (2-3); 327-338.
- Sportiche, D. 2008. In-situ subject relativization and subject extraction in French. UCLA ms.
- Sportiche, D. 2011. French Reative *Qui*. *Linguistic Inquiry* 42:1, 83-124.
- Taraldsen, K. T. 2002. The Que/Qui Alternation and the Distribution of Expletives. In P. Svenonius ed. *Subjects, Expletives, and the EPP*. Oxford: Oxford University Press; 29 – 42.
- Tekin, T. 1967. *A Grammar of Orkhon Turkic*: Uralic and Altaic Series, vol. 69. Bloomington: Indiana University Publications.
- Wasow, T. 1977. Transformations and the Lexicon. In Culicover, P.W., T. Wasow & A. Akmajian eds. *Formal Syntax*. New York: Academic Press; 327-360.



## What Did the Earliest Central Asian Languages Sound Like?

Christopher I. Beckwith  
*Indiana University*

### Abstract

For all but the most obscure modern spoken languages it is no longer much of a problem to get basic information on their phonology. We can find someone who speaks it and ask them to pronounce what we want, or we can find something on the internet. Even for medieval Central Eurasian languages we have a fairly good idea of their pronunciation once they begin to be written relatively widely and there are foreign transcriptions of the languages. But what about ancient and medieval languages that died many centuries ago, especially ones that do not have any known modern descendants? In fact, we often do not even know if they have any ancient relatives or modern descendants, partly because we are not sure what they sounded like.

This paper first briefly discusses the earliest known Central Eurasian steppe zone languages for which we have any linguistic data, namely Scythian, an Iranian language that has several descendants (including modern Ossetian) and Hsiung-nu, a language which has not been shown to be certainly related to any other language, though it is shown in a new article by Shimunek et al. (2015) that the brief 329 CE prophecy in Chinese transcription, which has been argued by some to be in Hsiung-nu, is actually in a slightly archaic, very interesting form of Turkic, while the people are actually referred to as Chieh (\*Kir) by the Chinese (i.e., they do not call them Hsiung-nu) and as Xwn “Huns” by the Sogdians.

The main focus of the paper is the sound of Tokharian (or “Tocharian”)—actually two languages (Tokharian A or “East Tokharian”, and Tokharian B or “West Tokharian”), which are preserved in early medieval manuscripts from East Turkistan (now the Xinjiang Uighur Autonomous Region of the People’s Republic of China), in eastern Central Asia. Unfortunately, when the Tokharians merged with the Uighurs late in the first millennium CE the Tokharian language became extinct. However, we luckily have many texts, some of them quite extensive, in Tokharian, which is written in an essentially alphabetic segmental script. Although some details of the pronunciation of the two languages remain uncertain, one of our most valuable sources for reconstructing Tokharian is its metrical system and music. Tokharian metrical patterns are now known, and it has been suggested by the Tokharianist G.J. Pinault and others that the patterns and their names are in fact tunes.

We know the metrical structure of the lone secular poem in Tokharian, the Anonymous Love Poem in Tokharian B, and we have some early medieval Tokharian orchestral music, known to the T’ang Chinese as the “music of Kucha” because its most famous practitioners came from Kucha, the capital of the Tokharian B-speaking people. Whole orchestras traveled to the Chinese capital, Ch’ang-an, where they performed their music for the T’ang court and completely reshaped Chinese music on Central Asian foundations. From China the new music made its way to Japan, where around a thousand years ago it was transcribed in a kind of tablature notation. It is still performed in Japan as part of the repertory known as gagaku (“Japanese Imperial Court Music”). The T’ang musical scores preserved in Japanese transcription, which have been studied by L. Picken and others, should give us a much better idea of how Tokharian verse sounded, including the Anonymous Love Poem in Tokharian B. If the Tokharianists and the musicologists reconstructing gagaku could only get together now, perhaps one day we might be able to sing the love poem in Tokharian, to the original Tokharian tune.

## **Turkish Suspended Affixation**

Jorge Hankamer  
*University of California Santa Cruz*

At the time of printing, the proceedings for this talk were not available. We apologize for any inconvenience.

## Contrast in Focus Persian and Turkish focus constructions

Simin Karimi  
*University of Arizona*

### Abstract

Persian and Turkish exhibit a number of striking similarities, including the following: (a) many verbal concepts are expressed by complex predicates (CPr) consisting of a non-verbal element (NVE) and a light verb (LV) (Folli et al. 2005, Megerdooimian 2002, Kornfilt 2003), (b) both languages exhibit scrambling, (c) they both mark direct objects for specificity (Enç1991, Karimi 2005), and (d) focus-sensitive elements may follow the non-specific object and the NVE in a hierarchical manner in both languages. There are, however, some important differences between these two languages: while contrastively focused arguments, including the non-specific object, can be displaced in Persian, they must be in-situ in Turkish. Similarly, the Persian NVE can be separated from the LV for emphasis, but not in Turkish. In this article, we concentrate on two types of focus known as information focus and identificational/contrastive focus (Kiss 1998), and try to account for the similarities and differences between these two languages in a principled way. We show that the focus-related similarities are accounted for by a post-syntactic operation that is based on specific structural properties in both languages. We further argue that the differences between the two languages with regard to contrastive focus boil down to the syntactic distinction they exhibit with respect to their clausal architecture.

*Keywords:* Complex Predicates; information focus; contrastive focus; post-syntactic operation.





Part II: Formal Linguistics and Research



# Uighur Consonants and Electropalatograph

Kaili Binu Abudureheman\*

Yu Hong Zhi

Wang Jian Bin

*Key Lab of China's National Linguistic Information Technology,  
Northwest University for Nationalities,  
Lan Zhou, China*

Mei Re Yi

*Foreign Language Institute,  
Northwest University for Nationalities,  
Lan Zhou, China*

## Abstract

This paper aims at analyzing the contact between tongue and hard palate during production of consonants and co-articulation in Uyghur by means of Electropalatograph (EPG). Under the same manner of place, plosives involve more tongue-palate contact than affricates, and affricates involve more contact with higher centrality than fricatives; laterals concentrate their tongue-palate contacts on alveolar, resulting high anteriority and centrality; nasal trill involve alveolar contact with high posteriority and centrality. Their tongue characteristics of onset are obtained: middle of tongue tip /d/, /t/ and behind of tongue /g/, /k/ involve much more tongue-palate contact; front of tongue tip /s/, /z/ and front of tongue /ʃ/, /ʒ/ involve more contact with higher centrality than front of tongue /tʃ/, /dʒ/; middle of tongue tip /l/ concentrate their tongue-palate contacts on alveolar, resulting high centrality; middle of tongue tip /n/ involve alveolar contact with high posteriority and centrality; middle of tongue tip /r/ contacts with tongue – palate, resulting high centrality and rearwardality.

*Keywords:* Electropalatograph, Uyghur, Coarticulation

## I. Introduction

Traditional Linguists have analyzed articulation places and manners of different consonants and coarticulations on the basis of auditory recognition, however, some of those founded theories are very abstract without convincing proves. Then a new subject, called experimental linguistics, emerged, which aims at convincingly explaining a series of phonetic phenomenon. With widely-using techniques and new research approach, the tide of experimental linguistics has brought a lot of new views towards traditional linguistics. In this paper, a useful approach research tool called electro-palatal will be introduced to deepen your understanding towards co-articulation in Uyghur.

## II. EPG and its artificial palatal

The increasing development of computer technique has changed human being's traditional mode of production and life completely. Radiography and palatograph have shown some basic information of consonant:

---

\*kelly751118@hotmail.com

its key palace of articulation and static linguopalatal contact, laying the foundation of the research on speech production. However, these equipment and methods cannot reflect the dynamic process of speech production. Fortunately, the invention of Electropalatography (EPG) makes up for the deficiency by recording detailed linguopalatal contact during speech production. EPG is a kind of artificial palatal used by researchers to analyze the speech production process of consonants. During speech production, the artificial palatal will transmit the tongue-palatal contact information to the computer. EPG involves a psedudo-palate with 62 electrodes or 96 electrodes to obtain the linguopalatal contact information. EPG has been used in the fields of correction of articulation disorders and foreign language teaching, and it is in recent years that EPG is widely used in the study of linguistic study. The psedudo-palate in this paper is embedded with 62 electrodes. It records 100 frames of linguopalatal contact information per minute.

The following is the structure of EPG artificial palatal.



Figure 1. Total contact of Uighur affricates.

Table I Syllable Type of Uyghur

Types of syllables	Total syllables
V	8
VC	192
VCC	249
CV	192
CVC	2868
CVCC	59
CVCV	168

From Fig. 1, it is can be observed that EPG artificial palatal is embedded with sixty two electrodes, which can acquire the contact information when tongue contacts the electrodes in the artificial palatal. All the sixty two electrodes will judge whether the tongue organ contacts them or not. Then a matrix included contact tongue-palatal information will be formed and stored as a single file.

### III. Experimental Program

#### A. Word list design

/a/, /i/, /u/ are the most special articulated vowels, for which /a/ is with the widest openness, /i/ is with the narrowest openness, and /u/ is with the most posterity. Then those three vowels are chosen to collocate with consoants in Uyghur. Table 1 is all the consonants that can be located after /a/, /i/, /u/.

Table II Collocation between vowels /a/, /i/, /u/ and consonants

conso- nants vowels	t	l	s	z	ʃ	q	h	j	ʧ	ŋ	ʁ	r	N
a	at	al	as	az	aʃ	aq	ah	aj	aʧ	aŋ	aʁ	ar	an
i	it	il	is	iz	iʃ	iq	ih	ij	iʧ	iŋ	iʁ	ir	in
u	ut	ul	us	uz	uʃ	uq	uh	uj	uʧ	uŋ	uʁ	ur	un

This paper will magnify the tiny differences among syllables with same vowel or same voiceless consonant by analyzing the collocation between vowels /a/, /i/, /u/ and consonants.



*B. Signals acquisition and processing*

Two speech producers, a young male and a young female are chosen to read our designed syllables in this experimental. And two speech producers are native Uyghur without speech production illness. EPG signals are recorded by Palatometer 6300, setting 100Hz as EPG sampling rate and 16000Hz as audio sampling rate.

After signal acquisition, all EPG and audio signals are processed by Matlab Program. The processing includes signal cutting and parameters compute. Signal cutting bases on the principle of choosing the most stable onset frame of EPG matrix as the representative of a consonant in a certain syllable. After signal cutting, the computing of chosen key matrixes is also practiced by Matlab to obtain some scientific parameters. And those parameters consist of two domains: contacted electrodes area percentage parameters and distribution parameters. In order to compare and analyze, all those parameters will be presented in percentage.

**IV. Coarticulation in Uighur VC syllable**

*A. Effect of vowel on consonant /t/*

Table 2 demonstrates the cut key matrixes of consonant /t/ in different syllables, however, all these syllables ends with the same consonant /t/, but begins with /a/, /i/ or /u/. From those matrixes, it can be observed clearly that there seems no obvious differences between male and female when they produce the same syllable. But there do exist vital differences among different syllables.

Table III EPG of consonants /t/ in different syllables

at		it		ut	
male	female	male	female	male	female

Among those matrixes, both /at/ matrixes is without full contact in the first two ranks, however, /it/ and /ut/ matrixes are nearly full contact in the first two rank. To make the comparison more vividly and scientifically, just move our focus on Fig. 2 and Figure3.

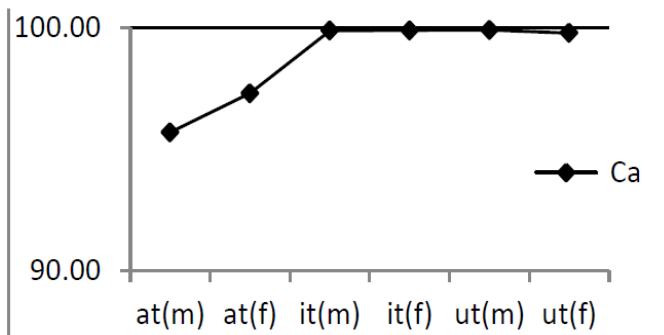


Figure 2. Contact Anteriority distribution percentage for consonant /t/.

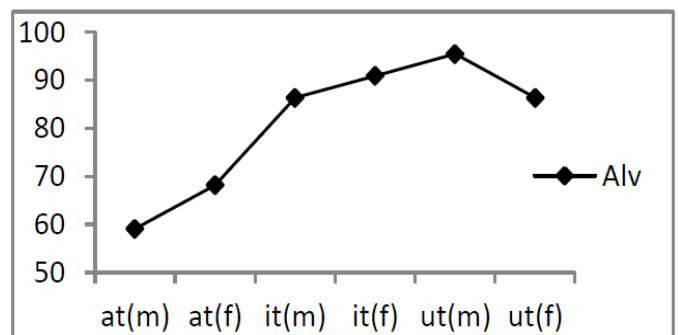


Figure 3. Contact alveolar distribution percentage for consonant /t/.

In Fig.2 and Figure.3, (m) means syllable pronounced by male, and (f) by female. It presents the tongue-palatal contact anteriority and alveolar distribution, which mainly includes the first three ranks of the matrix. Parameter Ca in syllable /at/ is relatively lower than that in syllables /it/ and /ut/. What's more, Param-

eter Alv is much more obvious in in their differences: Ca of /t/ in /at/ is much less than that in /it/ and /ut/. And this can be concluded: when producing /at/, speech producers don't constrain their tongue tip as strongly as producing /it/ and /ut/.

*B. Effect of vowel on consonant /s/*

Table 4 demonstrates the key matrixes of consonant /s/ in different syllable ended with the same consonant /t/. From those matrixes, male and female producers also indicate the speech production consistency in same syllable. But there do exist vital differences among different syllables. The biggest difference is velar contact among them.

*C. Effect of vowel on consonant /l/*

Fig. 4 demonstrates velar distribution percentage for consonant /l/, where the velar percentage of /us/ is higher than the others. That's because the tongue-palatal contact of /us/ is more in the last rank / than that of /as/ and /is/. In the last rank of the key matrixes, there exists more contact on its right in /us/, two electrodes, however, there only one contacted electrode on the right of the last rank in /as/ and /is/.

Table IV EPG of consonants /s/ in different syllables

as		is		us	
male	female	male	female	male	female
■□□□□□	□□□□□□	□□□□□□	□□□□□□	□□□□□□	□□□□□□
■ ■ ■ □ □ □ ■ ■	■ ■ □ □ □ □ ■ ■	□ ■ □ □ □ □ □ □	■ ■ □ □ □ □ ■ ■	■ ■ ■ □ □ □ □ □	■ ■ ■ □ □ □ □ □
■ □ ■ □ □ ■ ■ ■	■ ■ □ □ □ □ ■ ■	■ □ ■ □ □ ■ ■ ■	■ □ ■ □ □ ■ ■ ■	■ □ ■ □ □ ■ ■ ■	■ □ ■ □ □ ■ ■ ■
■ □ □ □ □ □ ■ ■	■ ■ □ □ □ □ ■ ■	■ □ □ □ □ □ ■ ■	■ □ □ □ □ □ ■ ■	■ □ □ □ □ □ ■ ■	■ □ □ □ □ □ ■ ■
■ □ □ □ □ □ ■ ■	■ □ □ □ □ □ ■ ■	■ □ □ □ □ □ ■ ■	■ □ □ □ □ □ ■ ■	■ □ □ □ □ □ ■ ■	■ □ □ □ □ □ ■ ■
■ □ □ □ □ □ ■ ■	■ □ □ □ □ □ ■ ■	■ □ □ □ □ □ ■ ■	■ □ □ □ □ □ ■ ■	■ □ □ □ □ □ ■ ■	■ □ □ □ □ □ ■ ■
■ □ □ □ □ □ ■ ■	■ □ □ □ □ □ ■ ■	■ □ □ □ □ □ ■ ■	■ □ □ □ □ □ ■ ■	■ □ □ □ □ □ ■ ■	■ □ □ □ □ □ ■ ■
■ □ □ □ □ □ ■ ■	■ □ □ □ □ □ ■ ■	■ □ □ □ □ □ ■ ■	■ □ □ □ □ □ ■ ■	■ □ □ □ □ □ ■ ■	■ □ □ □ □ □ ■ ■

Table 4 demonstrates the key matrixes of consonant /l/ in different syllable ended with the same consonant /t/. There seems exists no speech production difference in gender, but exist obvious differences among different syllables. Both sides of the last four ranks electrodes are contacted in syllable /il/ while it is a totally different situation in syllables /al/ and /ul/. However, the contact mode in syllables /al/ and /ul/ are almost the same.

To make the date more valid and convincible, the above key matrixes are computed, then the computed data is presented in Fig. 4, which demonstrate the post palatal distribution percentage for consonant /l/. It is can be observed apparently that consonant in syllable /il/ is with much higher post palatal distribution.

TABLE V EPG of consonants /s/ in different syllables

al		il		ul	
male	female	male	female	male	female
■ □ ■ ■ ■ □	■ □ ■ ■ ■ ■	■ ■ ■ ■ ■ ■	■ □ ■ ■ ■ ■	□ □ ■ ■ ■ ■	■ □ ■ ■ ■ ■
□ □ ■ ■ ■ □ ■ ■	□ ■ ■ ■ ■ □ ■ ■	□ ■ ■ ■ ■ □ ■ ■	■ ■ ■ ■ ■ □ ■ ■	■ ■ ■ ■ ■ □ ■ ■	□ ■ ■ ■ ■ □ ■ ■
■ □ ■ □ □ □ ■ ■	■ □ ■ □ □ □ ■ ■	■ □ ■ □ □ □ ■ ■	■ □ ■ □ □ □ ■ ■	■ □ ■ □ □ □ ■ ■	■ □ ■ □ □ □ ■ ■
■ □ □ □ □ □ ■ ■	■ □ □ □ □ □ ■ ■	■ □ □ □ □ □ ■ ■	■ □ □ □ □ □ ■ ■	■ □ □ □ □ □ ■ ■	■ □ □ □ □ □ ■ ■
□ □ □ □ □ □ □ □	□ □ □ □ □ □ □ □	■ □ □ □ □ □ ■ ■	■ □ □ □ □ □ ■ ■	□ □ □ □ □ □ □ □	□ □ □ □ □ □ □ □
□ □ □ □ □ □ □ □	□ □ □ □ □ □ □ □	□ □ □ □ □ □ □ □	■ □ □ □ □ □ ■ ■	□ □ □ □ □ □ □ □	□ □ □ □ □ □ □ □
□ □ □ □ □ □ □ □	□ □ □ □ □ □ □ □	□ □ □ □ □ □ □ □	□ □ □ □ □ □ ■ ■	□ □ □ □ □ □ □ □	□ □ □ □ □ □ □ □
□ □ □ □ □ □ □ □	□ □ □ □ □ □ □ □	□ □ □ □ □ □ □ □	□ □ □ □ □ □ ■ ■	□ □ □ □ □ □ □ □	□ □ □ □ □ □ □ □
□ □ □ □ □ □ □ □	□ □ □ □ □ □ □ □	□ □ □ □ □ □ □ □	□ □ □ □ □ □ ■ ■	□ □ □ □ □ □ □ □	□ □ □ □ □ □ □ □

## V. Summary

In Uyghur, the voiceless consonant would convey some articulation places and manners information of its anterior vowels within syllables. Under the same manner of place, plosives involve more tongue-palate contact than affricates, and affricates involve more contact with higher centrality than fricatives; laterals concentrate their tongue-palate contacts on alveolar, resulting high anteriority and centrality; nasal trill involve alveolar contact with high posteriority and situations lead to that the tongue characteristics of the same consonant after different kinds of vowels are different from each other by slightly changing their original articulation places to coarticulated with their anterior vowels, sometimes even changing their articulation manners. However, the degree of this kind of affection hasn't analyzed, which will be investigated in future research.

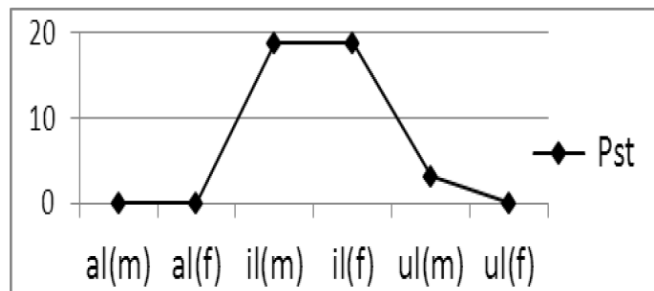
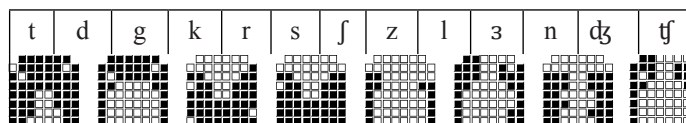


Figure 4. Post palatal distribution percentage for consonant /l/.



## Acknowledgement

This research was financially supported by the project of National Social Science Fund(Grant NO.14XYY013-XBMU-2014-AA-25).

## References

- Morgan, Barry RA. EPG from square one: An overview of electropalatography as an aid to therapy, *Clinic Linguistics and Phonetics*, 1989, 3(1):81—91.
- Farnetani, E. (1997). Coarticulation and connected speech processes. In Hardcastle, W. J. & Laver, J. (eds.) *The Handbook of Phonetic Sciences*, 371-404. Oxford: Blackwell publisher.
- Liu Jia. *Articulation place of consonants in standard Chinese: research based on EPG*. Zhejiang University, 2006.
- Li Yonghong, The Research on Tibetan Consonants in Xiahe Dialect by Means of Electropalatograph, *The 2nd International Conference on Multimedia Technology*, Vol.6, 201.

# Tungusic converbs in *-mi* from the perspective of linguistic area<sup>1</sup>

Sangyub Baek  
Hokkaido University

## Abstract

A frequently used converbal suffix *-mi* in Tungusic shares common forms and similar semantic functions throughout the Tungusic languages. Thus, it is a general view that *-mi* originated from the Proto-Tungusic. Previous studies, however, have a tendency to give a rather different description to *-mi* in each Tungusic language according to classification and areal distribution. This paper aims to clarify the functional differences among converbal forms in *-mi* in accordance with Tungusic distribution (Russian/Chinese territories) by employing typological parameters of converbs (Nedjalkov, V.P. 1995). In addition, we also apply the same parameters to functionally corresponding imperfective converbs of Russian and Mongolian to examine the possibility that the functional differences in *-mi* among Tungusic languages arise as a result of making a linguistic area with the imperfective converbs of the neighboring languages. In conclusion, the Tungusic converb in *-mi* is revealed to display remarkable functional distinctions between Russian and Chinese Tungusic in the following respects: 1) morpho-syntactic parameters: number marking, auxiliary construction, quotative index and 2) semantic parameters: conditionality. Given that most of the differences above between the converbs in *-mi* in Russian and Chinese Tungusic (i.e. auxiliary construction, quotative index, conditionality) correspond with differences between imperfective converbs in Russian and Mongolian (in *-ja* and *-j* respectively), the paper proposes that the functional differences among the *-mi* forms of different Tungusic languages occur because *-mi* forms a linguistic area with functionally corresponding imperfective converbs of Russian and Mongolian languages.

*Keywords:* converbal suffix *-mi*, Tungusic, Russian, Mongolian, linguistic area

## Introduction

This study aims not only to specify the functional differences of the converbal suffix *-mi* according to the geographic distribution of Tungusic languages (Russian and Chinese territories),<sup>2</sup> but also to raise the possibility that most functional distinctions of *-mi* among Tungusic languages arise because *-mi* creates a linguistic area with the functionally corresponding imperfective gerund *-ja* and imperfective converb *-j* in the neighboring Russian and Mongolian languages.

The Tungusic converbal suffix *-mi* presents common forms and similar semantic functions in all Tungusic languages. Therefore, the Tungusic *-mi* is generally considered a converbal form that stems from the Proto-Tungusic. Table 1 below summarizes existing descriptions of *-mi* in each Tungusic language. Firstly, the *-mi* form in the first group of Tungusic except Solon is mainly described as a conditional converb. However, this does not apply to Solon, which is spoken on the Chinese territory. As for the second and third group of Tungusic, the *-mi* in these languages is mostly defined as a simultaneous converbal form that connects two coincident events. Thirdly, the *-me* in Manchu (including Sibe), which is also distributed inside the Chinese border, is regarded as an imperfective or coordinative converbal form; in these characteristics, it bears similarities to Solon.

<sup>1</sup> This paper is a revised and extended English version of Baek (2014). Additionally, transcriptions, glosses, and English translations in this paper are my own and can be different from the original references. Any errors, of course, are solely my responsibilities.

<sup>2</sup> Following Ikegami (1974) and Kazama (1996), the Tungusic languages in this study are classified as follows: I: Evenki (Ek), Even (E), Negidal (N), Solon (S), II: Udihe (U), Orochi (Oc), Hezhen (Hz), III: Nanay (Nn), Olcha (Ol), Uilta (Ut), IV: Manchu (M).

In sum, judging from the previous studies on the converbal suffix *-mi*, we can confirm that descriptions of *-mi* vary according to the Tungusic classification and distribution.

Table 1. Description of converbal suffix *-mi* in previous literature

I	<b>Ek</b>	temporal-conditional converb (Konstantinova 1968: <i>-mī</i> , Nedjalkov, I. V. 1995: <i>-mi</i> ), conditional converb (Bulatova & Grenoble 1999: <i>-mī</i> )
	<b>E</b>	temporal converb? (Benzing 1955: <i>-mi</i> ), temporal-conditional converb (Novikova 1968: <i>-mi</i> ), conditional converb (Malchukov 1995: <i>-mi</i> , Kazama 2003 / Kim 2011: <i>-mi</i> )
	<b>N</b>	temporal-conditional converb (Kolesnikova & Konstantinova 1968: <i>-mi</i> ), conditional converb (Cincius 1982: <i>-mī</i> , Kazama 2002: <i>-mi</i> )
	<b>S</b>	imperfective converb (Poppe 1931: <i>-mi</i> ), coordinative converb (Chaoke et al. 1991: <i>-m</i> ), coordinative converb (Tsumagari 2009a: <i>-mi</i> )
II	<b>U</b>	present-tense converb (Shnejder 1936 / Girfanova 2002: <i>-mi</i> ), infinitive (Nikolaeva & Tolskaya 2001: <i>-mi</i> ), simultaneous converb (Girfanova 2002 / Kazama 2010a: <i>-mi</i> )
	<b>Oc</b>	simultaneous converb (Avrorin & Boldyrev 2001: <i>-mi</i> )
	<b>H<sub>z</sub></b>	coordinative converb (An 1986 / Li 2006: <i>-mi</i> ), simultaneous converb (Tamura 2008: <i>-mi</i> )
III	<b>Nn</b>	simultaneous converb (Avrorin 1961: <i>-mi</i> , Kazama 2010b: <i>-mi</i> )
	<b>O<sub>l</sub></b>	simultaneous converb (Sunik 1985: <i>-mi</i> , Kazama 2010c: <i>-mi</i> )
	<b>Ut</b>	simultaneous converb (Petrova 1967: <i>-mi</i> ), temporal-conditional converb ? [Ikegami 2001 (1959): <i>-mi</i> ], coordinative converb (Tsumagari 2009b: <i>-mi</i> )
IV	<b>M</b>	present-tense converb (Zakharov 1879: <i>-me</i> ), imperfective converb (Sung 1968 / Li 2000 / Gorelova 2002: <i>-me</i> ), simultaneous converb (Avrorin 2000: <i>-me</i> ), infinitive (Tsumagari 2002: <i>-me</i> )
	<b>S<sub>b</sub></b>	coordinative converb (Li & Chong 1986: <i>-me</i> ), imperfective converb (Norman 1974: <i>-me</i> )

### Previous Studies

There are two previous studies (Benzing 1956, Kazama 2003, 2010a, 2010b) on the functional differences of the converbal suffix *-mi* among the Tungusic languages. Benzing (1956) points out that the semantic function of *-me* in Manchu deviates from the ordinary semantic functions of *-mi* in the other Tungusic languages, providing an example of motion purposive interpretation.

- 1) *ara-me dosi-ka.*  
write-CVB enter-PTCP.PST  
“He entered to write.”

(Benzing 1956: 143)

In addition, Kazama (2003, 2010a, 2010b) states that the converbal suffix *-mi* slightly differs in functions among the Tungusic languages, observing that the *-mi* in the first group of Tungusic does not have alternative forms by number, unlike the third group of Tungusic, and generally conveys conditional meaning. However, both studies briefly mention the functional differences of *-mi* without elaborating on the cause of them.

### Approach

Referring to the previous accounts and textual materials (see Appendix 1) of Tungusic, this study applies the typological parameters of converbs listed in Figure 1 to determine the functional differences of *-mi* by areal distribution. These parameters are mostly based on Nedjalkov, V.P. (1995). However, some of the parameters, such as number-marking, auxiliary construction, and quotative index, are my own.

Moreover, in order to specify the functional distinctions of *-mi* by distribution (Russian vs. Chinese territories), this paper adopts Tsumagari's (1997) division by language territory and divides Tungusic into two groups: Russian Tungusic and Chinese Tungusic. In this study, we examine Evenki (I), Nanay (III) and Udihe (II) from Russian Tungusic, while Solon (I), Hezhen (II), Manchu (IV) are chosen from Chinese Tungusic (the Roman numerals in parentheses represent Tungusic classification in Table 1).

Figure 1 Typological parameters of converbs (based on Nedjalkov, V.P. 1995)

1. Morpho-syntactic parameters	2. Semantic parameters
a. person marking	a. simultaneity
b. number marking	b. anteriority
c. corefentiality	c. causality
d. auxiliary construction	d. purpose
e. quotative index	e. conditionality

## Tungusic

### *Morpho-Syntactic Parameters*

#### Person marking

In all Tungusic converbs in *-mi*, the person marker is not attached to the converbal form.

#### Number marking

The converbal form in *-mi* in Russian Tungusic can be marked by the number of the subject in the subordinate clause, as shown in Examples 2 and 3. However, there is a difference in marking plurality in *-mi* converbs between Evenki on the one hand and Nanay, Udihe on the other. Evenki uses the nominal plural suffix *-l* to express plurality, while Nanay and Udihe have alternative forms of *-mi* in accordance with number. The use of different forms of *-mi* according to number is common in the second and third groups of Tungusic except Hezhen in the Chinese territories. Regarding Chinese Tungusic, only *-mi* is used regardless of subject number in the antecedent clause, as Example 4 indicates.

*Evenki*

- 2) *ju-la-wer eme-mi-l jep-čo-tin.*  
 house-DIR-REF.PL come-CVB-PL eat-PST-3PL  
 "Having come home, [they] ate."

(Nedjalkov, I. V. 1995: 445)

*Nanay*

- 3) *em modan=tanii xupi-meeri=tenii, enin-či-ji jiju-xe-či.*  
 one time=CLT play-CVB.PL=CLT mother-DIR-REF.SG return-PTCP.PST-3PL  
 "One time playing around, [They] returned to one's mother."

(Kazama 2010b: 142)

*Hezhen*

- 4) *tigurun xofur xofur sokutu-mi omi-xe-ti.*  
 they OMTP get.drunk-CVB drink-PTCP.PST-3PL  
 "[They] drank while getting drunk."

(Tamura 2008: 43)



## Coreferentiality

Nedjalkov, V. P. (1995) mentions that converbs can be classified into three types (same-subject converbs, different-subject converbs, and varying-subject converbs) according to whether the subjects in the subordinate and main clauses are coreferential or not. In principle, the converbal suffix *-mi* in both areas of Tungusic is used in the same-subject situation. This is considered to be closely related to the fact that person is not marked in *-mi* converbs.

## Auxiliary construction

Auxiliary construction in this paper is defined as follows: verb 2 that follows verb 1 marked by *-mi* loses its original lexical meaning and functions as a grammatical element (expressing tense-aspect-mood characteristics). In this study, we examine the following four auxiliary constructions:

- a) progressive:  $V_1\text{-}mi + V_2$  ‘be’
- b) attempt:  $V_1\text{-}mi + V_2$  ‘see’
- c) benefactive:  $V_1\text{-}mi + V_2$  ‘give’
- d) perfective:  $V_1\text{-}mi + V_2$  ‘put’

Auxiliary construction with *-mi* is confirmed to be relatively restricted in Russian Tungusic, whereas it is highly developed in Chinese Tungusic, as Table 2 and Examples 5-15 demonstrate. Kazama (2014) also states that there are a lot of auxiliary constructions in Solon, Manchu, and Sibe that had or still have contact with Mongolian and/or Chinese, whereas such constructions have not developed in most Tungusic.

Table 2. Auxiliary constructions with converbal suffix *-mi*<sup>3</sup>

	Russian			Chinese		
	Ek	Nn	U	S	Hz	M
progressive: $V_1\text{-}mi + V_2$ ‘be’	-	+	+	+	+	+
attempt: $V_1\text{-}mi + V_2$ ‘see’	-	-	-	+	+	+
benefactive: $V_1\text{-}mi + V_2$ ‘give’	-	-	-	-	+	+
perfective: $V_1\text{-}mi + V_2$ ‘put’	-	-	-	-	-	+
Total	0	1	1	2	3	4

*Nanay 1*: progressive ( $V_1\text{-}mi + V_2$  ‘be’)

- 5) *xai ta-mi bi-i-si?*  
 what do-CVB be-PTCP.PRS-2SG  
 “What are [you] doing?”

(Kazama 2008b: 89)

*Udihe 1*: progressive ( $V_1\text{-}mi + V_2$  ‘be’)

- 6) *uti etete-ne-mi, jeu diga-mi bi-i-ti.*  
 that work-DIRINT-CVB what eat-CVB be-PTCP.PRS-3PL  
 “Coming to work, what are [they] eating?”

(Kazama 2004: 383)

<sup>3</sup> In this paper, the Tungusic and neighboring languages will be compared in terms of their degree of auxiliary construction as follows. If a language forms all four auxiliary constructions described above, its degree of auxiliary construction is four; if it only forms one of those auxiliary constructions, its degree of auxiliary construction is one, and so on.



*Hezhen 1: progressive (V<sub>1</sub>-mi + V<sub>2</sub> ‘be’)*

- 7) *su arki-we omi-mi bi-yi-su?*  
 you.PL.NOM liquor-ACC drink-CVB be-PTCP.NPST-2PL  
 “Are you drinking liquor?”

(Li 2006: 58)

*Hezhen 2: attempt (V<sub>1</sub>-mi + V<sub>2</sub> ‘see’)*

- 8) *bi edin-me gele-m-ičie-ø-mi.*  
 I wind-ACC seek-CVB-see-PTCP.PRS-1SG  
 “I try to seek wind.”

(Tamura 2008: 46)

*Solon 1: progressive (V<sub>1</sub>-mi + V<sub>2</sub> ‘be’)*

- 9) *aasi-mi bi-see.*  
 sleep-CVB be-PST  
 “[He] was sleeping.”

(Tsumagari 2009a: 10)

*Solon 2: attempt (V<sub>1</sub>-mi + V<sub>2</sub> ‘see’)*

- 10) *jege-we-n teŋteri-m iččee gunen.*  
 chin-ACC-3 stroke-CVB see.PTCP.PST say  
 “[They] say that [he] tried to stroke his chin.”

(Kazama 2008c: 49)

*Solon 3: benefactive (V<sub>1</sub>-mi + V<sub>2</sub> ‘give’)*

- 11) *aggu sibbaa-m buu-se gunen.*  
 wisdom give-CVB give-PTCP.PST say  
 “[They] say that [she] gave wisdom [for me].”

(Kazama 2008c: 21)

*Manchu 1: progressive (V<sub>1</sub>-me + V<sub>2</sub> ‘be’)*

- 12) *hūwang joo ba be tuwakiya-me bi-he se-me ala-mbi.*  
 Hwangcaoba ACC watch-CVB be-PTCP.PST say-CVB say-IMP  
 “[They] say that [they] were watching over Hwangcaoba.”

(Choi et al. 2012b: 111)

*Manchu 2: attempt (V<sub>1</sub>-me + V<sub>2</sub> ‘see’)*

- 13) *siyanseng leole-me tuwa-ø.*  
 teacher think-CVB see-IMP  
 “Teacher, try to think.”

(Sung 1968: 78)

*Manchu 3: benefactive (V<sub>1</sub>-me + V<sub>2</sub> ‘give’)*

- 14) *hahai jaka be haha de faksala-me bu-fi unggi-ø.*  
 man.GEN thing ACC man DAT divide-CVB give-ANT.CVB send-IMP  
 “Give this man his share and see him off.”

(Sung 1968: 79)

*Manchu 4: perfective (V<sub>1</sub>-me + verb ‘put’)*

- 15) *gemu saikan boo-de dosim-bu-me sinda-ø.*  
 all well house-DAT enter-CAUS-CVB put-IMP  
 “Let them all enter the house.”

(Choi et al. 2012a: 403)

## Quotative index

A speech verb in the converbal form cross-linguistically functions as a quotative or complement index. In Evenki, however, we cannot verify that a speech verb with *-mi* grammatically serves as a quotative index. According to Avrorin (1961), Nanay uses the clitic  $=(A)m$  as a quotative marker, as shown in Example 16. He speculates that this clitic originated from the speech verb *un-* ‘say’ plus *-mi*. Nevertheless, we do not consider it as a speech verb in *-mi* since the clitic no longer retains its original formation. As for Udihe, Nikolaeva and Tolskaya (2001) say that the particle *gumu* ( $< gune-$  +  $-u$ : say+PAS?) is used to mark a quotative or complement clause (see Example 17). On the other hand, a speech verb in the converbal form *-mi* functions as a quotative or complement index in most of Chinese Tungusic. Hezhen and Manchu use speech verbs with converbal suffixes *-mi* and *-me* respectively, for example, *ne-mi*: say-CVB (Hz), *se-me*: say-CVB (M), to mark quotation or complement, as shown in Examples 19-20. However, Solon does not utilize a speech verb with *-mi*, instead the particle *gunken* ( $<$  speech verb *gun-* ‘say’ plus fossilized simultaneous converbal suffix *-nAkAn*) is frequently used to introduce a quotation or complement clause (see Example 18). These facts are summarized in Table 3.

Table 3. Speech verbs plus converbal suffix *-mi* as quotative index

	Ek	Russian	U	S	Chinese	
		Nn			Hz	M
quotative index	-	$=(A)m$ (clitic)	<i>gumu</i> (particle)	<i>gunken</i> (particle)	<i>ne-mi</i> (say-CVB)	<i>se-me</i> (say-CVB)
result	-	-	-	-	+	+

### Nanay

- 16) *ča-do nee-mi aja bi-ǰeree=m un-dii-ni.*  
 that-DAT put-CVB good be-FUT=CLT say-PTCP.PRS-3SG  
 “[He] says: It will be good to put [it] there.”

(Kazama 2010b: 259)

### Udihe

- 17) *nuani diaŋ-ka eniŋe-tigi-i magajina-tigi ŋene-ǰeŋe-i gumu.*  
 he say-PST mother-DIR-REF.SG shop-DIR go-PTCP.FUT-1SG QUOT  
 “He said to his mother: I will go to shop.”

(Nikolaeva & Tolskaya 2001: 668)

### Solon

- 18) *sii ilee nenendi gunken minii exinbel aŋʊʊ-saa.*  
 you where go.PRS.2SG QUOT I.GEN older sister ask-PST  
 “My sister asked: Where do you go?”

(Kazama 2011b: 164)

### Hezhen

- 19) *esi eme nio=de ene-ø-i ne-mi xesu-rše-n.*  
 now one person=CLT go-PTCP.PRS-1SG say-CVB say-NEG.PRS-3  
 “Now even one person does not say: I go.”

(Tamura 2008: 51)

### Manchu

- 20) *doose geren i baru suwe omi-me ele-he-o se-me*  
 monk everybody GEN toward you.PL.NOM drink-CVB enough-PTCP.PST-Q say-CVB  
*fonǰi-ha de.*  
 ask-PTCP.PST DAT  
 “When the monk asked everybody: Have you had enough to drink?”

(Gorelova 2002: 274)

### Semantic Parameters

Nedjalkov, V. P. (1995) typologically classifies converbs into three types by semantic functions: specialized converbs, contextual converbs, and narrative converbs. Following this classification, *-mi* in Tungusic belongs to contextual converbs that can perform various semantic functions according to the context. As Table 4 illustrates, the converbal suffix *-mi* in each language is confirmed to perform multi-semantic functions such as simultaneity, anteriority, and causality. Contrary to Benzing (1956), the purposive construction, where the *-mi* form is followed by a motion verb, is also possible in most of Tungusic. However, there is a prominent difference between *-mi* in the conditional function between Russian and Chinese Tungusic.

Table 4. Semantic functions of converbs in *-mi* in Russian and Chinese Tungusic

	Russian			Chinese		
	Ek	Nn	U	Hz	S	M
simultaneity	+	+	+	+	+	+
anteriority	+	+	+	+	+	+
causality	+	+	+	+	+	+
purposive	-	+	+	+	+	+
conditionality	+	+	+	-	-	-

### Conditionality

In this study, we divided conditionals into two types as follows: real and counterfactual conditionals. Nedjalkov, I. V. (1995) observes that the converbal form in *-mi* in Evenki leads to a conditional reading when future tense, imperative, and subjunctive moods occur in the main clause. Examples 21-22 below indicate that the *-mi* in Evenki can create both real and counterfactual conditionals. The same phenomena are also confirmed in the *-mi* of Udihe, as presented in Examples 23-24, which are obtained from my fieldwork data. Nevertheless, note that the counterfactual conditional in Udihe is generally formed by the conditional marker *bisi*, as mentioned in Nikolaeva & Tolskaya (2001). As for Nanay, Kazama (2010b) states that *-mi* can create conditional constructions, as illustrated in Example 25. We can see that future tense in the main clause creates conditional interpretation. However, we cannot find an example of a counterfactual conditional with *-mi* in Nanay texts. The reason for this is presumed to be the fact that Nanay mainly uses the conditional marker *osini* to form counterfactual conditionals. As a rule, the converbal suffix *-mi* in Russian Tungusic forms same-subject conditionals. On the other hand, the *-mi* in Chinese Tungusic does not semantically function as a conditional in future, imperative, and subjunctive sentences. Instead, conditional converbs (Solon *-kki*, Hezhen *-ki* and Manchu *-či*), both with the same or different subject, are used to form real and counterfactual conditionals in Chinese Tungusic (see Examples 26-27).

#### Evenki

Real conditional (future tense, imperative mood in the main clause)

- 21) *aja-t hawa-l-mii-l, beje-l oo-jaŋaa-sun.*  
 good-INS work-INC-CVB-PL person-PL become-PTCP.FUT-2PL  
 “If [you] would start working well, [you] will become a person.”

(Bulatova & Grenoble 1999: 44)

Counterfactual conditional (subjunctive mood in the main clause)

- 22) *asatkan-me ajaw-mi, asila-mča-w.*  
 girl-ACC love-CVB marry-SUBJ-1SG  
 “If [I] loved this girl, [I] would marry her.”

(Nedjalkov, I. V. 1997: 54)

*Udihe*

Real conditional (future tense, imperative mood in the main clause)

- 23) *nuani maŋga bi-mi, uti sagdi ʃolo-wo uinde-mi mute-ʃeŋe-ni.*  
 he strong be-CVB that big stone-ACC lift up-CVB can-PTCP.FUT-3SG  
 “If he is strong, [he] will lift up that big stone.”

(Fieldwork data)

Counterfactual conditional (subjunctive mood in the main clause)

- 24) *nuani udie kewe-ni saa-mi, udie-ʃige-ʃi diana-mi mute-muse.*  
 he Udihe language-3SG know-CVB Udihe-PL-COM speak-CVB can-SUBJ.3  
 “If he knew the Udihe language, [he] could speak with Udihe people”

(Fieldwork data)

*Nanay*

Real conditional (future tense in the main clause)

- 25) *bumbie mana-mi, sumbie waa-nda-ʃaraa.*  
 we.ACC finish-CVB you.ACC kill-DIRINT-FUT.3  
 “If [they] finish us, [they] will kill you.”

(Kazama 2010b: 243)

*Manchu*Real conditional: *-či* (conditional converb)

- 26) *si gai-ki se-či, uthai gaisu.*  
 you.SG.NOM take-OPT say-COND.CVB then take.IMP  
 “If you want to take [something], then take [it].”

(Gorelova 2002: 297)

Counterfactual conditional: *-či* (conditional converb)

- 27) *aika duleke aniya adali elgiyan bargiya-ha bi-či, gemu ʃeterengge*  
 if last year like plentiful harvest-PTCP.PST be-COND.CVB everyone food  
*bu-ʃi ulebu-mbihe.*  
 give-ANT.CVB feed-SUBJ  
 “If [we] harvested plentiful like last year, [we] would give food to everyone and feed them.”

(Tsumagari 2002: 61)

## Summary

We have applied the typological parameters of converbs proposed by Nedjalkov, V.P. (1995) to the Tungusic *-mi* and summarized the results in Table 5 below.

Table 5. Tungusic *-mi* in Russian and Chinese territories by typological parameters of the converbs

		Russian			Chinese		
		Ek	Nn	U	S	Hz	M
morpho-syntactic	person marking	-	-	-	-	-	-
	number marking	+	+	+	-	-	-
	coreferentiality	+	+	+	+	+	+
	auxiliary construction	-	1	1	3	2	4
	quotative index	-	-	-	-	+	+
semantic	simultaneity	+	+	+	+	+	+
	anteriority	+	+	+	+	+	+
	causality	+	+	+	+	+	+
	purpose	-	+	+	+	+	+
	conditionality	+	+	+	-	-	-

The converb *-mi* in Russian and Chinese Tungusic displays noticeable distinctions in number marking, auxiliary construction, and quotative index functions. As for semantic functions, the conditional is considered a major difference between the two groups. Following these results, we can verify that the areal factor is strongly associated with the functional differences of the converbal suffix *-mi* in Tungusic. However, some differences, such as differences in ways of number marking, degree of auxiliary construction, formation of quotative index and conditional types, are also confirmed among Tungusic within the same area.

### *Neighboring languages (Russian, Mongolian)*

Russian and Mongolian are known to have influenced the Tungusic languages. The Russian imperfective gerund suffix *-ja* and the Mongolian imperfective converbal suffix *-j* functionally correspond to the *-mi* in Tungusic. We apply the same typological parameters of converbs to the imperfective converbal forms in the Russian and Mongolian languages.

### *Russian Imperfective Gerund in -ja*

The Russian imperfective gerund in *-ja*, used only in same-subject circumstances, is not marked for person and number. In Russian, we cannot confirm the quotative index and auxiliary construction with *-ja*. Concerning its semantic function, the imperfective gerund in *-ja* belongs to contextual converbs that serve various semantic functions such as simultaneity, cause, and conditional. Conditional readings, both real and counterfactual, are formed with *-ja* when the future tense and subjunctive mood appear in the main clause, as illustrated in Example 28 for the real condition (future tense < perfective verb in the main clause) and in Example 29 for the counterfactual condition (with subjunctive mood in the main clause).

- 28) *Raz'ezža-ja* po strane, on navedet spravki o syne.  
 travel.IMPF-CVB around country he will make inquiries about son  
 "If [he] travels around the country, he will make inquiries about his son."

(Boguslavskij 1977: 271)

- 29) *Zna-ja eti slova, vy mogli by vce perevecti.*  
 know-CVB these words you could SUBJ all translate  
 “If [you] knew these words, you would translate all.”

(Shirota 2010: 393)

*Mongolian Imperfective Converb in -j*

Mongolian converbs formed with the suffix *-j* are not marked for person and number. The converbal form in *-j* is most commonly used when the subjects of the main and subordinate clauses are the same. Mongolian can be classified as a language with a high degree of auxiliary construction since *-j* can form three out of four auxiliary constructions in this paper (see Examples 30-32). The speech verb *ge-* ‘say’ plus *-j* in Mongolian functions as a quotative or complement index, as Example 33 illustrates.

*Auxiliary construction*progressive:  $V_1-j + V_2$  ‘be’

- 30) *čaanaas mor'toj xün ir-j baj-na.*  
 beyond.ABL horse.COM person come-CVB be-NPST  
 “A person on a horse is coming from that side.”

(Yamakoshi 2012: 122)

attempt:  $V_1-j + V_2$  ‘see’

- 31) *tamxi tata-j üz-ex üüi?*  
 cigarette pull-CVB see-PTCP.FUT Q  
 “Do [you] try to smoke?”

(Yamakoshi 2012: 124)

benefactive:  $V_1-j + V_2$  ‘give’

- 32) *Bat ene blog-ijg nadad zaa-j ög-sön.*  
 Bat this blog-ACC I.DAT show-CVB give-PTCP.PST  
 “Bat showed this blog to me.”

(Yamakoshi 2012: 124)

*Quotative index*

- 33) *margaas arvan čag-t uulz'ja ge-j jari-lč-san.*  
 tomorrow ten hour-DAT meet.VOL say-CVB speak-RECP-PTCP.PST  
 “[We] said: Let’s meet at ten tomorrow.”

(Yamakoshi 2012: 128)

Concerning its semantic function, *-j* can convey simultaneity, anteriority, causality, whereas purpose and conditional sentences are not allowed. As with Chinese Tungusic, future tense, imperative and subjunctive moods in the main clause cannot lead to conditional readings. As Examples 34-35 demonstrate, real and counterfactual conditionals are formed by the conditional converbal suffix *-bal* and conditional particle *bol* (<\**bol-bol*: become-COND.CVB) respectively, both with same or different subject.

Real conditional: *-bal* (conditional converb)

- 34) *ene dugujg av-bal, bi avtobus-aar javaxguj.*  
 this bike.ACC buy-COND.CVB I bus-INS go.PTCP.FUT.NEG  
 “If [I] buy this bike, I won’t go by bus.”

(Kullman &amp; Tserenpil 1996: 162)

Counterfactual conditional: *bol* (conditional particle)

35) *ert irsen(sen) bol, bagš-taj uulzax bajžee.*  
 early come.PTCP.PST if teacher-COM meet.PTCP.FUT be.PST  
 “If [I] had come earlier, [I] would have met the teacher.”

(Kullman & Tserenpil 1996: 343)

## Summary

Results of the application of typological parameters of converbs to imperfective converbs in Russian and Mongolian are summarized in Table 6. In short, we can verify that the auxiliary construction and quotative index parameters are different between Russian and Mongolian imperfective converbs. Regarding semantic functions, anteriority and conditionality are major distinctions between the two languages.

Table 6. Russian and Mongolian imperfective converbs by typological parameters of converbs

		<b>R</b>	<b>Mo</b>
morpho-syntactic	person marking	-	-
	number marking	-	-
	coreferentiality	+	+
	auxiliary construction	<b>0</b>	<b>3</b>
	quotative index	-	+
semantic	simultaneity	+	+
	anteriority	-	+
	causality	+	+
	purpose	-	-
	conditionality	+	-

## Conclusion

We confirmed that the converbal suffix *-mi* indicates remarkable distinctions between Russian Tungusic (Evenki, Nanay, Udihe) and Chinese Tungusic (Solon, Hezhen, Manchu) in the following parameters: (a) morpho-syntactic: number marking, auxiliary construction, quotative index, (b) semantic: conditionality. As Table 7 indicates, most of these differences between Russian and Chinese Tungusic correspond with distinctions of the imperfective converbs (in *-ja* and *-j* respectively) in Russian and Mongolian. Thus, we propose that the functional differences that *-mi* shows in different Tungusic languages occur because *-mi* forms a linguistic area with the functionally corresponding converbs of the neighboring languages. However, we can also observe some differences in the way *-mi* functions in Tungusic within same area, such as differences in the ways of number marking, degrees of auxiliary construction, formation of quotative markers, conditional types. Further research is required to explain this variation.



Table 7. The Tungusic *-mi* and Russian / Mongolian imperfective converbs by typological parameters of converbs

		Russian				Chinese			
		R	Ek	U	Nn	S	Hz	M	Mo
morpho-syntactic	person marking	-	-	-	-	-	-	-	-
	number marking	-	+	+	+	-	-	-	-
	coreferentiality	+	+	+	+	+	+	+	+
	auxiliary construction	-	-	1	1	3	2	4	3
	quotative index	-	-	-	-	-	+	+	+
semantic	simultaneity	+	+	+	+	+	+	+	+
	anteriority	-	+	+	+	+	+	+	+
	causality	+	+	+	+	+	+	+	+
	purpose	-	+	+	+	+	+	+	-
	conditionality	+	+	+	+	-	-	-	-

### Abbreviations

1, 2, 3: first person, second person, third person  
 ABL: ablative  
 ACC: accusative  
 ANT: anterior  
 CAUS: causative  
 CLT: clitic  
 COM: comitative  
 COND: conditional  
 CVB: converb  
 DAT: dative  
 DIR: directive  
 DIRINT: directional intentional  
 FUT: future

GEN: genitive  
 IMP: imperative  
 IMPF: imperfective  
 INC: inchoative  
 INS: instrumental  
 NEG: negative  
 NOM: nominative  
 NPST: non-past  
 OMTP: onomatopoeia  
 OPT: optative  
 PAS: passive  
 PL: plural  
 PRS: present

PST: past  
 PTCP: participle  
 Q: question marker  
 QUOT: quotative marker  
 RECP: reciprocal  
 REF: reflexive  
 SG: singular  
 SIM: simultaneous  
 SUBJ: subjunctive  
 V: verb  
 VOL: voluntative  
 =: clitic boundary

### References

- An, J (1986) *Hezheyu Jianzhi* [=A Sketch of Hezhe Language]. Beijing: Minzu chubanshe.  
 Avrorin, V. A. (1961) *Grammatika Nanajnskogo Jazyka*, volume 2. Leningrad: AN. SSSR.  
 Avrorin, V.A. (2000) *Grammatika Man'chzhurskogo Pis'mennogo Jazyka*. Sankt-Peterburg: Nauka.  
 Avrorin, V.A. & Boldyrev, B.V. (2001) *Grammatika Orochskogo Jazyka*. Novosibirsk: izdatel'stvo. SO RAN.  
 Baek, S. (2014) Tsungus shogoni okeru fukudoushi *-mi* no souini tsuite [= Functional differences of converb *-mi* in Tungusic]. Network for Northern Language Studies (ed.). *Northern language studies* 4: 85-110. Sapporo: Graduate School of Letters, Hokkaido University.  
 Benzing, J. (1955) *Lamutische Grammatik*. Wiesbaden: Franz Steiner.  
 Benzing, J. (1956) *Die Tungusischen Sprachen: versuch einer vergleichenden grammatik*. Mainz: Akademie der Wissenschaften und der Literatur.  
 Boguslavskij, I. M. (1977) O semantičeskom opisanii russkix deepričactij: neopredelennost' ili mnogoznačnost'?. *Serija literatury i jazyka* 36.3: 270-281. Izvestija AN SSSR.  
 Bulatova, N. & Grenoble, L. (1999) *Evenki*. München / Newcastle: Lincom Europa.

- Chaoke & Tsumagai, T. & Kazama, S. (1991) *Solongo kihon reibunshuu* [=Solon basic sentences]. Sapporo: Faculty of Letters, Hokkaido University.
- Choi, D. et al. (2012a) *Cengeonogeoldaesinseok* [=Nogeoldae in Manchu]. Seoul: Pakmunsa.
- Choi, D. et al. (2012b) *Beye i cooha bade yabuha babe ejehe bithe* [=the book in which I wrote what I remembered in the war]. Seoul: Pakmunsa.
- Cincius, V. I. (1982) *Negidal'skij Jazyk: issledovanija i materialy*. Leningrad: Nauka.
- Girfanova, A. H. (2002) Udeghe München / Newcastle : Lincom Europa.
- Gorelova, L. M. (2002) *Manchu Grammar*. Handbook of oriental studies, Section 8. Central Asia vol. 7. Leiden / Boston / Köln: Brill.
- Ikegami, J. (1974) Versuch einer Klassifikation der tungusischen Sprache. *Sprache, Geschichte und Kultur der Altaischen Völker, Protokollband der XII. Tagung der Permanent international Altaische Conference 1969 in Berlin: 271-272*, Berlin: Akademie Verlag.
- Ikegami, J. (2001) The verb inflection of Orok. In: *Tsunguusugo Kenkyuu*: 24-72, Tokyo: Kyuukoshoin (First appeared in: 1959 The verb inflection of Orok. Kokugo Kenkyu No. 9: 34-73, Kokugakuin Daigaku Kokugokenkyukai, Tokyo).
- Kazama, S. (1996) Genetical position of Hezhen. *Gengo Kenkyu* 109: 117-139.
- Kazama, S. (2002) *Negidaarugo Tekisuto to Bunpou Gaisetsu* [=Negidal Texts and Grammar] (ELPR publication series: A2-021. Publications on Tungus languages and cultures 19). Suita: Faculty of Informatics, Osaka Gakuin University.
- Kazama, S. (2003) *Ewengo Tekisuto to Bunpou Gaisetsu* [=Even Texts and Grammar] (ELPR publication series: A2-030. Publications on Tungus languages and cultures 23). Suita: Faculty of Informatics, Osaka Gakuin University.
- Kazama, S. (2004) *Udehego Tekisuto (A)* [=Udihe Texts (A)] (Publications on Tungus languages and cultures 24/A). Tokyo: Research Institute for Languages and Cultures of Asia and Africa Tokyo University of Foreign Studies.
- Kazama, S. (2005) *Naanai no Minwa to Densetsu 8* [=Nanay Folk Tales and Legends 8] (Publications on Tungus languages and cultures 27). Chiba University.
- Kazama, S. (2006a) *Udehego Tekisuto 2* [=Udihe Texts 2] (Publications on Tungus languages and cultures 31). Tokyo: Research Institute for Languages and Cultures of Asia and Africa Tokyo University of Foreign Studies.
- Kazama, S. (2006b) *Naanai no Minwa to Densetsu 9* [=Nanay Folk Tales and Legends 9] (Publications on Tungus languages and cultures 32). Chiba University.
- Kazama, S. (2007a) *Udehego Tekisuto 3* [=Udihe Texts 3] (Publications on Tungus languages and cultures 35). Sapporo: Graduate School of Letters Hokkaido University.
- Kazama, S. (2007b) *Naanai no Minwa to Densetsu 10* [=Nanay Folk Tales and Legends 10] (Publications on Tungus languages and cultures 36) Sapporo: Graduate School of Letters Hokkaido University.
- Kazama, S. (2007c) *Sorongono no Minwa to Densetsu 1* [=Solon Folk Tales and Legends 1] (Publications on Tungus languages and cultures 37). Sapporo: Graduate School of Letters Hokkaido University.
- Kazama, S. (2008a) *Udehego Tekisuto 4* [=Udihe Texts 4] (Publications on Tungus languages and cultures 42). Tokyo: Research Institute for Languages and Cultures of Asia and Africa Tokyo University of Foreign Studies.
- Kazama, S. (2008b) *Naanai no Minwa to Densetsu 11* [=Nanay Folk Tales and Legends 11] (Publications on Tungus languages and cultures 40). Sapporo: Graduate School of Letters Hokkaido University.
- Kazama, S. (2008c) *Sorongono no Minwa to Densetsu 2* [=Solon Folk Tales and Legends 2] (Publications on Tungus languages and cultures 41). Sapporo: Graduate School of Letters Hokkaido University.
- Kazama, S. (2009) *Udehego Tekisuto 5* [=Udihe Texts 5] (Publications on Tungus languages and cultures 44). Tokyo: Research Institute for Languages and Cultures of Asia and Africa Tokyo University of Foreign Studies.
- Kazama, S. (2010a) *Udehego Ttekisuto 6* [=Udihe Texts 6] (Publications on Tungus languages and cultures 47). Tokyo: Research Institute for Languages and Cultures of Asia and Africa Tokyo University of Foreign Studies.

- Kazama, S. (2010b) *Naanai no Minwa to Densetsu* 12 [=Nanay Folk Tales and Legends 12] (Publications on Tungus languages and cultures 48). Tokyo: Research Institute for Languages and Cultures of Asia and Africa Tokyo University of Foreign Studies.
- Kazama, S. (2010c) *Urucha Koushou Bungei Genbunshuu* 5 [=Ulcha Oral Literature 5: A Collection of Texts] (Publications on Tungus languages and cultures 49). Tokyo: Research Institute for Languages and Cultures of Asia and Africa Tokyo University of Foreign Studies.
- Kazama, S. (2011a) Naanaigo no fukubunni tsuite: Jyouken keisikino tsukaiwakewo chuusinto site [=The complex sentence in Nanai: with a special focus on conditionals]. *Northern Language Studies* 1: 115-138. Network for Northern Language Studies ed. Sapporo: Graduate School of Letters Hokkaido University.
- Kazama, S. (2011b) Sorongoni okeru Mongorugono eikyou [=Mongolian influence on Solon]. *Gogaku Kyouiku Fooramu Vol 24*: 163-183. Teramura, M & T, Fukumori (eds.) Tokyo: Daito Bunka University.
- Kazama, S. (2014) Nihongono ruikeini tsuite: arutaigatagengono kaimeiwo mezashite [On the linguistic type of Japanese: toward the illumination of the "Altaic-type" languages]. Network for Northern Language Studies (ed.). *Northern language studies* 4. 157-171. Sapporo: Graduate School of Letters, Hokkaido University.
- Kim, J. (2011) *A Grammar of Even*. Seoul: Seoul National University Press.
- Kolesnikova, V. D. & Konstantinova, O. A (1968) Negidal'skij jazyk. *Mongol'skie, Tunguso-Man'chzhurskie i Jazyki*: 109-128. Skorik, P. R. (ed.) Leningrad: Nauka.
- Konstantinova, O. A. (1968) Evenkijskij jazyk. *Mongol'skie, Tunguso-Man'chzhurskie i Paleoaziatskie Jazyki*: 68-87. Skorik, P. R. (ed.) Leningrad: Nauka.
- Kormushin, I. V. (1998) *Udyxejskij Jazyk*. Moskva: Nauka.
- Kullmann, R. & Tserenpil, D. (1996) *Mongolian Grammar*. Institute of Language and Literature Academy of Sciences Ulaanbaatar Mongolia.
- Li, G. R. (2000) *Manchu A Textbook for Reading Documents*. Honolulu: University of Hawai'i Press.
- Li, L. (2006) *Hozengo no Doushi Kouzou* [=Verbal structure of Hezhen]. PhD. Dissertation (unpublished). Chiba University.
- Li, L. (2011) Hozengo no kaiwa tekisuto (1) [=Hezhen conversational text 1]. *Northern Language Studies* 1:197-216. Network for Northern Language Studies ed. Sapporo: Graduate School of Letters Hokkaido University.
- Li, L. (2012) Hozengo intabyuu tekisuto (1) [=Hezhen interview text 1]. *Northern Language Studies* 3: 183-216. Network for Northern Language Studies ed. Sapporo: Graduate School of Letters Hokkaido University.
- Li, S. & Chong, Q. (1986) *Xiboyu Jianzhi* [=A Short Description of Sibe Language]. Beijing: Minzuchubanshe.
- Malchukov, A. L. (1995) *Even*. München / Newcastle: Lincom Europa.
- Nedjalkov, I. V. (1995) Converbs in Evenki. *Converbs in Cross-Linguistic Perspective*: 441-463. Haspelmath, M. & Konig, E. (Eds.). Berlin / New York. Mouton de Gruyter.
- Nedjalkov, I. V. (1997) *Evenki*. London: Routledge.
- Nedjalkov, V. P. (1995) Some typological parameters of converbs. *Converbs in Cross-Linguistic Perspective*: 97-136. Haspelmath, M. & Konig, E. (Eds.). Berlin / New York: Mouton de Gruyter.
- Nikolaeva, I. & Tolskaya, M. (2001) *A Grammar of Udihe*. Berlin / New York: Mouton de Gruyter.
- Norman, J. (1974) A sketch of Sibe morphology. *Central Asiatic Journal* 18: 159-174.
- Novikova, K. A. (1968) Evenskij jazyk. *Mongol'skie, Tunguso-Man'chzhurskie i Paleoaziatskie Jazyki*: 88-108. Skorik, P. R. (ed.) Leningrad: Nauka.
- Petrova, T. I. (1967) *Jazyk Orokov* (Ul'ta). Leningrad.
- Poppe, N. (1931) *Materialy po Solonskomu Jazyku*. Leningrad: Izd-vo Akademii Nauk SSSR.
- Shirota, J. (2010) *Gendai Roshigo Bunpou* [=Modern Russian Grammar]. Tokyo: Touyoushoten.
- Shnejder, E. P. (1936) *Kratkij Udegejsko-Russkij Slovar'*. Moskva / Leningrad: Gosudarstvennoe uchebno-pedagogicheskoe izdatel'stvo.
- Sung, B. (1968) Manzueo muneoeu pudongsa eomi -mee daehayeo [=On Manchu converb ending *-me*]. *Hangeul* No.142: 72-102. Korean Language Research Society.
- Sunik, O. P. (1985) *Ul'cskij Jazyk*. Leningrad: Nauka.

- Tamura, K. (2008) Grammatical suffixes of Hezhen: with the text ‘Antu Mergen’ and its English and Japanese translations. *Bulletin of Aichi University of Education Humanities and Social Sciences* 57: 37-54. Aichi: Bulletin of Aichi University of Education.
- Tsumagari, T. (1981) Manshuugo no doushigobi-me [= Converbial ending -me in Manchu]. *Hoppo Bunka Kenkyuu* 14: 149-172 Sapporo: Hokkaido University.
- Tsumagari, T. (1997) Linguistic diversity and national borders of Tungusic. Shoji, H. & Janhunen, J. eds. *Northern minority languages: problems of survival (Senri Ethnological Studies No.44)*: 175-186. Suita: National Museum of Ethnology
- Tsumagari, T. (2002) *Manshuugo Nyuumon 20 kou* [=Manchu Introduction 20 Chapter]. Tokyo: Daigakushorin.
- Tsumagari, T. (2009a) A sketch of Solon grammar. *Journal of the center for Northern Humanities*, 2: 1-21 Center for Northern Humanities, Graduate School of Letters, Hokkaido University.
- Tsumagari, T. (2009b) Grammatical outline of Uilta (Revised). *Journal of the Graduate School of Letters* 4: 1-21. Sapporo: Graduate School of Hokkaido University.
- Vasilevič, G. M. (1936) *Sbornik Materialov po Evenkijskomu (tungusskomu) Fol'kloru*. Leningrad.
- Yamakoshi, Y. (2012) *Kuwashiku Wakaru Mongorugo Bunpou* [=Mongolian grammar]. Hakueisha.
- Zakharov, I. I. (1879) *Grammatika Man'chzhurskogo Jazyka*. St. Peterburg.

### Appendix 1. Linguistic Data

In this paper, following linguistic data are used to examine the functional differences of converbial suffix *-mi* by Tungusic distribution (Russian and Chinese territories) from the perspective of linguistic area.

#### • Russian Tungusic

##### Evenki (I)

Literature: Konstantinova (1968), Nedjalkov, I. V. (1995, 1997), Bulatova & Grenoble (1999)

Text: Vasilevič (1936)

##### Nanay (III)

Literature: Avrorin (1961), Kazama (2010b, 2011a)

Text: Kazama (2005, 2006b, 2007b, 2008b, 2010b)

##### Udihe (II)

Literature: Shnejder (1936), Kormushin (1998), Nikolaeva & Toskaya (2001), Girfanova (2002), Kazama (2010a)

Text: Kazama (2004, 2006a, 2007a, 2008a, 2009, 2010a)

Fieldwork data (conducted in 2013)

#### • Chinese Tungusic

##### Solon (I)

Literature: Poppe (1931), Chaoke et al. (1991), Tsumagari (2009a), Kazama (2011b)

Text: Kazama (2007c, 2008c)

##### Hezhen (II)

Literature: An (1986), Li (2006), Tamura (2008)

Text: Tamura (2008), Li (2011, 2012)

##### Manchu (IV)

Literature: Zakharov (1879), Sung (1968), Tsumagari (1981, 2002), Li (2000), Avrorin (2000), Gorelova (2002)

Text: Choi et al. (2012a, 2012b)



# Stress in Saudi Diaspora Uzbek and its relation to stress in Turkish

Hassan Bokhari and Jonathan North Washington  
*Indiana University*

## Abstract

This paper presents the first known analysis of the accentual system of Saudi Diaspora Uzbek (SDU).

While an increasingly extensive literature considers the accentual system of Turkish along with many of its nuances, there is little work done on the accentual system of other Turkic languages. The Uzbek accentual system is documented superficially by a number of sources, including Камол (1957), Кононов (1960), and Sjoberg (1963). However, even the existence of SDU does not appear to be documented anywhere, despite SDU being spoken by a rather large community in Saudi Arabia.

This paper considers data from SDU and makes generalizations about its accentual system. The Turkish stress system is used as a point of comparison. Drawing parallels with Turkish elucidates several important points of SDU's stress system; however, attention is drawn to ways in which SDU's stress system differs from that of Turkish.

Like Turkish, SDU has a regular stress pattern, along with a small handful of stems which violate this pattern. There are also exceptional suffixes which affect the regular stress pattern in predictable ways. Like Turkish, these include both monosyllabic and disyllabic suffixes. As with Özçelik's (2012) analysis of Turkish, these exceptional suffixes can be explained by positing an accompanying underlying foot, which the grammar makes trochaic. However, stress occurs on the syllable immediately before disyllabic suffixes in SDU, whereas stress occurs on the first syllable of exceptional disyllabic suffixes in Turkish. This difference in behavior is attributed to a difference in the alignment of feet in the two languages. This approach to an analysis of SDU stress has advantages over other approaches used for Turkish, such as that of Kabak and Vogel (2001) and Inkelas and Orgun (1998), as it allows for attested patterns of secondary stress, does not rely on phonological strata or cophologies, and does not predict the existence of Turkish-like initially-stressed disyllabic suffixes in SDU.

## I. Introduction

This paper analyses the stress pattern in Saudi Diaspora Uzbek (SDU), and compares it to the stress pattern of Turkish. While there is a growing body of literature on the accentual system of Turkish, there are only superficial descriptions of the Uzbek accentual system; to our knowledge, there are no formal analyses. Furthermore, there is no known previous documentation even of the existence of SDU.

Saudi Diaspora Uzbek is a set of varieties of Uzbek, a Turkic language, spoken mainly in the midwest part of Saudi Arabia in the region of Hijaz, including the cities of Mecca, Medina, Jeddah, and Taif. The authors estimate the number of speakers to be no less than 60,000. The speakers of this community are mostly descendants of immigrants after the Russian Revolution, people who went on Hajj and stayed, merchants, and recent immigrants from Afghanistan and other countries (sometimes via Turkey). While SDU is actively used among Uzbek communities in Saudi Arabia, there is hardly even mention of its existence outside of these communities. There are no known written sources on the variety, such as linguistic documentation or analyses of any sort.

The stress patterns of SDU and Turkish are very similar, but there are some points where they diverge. The main difference is that Turkish has disyllabic suffixes which exhibit the exceptional pattern of receiving stress on the first syllable, while exceptional disyllabic suffixes in SDU trigger stressing of the previous syllable. The analysis presented in this paper for SDU is much like the analysis of Özçelik (2012) for Turkish, but differs in how an underlying foot is aligned with these exceptional suffixes.

Data for SDU was compiled by examining grammars of Uzbek (Камол, 1957; Sjoberg, 1963; and Кононов, 1960) and comparing their accounts of stress in Modern Literary Uzbek to the pronunciation of the first author, whose is a native speaker of SDU.

Section 2 of this paper will present the data for SDU and Turkish, and section 3 overviews previous analyses of Turkish stress. Section 4 presents the present analysis, and section 5 concludes.

## 2. The data

### *Regular stress*

In Saudi Diaspora Uzbek, in cases of regular stress, the last syllable of a word, no matter how long or how many morphemes it contains, receives stress.

- (1) [tʰɔwɔ́ɪ] ‘plate’  
 [tʰɔwɔ́ɪlár] ‘plates’  
 [tʰɔwɔ́ɪlarím] ‘my plates’  
 [tʰɔwɔ́ɪlarímdá] ‘on my plates’  
 [tʰɔwɔ́ɪlarímdakí] ‘the one on my plates’  
 [tʰɔwɔ́ɪlarímdakilár] ‘the ones on my plates’

In Turkish, the pattern is identical. Words with the same glosses follow:

- (2) [tʰabákaʰ] ‘plate’  
 [tʰabaklár] ‘plates’  
 [tʰabaklarúɰ] ‘my plates’  
 [tʰabaklarúmdá] ‘on my plates’  
 [tʰabaklarúmdakí] ‘the one on my plates’  
 [tʰabaklarúmdakilár] ‘the ones on my plates’

Most SDU stems borrowed from other languages receive regular stress:

- (3) [dímfér] ‘Denver’  
 [koloradó] ‘Colorado’  
 [aŋqʰará] ‘Ankara’ (from Arabic [aŋqara])

### *Irregularly stressed stems*

Sjoberg (1963, 25) presents some proper nouns that receive exceptional stress in Uzbek, as follows:

- (4) [lutʰúlla] ‘Lutfullah’ (guy’s name)  
 [lɔ́la] ‘Lola’ (girl’s name)  
 [tɔ́ʃkent] ‘Tashkent’  
 [ɔ́lmónija] ‘Germany’

In Saudi Diaspora Uzbek, most of these words receive regular stress (e.g., [tɔ́ʃkánd]), though some names from Arabic may receive exceptional stress:

- (5) [lutʃúlla] ‘Lutfullah’ (guy’s name)  
[asadúlla] ‘Asadullah’ (guy’s name)

Sjoberg (1963, 24) and Камол (1957, 204) also present some adverbs and conjunctions borrowed from other languages that receive irregular stress, transcribed below in SDU:

- (6) [albátta] ‘certainly’  
[másala(n)] ‘for example’  
[ámma] ‘but’  
[bázi] ‘some’  
[hámma] ‘all’  
[lékin] ‘but’  
[hózir] ‘now’  
[ʃúunki] ‘because’

Additionally there are some adjectives in SDU formed with the Persian prefix /ser/- which are stressed on this prefix: [sérpul] ‘having a lot of money’, [sérujqu] ‘sleeping a lot’. There are also some interrogative words that receive irregular stress: [qájsi] ‘which’, [qándaj] ‘how / what kind’, [námaga] ‘why’. There is no generalisable pattern about the words that receive irregular stress in Uzbek. In many cases, these words preserve the stress of their language of origin.

In Turkish, many place names follow an irregular stress pattern as well. Instead of receiving stress on the final syllable, they receive stress on a heavy antepenult if the penult is light ([án.k<sup>h</sup>a.ra], [ʃév.ro.lɛ]); otherwise, the stress falls on the penult ([is.t<sup>h</sup>án.bul], [e.dír.nɛ], [a.dá.na]). Place names derived from non-proper nouns mostly follow this pattern ([ba.k<sup>h</sup>á.ɟak<sup>h</sup>], [t<sup>h</sup>ór.ba.lu], [bé.bek]). Some proper names from other languages ([k<sup>h</sup>ɛ.né.di], [p<sup>h</sup>a.p<sup>h</sup>a.do.p<sup>h</sup>ú.los]) also follow this pattern, regardless of their stress in the language of origin, but not all proper names from other languages receive this pattern of irregular stress ([man.dé.la], [bar.bá.ra]). Hence, it is possible to say that there is no consistent generalisable pattern about the accent in these words.

### *Pre-stressing suffixes*

In Uzbek, there is a class of morphemes which primary stress always precedes. The following list enumerates some but not all of them. Many of the morphemes were drawn from Sjoberg (1963), Камол (1957), and Кононов (1960), but not all of their examples are correct in SDU, and so were excluded, and not all the examples provided here are listed by these sources.

- Some postpositions - [nimá ʃʃin] ‘why?’, [keʃadán beri] ‘since yesterday’ (note that the preposition gets secondary regular stress)
- The negative verbal suffix -/ma/ - [bɔʃlá-ma-di-m] ‘I didn’t start’, [k<sup>h</sup>él-ma-gan] ‘hasn’t come’
- The equative suffixes -/daj/ and -/daqa/ - [ɔltındaj]/[ɔltındaqa] ‘like gold’, [ɔdam-lár-daqa] ‘like people’
- Some derivational suffixes dealing with number: [úʃʃala] ‘three in all’, [ózgina] ‘just a little’
- Some inflectional suffixes dealing with number: the classifier /ta/ [júz ta] ‘one hundred’,<sup>1</sup> [júz-ta-ʃʃa] ‘about one hundred’; the suffix -/ɟa/, as in [júz-lár-ɟa] ‘hundreds’, [júz-ɟa] ‘about a hundred’
- The null copula, as evidenced by the following:<sup>2</sup>
  - Personal copula particles: [ʃʃirjilɩk-miz] ‘we are beautiful’, [xunák-san] ‘you are ugly’, [ʃʃaq<sup>h</sup>ir-gán-man] ‘I invited’.
  - Modal particles: [u ʃʃirjilɩk ku] ‘s/he’s beautiful!’ (emphatic/intensifying, used when there’s doubt whether the interlocutor is aware of / sure about the information), [hasán dir] ‘it might be Hassan’

<sup>1</sup> Sjoberg (1963, 25) classifies this morpheme as prestressing, but in Saudi Diaspora Uzbek, this classifier can receive regular stress with some numbers, perhaps varying with intonational or syntactic differences, e.g. [ik<sup>h</sup>i tá], [uʃʃ tá], [ɔlt<sup>h</sup>i tá], [jeti tá], [sakk<sup>h</sup>iz tá], [toqq<sup>h</sup>iz tá]. This classifier is interesting to study on its own, but will mostly be ignored in this study because of its unpredictable behavior.

<sup>2</sup> See Kabak & Vogel (2001, §7) for a complete account of the null copula in Turkish.



(dubitive, used when speaker is unsure of the information), [u ʃiɾɔjliɾ ekan] ‘apparently s/he’s beautiful’ (hearsay) / ‘Wow, s/he’s beautiful’ (surprise at new information).

- Certain conjunctions: [qʰoʃini-m da] ‘my neighbor also’, [motʰor-imiz-gá da] ‘to our car also’,<sup>3</sup> [mahammád-u hasán kelifti] ‘Mahammad and Hassan arrived’
- Interrogative particles: [jangí-mi] ‘is it new?’, [kʰel-dí-ŋ-mi] ‘did you come?’, [sán-ʃi] ‘how about you?’
- The relativiser /ki/: [ajt-ti-ki] ‘he said that ...’.
- The personal suffixes of optative and softened imperative forms: [kor-éj-lik] ‘let’s see’, [vɔʃ-ij-lik] ‘let’s open it’, [kór-sin] ‘let him/her see’; [vɔʃ-ij-lar] ‘please open it (2pl)’, /bɔʃla-(i)j-lar/ [bɔʃlé-lar] ‘please begin (2pl)’, [tuʃun-ij-lar] ‘please understand (2pl)’
- The general/future tense morpheme /a/: [iʃán-a-man] ‘I believe’, [két-a-san] ‘you leave’, /bɔʃla-j-man/ [bɔʃleman] ‘I begin’.
- Causative morphemes, regular (-/tir/, -/dir/, -/t/) and irregular (-/ir/, -/kaz/, -/kiz/):
  - [qoʃ-ijáp-man] ‘I’m running away’ - [qoʃ-ir-ijáp-man] ‘I’m driving [someone] away’
  - [bɔla kúl-di] ‘the boy laughed’ - [men bɔla-ni kúl-dir-di-m] ‘I made the boy laugh’
  - [múz erí-j-di] ‘the ice melts’ - [múz-zi éri-t-a-man] ‘I melt the ice’
  - [eʃak ol-gán] ‘the donkey died’ - [eʃak-ki ol-tir-gán-man] ‘I killed the donkey’
  - [axmát-niŋ kasal-í ót-ti] ‘Akhmat’s disease spread’ - [axmát kasal-i-ni ót-kaz-di] ‘Akhmat spread his disease’
  - [qiz-lar uʃ-lar-i-ga jét-ti] ‘The girls got home’ - [qiz-lar-ri uʃ-lar-i-ga jét-kiz-di-m] ‘I brought the girls home’

In Turkish there are also quite a few pre-stressing suffixes. The following list was based off of Kabak & Vogel (2001, 328):

- The negative verbal suffix -/mA/ - [baʃlá-ma-du-m] ‘I didn’t start’.
- The -/dʒA/ adverbial suffix - [al-mán-dʒa] ‘in German’
- The coordinator /DA/ - [kom-ʃum da] ‘my neighbor also’.
- The relativiser [ki] - [sana-r-um-ki] ‘I think that ...’.
- Overt (shown in bold) and non-overt copula suffixes - [gyzæl-iz] ‘we are beautiful’, [ʃirkʰin-sin] ‘you are ugly’, [ʃaúr-muʃ-um] ‘I invited’, [ʃoʃúk-kʰæn] ‘when I was a kid’ [hastʰá-j-dum] ‘I was sick’.<sup>4</sup>
- Modal particles - [jení-mi] ‘is it new?’, [gʲæl-dí-n-mi] ‘did you come?’, [gid-ædʒæk-lær ja] ‘so you know how they’re going to come’
- Instrumental / commutative case /(j)lA/ - [arabá-m-la] ‘with my car’ (note the uncliticised form /ile/, [arabá-m ilè])

### *Disyllabic exceptional suffixes / stressed suffixes*

In Turkish, there is a small class of disyllabic suffixes that have exceptional stress. These suffixes carry stress on their initial syllable, and include only the morphemes -/(j)ÍndʒA/, -/(j)ÁrAk/, and -/Íjor/.<sup>5</sup>

SDU has no multisyllabic suffixes with initial stress, but it does have disyllabic exceptional suffixes. All disyllabic exceptional suffixes in SDU are pre-stressing. These include -/daqa/ and -/ala/, discussed above.

This paper will not directly deal with what happens when multiple stressed suffixes or an irregular-stressed word and a stressed suffix occur together. It will also not deal with compounds or reduplication. For a discussion on the stress patterns of reduplication in Uzbek, see КОНОХОВ (1957, §55, p. 55).

<sup>3</sup> Sjoberg (1963, p. 26) says that the conjunction /ham/ is pre-stressing, but in Saudi Diaspora Uzbek, it is incorporated into the phonological word and receives normal stress, e.g. [qʰoʃini-m ám], [motʰor-imiz-ga jám]. However, there is at least one example in Saudi Diaspora Uzbek where this conjunction is prestressing, i.e. [búnda jam] ‘here also’.

<sup>4</sup> While the copula itself /j/ is often not overt and is never syllabic when cliticized, the suffix following the copula, one of -/DI/, -/mIʃ/, or -/sA/, acts as the pre-stressing suffix.

<sup>5</sup> -/Ijor/ only takes exceptional stress in literary registers in Turkish; normally it receives regular stress (Kabak & Vogel, 2001, 332).

### 3. Analyses of Turkish stress

#### *Early analyses*

#### **Irregularly stressed stems**

Sezer (1981) presented one of the first accounts of irregularly stressed stems in Turkish. The analysis was summarized by Kaisse (1986) as follows: “mark final syllables extrametrical; at the right edge of the word, form a left-dominant, quantity-sensitive, binary foot.”

In irregularly stressed stems like [pénaltu] ‘penalty’, Inkelas (1994) and Inkelas and Orgun (1998) propose that a trochaic foot is prespecified, in a similar way to prestressing and stressed suffixes (see below). Inkelas and Orgun (1998) treat Sezer stems differently than this kind of stem, positing a cophonology that applies Sezer stress patterns to foreign and place names via a trochaic foot. Because final-stressing place names (e.g., [anadolú]) show regular stress patterns, they specify yet another cophonology to deal with these.

#### **Prestressed suffixes**

Lees (1961), as summarised by Kaisse (1986), analyzes pre-stressing suffixes in the following way: all regular words are marked in the lexicon as having word-final stress and regular suffixes are marked in the lexicon as receiving stress, e.g. /<sup>h</sup>abák-lár-úm-dá-kí-láér/. A rule removes stress from all but the rightmost of a string of stressed syllables, e.g. [<sup>h</sup>abaklarumdakilær] ‘the ones on my plates’. This ensures that the rightmost syllable of a word is stressed, as well as the syllable before a pre-stressed suffix, e.g. /gyzáel-láér-dáén-sin/ [gyzáellærdáénsin] ‘you are one of the beautiful ones’, since pre-stressed suffixes are not marked as stressed in the lexicon. To prevent syllables after a pre-stressed suffix from receiving stress, a separate rule removes any stress to the right of a suffix that does not receive stress, e.g. /baɟlá-ma-(j)án/ [baɟlámajan] ‘not started’.

Kaisse (1986, 235-236) takes issue with previous theories like Lees (1961), stating that they all require marking prestressed suffixes in some way in the underlying form. Kaisse (1986) instead notes that so-called pre-stressing or pre-stressed suffixes are poorly named because they don’t always cause stress immediately before them, and argues that they are in fact all “clitics placed at the edges of normally affixed words.” Kaisse (1985, 205) shows a pattern of “-en” adverbs (e.g., [náktæn] ‘in cash’, [tekæffulæn] ‘by surety’, [munhásuran] ‘specially’), originally noticed by Sezer (1983), where stress falls on the penult if it is strong, and on the antepenult if the penult is weak. Kaisse (1985 and 1986) proposes three lexical strata: the first is where feet are assigned to irregularly stressed stems, including the “-en” adverbs; the second is where normal word stress applies, and excludes clitics; and the third includes clitics (i.e., suffixes that are never stressed) and is the domain over which vowel harmony applies. By excluding clitics from the lexical stratum over which stress applies, Kaisse (1986) ensures that stress will occur before the first clitic.

For Inkelas (1994) and Inkelas and Orgun (1998), regular stress occurs via catalexis, where an empty syllable is inserted after the last syllable of a stem and a trochaic foot is formed, as shown in (7a). They posit a cophonology with an underlying foot in pre-stressing suffixes, as shown in (7), where the regularly stressed [tekmeledi] (7a) is contrasted with the form [tekmeleme] (7b), where the negative suffix /mA/ is prespecified as having a trochaic foot.

(7a)	(b)	F
/ tekmele-DI /	/ tekmele-mA /	
[ tek.mɛ.le.(di [σ]) ]	[ tek.mɛ(lɛ.mɛ) ]	

This is the first account of Turkish exceptional stress known to have presented a trochaic foot-based analysis.

### Stressed suffixes

Not many sources mention stressed suffixes. Inkelas (1994) and Inkelas and Orgun (1998, 371-372) posit an underlying foot in stressed suffixes, as shown in (8),<sup>6</sup> where the regularly stressed [japadʒák] is contrasted with the form [japáراك], where the verbal adverb suffix -(j)ArAk/ is prespecified as having a trochaic foot.

(8a)	(b)
$\begin{array}{c} / \text{jap-(j)AdʒAk} / \\ [ \text{ja.pa.dʒák} ] \end{array}$	$\begin{array}{c} \text{F} \\   \\ / \text{jap-(j)ArAk} / \\ [ \text{ja(pá.rak)} ] \end{array}$

*Kabak & Vogel (2001)*

### Regular stress

Kabak & Vogel (2001) present a rule for regular stress assignment (p. 324): stress the final syllable of a Phonological Word (PW). They define a PW as the root along with most suffixes.

### Irregularly stressed stems

Kabak & Vogel (2001) argue that the pattern noticed by Sezer (1981) does not account for all irregularly stressed stems. They propose an analysis where the stressed syllable of these stems is simply specified lexically, along with a rule (p. 325) that assigns stress to syllables lexically marked as stress-bearing. Their final stress rule is “Stress a lexically marked syllable; otherwise stress the final syllable of a PW” (p. 329).

### Pre-stressing suffixes

Kabak & Vogel (2001) present an analysis of Turkish where many morphemes, including the copula, question marker, and instrumental suffix are labelled in the lexicon as Phonological Word Adjoiners (PWA). These suffixes are “obligatorily excluded from the PW”. Since the PW is always stressed on the last syllable, this ensures that pre-stressing suffixes trigger stress on the previous syllable. An example of this is [baʃ.lá]<sub>PW</sub>-ma-du-m], where /mA/ is labelled in the lexicon as a PWA.

This approach is unified in that there is only one stress rule for regularly stressed words and pre-stressing suffixes. However, it is also like Kaisse’s (1986) analysis involving lexical strata, where the PW can be seen as the second stratum, and the morphological word can be seen as the third stratum.

### Stressed suffixes

While Kabak & Vogel (2001) do not directly address disyllabic stressed suffixes as a class of suffixes, they do notice that /Ijor/ sometimes receives exceptional stress on the first syllable. They deal with this by positing an underlying specification for stress, as with exceptionally stressed words.

---

<sup>6</sup> Despite the appearance of the representation in (8b), the foot is not aligned to any particular segment, but is simply specified underlyingly along with the morpheme.

Özçelik (2012)

### Regular stress

Özçelik (2012, §2.2) considers regular word-final stress in Turkish to actually be a boundary tone, based on phonetic and cross-linguistic phonological evidence. He accounts for this “final prominence” with a simple rule: “place a boundary tone at the end of a PwD” (p. 25).

### Irregularly stressed stems

Özçelik (2012), like Kabak & Vogel (2001), assumes that the stress in irregularly stressed stems is pre-specified.

### Pre-stressing suffixes

Özçelik (2012) assumes that pre-stressing suffixes are specified underlyingly with a foot. A rule that aligns the right edge of a foot in the UR with the right edge of a foot in the SR, along with a general specification that foot shape in Turkish must be trochaic, ensures that pre-stressing suffixes (which are all monosyllabic) will be footed as the second and unstressed syllable of the foot in the SR, with primary stress falling on the syllable preceding the suffix, as in /ʃaluʃ-(mA)<sub>F</sub>/ [ʃa(luʃ.ma)<sub>F</sub>] ‘don’t work’.

### Stressed suffixes

Özçelik’s (2012) analysis of stressed suffixes is very similar to that of pre-stressed suffixes, and requires no additional theoretical framework. Özçelik notices that all pre-stressing suffixes in Turkish are monosyllabic and all stressed suffixes are disyllabic. This makes it possible to posit an underlying foot for both classes of suffix, that aligns to the right of the suffix and is trochaic. An example is /ʃaluʃ-(ArAk)<sub>F</sub>/ [ʃa.lu(ʃá.rak)<sub>F</sub>] ‘while working’.

### Comparison to other accounts

Özçelik’s (2012) approach is very similar to that of Inkelas (1994) and Inkelas and Orgun (1998), in that feet are specified underlyingly with exceptional suffixes. However, there are some notable differences: in their analysis, there must be pre-stressing, exceptionally stressed, and regularly stressed cophologies; also, they specify trochaic feet underlyingly, whereas Özçelik (2012) specifies only the presence of a foot underlyingly and the grammar makes it trochaic. The advantage of Özçelik’s analysis is that all suffixes are subject to the same phonological parameters, and the only difference between exceptional suffixes (that is, pre-stressing and stressed suffixes) and regular suffixes is that exceptional suffixes are underlyingly footed. Özçelik’s analysis also explains the absence of stressed (as opposed to pre-stressing) monosyllabic suffixes as well as the absence of iambic disyllabic suffixes.

One advantage of Özçelik’s account over that of Kabak & Vogel (2001) is that, in their account, once a PW is ended with a PWA, there can be no further stress. Özçelik (2012, 41) points out examples of secondary stress, like [dɪnlɛmɛdɪ dɛ] ‘s/he didn’t listen either’, which Kabak & Vogel must treat as having only one stress, since the PW ends before the PWA -/mA/, disallowing further stress assignment after it. In Özçelik’s (2012) analysis there is no such restriction. Other examples of secondary stress that are accounted for in this analysis

are [baʃ(láma)dùm] ‘I didn’t begin’ and [an(láma)du(làrmu)] ‘did they not understand?’ Kabak & Vogel (2001) also have the weakness of Inkelas and Orgun (1998), in that there is nothing preventing the existence of iambic disyllabic stressed suffixes and stressed monosyllabic suffixes.

#### 4. Stress in Saudi Diaspora Uzbek

This paper will take an approach similar to that of Özçelik (2012) in its analysis of stress in SDU, with one difference necessary for SDU.

##### *Regular stress*

Like Özçelik (2012) did for Turkish, the present analysis considers regular word-final stress in Uzbek to actually be a boundary tone. While there is currently no good phonetic data available on this phenomenon in Uzbek, the phonological evidence is identical to that in Turkish. This “final prominence” can be accounted for with a simple rule: “place a boundary tone at the end of a PWd” (Özçelik, 2012, 25).

Based on Özçelik (2013, 2014 §3.1), since final prominence is crucial in Turkish for regular words, there can be no trochaic structure in these words; however, having trochees is also an essential part of Turkish grammar. The only other way to get final-prominence with feet in Turkish would be to have a monosyllabic foot; however, foot-binarity is also crucial in Turkish. Per Özçelik’s (2013, 2014) OT account, PARSE- $\sigma$  is ranked lower than these other constraints, and so final prominence may be achieved by simply applying a boundary tone to the final syllable of a PWd without parsing any syllables into feet. The present analysis will assume that Uzbek grammar is much like Turkish grammar in this respect.

In a word like /tɒwɒɛ-lar-im-da-ki-lar/ ‘the ones on my plates’, there is no footing (be it trochaic or iambic): \*[t<sup>h</sup>ɒwɒɛlarimda(ki.lar)]. Instead, a boundary tone is simply applied to the last syllable: [t<sup>h</sup>ɒwɒɛlarimdakilár].

##### *Irregular stress*

The literature seems to disagree as to whether patterns of irregular stress noticed in Turkish can be accounted for by Turkish-internal phonology or based on the preservation of source-language stress in borrowed words. Instead of coming up with multiple co-phonologies like Inkelas and Orgun (1998) or with multiple phonological strata, like Kaisse (1985), more recent accounts of Turkish, like Kabak & Vogel (2001) and Özçelik (2012), assume that irregularly stressed stems in Turkish are simply underlyingly specified for stress.

Saudi Diaspora Uzbek, much like Turkish, seems to preserve the stress in certain words from Persian and Arabic, but in many cases simply applies regular stress to borrowed words. Much like Özçelik (2012), this analysis assumes that stress can be specified underlyingly, and thus accounts for irregularly stressed stems. For example, /albátta/ is underlyingly marked as having stress on the penultimate syllable, and hence surfaces this way [albátta]. A word like /tɒwɒɛ-lar-im-da-ki-lar/ has no underlyingly marked stress, and so surfaces with regular stress. It is striking that SDU does not have a class of proper nouns with an irregular stress pattern.

##### *Exceptionally stressed suffixes*

The analysis presented in this study for exceptionally stressed suffixes in SDU is much like that presented by Özçelik (2012) for exceptionally stressed suffixes in Turkish, except that the alignment between the foot and affix is different.

As mentioned previously, Uzbek and Turkish both have disyllabic morphemes that participate in exceptional stress. In Turkish, these suffixes always receive stress on the first syllable of the suffix, as shown in example (9) for the suffixes *-(j)IndʒA/* and *-(j)ArAk/*.

- (9) a. [gʲæɫ-índʒɛ] ‘when [someone] came’  
 b. [ʃʲalɯʃ-úndʒa] ‘when [someone] worked’  
 c. [gʲæɫ-érek] ‘by coming’  
 d. [ʃʲæɫɯʃ-áراك] ‘by working’

These forms can be compared to regularly stressed disyllabic morphemes, such as *-(j)AdʒAk/*, as shown in (10).

- (10) a. [gʲæɫ-édʒék] ‘s/he will come’  
 b. [ʃʲalɯʃ-adʒák] ‘s/he will work’

In SDU, disyllabic suffixes also participate in exceptional stress. In these suffixes, unlike in Turkish disyllabic exceptional suffixes, the stress always precedes the affix. Examples are shown in (11) for */daqa/* and */ala/*.

- (11) a. [tór-ala] ‘four together’  
 b. [erkág-daqa] ‘like a man’  
 c. [erkag-lár-daqa] ‘like men’

These forms contrast with regularly stressed disyllabic morphemes, as shown for */daki/* in (12).

- (12) a. [koz-dakí] ‘the one in the eye’  
 b. [erkag-dakí] ‘the one at the man’

Özçelik (2012) posits a foot that is underlyingly specified together with the morpheme for this class of suffixes in Turkish, e.g. */((j)IndʒA)<sub>F</sub>/*. The language’s grammar makes the foot trochaic, resulting in the first of the two syllables of the suffix being stressed, as shown in (13).

- (13) a. [gʲæɫ(ín.dʒɛ)] ‘when [someone] came’  
 b. [ʃʲalɯʃ(úɯn.dʒa)] ‘when [someone] worked’

Regularly stressed suffixes have no foot specified underlying, and so a regular boundary tone is simply assigned word-finally if there is no foot already present.

Where Turkish aligns the underlying foot in exceptionally stressed suffixes with the rightmost syllable of the affix, SDU aligns the right edge of the foot to the left edge of the affix. This results in the syllable immediately preceding the suffix to be incorporated into the foot, and receive stress via the grammar assigning trochaic stress to feet, as shown in (14).<sup>7</sup>

- (14) a. */tor-(ala)<sub>F</sub>/* - [(tó.ra)là] ‘four together’  
 b. */erkag-(daqa)<sub>F</sub>/* - [er(kág.da)qà] ‘like a man’  
 c. */erkag-lar-(daqa)<sub>F</sub>/* - [erkag(lár.da)qà] ‘like men’

Note that regular stress is still assigned to the final syllable, which in SDU is not footed. Also note that these final syllables cannot be extrametrical due to this fact. More extreme examples are the words */tor-(ala)<sub>F</sub>-si-ni/* [(tó.ra)<sub>F</sub>la.si.ni] ‘the four of them (acc.)’ and */erkag-(daqa)<sub>F</sub>-lar-ni/* [er(kág.da)<sub>F</sub>qa.lar.ri] ‘those that are like a man (acc.)’;<sup>8</sup> in these examples, if [la] in */ala/* or [qa] in */daqa/* were extrametrical, regular stress would not be able to be assigned to any syllables beyond that one, preventing secondary stress on the final syllables of

<sup>7</sup> As in (8), the URs in (14) should only be understood to indicate that a foot is specified with the morpheme, and not that it is aligned in any particular way.

<sup>8</sup> The fact that nominal morphology can follow these suffixes is good evidence that they are not simply free particles, but are in fact bound morphemes.



these words.

Under Özçelik's analysis, since the foot is aligned with the right edge of the affix, monosyllabic exceptional affixes in Turkish become the unstressed syllable of the trochaic foot, and the syllable immediately preceding the affix receives stress, as shown in (15).

- (15) a. [baʃ(lá.ma)] 'don't start'  
b. [ha(sán da)] 'Hassan also'

In SDU the corresponding syllable receives stress, but formally, unlike in Turkish, the rightmost syllable of the foot is aligned with the leftmost syllable of the affix. Since the affix is monosyllabic, footing happens in the same way [on the surface] as in Turkish. Examples matching the Turkish ones in (15) are shown in (16) for SDU.

- (16) a. [baʃ(lá.ma)] 'don't start'  
b. [ha(sán da)] 'Hassan also'

This analysis predicts the absence of initially stressed disyllabic suffixes (like those in Turkish) in SDU.

## 5. Conclusion

This is the first study known to examine the phonology, or indeed address any aspect of the grammar, of Saudi Diaspora Uzbek. This paper has compared the stress pattern in Saudi Diaspora Uzbek with the stress pattern in Turkish, by suggesting that SDU also underlyingly specifies feet for exceptional suffixes, and that feet in the language are trochaic. Besides the striking absence of an entire class of exceptionally stressed proper names, the main difference between SDU and Turkish is that SDU aligns the second syllable of the foot to the leftmost syllable of exceptional suffixes, whereas Turkish aligns the second syllable of the foot to the rightmost syllable of exceptional suffixes. The account is based on that of Özçelik (2012), which is preferred over other accounts of Turkish stress because it accounts for secondary stress when there's more than one exceptionally stressed suffix, which is true for SDU as well. The resulting account also predicts the absence of initially stressed disyllabic suffixes in SDU, and presents everything as a unified analysis, with no cophologies or lexical strata.

An analysis of SDU stress in Optimality Theory that parallels Özçelik's (2013, 2014) analysis for Turkish is planned; however, since the different-edge alignment constraints that would be necessary to implement our preliminary analysis presented here are disfavored in the literature (cf. McCarthy, 2008, 214), a fairly different approach may be necessary. The authors would also like to examine more complex examples of stress in SDU, such as those involving compounding, reduplication, and secondary stress. Another interesting area of study is what happens with morphemes that can surface either with or without a vocalic nucleus, such as the SDU aorist (-/a/ or -/j/) and regular causative morphemes (-/Dir/, -/t/), which are both pre-stressing, but seem to behave differently in terms of whether the preceding vocalic nucleus or the preceding syllable receives stress; this issue appears not to have been addressed in the literature on Turkish, perhaps because Turkish does not have similar language-internal inconsistencies.

While the analysis presented is not fully developed, the authors hope that it will be a basis for future discussions of Turkic stress, especially SDU and other varieties of Uzbek, and how currently available work on Turkish fits into a wider typological distribution of phenomena.



## References

- Inkelas, Sharon (1994). Exceptional stress-attracting suffixes in Turkish: representations vs. the grammar. Utrecht University: Workshop on Prosodic Morphology.
- Inkelas, Sharon and C. Orhan Orgun (1998). Level (Non)Ordering in Recursive Morphology: Evidence from Turkish. In *Morphology and Its Relation to Phonology and Syntax*. Stanford, CA: CSLI. 360-392.
- Kabak, Barış and Irene Vogel (2001). The phonological word and stress assignment in Turkish. *Phonology* 18. Cambridge University Press. 315-360.
- Kaisse, Ellen (1985). Some theoretical consequences of stress rules in Turkish. In *Proceedings of the Chicago Linguistic Society* 21:1. 199-209.
- Kaisse, Ellen (1986). Towards a lexical phonology of Turkish. In *A Festschrift for Sol Saporta*. Seattle: Noit Amrofer. 231-239.
- Камол, Фахри (ed., 1957). *Ҳозирги замон ўзбек тили : лексикология, фонетика, графика ва орфография, морфология*. Тошкент: Ўзбекистон ССР Фанлар Академияси Нашриёти.
- Кононов, А. Н (1960). *Грамматика современного узбекского литературного языка*. Москва: Академия наук СССР.
- Lees, Robert B (1961). *The phonology of Modern Standard Turkish*. Uralic and Altaic Series 6. Bloomington: Indiana University Publications.
- McCarthy, John J (2008). *Doing Optimality Theory*. Oxford: Blackwell.
- Özçelik, Öner (2012). *Representation and acquisition of stress: the case of Turkish*. Doctoral thesis, McGill University.
- Özçelik, Öner (2013). Exceptions in Stress Assignment: Feet in Input. In *Proceedings of NELS 40*, v. II. Amherst, MA: GLSA. 107-120.
- Özçelik, Öner (2014). Prosodic faithfulness to foot edges: the case of Turkish stress. *Phonology* 31, 229-269.

# Why situation theory cannot explain the accusative case marking in Turkish

M. Ali Bolgün<sup>1</sup>

*Monterey Institute of International Studies\**

## Abstract

The accusative (ACC) is not a well-understood case in Turkish; it has been a challenge for linguists since Seaman (1670), if not earlier, who noted that the accusative marker signifies the direct object (DO) in Turkish. Because Turkish does not have any morphological determiners or a definite article, such as *the* in English (e.g., Underhill 1976; Erguvanlı 1984; Kornfilt 1997), ACC, one of the six cases in Turkish, has been characterized (generally speaking) either as corresponding to the definite article in English (e.g., Banguoğlu 2004; Crowley 1938; Ergin 1962; Erguvanlı 1984; Erguvanlı-Taylan 1987; Gencan 1970, 2001; Lewis, 2000; Mowle 1934; Nemeth 1962; Sebüktekin 1971; Swift 1963), as indicating referentiality (e.g., Dede 1986), or as indicating specificity (e.g., Swift 1963; Erguvanlı 1984; Enç 1991; Kornfilt 1997; Aygen-Tosun 1999; von Heusinger 2002; Aissen 2003). However, none of the approaches put forth so far seems to be able to fully capture the meaning and/or the function of the ACC marker. The presence or absence of ACC does not necessarily correlate with definiteness, referentiality, or specificity. DOs may have ACC but may be indefinite, non-specific, and/or non-referential (Bolgün 2005). Nakipoğlu (2009) investigates accented versus unaccented ACC marked definite objects, arguing that the former produces existential whereas the latter pragmatic presupposition respectively; however, it does not address the alternation between ACC versus no ACC on DOs (i.e., why some DOs take ACC in the first place). Kılıçaslan (2006) offers a situation-theoretic account of ACC and suggests that if the descriptive content is not part of what characterizes the situation described by the sentence, the DO bears ACC; otherwise, it does not. In this article, I argue and show through examples taken from METU Turkish corpus (Say et al. 2002) and Turkish newspapers, that the situation-theoretic account cannot explain the ACC marking.

*Keywords:* accusative; case marking; corpus-based; specificity; referentiality; situation theory; definiteness

## 1. Introduction and problem statement

In this paper, due to page limitations, I will address only Kılıçaslan's (2006) situation-theoretic account, leaving out (or touching very briefly on) other valuable contributions reported in various studies cited in the abstract. Researchers interested in this topic are encouraged to read those studies some of which are listed in the References section, which is by no means a complete list.

In linguistic literature on ACC in Turkish, there is usually a mention of four distinct DO types. These are illustrated in boldface in the following four examples (taken from Taylan and Zimmer 1994)<sup>2</sup> and will be referred to as Type I through Type IV (exemplified by (1) through (4) respectively).

- (1) Ali her gün **gazete-yi** oku-yor.  
 Ali every day newspaper-ACC read-PROG  
 'Ali reads the newspaper everyday.'

<sup>1</sup> M. Ali Bolgün. Tel.: +1-831-402-0338, *E-mail address:* alibolgun@gmail.com

\* The views expressed in this article are those of the author and do not necessarily reflect the official policies of the Monterey Institute of International Studies (MIIS). No funding was received from MIIS for the analysis reported in this article.

<sup>2</sup> Boldfacing is added; the gloss of the first example is slightly modified from the original, to fit the convention that is followed in this paper, and glosses have been added to examples (2), (3), and (4).

- (2) Ali her gün bir **gazete-yi** oku-yor.  
Ali every day one newspaper-ACC read-PROG  
'Ali reads a newspaper everyday.'
- (3) Ali her gün bir **gazete** oku-yor.  
Ali every day one newspaper read-PROG  
'Ali reads a newspaper everyday.'
- (4) Ali her gün **gazete** oku-yor.  
Ali every day newspaper read-PROG  
'Ali reads a newspaper/newspapers everyday.'

The boldfaced nouns in the above examples share a common feature: they all occupy the unmarked DO position, immediately before the verb. What is different about these DOs is that (1) has the accusative (ACC) marker, (2) has the ACC marker and is preceded by *bir*<sup>3</sup> 'one,' (3) is also preceded by *bir* but does not have the ACC marker, and (4) is in its so-called bare form; it neither has the ACC marker nor is it preceded by *bir*.

Given these different ways of expressing the (seemingly) same idea, the question arises as to what the difference is. The boldfaced noun in (1) is generally considered to be definite, in the sense that the hearer knows or can identify the *gazete* 'newspaper' being mentioned. In (4), with no ACC or a preceding *bir*, it is indefinite or even generic, in the sense that *gazete* refers to the category of newspapers in general. Therefore, while the presence and absence of ACC marker on the boldfaced nouns, in (1) and (4) respectively, can perhaps be explained with the concept of definiteness, it is not of much help, if any, in explaining the presence or absence of the ACC marker in (2) and (3), since the noun *gazete* in these is preceded by *bir*, which is considered by some to be the indefinite article in Turkish. These (i.e., the DOs preceded by *bir*) are explained by appealing to other notions some of which are listed in the abstract. However, none of them fully captures the meaning(s) and function(s) of ACC (See, for example, Bolgün 2005; Johanson 2006; Kornfilt & von Heusinger 2008; Özge 2011, among others).

## 2. Situation-theoretic account and why it cannot explain the ACC in Turkish

Kılıçaslan (2006) offers a situation-theoretic account of case-marking in Turkish. He states that, following situation theory, if the descriptive content is not part of what characterizes the situation described by the sentence, then the NP bears ACC; otherwise, it does not. Let us look at some of the examples presented in support of this account. (6) and (7), Kılıçaslan's (14) and (15), are assumed to be uttered as a reply to the question in (5)<sup>4</sup>, Kılıçaslan's (13).

- (5) Oda-ya gir-diğ-in-de, ne gör-dü-n?  
room-DAT enter-NOM-POSS-LOC what see-PAST-2SG  
'What did you see when you entered the room?'
- (6) Oda-da üç adam var-dı.  
room-LOC three man exist-PAST  
'There were three men in the room.'
- a. ??Adam-lar-dan bir-i bir elma-yı yi-yor-du.  
man-PL-ABL one-POSS one **apple-ACC** eat-PROG-PAST

<sup>3</sup> Taylan and Zimmer (1994) use the term 'indefinite article' to refer to *bir* 'one.' However, there is no consensus on this. For example, while Swift (1963), Lewis (1967) Tura (1973), Taylan and Zimmer (1994), Kornfilt (1997), and Lewis (2000) treat it as such in certain uses, others do not. Aygen-Tosun (1999) cites Crisma (1997) who argues that if a language has only one article, it is expected to be the definite article, and since Turkish does not have a definite article, *bir* is not likely to be an indefinite article.

<sup>4</sup> Examples (5), (6), and (7) are Kılıçaslan's (13), (14), and (15) respectively.

- b. Adam-lar-dan bir-i bir **elma** yi-yor-du.  
man-PL-ABL one-POSS one **apple** eat-PROG-PAST
- (7) Oda-da bir masa ve masa-nın üst-ü-nde üç karınca var-dı  
room-LOC one table and table-GEN top-POSS-LOC three ant exist-PAST  
'There was a table in the room and there were three ants on the table.'
- a. Karınca-lar-dan bir-i bir **elma-yı** yi-yor-du.  
ant-PL-ABL one-POSS one **apple-ACC** eat-PROG-PAST  
'One of the ants was eating an apple.'
- b. ??Karınca-lar-dan bir-i bir **elma** yi-yor-du  
ant-PL-ABL one-POSS one **apple** eat-PROG-PAST

Kılıçaslan argues that the reason (6a) and (7b) are odd when uttered in the actual world (as opposed to an imaginary one such as in a fairy tale) has to do with whether or not the 'apple' (to which is referred through the DO) is within the physical boundaries of the described situation. If it is within the physical boundaries of the described situation, then there is no need for ACC, and if it is not, then the DO will have to take the ACC. Therefore, in (6b), the apple falls within the physical boundaries of the parts of the man's body (including the mouth, the part of face surrounding the mouth and the hands or legs) involved in the act of eating the apple, and as such, there is no need for ACC. In (7b), however, the apple is not within the physical boundaries of the parts of the ant's body (the ant is much smaller than the apple), and as such, ACC is required; without ACC, it sounds odd.

While the explanation provided may seem to account for the alternation of ACC – no ACC, there are numerous examples where this does not hold. For instance, in the first of the following two examples with similar descriptive content, the DO *üniversite* 'university' in relation to 'finishing (i.e., graduating from) a university' is expressed with ACC, while in the second, it is expressed without. Please note that in both examples, the DO *üniversite* can be used with or without ACC without resulting in any oddity.

- (8) Artık **üniversite-yi** bitir-mek ve aynı kariyer-de  
anymore **university-ACC** finish-INF and same career-LOC

ilerle-mek çalışma yaşam-ı açısı-ndan  
progress-INF work life-CM angle-POSS-ABL

garantili bir yol değil.  
guaranteed one way not

'Graduating from a university and progressing in the same career is not a reliable way for the work life anymore.'

- (9) Gör-ül-düğ-ü gibi, dışarı-dan  
see-PASS-PTCL-POSS as out-ABL

bak-ıl-dığ-ı-nda doktor baba, sadık eş  
look-PASS-PTCL-POSS-LOC doctor father faithful spouse

ve **üniversite** bitir-miş, iş hayat-ı-na  
and **university** finish-MPST business life-POSS-DAT

at-ıl-mış delikanlı; toplum-un aile  
throw-PASS-MPST youngster society-GEN family

imaj-ı-na ve değer-ler-i-ne ne de  
image-POSS-DAT and value-PL-POSS-DAT what also

uygun bir aile.  
suitable one family

‘As can be seen, when looked at from the outside, a doctor father, a faithful spouse, and a youngster who has graduated from a university, and is already in the business life; what a family, just fitting the society’s family image and values.’

In the two examples above, *üniversite* is not a unique, or a particular university. The descriptive content is similar in both. In fact, in (8) the descriptive content (with *kariyer* ‘career,’ and *çalışma yaşamı* ‘work life’) is more pronounced in relation to the situation, and as such, following Kılıçaslan’s account, it should not be possible to use ACC here. Yet, the DO is used with ACC.

In (9), the DO *üniversite* is used without ACC. However, if used with ACC, the sentence would easily pass an acceptability judgment without any oddity prescribed to it.

Consider the following counterexample:

- (10) Siz-e yap-acağ-ım iğne bak-ın bu,  
you-DAT do-FUT-1SG injection look-2PL this
- di-yerek **bir ampül-ü** göster-miş, sonra da  
say-ADV **one ampoule-ACC** show-MPST then also
- çek-tiğ-i ilaç-ı yargıç-ın damar-ı-na  
extract-REL-POSS medicine-ACC judge-GEN vein-POSS-DAT
- ver-miş-ti.  
give-MPST-PAST

‘He had shown (him) an ampoule, saying ‘look, the injection that I am going to do is this,’ and then, had injected the medicine that he extracted into the judge’s vein.’

In example (10), the descriptive content (including, the *iğne* ‘needle; injection,’ *ampül* ‘ampoule,’ *ilaç* ‘medicine,’ and *damar* ‘vein’) is clearly part of what characterizes the situation. As such, by Kılıçaslan’s account, the DO *bir ampül* ‘an ampoule’ should not bear ACC. However, in the example, it is used with ACC.

Consider yet another counterexample:

- (11) Karne-ler öğrenci-ler için hazırla-n-ıyor.  
grade.report-PL student-PL for prepare-PASS-PROG
- Veli-ye ver-il-me-si hiç doğru bir yaklaşım  
parent-DAT give-PASS-NML-POSS never right one approach
- değil. Okul-lar **öğrenci-yi** belgelendir-ir.  
not school-PL **student-ACC** certify-AOR

‘Grade reports are being prepared for students. Giving them [the grade reports] to the parents is never a good approach. Schools document [give documents to] the student.’

In (11) as well, the descriptive content (i.e., grade report, students, parents, schools) is also part of what characterizes the situation, which, by Kılıçaslan’s account, should lead to no ACC on the DO. Yet, the DO *öğrenci* ‘student’ has ACC. Incidentally, in (11) *öğrenci* ‘student,’ is neither identifiable nor specific.

Especially challenging counterexamples to Kılıçaslan’s account (and perhaps to all other accounts put forth so far) are those in which the DO can be with or without ACC without leading to any oddities because it would be fairly difficult, if not impossible, to argue that the descriptive content simultaneously is and is not part of what characterizes the situation. For instance, in (12) below, the DO *bir şey* ‘something’ [lit. one thing] is used with ACC. However, it can perfectly be used without ACC. In fact in (13), in an almost identical clause, the DO *bir şey* is used without ACC.

- (12) Bir insan-ı yanlış yönlendir-ecek ne  
 one human-ACC wrong guide-FUT neither
- güc-üm ne de sihr-im var. Bir insan  
 strength-POSS/1SG nor also magic-POSS/1SG exist one human
- bir şey-i** bil-iyor-sa, bil-diği-nden asla  
**one thing-ACC** know-PROG-COND know-REL-ABL never
- vazgeç-mez.  
 concede-NEG

‘I have neither the strength nor the magic power to misguide a human being. If a human being knows something, he never concedes what he knows.’

- (13) Birisi **bir şey** bil-iyor-sa, ben-im yap-tığ-ı-m  
 someone **one thing** know-PROG-COND I-GEN do-NOM-POSS-1SG
- gibi belge-si-ni orta-ya koy-acak.  
 as document-POSS-ACC middle-DAT put-FUT

‘If someone knows something, he shall put forth its document, just like I have done.’

Note that in (13), the DO *bir şey* can perfectly be used with ACC, just as in (12) it can perfectly be used without.

### 3. Discussion

I propose (following Taylan & Zimmer 1994)<sup>5</sup> that the function of ACC is to individuate the entity denoted by the NP to which it is attached, and I also propose that it is the interplay of context, word order, point of view, and meaning of the verb, that leads to the determination as to whether or not the DO will take ACC. All other concepts (such as definiteness, specificity, and referentiality, etc.) come about as a result of the individuating function of ACC because individuation helps the referent of the DO noun to be seen, or thought of as an entity separate from all others around it. However, note that not all ACC-bearing DO nouns are definite or specific. For example, in (8), above, the ACC-bearing DO is not definite and in (12) it is not specific. On the other hand, DO can be specific even without ACC, as in *bir bebek* in (14), which refers to a specific baby; namely, Necla’s baby.

- (14) Korku-dan yaşa-dık-lar-ı-nı kimse-ye  
 fear-ABL live-REL-PL-POSS-ACC nobody-DAT
- anlat-a-ma-yan ve defa-lar-ca tecavüz-e  
 tell-ABIL-NEG-REL and count-PL-ADV rape-DAT
- uğra-yan Necla, hamileliđ-i-nin 7.  
 be.exposed-REL Necla pregnancy-POSS-GEN 7th
- ay-ı-nda ölü **bir bebek-Ø** dünya-ya getir-ince  
 month-POSS-ABL dead one baby world-DAT bring-WHEN
- ‘sır-ı orta-ya’ çık-tı.  
 secret-POSS out-DAT exit-PAST

<sup>5</sup> Taylan and Zimmer do not commit themselves to any particular definition of the term. In their seminal study of transitivity, Hopper and Thompson take individuation to refer “both to the distinctness of the patient from the A[gent] and to its distinctness from its own background” (1980: 253). The term individuation is used in a similar fashion in this paper, although as Hopper and Thompson acknowledge, individuation is more complex than it is often thought to be.



‘When Necla, who could not tell anyone about the things that she lived through and who was raped repeatedly, gave birth to a dead baby in the 7th month of her pregnancy, ‘her secret came out.’

Therefore, instead of analyzing ACC dichotomously, it should be perceived as being scalar, as proposed below.

Table 1 Individuation scale of DOs

			EXAMPLE
↑ Increasing  Decreasing ↓	Type I		
	Singular	<i>N-ACC</i>	<i>kalemi</i>
	Plural	<i>Npl-ACC</i>	<i>kalemleri</i>
	Type II		
Singular	<b>bir N-ACC</b>	<b>bir kalemi</b>	
Plural	<b>bir Npl-ACC</b>	<b>bir kalemleri</b>	
Type III			
Singular	<b>bir N</b>	<b>bir kalem</b>	
Plural	<b>bir Npl</b>	<b>bir kalemler</b>	
Type IV			
Plural	<i>Npl</i>	<i>kalemler</i>	
Singular	<i>N</i>	<i>kalem</i>	

Exploring ACC this way, one can more easily see how a proper name, which should have the highest level of individuation with ACC suffix (because a proper name refers to an entity that is quite distinct from everything around it), can be used in the lowest possible level of individuation (type IV – singular), as shown in the example below.

- (15) Yılmaz, “Kurtlar Vadisi”nde Necati Şaşmaz’ın Sharon  
 Yılmaz Kurtlar Vadisi-LOC Necati Şaşmaz-GEN Sharon  
 Stone’u öp-üş-me-si sahne-si-ne de  
 Stone-ACC kiss-REC-NML-POSS scene-CM-DAT also  
 ilginç bir yorum getir-di.  
 interesting one interpretation bring-PAST  
 Yılmaz “**Sharon Stone** öp-mek için 10 bin dolar  
 Yılmaz Sharon Stone kiss-INF for 10 thousand dollar  
 al-ır-dı-m.[...]” de-di  
 take-AOR-PAST-1SG say-PAST

‘Yılmaz also commented on the scene in which Necati Şaşmaz kisses Sharon Stone. Yılmaz said, “I would request 10 thousand dollars to kiss Sharon Stone [...]”

In (15), a Turkish comedian and actor, Cem Yılmaz, comments on a scene in which another actor kisses Sharon Stone in one of the episodes of a popular Turkish TV series. Unlike many people in the media who consider this to be amazing and hard to achieve (in fact, it is rumored that the producers have paid Sharon Stone a significant amount of money for a brief appearance), Cem Yılmaz thinks differently and says that he would have requested ten thousand dollars to kiss Sharon Stone or actors like her since, he argues, he is a strong

individual and also younger than she is. To express this generality (i.e., to include others and not just Sharon Stone), he does not use ACC with the name ‘Sharon Stone.’ The lack of ACC causes the name to be perceived as being generic and therefore less individuated. The intended and accomplished meaning is ‘both Sharon Stone and actors (or individuals) like her,’ instead of only Sharon Stone as an individual.

Regarding the verb meaning, it is shown (see Table 2) that some verbs, for example, always require ACC bearing DOs (Bolgün 2005).

Table 2 Verbs and the Percentages of ACC-bearing DOs They Take

Verb	Total number of DOs found /analyzed	Number of DOs w/ ACC	% of ACC
<i>andır-</i> ‘resemble’	77	77	100%
<i>vurgula-</i> ‘emphasize’	37	37	100%
<i>göze al-</i> ‘risk; venture’	36	36	100%
<i>ele al-</i> ‘consider’	30	30	100%
<i>ziyaret et-</i> ‘visit’	17	17	100%
<i>suçla-</i> ‘blame’	16	16	100%
<i>tanıt-</i> ‘publicize; introduce’	16	16	100%
<i>yönlendir-</i> ‘guide; direct’	16	16	100%
<i>kastet-</i> ‘mean’	15	15	100%
<i>selamla-</i> ‘greet’	9	9	100%
<i>mümkün kıl-</i> ‘make (sth) possible’	8	8	100%
<i>azarla-</i> ‘scold’	7	7	100%
<i>yala-</i> ‘lick’	7	7	100%
<i>yalanla-</i> ‘deny’	5	5	100%
<i>uyut-</i> ‘(cause sth to) sleep’	3	3	100%
<i>yerle bir et-</i> ‘destroy; level’	3	3	100%
<i>görüş-</i> ‘discuss; consider’	1	1	100%
<i>öngör-</i> ‘foresee’	1	1	100%
<i>paylaş-</i> ‘share’	81	80	98.76%
<i>çöz-</i> ‘solve; untie; undo’	78	75	96.15%
<i>azalt-</i> ‘reduce’	30	28	93.33%
<i>izle-</i> ‘follow; watch’	220	205	93.18%
<i>kaydet-</i> ‘record; state’	12	11	91.66%
<i>planla-</i> ‘plan’	11	10	90.90%
<i>anımsat-</i> ‘remind’	29	26	89.65%
<i>duyumsa-</i> ‘feel; sense’	9	8	88.88%
<i>kes-</i> ‘cut’	56	48	85.71%
<i>kaçır-</i> ‘miss; kidnap’	45	38	84.44%
<i>seyret-</i> ‘watch’	146	123	84.24%
<i>öğren-</i> ‘learn’	182	152	83.51%
<i>yakala-</i> ‘catch’	30	25	83.33%
<i>seç-</i> ‘choose; select; elect’	140	116	82.85%

<i>bil-</i> ‘know’	86	71	82.55%
<i>boşalt-</i> ‘empty’	18	14	77.77%
<i>tüket-</i> ‘consume’	18	14	77.77%
<i>yut-</i> ‘swallow’	31	23	74.19%
<i>sık-</i> ‘squeeze’	58	43	74.13%
<i>söyle-</i> ‘tell; say’	261	191	73.18%
<i>vur-</i> ‘hit’	68	46	67.64%
<i>sor-</i> ‘ask’	82	52	63.41%
<i>barındır-</i> ‘harbor’	22	13	59.09%
<i>duy-</i> ‘hear’	200	114	57%
<i>öde-</i> ‘pay’	42	22	52.38%
<i>tut-</i> ‘hold’	117	60	51.28%
<i>yaz-</i> ‘write’	82	36	43.90%
<i>tak-</i> ‘affix; plug’	23	23	43.39%
<i>patlat-</i> ‘(cause sth to) explode’	20	8	40%
<i>ye-</i> ‘eat’	46	15	32.60%
<i>kazan-</i> ‘win; earn’	180	42	23.33%
<i>yap-</i> ‘make; do’	204	41	19.71%
<i>yetiştir-</i> ‘produce; raise’	42	8	19.04%
<i>üret-</i> ‘produce; generate’	69	12	17.39%
<i>ver-</i> ‘give’	328	57	17.37%
<i>kazandır-</i> ‘cause to win / earn’	30	3	10%

As can clearly be seen in the table above, some verbs (such as *andır-* ‘resemble; remind of’; *vurgula-* ‘to emphasize’) clearly favor ACC-bearing DOs. The meaning of these verbs leads to the requirement that the DO take ACC. The reasons behind such requirement (i.e., why some verbs favor or require ACC-bearing DOs) call for further analysis. However, my initial observations suggest that this is in line with the individuation analysis. For example, with the verb, *andır-*, which means ‘to resemble; to remind of,’ the referent indicated by the DO would have to be individuated (have clear boundaries and seen as a complete, separate entity, some sort of a prototype) to be reminded of.

#### 4. Conclusions

One conclusion that can be made, other than that the situation-theoretic account as presented by Kılıçaslan (2006) cannot explain the use of ACC in Turkish, is that while syntax and semantics play a significant role in explaining the use of the ACC marker on nouns in DO position, they cannot fully do so without incorporating pragmatics into the explanation since in some cases, the same noun in DO position may or may not take the ACC marker within the same or a similar context.

It is interesting to note that when the verbs which have a high ACC-occurrence rate do not take ACC-marked DO, the reason seems to be because they are used with a ‘sense’ different from the one that takes ACC. For example, *izle-* ‘to follow’ can also mean ‘to watch.’ When, for instance, one is following a person or a thing (i.e., an individuated entity or entities) indicated by the DO, the DO takes ACC. However, when used with TV, for instance, as in *televizyon izle-* ‘lit. follow television,’ the person is not actually “following” the television (or the television set); he or she is “watching” it. What is being watched, in that sense, is different from an individu-

ated person or an object, and does not have clear boundaries. In this case, *televizyon* ‘television’ does not take ACC.

Further research is needed to examine verbs (including the ones listed in Table 2) in larger corpora to determine whether or not they follow a certain pattern and favor ACC-bearing DOs. Further research is also needed to analyze verbs that sometimes require ACC and sometimes do not. An analysis of the verbs by their meanings may enable us to make better generalizations regarding the function of ACC. The verbs were not distinguished according to the different meanings or senses they might have. For example, *kazan-* means ‘to win’ as in ‘to win a competition’ but it might also mean ‘to earn’ as in ‘to earn money.’ If the instances of this verb were to be analyzed based on its different senses that it has, the percentages regarding ACC in Table 2 might be different.

Point of view also plays a role in determining whether or not DO will take ACC, especially in situations where a DO with or without ACC is perfectly acceptable.<sup>6</sup> For instance, the following shows ACC emphasizing experiencer’s viewpoint (as well as highlighting the entity referred to by the DO noun). The first boldfaced DO *şey* ‘thing’ has ACC, and it reflects the experiencer (Cansın)’s viewpoint.

- (16) Cansın, 27 Haziran akşam-ı daha önce hiç  
Cansın 27 June evening-POSS more before never
- yap-ma-dığı bir şey-i yap-tı. İsveç’ten gel-en  
do-NEG-REL one thing-ACC do-PAST Sweden-ABL come-REL
- arkadaş-ı Umut Kanyılmaz’la site-de küçük bir  
friend-POSS Umut Kanyılmaz-COM site-LOC small one
- tur at-mak için izin iste-me-den  
tour throw-INF for permission request-NEG-ABL
- baba-sı-nın otomobil-i-nin anahtar-lar-ı-nı  
father-POSS-GEN car-POSS-GEN key-PL-POSS-ACC
- al-ıp direksiyon-a otur-du.  
take-PTCL steering.wheel-DAT sit-PAST
- Direksiyon-a geç-ince plan-lar deęiş-ti,  
steering.wheel-DAT pass-ADV plan-PL change-PAST
- site-de at-ıl-acak küçük tur birden büyü-dü,  
site-LOC throw-PASS-FUT small tour suddenly grow-PAST
- Londra Asfalt-ı’na taş-tı...  
Londra Road-POSS-DAT overflow-PAST

‘On June 27th, Cansın did something that he had not done before. He took the keys to his father’s car without permission and sat at the steering wheel to take a tour around the neighborhood with his friend Umut Kanyılmaz, who came from Sweden. When he sat at the steering wheel, the plans changed; the small tour that was to be taken around the neighborhood suddenly grew big and spilled over to the Londra Road...’

Example (16) is from a news item about a young boy named Cansın who dies as a result of something that he did for the first time. That ‘something’ is introduced into the discourse using *bir + NP-ACC* structure. The ACC in the above is optional, but its use, I believe, is intended to make us look at the events through the experiencer’s viewpoint.

<sup>6</sup> See Epstein (1994, 1998, 2001, and 2002) who argues that articles (in languages such as English and French) are essentially ‘multi-functional,’ and that other than their referential function, they also have an expressive function.

### Data Sources

- (1) Taylan & Zimmer (1994)
- (2) Taylan & Zimmer (1994)
- (3) Taylan & Zimmer (1994)
- (4) Taylan & Zimmer (1994)
- (5) Kılıçaslan (2006)
- (6) Kılıçaslan (2006)
- (7) Kılıçaslan (2006)
- (8) Sazak, Derya. (2003, May 19). Sohbet Odası. *Milliyet*. Retrieved August 15, 2014, from <http://www.milliyet.com/2003/05/19/siyaset/asiy.html>
- (9) METU Turkish Corpus. See *Say et al.* (2002).
- (10) METU Turkish Corpus. See *Say et al.* (2002).
- (11) Kahraman, Sibel. (2003, January 29). Karne öğrenciye mi veliye mi verilsin. *Milliyet*. Retrieved January 15, 2014, from <http://www.milliyet.com.tr/2003/01/29/guncel/gun08.html>
- (12) Altuntaş, Birsen. (2003, June 30). Yarışmacıyı yanlış yönlendirmedim. *Milliyet*. Retrieved August 14, 2014, from <http://www.milliyet.com/2003/06/30/magazin/mag02.html>
- (13) Çakırözer, Utku. (2013, August 07). Asıl darbe mağduru benim. *Cumhuriyet*. Retrieved August 24, 2014. [http://www.cumhuriyet.com.tr/koseyazisi/438334/Asil\\_Darbe\\_Magduru\\_Benim.html](http://www.cumhuriyet.com.tr/koseyazisi/438334/Asil_Darbe_Magduru_Benim.html)
- (14) Cebe, Özgür. (2003, January 29). Töre yine iki kurşunla işledi. *Milliyet*. Retrieved August 24, 2014, from <http://www.milliyet.com/2003/01/29/yasam/yas06.html>
- (15) Bağcıbaşı, Bora. (2005, December 23). On bin dolar almadan Sharon Stone öpmezdim. *Milliyet*. Retrieved August 16, 2014, from <http://www.milliyet.com.tr/2005/12/23/son/sonyas17.html>
- (16) Korap, Elif. (2003, June 30). Gençlik hatası, bir aileyi yasa boğdu. *Milliyet*. Retrieved August 17, 2014, from <http://www.milliyet.com/2003/06/30/guncel/agun.html>

### References

- Aissen, J. (2003). Differential Object Marking: Iconicity vs. Economy. *Natural Language & Linguistic Theory*, 21:3, 435-483.
- Aygen-Tosun, G. (1999). *Specificity and Subject-Object Positions/Scope Interactions in Turkish*. Presentation at First Manchester Conference on Turkic Languages.
- Banguoğlu, T. (2004). *Türkçenin Grameri* [Grammar of Turkish] (7th printing). Ankara: Türk Dil Tarih Kurumu. 1st printing: 1959.
- Bolgün, M. A. (2005). Accusative marking in Turkish (Doctoral dissertation). ProQuest. (AAT 3209425)
- Crisma, P. (1997). *L'articolo nella prosa inglese antica e la teoria degli articoli nulli*. Doctoral Diss. Università di Padova.
- Crowley, W. I. (1938). *A modern Turkish Grammar*. Harrogate, Tenn.: Pangloss Publications.
- Dede, M. (1986). Definiteness and referentiality in Turkish verbal sentences. In D. I. Slobin & K. Zimmer (Eds.) *Studies in Turkish Linguistics*. Amsterdam / Philadelphia: John Benjamins Publishing Company.

- Enç, M. (1991). The semantics of specificity. *Linguistic Inquiry*, 22, 1-25.
- Epstein, R. (1994). Variation and definiteness. In Katharine Beals et al. (eds.) *Papers from the 30th Regional Meeting of the Chicago Linguistic Society, vol. 2: The Parasession on Variation in Linguistic Theory*. Chicago: Chicago Linguistic Society, 61-75.
- Epstein, R. (1998). Reference and definite referring expressions. *Pragmatics & Cognition*, 6 (1/2): 189-207.
- Epstein, R. (2001). The Meaning of Definite Articles in Cross-Linguistic Perspective. In Enikő Németh T. (ed.) *Pragmatics in 2000: Selected Papers from the 7th International Pragmatics Conference*. Volume 2. Antwerp, Belgium. 2001.
- Epstein, R. (2002). The definite article, accessibility, and the construction of discourse referents. *Cognitive Linguistics*, 12(4): 333-378.
- Ergin, M. (1962). *Türk Dil Bilgisi*. İstanbul Üniversitesi Edebiyat Fakültesi Yayınları.
- Erguvanlı, E. E. (1984). *The Function of Word Order in Turkish Grammar*. University of California Press.
- Erguvanlı-Taylan, E. (1987). The role of semantic features in Turkish word order. *Folia Linguistica*, 21, 215-227.
- Gencan, T. N. (2001). *Dilbilgisi [Grammar]*. İstanbul: Ayraç Yayınevi.
- von Heusinger, Klaus. (2002). Specificity and Definiteness in Sentence and Discourse Structure. *Journal of Semantics*, 19, 1-30.
- Hopper, P. J. and Thompson, S A. (1980). Transitivity in Grammar and Discourse. *Language*, 56(2), 251-299.
- Johanson, L. (2006). "Two approaches to specificity". In *Case, Valency and Transitivity*, Kulikov, Leonid, Andrej Malchukov and Peter de Swart (eds.), 225–247.
- Kılıçaslan, Y. (2006). A situation-theoretic approach to case marking semantics in Turkish. *Lingua*, 116, 112-144.
- Kornfilt, J. & von Heusinger, K. (2008). Specificity and partitivity in some Altaic languages. In R. Vermeulen (ed.), *MIT Working Papers in Linguistics. Proceedings of the 5th Workshop on Formal Altaic Linguistics (WAFLL 5)*. Cambridge, MA: MIT.
- Kornfilt, J. (1997). *Turkish*. London; New York: Routledge.
- Lewis, G. L. (1967). *Turkish Grammar*. Oxford: Clarendon Press.
- Lewis, G. L. (2000). *Turkish Grammar*. Oxford: Oxford University Press.
- Mowle, A. C. (1934). *The new Turkish: An elementary grammar, vocabulary and phrasebook of the Turkish Language in the new Latin characters*. London: K. Paul, Trench, Trubner, & co., ltd.
- Nakipoğlu, M. (2009). The semantics of the Turkish accusative marked definite and the relation between prosodic structure and information structure. *Lingua*. Vol. 119, No. 9, pp. 1253-1280.
- Nemeth, G. (1962). *Turkish Grammar*. 'S-Gravenhage: Mouton & Co. English Adaptation of the German Original by T. Halasi-Kun, Ph.D.
- Özge, U. (2011). Turkish indefinites and accusative marking. In Simpson, A. editor, *MIT Working Papers in Linguistics #62: Proceedings of the 7th Workshop on Altaic Formal Linguistics (WAFLL7)*, 253-267, Cambridge, MA: MIT Press.
- Say, B., Zeyrek, D., Oflazer, K. & Özge, U. (2002). Development of a Corpus and a Treebank for Present-day Written Turkish." In K. İmer & G. Doğan (Eds.), *Current Research in Turkish Linguistics* (pp.183-192). Proceedings of 11th International Conference on Turkish Linguistics. Eastern Mediterranean University, Northern Cyprus.
- Seaman, W. [a.k.a. Gulielmo Seaman]. (1670). *Grammatica linguæ Turcicæ in quinque partes distributa*. Oxoniae [Oxford]. Excudebat Hen: Hall ... Prostant apud Edvardum Millington...
- Sebüktekin, H. I. (1971). *Turkish-English Contrastive Analysis: Turkish Morphology and Corresponding English Structures*. The Hague: Mouton.
- Swift, L. B. (1963). *A Reference Grammar of Modern Turkish*. Indiana University, Bloomington; The Hague, The Netherlands: Mouton & Co.
- Taylan, E. E. & Zimmer, K. (1994). Case Marking in Turkish Indefinite Object Constructions. In S.Gahl, A. Dolbey, & C. Johnson (Eds.) *Proceedings of the twentieth annual meeting of the Berkeley Linguistics Society* (February 18-21, 1994.). General Session. Berkeley Linguistics Society.
- Underhill, R. (1976). *Turkish Grammar*. Cambridge, Mass.: The MIT Press.



# Adverbial Clauses in Modern Turkish

Duygu Özge Demir  
Hacettepe University, Turkey

## Abstract

In this study, the topic of adverbial clauses within modern Turkish, the mostly spoken and processed language among Turkic language family in Altaic languages will be covered. Turkish type subordinate clauses has always been an attractive topic for western researchers and broad literature on these exists in the western world. Various evaluations of adverbial clauses are present in this literature. For example, in Altaic languages, the word order in an ordinary sentence is as subject-object-verb and subordinate clause precedes the main clause. However, during my PhD thesis, novel tendencies and constructions differing from these general observations are also determined. In this regard, word order of adverbial clauses in modern Turkish and marking of adverbial clauses as well as their position relative to superordinate clause will be discussed extensively in this study.

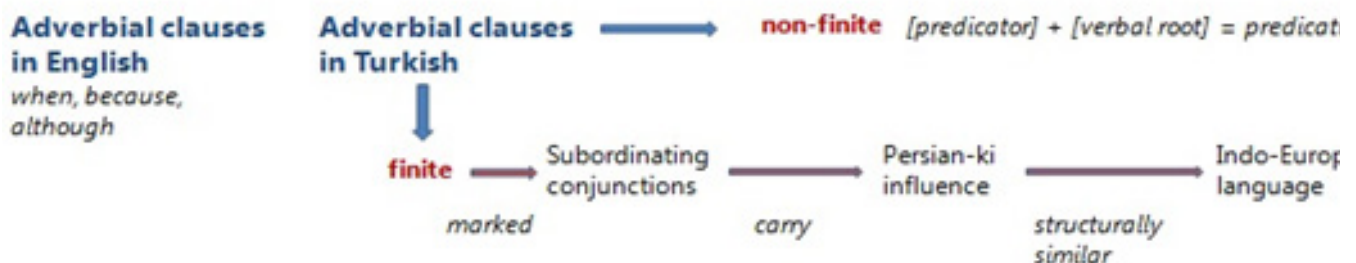
Keywords: Altaic languages, Turkish, syntax, adverbial clauses, word order, marking of adverbial clauses' position

## 1. Introduction

Turkish converb clauses correspond to adverbial clauses in European type languages in terms of function. However in European type languages, these are made by lexical elements, whereas in Turkish type languages converbs are made by morphological elements or complex units or syntactic means. Converbs are especially common in many non-European languages. Among these Turkic languages have attracted especial attention and converbs play a central role in their syntax. Furthermore Johanson states that all the more as Turkic languages –like Mongolian, Tungusic, Korean, Japanese, Dravidian etc. – exhibit particularly elaborate converb systems (1995: 313).

Example of this syntactic differences between Indo-European and non-European languages:

- (1) I fell in love when I saw you  
*when*: subordinative conjunction
- (2) Ali eve gelinece uyudu (Ali slept when he came home).  
*-ince*: suffixed subjunctor  
*gelinece* [predication] < *gel-* [verbal root] + *-ince* [predicator]



So far, the syntactic difference in forming adverbial clauses between Indo-European and Altaic type languages is simply pointed out. Here non-finite adverbial clauses were discussed only. The main topic of this study is to analyze the word order of adverbial clauses in modern Turkish and marking of adverbial clauses'

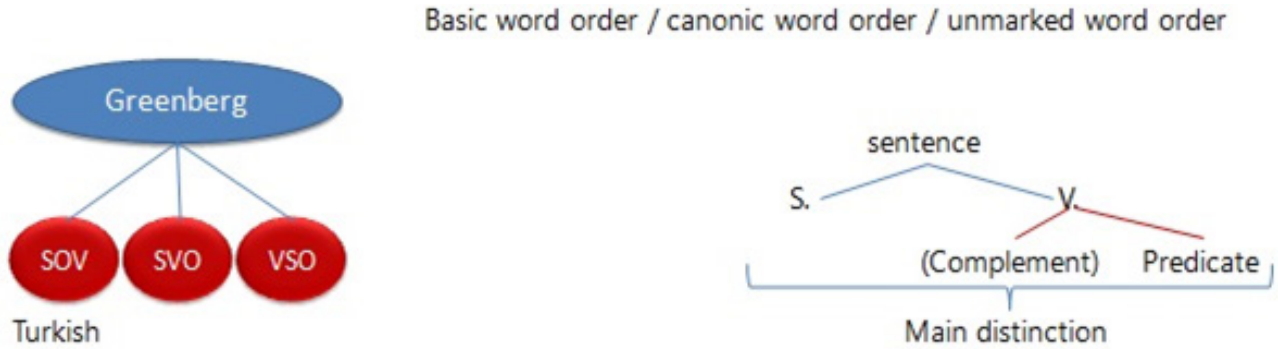
position relative to superordinate clause. There are three types of adverbial clauses in Turkish:

- Ones have certain morphology: formed with converbs: *-(I)p*, *-IncA*, *-ArAk*...
- Participle nominalization: participles, verbal nouns are marked with case or followed by postpositions: *-DIK+I + için*, *-AcAk+I + zaman*, *-DIk+I+n+dA*; *-AcAk+I+n+dAn*
- Subordination of finite clauses: *diye*, *ki*, *madem ki*, *-DI mI*, *clitic -dA*...

Since this topic is very comprehensive, I will only analyze adverb clauses formed by main converbs have certain morphology such as *-(I)p*, *-ArAk*.

In sample part of this investigation, a corpus involving a great number of different fields and types of written and spoken examples in modern Turkish was used. (<http://derlem.mersin.edu.tr/ikilitekrar/>).

## 2. Word Order of Adverbial Clauses in Modern Turkish



However, Turkish -especially at spoken language- deviates from this SOV order. This deviation is explained as “pragmatic priority at discourse”. Instead of constituents such as the subject or the object, pragmatic discourse constituents like topic/given (thema) and comment/new (rhema) play role in shaping the word order (Abraham 1994: 235). Accordingly, speaker can replace the elements with respect to the value of elements.

In Turkish, word order is flexible and not rigid as Turkish has an encoding in which words can move freely. However, this does not mean “you order words as you wish”. In Turkish, changing in basic word order should be functional. Erguvanlı Taylan presents a table in which there are three major syntactic positions, each of that has a corresponding pragmatic function (1984: 72):

Syntactic position	Pragmatic position
S- initial	topic
Immediately preverbal	focus
Post-predicate	backgrounding

Topic (T)= What is the sentence about? T is an information which has already existed in the discourse.

Focus (F)= F is a new information in the sentence. This new information generally is the most important information which should be focused on.

Background (B)= Compared to T and F, B is less important information. B is as a part of the context which is the rest of utterance.

### 2.1. SOV Word Order Pattern in Adverbial Clauses

Johanson states that adverbial clauses often start like a finite sentence, do not exhibit a special word order pattern and will not require any preparation that might cause difficulty while speaking (1991: 103). According to this, *topic* position is located in the very beginning of the sentence in Turkish. *Topic* position conveys major issue talked about. The *topic* of sentence is generally the subject of the sentence.

- (1) *Ben sen-i düşün-üp,* seni konuşmaktan başka bir şey yapamıyorum ciddi.  
I you-ACC think-IP.CONV

### 2.2. VOS Word Order Pattern in Adverbial Clauses

In such adverbial clauses, basic word order SOV changes according to pragmatic priority of discourse and sentence constituents become as V+O+S. In this word order pattern predicate is located “topic” position not at the end of the sentence:

- (2) *Anahtar paspasın altında durur, aç-ıp kapı-yı, girerler.*  
open-IP.CONV door-ACC
- (3) *Teknem var diye yur-up kafa-yı yatmayacaksın, her gün balığa çıkacaksın.*  
hit-IP.CONV head-ACC

### 2.3. SVO Word Order Pattern in Adverbial Clauses

Since T is generally determined depending on discourse, in the adverbial clauses below T is not existent, F and B are existent. In the examples below the elements which precedes adverbial clauses’ predicate are in focus position, because they convey the most important information. In addition, adverbial clause’s objects are also “background” elements, because they precede the adverbial clause’s predicate.

- (4) *Ben al-ıp baş-ım-ı gideceğim bir gün.*  
I take-IP CONV head-POSS1SG-ACC
- (5) *Aynanın kırıkları gözüne kaçan insan-lar ağla-yarak kendi-leri-ni kurtarabilirler.*  
human-PL cry-ARAK.CONV self-POSS3PL-ACC

### 2.4. OSV Word Order Pattern in Adverbial Clauses

In this word order pattern, object is located in *topic* position while the subject is located in *focus* position. Here, constituents in *focus* position are in the foreground:

- (6) *Kitab-ı, başka müşteri-ler gör-üp elimizden alacaklar, kapacaklarmış gibi, acele, veznenin yanındaki paketleme masasına koşturuyor, acele, parasını ödüyor, kitabevinden kaçırıyorduk adeta.*
- (7) *Parmak-ları-na dola-dığ-ı saç-ları-nı ben çek-ip alsaydım bilirdim; o mu değil mi?*  
finger-POSS3PL-DAT wind-DİK.PTCP-ACC hair-POSS3PL-ACC I pull-IP.CONV

### 2.5. OVS Word Order Pattern in Adverbial Clauses

As for this kind of adverbial clauses, based on pragmatic priority of the discourse, the object is located in the beginning of the clause which is sequentially followed by predicate and subject. Objects that immediately precede the predicate in the examples below convey the new information and are in the position of *focus*.

- (8) *Bazen öküz-ler de ağla-r di-yerek bay öküz, hem gözyaşı döktü*  
 sometimes ox-PL DA.PART cry-AOR say-ARAK.Conv mr. ox  
*yolun kıyıcığında, hem de sinirinden fısıldadı durdu.*
- (9) *Bu-nu öğren-ince Bedriye, Zilha 'yı yolladı.*  
 this-ACC learn-INCA.CONV Bedriye

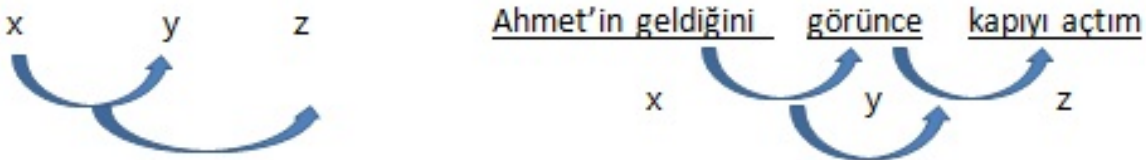
### 2.6. VSO Word Order Pattern in Adverbial Clauses

Contrary to the basic word order, predicate is in the *topic* position. According to the analyzed corpus, this word order pattern is the least frequently used one. As seen in example (10), this pattern is mainly used in poetry.

- (10) *Gel-ince sen yan-ım-a, deęişir birden ihtiyar dünya!*  
 come-INCA.CONV. you near-POSS1SG-DAT

## 3. Marking of adverbial clauses' position

One of the major characteristics of Turkish is that it has a left-branching syntax, in other words it has *rectum-regens norm*. "According to this rule subordinate constituent (rectum) precede the main constituent (regens) in terms of syntactic" (Johansen 1992, 254-255). For instance, if sentence extends, new constituents are put on left side of regens.



Kornfilt also states that in an unmarked word order, all types of adverbial clauses are placed at the beginning of the main sentence. However, due to the general flexibility of word order in Turkish, the adverbial clause can surface in any position (2000: 68). This is again related to *communication perspective*. As for analyzed corpus adverbial clauses are positioned three ways: initial position, final position, medial position adverbial clauses.

### 3.1. Initial position

In some languages are stated that a characteristic of adverbial subordinate clauses is their position. In Mandarin, Ethiopian Semitic, Turkish and many other languages, adverbial clauses must precede the main clause (Thompson and Longacre 1985: 174). However this is not a norm, which, adverbial clauses can be positioned by different ways as shown below.

Considering adverbial clauses' position relative to main clause there occurs three different phenomena with regard to word order and information structure: *topic*, *focus* and *background*. Namely, adverbial clause located in the initial position relative to main clause is a *topic* element since it presents on spoken information.

- (11) *Bekir sırtında gübre taşıyıp tarlayı beslerdi.*

- (12) *Oysa yataktan gökyüzüne bakınca, havayı bulutlandı sanmışım.*

### 3.2. Final position

Adverbial clauses take place after the main clause predicate. For this reason, adverbial clauses are *background* element. Therefore the important information is here beared by the main clause while the adverbial clause implies the least important information based on “pragmatic priority of discourse” norm. Aydemir also states that adverbial clauses founded in final position become defocused with regards to *communication perspective* (2010: 15).

- (13) *Sıktık birbirimizi, sarılıp.*
- (14) *Bu yazıyla ilgili bir haberi sen getirdin: Gazetenin Genel Yayın Yönetmeni Cevat Fehmi Başkut şaşırmış yazıyı alınca.*
- (15) *“Hoş geldin,” diyor bana bakarak.*

### 3.3. Medial position

Some adverbial clauses are found embedded into the main clause in the medial position. Such adverbial clauses entirely fuse with main clause. Since they precede main clause predicate, they generally convey the newest and most important information and are located as *focus* element.

- (16) *Onlar burada kendi köşelerinde “dilenci hayatı sürüp” dilenci edebiyatı yapıyor.*
- (17) *İçinden, o palayı alarak bu yabancı adama saldırmak geldi.*
- (18) *Amma yaptın Arif Ağa, biz olalım olmayalım, canın çekince girip istediğin kadar yiyeceksin.*

## 4. Results

- In Turkish, there is no specific word order pattern in adverbial clauses either, similar to main clauses.
- According to the pragmatic priority of discourse, sentence constituents can order flexibly. Johanson explain this as “communication perspective” (Mitteilungs-perspektive) and according to this, the speaker prefers an order which is formed by Thema in the beginning of the talk and Rhema at the course of the talk (1977:230). The important point here is what the speaker would like to give and how and to what extent it is given (Aydemir 2010: 10).
- Flexibility of Turkish word order can be expressed mathematically. Different orders may occur by the number of constituents. When are looked at the example above, there are six different orders:  $3! = 3 \times 2 = 6$
- The position of adverbial clauses are not fixed in Turkish. They also may appear in *initial*, *medial* or *final* position relative to main clause.
- Adverbial clauses may also be *topic*, *focus* or *background* element of the main clause.

## References

- Abraham, W. (1994). “Fokusgrammatik und Indefinitheit”, B. Haftka (ed.) *Was determiniert Wortstellungsvariation? Studien zu einem Interaktionsfeld von Grammatik, Pragmatik und Sprachtypologie*. Opladen: Westdeutscher Verlag, pp. 235-245.

- Aydemir, İ. (2010). "Tuvacada Zarf-fiil Cümlelerinin Söz Dizimine Dair". *Bellekten* 2010/2:7-18.
- Dik, S. (1997). *The Theory of Functional Grammar. Part 1: The Structure of the Clause*. (edited by Kees Hengeveld). Berlin, New York: Mouton de Gruyter.
- Greenberg, J. (1966). "Universals of Language". Cambridge: MIT Press.. pp. 73–113.
- Johanson, L. (1977). "Bestimmtheit und Mitteilungsperspektive im türkischen Satz". *Zeitschrift der Deutschen Morgenländischen Gesellschaft*, Suppl. III: 2. 1186-1203.
- Johanson, L. (1991). "Zur Typologie türkischer Gerundialsegmente", *Türk Dilleri Araştırmaları* 1991. 98-110.
- Johanson, L. (1992). *Strukturelle und soziale Faktoren in türkischen Sprachkontakten*. Frankfurt. (=Sitzungsberichte der Wissenschaftlichen Gesellschaft an der J. W. Goethe Universität Frankfurt am Main.)
- Johanson, L. (1995). "On Turkic converb clauses". In Haspelmath, Martin & König, Ekkehard. *Converbs in cross-linguistic perspective. Structure and meaning of adverbial verb forms – adverbial participles, gerunds. Empirical approaches to language typology* 13. Berlin & New York: Mouton de Gruyter. 313-347.
- Kornfilt, J. (2000). *Turkish*. London: Routledge.
- Taylan, E. E. (1984). *The function of word order in Turkish grammar*. University of California Press
- Thompson, S.A. and R.E. Longacre (1985) 'Adverbial clauses', in T.Shopen (ed.), *Language typology and syntactic description, vol. II. Complex Constructions*. Cambridge: Cambridge University Press, pp.171-234.



## Discourse Markers in Uzbek

Sharifa Djurabaeva  
Ball State University

### Abstract

This study explores the use of the Uzbek discourse markers *haligi* ‘like,’ *nima* ‘you know’, *shu* ‘well’ and their role in Uzbek discourse. To examine these discourse markers, conversations with individuals from Andijan, a region in southeast Uzbekistan, were recorded and transcribed. A total of eight individuals, four men and four women, participated in this study. The conversations were about food, school, video games, everyday life, and other interesting stories from the lives of participants. The goal of the study was to examine the function of discourse markers among the speakers of the region. The results showed that these discourse markers served many functions, such as marking the new information and placing focus on it, hedging, clarifying the upcoming topic in the discourse, and referring to the topic that was mentioned earlier in the discourse.

### Introduction

In this study I would like to explore some of the discourse markers used in Uzbek. Discourse markers, as Muller (2005) points out, “contribute to the pragmatic meaning of utterance and play important role in the pragmatic competence of the speaker” (p. 1). Discourse markers have been studied and labeled in different ways in research. As Fraser (1999) states, they were labeled as discourse markers, operators, discourse connectives, pragmatic connectives, sentence connectives, and cue phrases. In addition to these expressions, they were also named as conversation fillers, “verbal garbage,” Schourup (1985: 94) mentioned in Jean, Tree & Schrock (2002) and labeled as “slova-paraziti” (words-parasite) (Natalievich, 2010) in Russian that does not have any meaning in the sentence and distracts the listener.

Even though there are some articles about these discourse markers in Uzbek or Russian, there is little research published in English about Uzbek discourse markers. Therefore, the primary goal of this study is to analyze some of the most frequently used discourse markers in Uzbek and see if they function in the same way as in other languages. To study these discourse markers, I decided to limit my data sample to focus only on conversations from speakers of certain region, Andijan, because the same regional differences in the use and function of discourse markers.

### *Discourse Markers*

Fraser (1999) defines discourse markers as “a class of lexical expressions drawn primarily from the syntactic classes of conjunctions, adverbs and prepositional phrases” (p. 231). He states that they signal a relationship between the two segments where one segment introduces the second one (p.231). The definition of discourse markers depends on many factors, such as whether their grammatical function or pragmatic function is being analyzed in the discourse. As mentioned above, discourse markers are considered to be markers, connectors, particles, and others. Schiffrin (2001) states that expressions such as *like well*, *but*, *oh*, and *y’know* can be observed functioning in many domains such as cognitive, expressive, social, and textual domains. (p.54). Therefore, some studies consider them to be unnecessary elements in the discourse. But by exploring the discourse

markers *like* in English, Underhill (1988) states, “nonstandard *like* is neither random nor mindless.” (p.234) It functions as a marker of new information and focus, where he gives the thorough explanation of the functions of the *like*. There were many similar studies of discourse markers in various languages; however, this was not the case with Uzbek as there was a lack of studies examining of discourse markers in the speech of Uzbek speakers. This study will shed some light on discourse markers in Uzbek.

### *Uzbek*

The Uzbek language, currently the official language of Uzbekistan, belongs to the Qarluq branch of the family of Turkic languages. It has been influenced by Persian, Arabic and Russian. Uzbek is also spoken in Tajikistan, Kazakhstan, Kirgizstan, and Russia. The speech patterns of each region’s speakers differ from each other at some degree phonetically, lexically and syntactically. According to information from the Center for Turkic and Iranian Lexicography and Dialectology (2014, August 2) at Indiana University, Bloomington, Reshetov (1959) places the speakers of Andijan region into the Karluk-Chigil-Uyghur dialect Group retrieved from (<http://www.indiana.edu/~ctild/Uzbek/DialectClassification>), specifically in the Ferghana group.

### **Participants**

In the analyses of the discourse markers in Andijan dialect of Uzbek, eight volunteers, four men and four women have participated. The participants are between 20-40 years old, mainly from the Andijan region, including the city of Andijan and surrounding districts. All of the participants’ first languages are Uzbek spoken in Andijan. Their second language is Russian and the third language is English. All of the participants are educated professionals: three have high school diplomas, two have bachelor’s degrees, and two have and Master’s degrees.

### **Methodology and Data Analyses**

To collect the data, an announcement was sent through Facebook to approximately 100 people. People who showed an interest to participate in the project were selected to have a conversation with each other or with the researcher over the phone or Skype. The conversations were recorded on Soundcloud.com and transcribed. The conversation themes consisted of the topics such as attending an American universities, food, teaching, giving birth, learning, banking, and video games. The purpose of having different topics was to see the frequency of usage of discourse markers in the speech of these participants. The results suggest demonstrate that the frequency of the discourse markers used is related to the topic. The most frequently used discourse marker was *shu*, ‘well’ then *haligi*, ‘you know’ and the last one is *nima* ‘you know.’ However, there might be other factors that contribute to the usage of these discourse markers.

### **The usage of haligi ‘that’ or ‘like’ in the discourse**

The original function of the *haligi* ‘that’ ‘that one’ is to point out something in the sentence. For example, *Haligi odam bizinikiga keldi*, ‘*That person came to our house.*’ But in the discourse of the participant’s speech it functioned differently, as “a marker of new information and focus” as stated by Underhill, (1988) and functioned similar to English ‘*like*.’ (p.238) There are some examples of marking new information with the discourse marker *haligi*. English translations of these conversations might be interpreted slightly differently without the full context. Uzbek texts in these analyses represent only the spoken version of it. Below are three examples of haligi being used to introduce new information in the discourse.

(1)  
Menga .....bo'gan, shunaqa ediki, **haligi** berda Kolumbiyada, har bitta **aligi** Ivy League degan **haligi** sakkizta ta universitet bor. O'shalardi har bittasini spetsializatsiyasi bor. Biznes sohasi bo'yicha, misoli uchun, Wharton Business School ham judayam kuchli. Penselvaniyada... U moliya sohasida judayam kuchli. Harvard management bo'yicha kuchli.

*To me ... happened, it was like that, here **like** in Columbia, each **like** Ivy League schools, there are **like** eight universities exist. Each of them has their own specializations. For example, in business, the Wharton Business School is very strong. It Pennsylvania... It is very strong in the field of finance. Harvard is strong in the filed of management.*

(2)  
Yani, Derivatives degani, aaa o'zini anaqasini, qimmatini, **haligi**, it derives its value from something else.

*In other words, derivatives means aaa its that one, value, **like**, it derives its value from something else.*

(3)  
U erdan ip chiqqandan keyin, uni **aligi**, material to'qiydigan tsexta junatdim.

*After the yarn was produced, that one, **like**, I sent to the department that weaves a fabric.*

So, in (1) it is used to introduce *Columbia, Ivy League* and *eight universities*, the information that was not previously mentioned in the conversation. The same function can be seen in (2) where the new information, 'the process,' is introduced with **haligi** by telling the listener about the process of derivation. In (3) the speaker is introducing the new topic of the department 'that weaves the fabric.'

In addition to introducing the new information, these discourse markers function as the markers of focused information. Below are two examples where **haligi** marks the focused information.

(4)  
A: Uuuuuu zurakanu! Man bopakanda u!  
B: Aaaa kein divanlari bor. **Aliyi**, nima alohida xonachalari bor. Aaa yaxshi tomoni aliyi study roomga o'xshagan. O'zizga kirvolib, hech kim xalaqit bermaydi.

*A: Ohhh cool! That's exactly what I like!*

*B: Aaa then there are couches. It is that **like**, there are separate rooms. Aaa good side is that there are rooms similar to like study rooms. You can enter and nobody bothers you.*

(5)  
Misol uchun Agrobank bu faqat paxta, bug'doy bilan ishlaydi. Keyin **aliy**, aaa paxtani puli Agro bankka tushadi. Agro bankda aylanadi.

*For example, Agro bank works only with cotton and wheat. Then **like**, aaa the money paid for cotton goes to Agro bank. It evolves in the Agro bank.*

In example (4) the discourse marker is marking the focused information 'room' in two languages: Uzbek and English: **xona**-Uzbek and study room-English. In example (5) the focus is on *cotton*. The '**paxta**' *cotton* is mentioned in the previous sentence and focused again in the next sentence. So we can see that in above-mentioned examples, speakers either mark the new information or focuses on previously mentioned information by using the discourse marker **haligi**.

### The usage of discourse marker nima 'you know,' nima edi (deydi)? What was that?

#### Hedging

Another discourse marker **nima** 'you know' is frequently used in hedging. Originally, **nima** 'what?' is used in forming a question. For example, *Bu nima (dir)?* 'What is it?' *Nima qilasan?* 'What will you do?' But in discourse they have different functionality roles, such as hedging or softening the information. Often times,

when the person is too shy to say some information directly, they use *nima* a lot in the discourse. It is worth noting that hedging with *nima* occurred more among women speakers, but more data analyses are needed to prove any difference in frequency between male and female speakers. That said, most of the examples provided here come from women's speech.

(6)

Woman teacher is explaining about her participation in the contest.

A: A...

B: Shu 5, 6 IELTSs ga to'g'ri kelar ekan. Shu bo'yicha imtihon qildi. Biz topshirib, xudo hohlasi muvaffaqiyatli o'tiv oldik shundan. Hali aaa... *nima* sertifikatlar berilgan yuq. Shu kamchilik boptiy-u, lekin shularni ichida bormiz xudoga shukur.

A: Hm....

B: It is 5, 6 according to IELTS's score. They examined us on it. We have taken it and thanks to God we successfully passed it. Still hmm... *you know*, the certificates haven't been given yet to us. It was the shortcoming, but thanks to God we are among the participants.

(7)

A: Aha. Aha

B: Uydagilaram qullab-quvvatlab ancha tajribamizni oshirishga harakat qildik. Bilmadik, baribir, *nima* endi kamchiliklar bo'gan bo'sa kerak shu bumadi endi. Yana bitta programmalar borakan Harbert Humferi deb yaqinda keldi. Shuni DDDon o'zi tavsiya qip qoldi.

A: Aaa. Aha

B: We tried to gain an experience and our family members supported us. We didn't know, anyway, *you knooooow* there must have been some shortcomings, it didn't work out. There was another program named Hubert Humphery, which we heard. D has recommended it for us himself.

(8)

A: Ey qomayapsizmi, yana uch oyga?

B: Yooo oshani, *nimaga*, cholimga aytsam, endi ashu etmay turuvdi dedi.

A: Are you going to stay for three more months?

B: Noooo, I told that to *you know*, my husband. But he refused saying that's not what he was expecting.

(9)

A: Nega?

B: Endi... esayiz dedim, haliyi. Ichim quruq ketayotganimda. Aligi, qornimdaginiyam *nima* qildim. Bo'shanib ketdiim. Egim kemadi, chunki endi andaqada... Ichiz toza bo'sa yaxshi-da. Kuchanasiz.

A: Why?

B: Well..., if eat I thought, like. My stomach was empty when I was leaving. Like, the one that was in my stomach, *you know*, I did that. Emptied it. (She went to the restroom). I didn't want to eat because... well, it is like that... If your stomach is clean it is good. You need to push.

(10)

Keyin aytdi, *nima* qilamiz... Brinchi dediiiiiii. Hhhhhhhh anaqani matkani ochadigan ukol qilamiz dedi.

Then she said we would *you know* do... First they said. Hmmm like that... we will give a shot to open up the womb, they said.

These examples demonstrate how *nima* is used for hedging, similar to the discourse marker '*you know*' in English. Women feel uncomfortable discussing or saying certain words or information directly in the discourse; therefore, to soften the conversation they use this discourse marker. In examples (6) and (7) the participant is hesitating to criticize that she still didn't get a certificate and the shortcomings that she encountered. In example (8) the woman is shy to say 'my husband', so instead she is using the hedging marker '*nima*' first and then explains what she means. In examples (9) and (10) the woman is kind of embarrassed to talk about her conditions before birth, therefore she is using *nima qildim* '*you know I did that*' which means that she went to the bathroom to clean her stomach, but culturally Uzbek women do not talk about this topic explicitly. She is also not saying directly that her womb will be opened, but using the discourse marker *nima qilamiz* '*you know*, we

would do that' and pauses slightly in the discourse or hesitates about the formulation of the upcoming information in the speech. So, *nima* 'you know' is used a lot in hedging, and based on the data collected here, is used more frequently by women.

### Clarification

In addition to hedging, the question form *nima* 'what' also functions as a clarifying question in the conversation or as an engagement question in the discourse. Speakers are using the question *nima edi (deydi)?* 'What was it?' or 'what is said?' to clarify the topic or to engage the interlocutor for the conversation.

(11)

B: U endi prosesni tushintirib beradigan bo'sam, ko'proq aql bilan ish ko'rilishk kerak. Aligi, bu ko'proq aligi *nima desa bo'ladi?* Aa iqtisodchilarga rosa qo'l keladi bu.

*B: Well, if I explained the process, one needs to work wisely. Like, this one like, what was it? (What can I say?) Hmm it is really good for economists.*

(12)

A: Ohhh zurku!

B: Mana aligi, *nima deydi?* Sizga xalaqit bermoqchi bo'gan narsayam pul bo'p qola-dida sizga.

*A: Ohhh this is great!*

*B: Here, like, what do we say? The thing that can be an obstacle on your way becomes a profit for you.*

(13)

A: Hhhhh, pupovinasini. Kindigini.

B: Hhhhhaaa kesishdi hohlaysami duvdi? Qanday kesaman deudi, kes, kes, davay deb, andaqa, *nima deydi?* Kuchli, bud silnee, deydi-yu.

A: Aha

B: Ina. Qo'rqma, qo'rqma deb, keshdirishdi keyin o'ziga.

*A: Hmm, belly button. Belly button.*

*B: Yeeees they asked if he wanted to cut it. And he asked how he would cut it, they kept telling 'cut, cut...'. Well, what do they say? They say strong, be strong*

A: Aha

*B: Then, they made him to cut by telling him to be strong.*

These examples show that speakers want to clarify the information or the topic or they want to engage their interlocutors. While in many cases *nima* may function as a placeholder when the speaker has difficulties in finding the right word, it also play the role of clarification and engagement of the interlocutor. In all three examples above the speaker is asking herself, trying to clarify the information to make sure that the interlocutor has understood the conversation and engaging the other speaker in the conversation.

### The usage of shu 'well' in the discourse

In general *shu* 'this' is a pronoun that points at something or someone. For example, *shu bola meni urdi* is translated as 'This child hit me.' In spoken discourse it plays other roles, such as sharing information, introducing, explaining, or clarifying. It is translated into English as 'well.' The examples below demonstrate their role in shared information.

(14)

B: Voy kim bilsin, zo'r bolardi. Biz bilmapmizda bunaqa... Bugun Dbek aytib qoldi-da, *shu* Sharifa opamiz bilan gaplashinglar opa deb, shunaqa devdi. Voy jon deymiz, qaniydi ertaroq anayqa qimagan ekan bilmapmiz bizam.



*Oh, who knew, it would have been great. We didn't know... Today Dbek told us that, well, you needed to talk to Sharifa sister. Oh, we would be happy to talk her, but we didn't know about it.*

(15)

A bo'sh vaqtim hozir... a **shu**, D aka bilan o'qiymiz. Keyin boshqa vaqtlarda debat kurslari bor... Debat kurslariga borardim oldin.

*Hm my free time now... well, we study with brother D. Then in other times, there are debate courses... I used to go to debate courses.*

In (14) and (15) **shu** is taking the same function as English well in the discourse. Fraser (1999) referring to Labov and Fanshel (1977:156) points out that “well” refers to some topic that was shared among the participants of the speech. In all these examples the ‘well’ refers to a topic that is shared among the participants. For instance, in (14) *Bugun Dbek aytib qoldi-da, shu Sharifa opamiz bilan gaplashinglar opa deb, shunaqa devdi.* ‘Today Dbek told us that, well, you needed to talk to Sharifa sister discourse marker **well** is referring to the person that is familiar to both speakers that is *Sharifa sister*.

## Discussion

Discourse markers at the functional level can work on numerous linguistic levels at the same time points out Jucker and Ziv, 1998, therefore, they can be counted in many other ways. Sometimes they can function in its original role and sometimes they can mark particular information in the discourse. The discourse markers **haligi** ‘like,’ **nima** ‘you know’, **shu** ‘well’ may operate in different ways. In analyzing the discourse markers the focus of the study is important. How these discourse markers are being analyzed and what is being paid attention to. In these analyzes, I tried to briefly look at their functions in the discourse, but they all need more deeper analyses of its own. These analyses were made on the bases of only one regional speaker’s utterances. The overall educational level of the speakers, their language contact with Russian, English and with other dialects may have had an influence in their speech in Uzbek. Therefore, other factors should be examined too.

## Conclusion

This study briefly examined the discourse markers in the speech of individuals from the southeastern part of Uzbekistan. In addition to their direct functions, these pronouns functioned in different ways in the discourse. Discourse marker **haligi** ‘like,’ operated as a marker of new information and focus. The discourse marker **nima** ‘you know’, functioned as hedging and clarification. And the most used discourse marker, **shu** ‘well’ functioned as a marker of the information that was already mentioned in the discourse. These discourse markers may be used in other roles, but more data needs to be analyzed in the future.

## References

- Center for Turkic and Iranian Lexicography and Dialectology. Uzbek Dialect Classification: Reshetov’ Classification 1959 (2014) Retrieved from <http://www.indiana.edu/~ctild/Uzbek/DialectClassification>
- Fraser, B. (1999). What are discourse markers? *Journal of Pragmatics*, 31, 931-952.
- Jean, E. Tree, F & Schrock, J. C. (2002). Basic meanings of *you know* and *I mean*. *Journal of Pragmatics*. 34, 727-747.
- Jucker, A. H. & Ziv, Y. (Eds.) (1998) *Discourse markers descriptions and theory*. John Benjamin Publishing Company. Amsterdam/Philadelphia.
- Muller, S. (2005). *Discourse markers in native and non-native English discourse*. John Benjamin Publishing Company. Amsterdam/Philadelphia.



- Natalevich, O. (2010). Slova paraziti i rechevaya kultura (Words-parasite and speech culture). Registr. *Informatsionno-pravovoe agentstvo*. 6. Retrived from <http://www.profmedia.by/pub/sec/art/41972/>
- Schiffrin, D. Tannen, D. & Hamilton, H. E. (Eds.) (2001). *The handbook of discourse analysis*. Blackwell Publishers. Malden. Massachusetts.
- Underhill, R. (1988). Like is like, focus. *American Speech*. 63 (3), 234-246.

## Uzbek's suffix expressing possession -li

Shinsuke Hidaka  
 hidaka.shinsuke.g0@tufs.ac.jp  
 Tokyo University of Foreign Studies

### Abstract

In Uzbek, *-li* is an adjective-forming suffix. An adjective derived with this suffix expresses possession of the entities expressed by its base noun stem (Kononov 1960: 147, Bodrogligeti 2003: 372, Берматов ва бош. 2008: 583). However, *-li* overlaps functionally with *bor*, which expresses existence as well as possession. I conducted elicitation using these suffixes with two native Uzbek speakers. From the results, I concluded that *-li* and *bor* differ in two ways. The first difference is in “possession cline” (Tsunoda 1995: 567): if N is animate, the border between *bor* and *li-* is on “kin.” Moreover, *-li* is permitted if N is “wife” or “child”; however, *-li* is not permitted if N is “elder brother.” The second difference is in information structure: if N is inanimate, speakers use different expressions. When speakers refer to a previously known event, they use *-li*, but when they refer to a previously unknown event, they use *bor*.

*Keywords:* Uzbek, possession, possession cline, information structure

In Uzbek, *-li* is an adjective-forming suffix. An adjective derived with this suffix expresses possession of the entities expressed by its base noun stem. However, *-li* overlaps functionally with *bor*, which expresses existence as well as possession, as shown in the following examples (judged for accuracy by informants).

- (1) a. *U odam bola-li-ø.*  
       that person child-PROP-3SG
- b. *U odam-ning bola-si bor.*  
       that person-GEN child-3SG.POSS exist  
       “That person has a child.”

In this paper, I examine whether the Uzbek suffix *-li* exhibits the features described in Kazama (1999). Kazama (1999) examined possession suffixes in several Altaic languages (Mongolian, Turkish, Sakha, and Tungus) in terms of the following three points, (of which points 1 and 2 refer to the possessor's character): 1. “possession cline” (body part > inherent attribute > clothing > kin > pet animal > product > other possessee) (Tsunoda 1995: 576); 2. “character of place” (Teramura: 1968); and 3. information structure.

### Literature Review

#### *The usage of -li*

The suffix *-li* generates an adjective that expresses possession.

- (2) *Talant-li* (< *talant* + *-li*)  
       “talented”      talent      PROP

A word derived using *-li* may be used in two patterns. In the first (3), *-li* is used as a noun; the noun *odam*, meaning “person,” is not present. As may be seen in the example, *-li* accepts noun morphology in this usage. In the second pattern (4), *-li* is used as an adjective; *odam* is present in this case.

- |   |  |
|---|--|
| (3) Noun:<br><i>Baxt-li-lar-dan</i><br>fortune-PROP-PL-ABL<br>“from a lucky [person]” | (4) Adjective:<br><i>bola-li odam</i><br>child-PROP person<br>“person who has a child” |
|---|--|

### The usage of *bor*

The first pair of examples demonstrates the nominal usage of *bor*. In (5a), *bor* is followed by the possessive suffix *-i*, expressing the 3rd person singular, and the locative case suffix *-da*. As may be seen in the example, *bor* accepts noun morphology in this usage. In (5b), *bor* modifies the noun *joy*, meaning “place.”

- |   |   |
|---|---|
| (5) Nominal usage (Sjoberg 1963: 57, 61):   |   |
| a. <i>bor-i-da</i><br>existent-3SG.POSS-LOC<br>“during his lifetime”<br>(lit. during-his-existence) | b. [ <i>suv bor</i> ] <i>joy</i><br>water existent place<br>“a place with water”<br>(lit. water-existing place) |

The second pair of examples demonstrates two existential/possessive sentence structures using *bor*. In (6a), the number and person of the possessor are shown by the possession suffix attached to *soat*, meaning “watch.” Boeschoten (1998: 372) claims that alienable possession can also be rendered using the locative, as in (6b).

- |  |  |
|--|--|
| (6) Existential/possessive sentence (Boeschoten 1998: 372):  |  |
| a. Possessed-Possessive suf.+ <i>bor</i> :<br><i>Soat-ing bor=mi?</i><br>watch-2PL.POSS existent=Q<br>“Do you have a watch?” | b. Possessed+Possesser-LOC+ <i>bor</i><br><i>U kitob men-da yoq.</i><br>that book 1SG-LOC no<br>“I do not have that book.” |

### How do *bor* and *-li* differ?

As previously mentioned, statements such as “That person has a child” may be rendered using two patterns in Uzbek. (1a) provides an example using *-li*, while (1b) provides an example using *bor*. The following subsection provides two answers to the question of how *bor* and *-li* are different. The first answer, that they differ in possession cline, was introduced by Tsunoda (1995). The second answer is that they differ in information structure: when a speaker regards an event as old information, *-li* is chosen.

### Investigation

I have used the conceptual viewpoints of Kazama (1999) in the present investigation. Kazama (1999) examined possession suffixes in a range of Altaic languages (Mongolian, Turkish, Sakha, and Tungus) in terms of the following three points.

Possession suffixes in Altaic languages (Mongolian, Turkish, Sakha, Tungusic)

1. “Possession cline” (Tsunoda 1995: 576)
2. “Character of place” (Teramura: 1968)
3. Information structure

When the possessor is animate, I have considered first point, while when possessor is inanimate, I have considered point 3. I do not discuss point 2 (“Character of place”) in this presentation, as it is not involved in the difference between *-li* and *bor*.

### *Possession cline*

I will first discuss the difference between *bor* and *-li* in terms of Tsunoda’s (1995: 576) “possession cline.” This cline indicates “the degree of closeness/attachedness between possessor and possessee,” body part > inherent attribute > clothing > kin > pet animal > product > other possessee.

Tsunoda (1995: 592) also examines “possessor ascension.” The possessor can ascend in example (7), while in example (8), the possessor cannot ascend. The possession cline provides an explanation for this difference. The possessor (7), “leg,” is high on the hierarchy, permitting (7b); the possessor in (8), “wine bottle,” is low on the hierarchy and does not permit (8b). The possession cline is useful for discussing grammatical phenomenon.

- |   |  |
|---|--|
| (7) a. <i>I hit his leg.</i><br>b. <i>I hit him on the leg.</i> | (8) a. <i>I hit his wine bottle.</i><br>b. <i>*I hit him on the wine bottle.</i> |
|---|--|

### *Information structure*

Kazama (1999: 95-96), quoting from Lyons (1968), explains that the following sentences differ in term of information structure. In example (9), single-underlined words represent the topic, and double-underlined words represent comment or new information.

- (9) a. I have a book. / I have a book.  
b. The book is mine.

As shown in the following investigation, this distinction plays a role in the difference between *bor* and *-li*.

Kazama’s analysis (1999) did not include Uzbek. For this reason, I conducted an investigation of Uzbek using Kazama’s work on the possession cline and information structure. In this investigation, I conducted elicitation sessions with two native speakers of Uzbek. I constructed sentences in Uzbek and had native speakers judge whether or not those sentences were permissible.

## **Examination**

I examined the differences of *bor* and *-li* in terms of possession cline and information structure.

### *Possession cline*

Table 1 displays the results for cases of predicative use. The check marks indicate which properties can be used with *-li* or *bor*. Table 2 displays the results for cases of attributive use.

Table 1: Possession cline in the predicative use of *-li* and *bor*

	Body part/inherent attribute		Clothing	Kin		Pet	Other
	Everyone	Not everyone		Child, wife	Elder brother		
Except below	✓		✓				
<i>N-li</i>	✓	✓	✓	✓			
<i>bor</i>		✓		✓	✓	✓	✓

Table 2: Possession cline in the attributive use of *-li* and *bor*

	Body part/inherent attribute		Clothing	Kin		Pet	Other
	Everyone	Not everyone		Child, wife	Elder brother		
Except below			✓				
<i>N-li</i>	✓	✓	✓	✓			
<i>bor</i>		✓		✓	✓	✓	✓

Table 3 combines the results of Tables 1 and 2. The suffix *-li* is relatively high on the hierarchy, while *bor* is relatively low on the hierarchy.

Table 3: Possession cline in the attributive use of *-li* and *bor*

	Body part/inherent attribute		Clothing	Kin		Pet	Other
	Everyone	Not everyone		Child, wife	Elder brother		
<i>N-li</i>	✓	✓	✓	✓			
<i>bor</i>				✓	✓	✓	✓
ex.	(10)		(11)	(12)	(13)	(14)	

I have provided examples of the possession cline in the following section. The sentences using *-li* are permitted through example (11), but the sentences using *bor* are not permitted.

## (10) Body part/inherent attribute

a. *U qiz uzun soch-li-ø.*  
that girl long hair-PROP-3SG  
“That girl has long hair.”

b. *\*U qiz-ning uzun soch-si bor.*  
that girl-GEN long hair-3SG.POSS existent

## (11) Clothing

a. *U odam ko'k ko'ylak-li-ø.*  
that person blue shirt-PROP-3SG  
“That person wears a blue shirt.”

b. *\*U odam-ning ko'k ko'ylag-i bor.*  
that person-GEN blue shirt-3SG.POSS existent

Interestingly, in case of a child, both sentences are permitted, as shown in (12).

- (12) Kin; child, wife  
 a. *U odam bola-li-ø.*  
 that person child-PROP-3SG  
 “That person has children.”  
 b. *U odam-ning bola-si bor.*  
 that person-GEN child-3SG.POSS existent

However, in the case of an elder brother, *-li* is not permitted.

- (13) Kin; elder brother  
 a. \**U odam aka-li-ø.*  
 that person elder-PROP-3SG  
 b. *U odam-ning aka-si bor.*  
 that person-GEN elder-3SG.POSS existent

For a pet animal or other possessee, *-li* is not permitted, while *bor* is permitted.

- (14) Pet animal; other possessee  
 a. *U odam-da ruchka bor.*  
 that person-LOC pen existent  
 “That person has a pen.”  
 b. *U odam-ning ruchka-si bor.*  
 that person-GEN pen-3SG.POSS existent

### *Information structure*

I will next explain my findings regarding information structure. If a speaker is referring to old information, *-li* is chosen, while *bor* is chosen when referring to new information. I have provided examples in the following section.

If a speaker already knows that “That room has two windows,” the statement is expressed as shown in example (15).

- (15) *U xona ikki deraza-li-ø.*  
 that room two window-PROP-3SG

This sentence expresses old information and uses *-li*. Conversely if the speaker is entering the room for the first time, the statement is expressed as shown in example (16).

- (16) *U xona ikki deraza bor.*  
 that room two window existence.

This sentence expresses new information, and uses *bor*.

### **Summary**

Finally, I have summarized the properties of *-li*.

#### 1. Possession cline

Body part, inherent attribute, clothing, kin (child, wife)

However, kin (elder brother) is not suitable.

#### 2. Information structure

*Old* information



**Abbreviations**

-	suffix border	LOC	locative	PROP	proprietary
1, 2, 3	1 <sup>st</sup> , 2 <sup>nd</sup> , 3 <sup>rd</sup> person	PL	plural	Q	question maker
ABL	ablative	POSS	possessive	SG	singular
GEN	genitive				

**References**

- Бегматов, Э., А. Мадавалиев, Н. Махкамов, Т. Мираев, Н.Тўхлиев, Э. Умаров, Д. Хутойберганава, А. Ҳожев. (2008) *Ўзбек тилининг изоҳли луғати : 80000 дан ортиқ сўз ва сўз бирикмаси: 5 жилд*. Тошкент: “Ўзбекистон миллий энциклопедияси” Давлат илмий нашриёти.
- Bodrogligeti, András J. E. (2003) *An academic grammar of Modern Literary Uzbek*. München: LINCOM EUROPA.
- Kazama, Shinjiro (1999) Proprietary Affixes in some Altaic Languages. (アルタイ諸言語のいくつかにみられる所有／存在を示す一形式について) *Altai Hakpo* 9: 93-124.
- Kononov, A. N. (1960) *Grammatika sovremennogo uzbekskogo jazyka*. Moskwa, Leningrad: Izdatel'stvo Akademii Nauk SSSR.
- Teramura, Hideo (1968) Subcategory of Japanese noun. (日本語の名詞の下位分類) *Japanese Language Education*. (日本語教育) 12: 42-57.
- Tsunoda, Tasaku (1995) The possession cline in Japanese and other languages. Hilary Chappell and William McGregor (eds.), *The grammar of inalienability [:] A typological perspective on body part terms and the part-whole relation*, 565-630. Berlin: Mouton de Gruyter.

## Pamiri Languages and Problems of Written Language

M. Iolov

P. Iolova

Sh. P. Yusufbekov

*Academy of sciences of the Republic of Tajikistan*

### Abstract

Pamirian languages date back to the ancient Eastern Iranian dialects, unlike Tajiki, an Iranian language of the western group. Historical comparative analysis of phonetics and structural- typological similarities of the modern Pamiri languages allows us to assume that these are areal groups of languages. This Union originated after the arrival of the ancestors of the Pamirians in the area due to the mutual influence and convergent evolution of the languages, as well as the overall interaction with the substrate side and the superstructure of the Tajik language.

Keywords: Shughni language, Pamir, written languages

### 1. Introduction

The scientific study of the non-written Iranian languages of the Pamir and Eastern Hindi Kush has a rich history. First of all, it should be noted that European, Russian and Soviet linguists and historians made a great contribution in the linguistics research. Such classical works by R. Shaw, K.G. Zaleman, R. Gauthiot, G. Morgenstierne, A. A. Bobrinskoi, I.I Zabubin, A. E. Snesev, M. S. Andreev, V.A. Lifshitz, T. N. Pakhalina, A.K. Pisarchik, V.S. Sokolova, etc... form the basis of modern Eastern Iranian linguistics. Tajik linguists had also contributed to this science. Their names are R.Kh.Dodikhudoev, D.K.Karamshoev, M. Faizov, N. Karamkhudoev, B.B. Lashkarbekov and many others.

In the works of the above researchers, synchronous description of all Pamir languages is given. They analyzed in depth the lexicology and studied historical phonetics and grammar of the Pamir languages. Proto-language conditions of some of these languages were reconstructed at the level of ancient Eastern-Iranian dialects. And some of the similarities with the common Iranian dialects was revealed. All of their works allowed to identify the genetic relationships between some of the Pamir and Eastern Iranian languages and to determine their place in the family tree among the same Eastern Iranian branch.

However, these studies are often purely linguistic and ethno-linguistic problems are barely covered. At this stage of overall linguistics development, the interest in ethno-linguistic problems increases on a daily basis. The urgency of this problem is also related to the fact that rapid globalization of the modern world can lead to a drastic change in ethno-linguistic situation, even in remote parts of the world, and to the disappearance of the various ethnic communities that tend to leave their traces with all ancient civilizations.

Scientific analysis of the ethno-linguistic history of smaller groups of people of the Pamirs can also contribute to the number of controversial issues of ethnicity of the peoples of Pamirs. Such modern authors as B.B.Lashkarbekov, Sh. P.Yusufbekov and L.R. Dodikhudoeva contributed significantly to the above discussions with their recent publications.

## 2. Ethno-linguistic situation in Pamir-Eastern Hindu Kush region

In the monotonous region of the Pamir-Eastern Hindu Kush, located in the heart of the Central Asia, live nations and ethnic groups that speak both Indo-European and non-Indo-European languages. Indo-European languages include Iranian, Indian, Dazdian and Nuristanian languages while non-Indo-European languages are Turkic and Burushaski.

We are interested in the ethno-linguistic history of small indigenous peoples speaking Eastern Iranian languages that include:

- Munji (whith igda dialect)
- Wakhi
- Closely related Ishkoshimi and Sanglichi as language-dialects
- Shughni-Rushani group (Shughni, Rushani, Bartangi, Roshorvi and Sarikuli language-dialects)
- Yazgulomi

### 2. a Geography of Eastern-Iranian languages

Native speakers of all these languages live in four different countries: Tajikistan, Afghanistan, China and Pakistan. More precisely, in the Tajik, Afghan and Chinese parts of the Pamirs and in some valleys of the Eastern Hindu Kush that are parts of Afghanistan and Pakistan.

According to Gruenberg and Steblin-Kamenskiy, it is a separate ethno-linguistic region, because these people have much in common not only in language but also in material and spiritual culture. In this regard, the ethno-linguistic region referred to herein should be explored not otherwise as historically established integer. Just covering all ethno-cultural space, it is possible to recreate a realistic picture of local ethno-linguistic communities.

It should be noted that until now these Iranian-speaking peoples (nations), despite the emerged disunity and political fragmentation are a kind of unity that created their own special regional Pamir-Eastern Hindu Kush culture.

### 2. b Ethno-linguistic map of Eastern-Iranian languages

Here is some statistics on dissemination of Eastern Iranian languages speakers in the region of Pamir-Eastern Hindu Kush:

- 1) Shughni (105-110 thousands) live in Khorog city, Shughnan and Roshkala districts of the Republic of Tajikistan. Shughnan viloyati of Afghanistan.
- 2) Wakhi (65-70 thousands) live in Wakhan valley of Ishkashim district of Tajikistan, in Tajik autonomous viloyati Qashqadaryo of Xinyang-Uyghur Autonomous region of China and Khunza and Ishkoman valleys in Northern Pakistan in the foothills of the Hindu Kush.
- 3) Rushani and Bartangi (near 30 thousands) live in Rushan district of the Republic of Tajikistan.
- 4) Sariquli (25 thousands) live in Tashqurgan viloyati Qashqadaryo of Xinyang-Uyghur Autonomous region of China.
- 5) Yazgulomi (8-10 thousands) live in Yazgulom valley of Wanj district of Tajikistan.

- 6) Munji (near 4 thousands) live in Khunza valley in northern Pakistan.
- 7) Ishkoshimi (1.5 thousands) live in Rin and Sumchin villages of Ishkoshim district of Tajikistan.
- 8) Sanglichi (from 100 to 150) live in Badakhshan province of Afghanistan.

Hence, the total number of Iranian-speaking of Pamir-Eastern Hindu Kush reaches 240-250 thousands. Should be noted that peoples of the Pamirs live in the capital of Tajikistan, Dushanbe city and in some other areas of the south Tajikistan (Qumsangir, Jillicul, Qabodiyon), north Tajikistan (Khujand, Chkalovsk), in Kirgizia (Osh city) and in Russia (Moscow, Krasnodar, Vladimir, Samara, Sankt Petersburg, Ekaterinburg) as compact communities. There isn't an exact data on the population of these communities. From non-official data we can find that about 50 thousand of the Tajik Pamiris live in Post-Soviet countries outside of Badakhshan, Tajikistan.



### *2.c Definition of Pamir languages*

Definition of the Pamir languages: Pamir languages the concept is likely to be geographical, because in genetic aspect they are generally not a separate branch of the Eastern Iranian sub-groups. There exists a secondary commonality between the northern Pamir languages such as Shughni-Rushani sub-group.

### **3. The initial period of the spread of Iranian languages in Pamir - Eastern Hindu Kush region**

Linguistic research has shed light on some features of the eastern Iranian tribes settled in this region in the first half of the first millennium BC, to identify ways and stages of their migration and settlement. The language can be judged by the characteristic features of the material and spiritual cultures of ancient settlers.

#### *3a. Migration and Settlement*

Studying the genetic relationships of the Pamir languages most researchers agree that in this case the common linguistic state could not exist. This means that the ancient settlers of the Pamir - Eastern Hindu Kush were independent from each other from the beginning of the development of the region. Perhaps the Iranian

settlers in these places came as waves at different times from different places. From the beginning these were tribes with distinct ethno-linguistic attributes. To penetrate into the region as they could from the east through the Alai valley and from West passing through the Darvaz ridge. It is possible that some tribes reached Pamir from the south through the current Afghan Badakhshan a more accessible path than that of the West and East.

Research on history of phonetics and grammar indicate on that with the collapse of the south-eastern sub-groups some of the protopamir languages developed autonomously. It has been discovered that genetic commonality between Yazgulomi language and Shughni –Rushani group and between Munji language and Yazgulomi-shighni group. The first inhabitants of the Pamir coming from the East according to the archeological data belong to nomadic saak tribes. Perhaps they include protowakhis separated from the main mass of the saaks. The name of the first top Wakhi village on the Pamir river *Ratm* came from the ancient Iranian < \*fra-tama which means “first,” “initial” indicates that Saak nomads entered the Wakhan valley exactly from the Eastern side.

### *3b. Ethno-confessional relations and characteristics*

In the names of the individual religious buildings reflected in the beliefs of the ancient inhabitants, the ancestors of the contemporary Pamirians. For example, ancient Wakhis were sun worshipers and constructed religious buildings in honor of the heavenly bodies (wakhi veriz – the name of a place in Langar village where there is a religious building came from the ancient Iranian hvara-yaza “a place for sun worship”; vendat, name of a village in Afghan Wakhan came from ancient Iranian hvan-data “sun-given.” Later with the spread of Zoroastrianism the ancestors of ancient Wakhis accepted this new belief but brought in new elements of their own faith. Furthermore, the word sun was acquired from the Supreme Deity of Zoroastrians – Ahuramazda ( ahura, sun is ir in Wakhi and khir in Shughni.)

This suggests that the ancient settlers of the Pamir and Hindu Kush worshiped the sun and in the period of Zoroastrianism belonged to the same religion.

## **4. The process of the emergence of the new ethno-linguistic communities**

The collapse of the ethno-linguistic Eastern Iranian communities began very early in the period of tribal system when language of a tribe was still a dialect of an ancient Iranian language. Southern and northern sub-groups of languages could not keep the linguistic unity and very soon broke up into new dialects. Southern sub-group at early stages split into dialects which formed the basics for the Khotan-saak and Wakhi languages, likely due to expressive mobility of the nomadic tribes of Saaks. The relocation of the protowakhis in the Pamirs was the cause for the split of these commonalities. This separation could have occurred in 7th - 6th centuries BC. The transition to a sedentary lifestyle opens a direct path to the formation of an independent Wakhi language in the third stage of language development. Chinese historical chronicle gives information on Wakhan at the turn of our era. It has been indicated that the beliefs of the ancient Wakhis represent a synthesis of the Iranian saaks and local pre-Iranian traditions then transformed into one of the varieties of Zoroastrianism. In the Middle Ages in Wakhan penetrates Buddhism. All this has left its mark in Ismailism, that spread in Pamir near 1100 years ago.

## **5. Modern and Contemporary Ethno-linguistic History of small nations of the Pamir- Eastern Hindu Kush**

In 1895 an Agreement on the definition of the boundaries of two powers was signed by the two Russian and British Empires. De jure put an end to the Big Game. De facto allowed for the so-called “Pamir question.”



Thus this led to the establishment of the colonial rule over Central Asia on the one hand and northern India and Afghanistan on the other.

Qualitatively new political conditions changed the riverbed ethno-linguistic processes and the new epoch can be formally divided into three stages.

### *5a. Colonial period*

Colonialism put an end to the oldest surviving in the Central Asian ethno-linguistic entities that were the Pamir and Eastern Hindu Kush leadership. Historic Wakhan, Shughnan, Rushan and Darvaz arranged in a sequence one after another in the upper reaches of the Amu-Darya, where dissected: the left bank of the village moved to a dependence on Britain to Afghanistan and the right bank to Russia and Bukhara.

Wakhan settlement in Hunza and Ishkoman became part of British India (now Pakistan.) The region has changed the political and living conditions which also led to the formation of new traits in local groups of the same people living on different sides of the border. So under the influence of language Burushaski, Urdu, and English in northern Pakistan formed the new Wakhani speech. We have different ethno-cultural environment in which changing the mentality of people, the evolution of traditions, customs, and dress code, music and performing arts. Such changes are observed in the traditions and customs of Wakhian, Ishakishimian, Shughnian, and Rushanian of Afghanistan and Tajikistan. The Pamir languages are strongly influenced by Dari and Pashto in Afghanistan, Tajiki and Russian in Tajikistan. Particularly, lexical elements of these languages were intensively borrowed from other languages.

### *5b. Soviet period*

Soviet period left an imprint in languages and cultures of the people and defined an advantageous condition for many years ahead. It was the most favorable period in the history of Pamirs in the modern times, despite all the negative effects caused by ideological attitudes. In the late 30s Shughni language acquired a written form and became the subject of schooling. Unfortunately, such a development in language and culture did not last very long. In the post war years and further for the people of the Pamir it was ceased to recognize the independent status of this ethnic group. Pamir languages were prohibited to be used in public places, kept from learning, were not allowed to perform, sing songs or pass them on to the radio.

### *5c. Independent period*

In the post-Soviet era many restrictions were removed that were raised out of previous political and ideological installations. In 1991, ABC Shughni language was re-introduced. There were a lot of creative groups and ethnographic ensembles performing pamirian songs and dance. Population of the region came out of spiritual isolation while establishing ethno-confessional connection with their spiritual leader the Aga Khan IV and coreligionists worldwide.

However, in recent years due to the economic turmoil in the country, people in the Pamir faced with discrimination based on ethno-linguistic and religious indications. Youth population at the Pamir leaving their homes in search of work and wondering around Russia, other CIS countries and abroad. One can only hope that these challenges carry a temporary character and the society is able to overcome the difficulties.



# Linguo-Methodical Objectives and Frequency-Specific Approaches for Kazakh Language Instruction

Aliya Kuryshzhanova  
 aliya.kuryshzhan@gmail.com  
*Hankuk University of Foreign Studies*

## Author's Note

This work was supported by the Hankuk University of Foreign Studies Research Fund.

Correspondence concerning this article should be addressed to:

Aliya Kuryshzhanova  
 Department of Central Asian Studies,  
 HUFs, Global campus, Wangsan-ri, Mohyeon-myeon, Cheoin-gu,  
 Yongin-si, Kyeonggi-do, 449-791, Korea

## Abstract

Frequency-specific lexical lists provide objective criteria on which to base accurate level-specific content distribution, quantitative characteristics, adequate dosing and distribution of teaching materials by levels, objective assessment and others. This technique also makes the teaching objectives more concrete and visible. The modern Kazakh language and its contemporary vocabulary are reflected above all in its spoken form. Therefore, a beginning frequency list has been conducted of current Kazakh speech based on types of modern communication. In particular, we analyzed spoken conversations in movies (25000 units), mass media interviews and dialogues from literature essays (30000 units). Selected frequency-specific vocabulary helps to differentiate content for various educational levels differentiation, identifying nuclear / basic high frequency vocabulary for specific area and also identified objects for accurate assessment. Use of frequent vocabulary and grammar minimums allows reducing and detailing the studied content to what is actually needed within the usually very limited time of language courses. This way provides possibilities to solve multiple needs of Kazakh language researchers, methodologists and practice teachers.

Keywords: teaching methods, Kazakh language, frequency list

## I. Introduction

Demographic growth of the Kazakh ethnic group, major changes in the community, and urban globalization have significantly increased the spheres of use of the state language. This in turn has caused major changes in the language's vocabulary. New expanded functions have to be created by use of the state language in such include services KazNet.kz, Uikibilim.kz, Google.kz, e-gov.kz, text semantical analyzers and SEO formalizers, search machines / searchers, and many others. Kazakhstan's language policy has stimulated the use of the state language by both native and other speakers, thus demanding a search for new corpus and status development of the language and new methods of instruction.

State policy seeks to achieve mastery of Kazakh by all citizens; full-fledged social functioning of the language, production of linguistic information and communication technologies, and many other goals. Al-

though state language policy identifies approaches to achieve many current linguistic objectives, problems of studying the state language will remain unfulfilled until the country attains the state of socially-comfortable poly-lingualism.

Although much research has been conducted about Kazakh linguistics, greater demand for «Kazakh for foreign / Russian-speakers» has become manifest since independence. Kazakh language methodology reviews the training problems from the following perspectives: a) Kazakh as L1KZ (first / native language, i.e. Kazakh for Kazakh-speaking) and b) Kazakh as L2KZ (second / non-native / foreign – for those speaking other language). The current language situation makes L2KZ particularly relevant, and this issue requires specific approaches to language acquisition.

The increased attention to methodological issues requires major substantiation of teaching parameters that affect the result: among these are detailed objectives of training, level-specific training and specialization, as well as such parameters as language environment expanding.

This means that it is necessary to various different methodologies a) when Russian-speaking Kazakhstan citizens study the state language as L2KZ in the language environment from the methodology required b) when Kazakh language is studied as L2KZ outside the language environment. For instance, specific methodologies are required for Kazakh language courses in Kazakh communities, for language training of foreign companies' employees, and for students of foreign universities. When preparing training courses and selecting methodologies, identification and using focused in frequency parameters of language leads towards more efficient performance results.

In order to find an efficient methodology, it is necessary to identify the objective criteria for a course, an accurate level-specific definition of the training content, quantitative counting and then the statistically distributed language material by study levels, objective assessment, this is best accomplished by applying frequency-specific language characteristics (frequency of the most often used words). Hierarchic frequency-specific lexical lists of educational language materials make the training objectives more concrete. However, existing Kazakh frequency vocabularies were created based on written texts, and they are outdated, since they were compiled in the last quarter of the last century. Meanwhile, modern Kazakh, with its more up-to-date vocabulary reflects the current state of the language. This first reflected in the verbal form, because verbal use more promptly reacts to changes and more adequately reflects real functioning of the current language.

The research for this paper has been conducted in order to analyze the current Kazakh language based on movie materials, mass media interviews and dialogues of Kazakh writers' literature. Data have been processed with specific software codes to formalize and create an algorithm of the language, and form ranked vocabulary databases for the verbal Kazakh language.

The research for this theme has been conducted by processing modern Kazakh speech and identifying the most frequently used vocabulary; these items were used to develop a primary education course. In selecting educational materials based upon frequency parameters it becomes obvious that various given characteristics of the Kazakh language significantly affect developing frequency vocabulary.

Let us compare frequency lists made primarily for modern Kazakh oral speech with the frequency vocabulary of Abai Kunanbayev's written texts. Significant differences of parameters of the applied vocabulary become obvious in comparing these two frequency lists. In this case the given parameters also define the nature of this text frequency vocabulary: written or spoken language. Tables 1 and 2 demonstrate the use of the first 20 most high-frequency lexical units, which differ by oral / written speech parameters.

The first table of these statistic materials are a part of our research for creating the frequency list of oral communication, which was generated by modern Kazakh movies spoken dialogues, and the second table is a part of frequency list of Dictionary of Abay Kunanbaev, located at Kazakh language portal (<http://til.gov.kz/wps/portal>).

Table 1 Modern Kazakh movies dialogues vocabulary

#	lexeme	Quant.	Frequency	Freq. %
1.	ал / ал(у)	662	0,265	26,48
2.	ай	602	0,241	24,08
3.	не	484	0,194	19,36
4.	ол	420	0,168	16,80
5.	бол(у)	370	0,148	14,80
6.	ғой	267	0,107	10,68
7.	бір	250	0,100	10,00
8.	жоқ	245	0,098	9,80
9.	де / де(у)	228	0,091	9,12
10.	ой	223	0,089	8,92
11.	ма	218	0,087	8,72
12.	бар / бар(у)	211	0,084	8,44
13.	сен / сен(у)	203	0,081	8,12
14.	қыз	200	0,080	8,00
15.	ақ	194	0,078	7,76
16.	өз	185	0,074	7,40
17.	кел(у)	182	0,073	7,28
18.	ба	177	0,071	7,08
19.	он	175	0,070	7,00
20.	мен	167	0,067	6,68

Table 2 Dictionary of Abay Kunanbayev

#	Frequency	lexeme
1.	69	болу
2.	56	айту
3.	56	алу
4.	50	көру
5.	45	білу
6.	44	келу
7.	40	қалу
8.	39	беру
9.	39	кету
10.	35	түру
11.	34	жүру
12.	32	деу
13.	25	өз
14.	23	жату
15.	23	сөз
16.	23	қою
17.	22	есту
18.	22	жету
19.	22	отыру
20.	22	салу

In this kind of collation the following characteristic becomes visible this distinction and it proves the most important in this situation: this vocabulary use oral or written speech. These two tables of the most often used 20 words show that spoken and written Kazakh use different vocabulary or even there are the same words they are used with different range of frequency. Besides, other important parameters are: obsolete vocabulary or current neologisms; topics chosen, style and text content as well as some other specific ones.

It is worth mentioning that because these findings demonstrate frequency analysis findings, those conducted at various period, based on various methodologies and with various technologies (e.g. analysis of the movies language became possible due to new technology methods), then parameters of these lists have different characteristics. Thus, a frequency list of Kazakh movies demonstrate more detailed and disseminated frequency characteristics, which provide more information for further use. Although such comparison is not absolutely precise, this comparison of various type frequency vocabularies is obvious and provides more detailed materials for practitioners. Besides, previous years' data are valuable in serving the object of synchronic and diachronic analysis.

The case illustrated above shows the difference between the oral and written forms of the language. The lists for both contain similar lexeme units of the most frequent vocabulary, which falls into top 20; however, these lexical units indicate various frequency ranges, thus characterizing their various application in oral and written communication.

Awareness of this information is important for many practical applied objectives, one of them is the teaching Kazakh. In the present case we have looked within the focus of educational purposes, where definition of lexical minimum for levels of education, their adequate control and evaluation are always critical. If teachers possess such information as what vocabulary is important for a particular level; what specific language is re-

quired for a particular profession; or what vocabulary is important for particular audience or goals – they obtain more opportunities to develop more targeted and time-efficient educational course, which is critical within the limited length of an educational course.

Besides, frequency lists based on written speech texts and oral speech, obviously indicate what has to be included into the oral communication course, which vocabulary has to be emphasized, e.g. within the academic writing course – differences of frequency vocabularies of oral and written speech are obvious enough. Therefore, besides the obviously most nuclear vocabulary, the specific objective of teaching – oral communication or writing – should be considered while developing every language course.

It is also worth mentioning the widely disseminated practice is that the teachers emphasize semantically important vocabulary for the topic of a lesson or course. This is usually justified with a significant rule of thumb. However, frequency analysis findings indicate that semantically zero or less significant, or semantically insignificant connective words are the most applicable. For language teachers this means that equal attention should be paid to teaching this particular vocabulary too. Work on structurally formalized and abstract models of statements, usually built upon most frequent connective words, is a good example of this: e.g. if the students have been taught to understand and work using the «Kim ony zhasajdy?» / «Neni kalaj zhasajdy?» / «Kim ony kalaj zhasajdy?» and other models, then the following semantically lexical replenishment and model variation gets mastered by them more successfully.

Let us now compare frequency lists for Kazakh language (the analysis gives data for current Kazakh mass media language – Table 3) and, for instance, frequency educational dictionary of Russian language (N. J. Brown. Russian Learners' dictionary 10000 words in frequency order. 2003 – Table 4). Both the first and second examples demonstrate high ranges of frequency of helping words in particular:

#	lexeme	Frequency
1.	және	344
2.	мен	325
3.	бұл	308
4.	да	250
5.	деп	246
6.	астана	200
7.	бір	198
8.	baq.kz	186
9.	үшін	179
10.	бойынша	175
11.	де	167
12.	қазақ	167
13.	тамыз	161
14.	Қазақстан	156
15.	осы	154
16.	ал	143
17.	бар	141
18.	ол	136
19.	деген	134
20.	оның	128

#	lexeme
1.	и
2.	в (во)
3.	не
4.	на
5.	я
6.	он
7.	что
8.	с (со)
9.	это
10.	быть
11.	а
12.	весь (вся, все, все)
13.	они
14.	она
15.	как
16.	мы
17.	к (ко)
18.	у
19.	вы
20.	этот (эта, это, эти)

First table shows frequency list of the most 20 high ranging lexical units as a part of the frequency dictionary created by materials of modern Kazakh mass media for 2012-2013 (Kuryshzhan A.A., Omirzakova A.K., 2013) and second one demonstrates first 20 units of the most often used Russian language words for Beginners. Almost of these 20 units of both tables are ancillary and auxiliary words in Kazakh or in the Russian. Same situation is for English frequency dictionary (COCA, 2012).

Similar comparison has been conducted for frequency lists of dialogues in the belles-lettres literature of various periods, mass communication media language and TV, direct oral and mediated written interviews, talk-shows and aitys – oral traditional Kazakh poetries competitions (Boranbaev S., Danaev N., Zhumabaev A., Kuryshzhanova A., Omirzakova A., 2013). Findings of such frequency lists and developed methodologies are anticipated to be made accessible at the end of the research. However, findings received so far indicate new application opportunities to address educational objectives: thus, identifying frequency vocabulary provides for the opportunity to distribute the content and differentiate educational levels, identifying nuclear high-frequency vocabulary and less frequently used vocabulary for the professional or specialized education, calculating parameters for accurate and objective evaluation and control.

An adequate testing system based on selected frequent grammar minimums can help to organize the following steps of correction and improving teaching process more properly and exactly: to correct the lexis or grammar, to pay attention to words forms or to models of sentences. Use of frequency vocabulary and grammar minimums at particular level or specialty allows narrowing and detailing the studied content till actually necessary within the limited courses. A developed program and research-based toolkit provides for the opportunity to address multiple specific objectives of Kazakh educators, methodologists and teachers.

Even the preliminary results of this research show possible use intended to solve educational objectives: thus, selected frequency-specific vocabulary helps teaching content distribution, specify the language levels differentiation, identifying nuclear high-frequency vocabulary for basic, intermediate or advanced courses, devise general and professional or specialized education, and clearly calculates parameters for accurate assessment and next correction of found difficulties. Use of frequent vocabulary and grammar minimums for a particular level or specialty allows reducing an amount of necessary materials and detailing the studied content until the one actually needed within the limited courses. Software-research toolkit developed provides possibilities to solve multiple objectives of a Kazakh linguistic theory, applied language products and teaching practice of Kazakh.

## References

- Boranbaev S., Danaev N., Zhumabaev A., Kuryshzhanova A., Omirzakova A. Kazakh tili koldanbaly lingvistikasynyn zhijilik zertteu maseleleri / The issues of frequency research of applied Kazakh linguistics. (2013). Proceedings from “XLIII scientific-methodical conference (ISBN 978-601-247-843-3, Vol.2, p.100-104). Kazakhstan, Almaty: KazNU.
- Kuryshzhan A.A., Omirzakova A.K. Kazak tildi BAQ vocabuljarynyn zhijilik taldauy turaly / About the frequency analyze of Kazakh mass media. (2013). Proceedings from “Actual problems of State language policy” International scientific-methodical conference (ISBN 979-601-7344-02-8, Vol.1, pp. 262-266). Kazakhstan, Almaty: Educational-Methodical Center “Til” of Almaty Department of Languages Committee of the Ministry of Culture of Republic of Kazakhstan.
- N. J. Brown (2003) Russian Learners’ dictionary. 10000 words in frequency order. London-New York: Routledge.
- Portal of the Official Language, Committee on Languages development and public-political works of the Ministry of culture and sport of the Republic of Kazakhstan <http://til.gov.kz/wps/portal>. Lexical search.
- Dictionary of Abai Kunanbaev. Frequency dictionary (2014). Online Web site: <http://goo.gl/NShty1>.
- The Corpus of Contemporary American English (COCA) 450 mln words, 1990-2012. Online Web site: <http://corpus.byu.edu/coca/>.



# Development of the progressive construction in Modern Persian

Narges Nematollahi  
Indiana University

## Abstract

Since its first mention in 1888, the Modern Persian progressive construction with *dāštan* ‘to have’ has received little attention as far as its tense domain and source are concerned. Based on an analysis of 143 cases of present and past progressive tenses collected from several literary works between 1907 and 2010, this paper presents an account of the development of this newly formed yet ever-increasingly used construction in Modern Persian. Applying Vendler’s (1967) framework to classify the situations, this analysis shows that the Modern Persian progressive construction is used with achievement, accomplishment and activity situations, mostly denoting an imminent action in the case of achievements, and an ongoing action in the case of accomplishments and activities. Furthermore, in the light of the distinctive structure of this construction, which is identified as a Serial Verb Construction (SVC), the SVC with *bar dāštan* ‘to set off’ is proposed as the source of the construction in question.

Keywords: modern Persian; progressive tenses; serial verb construction; periphrastic construction

## 1. Introduction<sup>1</sup>

This paper deals with a newly formed verbal construction in Modern Persian, which uses the auxiliary verb *dāštan* ‘to have’ together with the three imperfective forms of the main verb: present, past, and evidential. Both the auxiliary and the main verb appear as finite verbs, inflected for person and number, and taking the same mood. Examples (1)-(3) represent these three progressive constructions:

(1)			
ādam-hā	dār-and	mi-āy-and	
person-PL	have-3PL.IMPF.PRES.	come-3PL.IMPF.PRES	
People are coming.			

(2)		
dāšt	mi-mord	
have-3SG.IMPF.PT	die- 3SG.IMPF.PAST	
He was dying.		

<sup>1</sup> Throughout this paper the terminology used for tense, aspect and mood is based on the description of the verb system of Modern Persian given in Windfuhr (2009: 446-62). The abbreviations used in this paper are: SG (singular), PL (plural), IMPF(imperfective), PERF (perfective), RES/STAT (resultative/stative), PRES (present), PT (past), EV (evidential),PART (participle), INF (infinitive), COP(copula), INDEF (indefinite), MK (marker), DIR.OB. (direct object), NEG (negative), SUBJ (subjunctive), IMPV (imperative), and intr (intransitive).

The Persian words and phrases cited are given in a phonetic transcription; [s], [ʃ], and [ʒ] are all represented by /s/, [z], [ʒ] and [ʒ] by /z/, [t] and [t̪] by /t/, [h] and [h̪] by /h/, [q] and [gʰ] by [q]. /č x ž š/ represent the sounds usually transliterated as *ch, kh, zh, sh*. /i e a/ represent the front vowels from high to low, and /u o ā/ their back counterparts.



(3)			
zāheran	dāšte	sabzi	mi-xaride
apparently	have-3SG.IMPF.EV.	vegetable	buy-3SG.IMPF.EV.
Apparently he was buying vegetables.			

Perhaps due to the fact that the progressive construction “has not yet fully integrated into literary Persian” (Windfuhr 2009:461), this construction has received but little attention since its first mention in 1888. However, the scholars are called upon to do a thorough study of this construction in (Keshāvarz 1962) and (Dehghan 1972), and the topic has been considered as “an interesting case for diachronic and sociolinguistic research” in (Windfuhr 1979:102).

In this paper, first previous accounts of the Modern Persian progressive construction are reviewed. Then, the literary works that have been used in collecting data are introduced, and an analysis of the collected data is presented. The result of the analysis is then compared to the previous accounts, showing that none of the accounts could be considered as adequate. Finally, a possible origin for this construction is proposed, and the morphological and semantic development of the progressive construction from that origin is delineated.

## 2. A review of the previous literature on the Modern Persian progressive construction

Zhukovskij (1888) seems to have been the first who reports on the use of this construction in colloquial Persian. He mentions the progressive present and past tenses only, calling the former as *aoristus* and the latter as *praeteritum*. According to Zhukovskij, progressive present refers to an immediate future action, and can be compared with the French present of *aller* with an infinitive. Progressive past, on the other hand, has the meaning of a past which has just been completed, and is comparable to the French present of *venir de* with an infinitive.<sup>2</sup> Later on, Lorimer (1916:469-70) points to the use of the progressive tenses with the verb ‘to have’ in the Gabri dialect of Persian, the dialect then spoken by the Zoroastrians in Yazd and Kerman. He calls the construction as “a special idiom” which is used when “stress is laid on the actual course of the action”. He also mentions that the same construction is common among the non-Zoroastrians of Kerman too.

Looking into the grammars of Persian, *Persian grammar* by Lambton, first printed in 1953, seems to be the first one which mentions this construction. Lambton (1963:160) points to the progressive present and past constructions, calling them continuous present and past, and translates them just as Lorimer. *Grammair du person contemporain* by Lazard, first printed in 1957, is the only grammar that in addition to the progressive present and past, lists the progressive evidential form too. Lazard (1992:160) calls the latter as the “completed past”. Among the grammars of Persian in the Persian language, Vahidiyān (1963), Bassārī (1969) and Farshidvard (1969) are perhaps the first ones that mention this construction, for which different terms such as *malmus* ‘realized’, *modāvem* ‘continuous’ and *jāri* ‘ongoing’ are suggested (the last two by Yarshater(1970: 670 fn. 6)).

Keshāvarz (1962) and Dehghan (1972) have two scholarly articles dedicated to this construction. The former claims that there has been no record of this construction until a century ago, mentioning a couple of instances in classical and early modern Persian where other constructions, including imperfective present and past, were employed in the sense of progressive. He then examines the progressive construction in Tajik Persian as well as in some dialects of Iran including Gilaki and Talyshi, and finds out that none of them use the auxiliary verb ‘to have’. One should note that they are also different from the Modern Persian progressive construction in that none of them have both the auxiliary and the main verb inflected. Rather, the participial or infinitive form of the main verb is used together with the inflected form of an auxiliary, which is ‘to stand’ in Tajiki and ‘to

<sup>2</sup> As regards the form, his reported form for the present progressive tense is the same as that shown in (1) above, but that of the past progressive is different from (2); according to him, both the verb *dāstan* and the main verb appear in perfective past. Given the fact that in *dāstan*, the forms of imperfective and perfective past are identical, there only seems to be an error in recording the main verb in perfective, rather than in imperfective past

be' in Gilaki and Talyshi.<sup>3</sup> Keshāvarz continues that the progressive construction with the verb 'to have' is not found in any of the foreign languages known to Iranians either, and thus, it could not have entered the language through borrowing.

Calling the progressive present, past and evidential forms as progressive present, past, and perfect respectively, Dehghan (1972) describes the first two as follows:

“The progressive present denotes “(a) an action in the process of being completed at the time of speaking, although it may have begun in the past; (b) an action which will be going on in the future before some other action or state of being (rare) e.g. ‘when you return, I will be (in the process of) writing’; (c) an action which will be completed right away, i.e. in the very near future (very rare) e.g. ‘I am about to come; I come right away’, and the progressive past “is used to express an action that had begun in the remoter past, was in the process of being performed at the time spoken of, and may either have ceased by the commencement of some other action or may have continued for some time afterward, e.g. ‘when I came away he was [still] (in the process of) writing’ (the action continued); ‘I was writing [when] the light went off’ (the action ceased)”<sup>4</sup>. (Dehghan 1972:199-200)

The further remarks made by Dehghan (1972: 200-1) are rearranged by Windfuhr (1979: 102-3) as follows:

- 1- In this construction, negation is blocked
- 2- The progressive construction is impossible with the stative verbs, such as “to have” and “to be”;
- 3- With verbs of progress such as *mordan* ‘to die’ and compound verbs with *šodan* ‘to become’ the progressive construction takes the inchoative meaning ‘to be about to’.

Windfuhr (2009:452) lists the following functions for imperfective tenses, i.e. imperfective present and past, in Modern Persian:

- Habitual-iterative and generic, ‘they (always, generally) go, leave’ in present, and ‘they (always, generally) went, left’ ~ ‘would go, leave’ ~ ‘used to go, leave’ in past;
- Progressive, ‘they are/were going, leaving’;
- Intentional, ‘they are/were about to go, leave’;
- Future ‘they will go, leave’ in present, and ‘they would go, leave (the next day, etc.)’ in past.

In his section on the extended verb system of Modern Standard Persian, then, Windfuhr (2009:461-2) points to the progressive, remarking that “functionally, it (=progressive) disambiguates the progressive and imminent-future functions of the imperfective, but has not yet emptied the latter of that function”.

### 3. Description of data collection and the methodology used to analyze the data

Dehghan (1972:202) mentions that his examination of “a substantial amount of nineteenth century Persian writing”, including the writings which seem to reflect the colloquial language of the time<sup>5</sup>, does not show any progressive construction with *dāštan*. As mentioned above, the use of this construction in colloquial Persian

<sup>3</sup> See Tajiki: xānd-e                      istād-e ast                      “He is reading”  
read-PART                      stand-3SG.RES/STAT.PRES

Gilaki: sho-on                      dar-am                      “I am in the action of going = I am going”  
go-INF                      in -COP.1SG.PRES

<sup>4</sup> In favor of space, the Persian transcription of the examples, which seem to be made-up ones, is excluded from the quotation.

<sup>5</sup> He particularly mentions the Qarāchadāghi’s translations into Persian of several plays by Akhundoff, and the Persian version of James Morier’s *Hajji Baba of Isphahan* by Mirzā Habib of Esfahān.

was first briefly reported in the late 19th century by Zhukovskij, who in addition to two made-up examples of the progressive past and present, records a progressive present form in a satirical folk song popular around the same time. The earliest attestation of this construction in Persian writings seem to occur in *Čarand parand*, literally ‘fiddle-faddle’, the collection of satirical essays by Dehkhodā (1879-1956), which were published in the newspaper Sur-e Esrāfil in the years 1907-1908.

A brief description of the Persian literature in late 19th and early 20th centuries is in order. Soroudi (1993:214) mentions that the Persian prose was “ornate and abstruse” at the beginning of the 19th century, but gradually a tendency towards simpler styles was observed among many writers of the time. Different factors such as the introduction of printing in 1816-17, the appearance of the first newspaper in 1837, and extensive contacts with European countries should be held responsible in forming such a movement. Therefore, simple prose style and the use of everyday idioms and expressions, according to Soroudi, characterize the innovative literary works of late 19th century. After the Constitutional Revolution of 1906, this simplifying trend was developed further by writers, poets, and political activists, with ‘Ali Akbar Dehkhodā being one of the most prominent ones. Dehkhodā’s style in *Čarand parand* is described by Yusefi (1990:793) as embracing the modes of spoken language, using popular idioms, vocabulary, and manners of expression. This newly formed simple prose style, more closely related to the colloquial language, was to be followed by later writers, including Mohammad ‘Ali Jamālzāde (1892-1997), who wrote the first Modern Persian short story, published in 1921.

The following books/short stories were examined in search for the progressive construction with *dāštan*: the Qarāchadāghi’s translations into Persian of three plays by Akhundoff as appeared in (Rogers, 1890), *Čarand parand* by Dehkhodā, as published in (Dehkhodā, 1983); the total of 34 short stories and 2 excerpts of novels by Jamālzāde written in 1921-1974, consisting of the 6 short stories in *Yeki bud yeki nabud*, first published in 1921, as appeared in (Jamālzāde 1966), 9 short stories and 2 excerpts of novels, originally published in the period of 1942 to 1964, as appeared in (Jamālzāde 1999), the 7 short stories in *talx va širin* first published in 1955, as appeared in (Jamālzāde 1955), the 12 short stories in *qessehā-ye kutāh barā-ye baččehā-ye rišdār* first published in 1974, as appeared in (Jamālzāde 2001); two plays written by Ya‘qubi, *zemeštān-e 66* and *neveštān dar tāriki*, performed respectively in 1998 and 2010, as appeared in (Ya‘qubi 1998) and (Ya‘qubi 2010).<sup>6</sup>

A common characteristic of the texts mentioned above and examined in this research is their close connection to the colloquial language of the time. Soroudi (1993:215) mentions that the plays of Akhundoff use “different levels of the spoken language in the original Turkish, a feature that was emulated in the Persian versions”. She continues that Dehkhodā in his *Čarand parand* “adopted colloquial language and storytelling techniques” (Soroudi 1993:216), and thus laid the foundations of modern Persian prose literature. Jamālzāde, on the other hand, in the preface to his first book, *Yeki bud yeki nabud*, praises the European writers for their simple style which is closely related to the vernacular language of their people, and calls upon the Persian writers to denounce the ornate language of the elite and develop the same simple communicative prose style. As regards the most recent works mentioned above, both Ya‘qubi’s plays, *zemeštān-e 66* with its direct historical references to the Iraq-Iran war (1980-88), and *neveštān dar tāriki* referring to the 2009 Iran presidential election, are meant to depict the society of contemporary Iran. The main characters in both plays represent the youth of Iran today, and both plays were performed in Tehran in the recent years. Considering all these factors, one can make sure that the language of the plays should reflect the colloquial language of contemporary Iranian people, particularly that of the youth.

The data is described in the framework of Vendler’s classification of situations, which distinguishes between “activities”, “accomplishments”, “achievements”, and “states” (Vendler 1967: 97-121).

---

<sup>6</sup> These works are henceforth abbreviated as TPP (*Three Persian plays*), CP (*Čarand parand*), YK (*yeki bud yeki nabud*), CJ (the collection of Jamālzāde’s works appeared in (Jamālzāde 1999)), TSH (*talx va širin*), BR (*qessehā-ye kutāh barā-ye baččehā-ye rišdār*), Z66 (*zemeštān-e 66*), NT (*neveštān dar tāriki*)

#### 4. Results of data collection

Table 1 shows that whereas the progressive forms do not appear at all in the translations of the *three Persian plays* of Akhundoff, which arguably (see above) reflect the Persian language of the late 19th century, the two plays of Ya'qubi, having approximately the same number of words as that of the *three Persian plays* and belonging to the same genre, are filled with these forms, containing 66 cases of progressive present and 13 of progressive past. Although Čarand parand and the short stories of Jamālzāde do not belong to the same genre, Dehkhodā, as mentioned above, adopts storytelling techniques in his Čarand parand, which would justify a comparison between this work and Jamālzāde's collection of short stories. As the table shows, progressive forms appear in the short stories of Jamālzāde five times more frequently than in Čarand parand.

No instance of progressive evidential tense was found in the examined data, and while it is noteworthy that in Čarand parand, which has apparently the earliest attestations of the progressive construction in Persian literature, both progressive present and past are attested, in other works, progressive present clearly outnumber the progressive past.

	Year of publication	Genre	Total number of words (approximately)	Total number of PROG. forms	Frequency of PROG forms <sup>7</sup>
<i>The three Persian Plays</i>	1890	Play	23500	0	0
<i>Čarand parand</i>	1907-1908	Satirical essays	25500	2 1 PRES 1 PT 0 EV	0.008
<i>The collection of Jamālzāde's works</i>	1921-1974	Short story	169000	62 41 PRES 21 PT 0 EV	0.037
<i>Two plays by Ya'qubi</i>	1998, 2010	Play	21000	79 66 PRES 13 PT 0 EV	0.376

Table 1

#### 5. Discussion of the results

The following observations can be made regarding the 35 instances of progressive past found in the examined works:

- All of them are in the indicative mood.
- None of them is negative.
- The distribution of the cases among the four kinds of situations is as follows: 13 cases of accomplishments, 12 cases of achievements, and 10 cases of activities.
- **List of accomplishment situations (13 cases)<sup>8</sup>:** (1) *be šahr āvardan* 'to bring to the city'; (2) *man rā divāne kardan* (2 times) 'to make me mad'; *šekam rā pāre kardan* 'to rip the belly apart'; *yād dādan ke*

<sup>7</sup> This is calculated as (total number of PROG forms/total number of words)\*100.

<sup>8</sup> In the lists that follow, situations are categorized chronologically into three groups; (1), (2) and (3) indicate that the situations appear in Dehkhodā, Jamālzāde and Ya'qubi's works respectively.





*o gu kardan* ‘to have a conversation’; *zendegi kardan* (3 times) ‘to live’; *ziyāde ravi kardan* ‘to exceed one’s limits’; *qadam zadan* (2 times) ‘to stroll’; *towhin kardan* ‘to insult’; *be...saxt gereftan* ‘to treat... severely’; *mahrum kardan* ‘to deprive; (in the context in question) to misbehave’; *jafā kardan* (2 times) ‘to misbehave’; *zabt kardan* ‘to record’; *kār kardan* ‘to work’; *zadan* ‘to beat’; *e’terāf kardan* ‘to confess’.

- **The list of achievement situations (23 cases):** (2) *aql az sar-am paridan* ‘to lose one’s mind’; (*howsele sar raftan* ‘to run out of patience’; *koštan* ‘(in the context in question) to kill by shooting’; *xaffe shodan* (2 times) ‘to choke (intrs.)’; *falaj shodan* ‘to become mutilated’; (*dokān) dar o taxte šodan* (2 times) ‘to go bankrupt’ (used for a business); *mordan* (1 times) ‘to die’; *az hāl raftan* ‘to faint’; *oftādan* (2 times) ‘to fall’; *mahkum kardan* ‘to find guilty’; *residan* ‘to arrive’; *az miyān raftan* ‘to become extinguished’; *bačče šodan* ‘to become a child’; (3) *mordan* (2 times) ‘to die’; *didan* (2 times) ‘(in the context in question) to come to realize’; *qushi rā gozāstan* ‘to hang out the phone’; (SMS) *ferestādan* ‘to send (used for a text)’; *šart-bandi kardan* ‘to bet’;
- **List of accomplishment situations are (21 cases):** (1) *āmadan* ‘to come’; (2) *be qabrestān bordan* ‘to take to the cemetery’; *bordan* (2 times) ‘to take (someone/something away)’; *māsīdan* ‘to accomplish’; *pāšne-ye sabr va howsele-ye man rā az jā kandan* ‘to make me lose my patience’; *mesl-e barf āb šodan* ‘to become melted like snow’; *xarāb kardan* (2 times) ‘to destroy’; (*dandān) dar-āmadan* ‘(teeth) to come out’; (*āftāb) qorub kardan* ‘(sun) to set’; *deqkosh kardan* ‘to kill someone by making them so sad’; *kār-e xod rā kardan* ‘to accomplish one’s job’; (3) *telefon rā vasl kardan* ‘to plug in the phone’; *yād-am āmadan* ‘to come to my mind’; *enteqām gereftan* ‘to take revenge’; *āmadan* ‘to come’; *šam rā xāmuš kardan* ‘to blow out candles’; *qāne kardan* ‘to persuade’; *tahmil kardan* ‘to impose one’s opinion’; *jā oftādan* ‘to settle down’.
- In most cases (60 cases of activities, and 20 cases of accomplishments), the progressive present used with activities and accomplishments describe an ongoing action (as in 7).

(7)

<i>did-im ...</i>	<i>javāni</i>	<i>rā ....</i>	<i>dār-and</i>	<i>mi-āvar-and</i>	<i>be taraf-e</i>	<i>mahbas</i>
see-	young	DIR.OB.MK	have-	bring-	toward	prison
1PL.PERF.PT				3PL.IMPF.PRES	3PL.IMPF.PRES	

We saw that they are bringing a young man to the prison (YK:40).

- As regards the 4 remaining cases of activities and 1 case of accomplishment, which all occur in the most recent works, i.e. the Ya‘qubi’s plays, 2 cases denote a habitual action, (8) and (9), 1 case denotes a near future action (10), 1 case expresses a historic (progressive) present (11), and 1 case denotes an iterative action (12).

(8)

<i>sāl-hā-st</i>	<i>ke</i>	<i>in tu</i>	<i>dār-am</i>	<i>kār mi-kon-am</i>
year-PL.MK-COP.3SG.PRES	that	inside	have-1SG.IMPF.PRES	work-1SG.IMPF.PRES

It is years that I have been working inside (here). (NT:62)

(9)

<i>bišter-e</i>	<i>vaqt-hāke</i>	<i>man</i>	<i>dār-amharf mi-zan-am,</i>	
most –of	time-PL.Mk	when	I have-1SG.IMPF.PRES	talk-1SG.IMPF.PRES

<i>to</i>	<i>aslant</i>	<i>havās-et</i>	<i>be</i>	<i>man</i>	<i>nist.</i>
you	at all	attention-your	to	me	COP.3SG.PRES.NEG.

Most of the times, when I am talking, you don’t pay attention at all. (Z66:42)

(10)

<i>beh-eš</i>	<i>goft-am</i>	<i>mā</i>	<i>dār-im</i>	<i>mi-ri-m</i>	<i>holand</i>
to-him	say	we	have-	go-1PL.IMPF.PRES	the Netherlands
	-1SG.PERF.PT		1PL.IMPF.PRES		

I told him we are about to go to the Netherlands. (NT:35)



(11)  
*vaqti* *čini-hā* *dār-an* *tabbati-hā* *rā...* *mi-zan-an,*  
 When Chinese have-3PL.IMP.F.PRES Tibetan-PL.MK DIR.OB.Mk beat-3PL.IMP.F.PRES

*Yeho* *ye* *Budāyi* *xodesh* *rā* *mi-suzun-e*  
 suddenly one Buddhist himself DIR.OB.MK burn-3SG.IMP.F.PRES  
 When Chinese are beating Tibetans..., suddenly a Buddhist man burns himself. (NT: 60)

(12)  
*Nāhid* *dār-ad* *šam'-hā* *rā* *yeki yeki* *xāmuš* *mi-kon-ad*  
 Nahid have-3SG.IMP.F.PRES candle-PL.MK DIR.OB.MK one one still kardan-3SG.IMP.F.PRES  
 Nahid is blowing out the candles one by one (Z66:14)

- In most cases (20 cases out of 23), the progressive present used with achievements refers to an action in the near future (as in 13).

(13)  
*did* *howsele-ye* *man* *dār-ad* *be koli* *sar mi-rav-ad*  
 see-3SG.PERF.PT patience-of my have-3SG.IMP.F.PRES totally run out-3SG.IMP.F.PRES  
 He saw that my patience is running out (YK:59)

- In the remaining 3 cases of achievements, in 2 cases, the use of the adverbs *rafte rafte* “gradually” (BR:33), and *ruz be ruz* “day by day” (BR:229), together with the progressive present gives the sense of an ongoing action to the achievement situations in question, and in 1 case, the adverb *tekke tekke* “part by part” (TSH:180) gives the sentence an iterative interpretation.

To sum up, the progressive past and present in the data are mainly used to refer to i) an ongoing action, ii) an imminent action. Few cases of the present progressive exhibit three further functions, i.e. habitual, iterative, and historic present. Table 2 shows the distribution of the cases based on their functions.

	ongoing	imminent	habitual	iterative	historic present
Progressive Past (35 cases)	23	12	0	0	0
Progressive Present (108 cases)	82	21	2	2	1

Table 2

Based on Table 2, we could now evaluate Dehghan and Windfuhr’s description of the progressive tenses mentioned above. Dehghan’s account fails to include the imminent function of the progressive past, which constitute one third of the cases in our data, and both his examples, exhibiting the ongoing function of the progressive past, have a reference to a perfective past action, whereas in our data, two third of the cases use the progressive past absolutely. As regards his description of the present progressive, function (a) corresponds to ongoing function of the progressive present, function (b) is non-attested in our data, and function (c), which is described as “very rare”, corresponds to the imminent function of the progressive present, which constitutes %20 of the cases. Our results are in more accord with Windfuhr’s description, where the progressive tenses are said to “disambiguates the progressive and imminent-future functions of the imperfective” (Windfuhr 2009:462). In this description, however, it is not clear what is exactly meant by “imminent-future function of the imperfective”, since in his section on the imperfective past and present, Windfuhr lists four functions, none of them named explicitly as imminent-future.

Establishing the functions of the progressive tenses exhibited by our data, it is now possible to look into the earliest works, especially the *Three Persian Plays*, where there is no instance of such construction, in order to find out what parallel constructions represent ongoing and imminent actions there. It is observed that in most cases imperfective tenses, past and present, are used, as in (14) and (15), while in two cases, (16) and (17), perfective present<sup>9</sup> and past are employed.

<sup>9</sup> Windfuhr (1979:90) recognizes this form, i.e. inflected form of *xāh* “to want” followed by the apocopated infinitive of the main verb, as one of the two forms of the perfective present. In his later book, i.e. (Windfuhr 2009), however, he does not mention this form.

(14)  
*Aruj jelo be-rav-im be-bin-im ki-st mi-āy-ad*  
 Aruj forward go-1PL.SUBJ see-1PL.SUBJ who-COP.3SG come-3SG.IMPF.PRES  
 Aruj, let's go forward, see who is coming. (TPP:93)

(15)  
*bābā-t pošt-e xāne bā čupān-ān harf mi-zad*  
 father-your back-of house with shepherd-PL.MK speak-3SG.IMPF.PT  
 Your father was speaking with the shepherds at the back of the house (PTT:134)

(16)  
*bar xiz az injā bo-ro Divān beygi injā xāhad āmad...*  
 stand up-2SG.IMPV from here go-2SG.IMPV Divān Beygi here come-3SG.PERF.PRES

(Tārrudi pā mi-šav-ad mi-rav-ad. Dar in hāl Divān beygi mi-ras-ad)  
 Tārrudi stand up-3SG.IMPF.PRES go-3SG.IMPF.PRES in this moment divān Beygi arrive-3SG.IMPF.PRES  
 Stand up, go out of here! Divān Beigi is coming here. (Tārrudi stands up, goes. In this moment, Divān Beygi arrives) (TPP:104)

(17)  
*be dād-am be-ras-id ke xaffe šod-am*  
 to cry-my come-2PL.IMPV that choked become-1SG.PERF.PT  
 Come to my help since I am becoming choked.

## 6. The origin of the progressive construction in Modern Persian

So far the attempts to establish the origin of the newly formed progressive construction in Modern Persian have been focused on the use of the verb 'to have' in this construction (Keshāvarz 1962, Dehghan 1972, Vafaeian 2012). Not finding similar progressive constructions with the verb 'to have' in any of the Old and Middle Iranian languages, different dialects spoken in Iran, and foreign languages known to Iranians such as English, French, Russian and German, therefore, the attempts have not been successful in proposing a possible source for this construction. It seems, however, that the more distinctive feature of this construction, other than the choice of the auxiliary verb, is the fact that both the auxiliary and the main verb appear as finite verbs, i.e. both of them are inflected for person and number. This feature distinguishes the progressive construction from other periphrastic verbal constructions of Modern Persian, such as the resultative-stative tenses and the perfective present tense<sup>10</sup>, where only the auxiliary verb gets inflected, and the main verb appears in the form of a participle or apocopated infinitive, as in (18) and (19).

(18)  
*xaride -am*  
 buy-PART COP.1SG  
 I have bought

(19)  
*xāh-am xarid*  
 want-1SG.PRES. buy-APOCOPATED INF.  
 I will buy.

Furthermore, it is observed that in the progressive construction, both auxiliary and the main verb take the same mood, which distinguishes this construction from "modal constructions and subordinate clauses implying potential actions and states" (Windfuhr 2009:457), where the second verb always appears in the subjunctive mood, as in (20).

<sup>10</sup> See footnote 9.

(20)  
*mi-xāst-am*      *xāne*      *be-rav-am*  
 want-1SG.IMP.F.PT    house      go-1SG.SUBJ  
 I wanted to go home (literally: that I go home) (taken from Windfuhr 2009:457)

On the other hand, many instances of similar construction as the progressive one, i.e. clauses which apparently contain two finite verbs sharing the same tense and mood inflection, as in (21)-(24)<sup>11</sup>, are found in the examined works.

(21)  
*gozāšt-i*      *dar raft-i*      *be*      *hend*  
 leave-2SG.PERF.PT    escape-2SG.PERF.PT    to      India  
 You left, escaped to India

(22)  
*mi-rav-i*      *doxtar...*      *rā...*      *nāxoš*      *mi-kon-i*  
 go-2SG.IMP.F.PRES    daughter    DIR.OB.MK    sick      make-2SG.IMP.F.PRES  
 You go, make the girl sick (YK:105)

(23)  
*to*      *bāyad*      *be-šin-i*      *še 'r*      *tarjome kon-i*  
 you      should      sit-2SG.SUBJ      poem(s)      translate-2SG.SUBJ  
 You should sit, translate poems (NT:49)

(24)  
*biy-āy-id*      *be-r-im*      *tu-ye*      *ye*      *panāhgah*  
 come-2PL.SUBJ    go-2PL.SUBJ    inside-of    one    shelter  
 Let us come, go inside of a shelter (Z66:23)

In the absence of a better term to call these constructions, they are henceforth called Serial Verb Constructions (SVC)<sup>12</sup>. Given the fact that the progressive construction with the verb ‘to have’ can be basically considered as a SVC, one might suggest that the grammaticalization of a SVC, whose first verb is formally related to *dāštan*, could be held responsible for the rise of such construction in Modern Persian. In view of this suggestion, I found 7 instances of a SVC with *bar dāštan*<sup>13</sup> as the first verb in the collected data, 1 case in the Three Persian Plays and 6 cases in Čarand parand, 3 of which are shown in (25)-(27). Dehkhodā (1958:858) lists different meanings for this verb, among which ‘azm-e rahil kardan ‘to set off (intr.)’<sup>14</sup>, which has been attested as early as in the 10th century, seems to work more appropriately in the case of our examples.

(25)  
*bar dār-im*      *be-rav-im*      *sar-e*      *morāfe 'e*  
 set off-1PL.SUBJ    go-1PL.SUBJ    head-of      trial  
 Let us set off, go to the trial (TPP:41)

(26)  
*be*      *farangi-hā*      *na-guy-id*      *ke*      *bar dār-and...*      *be-nevis-and...*  
 to      foreigner-PL MK      tell-2PL.IMPV.NEG      that      set off-3PL.SUBJ      write-3PL.SUBJ  
 Don’t tell (this) to the foreigners or they (will) set off, write that ....(CP:183)

<sup>11</sup> In order to make the case clearer, these examples are translated literally.

<sup>12</sup> Sebba (1987:1) states that the authors usually apply Serial Verb Construction “fairly indiscriminately to constructions in which there is a sequence of the form V NP V NP or V NP V, where V is not obviously an infinitive”, a condition which is fulfilled in the case of our examples.

<sup>13</sup> *Bar* ‘up’ is considered by Windfuhr (2009:448) as one of the three most frequently occurring preverbs in Persian.

<sup>14</sup> In this sense, *bar dāštan* is perhaps the shorter form of *rāh bar dāštan* ‘(literally) to take the way’, also listed in (Dehkhodā 1958:159, 227), which has just the same meaning as ‘to set off’.

(27)  
*in bud ke ān-hā ... bar dāšt-and telegrāf kard-and...*  
 this be-3SG.PERF.PT that he-PL.MK set off-3PL.PERF.PT telegraph-3PL.PERF.PT  
 Thus, they set off, telegraphed (that) .... (CP:168).

From the semantic point of view, the progressive constructions with *dāštan* are related to the SVCs with *bar dāštan* in that both can have a reference to an imminent action. As Table (2) shows, the imminent function of the progressive forms account for one fifth of the cases, and it is noteworthy to mention that in the earliest works, i.e. in Čarand parand and the collection of Jamālzāde's works, more than one third of the cases (22 out of 64) express an imminent action. In this regard, one is also reminded of the first account of the progressive construction, given by Zhukovskij, where denoting an immediate future action is mentioned as the only meaning of the present progressive tense.

## 7. Conclusion

Assuming the grammaticalization of SVCs with *bar dāštan* as responsible for the formation of the progressive construction in Modern Persian, the morphological and semantical development of the construction can be described as follows: as the SVCs in question undergo grammaticalization, the verb *bar dāštan* gets shorter by losing the preverb *bar*, and changes into *dāštan*. In the domain of semantic, the original meaning of the SVCs, i.e. denoting an imminent action, is preserved in the emerging progressive constructions, which gradually take on some other functions, more importantly denoting an ongoing action, too.

## References

- Bassārī, T. 1969. *Dastur-e zabān-e Fārsi* [Persian grammar]. Tehrān: Ketābkhāne-ye Tahūrī.
- Dehghān, Iraj. 1972. *Dāštan* as an auxiliary in contemporary Persian. *Archiv Orientální* (Praha) 40: 198-205.
- Dehkhodā, 'Alī Akbar. 1958. *Loqāt nāme* [dictionary]. Tehrān: Tehran University Press
- \_\_\_\_\_. 1983. *maqālāt-e Dehkhodā*, vol.1. Tehrān: Shirāzhe
- Farshidvard, Kh. 1969. *Dastur-e emruz* [grammar of contemporary Persian].
- Jamālzāde, Mohammad Ali. 1955. *talkh va shirin*. Tehrān
- \_\_\_\_\_. 1966. *yeki bud yeki nabud*. Tehrān: Kānūn-e Ma' refat.
- \_\_\_\_\_. 1999. *khāterāt-e seyzed Mohammad 'Alī Jamālzāde* [Memories of Seyyed Mohammad 'Alī Jamālzāde]. Tehrān: Nashr-e Shahāb, Sokhan.
- \_\_\_\_\_. 2001. *qessehā-ye kutāh barāye bachchehā-ye rishdār*. Tehrān: Sokhan.
- Keshāvarz, K. 1962 *mozāre' va māzi-ye malmus* [The continuous present and past]. *Rāhnemā-ye Ketāb* 5. 687-94.
- Lambton, Ann K. S. 1963. *Persian grammar*. Cambridge: Cambridge University Press.
- Lazard, G. 1992. *A grammar of contemporary Persian* (Translation of *Grammaire du persan contemporain*). Costa Mesa, Calif.: Mazda Publishers, in association with Bibliotheca Persica.
- Lorimer, D. L. R. 1916. Notes on the Gabri dialect of Modern Persian. *Journal of the Royal Asiatic Society of Great Britan and Ireland* July.423-89.
- Roger, A. 1890. *Persian Plays: three Persian plays with literal English translation and vocabulary*. London: W. H. Allan & Co.
- Sebba, M. 1987. *The Syntax of Serial Verbs: an investigation into serialization in Sranan and other languages*. Amsterdam; Philadelphia: John Benjamin Publication Company.
- Soroudi, S. 1993. Constitutional Revolution vii. *The Constitutional Movement in Literature*. *Encyclopedia Iranica*, ed. by E. Yarshater, 6.212-16.

- Vafaeian, Gh. 2012. Progressive constructions in Iranian languages. Proceedings of the Doctoral Festival 2010, Department of Linguistics, University of Stockholm. Available at <http://www.diva-portal.org/smash/get/diva2:634997/FULLTEXT01.pdf> (accessed at 15 November 2013).
- Vahidiyān, T. 1963. *Dastur-e zabān-e ‘āmiyāne* [grammar of colloquial Persian].
- Vendler, Zeno. 1967. *Linguistics in Philosophy*. Ithaca, New York: Cornell University Press.
- Windfuhr, Gernot. 1979. *Persian grammar: history and state of its study*. The Hague, New York: Mouton.
- \_\_\_\_\_. 2009. Persian and Tajik. *The Iranian languages*, ed. by Gernot Windfuhr, 416-544. London; New York: Routledge.
- Yarshater, E. 1970. Iran and Afghanistan. *Current Trends in Linguistics*, ed. by Thomas A. Sebeok, 6. 669-89. Paris; The Hague: Mouton.
- Ya‘qubi, M. 1998. *zemestān-e 66*, available at <http://www.takbook.com/1228-biography-film/%D9%86%D9%85%D8%A7%D9%8A%D8%B4%D9%86%D8%A7%D9%85%D9%87-%D8%B2%D9%85%D8%B3%D8%AA%D8%A7%D9%8666/> (accessed at 15 November 2013)
- \_\_\_\_\_. 2010. *neveshtan dar tāriki*, available at <http://www.yaghoubee.com/915NeveshtanDarTaariki/1.play/910219NevDarTaarikiComless.pdf> (accessed at 15 November 2013).
- Yusefi, Gh. 1990. *Charand parand*. *Encyclopedia Iranica*, ed. by E. Yarshater, 4.792-95.
- Zhukovskij, Valentin. 1888. *Osobennoe znachenie glagola dāstān v persidskom razgovornom yazyke* [Special meaning of the verb *dāstān* in spoken Persian]. *Zapiski Otdeleniya Imperatorskago russkago arkeologicheskago obshchestva Vostochnago* 3.376-77.

## Metaphorical approach to mental verbs in Turkmen Turkish

Savaş Şahin  
savassahin06@hotmail.com  
Amasya University

### Abstract

Mental verbs express the verbs related to the cognition, emotion and perception of the mind. They completely express cognitive and mental situations that do not include physical actions. People can use different words while expressing their perceptive outputs in mental or emotional situations. Concept intelligence is the same in all people, however its interpretation and linguistic way of expression is different. The number of mental verbs and metaphoric usage can show differences between cultures.

Metaphors expand the conceptual abstract and concrete usage area of words. They are expressions that show how people evaluate the objects, events and situations existing around and how they perceive the world they live in. The way of expression and perception of schemes may differ from culture to culture. Moreover, metaphors can convert abstract image schemes into concrete concepts. The expansion of the meaning area of abstract mental verbs related to mental area can generally be realized by metaphors.

Keywords: Turkmen Turkish, mental verbs, metaphor

### Introduction

The mind analyzes the meanings of external and internal stimulus, perceptions and emotions and tries to acquire them. The verbs related to cognition are based on thinking. The most important output of a mental process is cognitive verbs. The first stage of the mental processes is acquisition. The duty of acquiring is to make the knowledge coming from different senses reach language centers in brain. So the function of language processes starts with acquisition. The second stage of the mental process of a language is commenting and understanding perceptions. That is, information coming from the different senses is put into operation. At this stage, connections of knowledge coming from outside and knowledge that is coded in brain and the stage of signification is realized. After this stage, the stage of reaction or self-expression starts. This is the third and the last stage of the mental process. We call the verbs regarding cognitions, perceptions and feelings that are occurred as the result of a mental process a mental verb. Mental verbs include many cognitive processes. Mental verbs are about cognitions, perceptions and feelings. Another output of the mental process is feelings. Emotions are a characteristic proper to people. Emotions are occurred as the result of a mental process and the beginnings and the effects of emotions differ from person to person.

Researches (Shatz, 1983) have investigated to explore and understand the world of children's inner mental field (beliefs, thoughts, intentions) and in and internal world. Children's perceptions, emotions and cognition in general, the areas belonging to the words be termed by mental verb or cognitive words (Booth, 1997).

Morley (2000) states that there are only three main processes of verbs i.e. material, mental, and relational. Mental processes are processes of sensing, in that in place of an actor they involve a senser and a phenomenon in processes of *perception* (e.g. see, hear), *affection* (e.g. like, fear) and *cognition* (e.g. think, know).

Lock (2005) states that mental processes are recognized in four subtypes. The first type is *perception* (seeing, hearing, noticing, feeling, tasting, and smelling). The second type-*affection* includes processes such



as liking, loving, admiring, missing, fearing, and hating. The third type, *cognition* includes processes such as thinking, believing, knowing, doubting, remembering and forgetting. The fourth type *volition*, includes processes such as wanting, needing, intending, desiring, hoping, and wishing.

In the research, it has been tried to explain the mental verb concept and mental verbs were analyzed in three types. After that, the examples from Turkmen Turkish were given. In the second part of the research, the metaphor concept and its relation with mental verbs were explained. In the final part of the research, the examples from Turkmen Turkish for the verbs related cognition, emotion and perception verbs.

Since the conceptual area about cognition, emotion and perception is very large, we limited the subject by analyzing only main mental verbs. As analyzing emotion verbs, we examined the main feelings- happiness, sorrow, anger and fear in terms of metaphors. Explaining the mental verbs related to the mental concept area through the metaphor concept related again to the mental area has lead interesting results.

## 1. Metaphor

A metaphor, as defined a figure of speech in which an implied comparison is made between two unlike things that actually have something important in common. The word *metaphor* itself is a metaphor, coming from a Greek word meaning to “transfer” or “carry across.” Metaphors “carry” meaning from one word, image, or idea to another. Metaphors depends on perceptual experience such as sensation and imagery, and sensory experience.

Reddy (1993) drew attention to a cluster of metaphors which motivate a substantial amount of our talk about language and about the role of language in communication (Taylor, 2003). Reddy shows that it is very difficult to talk about communication without using linguistic metaphors from the source domain of physical transfer. He also argue that this is a reflection of our ways of thinking about communication, which are inherently metaphorical.

Taylor (2003) states that the study of metaphor a central topic in Cognitive Linguistics. The basic insight has been that metaphor is not just a manner of speaking. Metaphors are element that get vigor and deep to expression and display variety to another culture from a cultures and are not only means by literal language but that they are actually part of our every-day communication. They have been taken over, mostly, from folk conceptions of language. Metaphors appear through our physical, cognitive, social and cultural, experiences. Also we can say that metaphors are not only part of our every-day language but also of our mind. By using it maybe even better to use metaphorical expressions to explain difficult terms. The study of metaphor has been a central topic in cognitive linguistics.

There are also many different views about the types of metaphors. Lakoff and Johnson called them *orientational metaphors* because of most of their majority’s obligation of being realized with spatial orientation. Metaphors based on physical objects, substances and especially experiences about our own bodies are called ontological metaphors. Container metaphors are examined in three type land areas, the visual field and the metaphors that express events, movements, activities and situations. (Lakoff & Johnson, 1980).

Conceptual metaphor is used in a lingusitic expression, as can be demonstrated by a brief glance at the by now classic list of references Lakoff&Johnson (1980), Johnson (1987), Lakoff (1987), Turner (1987), Lakoff&Turner (1989), and Lakoff (1993) (Steen, 1999).

Conceptual metaphor theory (Deignan, 2005) rejects the notion that metaphor is a decorative device, peripheral to language and thought. Instead, the theory holds that metaphor is central to thought and therefore to language. From this starting points a number of tenets are derived, which are discussed here with particular reference to language. These tenets are: Metaphors structure thinking, metaphors structure knowledge, to abstract language, metaphors is grounded in physical experience, metaphor is ideological.

In the field of cognitive linguistics, conceptual metaphor is usually described as understanding one conceptual domain in terms of another conceptual domain or to put it simply: A IS B. B is called the source domain which is used to describe and structure the target domain A. Between A and B exist further a set of correspondences which are often referred to as mappings, whereby elements of the source domain are mapped onto elements of the target domain. metaphor can be characterized with the formula *A IS B*, where the target domain (a) is comprehended through a source domain (b). This comprehension is based on a set of mappings that exist between elements of a and elements of b. To know a conceptual metaphor is to know this set of mappings. This basis is formulated as the ‘Invariance Hypothesis’ in Lakoff (1989, 1990).

Kövecses state that *conceptual metaphors* can be classified according to the cognitive functions that they perform. On this basis, three general kinds of conceptual metaphor have been distinguished: *structural*, *ontological* and *orientational*. These kinds of metaphor often coincide in particular cases (Kövecses, 2010). *Structural metaphors* map the structure of the source domain onto the structure of the target and in this way allow speakers to understand one domain in terms of another. *Oriental metaphors* have primarily an evaluative function. They make large groups of metaphors coherent with each other. *Ontological metaphors* provide extremely fundamental but very crude understanding for target concepts (Kövecses, 2010).

*Ontological metaphors* are our way of commenting on the stimulus of the external world relying on our experiences. This commenting matter is totally about people’s mental acquisition manner.

The ontological metaphors, are based on the experience with physical objects. Those experiences can be identified and categorised as entities restricted by a surface. So we can categorise those things that normally do not have such a boundary by using those entities. We set up artificial boundaries (Lakoff and Johnson, 1998)

There is another important kind of metaphorical concept, one that does not structure one concept in terms of another but instead organizes a whole system of concepts with respect to one another. We will call these *orientational metaphors*, since most of them have to do with spatial orientation: up-down, in-out, front-back, on-off, deep-shallow, central-peripheral. These spatial orientations arise from the fact that we have bodies of the sort we have and that they function as they do in our physical environment. *Oriental metaphors* give a concept a spatial orientation; for example, HAPPY IS UP. The fact that the concept HAPPY is oriented up leads to English expressions like “I’m feeling *up* today (Lakoff&Johnson, 2003).

## 2. Metaphors Related Cognition Verbs in Turkmen Turkish

It is considered that in Turkmen Turkish compound verbs belonging to the cognitive meaning field are generally used with the words such as “notion”, “heart”, “mind”, “memory”. In Turkmen Turkish the mind is considered as a container We mostly see these metaphors in the metaphoric usage of cognitive verbs: THINKING IS SINKING/ THINKING IS DIVING/ THINKING IS DROWNING/ THINKING IS RETURNING/ THINKING IS MOVING.

In Turkmen Turkish the words having directional and motional meanings are frequently used in the metaphoric usage of cognition verbs. In Turkmen Turkish the words “to think”, “to sink”, “to fall down” “to turn” express movements that go down. In these verbs there are the metaphors- THINKING IS FACING DOWN/ THINKING IS MOVING. According to Lakoff & Johnson what is rational and conscious is up and what is unconscious is down. RATIONAL IS UP; EMOTIONAL IS DOWN / CONSCIOUS IS UP; UNCONSCIOUS IS DOWN (Lakoff&Johnson, 1980). This situation conflicts with Turkmen Turkish. Because in Turkmen Turkish what is conscious and rational is down. I think Turkmen people have considered the sky as a holy thing since ancient times. God lives up and he thinks better than people. But a human being lives down and he cannot think better than God. The thought is granted to the human being living down, on the Earth by the God living up.

The verb “to pore” is consisted of the words “Notion” and “go”. The verb “go” expresses going far away. When a person concentrates on something and thinks he isolates himself from the world and loses him-

self. There are the metaphors of THINKING IS PASSING OUT/ THINKING IS GOING AWAY.

The verbs *pikir deryasında gark bolmak*, *pikir deryasında yüzmek* have also the same meaning as “to think”. *Pikir* means “thought”, *deryada gark bolmak* means “to drown in the sea” and *deryasında yüzmek* means “to swim in the sea”. In these verbs the metaphors- MIND IS LIQUID/MIND IS A SEA/THINKING IS SWIMMING are seen.

The verbs *yada salmak*, *aklına getirmek*, *yadına düşmek*, *gövnüne düşmek*, *gövnüne gelmek* are the verbs expressing “remembering”. In these verbs the words “mind” and “heart” are used as containers, it means as matters. With the verbs “to insert”, “to bring” and “to fall down”, the time realizing abstract mind concept’s remembering, in other words it is assumed that the mind which is outside again comes to the container. In these verbs we see the metaphor REMEMBERING IS AN ACTION GOING TOWARDS THE MIND.

In the verb *yürege düşmek* the verb *düşmek* expresses the action of going downwards. In this verb the heart is also considered as a container and thinking that abstract mind concept falls down to the container it is concretized and used in the meaning of “wish”. There are WANTING IS AN ACTION GOING TOWARDS THE MIND/ WANTING IS FALLING DOWN metaphors in this verb.

In the verb *aklıñ kesmek* “to understand” the verb *kesmek* is in the meaning of “to cut” and the mind is considered as a body organ. It is thought that the mind is a sharp object and a sharp thing is strong and dangerous. Thought is like a sharp thing, strong and dangerous. In this verb the metaphors THE MIND IS A BODY/ THINKING IS DANGEROUS AS THE SHARPNESS are seen.

### 3. Metaphors Related Emotion Verbs in Turkmen Turkish

Emotions are considered with inner “feelings” and are regarded as mental phenomena with their essence being of subjective or introspective characters. People need to respond emotionally to the events, situations or effects that happen around them and affect them. We predicted that people would be more likely to use metaphors and metaphorical comparisons when describing how they felt when they were experiencing an emotion than when describing what they did when they experienced it.

People are in fact especially likely to use metaphors when describing emotion, and use more metaphors for intense emotions than for mild emotions. Thus, metaphors are particularly useful in conveying subtle nuances of emotional experience (Gibbs et al., 2002). Researchers (Larsen & Fredrickson, 1999) have used the term emotion in many different senses, without a widely adopted definition.

The emotion concepts that have received attention from a variety of scholars in this tradition include anger, fear, happiness, sadness, love, lust, pride, shame, and surprise. Anger is perhaps the most studied emotion concept from a cognitive semantic point of view. Kövecses (1986) and Lakoff and Kövecses (1987) found a number of metaphorical source domains that characterize anger. Some further ones can be added to those source domains, such as AN ANGRY PERSON IS A FUNCTIONING MACHINE and ANGER IS A SOCIAL SUPERIOR.

In English in expressing feelings, the happiness is up and the sorrow is down: HAPPY IS UP; SAD IS DOWN “*I’m feeling up. That boosted my spirits. My spirits rose. You’re in high spirits. Thinking about her always gives me a lift. I’m feeling down.etc.*” (Lakoff & Johnson, 2003).

Figurative expressions of specific emotions reflect aspects of the bodily experience those of emotions (Kövecses, 1990). Consider the examples of anger and fear, two emotions that vary greatly in their physical experience, as well as their conceptual structure. When angry, for example, people talk about letting off steam, losing their cool, being ready to explode, and so on. For those expressions, HEAT OF FLUID IN A CONTAINER is the source domain of the metaphor for which various “entailments” (Lakoff & Johnson, 1980) follow. In addition to physical sensations, people also use spatial terms to describe emotions. For instance, the basic orientation of the human body in space (certain things are up or down, relative to the body) is used when

metaphorically talking about feeling up or down. These expressions are correlated with what goes on with the human body when one feels a certain emotion: An upright, relaxed posture when happy, versus a slumped, drooping posture when depressed. In addition to, the subject of much discussion for metaphor theorists, who argue that many of the terms used to talk about emotions are mapped from the domain of temperature (Kövecses, 2000,2002).

### 3.1 Metaphors Related Sorrow Verbs in Turkmen Turkish

In Turkmen Turkish, most of the verbs expressing sorrow are idioms and there are not any prepositions in the verbs. The verbs formed by metaphoric meanings and usages are perceived under the heel of a word in itself in terms of their meanings (Erdem, 2003). Since idioms are the expression ways of a nation's perception and interpretation manner of the world, it is also abundant in metaphors.

A person's feeling of desperation against the events and situations he faces in unexpected time may be described as the feeling of sorrow. Besides, sorrow can be used for the situations when unwanted creatures, places or conditions make an active and cheerful person quiet, withdrawn and calm. Sorrow may as well be felt as the result of a physical destruction.

The verbs *Yüregi gıyılmak*, *gövni gıyılmak*, *yüregi yaralanmak*, *yüregi gısmak*, *yüregi yanmak*, *yürek daglamak*, *gövni çökmek* are some verbs in Turkmen Turkish which express sorrow. The word "Yürek" means "heart". In the verbs *yüregi gıyılmak*, *yüregi gısmak*, *yürek daglamak*, which express sorrow, HEART IS SUBSTANCE, the top metaphor, is observed. The words *gıyılmak* (shatter) *gısmak* (narrow), *giñemek* (enlarge), *daglamak* (burn), *yaralanmak* (get hurt), *yanmak* (burn) also express characteristics of a matter.

Considering the indicators that meet a sorrow concept and the words they are with in collocations in Turkmen Turkish, we can examine the metaphor SORROW IS A PHYSICAL DAMAGE under four subtitles: GETTING SAD IS GETTING INJURED/GETTING SAD IS CRUMBLING/GETTING SAD IS BURNING/GETTING SAD IS BEING BROKEN/GETTING SAD IS GETTING NARROW/GETTING SAD IS SQUASHING.

In the compound verbs *gövni çökmek ve gama batmak* "getting extremely sad", the verbs *batmak* and *çökmek* express a decreasing movement that goes downwards. The words *gam* and *gövni* are abstract nouns. With the verbs "to sink" and "to collapse", the feeling of sorrow has concretized. SORROW IS A MOVEMENT GOING DOWNWARDS/ SORROW IS SINKING are risen from this verb. The verb "gama batmak" (sink) takes us to the metaphor, SORROW IS LIQUID. The word "govun" is an abstract word in the verb *Gövni sınmak*. The verb "sınmak" is a concrete verb in the meaning of "to be broken" and it takes us to the metaphor SORROW IS BEING BROKEN.

### 3.2 Metaphors Related Happiness Verbs in Turkmen Turkish

Happiness is feeling relaxed bodily, spiritually, physically and mentally. It is a positive feeling coming up as the result of situations such as a person's getting rid of the problems, having the events in his environment that realize on his side, his requirements' being met and becoming in a good situation in his environment.

Kövecses illustrated (2004) happiness metaphors in his work "Metaphor and Emotion" in this way: HAPPY IS UP/BEING OFF THE GROUND/ BEING IN HEAVEN/LIGHT/ VITALITY/ WARM/ HEALTH/ LIVES WELL/ PLEASURABLE PHYSICAL SENSATION/ CONTAINER/ CAPTIVE ANIMAL/A RAPTURE/ INSANITY/ NATURAL FORCE.

In the mental verbs related happiness it is observed that the words *gövni* (soul) and *yürek* (heart) are used very often just as in sorrow verbs. In Turkmen Turkish in the indicators meeting happiness verbs the



orientational metaphor BEING HAPPY IS BEING FAR AWAY THE GROUND is generally observed. In the verbs such as *gövünden turmak*, *gövün galkmak*, *gövnüň göge yetmek*, *gövnün göge uçmak*, *gövnünü götermek*, the verbs *turmak*, *galkmak*, *göge yetmek*, *göge uçmak*, *götermek* express rising up. The sky is valuable for Turkish people, because it is holy for them. That's why happiness and abundance indicators are up. The verbs *gövnüň doymak*, *gövnü suv içmek*, are some of the remarkable verbs related happiness in Turkmen Turkish. The verb “doymak” in these verbs express getting enough food and the verb “suv içmek” expresses drinking water. A person who is full feels happy. Heart's being full and its drinking water make a person spiritually happy. The metaphor HEART AND BODY ARE PHYSIQUE is seen in these two verbs. In addition, we see the metaphors HAPPINESS IS BEING FULL/ HAPPINESS IS DRINKING WATER.

In the verbs *yüregi giňemek*, *gövnüň giňemek*, *gövün açmak* the verbs *giňemek*, *açmak* express expansion. In these verbs, the metaphor HAPPINESS IS EXPANSION is observed.

The compound happiness verbs “beğenç atına atlanmak” is used in the meaning of flying to the sky because of happiness. The word *beğenç* is in the meaning of liking. We think that the reason for using the verb “ata atlanmak” in this verb is that there is a big role of horses in Turkish people's lives. Turkish people are happy when they are on horses and they like riding horses. A horse is a fast-running animal. A person riding a horse goes as if he was flying. He loses himself because of the happiness. In the verb “beğenç atına atlanmak” there is the metaphor HAPPINESS IS RIDING A HORSE.

### 3.3 Metaphors Related Anger Verbs in Turkmen Turkish

Generally, anger is an emotional reaction given to unmet expectations, unwanted results and unsaturated desires. Like other feelings, anger is possibly corrective feeling for the interpersonal relations when it is expressed completely naturally, universally and healthily. Some of the verbs expressing anger in Turkmen Turkish express the physical situation of a person besides his mood.

The verbs *gazaba çıkmak*, *gahar- gazabı depesine çıkmak* expressing the feeling of anger in Turkmen Turkish the verb “çıkma” is a verb that describes rising up. The meaning of the word “gazab” is anger. In these verbs there are the metaphors ANGER IS AN UPWARDS MOVEMENT/ANGER IS RISING. We can conclude from these verbs that the direction of the feeling of anger is up in Turkmen Turkish. This situation conflicts with the metaphors of English which are CONTROL IS UP; LACK OF CONTROL IS DOWN (Kövecses, 2010) and HAVING CONTROL OR FORCE IS UP; BEING SUBJECT TO CONTROL OR FORCE IS DOWN (Lakoff & Johnson, 2003) In Turkmen Turkish, the direction of the feeling of uncontrolled anger is upwards.

In the idioms such as *gazaba müňmek*, *gahara münmek*, *gazap atına atlan*, there are the metaphors ANGER IS A HORSE/ GETTING ANGRY IS RIDING A HORSE. It is possible to reach the top metaphor of ANGER IS A WAR from these metaphors. Anger is a reason for a war. The metaphors of ANGER IS A HORSE/ ANGER IS RIDING A HORSE that are seen in these frequently used idioms among Turkmen people today, not only show that Turkmen people ride horses and go to wars when they get angry, but also give information about the action areas that Turkish/Turkmen society structure relate with horses and about Turkish war techniques (Erdem, 2003).

In the verbs *gaharını getirmek*, *gaharı gelmek*, the verbs “to come” and “to bring” express a correct direction to us. People damage mostly themselves when they get angry. Therefore, the direction of the anger is towards us. The metaphors ANGER IS A MOVEMENT COMING TOWARDS US/ ANGER'S DIRECTION IS TOWARDS OURSELVES are observed in these verbs.

The verb *gızmak* in the verbs *ganı gızmak* expresses the temperature of a matter. There is the metaphor ANGER IS HOT in this verb.

In the verb “yüregini çışirmek” the verb “çışirmek” describes expansion. When we get angry and take a breath, our stomach area expands. Considering this situation, a relation is made between anger and expansion. There is the metaphor ANGER IS EXPANDING in this verb.

In Turkmen Turkish, anger is considered as a hot and melting object. In the verb *Yüregi eremek* yürek “heart” is assumed as a container and anger is considered as a melting matter in it. There are the metaphors ANGER IS A MELTING OBJECT/ ANGER IS MELTING in this verb.

### 3.4 Metaphors Related Fear Verbs in Turkmen Turkish

Fear is a feeling felt against a possible danger that threatens the safety. The development of the feeling of fear is proportional with the development of mind. When a human starts to become conscious, the feeling of fear starts by the organism’s being stimulated. In the verbs *hovsala düşmek*, *hovsala salmak*, *yüreği agzından çıkmak* which are related the feeling of fear in Turkmen Turkish, *düşmek* describes a downwards movement, *salmak* describes a forward movement and *çıkmak* describes an outwards movement. In the verb *yüreği yarılmak* the verb *yarılmak* describes squashing.

Considering the indicators that meet a fear concept and the words they are with in collocations in Turkmen Turkish, we can examine the metaphors that are similar to the other emotion verbs: FEARING IS FALLING/FEARING IS MOVING FORWARD/FEARING IS MOVING OUTWARDS/FEARING IS CRUMBLING.

## 4. Metaphors Related Perception Verbs in Turkmen Turkish

A function of a sensation that is to make a meaning becoming integrated with other sensations and memories understands. Impression and information created by the function of understanding is a meaning and a perception. Sensations are easy meanings and nominative conscience. Acquisition verbs constitute a part of a mental process that is related to senses. Perception verbs are the verbs about touching, hearing, smelling, tasting and seeing (Şahin, 2013).

Perception is a person’s signification of the stimulus around him. These stimulus are in a certain place. Therefore, we can generally express the verbs related perception by the visual field metaphors. The visible visual field metaphors are the metaphors that are the results of conceptualization of the field-of-view as a creature having a coverage area and specifying sighted creatures as creatures located in the field-of-view. This kind of metaphors are natural metaphors rised by the determination of the limit of the land, in other words, visible parts of the field-of-view while looking at any land (Lakoff-Johnson, 2003). In opposition to place-area metaphors, visual field metaphors are the metaphors which do not have any certain geographical borders and whose borders are determined by the perception of eye. In Turkmen Turkish in the mental verbs related hearing and seeing, direction is down, froward or stil in metaphoric usages.

In Turkmen Turkish compound verbs formed with the noun “eye” are used as a cognition verb expressing the feeling of liking such as in the verb *goze yakmak*, the feeling of sorrow such as in the verb, comprehension such as in the verb *göz önüne getirmek*. Compound verbs formed with the noun “eye” mostly describe a mental situation.

In Turkmen Turkish the main verbs about seeing are the verbs *görmek* and *seretmek*, however a number of compound verbs are used. Eye is a movable concrete organ, therefore in the idioms related seeing the reason for the usage of the action verbs such as “*aylamak*, *döndürmek*, *çevirmek*, *salmak*, *atmak*, *fırlatmak*, *gezmek* are that eye is a movable organ. We can walk and move our glances in outer space in a comfortable way. In the verbs related seeing, the source area is an organ and the target area is a movement. The organ eye is scheduled on movement. This scheme takes us to the top metaphor of SEEING IS THE MOVEMENT OF EYE.



In the verb *göz düşmek*, *düşmek* expresses a movement towards the ground and it takes us to the sub-metaphor SEEING IS EYE'S DOWNWARDS MOVEMENT.

In the verbs *göz aylamak*, *gözünü gıtağını aylamak* the verb *geçen aylamak* means “to rotate”, “to turn” and it takes us to the sub-metaphor SEEING IS TURNING. In the verbs *gözeg “bakış” salmak*, *nazar “bakış” salmak* SEEING IS EYE'S LEAPING UP is emerged.

The verbs “*salmak*, *atmak*, *fırlatmak*” that create compound verbs with the noun “eye” in the mental verbs related seeing, they form compound verbs with the noun “ear” in verbs related listening and metaphoric usages are observed.

The main meaning of the verb *gulak salmak* describes an conscious hearing situation. The Source field is an organ and the target area is a movement in mental verbs related hearing. Ear is a substance. The metaphors LISTENING IS A FORWARD MOVEMENT OF EAR/ LISTENING IS A FORWARD MOVEMENT can be risen.

The verb *gulak kabartmak* describes a conscious hearing situation. In the verb *gulak kabartmak* the verb *kabartmak* is a verb expressing expansion. The metaphor of LISTENING IS LOATING THE EAR is seed.

The verb *gulaga ilmek* is a mental verb which means “to hear”. The verb *ilmek* means “to fall down”. In the verb *gulaga gelmek*, there is a meaning of unwanted information's falling down on our ears. As the information that we do not want to hear and that we hear involuntarily, the direction is down. HEARING IS INFORMATION'S FALLING DOWN ON OUR EARS/ UNWANTED INFORMATION IS A THING FALLING DOWN metaphors are risen.

## Conclusion

The mind utilizes mental verbs for the purpose of running, arranging and maintaining its relation with the environment in the process starting with perception. Mental verbs in Turkmen Turkish are classified in three types which are cognition, emotion and perception verbs. In Turkmen Turkish it is possible to explain the mental verbs such as “to think”, “to understand”, “to know”, “to fear”, “to hear” and ect. with more than one word. The remarkable point in mental verbs in Turkmen Turkish is that the number of main mental verbs is little and the number of compound mental verbs is much.

Abstract mental verbs are wanted to be uttered by concretizing them in order to add power and beauty to the expression. These mental verbs are mostly expressed through metaphors. In Turkmen Turkish, abstract mental verbs have gained a metaphoric phraseology by simile and concretization. While concretizing and similing, elements from Turkish culture were used. Cultural differences between societies have a big role in the differences between the number of words about the mind among languages. In addition, their way and style of perception of the nature and the world is the most important factor that determine this difference. In Turkmen Turkish, we see that the compound verbs that belong to the cognitive meaning area are generally used with the words such as “notion”, “*heart*”, “mind”, “memory”. The words that have direction and action meaning are often used in the metaphoric usage of the mental verbs in Turkmen Turkish like “to sink”, “to fall down”, “to dive”, “to turn”. We mostly see these metaphors in the metaphoric usage of cognition verbs : THINKING IS SINKING/ THINKING IS DIVING/ THINKING IS RETURNING/ THINKING IS MOVING .

Regarding the indicators meeting a sorrow concept in Turkmen Turkish and the other words that they are together in sorrow indicators, we can examine the conceptual metaphor “sorrow is a physical damage” under seven subtitles: BEING SAD IS GETTING HURT/ BEING SAD IS CRUMBLING/ BEING SAD IS BURNING / BEING SAD IS BEING BROKEN/ BEING SAD IS GETTING NARROW/ BEING SAD IS SQUASHING.

It is considered that in mental verbs related happiness, the words “soul” and “heart” are used very often just as in sorrow verbs. As the sky is holy and valuable for Turkmen people, happiness and abundance indicators

are up. That's why we see BEING HAPPY IS BEING FAR AWAY THE GROUND. Because Turkmen people used to live in moorlands and to fool around there, in mental verbs the metaphor HAPPINESS IS EXPANSION is seen.

The metaphors ANGER IS MOVING UPWARD/ ANGER IS RISING is seen in the metaphoric usage of the anger verbs. This situation forms conflict between the metaphors in English CONTROL IS UP; LACK OF CONTROL IS DOWN. Metaphors of direction are also often used in the metaphors expressing the fear. At the same time it is seen that the words related physical breakup are also used often: FEARING IS FALLING DOWN / FEARING IS MOVING FORWARD / FEARING IS MOVING OUTWARD/ FEARING IS SQUASHING.

We see that in Turkmen Turkish there are mostly the metaphoric usage of the verbs related seeing and hearing in perception verbs. The fact that there are the usage of action verbs such as "to rotate", "to throw", "to travel" is because the eye is a moving organ. And it takes us to the metaphor SEEING IS THE EYE'S MOTION.

### References

- Booth, J. R. & Hall W. S. & Robinson, G. C., Kim, S.Y. (1997). Acquisition of The Mental State Verb Know by 2- to 5-Year-Old Children. *Journal of Psycholinguistic Research*, 26(6).
- Deignan, A. (2005). *Metaphor and Corpus Linguistics*. Philadelphia: John Benjamin Publishing Company.
- Erdem, M. (2003). *Türkmen Türkçesinde Metaforlar*. Ankara: Köksav.
- Gibbs, Raymond W., Jr. (1990). Psycholinguistic Studies on The Conceptual Basis of Idiomaticity. *Cognitive Linguistics*, 1, 417-451.
- Kövecses, Z. (1990). *Emotion Concepts*. New York: Springer Verlag.
- Kövecses, Z. (1995). *The Container Metaphor of Anger in English, Chinese, Japanese and Hungarian*. Radman, Zdravko, ed. From a Metaphorical Point of View.
- Kövecses, Z. (2004). *Metaphor and Emotion, Language, Culture and Body in Human Feeling*. Paris: Cambridge University Press.
- Kövecses, Z. (2005). *Metaphor in Culture*. New York: Cambridge University Press.
- Kövecses, Z. (2010). *Metaphor By Live*. New York: Oxford University Press.
- Lakoff, G. (1999). *Philosophy in The Flesh. The Embodied Mind and Its Challenge to Western Thought*. New York: Basic Books.
- Lakoff G., Johnson M. (2003). *Metaphors We Live By*. London: The university of Chicago Press.
- Lakoff, G. (2006). The Contemporary Theory of Metaphor. *Cognitive Linguistics Research* 34. (Honorary Ed. Langacker, R.W.) Berlin, New York: Mouton de Gruyter.
- Larsen, R. J., & Fredrickson, B. L. (1999). Measurement issues in emotion research. In D. Kahneman, E. Diener & N. Schwarz (Eds.). *Well-being: The Foundations of Hedonic Psychology*, 40-60, New York: Russell Sage Foundation.
- Lock, G. (2004). *Functional English Grammar*. New York: Cambridge University Press.
- Morley, G. D. (2000). *Syntax in Functional Grammar*. An Introduction to Lexicogrammar in Systemic Linguistic. London, New York: Continuum.
- Taylor, J., Cuykens, H., Dirven, R. (2003). *New Directions In Cognitive Lexical Semantic Research*. Cognitive Approaches to Lexical Semantics. Berlin, New York: Mouton de Gruyter.
- Shatz, M. & Wellman H. M. & Silber, S. (1983). The acquisition of mental verbs: A systematic investigation of the first reference to mental state. *Cognition*. 14 (1983) pp. 301-321. University of Michigan.
- Steen, G. (1999). From Linguistic To Conceptual Metaphor In Five Steps. *Metaphor In Cognitive Linguistics Selected Papers From The 5th International Cognitive Linguistic Conference Amsterdam 1997*. Amsterdam: John Benjamins Publishin Company.
- Şahin, S. (2012). Mental Verb Term and Mental Verbs in Turkmen Turkish. *International Journal of Turkish Literature Culture Education*. Volume 1/4, pp. 45-62.

# Implementing a Machine-Readable Grammar of Uyghur (UIG)

Alexander Sugar  
*University of Washington*

## Abstract

Creating a machine-readable grammar for the Uyghur language presents various challenges. The portion of Uyghur grammar under discussion was based on code provided by the LinGO Grammar Matrix, which uses a syntactic formalism based on Head-Driven Phrase Structure Grammar (HPSG). HPSG sentences are considered “well formed” when they satisfy a series of phrasal and lexical constraints. Most information about a word—its semantic properties, potential for morphological changes, and the types of words or phrases with which it combines—is stored in lexical entries. A test suite based on authentic Uyghur texts and native speaker judgments led to the formulation of lexical types and lexical and phrasal rules, all organized into type hierarchies. Specific grammatical properties of Uyghur, including case marking and verb morphology, were modeled using HPSG.

Keywords: Uyghur, HPSG, grammar engineering

## O. Introduction

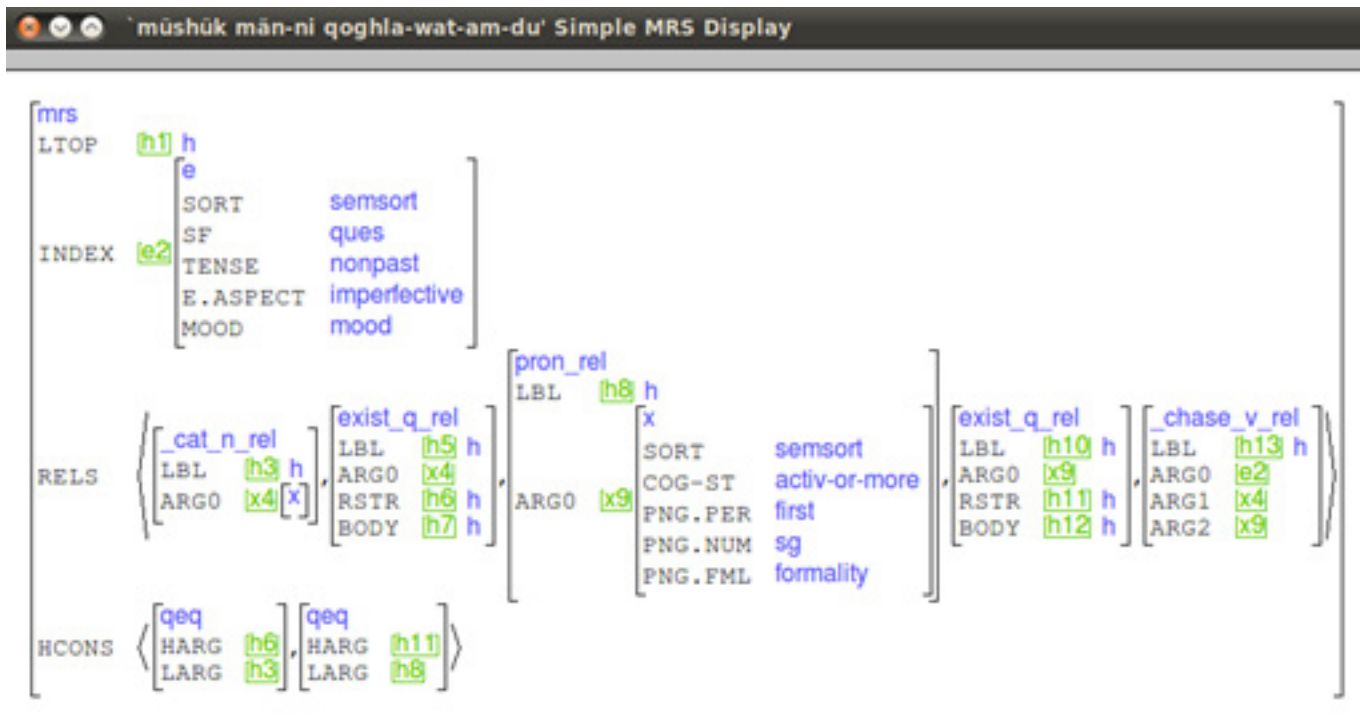
While the syntax of Uyghur has received some scholarly attention over the last decade (Asarina 2006, Bridges 2008, Engesæth et al 2010 among others), it has yet to have been studied, to this author’s knowledge, through the formalism of Head-Driven Phrase Structure Grammar (HPSG). The Uyghur language has also been underrepresented in linguistic technology innovations like search engines and translation software. This paper describes the beginning of a project aiming to address both of these deficiencies by building a machine-readable or precision grammar of Uyghur. Section I describes the formalism and methodology of the project, including A) grammar engineering, B) the Grammar Matrix and C) HPSG. Section II includes examples of notable phenomena of Uyghur syntax and morphology and how they were dealt with in HPSG. Section III lists current coverage of the grammar and concludes with a look to future work.

## I. Formalism and Methodology

### A. Grammar Engineering

Grammar engineering is the process of making formal rules of linguistic knowledge precise enough that they can be encoded on a computer (Bender et al 2011). In other words, the grammar engineer writes into a series of files all of the phrasal rules, lexical rules, lexical entries etc. that a computer will need to accurately process a given language. Processing here means both parsing and generating. A computer should be able to successfully parse grammatical sentences, connecting strings of words to only their correct semantic representations; it should fail to parse ungrammatical strings because they violate specific grammatical constraints. In the case of Uyghur, a computer should be able to take the string “müshük meni qoghliwatamdu” (is a cat chasing me?) and produce a semantic representation with a chasing event in progressive/imperfective aspect. The chaser is a cat, the thing being chased is first person singular, and the speaker is asking a yes-or-no question about the

event's occurrence. This semantic representation is shown below using a Minimal Recursion Semantics (MRS) graphic (Copestake et al 2005).



Conversely, a computer should be able to produce one or more sentences that convey the desired meaning in a language given a target semantic representation. A computer is able to perform these tasks by having a model of the linguistic knowledge possessed by a speaker of a given language.

While there is an interest in comparing knowledge bases of distinct languages to see where they overlap and differ, the process of grammar engineering is bottom up. It begins by focusing only on the facts available about a specific language, with the collection of a test suite. The test suite consists of grammatical and ungrammatical sentences from primary documentation of the language, ideally vetted by native speakers. The test suite should cover a broad range of linguistic phenomena from basic word order, to tense and person marking, to embedded clauses. Sentences are listed in a file with information about their source, grammaticality, phenomena exhibited, English glosses and rough translations, as in the sample entry below.

```

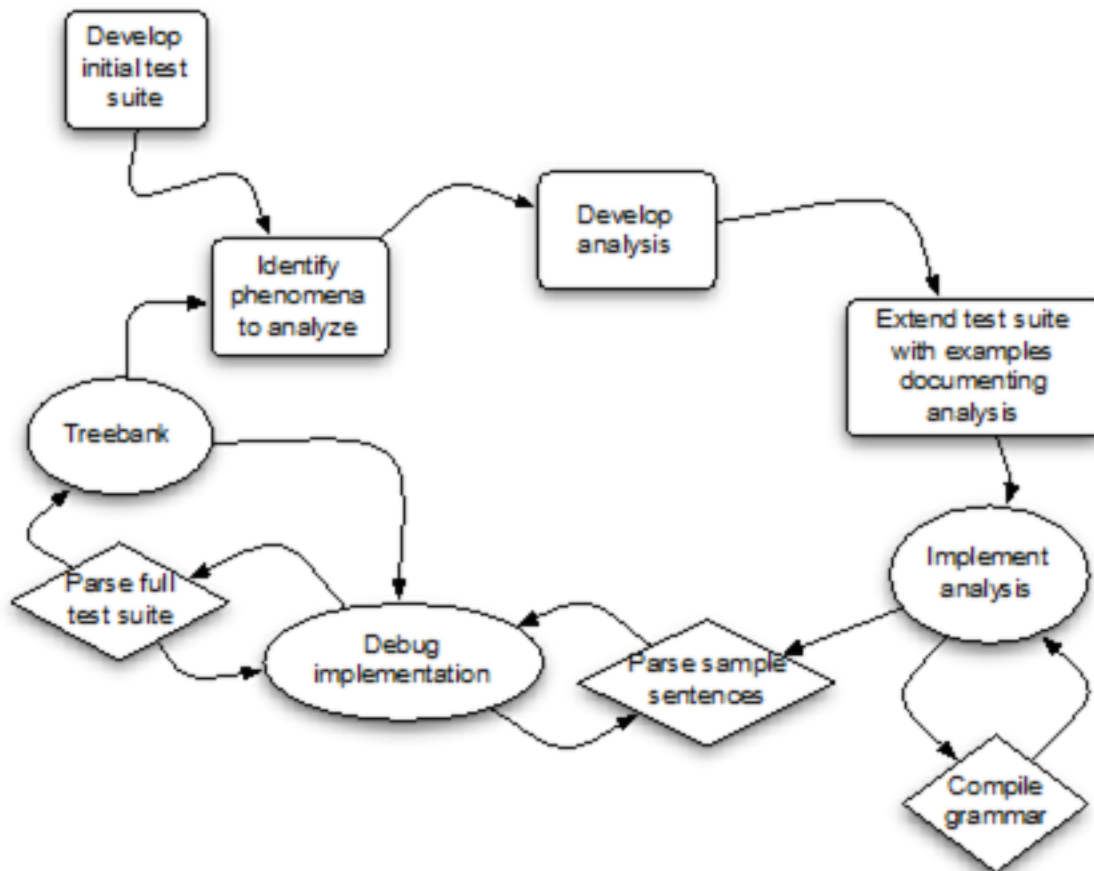
#1 SOV word order
Source: elicited
Vetted: t
Judgment: g
Phenomena: wo
Mān kitab oquymān
Mān kitab oqu-y-mān
1SG.NOM book read-NPST-1SG
'I will read a book.'

```

The grammar engineer will then formulate phrasal and lexical rules that account for the grammatical sentences while ruling out the ungrammatical sentences. Progress in accurately covering test suite sentences can be monitored using software like [incr tadb()] (Oepen 2001). Finally, the grammar engineer will evaluate the coverage of the grammar over a test corpus, a collection of authentic text in a given language held out from the initial test suite.

Testing a grammar over sentences in test suites and corpora, the linguist will inevitably need to add or remove constraints from existing rules, and even write new ones. Because grammars are written to cover all the linguistic knowledge of languages, rather than targeting certain aspects of language, any change made to a part of a grammar can affect the way it treats other aspects of a language. Thus, an important part of the grammar engineering process is reanalyzing coverage after each change is made to the grammar, to see whether a change made with one problem in mind has caused a new problem or even solved what seemed to be an unrelated problem. A single rule governing the relationship between heads and complements, for example, can apply to both verb phrases and preposition phrases.

At various stages of achieving coverage, the grammar engineer may wish to mark the preferred analyses of linguistic phenomena by treebanking before finding new phenomena to analyze. This cyclic workflow is summarized below (from Bender et al 2011).



### B. The Grammar Matrix

Bender et al (2010) devised a starter kit known as the Grammar Matrix to assist in the building of precision grammars. For the linguist user, the Matrix takes the form of an online questionnaire (located at <http://www.delph-in.net/matrix/customize/matrix.cgi>) to be filled out with information about a specific language. The Matrix attempts to capitalize on cross-linguistic commonality by allowing many encoded linguistic rules to be shared across grammars (Bender 2008). Where languages may vary, the Matrix is divided into “libraries” each devoted to a different linguistic phenomenon. So far phenomena addressed by Matrix libraries include negation, adjectives and coordination. The linguist user is given a series of options from which to select and in some cases may be asked to provide specific instances of lexical types within a library to best represent their language. The image below shows a first person singular pronoun being defined in the matrix.



▼ 1SG-pronoun (noun2)

✕ **Noun type 2:**

Type name:

Supertypes:  ▼

Features:

✕ Name:  Value:  ▼

✕ Name:  Value:  ▼

For nouns of this type, a determiner is  obligatory  optional  impossible

Stems:

✕ Spelling:  Predicate:

After completing the questionnaire, the user can download files documenting their choices and the rules produced by the Matrix in Type Description Language (TDL) (Copestake 2002). While hand editing of the TDL may be needed as more facts emerge about the language one wishes to model, the Grammar Matrix provides a significant jump-start to the grammar engineering process.

### C. HPSG

The TDL files produced by the Grammar Matrix and used in the grammar engineering process follow the formalism of HPSG (Pollard and Sag, 1994). HPSG is fundamentally a constraint-based grammar, in which each word is stored as a lexical entry inheriting a specific set of constraints according to its type. Lexical types are organized into type hierarchies so that constraints shared by multiple types need not be restated on each entry. In order to be realized as a word in a sentence, a lexical entry must go through one or more lexical rules, which add information about the word and its behavior in a sentence, and may also add an affix.

A noun in the Uyghur grammar like *alma* (apple) is a stem of a type called normal-noun-lex. This lexeme type inherits constraints from a broader type noun-lex (including the constraints that it needs case inflection and can take a determiner), and adds the constraints that it is always third person and must be inflected for number via a lexical rule.

```

alma := normal-noun-lex &
  [ STEM < "alma" >,
    SYNSEM.LKEYS.KEYREL.PRED "_apple_n_rel" ].
normal-noun-lex := noun-lex &
  [ SYNSEM.LOCAL.CONT.HOOK.INDEX.PNG.PER 3rd,
    INFLECTED.NUM-FLAG - ].

```

*Alma* must go through a number lexical rule before entering a sentence as a word, and it has two options. The singularnoun-lex-rule and the pluralnoun-lex-rule will mark the noun's NUM value as singular or plural, respectively. Both rules inherit the instruction to set the NUM-FLAG as +. While the singularnoun-lex-rule inherits from a rule licensing constants (adding no morphemes to the stem), the pluralnoun-lex-rule inherits from a rule licensing inflections, and is associated with a rule adding the plural suffix *-lar*. Both rules are shown below.

```

singularnoun-lex-rule := number-lex-rule-super & const-lex-rule &
  [ SYNSEM.LOCAL.CONT.HOOK.INDEX.PNG.NUM singular ].

pluralnoun-lex-rule := number-lex-rule-super & infl-lex-rule &
  [ SYNSEM.LOCAL.CONT.HOOK.INDEX.PNG.NUM plural ].

```



Words are put together through phrase structure rules, which specify which types may combine, and what each constituent word contributes to the phrase it helps form. Like lexical types, lexical and phrase structure rules are organized into type hierarchies such that the most widely shared constraints are located in higher nodes, while lower nodes contain constraints specific to a smaller set of instances. Let us suppose the noun *alma* has gone through the pluralnoun-lex-rule to become *almilar* (*alma* + *lar*, ‘apples’) and is ready to go through a phrase structure rule. *Almilar* can be part of a verb phrase if we combine it with the verb *yäydu* (*yä* + *y* + *du*, ‘(s)he/they eat(s)’), which will have been licensed through a different set of lexical rules applying to verbs. These two words will be put together by the comp-head-phrase rule, which specifies that the second of its two daughters becomes the head of the phrase: the daughter whose semantic index is shared with its mother’s. The result is *almilar yäydu* (‘(s)he/they eat(s) apples’).

An HPSG grammar also records a specific list of constraints that a sentence must satisfy in order to be considered well formed. The ultimate goal of HPSG is to map syntactic strings to semantic representations, using MRS.

## II. Sample Phenomena

### A. Accusative Case Definiteness

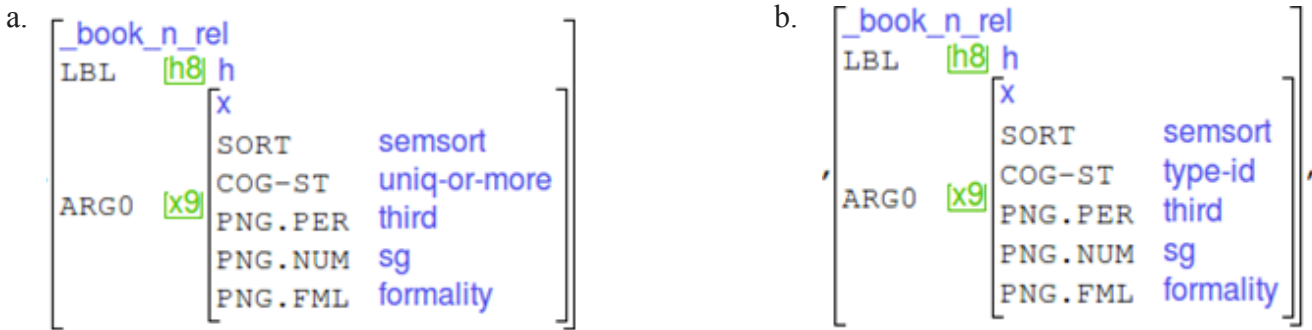
In Uyghur, direct objects may appear either with or without a visible accusative case suffix (*-ni*). The presence or absence of this suffix affects the definiteness of the object. Consider the following sentences:

- (a)        *Kitabni oquymän*  
              *Kitab-ni oqu-y-män*  
              book-ACC read-PF-1SG  
              “I will read the book.”
- (b)        *Kitab oquymän*  
              *Kitab oqu-y-män*  
              book.ACC read-PF-1SG  
              “I will read a book.”

With (*-ni*) present in (a), the speaker asserts that they will read a specific book. Without this suffix in (b), the speaker merely asserts that they will read a book, without a specific book in mind.

Every noun in the Uyghur HPSG grammar must go through a case-assigning lexical rule, which will mark it as eligible for playing certain roles in a sentence. The first question here is what case a bare indefinite object like ‘*kitab*’ should be assigned. I opt to consider ‘*kitab*’ as having accusative case without a suffix, which clarifies its role as the direct object in sentences like (b) and avoids the unwanted ambiguity of possibly being the sentential subject (if it were marked with nominative case).

To encode this analysis, two separate accusative case rules must be created: one which adds the suffix (*-ni*) and specifies that the noun refers to a definite object, and the other which adds no suffix and specifies that the noun refers to a general type of object. These two rules can inherit all other information associated with accusative case from the same parent lexical rule. The definiteness contrast is recorded in the MRS through a feature of nouns called COG-ST (cognitive status). The accusative case rule which adds the (*-ni*) suffix gives ‘*kitab*’ a COG-ST value of *uniq-or-more* (that there is a unique item being referenced), as shown in (a); while the rule which does not add a suffix gives ‘*kitab*’ a COG-ST value of *type-id* (that its type is identifiable), as shown in (b).



## B. Verb Morphology

Verbal suffixes in Uyghur occur in a relatively fixed order (e.g. negation, tense, aspect, person), but only person needs to be overtly present on every verb. To track a verb’s morphological realization, a flag system is used (Goodman 2013). Suffix morphemes are organized into position classes based on their appearance in relation to the stem and other suffixes; each position class is represented by a flag. Flags have *luk* values (ternary and disjunctive) indicating that a given position class needs to be realized (-), has been realized (+), is not relevant in the given context (*na*), or is either *na* or one of + or -. In order for a word to be considered properly “inflected” for participation in a phrasal rule, all of its flags must be set to *na-or-+*; i.e., a word with a position class that needs filling is not considered properly inflected.

For a concrete example, consider the following paradigm of tense morphology in Uyghur. A standard nonpast verb with no special aspect marking will be marked by the -y suffix, as shown in (a).

- (a) Itlar uxlaydu  
 It-lar uxla-y-du  
 Dog-PL.NOM sleep-NPST-3  
 “Dogs will sleep.”

When the progressive aspect suffix *-wat* is added, however, then the -y suffix is not added, although the sentence is still interpreted as nonpast (Engesæth et al 2010).

- (b) Itlar uxliwatidu  
 It-lar uxla-0-wat-du  
 Dog-PL.NOM sleep-NPST-PROG-3  
 “Dogs are sleeping.”

Yet when the negation suffix *-ma* is added along with the progressive suffix *-wat*, then the nonpast -y suffix becomes overt once again.

- (c) Itlar uxlimaywatidu  
 It-lar uxla-ma-y-wat-du  
 Dog-PL.NOM sleep-NEG-NPST-PROG-3  
 “Dogs are not sleeping.”

This peculiar contrast can be modeled in HPSG using flags. Relevant to the above examples, the flags TENSE-FLAG, COVERTNONPAST-FLAG, PROG-FLAG and NEG-FLAG are all defined under INFLECTED. When verb-lex is instantiated, TENSE-FLAG is set to - (since every verb must be inflected for tense), while the other three flags are set to *na* (since it is not necessary that a verb in a vacuum have unmarked nonpast tense, progressive aspect or negation.) For sentence (a), the verb *uxla* will simply go through the nonpast-lex-rule, setting its TENSE-FLAG to +, and the word will then be ready to go through another lexical rule assigning it a person suffix.

A verb that has received the -y suffix will not be eligible to receive the progressive -wat suffix, because

the progressive-lex-rule requires its daughter to have a COVERTNONPAST-FLAG + value. Thus the verb in (b) first goes through the covertnonpast-lex-rule, which does not add a suffix, sets the COVERTNONPAST-FLAG and the TENSE-FLAG values to +, and also sets the PROG-FLAG to -, requiring that the verb now go through the progressive-lex-rule.

In (c), the verb *uxla* first goes through the neg-lex-rule, which adds the suffix *-ma* and sets the NEG-FLAG to +. The verb is now eligible for the negprogovert-nonpast-lex-rule, which adds the *-y* suffix but sets the COVERTNONPAST-FLAG value to +, “tricking” the progressive-lex-rule into adding the *-wat* suffix.

The solution provided here to reproducing this inflectional paradigm is only one of many possible ways to model the morphological reality observed in the Uyghur language.

### C. Nominal Predicate

Unlike English, Uyghur allows nouns to serve as predicates without the linking of a copula verb. Thus Uyghur can have complete sentences like,

Bu zhurnal  
DEM magazine  
“This is a magazine.”

To model this reality, a noun-predicate-rule was created that takes a noun in nominative case as a daughter and converts it into a verb. The rule also adds a “be” predication, linking the first argument with the sentential subject and the second argument with the index identifying the rule’s head daughter noun. This rule allows sentences that seemingly “don’t have a verb” to be licensed while still observing the requirement that a well-formed Uyghur sentence have a verb.

### D. Transformation Rules

Since machine translation is a valuable natural language processing (NLP) application, it is important to engineer a grammar that will be mutually intelligible with other grammars. Transfer rules ensure that components of one grammar can be interpreted in another. Their role is to match potential input from other grammars with a grammar’s own types. Most languages have a unique way of encoding tense and aspect, for example. A language like Chinese has no tense morphology, but aspectual morphemes allow the interlocutor to infer tense. The perfective marker *guo*, for example, indicates that an event occurred in the past. A transfer rule that maps perfective markers to past tense would then be a useful transfer rule to include in an English grammar. Transformation rules are also used when predicates don’t exactly match between languages. The meaning of the verb *hurt* in English is morphologically expressed as *cause pain* in Uyghur (*aghir-t*, lit. pain-CAUS). Transfer rules can encode these relationships. It would be preferable to develop a systematic way of identify these predicate correspondences, but doing so is beyond the scope of this project.

## III. Conclusion

The process of building a precision grammar of Uyghur has been described with the accusative object definiteness contrast, one verb morphology paradigm and nominal predicates given as examples of specific properties of Uyghur grammar that can be treated in a machine-readable HPSG approach.

As measured using [*incr tsdb()*], this Uyghur grammar currently provides 75.3% coverage over a test suite containing 253 items. Many of the sentences lacking a parse exhibit grammatical phenomena (such as dative case) yet to be defined in the grammar. Ambiguity (in the form of multiple parses per string) often arises

because a Uyghur speaker can switch the order of subject and object in some sentences to focus on the latter. As the grammar expands to cover more phenomena, so too will the test suite grow to exhibit a greater variety of phenomena, prompting additional modifications to the grammar in a cycle of discovery.

Grammar engineering provides the ability to parse and generate sentences in real time, allowing for quick testing of linguistic hypotheses. Precision grammars can also easily be compared to determine what properties are or are not common to language families or even universally shared by languages. Additionally, machine-readable grammars can be used in a variety of NLP tasks from machine translation to information retrieval. This project represents early steps in the vast process that is building up a precision grammar of a language. This Uyghur grammar is currently being used to model a class of Uyghur auxiliaries that contribute aspectual information rather than an independent predicate.

### References

- Asarina, Alevtina. (2006). *Case in Uyghur and Beyond*. PhD Diss., Massachusetts Institute of Technology.
- Bender, Emily M. (2008). "Evaluating a Crosslinguistic Grammar Resource: A Case Study of Wambaya." In *Proceedings of ALC08:HLT*. Columbus, OH: Association for Computational Linguistics.
- Bender, Emily M., Scott Drellishak, Antske Fokkens, Laurie Poulson, and Safiyyah Saleem. (2010). "Grammar Customization." *Research on Language and Computation* 8(1):23-72.
- Bender, Emily M., Dan Flickinger and Stephen Oepen. (2011). "Grammar Engineering and Linguistic Hypothesis Testing: Computational Support for Complexity in Syntactic Analysis." In *Language from a Cognitive Perspective*, ed. Emily M. Bender and Jennifer E. Arnold, pp. 5-30. Stanford: CSLI Publications.
- Bridges, Michelle. (2008). *Auxiliary Verbs in Uyghur*. MA Thesis, Kansas University.
- Copestake, Anne. (2002). "Definitions of Typed Feature Structures." In Oepen, Stephan, Dan Flickinger, Junichi Tsujii and Hans Uszkoreit, eds., *Collaborative Language Engineering*. Stanford: CSLI Publications. pp. 227-230.
- Copestake, Ann, Dan Flickinger, Carl Pollard and Ivan Sag. (2005). "Minimal Recursion Semantics: An Introduction." *Research on Language and Computation* 3(4), 281-332.
- Engesæth, Tarjei, Mahire Yakup and Arienne Dwyer. (2010). *Greetings from the Teklimakan: a Handbook of Modern Uyghur*. Lawrence, KS: University of Kansas Scholarworks.
- Goodman, Michael Wayne. (2013). "Generation of Machine-Readable Morphological Rules from Human-Readable Input." In *University of Washington Working Papers in Linguistics* (13).
- Oepen, Stephan. (2001.) [*incr tsdb()*] — *Competence and performance laboratory. User manual*. Technical report. Saarbrücken, Germany.
- Pollard, Carl and Ivan A. Sag. (1994). *Head-Driven Phrase Structure Grammar*. Studies in Contemporary Linguistics. Chicago, IL and Stanford, CA: The University of Chicago Press and CSLI Publications.

## How Can We Define The “Participles” in Mongolic Languages: Two Problems In Shinekhen Buryat

Yasuhiro Yamakoshi  
*Tokyo University of Foreign Studies*

### Abstract

This paper aims to examine the function of “participles (or also called as “verbal nouns”)” in Shinekhen Buryat (SB; one of the Mongolic languages) and to suggest the following two points: 1) It seems to be difficult to define the prototypical participles in SB as inflectional verb > adjective transposition; 2) we should recategorize the participles in SB. In SB, Some of the participles are more verbal and more inflectional. These are “participles” of the inflectional paradigm of verbs. On the other hand, the other forms are less verbal and less inflectional, i.e., more derivational. We should categorize those suffixes into syntactic derivational suffix.

We can find the same problem about participles on the other Altaic languages. Participles in Altaic languages are different from those in Indo-European languages on their verbal properties. It would be important for not only linguistic description, but also for language-teaching to examine the function of each “participle” found on previous grammar.

*Keywords:* participles, verbal nouns, inflection, derivation, Mongolic languages.

This paper aims to examine the function of participles in Shinekhen Buryat, which is one of the Mongolic languages spoken in Manchuria, and suggests the following two points: first, we cannot define the prototypical participles in Shinekhen Buryat as inflectional verb to adjective transposition, since they contain certain features that differ from adjectives; second, we recommend a recategorization of the participles in Shinekhen Buryat as they differ with respect to the degree of deverbalization.

### Linguistic outline of Shinekhen Buryat

Shinekhen Buryat is spoken by the Buryat and the Khamnigan peoples around the basin of the Shinekhen River in Manchuria. The ancestors of speakers were in Russian territory; however, they immigrated to China from 1917 until 1932.

Shinekhen Buryat is a Mongolic language, and as such its grammatical features share many similarities with other Mongolic languages. Such features include agglutination, suffix-dominant morphology, head-final structure, SOV word order, and subordinate clauses preceding main clause. It has rich allomorphs due to vowel harmony. Adjectives are classified into the nominals, which comprise substantives, adjectives, numerals, and so on, as adjectives are able to tolerate every case suffix. For example, the word “*hain*” in (1) functions as an adjective. On the other hand, “*hain*” in (2) takes an accusative case suffix, i.e., the word “*hain*” functions as a noun in (2).

- |     |   |                                     |     |  |                                  |
|-----|---|-------------------------------------|-----|--|----------------------------------|
| (1) | <i>hain</i><br>good(A)<br>“good person” | <i>xun</i> . <sup>1</sup><br>person | (2) | <i>hain-ii=n</i><br>good(A)-ACC=3POS<br>“Choose the better one.” | <i>ab-ii=s</i> .<br>take-IMP=2sg |
|-----|---|-------------------------------------|-----|--|----------------------------------|

<sup>1</sup> All examples of Shinekhen Buryat has been examined by native Shinekhen Buryat speakers. Examples without any sources such as (1) are collected through elicitation.

Furthermore, adjectives also function as adverbs when they appear without any derivational suffixes. For example, the word “*hain*” in (3) modifies the verb “*jab-aarai* (to go later)” directly, acting like an adverb.

- (3) *hain jab-aarai=t.*  
 good(A) go-IMP.FUT=2pl  
 “Good-bye (lit. go well.)”

In the main clause, the person and number of the subject are indicated by enclitical personal predicative particles. Example (4) shows a verbal predicate. The particle “=*b*” agrees with the subject “*bii*.” In (5), this sentence has a nominal predicate. In this sentence “=*b*” also agrees with the subject “*bii*,” as in (4).

- (4) *bii jab-na=b.* (5) *bii Dɔndɔg=b.*  
 I go-PRS(V)=1sg I PSN(N)=1sg  
 “I go.” “I am Dondog.”

In the subordinate clause, subject-agreement is indicated by enclitical personal possessive particles. In (6), in the subordinate clause “*jabxada=mni* (when I went there),” =*mni* agrees with the subject of the clause. In this position, personal predicative particles are never used.

- (6) *jab-xa-da=mni xen=sje ugui=hen.*  
 go-P.FUT-DAT=1sgPOS who=also no=PFV  
 “When I went (there), I couldn’t find anyone.”

Next, we examine the verbal morphology of Shinekhen Buryat (see Figure 1). The verb stem is comprised of a root and some derivational suffix, which relates to voice and aspect. The stem takes inflectional suffixes such as finite suffixes, participle (or also called as verbal noun) suffixes, and converb suffixes. Finite and participles can also take the negative suffix “-*gui*” when needed.

Figure 1. Verbal Morphology of Shinekhen Buryat

V STEM-					
Root-	(-derivational-)		-inflectional	(-negation)	(=person/NB)
	VOICE	Asp			
	-CAUS- -PSS- -RCP- etc.	-PRG- -PFV-	-FIN -P -CV	-NEG (attached to FIN and P)	=PRED =POS

The inflection of verbs consists of three categories: *finite forms*, *participles*, and *converbs*. Finite forms are only used for the predicate of the main clause. Converbs are only used for the predicate of the adverbial clause. Participles, on the other hand, are multi-functional. They are used for the predicate of the adjectival, nominal, main clause, and the adverbial clause.

(7)-(10) provide examples of participles used in various clauses. In (7), the participle “*jab-aa* (going)” modifies noun, “*zon-ood* (people).”

- (7) *tende jab-aa zon-ood xed=be.*  
 there go-P.IPFV people-PL who(PL)=Q  
 “Who went there? (lit. Who is the people who went there?)”

In (8), a participle is used as the predicate of the nominal clause.

- (8) *tamɨʃa tat-xa=sni bijen-de=sj moo.*  
 cigarette pull-P.FUT=2sgPOS body-DAT=2sgPOS bad  
 “Smoking is bad for your body.” [Yamakoshi 2006: 156]

In (9), a participle is serving as the predicate of the main clause.



- (9) *bii* *axai-d-aa* *zaxidal* *bisi-ee=bi*.  
 I elder.brother-DAT-REFL letter write-P.IPFV=1sg  
 “I wrote a letter to my brother.” [Yamakoshi 2006: 165]

Furthermore, participles can function as the predicate of the adverbial clause (e.g. (10)). In such a case, however, a personal possessive particle should be attached to the participle.

- (10) *tereen-ii=n* *s'ax-aad* *garga-ha=moni* *onta-z'a* *bai-na*.  
 that-ACC=3POS push-CV.PFV put.out-P.PFV=1sgPOS sleep-CV.IPFV be-PRS  
 “Although I pushed out that, (he) was sleeping.” [Yamakoshi 2014: 192]

### Are participles “the inflectional V > A transposition?”

As previously demonstrated, such multi-functionality of participles is similar to that of adjective we have seen above. Adjectives also have four uses, *i.e.*, the predicative, adnominal, nominal, and adverbial. Haspelmath & Sims (2010) propose that participles are the inflectional Verb-to-Adjective transposition. This definition seems to be correct in Shinekhen Buryat due to the syntactic multi-functionality of participles, similar to that of adjectives.

However, the syntactic behavior of participles is not equal to that of adjectives. Firstly, participles must take any particles, e.g., a personal possessive particle as in the sentence (10) when participles appear in the predicate of adverbial clauses. On the other hand, adjectives never take any particles as seen in (3).

Secondly, participles can take the negative suffix, “-*gui*,” one of the modal suffix. However, adjectives cannot take the negative suffix.

- (11) *idj-xe-gui* / *idj-ee-gui*. (12) *\*hain-gui*  
 eat-P.FUT-NEG / eat-P.IPFV-NEG good(A)-NEG  
 “Not eating / Not having eaten”

We can account for these differences by proposing that participles maintain verbal properties whereas adjectives lack these features. These differences are the reason why participles are not equal to adjectives.

### Participles as the deverbal nominalization

Next, we identify the participles in the Buryat inventory. Table 1 shows forms used to classify Buryat participles in previous studies.<sup>2</sup> We can see that the participles contained in each description differ between the various studies. This demonstrates lack of agreement concerning the definition of participles. In order to resolve this discrepancy, we must examine the ways in which each form differs from the others.

<sup>2</sup> The small number at each end of morpheme (e.g. 3 in *-han<sub>3</sub>*) indicates the number of allomorphs due to vowel harmony. For example, *-han<sub>3</sub>* has three allomorphs such as *-han*, *-hen*, *-hɔn*.

Table 1. Classification of participles in previous researches.

	term	participial suffixes
Poppe (1960)	Verbal Noun	<i>-han</i> <sub>3</sub> PFV; <i>-aa</i> <sub>4</sub> IPFV; <i>-xa</i> <sub>3</sub> FUT; <i>-dag</i> <sub>3</sub> HAB; <i>-gsja</i> <sub>3</sub> AGT; <i>-aasja</i> <sub>4</sub> AGT; <i>-aŋxai</i> <sub>3</sub> RSL; <i>-aatai</i> <sub>4</sub> PSS
Kuribayashi (1992)	<i>Jap. Keidoushi</i> (Adjectival-verb)	<i>-han</i> <sub>3</sub> ; <i>-aa</i> <sub>4</sub> ; <i>-xa</i> <sub>3</sub> ; <i>-dag</i> <sub>3</sub> ; <i>-gsja</i> <sub>3</sub> or <i>-aasja</i> <sub>4</sub> ; <i>-aŋxai</i> <sub>3</sub>
Darbeeveva (1997)	<i>Rus. Prichastie</i> (Participle)	<i>-han</i> <sub>3</sub> ; <i>-aa</i> <sub>3</sub> ; <i>-xa</i> <sub>3</sub> ; <i>-dag</i> <sub>3</sub> ; <i>-gsja</i> <sub>3</sub> ; <i>-aasja</i> <sub>4</sub> ; <i>-aŋxai</i> <sub>3</sub> ; <i>-aatai</i> <sub>4</sub> ; <i>-maar</i> <sub>4</sub>
Skribnik (2003)	Participle	<i>-han</i> <sub>3</sub> ; <i>-aa</i> <sub>4</sub> ; <i>-xa</i> <sub>3</sub> ; <i>-dag</i> <sub>3</sub> ; <i>-gsja</i> <sub>3</sub> ; <i>-aasja</i> <sub>4</sub> ; <i>-aŋxai</i> <sub>3</sub> ; <i>-aatai</i> <sub>4</sub> ; <i>-xaar</i> <sub>4</sub> ; <i>-maar</i> <sub>4</sub>

In order to analyze the differences in the various forms, I apply Malchukov's proposal of hierarchy constraints on transcategorial operations. Malchukov (2006) points out that when verbs are nominalized, both deverbalization and nominalization occur independently along the hierarchy. When a verb is nominalized in various ways, verbal properties contained within the verb will be lost, such as illocutionary force, agreement, mood, tense, aspect, voice, and valency. Simultaneously, the word will first acquire "case" property, then receiving the determiner feature, possessive marker, number and classifier in this order.

To better illustrate this process, let us examine both the deverbalization and nominalization of deverbal nouns. *id'emxei* appeared in sentence (13), is a noun derived from the verb *id-* meaning "eat." This *id'emxei* can take case suffixes such as accusative, possessive markers, and a plural suffix.

- (13) *id-emxei-nuud-ii id'eel-uul-xe-gui. (+NB)*  
eat-ADJV-PL-ACC eat-CAUS-P.FUT-NEG  
“(S/He) doesn't feed greedy eaters.”

Degree of both the deverbalization and nominalization of deverbal nouns is indicated as in (14)

- (14) Deverbalization  
-Valency >> -Voice >> -Aspect (>> -Tense) >> -Mood >> -AGR >> -IF  
No properties!  
Nominalization  
-Case (>> -Det) >> -Pos >> -NB (>> -CL)<sup>3</sup>  
----->

Thus, we can say this word is highly nominalized. On the other hand, this "*id'emxei*" cannot take any verbal suffixes which indicate Mood, Voice, Aspects and cannot take any argument (subject or object), *i.e.*, this form lacks all verbal properties, from illocutionary force to valency. Deverbal nouns maintain their lexical meanings, however, they lose their verbal properties.

### Verbal Properties of Prototypical Participles

Prototypical participles, including "future," "imperfective," "perfective," and "habitual," are not so "deverbalized," compared to derivational nouns. These participles only lack the illocutionary force in the cline of verbal properties. They maintain many properties such as, Valency (they take subject and/or object arguments), Voice (they can attach voice suffixes), Aspect (they also attach aspectual suffixes), Mood (they can take negative suffix "*-gui*"), Agreement (they indicate person and number by personal particles). Additionally, these participles also have rich nominal properties, with the exclusion of number. They can take case suffixes and possessive markers.

<sup>3</sup> I parenthesized Tense, Det(eterminer), and Cl(assifier) since those properties are not indicated by participles in Shinekhen Buryat.

- (15) *id'eel-uul-deg-gui=sj* *moo.* (+Voice, +Mood, +AGR)  
eat-CAUS-P.HBT-NEG=2sgPOS bad  
“It is bad that you don’t feed someone. (*lit.* Not your feeding (someone) is bad)”
- (16) *ab-xa-jii=mni* *tere med-ne.* (+Case, +Poss)  
take-P.FUT-ACC=1sgPOS s/he know-PRS  
“S/He knows what I will buy.”

Degree of both the deverbalization and nominalization of the four prototypical participles is indicated as in (17).

- (17) Deverbalization  
-Valency >> -Voice >> -Aspect (>> -Tense) >> -Mood >> -AGR >> -IF  
----->  
Nominalization  
-Case (>> -Det) >> -Pos >> -NB (>> -CL)  
----->

### Verbal Properties of Agentive Participle *V-gs'a3*

However, we cannot say that all participles in previous descriptions maintain such rich verbal properties. The agentive participle *V-gs'a3* lacks many verbal properties, such as Illocutionary force, Agreement, Mood, and Aspect. However, this form can take a plural suffix, so we can say that it is nominalized higher than the prototypical participles.

- (18) *ɔrci onta-gs'a=sjni* *juu g-ees'e=b.* (-AGR, +Case, +pos)  
late sleep-P.AGT=2sgPOS what say.that-P.AGT=Q  
“Why is it that you sleep so late? (*lit.* Thy late sleeping is what?)” [from Poppe 1960: 67]
- (19) *\*tere basgan sora-gs'a-gui.* (OK: *sora-gs'a bis'*) (-Mood)  
that girl:NOM study-P.AGT-NEG study-P.AGT NOT  
“That girl is not a student”

Degree of both the deverbalization and nominalization of the agentive participle *V-gs'a3* is indicated as in (20).

- (20) Deverbalization  
-Valency >> -Voice >> -Aspect (>> -Tense) >> -Mood >> -AGR >> -IF  
----->  
Nominalization  
-Case (>> -Det) >> -Pos >> -NB (>> -CL)  
----->

### Verbal Properties of Agentive Participle *V-aas'a4*

Another example, *-aas'a4*, preserves the verbal properties of Mood and Aspect. It can take the negative suffix just like the prototypical participles.

- (21) *tam-da ona-z'ai-g-aas'a=haa,* (+Valency, +Aspect)  
hell-DAT fall-PROG-E-P.AGT=COND  
“If s/he were in the hell, (*lit.* If (s/he) is the person being in the hell,)”
- (22) *jab-aas'a-gui-de.* (+Mood, +Case)  
go-P.AGT-NEG-DAT  
“To someone who does not go.”

Degree of both the deverbilization and nominalization of the agentive participle *V-aasja*<sub>4</sub> is indicated as in (23).

- (23) Deverbilization  
 -Valency >> -Voice >> -Aspect (>> -Tense) >> -Mood >> -AGR >> -IF  
 ----->  
 Nominalization  
 -Case (>> -Det) >> -Pos >> -NB (>> -CL)  
 ----->

### Verbal Properties of Agentive Participle *V-anxai*<sub>3</sub>

Example (24) shows the example of resultative participle *-anxai*<sub>3</sub>. This form can take neither the negative suffix (–Mood) nor the plural suffix (–NB). This form lacks more verbal properties than the prototypical participles.

- (24) *jaa-han*                      *es-enxai*                      *s'aaxai-tai=s'*                      (+Valency, +Aspect)  
 do.what-P.PFV                      be.tired-P.RSL                      shoe-PROP=2sg  
 “What wear-out shoes you put on!”

Degree of both the deverbilization and nominalization of the resultative participle *V-anxai*<sub>3</sub> is indicated as in (25).

- (25) Deverbilization  
 -Valency >> -Voice >> -Aspect (>> -Tense) >> -Mood >> -AGR >> -IF  
 ----->  
 Nominalization  
 -Case (>> -Det) >> -Pos >> -NB (>> -CL)  
 ----->

### Verbal Properties of Agentive Participle *V-aatai*<sub>4</sub>

(26)(27) are examples of passive participle *-aatai*<sub>4</sub>. I propose that this form is analyzed as the imperfective participle with a proprietive suffix *-tai*<sub>3</sub> since the negative form of *V-aatai*<sub>4</sub> is realized *V-aa-gui*<sub>4</sub>, not *V-aatai-gui*<sub>4</sub>.

- (26) *saarhan*                      *deer*                      *bisi-eetei*                      *bisjeg*                      (+Valency, +Aspect)  
 paper                      on                      write-P.PSS                      letter  
 “a letter / letters written on a paper.”
- (27) *\*bisi-eetei-gui./*                      *bisi-ee-gui.*                      (–Mood)  
 write-P.PSS-NEG/                      write-P.IPFV-NEG  
 “Not being written”

Degree of both the deverbilization and nominalization of the passive participle *V-aatai*<sub>4</sub> is indicated as in (28).

- (28) Deverbilization  
 -Valency >> -Voice >> -Aspect (>> -Tense) >> -Mood >> -AGR >> -IF  
 ----->  
 Nominalization  
 -Case (>> -Det) >> -Pos >> -NB (>> -CL)  
 ----->

### Difference among those “participles”

Let us observe the degree of both verbal and nominal properties of each participle outlined in the previous description. (29) shows the verbal properties contained within each participle. From this cline, we can say that three participles *V-gsʹa<sub>3</sub>*, *V-aŋxai<sub>3</sub>* and *V-aatai<sub>4</sub>* are more deverbalized than the prototypical participles, as they cannot take some verbal suffixes, i.e., they lack certain properties. On the other hand, the nominal properties (30) that each form has are not so different from those contained in derivational nouns.

(29) How much is each form deverbalized?

	-Valency >> -Voice >> -Aspect (>> -Tense) >> -Mood >> -AGR >> -IF
Prototype	----->
<i>V-aasʹa<sub>4</sub></i>	----->
<i>V-gsʹa<sub>3</sub></i>	----->
<i>V-aŋxai<sub>3</sub></i>	----->
<i>V-aatai<sub>4</sub></i>	----->
V-NMLZ	No properties!

(30) How much is each form nominalized?

	-Case (>> -Det) >> -Pos >> -NB (>> -CL)
Prototype	----->
<i>V-aasʹa<sub>4</sub></i>	----->
<i>V-gsʹa<sub>3</sub></i>	----->
<i>V-aŋxai<sub>3</sub></i>	----->
<i>V-aatai<sub>4</sub></i>	----->
V-NMLZ	----->

Furthermore, the productivity of each participle is different. All verb-stems can take the four prototypical participles. Additionally, almost all verbs can take the agentive *-aasʹa<sub>4</sub>*. However, certain verbs cannot tolerate the agentive *-gsʹa<sub>3</sub>*, the resultative *-aŋxai<sub>3</sub>* and the passive *-aatai<sub>4</sub>*. It is important to note that derivational suffixes are less productive than inflectional suffixes, a pattern observed cross-linguistically.

(31) What kind of verbs can take suffixes:

Prototype	ALL verbs	<- Highly productive
<i>-aasʹa<sub>4</sub></i>	Almost all verbs	<- Highly productive
<i>-gsʹa<sub>3</sub></i>	Verbs without some state verbs	<- Less productive
<i>-aŋxai<sub>3</sub></i>	Verbs without some state verbs	<- Less productive
<i>-aatai<sub>4</sub></i>	Verbs without intransitive verbs	<- Less Productive

### What are “participles” in Shinekhen Buryat?

Then, what are the participles in Shinekhen Buryat? Regarding Altaic languages, Kazama (2003) mentions that “verbal nouns or participles in European languages lack more verbal properties than those in Altaic languages.” From this perspective, we may also define *-gsʹa<sub>3</sub>*, *-aŋxai<sub>3</sub>* and *-aatai<sub>4</sub>* suffixes as participles. However, these suffixes are less verbal.

Participles in Altaic languages are more verbal than those in Indo-European languages. If we include participles as well as finite and converbs into the inflectional paradigm<sup>4</sup>, we should recognize only the four prototypical participles (*V-han<sub>3</sub>*; *V-aa<sub>4</sub>*; *V-xa<sub>3</sub>*; *V-dag<sub>3</sub>*) and the agentive participle *V-aasʹa<sub>4</sub>* as participles. On the other hand, the other suffixes are less inflectional (i.e., less productive and less verbal). However, they preserve the verbal property, Valency. Therefore, I propose that we recategorize these incomplete-inflectional suffixes as syntactic derivational suffixes that Vinokurova (2005) proposed in Sakha.

## Conclusion

In conclusion, I suggest two points. First, participles in Shinekhen Buryat are not accurately captured through the inflectional Verb to Adjective transposition. Second, previous research varies with respect to participle classification, and a recategorization is demanded.

In Shinekhen Buryat, Some of the participles (V-*han*<sub>3</sub> PFV; V-*aa*<sub>4</sub> IPFV; V-*xa*<sub>3</sub> FUT; V-*dag*<sub>3</sub> HAB; V-*aas/a*<sub>4</sub> AGT) are more verbal and more inflectional. These are “participles,” the inflectional paradigm of verbs. On the other hand, other forms (such as: V-*gs/a*<sub>3</sub> AGT; V-*anxai*<sub>3</sub> RSL ; V-*aatai*<sub>4</sub> PSS) are less verbal and less inflectional, *i.e.*, more derivational. These suffixes should be categorized as syntactic derivational suffixes.

We observe the same problem concerning participles in other Altaic languages, which differ from Indo-European languages with respect to their verbal properties. It is important for not only linguistic description, but also for language instruction, to examine the function of each “participle” found in previous grammar.

## Abbreviations

-: morpheme boundary, =: clitic boundary, 1,2,3: person, A: adjective, ACC: accusative, ADJV: adjective-derivational suffix, AGR: agreement, AGT: agentive, CAUS: causative, CL: classifier, CV: converb, DAT: dative, Det: determiner, FIN: finite, FUT: future, HAB: habitual, IF: illocutionary force, IMP: imperative, IPFV: imperfective, N: noun, NB: number, NEG: negative, NMLZ: noun-derivational suffix, NOM: nominative, P: participle, PFV: perfective, pl/PL: plural, POS: possessive, PRED: predicative personal particle, PRG: progressive, PROP: proprietive, PRS: present, PSN: person name, PSS: passive, Q: interrogative particle, RCP: reciprocal, REFL: reflexive possessive, RSL: resultative, sg: singular, V: verb.

## References

- Darbeeveva, A. A. (1997). Buryatskii Jazyk. *Jazyki Mira: Mongol'skie jazyki, Tunguso-man'čžurskie jazyki, Japonskii jazyk, Koreiskii jazyk*. pp. 37-51, Moscow: Izdatel'stvo “Indrik.”
- Haspelmath, M. & Sims, A. D. (2010). *Understanding Morphology* [2nd ed.]. London: Hodder Education Publishers.
- Janhunen, J. (2003). Proto Mongolic. In: Juha Janhunen (ed.) *The Mongolic Languages*. pp. 1-29, London and New York: Routledge.
- Kazama, S. (2003). Arutai shogengo no 3 gurupu, oyobi Chosen go, Nihon go no bunpo ha honto ni niteiru no ka: taisho bunpo no kokoromi. (Do the Three Groups of the “Altaic” Languages (Turkic, Mongolic, and Tungusic), as well as Korean, really Resemble Japanese Grammatical Structure?: An Attempt at a Contrastive Grammar Analysis) In: Alexander Vovin and Osada Toshiki (eds.) *Nihongo keitoron no genzai. (Perspectives on the Origins of the Japanese Language)* pp. 249-342. Kyoto: Kokusai Nihon Bunka Kenkyu Center.
- Kuribayashi, H. (1992). Buriyaato go. (the Buryat Language) In: Kamei Takashi, Kono Rokuro and Chino Ei-ichi (eds.) *Gengogaku Daijiten. (The Sanseido Encyclopaedia of Linguistics)*. (vol. 3, pp. 814-27). Tokyo: Sanseido.
- Malchukov, A. L. (2006). Constraining nominalization: function/form competition. *Linguistics*. 44(5), 973-1009.
- Poppe, N. (1960). *Buriat Grammar*. Bloomington, Indiana University Press.
- Skribnik, E. (2003). Buryat. In: Juha Janhunen (ed.) *The Mongolic Languages*. pp. 102-128. London and New York: Routledge.
- Vinokurova, N. (2005). *Lexical Categories and Argument Structure: A Study with Reference to Sakha*. Utrecht, LOT. Retrieved from <http://igitur-archive.library.uu.nl/dissertations/2005-0325-013011/full.pdf>



- Yamakoshi, Y. (2006). Shinehen-Buriyaato go Tekisuto: Nichijo-kaiwa wo Daizai ni Shita Kihon Bunrei-shuu. (Shinekhen-Buryat Text: Basic Sentences in Daily Conversation). *Kan-kitataiheyoo no gengo (Languages of the North Pacific Rim)*. 13, 139-180.
- Yamakoshi, Yasuhiro. (2014). Shinehen-Buriyaato go Tekisuto: Tenegu Taribu hoka 2 hen. (Shinekhen Buryat Text (4): “Teneg Tarib” and Others). *Hoppo Gengo Kenkyu. (Northern Language Studies)* 4, 185-198. Retrieved from <http://eprints.lib.hokudai.ac.jp/dspace/handle/2115/55129>

### Author’s Note

Yasuhiro Yamakoshi, Research Institute for Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.

This research was supported by JSPS KAKENHI Grant Number 22720163

Corresponding concerning this article should be addressed to Yasuhiro Yamakoshi, Research Institute for Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies, Fuchu-shi, Tokyo, JAPAN 1838534.

Contact: [yamakoshi@aa.tufs.ac.jp](mailto:yamakoshi@aa.tufs.ac.jp)





Part III: Pedagogy and Teaching



## Developing a Learner-Centered Pashto Curriculum for Specific Purposes: A Case Study

Marla Federe  
Clay Leishman

*NSW Foreign Language and Culture Program*

### Abstract

Writers will present a Pashto curriculum developed by Naval Special Warfare for Navy SEALs. The development process of this language for special purposes (LSP) curriculum, and the conceptual framework behind it, will be outlined. Aspects central to this curriculum will be examined including learner-centeredness, criterion-referenced objectives, as well as formative, summative, authentic and transparent assessment. Writers' objective is to provide a model for Pashto instructors and curriculum developers which will assist them in focusing their curricula and instruction on students' goals and learning needs.

*Keywords:* learner-centered, formative assessment, instructional design

### Background

Recently, the Department of Defense (DOD) redefined its use of language for Special Operations Forces (SOF) in the 21st century. To that end, Naval Special Warfare Command was tasked to develop and administer programs capable of preparing Naval Special Warfare (NSW) Operators to maintain a basic level of linguistic competence appropriate for a globally-employed force. The desired outcome of these programs is improved counterpart relations and decreased dependence on interpretation/translation by non-SOF linguists or interpreters. Also inherent in the language training mission is an increased level of cultural sensitivity and knowledge that contributes significantly to situational awareness, safety, and security.

At the onset of this initiative in 2009, the curricula available focused on a generic, global, language proficiency which was primarily thematically based (e.g., numbers, greetings, hobbies, family, etc.). It was apparent that a gap existed between available curricula and our needs. Logistically, this required an in-depth needs assessment, development of standardized courses of instruction, course objectives, course schedules, assessment systems, course data management systems, and a course evaluation system. However, with an immediate requirement to start training Operators, the decision was made to combine the existent curricula with best practices teaching methods (Brown, 1994; Brown, 1993; Wilkerson & Irby, 1998) and an integrated system of instructor, student, and course evaluation (Evans, 2013; Sadler, 2010). To ensure that these practices were adhered to, the teaching standards were clearly outlined during the instructors' pre-service training in an internal Quality Assurance Surveillance Process (QASP) (Federe, 2008) which included eight basic tenants (Figure 1). These eight tenants formed the cornerstone of our conceptual framework.

Although the existing curricula did not meet the requisite needs due to their focus on general proficiency, students were able to learn enough to score well on the Defense Language Proficiency Test (DLPT) which tests listening and reading proficiency and the Oral Proficiency Interview (OPI) which tests a student's speaking ability. Both tests assess global proficiency as measured by the Interagency Language Roundtable (ILR) proficiency scale, which measures a student's proficiency level on a scale of 0 to 5. However, these tests do not measure the attainment of the goals outlined by NSW leadership and the curricula did not cover the skills needed to meet leadership's goals. With a limited amount of time to train, in our case 12 weeks were allotted by the command for initial acquisition training (IAT), and very specific goals, it became essential that precise language objectives, curricula, assessments, and evaluations to meet these needs be defined (Brown, 2010; Tyler, 2013).

**Quality Assurance Surveillance Process (QASP) Teaching Criteria**

- Provide a positive, supportive learning environment, which encourages all students to participate actively in class.
  - Ensure teaching is learner-centered.
  - Be flexible and adaptive to students' learning styles and learning needs.
  - Engage students in authentic, contextualized, high-interest, relevant tasks which are focused on valid operational language objectives.
- Make students active agents in their own learning process.
- Maintain a balance between fluency and form, and between grammatical, discourse, sociolinguistic, and strategic competencies.
- Provide frequent, constructive feedback and assessment as to how students are progressing, encouraging/training students to self-assess and thus become active in their own learning process.
- Ensure lessons have clearly articulated goals, objectives, and outcomes.
- Focus on use of language for real-life, contextualized communication tasks.
- Maximize target language comprehensible input, output, and interaction.
- Encourage critical analysis and synthesis.

Figure 1 QASP Criteria

### *Framing the Problem*

Leadership had specified the need to use language in order to communicate directly with indigenous peoples, improve counterpart relations, enhance operational capability, decrease dependence on interpreters/translators and for an increased level of cultural sensitivity and knowledge that contributes to situational awareness, safety, and security. While there was a clear vision of what leadership wanted and we had the basic cornerstone for the conceptual framework, a gap existed between the curricula and assessments in use and the ones needed to accomplish these goals (Messick, 1990). The decision was made that it would be less expensive to develop and implement a curriculum that fit the Operators' needs precisely than to teach from a general-proficiency, generic curriculum that students might score well on in standardized tests, but did not meet mission-specific needs.

Another concern was the effect that high-stakes tests had, and would have, on instruction and learning. Namely, the knowledge of the final tests could affect what the students were learning in the classroom, and could have a counter-productive washback (Messick, 1996) on learning. In particular, if the students need to be able to speak with counterparts, a general-proficiency test which focuses on academic listening and reading would only send a confusing message to the students and wouldn't support our specific learning needs and objectives. To be effective and efficient, it would be necessary to develop a curriculum, assessment, and policy system which were in alignment. In fact, the requirements necessitated the development of such a curriculum in seven languages: Pashto (PU), Farsi (PF), Arabic (AD), Tagalog (TA), French (FR), Spanish (QB), and Swahili (SW).



*Solving the Problem*

To address the gap, it became essential that the NSW-specific needs be defined in detail. In order to detail our requirements and develop the necessary curricula, Brown’s (1995) model for systematic curriculum design was utilized and adapted. This model included a defined needs analysis, objectives, testing, materials, teaching, and evaluation, with the flow between these items going both ways (Figure 2). The Brown framework was then adapted (Figure 3) in accordance with previous projects (Federe, 2005) including the addition of vision, goals, and conceptual framework as there was a need to keep leadership and stakeholders’ ultimate vision central to the process, a shift from “testing” to “assessment” to reflect our view of assessment as information-rich feedback embedded throughout the curriculum and teaching methodology, and the addition of policies and procedures as this would be necessary in order to ensure continual alignment of vision and curriculum and to capitalize on student motivation (Bandura, 1993).

The overall goal for this project was to develop a 12-week (360-hour) course which produced the capabilities required for NSW mission sets, was learner-relevant (Hutchinson, 1987), had explicit and transparent course and learning objectives (Brown, 1995), was efficient (Tumposky, 1984), flexible, adaptable, updatable, and allowed for monitoring and accountability for students, instructors, and the program as a whole (Valette, 1980).

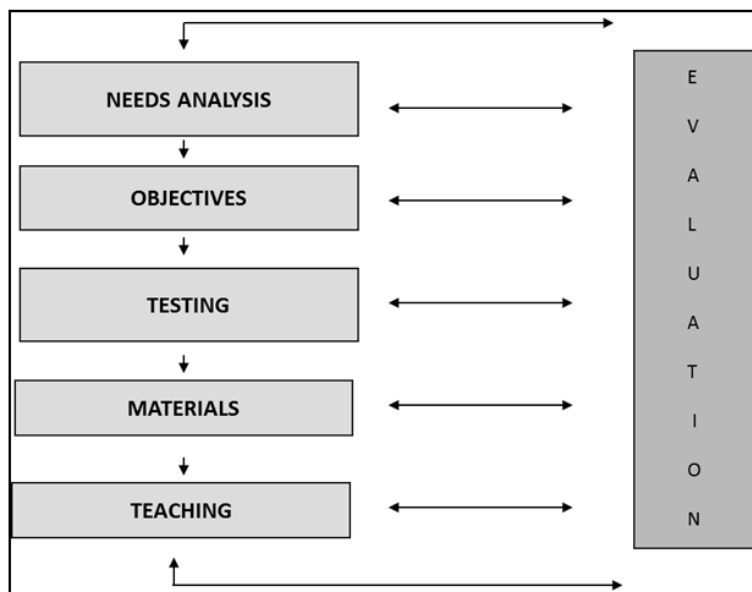


Figure 2 Brown Model

**Development Process**

The development process included eight basic components. The needs assessment came first, followed by a refinement of our conceptual framework and detailed definition of the target objectives and outcomes. The next four components (assessment, methods, materials, and policy/procedures) were intertwined, but always referred back to the objectives, outcomes, and conceptual framework. While iterative evaluations (represented by the horizontal two-way arrows in Figure 3) were part of the development process, the final step of the curriculum project was the development of a formal evaluation process to continue monitoring and updating the curriculum as SEAL mission sets adapt and evolve.

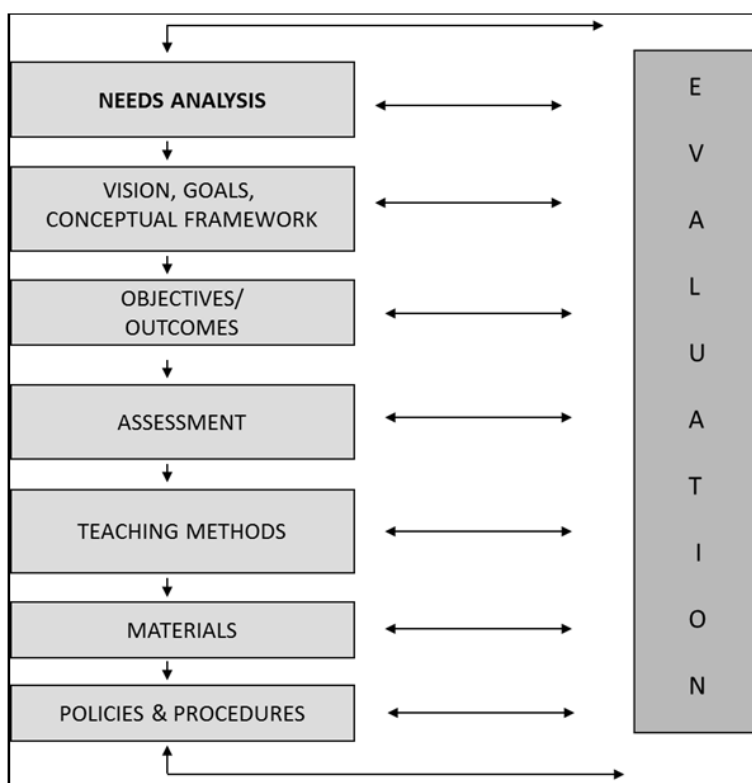


Figure 3 Federe Adaptation

### *Needs Assessment*

With a limited amount of time to train, it was essential that we identified our students' exact linguistic needs. Our goal at the onset was to identify situations in which the students needed to use the target language, break those situations down into tasks, and then analyze the tasks in order to identify the associated language functions. In other words, "(l)anguage varies from one situation of use to another, it should be possible to determine the features of specific situations and then make the features the basis of the learners' course." (Hutchinson, 1987)

Our aim was to join best practices from second language acquisition theory with best practices from industrial organizational psychology (Klimoski & Zukin, 2003; Münsterberg, 1913; Surface, 2012) to complete a thorough language-focused task analysis. Industrial organizational psychologists conduct KSA (knowledge, skills, and abilities) analysis on job tasks on a regular basis. Our aim was to apply that methodology through the lens of second language acquisition.

Four Navy doctrinal manuals which detail the tasks a SEAL must be able to perform were utilized. The cumulative list contained 834 items. Then, focus groups of qualified SEALs were formed. In order to qualify for the focus group, an Operator had to have a minimum number of deployments, number of years of service, specified ranks, and had to have deployed to at least 2 of our target language regions on a regular basis.

These focus groups were then asked to complete detailed surveys, which included all 834 tasks. In these surveys, the tasks were presented one at a time and the Operators were asked whether they had ever done the task, how frequently, how important the task was to their mission, whether or not it required language or culture, at what level of engagement (tactical, operational, or strategic) the target language/culture was used, whether they used the target language to do the task themselves, or to train others to do the task, and whether or not it required an interpreter. The results of these surveys were tallied, and we derived 613 mission-based tasks that required use of the target language and/or culture to some degree. All data were then put into a comprehensive database.

Next, a second focus group was formed. This focus group had to meet the qualification criteria of the first group, and also had to have a record of using the target language operationally. This group of Operators was matched up with a group of language subject matter experts (SMEs) from the American Council of Teachers of Foreign Languages (ACTFL). The Operators then were able to group the tasks into three main categories. This grouping made it easier to evaluate the 613 tasks. Then, the language SMEs interviewed the Operators asking them to describe how they would use language and/or culture in each of the tasks. As the Operators explained the process, the language SMEs distinguished the language functions associated with each of those tasks. All of this information was put into the database so that in the end it contained all of the tasks which require language and/or culture with their associated linguistic functions, level of conflict, whether done by oneself or involved training others to do, frequency and importance, levels of socio-cultural knowledge required, communicative modalities required, and a target proficiency and performance level required to complete the task successfully.

### *Conceptual Framework*

While a working copy of our conceptual framework was developed before the needs assessment process was conducted, after the needs assessment, we reviewed, updated and elaborated on it so that curriculum developers could clearly understand our intent. Central to our framework is the concept of learner-centeredness since our goal is to equip students with a mission-focused language capability. "The learner-centered view tends to view language acquisition as a process of acquiring skills rather than a body of knowledge." (Nunan, 1996).

Likewise, the focus of this course and curriculum was on developing a skill that would be applied, or in other words, our focus was the development of a capability.

Two key tenants of learner-centered instruction (Blumberg, 2009; Weimer, 2013) in the second language classroom are adult learning theory (Knowles, 1983; Knowles, 1980; Merriam & Caffarella, 1999; Mezirow, 1997; Mezirow, 1981) and communicative language teaching (Brown, 1994; Nunan, 1996; Nunan, 1987). According to adult learning theory, adult learners have needs that are fundamentally different than K-12 learners, which changes how learning activities should be developed and used. Characteristics common to adult learners (Knowles, Holton, & Swanson, 2005) include: needing to know the reason(s) behind their learning, maintaining an element of control in the decision-making process, having previous experience(s) recognized, having an eagerness to learn things that will help them succeed in real-life situations, and that they are more often motivated by internal (ability to perform well on their jobs) rather than external (pay, promotions) motivators.

Communicative language teaching is based on the development of communicative competence for real-life application as opposed to mere memorization of discrete knowledge (Cody, 1997) or rules in isolation of context (Krashen, 1982; Mahn, 2013). Simply being able to create grammatically correct structures in a language does not necessarily enable a learner to use language to carry out various real-life tasks (Widdowson, 1978). It is also believed that there are strong and weak versions of communicative language teaching (Howatt, 1984), which is still seen today, more than thirty years later. Almost all language instructors will claim that they use communicative language teaching methods and yet many use what Howatt describes as the “weak version”. In particular, “(t)he weak version which has become more or less standard practice in the last ten years, stresses the importance of providing learners with opportunities to use their English for communicative purposes and, characteristically, attempts to integrate such activities into a wider programme of language teaching.” (p. 279) Howatt goes on to say “(t)he strong version of communicative language teaching, however, sees language ability as being developed through activities which actually simulate target performance. In other words, class time should be spent not on language drills and controlled practice leading towards communicative language use, but in activities which require learners to do in class what they will do outside.” (ibid.)

Blumberg (2009) details five dimensions of learner-centered teaching. These include: the function of content which should give the student the ability to apply content and understand why they are learning it, the role of the instructor as a facilitator of learning, the responsibility for learning which should be that of the students with proactive support and expert guidance from the instructor, the purpose of assessment which should be to provide constructive feedback and information to the learner about where they are and where they need to be, and the balance of power shifts to give students and instructors input as to what is to be covered and how.

These concepts of adult learning theory, communicative language teaching, and learner-centered instruction in general, form the foundation of our teaching criteria as outlined earlier through our Quality Assurance Surveillance Process, and form the conceptual framework which undergirds this curriculum. It is vital to provide a positive, supportive learning environment, ensure teaching is learner-centered, maintain a balance between fluency and form, provide constructive feedback and assessment, ensure lessons have clearly articulated objectives, focus on use of language for real-life, contextualized tasks, maximize target language comprehensible input, output, and interaction, and encourage critical analysis and synthesis.

### *Objectives and Outcomes*

From the extensive data collected from the needs assessment, the learning objectives were generated. Specific instructional objectives have three essential characteristics: (Mager, 1984) performance (what tasks the Operator needs to be able to perform), condition (important conditions under which the performance is expected to occur), and criterion (the quality or level of performance that will be considered acceptable). Our learning objectives were stated in terms of outcomes and all included the following elements: mission content domain and task, purpose (do self or train others), level of conflict (tactical, operational, or strategic), frequency and

importance, linguistic function needed for this task, level of socio-cultural knowledge needed for this task, communicative modalities used in this task, and a target proficiency and performance level. In our case we defined these targeted proficiency levels in terms of the ILR scale; namely, we mapped our tasks, conditions, and standards to the ILR proficiency scale (i.e., what level of proficiency did the Operator need to be able to accomplish the specified task). Every identified task was coded according to these criteria. The tasks were then ordered by linguistic function, so that they came together into a scope and sequence ordered by linguistic function with simpler functions building to more complex functions. Developers then had a number of contextualized tasks to choose from for each linguistic function, and were directed to choose those marked as most important first. Thus, requisite linguistic functions were contextualized.

#### *Assessments, Methods, Materials, Policy and Procedures*

The four aspects (assessments, methods, materials, policy and procedures) were closely intertwined and developed together once the other foundational items had been established. Some of the methods central to this approach are: spiraling, and scaffolding, a focus on objectives, feedback, student reflection and responsibility, and performing a task only up to a pre-determined level.

The objectives are clearly detailed at the start of the course, each day, and each hour/block. All class activities are based on meeting these objectives. Periodically (start of day, during the day, at the end of the day), the students are asked to reflect on how they would use the daily objectives operationally and how confident they felt in accomplishing the tasks operationally. Students have the responsibility to actively reflect on how they would use the materials operationally. Feedback, quizzes, and end-of-week tests are all based on the ability to accomplish the tasks at the level, and under the conditions specified in the objectives.

For assessment we aimed to be formative, authentic, and transparent (Sadler, 1989). We view assessment as information-rich feedback providing students with information regarding their progress towards attainment of the objectives. Thus, assessment takes place continually in class and the teacher provides the student constant feedback. With task-based exercises in class, practical exercises in our simulated village (where students interact with role-players performing the type of tasks presented in class), as well as traditional, weekly computer, oral, and paper-based tests, students are provided with rich, constructive feedback on their overall progress. Scaffolding is embedded throughout the curriculum which requires students to do tasks over time with less and less support and more independence at higher levels of performance. This approach requires the teacher to give a lot of informative guidance and feedback so that the students clearly know where they need to be in order to perform independently without the benefit of a teacher's support.

One aspect of this curriculum that is very different for teachers is a focus on accomplishing specified tasks under specified conditions and only up to predetermined standards. Many teachers are accustomed to teaching a grammar point or a notion, from beginning to end until mastered and then moving on to the next point or topic. Because this curriculum is task-focused, it requires teachers to teach only what is needed in order to accomplish the particular task at hand. Due to the nature of this task focus, we carefully and deliberately sequenced linguistic functions, grammar, and notions such as numbers, dates, etc. All major aspects for the target proficiency level (grammatical, lexical, etc.) are presented in the curriculum. However, the students do not get them all at once; rather, they get them in context as they become needed to complete a specific task. For this reason, a lot of spiraling (revisiting tasks or functions with different applications) and scaffolding (revisiting tasks with less and less support as more responsibility is put on the student over time) is necessary. All of this was built into the underlying framework of the curriculum plan so that developers had a clear map as to what needed to be covered where and to what degree. To that end, students do develop a general proficiency, yet the curriculum is not proficiency-based. The proficiency items are interwoven within the context of the tasks and linguistic functions needed to perform the job of a Navy SEAL.

The guiding policies and procedures for this program are an integral part of keeping all of the items

aligned. Due to its intricately interwoven design, teachers, students, and administrators must follow policies and procedures closely to keep it true to its purpose.

### *Evaluation*

The evaluation component of the approach we adopted (Federe, 2005) calls for evaluation throughout the entire development process. We opted for a systematic and multi-faceted process that involved pilot runs of Units 1 and 2 (four- and eight-week iterations), formal and informal evaluations and feedback from teachers and students, group discussions, and an open-door policy for students and teachers alike. Mirroring the recursive and spiraled design of the curriculum, we also developed an evaluation plan for the courses as a whole. This plan includes a year-by-year plan for the next five years, as well as a basic evaluation protocol to ensure currency, relevancy, and validity. In this plan, all the stakeholders are identified along with the stakes, as well as the methodology to ensure continued evaluation against those factors.

### **Results**

Weekly quizzes, unit tests, and an end-of-course test are built into the curriculum. All of these exams are formative in order to assist the students, teachers, and administrators evaluate individual attainment of the course objectives. In addition, we administer the Defense Language Proficiency Test (DLPT) and the Oral Proficiency Interview (OPI). While the DLPT and OPI do not adequately assess our goals and objectives, they are currently the only official test of record per Navy and Department of Defense guidance. In the next phases of this project, one of our goals is to develop an assessment which more directly measures attainment of an operational capability. We are at the development stage of this next phase.

Considering our goals and objectives are not tied to the current, high-stakes DOD proficiency exams (the DLPT or OPI) in any way, comparison of the results (Figure 4) is statistically noteworthy and implications for second language curriculum design can be extrapolated.

The first set of data (Figure 4) compares the results of the general-proficiency, Pashto Initial Acquisition Training (IAT) from 2010 to present, to the results of an IAT class piloting Unit 1 (first four weeks) of the NSW-specific, Pashto curriculum. The column titled Old Curriculum shows DLPT and OPI results for students that were a product of the commercial curricula available at the time. The column titled New Curriculum shows DLPT and OPI results for a class of students that piloted Unit 1 of the new curriculum and then transitioned to the traditional texts for the remaining eight weeks. On the publish date of this case study, the data for Pashto were limited to this four-week pilot due to current operational demand, but the positive results can already be seen in terms of the students' reading scores on the DLPT. It should be noted that the curriculum was just completed in December 2013.

When comparing classes from the same DOD categories of difficulty (CAT III/IV), one could extrapolate like results for Pashto as depicted in figure 5 for Farsi and Arabic had we been able to run the full, 12-week course of instruction (COI) for Pashto.



<b>Pashto Old Curriculum Results                      Compared to                      Pashto 4-week Pilot Class                      Overall Pass Rate*</b>		
<b>Modality</b>	<b>Old Curriculum Pass Rate</b>	<b>New Curriculum Pass Rate</b>
OPI-Speaking ( S )	99.15%	100.00%
DLPT-Listening ( L )	65.81%	66.67%
DLPT-Reading ( R )	49.57%	66.67%
*ILR 0+ or higher on DLPT and/or OPI for CAT III/IV Languages ILR 1 or higher for CAT I/II Languages		

Figure 4 Pashto Only Results

The data in figure 5 depict the results of 12-week IATs including the six target languages chosen for the curriculum project—Tagalog (TA), French (FR), Spanish (QB), Persian-Farsi (PF), Pashto (PU), and Modern Standard Arabic (AD) from 2010 to present. The column titled Old Curriculum shows DLPT and OPI results for students that were a product of the commercial, general-proficiency curricula available at the time. The column titled New Curriculum shows DLPT and OPI results for students that completed the full 12 weeks of the new curriculum. These final scores are statistically striking for several reasons; particularly as the students did not follow a general-proficiency curriculum, which the DLPT and OPI assess, but rather a task-based, mission-focused curriculum. Their scores also exceeded those normally associated as an outcome of a general-proficiency course.

We will continue our analysis as the data set grows over time, but of particular interest will be the effect of a performance-based curriculum on the development of general proficiency. If preliminary results show that this performance-based curriculum was capable of producing not only the level of performance needed by our command, but also a higher, general proficiency, a tremendous amount of time and money could potentially be saved on language acquisition, and most importantly, the production of a student capable of doing more with the language.

<b>TA, FR, QB, PF, PU, and AD Old Curriculum Results                      Compared to                      New Curriculum Classes                      Overall Pass Rate*</b>		
<b>Modality</b>	<b>Old Curriculum Pass Rate</b>	<b>New Curriculum Pass Rate</b>
OPI-Speaking ( S )	92.99%	100.00%
DLPT-Listening ( L )	48.77%	70.83%
DLPT-Reading ( R )	58.58%	79.17%
1/1/1 All Modalities ( L, R, S )	58.58%	79.17%
*ILR 0+ or higher on DLPT and/or OPI for CAT III/IV Languages ILR 1 or higher for CAT I/II Languages		

Figure 5 New Curriculum vs. Old Curriculum Results



## **Significance**

The significance of this work is threefold. In the context of strategies, one could argue this case study has significance at the Tactical (for Teachers), the Operational (for Programs), and Strategic (For Future Research) levels.

### *For Teachers*

At the Tactical (practical) level, this study outlines a methodical approach for teachers to delineate clear objectives, present those objectives in an engaging manner, and measure those objectives in meaningful and productive ways. The recursive, and spiraled approach of both content delivery and formative assessment, builds a firm foundation for both students and teachers. It is this process of iterative reflection, application, refinement, expansion, and further reflection that allows for growth not only for the student, but also the instructor and his/her teaching.

### *For Programs*

At the Operational (developer) level, this study demonstrates the utility of a thorough needs assessment coupled with clearly defined objectives, and a deep knowledge of learner and language learning theory. The resulting product yields rewards not only for the student, but the program as a whole. Maintaining a learner-centered approach as outlined in this paper, is instrumental to a program's success. Such an approach also allows for maximum flexibility and the opportunity to adapt and update the curriculum as needed for the students. Echoing the reflective practices outlined for the teachers and students, the program too, must undergo a process of reflection and refinement to ensure the content presented in the curriculum continues to meet the needs of the students over time.

### *For Future Research*

At the Strategic level, there are several implications of this case study on future research. If one were to take this curriculum project to its logical conclusion, there is one piece of this whole process that remains misaligned. The objectives were based on a thorough needs assessment. The lessons were based on those clearly defined objectives and the internal assessments, and reflective exercises were designed to measure and reinforce those objectives. The formal, high-stakes assessment piece, however, remains an external construct that lacks validity (Messick, 1990; Tyler, 1949) within the context of our defined needs and learning objectives. While our students performed surprisingly well on these standardized, proficiency-based assessments, it remains to be seen if an equally recognized form of final assessment can be created based on the authentic tasks they need to perform operationally and how well they are able to perform them. Until such an assessment is created, we cannot fully measure our students' capabilities vis-à-vis a formally recognized standard.

## **Special Thanks**

Those of us in the Naval Special Warfare, Language Regional Expertise and Culture Program, would like to express our tremendous gratitude to those non-government organizations that made this curriculum project successful including Surface, Ward, and Associates Consulting (SWA), American Council on the Teaching of Foreign Languages (ACTFL), Diplomatic Language Services (DLS), and Mid Atlantic Professionals- SSI (SSI).

## References

- Anderson, L. M., Blumenfeld, P., Pintrich, P. R., Clark, C. M., Marx, R. W., & Peterson, P. (1995). Educational psychology for teachers: Reforming our courses, rethinking our roles. *Educational Psychologist*, 30(3), 143-157.
- Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist*, 28(2), 117-148.
- Blumberg, P. (2009). *Developing learner-centered teaching*. San Francisco: Jossey-Bass.
- Brown, H. D. (2010). *Language assessment: Principles and classroom practices*. Longman Pub Group.
- Brown, H. D. (1994). *Teaching by principles: An interactive approach to language pedagogy* (Vol. 1, p. 994). New Jersey: Prentice Hall Regents.
- Brown, H. D. (1993). After method: Toward a principled strategic approach to language teaching. Georgetown University round table on languages and linguistics, 1993, 509-520.
- Brown, J. D. (1995). *The elements of language curriculum*. Boston, MA: Heinle & Heinle.
- Coady, J., & Huckin, T. (Eds.). (1997). *Second language vocabulary acquisition: A rationale for pedagogy*. Cambridge University Press.
- Dewey, J., & Small, A. W. (1897). *My pedagogic creed* (No. 25). EL Kellogg & Company.
- Evans, C. (2013). Making sense of assessment feedback in higher education. *Review of Educational Research*, 83(1), 70-120.
- Federe, M. (2008, August). Quality assurance surveillance program: Course evaluation. Presented at the Annual SOFLO Conference, Ft. Walton Beach, FL.
- Howatt, A. (1984). *A history of English language teaching*. Oxford: Oxford University Press.
- Hutchinson, T. (1987). *English for specific purposes*. Cambridge University Press.
- Klimoski, R. J. and Zukin, L. B. 2003. Psychological Assessment in Industrial/Organizational Settings. *Handbook of Psychology*. Two:317-34
- Knowles, M. S., & Holton, E. F. III, & Swanson, RA (2005). *The adult learner: The definitive classic in adult education and human resource development*.
- Knowles (1983). *The adult learner: A neglected species*. Houston: Gulf Publishing.
- Knowles, M. S. (1980). *The modern practice of adult education: From pedagogy to andragogy*. (2nd ed.) New York: Cambridge Books.
- Krashen, S. (1982). *Principles and practice in second language acquisition* (pp. 65-78). Pergamon: Oxford.
- Mager, R. (1984b). *Preparing instructional objectives* (2nd ed.). Belmont, CA: Pitman Management and Training.
- Mahn, H. (2013). Vygotsky and second language acquisition. *The Encyclopedia of Applied Linguistics*. Oxford, UK: Wiley-Blackwell.
- Merriam, S. B., & Caffarella, R. S. (1999). *Learning in adulthood: A comprehensive guide*. (2nd ed.). San Francisco, CA: Jossey-Bass.
- Messick, S. (1996). Validity and washback in language testing. *Language testing*, 13(3), 241-256.
- Messick, S. (1990). Validity of test interpretation and use.
- Mezirow, J. (1997). Transformative learning: Theory to practice. *New directions for adult and continuing education*, 1997(74), 5-12.
- Mezirow, J. (1981). A critical theory of adult learning and education. *Adult education quarterly*, 32(1), 3-24.
- Münsterberg, H. (1913). *Psychology and industrial efficiency* (Vol. 49). Continuum.
- Nunan, D. (1996). *The learner-centered curriculum*. New York: Cambridge University Press.
- Nunan, D. (1987). Communicative language teaching: Making it work. *ELT journal*, 41(2), 136-145.
- Peer, K. S., & McClendon, R. C. (2002). Sociocultural learning theory in practice: Implications for athletic training educators. *Journal of Athletic Training*, 37(4 suppl), S-136.
- Sadler, D. R. (2010). Beyond feedback: Developing student capability in complex appraisal. *Assessment & Evaluation in Higher Education*, 35(5), 535-550.
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18(2), 119-144.
- Strauss, S. (1996). Confessions of a born-again constructivist. *Educational Psychologist*, 31(1), 15-21.

- Surface, E. A. (2012). Training needs assessment: Aligning learning and capability with performance requirements and organizational objectives. *The handbook of work analysis: The methods, systems, applications and science of work measurements in organizations*, 439-464.
- Tumposky, N. R. (1984). Behavioral objectives, the cult of efficiency, and foreign language learning: are they compatible?. *TESOL quarterly*, 18(2), 295-310.
- Tyler, R. W. (2013). *Basic principles of curriculum and instruction*. University of Chicago press.
- Valette, R.M. (1980). Evaluating the second-language learning program. In F.M. Grittner (Ed.), *Learning a Second Language*. Chicago, University of Chicago Press.
- Vygotski, L. L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Weimer, M. (2013). *Learner-centered teaching: Five key changes to practice*. John Wiley & Sons.
- Widdowson, H.G. (1978). *Teaching language as communication*. Oxford: Oxford University Press.
- Wilkerson, L., & Irby, D. M. (1998). Strategies for improving teaching practices: a comprehensive approach to faculty development. *Academic Medicine*, 73(4), 387-96.

## **Afghan Languages (Dari and Pashto) as a Source of Unity Rather Than Division**

Johnathan Krause  
*Western Carolina University*

### **Abstract**

The government in Afghanistan, and more specifically the ministry of Education, clearly has generated a certain level of conflict between linguistic communities through its educational policies. There has been linguistic tension in Afghanistan for quite a long time and Dari-speaking students at the University of Kabul have demonstrated these tensions through protests. The protestors demanded that Kabul University (and Mazar University) use the word for university in both languages, *Daanish Gah* in Dari and *Pohantoon* in Pashto. However, the government, like its predecessors, insisted on being monolingual and showing deference to Pashto. Reaction to the protests triggered clashes with police and reignited ethnic tensions, as groups of Pashtuns heckled the students in the latter language. Indeed, language is a divisive issue in Afghanistan.

Language, however, should be a source of unity. According to the Asia Foundation's most recent statistics, nearly half of all Pashtuns report being speakers of Dari. Other experts believe that the percentage of Dari-speaking Pashtuns is higher.

Statistics tell one story, but the commonalities between the two languages are striking and present a better case for unity. When people observe a phenomenon, it has a greater effect than simply reading about it. University of Wisconsin Professor Robert Kaiser, in his research on election fraud in Russia, has noted for instance that seeing officials commit fraud on television had a far greater effect than reading about it. My paper will present these points of intersection with the tacit intention of showing that seeing is believing: seeing commonalities between the languages.

In the same vein, I want to emphasize the commonalities of the two languages through visual aids and go on to make suggestions from a linguistic and not a political scientist point of view, as to why these linguistic divisions need not exist (Mariam Alamyar, a Purdue Alumnus, wrote a similar thesis on this issue, but she focuses mainly on power politics).

In doing so, I include a sample of *Tabestaney Zalmay* (Zalmay's summer) in the text, a short narrative about a Pashtun college student and his family. The narrative will show numerous words that are common to Pashto and Dari in red and few words in black that are unique to Dari. In particular, this paper will show a surprising observation about the two words, *Danesh Gah* and *Pohantoon*, which have caused linguistic division. Both words appear early in the narrative.

In addition to the narrative, I will also show that there is a high amount of positive transfer between the two languages. Both share similar word orders and have a plethora of common verbs and subject matter. There is also negative transfer between the two languages but whether there is more negative than positive transfer is not a concern of this paper. Scholars tend to argue that there is more of the former than the latter between languages. To do so would deviate from the topic. However, I have to briefly mention both forms of transfer since the paper deals with comparing two languages.

### **The background behind linguistic tension in Afghanistan**

For years Afghanistan has had many divisive issues. The most over looked issue, aside from corruption, poverty, and the insurgency, is that of language. Indeed, there has been linguistic tension in Afghanistan for a long time. The government and its predecessors have clearly generated conflicts between linguistic communities (Dari and Pashto) through their educational policies. Often, the governments in Kabul (or different language groups within the government) have tried to promote their language over the other (Alamyar, 2010).

From the reign of Zahir Shah to the Taliban Pashto was promoted as the medium for instruction and administration (Alamyar, 2010). The first government failed to utilize Pashto in different settings and thus the Ministry of Education recommended a compromise (Alamyar, 2010). The government created Dari schools in areas where the majority of people were Dari speakers and vice versa with Pashto speakers (Alamyar, 2010). Yet, higher instruction would be in Dari. Implementation of this program failed and teachers were unable to teach in Pashto (Ibid). Alamyar (2010) notes in her thesis that different types of conflicts were created, especially as another administration encouraged people to raise the issue of language.

Currently, the Ministry of Education’s curriculum and the Constitution’s language policy calls for language segregated schools (like those of earlier governments) in areas where one group is prevalent (Alamyar, 2010). Even within the government workplace, language groups generate linguistic tension by discriminatory practices. Alamyar (2010) notes that some Farsi Dari (or simply Dari) and Pashtu speakers in different governmental (and NGOs) discriminate against each other based on the language they speak.

Sometimes Dari speakers have reacted with resentment to these policies, especially when one considers Mustafa Siddiqi’s experience, as included in Alamyar’s thesis (2010). He attended Kabul University when the Taliban were in power. He had serious issues with a teacher who was sent by that regime. Siddiqi tells of how the teacher refused to teach in Persian and cursed at the students when they made that request (Alamyar, 2010). His contempt for the teacher is clearly expressed as Siddiqi referred to the later as a “turbaned, inexperienced, and inept teacher-I can’t call him teacher...he was really non-professional (Alamyar, 2010).”

Resentment even spilled out into the streets. Arguably the most significant source of linguistic tension stems from the dispute over the word university. Government policy favors the Pashto word for university پوهنتون instead of the Dari word دانشگاه. The first word is Pohantoon and the second word is Daanish Gah. Dari-speaking students at Kabul University and Balkh University in 2008 demanded that the latter term be included on university signboards (Alamyar, 2010) Their attempts to hang a sign in Dari were unsuccessful as police in both instances prevented this from happening (Alamyar, 2010). The protests turned violent as the police and protesters clashed with each other, resulting in several injuries. The protests also reignited ethnic conflict as an Afghan who listened to Tolo TV blogged that “There was another mob talking in pashtu speaking against them (A. Muslim, 2008)”. He means a group of Pashtuns were heckling the Dari-speaking protestors.

Physical clashes over the word were not limited to Dari-speaking students and police, but another group. A bill regarding higher education was left unresolved because of a dispute of the word university. Ariana News (2013) reported last year that:

the Wolesi-Jirga has not been able to come to agreement on whether to use both the Pashto word for University (Puhantun) and the Farsi word (Danishga) in the portion of the law that pertains to universities.

Some of the MPs in the Lower House of Parliament wanted to only use the word Pohantoon in legislation that dealt with universities. There were those who wanted to only use Daanish Gah and others who wanted to use both. Ariana News reported that the disagreements led to loud shouting and physical assaults (Ibid). The picture on my presentation poster effectively shows that language can be a source of tension.



Photo from Ariana News

### **Afghan Language as a source of unity**

Language, however, should be a source of unity. According to Keith Shawe, who worked on the Asia Foundation’s most recent Survey of the Afghan People (2013) nearly half of all Pashtuns who participated in the survey reported that they



could speak Dari. Based on Alamyar's statistics, 90% of Pashtu speakers understand and speak Farsi Dari and 50 to 65% of Farsi Dari speakers can understand Pashtu (Alamyar, 2010). Certainly, there are reasons for ethnic strife that transcend linguistics, but my study illuminates the fact that linguistic differences alone need not be a significant source of division.

One would take the same idea from Alamyar's thesis (2010). She asserts that language issues are among those who are high authorities and have power, not ordinary people. Indeed, she even notes that Afghan hospitality erases the language issue completely. She claims that when a Pashtun visits a Farsi Dari speaking-family, he need not speak Farsi Dari or, if the host family knows Pashtu, they should speak their guest's language. Likewise, when a Farsi Dari speaker visits a Pashtun family, the Pashtun family should speak in Farsi Dari to give more respect and show distinguished behavior towards the visitor (Alamyar, 2010). Indeed, the majority of her interviewees, which included a sizeable amount of Pashtun students, feel "comfortable" or "very comfortable" receiving instruction in Dari.

### **An approach to solving linguistic tension**

Alamyar's interviews, statistics (as well as those from the Asia Foundation), and discussions of the cultural and political dynamics tell one story, but my study will show points of intersection of between Dari and Pashtu through linguistic evidence with the intention that seeing is believing: seeing commonalities between the two languages.

When people observe a phenomenon, it has a greater effect than simply reading about it. University of Wisconsin professor Robert Kaiser, in his research on election fraud in Russia, has noted for instance that seeing officials commit election fraud on television had a greater effect than reading about it (2012). In the same vein, I want to emphasize the commonalities of the two languages through visual aids and go on to make suggestions from a linguistic and not a political scientist point of view, as to why these linguistic divisions need not exist.

In doing so, I will include sample pages from Tabstaney Zalmay (Zalmay's summer), a short narrative in Dari about a Pashtun college student and his family. The narrative will show numerous words that are common to Pashtu and Dari in red and few words in black that are unique to Dari. In particular, the narrative will show a surprising observation about the two words, Daanesh Gah and Pohantoon, which have caused linguistic division.

In addition, I will show that the two languages have a high amount of positive language transfer through Tabestaney Zalmay and examples taken from my language courses at Indiana University and the University of Wisconsin. Positive Language transfer is when the forms, structures, vocabulary, and rules of the L1 and L2 coincide (Winer, 1989).

Before continuing it is important to mention that there is negative transfer between both languages. From my experience Pashtu requires learning more rules and sentence structures that can take more forms than Dari. Although Alamyar (2010) notes that there are both similarities and differences between Pashtu and Dari her interviewees believed that Dari speakers have had difficulty learning Pashtu and Pashtu instructors have had trouble themselves with the use of "ya" (a Pashtu letter that is similar to y in English). Scholars generally believe that there is mostly negative transfer between languages. Lise Winer (1989) argues that real and perceived closeness of language distance between similar languages (Trinidadian English and Standard English in her article) contributes greatly to persistent difficulties students have in writing English (standard).

I will not dispute Winer or Alamyar's findings nor question the interviewees on their beliefs about language. Nevertheless, I will make a strong case for positive transfer based on the commonalities in Tabestaney Zalmay and numerous examples in the last section. In addition, my language courses greatly assisted subsequent languages that I took. Pashtu provided the foundation for Tabstaney Zalmay, which was compiled from several



lists of common words between Dari and Pashto. A Dari speaker and/or linguist learning Pashto could learn a lot from a similar list and build a strong foundation.

### The case against linguistic tension: Tabstaney Zalmay

I wrote Tabstaney Zalmay last summer (2013) at Indiana University. It is truly a blend of linguistic syncretism. It was my project for Dari where I took as many words that are common to Dari and Pashto and made them into a narrative. In addition it focuses on a Dari-speaking Pashto family. Usually, language and ethnicity are correlated in Afghanistan (Alamyar, 2010).

Most importantly, it visually (and thus effectively) shows how Dari and Pashto are sources of unity due to the abundance of common words. Some words in both languages are spelled exactly the same; others have slight differences but are the same word and have the same meaning. Nevertheless, linguists and non-linguists alike can clearly see the commonalities in red (pages 6-9):

زلمی دروازه خانه اش را باز کرد و داخل شد. مادرش کابلی پلو پخته کرده بود و پدرش يك چوکی را جور می‌کرد. زلمی به پدرش گفت: "پدر! سلام علیکم. مانده نباشید، شما چطور استین؟" پدر جواب داد: "وعلیکم سلام زلمی جان، زنده باشی من خوب هستم. تو چطور استی؟" زلمی گفت: "مه هم خوب استم تشکر. سلام علیکم مادر، خسته نباشید، شما چطور هستین؟" مادرش جواب داد: "وعلیکم سلام بچیم، سلامت باشید من خوب هستم تشکر. تو چطور استی؟ زلمی جواب داد: "مه هم خوبم. از دیدن همه شما زیاد خوشحال هستم." خواهر زلمی گفت: "سلام زلمی!" زلمی گفت: "سلام مریم، تو چطور هستی؟" مریم جواب داد: "مه خوب هستم تشکر لالا جان."

بعد از احوالپرسی آنها در میز نان شب را خوردند. پدر زلمی گفت: "ما صبا به بازار خواهیم رفت و میوه، ترکاری و گوشت مرغ خواهیم خرید." زلمی جواب داد: "خیر، خوب است."

روز بعد زلمی، پدرش و کاکایش که پاچا نام دارد، به بازار رفتند. آنها در يك دوکان بسیار میوه های تازه را دیدند. در آن دوکان، ام، انگور سرخ، انار، شفتالو، مالت و کیله بود. اما میوه ها بسیار قیمت بودند. آنها میخواستند که از دوکان بیرون بروند که دوکاندار صدا کرد: "صبر کنید، صبر کنید، من به شما ارزان تر خواهم فروخت." پدر زلمی گفت: بسیار خوب. پس آنها ازان دوکان پیاز، متر، کچالو و سبزی پالك هم خریدند و باز به دوکان دیگر رفتند. ازان جا گوشت مرغ و گوشت گاو خریدند. آنها بعد از بازار به خانه شان رفتند.

The narrative also questions the validity of linguistic tensions over the word university. It includes the controversial words Daanesh Gah and Pohantoon that are common to both languages. Logically speaking, since Gah گاه (place) is common to both languages, as seen in Daanesh Gah (دانشگاه) in Dari means university (found on page 2 of the narrative)) and Lashkar Gah لشکرگاه (Pashto (and Dari) word for a city in Helmand Province (found on page 4 of the narrative)) then there should be little, if any dispute over the word for university. In addition, Pohantoon پوهنتون (on page 2) is a Pashto word commonly used in Dari, as seen in Dari dictionaries (C. Bulkin, 2012) and Dari language texts (R. Arman, 2012).

These two words are manifested together in the narrative on page 2:

Zalmay is a student in (at) Kabul University (excerpt from page 2).

## پوهنتون کابل / دانشگاه ی



Photo from Afghan Ministry of Higher Education Website

## زلمی در پوهنتون کابل (دانشگاه ی کابل) محصل است.

The narrative also contains university-related words that are indisputable and common to both languages. I interviewed Omid Wali Shirzad, assistant professor at Nangarhar University and a Fulbright FLTA at the University of Georgia. He mentioned that all other university-related terms are accepted by both speakers of the two languages (Shirzad, 2014):

a)Table 1

Adjunct Instructor (page 12)	نامزد یو هیالی
Junior Teaching Assistant	یوهنیار
Senior Teaching Assistant	یوهنمل
Professor (page 12)	یوهاند / استاد

Thus, the narrative shows that there should not be any linguistic tension over university terms since the vast majority is accepted by Dari and Pashto speakers. Even the word for university is logically a source of unity because it is common to both languages.

### The case against linguistic tension: positive transfer in both languages

The narrative, in visually portraying the abundance of common words, shows that there is a high amount of positive language transfer between the two languages. Indeed, Pashto and Dari share many common lingu-

tic elements. This section will visually illustrate those elements. It will have minimal discussion because the examples will speak for themselves. It will read like a math text book.

I. There is a high amount of Positive Transfer between the two languages (Krause, 2009-2013):

A) There are many common words and subject categories between the two languages

1. Words pertaining to family are similar:

a) Table 2

English	Dari	Pashto
father	پدر	پلار
mother	مادر	مور
sister	خواهر	خور
paternal uncle	کاکا	کاکا

b) Dari and Pashto share many words for geographical places (i.e. “cave” غار), colors (i.e. “red” سرخ), fruits (i.e. “peach” شفتالو), and numerous other categories.

B) Both languages share word order

1. Both follow (Subject-Object-Verb)

a) I eat grapes.

من انگور ها ميخورم. (Dari)

Verb      Object Subject

زه انگورونه خورم. (Pashto)

C) Some verb tenses have the same structure and placement in both languages.

1. They share three tenses.

a) Present simple

زه اوبه څښم. (“I drink water” in Pashto.)

Personal ending for first person and present stem

من کار ميکنم. (“I am working” in Dari.)

b) Near Past

ما شربت څښلې دے. (“I have drunk cough syrup” in Pashto.)

To be Main Verb

او دري را درس داده است. (“He has taught Dari” in Dari.)

c) Far Past

ما شربت څښلې وو. (“I had drunk cough syrup” in Pashto.)

To be Main Verb

او دري را درس داده بود. (“He has taught Dari” in Dari.)

D) Placement of Adverbs is similar to both

1. “very” is similarly placed

دا هر څه هلته ډیر ارزان او تازه وو. (“All these things are very cheap and fresh over there” in Pashto.)

Descriptive word very

این کیک بسیار زیبا است (“That cake is very pretty” in Dari.)

E) Both languages contain similar verbs

a) Table 3

English	Dari	Pashto
To arrive	رسیدن	رسیدل
To cook	پختن کردن	پخول
To work	کار کردن	کار کول
To eat	خوردن	خورل

F) Negative transfer between Dari and Pashto.

1. Words that look the same have different meanings.

a) به in Dari means “to” and in Pashto means “will (do)”.

b) او in Dari means “he/she/it” and in Pashto means “and”

c) ما in Dari means “we” and in Pashto is the past tense for “I”

2. The gender of the verb in Pashto will be dependent on the object and Dari does not have this rule.

a) مایو خټکے او خورلو (“I ate a melon” in Pashto.)

Masculine ending of object influences verb ending

b) Not in Dari

ما اب ریختیم (“we pour water” in Dari)

3. Dari uses prepositions while Pashto use circumpositions

a) Table 4

English	Dari	Pashto
after	بهداز (word)	د (word) نه پس
in	در (word)	په (word) کښی

G) However, colloquialisms can turn negative transfer into positive transfer

1. Circumpositions can be omitted in Pashto colloquial speech and writing. As a result, both languages can have a similar sentence structure in regards to preposition placement.

a) زه د پښور يم ("I am from Peshawar" in Pashto.)

From

مه از پشاور هستم. ("I am from Peshawar" in Dari.)

There are many other examples of negative and positive transfer but for the sake of this paper all of them will not be listed.

## Conclusion of Results

Despite the linguistic tensions that Afghanistan has experienced, this paper has proven that Afghan languages (Dari and Pashto) are a source of unity rather than division due to the large amount of positive language transfer between them, as seen in Tabestaney Zalmay and the linguistic evidence shown in section 3. Thus, there should be no linguistic tension in Afghanistan and Alamyar points this fact out as well. However, she takes another approach to the issue and focuses on a political, instead of a linguistic, point of view.

**Disclaimer:** This paper is current up to the Karzai Administration. It was written before the last presidential election. President Ghani's policies are unknown to the author at this time.

## References

- Alamyar, Mariam (2010). "Language controversy in Afghanistan: The Sociolinguistics of Farsi Dari and Pashtu at Kabul University". A Thesis Submitted to the Faculty of Purdue University. Purdue University, West Lafayette, Indiana.
- Arman, Rahman (2012). *Dari, An Elementary Textbook*. CelCAR. Bloomington, IN.
- Bulkin, Carleton (2012). *Dari practical dictionary*. 2nd ed. Hippocrene Books, Inc. New York.
- Kabul University. Kabul University Ministry for Higher Education webpage. <http://www.mohe.gov.af/?lang=en&p=gov&nid=30> Copyright © 2014. Ministry of Higher Education.
- Kaiser, Robert. (2012). "CREECA Faculty Roundtable: Russian Presidential Elections". University of Wisconsin. Spring Semester. March 8, 2012.
- Krause, Johnathan. (2009-2013). Classroom notes from Pashto and Dari courses. Indiana University and the University of Wisconsin.
- Muslim, Abu. "Daneshgah/Pohantoon Protests in Kabul?" November 17, 2008. Accessed May 13, 2014. <http://tajikam.com/forum/index.php?topic=4928.0>
- No Author. "Lower house disapproved the draft higher education plan, quarrels continued." Ariana News. May 26, 2013. Accessed May 13, 2014. <http://ariananews.af/regional/lower-house-disapproved-the-draft-higher-education-plan-quarrels-continued/>
- Shawe, Keith. (2013). Conference discussion. 12/11/2013. Additional information from the Asia Foundation's A Survey of the Afghan People (2013). San Francisco and Washington.
- Shirzad, Omid Wali. Interview. The University of Georgia. April 29, 2014.
- Winer, Lise. (1989). "Variation and Transfer in English Creole-Standard English Language Learning." *The Dynamic Language Learning*. Miriam R. Eisenstein. Plenum Press. New York. 115-184.

# Functional-Semantic Fields and Teaching of Kazakh Language<sup>1</sup>

Zeinekhan Kuzekova  
Nazarbayev University

## Abstract

This paper describes the systematic functional approach (with a particular focus on the functional-semantic approach) for integrating language learning in a second language context. Using Kazakh as an example, I discuss how this framework has been successfully integrated into teaching both second-language learners and heritage speakers. This approach is especially useful in less commonly taught languages, of which Kazakh is a prime example. Among its numerous advantages include: intentionally building upon students' prior learning, fostering students' communication ability, and avoiding rote memorization of limited grammar and vocabulary. In addition, students learn topics which are communication based, and thus the structure of grammar and the grammar system are reinforced by connecting from the beginning to advanced levels based on the same topic. This method not only places the language in a contractual discourse, but also teaches heritage speakers how to use the language in a powerful way, preserving and enriching their language in different social contexts. In support of the functional-semantic approach, the textbook is organized according to the communicational and functional characteristics of the language.

Keywords: Functional-semantic field, Kazakh language, functional grammar, teaching Kazakh language

## Introduction

The motivation of language learning not only focuses on establishing vocabulary and grammar, but also using these in active communication. The most critical problem in learning a new language is how language learners become more functional in the use of the language when they are acquiring the language. In other words, gaining functionality in a language means not only mastering grammar, but also internalizing a variety of forms and actively employing a wide range of vocabulary and grammatical constructions in one's speech and writing. The theory of semantic fields first appeared in German linguistics (Guliga et al, 1969; cited from Bondarko, 2005) under the influence of similar understandings of systemization in the fields of physics and the other natural sciences. There has been much research on the functionality of language learning (Shur, 1974; Zolotova, 2004; Sidorova, 1994; Bondarko, 1983, 2003. Rysaldy, 1980, 2011; Kuzekova, 2010, 2012). Bondarko (2005) applied this theory systematically into Russian language teaching and extended it to a more in-depth level. The author established the bases of the functional systemic grammar which emphasizes the systematization of grammar in textbook materials through semantic fields. This has resulted in new methods for mastering structural grammar, which are founded upon a linguistic basis of language learning and which create a functional syntax by grouping functions of a language. In other words, students are exposed to grammatical constructions and vocabulary within specific settings and situations. Moreover, the various semantic categories of a language are identified and presented in a systematized manner in pedagogical materials, so that as students

---

<sup>1</sup> This research was supported in part by a grant from Seed grant of Nazarbayev University.



makes progress, they are introduced to an ever widening range of options for expressing the same meaning in different social contexts, at different rhetorical levels, and with different levels of complexity.

Functional grammar uses the “field” approach to the language system (Parmenova, 2002). According to Parmenova, the concept of the field is in its widest application learning grammar from the pedagogical field. Language is not seen as a system of levels (phonetics, vocabulary, phraseology, morphemes, morphology, syntax), but as a system of functional-semantic fields (FSF), each of which combines multi-level language tools, expressed in language similar or similar values. Within each field there is a certain order, a hierarchy of different levels, which are united not only by the similarity of values, but also by the correlation of functions

The priority of the center and periphery within the FSF helps one comprehend the rules for the selection of a particular language unit for a specific act of speech. In the practice of language teaching (especially in schools), it is not necessary to give up the traditional level system of language education. Rather, the use of recent theoretical research in the field of functional and communicative grammar can and should give a positive impetus to the improvement of language education.

### **Usage in Kazakh language**

The use of FSF in language-learning and language-teaching materials is widespread in the teaching of many of the major world languages as second or foreign languages (cited in Matkhanova, 2005; Rodionova, 2005, pp 158 - 165). However, functional strategies have yet to be systematically incorporated into the teaching of Kazakh language as a foreign language. Historically and currently, the standard approach to teach Kazakh as a foreign language has been to stress memorization of basic grammar while using limited forms in communication. Once students can recognize and use this limited number of forms, they are considered to have mastered the language. This approach is problematic given that the grammatical forms and vocabulary that students have been exposed represent only a small fraction of the forms and vocabulary that make up spoken and literary Kazakh. The result is that, while students may express themselves in basic terms, they continue to struggle to understand native speakers in natural speech and authentic texts. Likewise, their speech and writing lack the subtlety and flexibility that are expected in the utterances of a person considered “functional” in the language. This is, indeed, currently the case among non-native speakers of Kazakh who have attempted to learn the language through school and work. Moreover, in our experience we have found that many heritage speakers of Kazakh who are studying the language struggle to read novels and articles in Kazakh and have considerable troubles expressing their ideas or opinions in class discussions and written assignments.

There are two main reasons why FSF remains little used in Kazakh language pedagogy. The first is that research that divides the Kazakh language into semantic fields remains incomplete. The second is that the semantic fields in the Kazakh language have been yet integrated into language textbooks. Thus, our goal in this article is twofold: first, the extension of the current research on defining semantic fields in the Kazakh language; and, second, creation of a new series of textbook in Kazakh-language that structured around these semantic field textbooks for non-native speakers. The grammatical structure of the Kazakh language is complicated and multifaceted, but at the same time it possesses a unified system. In teaching Kazakh, not all aspects are included in the teaching system by using traditional teaching methods, which is structural approach. The objective investigation of the system demands new techniques, such as the most of effective of which is, in our view, the modern theory of functional semantic approach.

Recently, functional semantic theory has been applied to some fields of the Kazakh language in the works of the author who proposes this project (Kuzekova, 2010, 2012). Our goal is to widen the ranges of topics analyzing on the basis of actual linguistic material and to use the results of our research as a basis for creating a new system of tutorials for teaching Kazakh language. We should find the fields of names semantic categories in Kazakh language.

Kazakh language teaching will be organized by following topics (Kuzekova, 2010): expressions of the subject-predicate relationship; expression of object relations; characteristics of action; attributive relations; expressions of attributive and adjectival relations; expressions of spatial relations; time relations; expressions of comparison; expressions of conditional relations; expressions of negation; expressions of objections; expressions of causal relationship; expressions of purpose; expressions of different phases of action; expressions of human condition and condition of the environment; expressions of the nature of action; expressions of uncertainty; interrogative expressions; indirect speech; language communication tools; the concept of numbers, etc. We already listed 25 different categories' of semantic clusters based on the functional usage, but we think these are not enough to cover all aspects of Kazakh language. In categorizing semantic fields, we focused on the basis of spoken materials.

The structure of the functional-semantic field consists of a core part and a periphery (Bondarko, 1987). The core is the heart of the constructed grammatical categories, which possess a full set of distinguishing characteristics in their maximum concentration. The periphery includes the members of a functional semantic field in which its functional burden is less, but it is close to core part of the semantic categories. Therefore, the periphery is also supplementary for the core; in other words, the periphery will work to extend the core under the same topic. The periphery is usually based on natural communications.

We arranged the grammar and vocabulary in specific settings and situation according to functional-semantic bases and thus, these settings and situations will be extended into four progressive levels (A1, A2, B1, B2). Moreover, this approach will require students to actively engage linguistic material through interactive speaking, listening, and writing activities. In addition to these two characteristics, it will also encourage students to associate particular vocabulary, language patterns and communication styles with particular contexts. These advantages of the new textbook make sharply contrast with the current method of teaching Kazakh language, which relies heavily on the memorization of decontextualized materials. For example, if the topic is about expressing attributive relationships kills by level (Kuzekova, 2010), we will construct the linguistic elements by different levels including as flowing structures: (1) level 1 includes describing items by quantities and quality, color, and size etc.; (2) level 2 focuses on topics beginning with describing shape, the presence or absence of the property to expressing the relationship between properties or items etc.; and (3) in level 3, which is the advanced level, students describe the attributive relationships via idiomatic patterns or at a more academic level.

The core of this set of textbooks will be a group of defined semantic fields that students might typically encounter at various levels of complexity over the course of their language education. Thus, in contrast to the present system of language education, which offers a student only one or two standardized forms for expressing a particular idea, this new series of textbooks will focus on introducing students to an ever-broadening range of choices for expressing the same concepts and ideas as from basic level to advanced level through their language training. In other words, this method puts the emphasis on the recycling and accumulation of previous language components into higher level learning. In this way, by the time students reach an advanced level of Kazakh, they will command not only the basic forms (i.e. simple past, present, and future tense aspects), but also a range forms that allow them to understand native speakers and incorporate greater nuance into their own speech.

The theory of functional semantic fields is considered the basis of one of the most modern techniques for the description of language. Furthermore, in our opinion, the issue we discussed here can be applied to the study of any objective phenomenon in addition to linguistics. In the structural approach, which is based on grammar and linguistic knowledge, language learners can only acquire grammar, but they cannot effectively incorporate the skills into their communication, and, thus, their acquisition levels still remain tightly bound to limited knowledge of the grammar rules. This prevents students from uncovering and appreciating the richness of the language.

## **Advantages of this approach**

When compared to the structural approach, the advantages of using the functional-semantic approach in language teaching are based on the fact that language is considered as a functional-semantic system. The purpose of the functional-semantic approach is to systematize vocabulary and syntactic structures. For example, in the teaching of the expression time through the functional-semantic approach, we minimize the grammar tools at the same time that we combine the communicative-functional characteristics of the time expressions in teaching. This topic spans the beginning to advanced levels of Kazakh. At the beginning level, the topic includes the expression of simple clock time (1 p.m., 2:35 p.m., etc.), expressing dates (29th of April, etc.), expressing the time of day (morning, day time, evening) and meal times (breakfast, lunch, etc.). Next, in the intermediate level, we include the expressions of habitual action (daily, every year), intervals (from 9 to 5), expressions of beginning and ending (since the fall, starting at 5), and expressions of indefinite time (around 5:00, for about 3 hours) etc. At the advanced level, we include more complicated topics, such as expressions of time through idioms. Each level has sublevels, and sublevels have their own hierarchy status in using the functional-semantic characteristics of the topics. The same topic will be repeated at different levels of language acquisition by adding extended functional-semantic characteristics.

As currently planned, the Kazakh textbook created using the functional-semantic approach will be useful in the Philology Department of a foreign language program with future teachers of Kazakh language. The materials will be assembled specifically for achieving our primary goal, which is the teaching of the syntax of the Kazakh language through functional-semantic and structural-semantic approaches. We recognize that both approaches are necessary in order to understand the mechanism of the creation of the syntactic units and their function. At the very least, this project will encourage the research of the language from teachers' perspective and from that of the students. It will provide a platform (systematic teaching and learning) for both teacher and students on the topics, which were lost in the previous teaching system and have not been well studied. This approach is useful not only for heritage Kazakhs speakers, but also for non-native speakers of Kazakh language. Unfortunately, current teaching practices at the university level for heritage and non-native speakers of Kazakh continue to rely heavily on the structural (grammar based) teaching system. As a result, presentation of the syntax of the Kazakh language in textbooks remains fragmentary and haphazard. Students are confused and generally lose interest in learning the Kazakh language. To rectify this situation, materials will target not only non-native-speaker audiences, but will also be in demand by teachers and researchers in philological and Kazakh study programs.

This approach will bridge the current gap between new Kazakh language-teaching methodology and the existing textbooks/teaching materials. It will look to existing scholarly work on Kazakh linguistics and, in particular, FSF, to design a new series of textbooks for use in the school and university classroom. Thus, this research will simultaneously contribute to theoretical work on Kazakh language and bring those contributions from the scholarly world to the classroom.

## **Conclusion**

We use the semantic-functional approach for teaching Kazakh language for heritage and non-native speakers. In so doing, we expect to reach the following results: (1) improve the logical and content relations of the language; (2) develop associative thinking skills and logical reasoning of language learners; (3) help find semantically close components and help students apply the language appropriately; and (4) build a teaching system taking into account the connection between different language units. This approach not only fits the study and teaching of the Kazakh language, but other languages, as well. This is because the functional-semantic approach captures universal characters of languages, which include the functional semantic usage based in every language.

## References

- Bondarko, A.V., (1987). *Теория функциональной грамматики*. [The theory of functional grammar]. A.V. Bondarko (Ed.). Leningrad: Science.
- Bondarko A.V (2005). Проблема функциональной грамматики. [Problems of functional grammar.] A.V. Bondarko (Ed.). Sankt - Petersburg: Science.
- Kuzekova Z.S. (2010). Қазақ тілі практикумы [Grammar exercises]. Astana: IC-Cervic Publishing House.
- Kuzekova Z.S. (2012) Қазақ тілін оқытуға арналған типтік бағдарлама. [Typical training program for Kazakh language]. Astana: National Center for Testing.
- Parmenova T.V. (2002). Функциональный подход к изучению грамматики в школе [Functional approach to the study of grammar school], газета Русский язык [The newspaper of Russian language] №24.
- Matkhanova I.P. (2005). Поле состояния в современном русском языке. [The status field in the modern Russian language: a prototype and its environment]. Ed. Bondarko A.V. Проблема функциональной грамматики. [Problems of functional grammar.] (pp.103-113) Sankt - Petersburg: Science.
- Rodionova С.Е. (2005). Семантика интенсивности и ее выражение в современном русском языке [The semantics of the intensity and its expression in modern Russian]. Ed. Bondarko A.V. Проблема функциональной грамматики. [Problems of functional grammar.] (pp.150-166) Sankt - Petersburg: Science.
- Shur G.C. (1974). Теория поля в лингвистике. [Field theory in linguistics]. Moscow: Science.
- Rysaldy, K.T. (1980). Категория интенсивности в немецком языке. [Category intensity characteristic in modern non-German language]. Ph.D. Dissertation. Moscow.
- Rysaldy, K.T. (2011). Категория интенсивности в немецком языке. [Category intensity characteristic in modern non-German language]. monograph. Alma-Ata.
- Zolotova G.A. , Onipenko N.K., Sidorova M.U. ( 2004) Коммуникативная грамматика русского языка. [Communicative Grammar of the Russian Language], издательство русского языка им. В.В. Виноградова РАН., [Publishing House Russian Language Institute. V.V Vinogradov Russian Academy of Sciences], Moscow.
- Sidorova M. (1994) Функционально - семантические свойства имен прилагательных в современном русском языке. [Functional - semantic properties of adjectives in modern Russian]. Ph.D. Dissertation. Moscow.

# Determinant Factors in Effective Teaching and Learning of Dari and Pashto languages: An Action Research Paper

Farid Saydee, Ph.D.  
*San Diego State University*

## Abstract

The field of foreign language teaching is increasingly informed by research on the determinant factors in second language acquisition (SLA) (Shrum & Glisan, 2010). Understanding how languages are learned or acquired is the foundation for pedagogic recommendations and assists teachers in developing their lesson plans accordingly (Ellis, 1990). Some studies have demonstrated that there is a discrepancy between teachers' practices and effective teaching methods proposed by researchers. In other words, it appears that academic research has little effect on teachers' behavior inside their classrooms (Freeman, 1989).

To resolve the discord between academic research and teachers' practices, Farrel (2012) advocates reflective language teaching. He encourages teachers to be self-reflective, identify problems, develop action research projects, and use the results of their action projects to make informed instructional decisions.

The purpose of this mixed method study was to reveal the elements of an action research project that affected Dari and Pashto language acquisition in six-eight week intensive courses at a Language Resource Center (LRC) at a public university in Southern California. Students entered the program with no prior knowledge of Dari and Pashto and completed the program with a Novice High or above level of proficiency.

The researcher collected data over a period of three months. He reviewed students' responses to an online survey, gathered their test scores and class works, and interviewed seven of them. He further interviewed four teachers and conducted four formal classroom observation and frequent 3-minute walkthrough classroom observations. The data was analyzed quantitatively and qualitatively. The findings were triangulated to validate what the teachers and program supervisor believed were the determinate factors of instructional improvement at the LRC.

This paper will provide insight on the elements that affect teaching and learning of Dari and Pashto in an intensive program and encourage teachers to be self-reflective, conduct action research projects, and make informed decisions about their methods of teaching.

## Introduction

Since the late 1970s researchers have tried to understand how languages are acquired, what are the best practices and determinant factors in second language learning and acquisition (Shrum & Glisan, 2010). During the last two-three decades, foreign language based instruction has gone through tremendous changes. Language learning and acquisition approaches have been influenced by research in the areas of linguistic, psychology, and second language acquisition (Shrum & Glisan, 2010).

Understanding how languages are learned or acquired is the foundation for researchers to form pedagogic recommendations and informs teachers how to develop their lesson plans accordingly (Ellis, 1990). Teaching has shifted from emphasizing reading and writing skills to more interactive purposes such as using language in real-life situations (Shrum & Glisan, 2010). According to Ellis (1990), informal second language learning occurs with observation and direct participation in conversations (p. 2).



To understand what actually happens inside the classrooms that affect second language acquisition (SLA), it is important to gather and analyze data. Discovering the root problems enable educators to develop and implement intervention plans for better instruction and student outcomes (Frey & Fisher, 2006).

This research has been organized in the following manner: first, the researcher states the problems; second, he reviews the literature on SLA; third, he explains the research methodology; fourth, he presents his findings; and fifth, he explains the process of the action research project and its implementation and reveals some data to support the effectiveness of the intervention.

## **I. Problem Statement**

In order to be able to carry out their mission in Afghanistan successfully, the US government requires service men and women to develop language and cultural competency of the local community in Afghanistan where they will be deployed. Prior to their deployment, students are placed in Dari and Pashto language courses offered by Language Resource Centers (LRC). Intensive courses, eight weeks (320 hours) for Pashto (category IV language) and six weeks (240 hours) for Dari (category III), are designed to help students acquire an intermediate level of language proficiency. Category III languages have different linguistic and cultural aspects than English and take about 1100 hours of instruction for students to develop a superior level of proficiency, being able to converse formally and informally using extended discourse (ACTFL, 2012). Category IV languages have significant linguistic and cultural differences with English and requires learners about 2200 instructional hours to develop a superior level of proficiency (AWL, 2014). “Proficiency level” has been defined based on the criteria set by the American Council of the Teaching of Foreign Language (ACTFL). For instance, ACTFL considers an intermediate level speaker (Level 1/Level 1+) based on how he or she is able to create with language and produce sentence level discourse.

Novice High or Intermediate Low level of proficiency was the goal of the Dari and Pashto program. The researcher, who was a Dari and Pashto teacher and the director of the Program, as part of his responsibilities, frequently collected data to assess the effectiveness of the program and implement intervention plans, when necessary.

The data, collected between August-October 2012, revealed to him that the majority of students in the Dari and Pashto programs were not able to retain vocabulary and converse in the target language using short and simple sentences, by the end of program. The researcher found the teachers were not using communicative teaching approaches. To provide some context, the Dari and Pashto teachers had mostly received their formal education in Afghanistan and taught in their home country for some time where traditional teaching (grammar based) dominated the educational system.

Although, the LRC required its instructors to complete a series of teacher and professional development workshops before teaching, some teachers maintained their traditional ways of teaching and asking students to complete mundane drills. The majority of the students exited the program with novice (Level 0+) of language proficiency. In addition, there was no standard curriculum and each instructor used his or her own teaching materials in class.

The program continuously collected data, but had not taken any solid steps to analyze and use the findings for better achievement of student outcomes. If the method of instruction does not change, the program will not be able to achieve its goal – improving the level of students’ language proficiency. Consequently the students will be unable to learn and acquire the language; the program may lose its stakeholders’ support and, eventually the funding for the project would cease.

In the following sections, the researcher first clearly states the objectives for this research action plan and moves on to the purpose statement and research question:



## **II. Objectives**

### *1. Publically Validated:*

The student proficiency level of 90% for the 320-hour (Pashto) and 240-hour (Dari) course will increase from novice mid (Level 0+) to at least intermediate low (Level 1) based on the ACTFL/ILR proficiency criteria. They will be able to create with language, ask and answer basic questions and become engaged in short conversations.

### *2. Outcome through students' monitoring process*

- a. 90% of the students' weekly quiz scores will improve to at least an "A-"
- b. 90% of students will complete their homework on a daily basis.

## **III. Purpose Statement and Research Questions**

The purpose of this paper was to investigate the elements that affect acquisition of Dari and Pashto in a short-term intensive course. Specifically, the researcher was interested to know why (root causes) some students were not able to retain vocabulary and engage in short conversations with native speakers of the target language by the end of the program. The following research questions guide this research:

### *Research Questions*

1. What are the main factors that affect learning and acquisition of Dari and Pashto in an intensive six – eight weeks program?
2. How do the implementation of action plans affect student outcomes?

## **IV. Literature Review**

Some practitioners, such as behaviorists, believe that language learning occurs the same way that other types of learning occur in the classroom (Skinner, 1992). However, Chomsky (1965) claims that language learning is different in many aspects from the general theory of learning and engages a separate mental faculty. According to him, language learning more naturally takes place through social interventions with speakers of the target language (Ellis, 1990, p .3).

Behaviorists believe that habit formation plays a crucial role in language learning, especially when it comes to "error correction." However, according to Chomsky (1967), errors could be tolerated as long as it does not hinder comprehension. Krashen (1987) also agrees that "error correction" has little or no effect on language acquisition. Krashen (1987) cites Burt and Kiparsky (1972) and explains that teachers should correct errors that: (a) interfere with communication; (b) are most stigmatized; and (c) occur most frequently (p. 119).

Krashen (1987) further refers to "meaning" as the most important element in language acquisition. He claims that language acquisition occurs when learners understand the meaning of a message not when they produce grammatically correct forms. For instance, even babies, acquire vocabulary if they are given the opportunity to make connections between words and their meanings such as "this is your milk." (Shrum & Glisan, 2010).

Krashen (1987) downplays the rule of grammar as the major factor in language acquisition. He claims that grammar-translation classes are mostly conducted in the learner's first language and involve explanation of rules with examples and exercises. They are designed in a way to provide conscious practice of the grammar and vocabulary of the lesson (p. 127). Krashen introduces the five most influential hypotheses in language acquisition:

**1) The acquisition-learning hypothesis** in which Krashen refers to "acquisition" and "learning" as two different ways that adults learn a second language. He believes that adults acquire second languages unconsciously the same way that children pick up their first language (Altakhaineh, 2012). Krashen (1987) argues that generally learners are not consciously aware of the grammatical rules of the language. Learners usually acquire competence unconsciously and feel for correctness or wrongness of grammatical sentences. Krashen argues that "language acquisition" is more powerful and effective than "language learning" (p.10). On the other hand, adults learn a language consciously by understanding the forms and rule of the language (Lightbown & Spada, 2000). Krashen believes that only unconscious learning of a language will lead to fluency, whereas, those who learn the rules and forms of a language, may not be able to speak that language fluently (Lightbown & Spada, 2000).

**2) The monitor hypothesis:** Krashen argues that the learned system of a learner acts as an editor or 'monitor.' According to this hypothesis, learners focus on being correct rather than on what they have to say. Krashen claims that the monitor helps learners to do better in writing when they have more time to search for words and proper vocabulary. However, over application of the monitor slows learners' speaking ability in a conversation (Lightbown & Spada, 2000). Krashen (1987) claims that monitoring occurs under at least three conditions in order for students to apply conscious grammar rules to their utterances such as; time, focus on form, and knowing the rule. According to Krashen, second language learners need time to think about applying rules when speaking. He claims that a normal conversation does not allow sufficient time for this condition (p. 16). Dulay and Burt (1978) argue that in order to use the monitor effectively, performers must also focus on the form and think about correctness. Nevertheless, Krashen explains that heavy use of the monitor disturbs the natural order (p. 18).

**3) The natural order hypothesis:** Based on this hypothesis, second language is acquired in predictable sequences (Lightbown & Spada, 2000). Krashen (1987) argues that certain grammatical structures are acquired in early stages. For instance, Brown (1973) reported that English language learners learn the "ing" form of the verb and plural form of nouns "adding 's'" sooner compared to verb-form for the third person, adding 's' (he works) and the ('s) as showing ownership "John's hat" (p. 12).

Krashen (1987) claims that based on natural order hypothesis, instructors primarily provide input for acquisition and teachers encourage use of the target language in a low-affective filter environment. Students will not be required to speak until they are ready to do so (p. 139).

**4) The input hypothesis:** By introducing this hypothesis, Krashen (1987) tries to cast light on the question of "how languages are acquired." He explains that learners learn languages only by exposure to input that is just beyond the learner's current level of competence ( $i + 1$ ), where  $i$  is the current level of language knowledge and 1 will be the amount beyond learners' understanding (Lightbown & Spada, 2000).

In addition, Krashen encourages activities that are designed to develop listening skills. He claims that learners, especially children, build up competence in the SL by first listening to the language, which will ultimately lead to speaking ability; once sufficient listening in the form of comprehensible input is developed. According to Krashen (1987), learners need time to process and comprehend the message before beginning to produce (silent period). He claims that in a formal classroom, where languages are learned instead of acquired, learners are not given sufficient time to process and are not allowed to go through the silent period. Teachers who are not familiar with the concept of the "silent period" often ask their students to produce in the early stage of learning.

Krashen (1987) cites Newmark (1966) and argues that, in this situation, language learners usually refer to the rules of their first language for production in the second language (p. 27). Krashen claims that using L1 rules has some short-term advantages, whereas its disadvantages will be quite serious. Krashen refers to “positive transfer” when learners apply some identical rules of L1 to L2 (the contrastive analysis hypothesis). Positive transfer may help learners temporarily enhance production, but the progress will not be real. On the other hand, if there is fall back to rules of L1 to produce L2 and rules are not similar “negative transfer,” students may become frustrated.

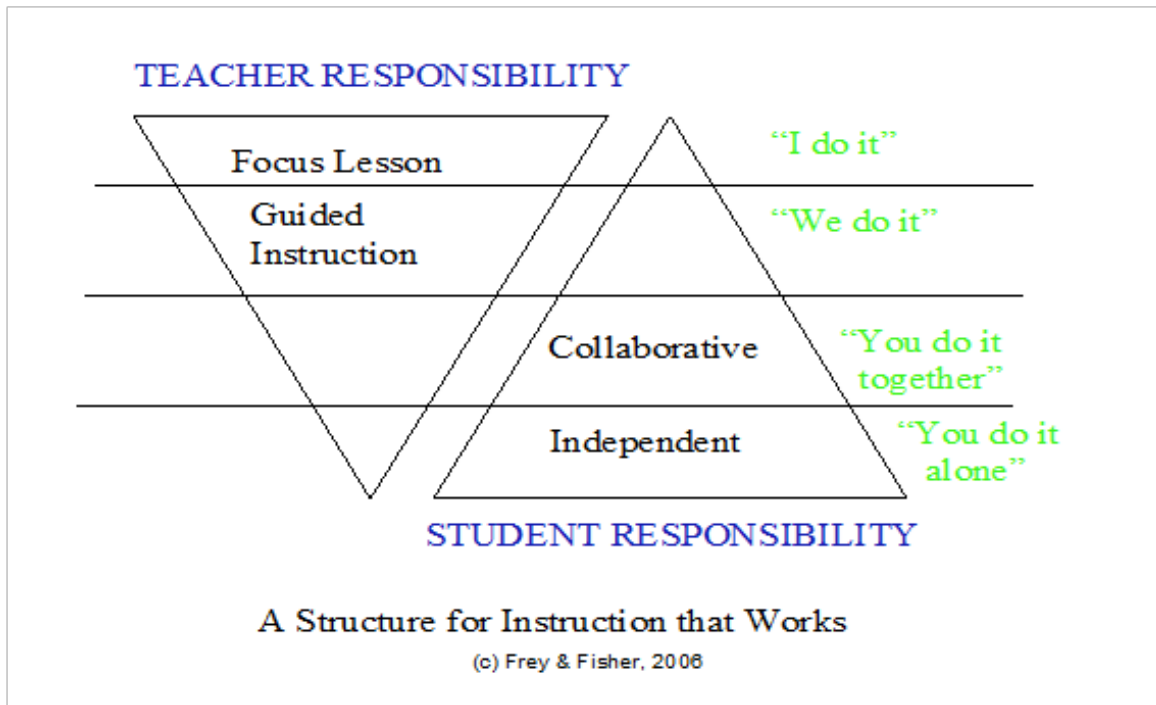
Krashen refers to “reading” as a great source of comprehensible input, however he warns that usually readings are hard to be comprehended by students. In addition, the irrelevancy of some readings to students’ interests makes the reading an ineffective means of comprehension (Lightbown & Spada, 2000). According to Krashen (1987), readings are designed to reinforce forms of the lesson and purposely sequenced grammatically. This situation violates every theory of the Input Hypothesis including “natural order,” (p. 129). Krashen (1987) claims that slower rates and clearer articulation, more use of high frequency vocabulary, fewer slang words or idioms, and shorter sentences greatly aid in the acquisition of second language (p. 55). In addition, discussing topics familiar to students will also aid in comprehension, as topics that are unknown to students will make the topic difficult to understand, and students may lose their interest (p. 67). Krashen further refers to “negotiating of meaning” as another source of comprehensible input. Some other researchers refer to this concept as “negotiating of meaning” in which explanation is given to students in a different and modified way. Krashen (1987) believes that modified and simplified language will assist learners in understanding the SLA. He refers to “modified language” as foreigner-talk and teacher talk. Krashen (1987) sees input hypothesis as necessary and sufficient for second language acquisition and ignores the role of output. He claims that learners acquire fluency not by speaking but by listening, reading, and understanding input (p. 61). Swain (1985, 1995), on the other hand, supports the output hypotheses and claims that production assists learners to observe their progress and realize the gap between “what they want to say” and “what they can actually say,” (Shrum & Glisan, 2010, Menees, 2012). Swain refers to this concept as “noticing” and argues that noticing has a significant role in SLA (Menees, 2012). Nevertheless, Krashen claims that forcing students to speak before they are ready to do so, will create anxiety and reduce language acquisition. He argues that “participation in conversation” helps language acquisition (Krashen, 2012).

**5) The affective filter hypothesis:** Krashen (1987) refers to the “affective filter” as an imaginary barrier that prevents SLA. These affects include motives, needs, attitudes, self-confidence, and emotional states (p. 30). According to this hypothesis, languages will not be acquired if the learning environment is tense and stressful (Lightbown & Spada, 2000). The concept of the Affective Filter was first introduced by Dulay and Burt (1977) as the most directly related concept to second language achievement (Krashen, 1987, p. 31). Stevick (1976), supporting the hypothesis of the Affective Filter, adds that classrooms with low affective filters provide students with better learning environments and keep students “off the defensive.”

Another factor that is considered important in SLA is “acculturation.” Krashen (1987) believes that acculturation lowers the filter and makes input more comprehensible (p. 45). Krashen (1987) cites Schumann and argues that with social integration and interaction with native speakers of the target language, learners become physiologically open to the target language and acquire it easier.

Lev Vygotsky, a Russian psychologist, acknowledges the significance of “acculturation” and adds that all cognitive development including language development occurs as a result of social interactions between people (Lightbown & Spada, 2000). Vygotsky wrote: “the source of learning and development is found in social interaction rather than solely in the mind of an individual,” (Swain, Kinnear & Steinman, 2011, p. x). To him, a learner can better learn a language if he or she interacts with adults/or with other learners. He refers to this concept as Zone of Proximal Development (ZPD). According to Vygotsky (1986), ZPD is the difference between what students can achieve alone and what they can achieve with some help of adults or more experienced peers. Based on this concept, learners learn better if they receive support from interaction with more advanced inter-

Figure 1: Scaffolding



locutors (repetition, simplification, and modeling). Another similar concept is Scaffolding. Wood et al. (1976), cited by Swain, Kinnear and Steinman (2011), defined scaffolding as “a kind of process that enables a child or novice to solve a problem, carry out a task, or achieve a goal which would be beyond his unassisted efforts,” (p. 26). Frey and Fisher (2006) explain the concept of scaffolding with the diagram in Figure 1.

Long (1983) suggests that an interaction must be modified (e.g. elaboration, slower speech rate, gesture, etc.) in order for an input to become comprehensible (Long, 1983; Lightbown & Spada, 2000). He further explains that linguistic form must be consistently modified until learners demonstrate signs of understanding (Lightbown & Spada, 2000).

Frey and Fisher (2006) claim that improving instruction is a shared responsibility and all stakeholder including leaders, staff, teachers, students, and the community must contribute to the cause. They refer to the Action Research theory in language teaching as the remedy for improving instruction and student learning outcomes.

## Theoretical Framework

### 1. Theory of Action:

This theory is based on: 1) standard-based curriculum (national); 2) building professional learning communities (PLC); and 3) professional development and capacity building of educational staff. Curriculum formed by national standards focuses on student-centered class – meaning more students participate and collaborate. Some may argue that a standard-based curriculum takes away teacher autonomy in the classroom. However, standard-based curriculum unifies programs and teachers, and promotes teacher collaboration and professional development. Ultimately, standards-based curriculum translates to schools and programs success (Darling-Hammond, Aneess & Ort, 2002) cited by Frey and Fisher (2006).

## 2. Curriculum Theories and Practice and their Implications

This theory explains that learning is planned and guided. It tries to answer “what is taught?” and “how it is taught?” In this paper the researcher reviews the curriculum from the perspective of Catherine Cornbleth (1990, p. 5) who explains: curriculum is what actually happens in classrooms e.g. “an ongoing social process comprised of the interactions of students, teachers, knowledge and milieu.” With this lens the study gathers information about the type of the curriculum implemented by the program.

## V. Method

### 1. Participants

#### a) Students:

Data was collected in the form of work samples and classroom observations between August and October 2012 from 43 students (13 Dari students and 20 Pashto students). However, eight students dropped from the study as they left the program early.

Table 1

#### *Descriptive Statistics*

Measure	N	Minimum	Maximum	Mean	Std. Deviation
Age	35	19.00	42.00	28.49	5.97
Highest level of Education	35	1.00	4.00	2.60	1.08
No. of Foreign Lang	35	0.00	3.00	0.71	0.86
Living Exp. abroad	35	1.00	2.00	1.30	0.46
Living Exp. Afghan.	35	1.00	2.00	1.70	0.48

*Education level: 1 = High School, 4 = Master*

The participants were all adults who were admitted into the language program as a job requirement. These participants entered the program with no prior knowledge of the language. Their formal education ranged from high school to graduate level degrees, and their ages ranged between 18 and 50. Some had experience living in a foreign country and knew at least one foreign language.

#### b) Teachers:

Also, four Dari/Pashto teachers who were native speakers of Dari and Pashto took part in the study. Three of them received their formal education in Afghanistan and only one completed her undergraduate level education in the United States. These teachers completed professional development and teacher training workshops offered by STARTALK (a federally funded educational program) and California World Language Project (CWLP). Their ages ranged from 26-45.

*Data Collection*

Data was collected in the form of surveys, interviews, class observations, and student work sample coupled with test scores.

**a) Student's survey**

Students were surveyed at the beginning of the course using an instrument that is attached as Appendix.

**b) Students - Interviews**

Seven students were randomly selected and interviewed from the Dari and Pashto courses. Students were asked to define "best practices" in second language acquisition and talk about their learning experience in the program (Appendix 2). Each interview lasted for about 30 minutes. All seven students knew at least one more foreign language and had lived outside of the United States. They were highly motivated and enjoyed learning languages. The interview offered an opportunity for students to express themselves and talk openly about their experience in the program. The interview (soft data) helped the researcher to better understand the significance of some factors such as one-on-one mentoring in SLA. The researcher agrees with Frey and Fisher (2006) who claim that soft data answers the question of "why" and puts a human face on the numbers.

**c) Teachers – Interviews**

Four teachers were interviewed and asked to talk about their experience in the program. Each interview lasted between 20-25 minutes. The teachers were specifically asked to reflect on their challenges. A semi-structured interview protocol (Appendix 3) was used to keep the interview organized.

**d) Classroom observations**

- **Three-minutes classroom walk-through:**

Based on the principals offered by Downey, Steffy, English, Frase, and Poston (2004), the researcher conducted many short and informal classroom observations during the reported period and collected evidence of instruction delivery and students' interaction.

- **Formal classroom observation**

During the reported period, the researcher observed four classes taught by different instructors. The purpose of the observation was to gather evidence of variation in types of instruction and class delivery. Specifically, he was interested to see teachers' instructional pedagogies and practices. He also observed student interactions and participation during class. He used an altered rubric from the STARTALK program to keep his observations organized. A copy of the rubric is attached as Appendix 4.

- **Weekly progress report**

Soft data such as homework, attendance, student's behavior and performance also collected from student weekly progress reports (Appendix 5) were entered into SPSS (quantitative research software) for analysis.



### **e) External OPI scores**

At the end of the course some students were also required to take an official OPI administered by Language Testing International LTI, a branch of ACTFL. The test score was also used to further validate the finding of this study. However, as mentioned earlier, not all students took this test.

### **f) Exam score**

Students final exam scores were entered into the SPSS to be analyzed, to find out if a correlation existed between variables and student progress.

## *2. Variables*

Variables such as age, ethnicity, education, information about foreign language, perception of Afghan culture, prior exposure to the languages, amount of tutoring, grades, OPI score, attendance, student weekly progress report, and class participation were collected over a period of three months (August – October 2012). The data was entered into SPSS and analyzed quantitatively and qualitatively.

External OPI scores were also collected to match with internal data to further establish validity and reliability of the research outcome. However, this data was not available for all students.

## **VI. Findings**

### *1. Data from SPSS*

Two OPIs were conducted internally by those teachers who were ACTFL certified Dari/Pashto testers. It is important to mention that the teachers were not testing their own students. The 1st OPI was conducted mid-program and the 2nd OPI was conducted at the end of the program. The 2nd OPI scores (internally conducted) were matched with other variables to determine if there was any relationship or correlation among them.

The figures from table 2 indicate that there was not a statistically significant correlation between age and 2nd OPI Score ( $r=.36$ ) and living experience abroad ( $r=.48$ ). However, the data revealed a slight negative correlation between the 2nd OPI score and level of education and the belief about the significance of grammar in SLA, which means students who had greater education and were more concerned with accuracy (grammar) had lower oral proficiency ratings compared to their counterparts. This finding is important as it endorses the Monitor Hypothesis that Krashen (1982) has proposed.

The data further shows a slight positive correlation between the 2nd OPI score and perception toward feedback ( $r=.06$ ) as well as with the number of foreign languages that the students knew ( $r=.07$ ).

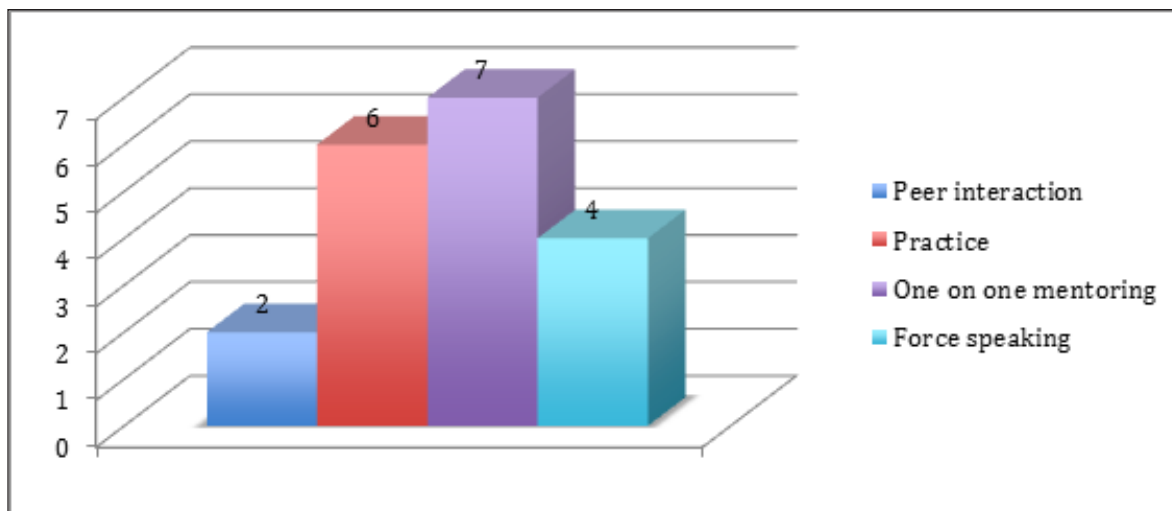
Table 2  
*Correlations Between 2nd OPI Score and other Measures*

Measure	N	Pearson Correlation
Age	35	.36
Highest level of Education	35	-.03
Perception toward Grammar	35	-.02
Perception toward Feedback	35	.06
No. of foreign language know	35	.07
Living experience abroad	35	.48

## 2. Interview - Students

As the findings from the interview with the seven students indicate (Figure 2), the variables that contribute the most to language learning and acquisition are: 1) one-on-one mentoring with instructor; 2) continued opportunities for practice; 3) and forced speaking in the target language (Dari/Pashto). However, fewer students believed that peer interaction assists in learning a second language. They believed that peers might not be in a position to correct each other and provide feedback for improvement. “Peer interaction” was rated the least important factor in learning a foreign language.

Figure 2. Students’ beliefs about effective teaching strategies



Students were also in favor of a complete immersion program where Dari/Pashto language is used in the class more than 95 percent of the time. One of the students referred to the immersion program as a “sink or swim environment.”

## 3. Classroom Observation

Four classrooms were observed formally, meaning that, the researcher spent more than 30 minutes collecting evidence of instruction delivery, student-teacher interaction using an observation rubric. The rubric

(Appendix 4) was modified from its original form to meet the specific goals of the program. He also observed classes informally (three-minute walk-through) every day to collect data. The purpose of the informal class observations were not necessarily to evaluate the teacher; rather they were to collect data on curriculum, instructional, and student level of engagement (Downey, 2004). In the following sections, the researcher first reveals his observations about the curriculum and then moves to student work sample, and formal classroom observations.

## **a. Curriculum**

**1) Data from Survey Monkey:** The majority of students were not satisfied with the curriculum used by the program. The following are quotes from the students in response to a question in the online survey about the effectiveness of the curriculum:

- Random lessons pop up in the curriculum
- The method of instructions ok. I don't really like how we have a different teacher every hour
- It was kind of hard switching instructors every hour as well as subject
- Some of the curriculum seems disjointed
- I often found it confusing and difficult to associate the homework with the lesson plan in the book

The survey form can be found in Appendix 1.

**2) Data from Teachers' Interview:** Developing standardized curriculum was the area that all teachers emphasized in their interviews. All teachers agreed that at the beginning of the program, there was not a set curriculum and they had to develop their own teaching materials. A teacher indicated that he was very nervous entering the classroom as he did not have a solid lesson plan. He was always thinking what he should do for the next session. Another instructor pointed to students' complaints and feedback as a source of frustration. According to him, students were not happy with the instructors entering and exiting the classroom to teach different sessions without coordinating with each other, everything was mixed and there was no clear instructions. Another teacher complained that the curriculum was dictated by the administration to teachers "top-down." The interview protocol is attached as Appendix 3.

## **b. Students Work Sample**

**1) Homework:** Student learning output demonstrates what and how the contents are taught (Frey & Fisher, 2006). The review of students' homework from the Dari and Pashto classes revealed pronunciation, rules of first language, and instructor feedback as the areas that affected student learning negatively.

- i. Pronunciation:** Student writing sample indicated that most of them struggled with differentiating between letters' sounds. For instance, some were not able to distinguish the sound of soft "ah" as in Yeah from "aa" as in Australia.
- ii. Negative transfer:** Transfer of rules from English (students' first language) was evident, especially in the use of preposition such as "to." For instance, students mostly translated "to school" exactly in the same order as in English, while in Pashto the order switches "school to." Another student translated "tailor sews clothes" exactly in the same order as in English, while in Dari and Pashto the formula for a simple sentence is "subject + object/complement + verb." The above sentence should have looked like: "tailor clothes sews."

**iii. Instructor's feedback:** Review of student homework revealed that instructors did not provide thorough and constructive feedback to the students. In some parts instructors completely ignored questions without providing any feedback, which may have led to misunderstanding on the part of students who might have thought their answers were correct. On another occasion, an instructor had completely crossed out (marked incorrect) a student's translation without offering any comments or feedback.

### c. Classroom Observation

**i. Class Management:** The classes were taught with one lead and a teacher aide. During the 30 minute four formal classroom observations, the researcher found that the lead teacher usually started the lesson with a lesson plan and the teacher aide waited for opportunities to interject. While waiting, the teacher aide would quite often get involved in side conversations with students. Sometimes, the teachers would walk out of the classrooms to attend their phone calls. Lack of prior coordination between lead teachers and teacher aides was evident. In one instance, a lead teacher asked the teacher aide if he agreed with him to give a certain activity to students. In a couple of sessions the teacher aides joined the second half of the class and did not know what was taught earlier. When the teachers started small group activities, there was a lack of their awareness about the first half of the session which led to confusion on the part of students because the teacher aides did not use the same vocabulary that students had learned in the first half of the session. In one of the classes, a teacher aide walked into the class and provided clarification to students while the lead teacher was lecturing. The observation rubric can be found in Appendix 4.

#### ii. Instruction

1) **Handouts** had many typos; for instance, the transliteration was not consistent.

2) **Error correction:** Teachers corrected students' errors immediately/or moved to another student to get the right answer. In one instance, a student completely stopped producing as soon as he was asked to correct his sentence structure (subject + object + verb) when he was trying to provide an answer to a question posed by the instructor:

Teacher: How do you say, "I go to class?"

Student: "May-rawam (go) Man (I)..."

Teacher: "No, no, you have to use the verb (May-rawam -go) at the end of the sentence"

Student: Um, Um [gave up]

Ellis (1990) explains that learners build fluency if errors are to be tolerated.

3) **Use of English:** The instructors were falling back to English instead of sticking to the target language e.g. Dari/Pashto to explain the grammatical concepts to the students.

4) **Interaction:** Instructors were mostly interacting with a few students who were active participants.

5) **Vocabulary:** was introduced and practiced with students in the form of traditional drills such as "repeat after me." For instance, on one occasion a teacher was pronouncing the infinitive form of the verb and asking students to repeat after him:

Teacher: "Nashas-tan" (to sit)

Students: "Nashas-tan"

Teacher: "Nashas-tan"

Students: "Nashas-tan"

6) **Lack of comprehensible input:** to introduce vocabulary was evident. For example, no visual aid, realia, modeling, or Total Physical Response (TPR) was used to introduce and practice vocabulary.

7) **Scaffolding:** was also not evident as the lesson was mostly presented in the form of lecture where the teacher spoke the majority of the time. Students received fewer opportunities to participate in their learning.

### **VIII. Develop and Implement and Monitor Action Plan**

Once the data has been collected, the teachers created a Professional Learning Community (PLC) team to analyze and reflect on the data to make a data-driven decision and take action. During data collection, the teachers noticed situations that needed attention and intervention. The researcher agrees with Frey and Fisher (2006) who argue that understanding the root causes of problems makes it easy for the educators to develop goals and objectives to eliminate them. After each observation, teachers were asked to meet as a PLC and discuss and address an observed problem. They were basically asked to discuss what they want their students to know and how to collect evidence of student learning (Frey & Fisher, 2006).

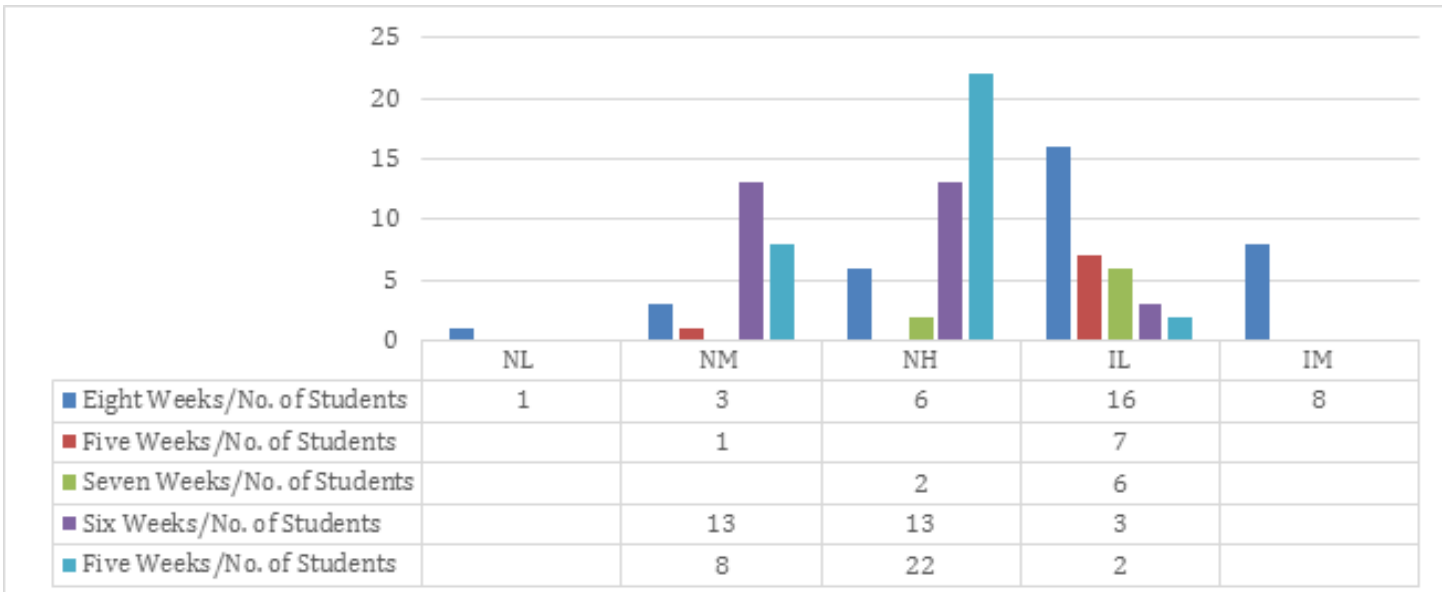
Based on the findings of this study the teachers identified non-availability of a standardized curriculum as the main factor that affected student learning outcomes negatively. Teachers decided to talk about a curriculum that pertains to the specific needs of their students. They decided to develop a curriculum that supports each individual session and prepares students for the next session(s). They also decided to involve students in the process by offering them a list of topics to choose from as major themes and sub-themes for the duration of the course. The teachers also decided to share the draft of curricula with their students for feedback. In addition, the teachers decided to match materials' appropriateness to students' level based on ACTFL proficiency guidelines (ACTFL, 2012) and include the following components in it:

- Conversation and practice
- Teaching grammar explicitly to students through homework assignments in order to save instructional time.
- Increasing use of Dari and Pashto language in class.
- Providing opportunities for students to get exposed to the language outside of the classroom.

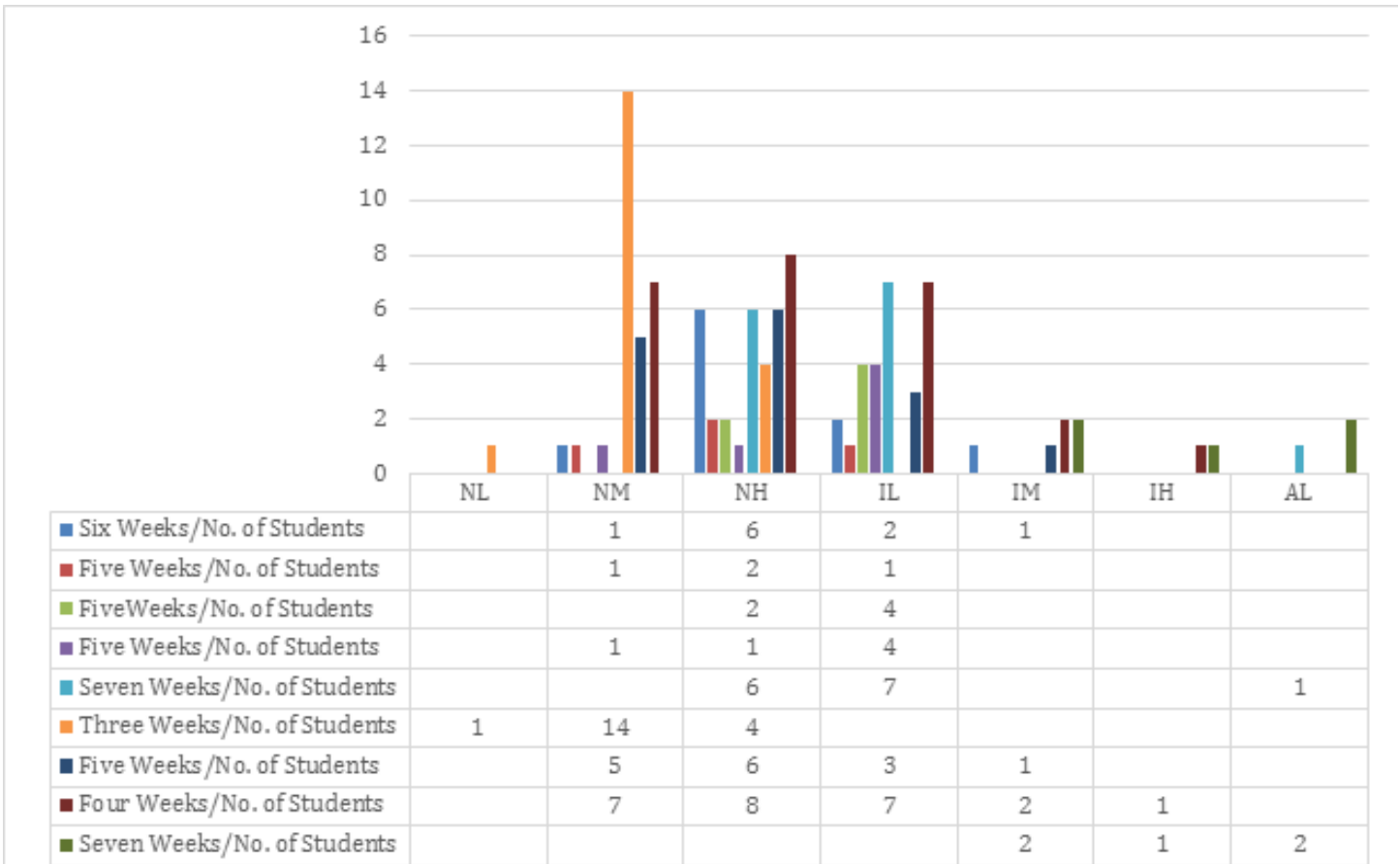
#### *The Current Curriculum*

The current curriculum has been designed and developed based on the principals of Backward Design (Eddy, 2012), California five-step lesson plan (CWLP, 2013), and ACTFL oral proficiency guidelines (ACTFL, 2012). The Lesson Plan Generator (LPG), a software developed by LRC, has been used to develop these lessons (LARC, 2014). Intensive elementary Pashto (eight weeks) now has 280 lesson plans and activities and Intensive elementary Dari (six weeks) has 218 lesson plans and activities. The lessons are thematic and begin with introduction of vocabulary in the form of a conversation, which enhances students ability to acquire the language in a short time and be able to create with the language. As the data in graph 1 and 2 demonstrate, the students' OPI scores improved tremendously after the implementation of the new curricula. The scores have been assigned by external testers of ACTFL.

Graph 1. Elementary Pashto course - OPI Scores officially assigned by ACTFL  
January – September 2013



Graph 2. Elementary Dari course - OPI scores officially assigned by ACTFL  
January 2013 – May 2014



N (Novice), I (Intermediate), A (Advanced), L (Low), M (Mid), H (High)



For monitoring and assessing the effects of the action plan, the team of instructors decided to establish a system of promoting the culture of PLC and reflection on instruction among teachers. This would entail audio-taping and critiquing instruction, continued classroom observations and providing feedback for improvement from more experienced peers. The teachers also decided to monitor student progress reports on a weekly bases using a rubric (Appendix 5).

## Conclusion

Action research projects serve the needs of teachers because teachers themselves get involved in systematically collecting information on issues and then taking action using their data. In this research action project, the teachers identified areas that hold students back from developing competency and fluency in speaking Dari and Pashto in intensive classes. The teachers decided to develop more coherent curriculum and assess and monitor its implementation by videotaping classes, observing each other's instruction and providing feedback, and reflecting on their instruction and student learning outcomes regularly using the PLC format.

## References

- ACTFL. (2012). *ACTFL oral proficiency interview familiarization manual*. Retrieved from <http://www.language-testing.com/wp-content/uploads/2012/07/OPI.FamiliarizationManual.pdf>
- AWL. (2014). *Language learning difficulty*. Retrieved from <http://aboutworldlanguages.com/language-difficulty>
- Altakhaineh, A. (2012). *The Main Theories in Second Language Acquisition (SLA)*. Retrieved from <http://www.eltworld.net/times/2011/11/the-main-theories-in-second-language-acquisition-sla/>
- Brown, R. (1973). *A first language*. Cambridge, MA: Harvard University Press.
- Cornbleth, C. (1990) *Curriculum in Context*, Basingstoke: Falmer Press.
- Chomsky, N. (1967). *A Review of B. F. Skinner's Verbal Behavior*. Retrieved from [http://www2.uni-wuppertal.de/FB4/anglistik/multhaup/methods\\_elt/pop\\_ups/chomsky\\_review\\_of\\_skinner.htm](http://www2.uni-wuppertal.de/FB4/anglistik/multhaup/methods_elt/pop_ups/chomsky_review_of_skinner.htm)
- Chomsky, N. (1965). *Aspects of the theory of syntax*. The MIT Press: Cambridge.
- CWLP. (2013). *5-Step lesson plan template*. Retrieved from <http://reffor.us/>
- Downey, C. J., Steffy, B., English, F., Frase, L. & Poston, W. K. Jr., (2004). *The three-minute Classroom walk-through: Changing school supervisory practice one teacher at a time*. Thousand Oaks, CA: Corwin Press
- Dulay, H. & Burt, M. (1978). *Some guidelines for the assessment of oral language proficiency and dominance*. TESOL Quarterly 12: 177-192.
- Eddy, J. (2012). *Backward Design: Curriculum planning for teacher programs*. Retrieved from: <https://startalk.umd.edu/2012/meetings/NFPE/pre/presentations/BackwardDesignCurriculumPlanningForTeacherPrograms.pdf>.
- Ellis, R. (1990). *Instructed Second Language Acquisition*. Malden, MA: MIT Press.
- Farrell, T. (2012). *Reflecting on reflective practice: (Re) visiting Dewey and Schon*. TESOL International Journal, 3 (1)
- Fisher, D., & Frey, N. (2008). *A Structure for Instruction That Works*. Retrieved from <http://www.azed.gov/leadingchange/files/2013/06/fisherkeynotepp.pdf>
- Freeman, D. (1989). *Teacher training, development, and decision making: A model of teaching and related strategies for language teacher education*. TESOL Quarterly, 23(1)
- Krashen, S. (1982). *Principles and Practice in Second Language Acquisition*. Oxford: Pergamon.
- Krashen, S. (1987). *Principles and Practice in Second Language Acquisition*. New York, N.Y.: Prentice-Hall International.
- Krashen, S. (2012). *Second Language "Standards for Success"*. Retrieved from <http://www.sdkrashen.com/articles/standards/all.html>
- LARC (2014). *The lesson plan generator*. Retrieved from <https://lpg.sdsu.edu>

Lightbown, P., & Spada, N. (2000). *How Languages are Learned* (3rd ed.). Oxford, N.Y.: Oxford University Press

Long, M. (1983). *Native speaker/non-native speaker conversation and the negotiation of comprehensible input*. *Applied Linguistics*, 4, 126-141.

Menees, V. (2012). *Second Language Acquisition: From Main Theories to Complexity*. Retrieved from <http://http://www.veramenezes.com/slatheory.pdf>

Overview of Second Language Acquisition Theory. (2003). Retrieved from <http://www.nwrel.org/request/2003may/overview.html>

Shrum, J., & Glisan, E. (2010). *Teacher's Handbook* (4th ed.). Boston, MA: Heinle.

Skinner, B. (1992). *Verbal behavior*. Cambridge: Prentice-Hall, Inc.

Stevick, E. (1976). *Memory, Meaning, and Method*. Rowley, Ma.: Newbury House

Swain, M., Kinnear, P., & Steinman, L. (2011). *Sociocultural Theory in Second Language Education*. Tonawanda, N.Y.: mm textbook.

Swain, M. (1985). *Communicative competence: Some roles of comprehensible input and comprehensible output in its development*. In S. Gass & C. Madden (Eds.), *Input in second language acquisition*. Rowley, MA: Newbury House.

Swain, M. (1995). *Three functions of output in second language learning*. In G. Cook, & B. Seidlhofer (Eds.), *Principles and practice in the study of language: Studies in honor of H. G. Widdowson*. Oxford: Oxford University Press.

Vygotsky, L. (1986). *Thought and language*. Cambridge: The MIT Press.

## Appendices

### Appendix 1

#### *Pre-course Survey - Students*

Please select a four-digit number that can be used as an identification code for post survey : \_\_\_\_\_

1. How old are you? \_\_\_\_\_
2. What is your gender?        Male / Female
3. How do you define your ethnicity? \_\_\_\_\_
4. What is the highest level of your education? \_\_\_\_\_
5. What is your native language? \_\_\_\_\_
6. How many other languages do you speak?    0        1        2        3        4/ more
7. Do you have any experience living in a foreign country? Yes / No  
If yes, where and how long? \_\_\_\_\_
8. Do you have any experience living in Afghanistan? Yes / No  
If yes, how long? \_\_\_\_\_
9. How do you visualize Afghans and Afghanistan? /what are your perceptions? Please be as detailed as possible: \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

- 
- 
10. Do you have any prior exposure to Dari/Pashto? Yes / No:  
If yes, where and how long? \_\_\_\_\_
11. Do you have any outside classroom exposure to Dari/Pashto? Yes / No:  
If yes, how and for how long? \_\_\_\_\_
12. Please define best practices in learning a foreign language:  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_
13. Please define best practices in teaching a foreign language:  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_
14. How important is it to learn *grammar* explicitly?  
Very Important    Important    Less important    Not important
15. How do you see the role of *homework* in learning the contents of the lesson?  
Very Important    Important    Less important    Not important
16. How do you see the role of *feedback* in learning the contents of the lesson?  
Very Important    Important    Less important    Not important
17. Please provide any other comments that you may have:

## Appendix 2

### *Interview Questions - Students*

1. Tell me about your background as a language learner.
2. Tell me about your experience in this course
3. How do you define “best practices” in a second language class?

### Appendix 3

#### *Interview Questions - Teachers*

1. Tell me about your experience as a teacher of Dari/Pashto in this program
2. What challenges have you faced? And how did you deal with them?

### Appendix 4

#### *Class Observation Rubric*

	Strongly agree	Agree	Disagree	No opportunity to assess/ Not observed
(1) The observed class matches the written descriptions of the curriculum				
(2) The teacher had lesson plans that clearly communicate the learning goals and activities for each day.				
(3) Students clearly understand the goals of the lesson.				
(4) The environment is conducive to language learning.				
(5) The teacher to student ratio is appropriate to ensure the success of the program				
(6) The target language is used at least 90% of the time for communication and instruction. English is used only when necessary.				
(7) The input is comprehensible. Students demonstrate comprehension of the target language.				
(8) The teacher checks frequently for student understanding.				
(9) Instruction is designed to facilitate learner-centered learning.				
(10) Instruction allows for meaningful interaction in the target language.				
(11) Instruction integrates language and culture.				
(12) Students use technology when available to meet the goals.				
(13) Authentic resources are used effectively to support the class.				
(14) Learning experiences address the interpretive mode of communication.				
(15) Learning experiences address the interpersonal mode of communication.				
(16) Learning experiences address the presentational mode of communication.				
(17) Instructional experiences build toward opportunities in each lesson for meaningful, unrehearsed communication. A balance exists between meaningful guided and independent practice.				
(19) Instructional time is used effectively in each lesson.				

**Appendix 5**

*Weekly Progress Report*

Teacher's Name: \_\_\_\_\_

Date: \_\_\_\_\_

Week # \_\_\_\_\_

	(Student Name)	(Student Name)	(Student Name)	(Student Name)	(Student Name)	(Student Name)
<i>Homework</i>	M T W T F □ □ □ □ □	M T W T F □ □ □ □ □	M T W T F □ □ □ □ □	M T W T F □ □ □ □ □	M T W T F □ □ □ □ □	M T W T F □ □ □ □ □
<i>Attentiveness</i>	M T W T F □ □ □ □ □	M T W T F □ □ □ □ □	M T W T F □ □ □ □ □	M T W T F □ □ □ □ □	M T W T F □ □ □ □ □	M T W T F □ □ □ □ □
<i>Class Participation</i>	M T W T F □ □ □ □ □	M T W T F □ □ □ □ □	M T W T F □ □ □ □ □	M T W T F □ □ □ □ □	M T W T F □ □ □ □ □	M T W T F □ □ □ □ □
<i>Classroom behavior</i>	M T W T F □ □ □ □ □	M T W T F □ □ □ □ □	M T W T F □ □ □ □ □	M T W T F □ □ □ □ □	M T W T F □ □ □ □ □	M T W T F □ □ □ □ □
<i>Homework</i>						
<i>Participation</i>						
<i>Classroom behavior</i>						

**General Remarks: Student will be evaluated based on the following scales:**

Scales	Attentiveness	Class Participation	Classroom behavior
1 – 3	Does not listen	Does not participates	Poor behavior (busy with other stuff during the class times or disrespecting the teachers)
4 – 6	Sometimes listens	Sometimes participates	Good behavior
7 - 10	Always listens	Always participates	Always well behaved

# Trends in US-based Central Asian Language Instruction: The Case of Indiana University's Summer Language Workshop

Ariann Stern-Gottschalk  
*Indiana University*

## Abstract

This paper focuses on the administrative aspects of Central Asian language intensive summer language training, drawing chiefly on data, examples, syllabi, and other materials from the Indiana University Summer Language Workshop, which, for over twenty years, has offered intensive courses in a broad range of languages of Central Asia. The national picture of US-based summer programs for Central Asian language teaching is reflected in this case study of one of the most historically robust of these programs. Discussion includes factors affecting language offerings each year, an overview of funding mechanisms that support this programming and impact the state and availability of Central Asian language instruction, enrollment and hiring trends, student demographics, instructional support mechanisms, and language materials.

## Introduction

Intensive instruction in the Central Asian languages has been a feature of the Indiana University (IU) Summer Language Workshop (the Workshop) curriculum for over 20 years. This paper focuses on the administrative aspects of Central Asian language training, drawing on data, examples, and other materials from the Workshop. Specifically, this paper provides an overview of the program with a focus on languages taught and the decisions that drive course offerings, enrollment trends, students, funding, faculty, instructional support, and language materials. In this paper, "Central Asian languages" include Iranian (Dari, Pashto, Persian) and Caucasian (Azerbaijani and Georgian) languages.

## Language Offerings

### *Historical overview:*

Georgian was the first Caucasian language offered in the Workshop, followed by Uzbek, which was first offered beginning in 1991. Georgian was offered continuously until 2013 and Uzbek was offered uninterruptedly until 2014. In 1993, Azerbaijani and Kazakh were added. Kazakh was taught continuously from 1993 until 2013, while Azerbaijani was offered through 2012 (and is likely to be offered again beginning in 2016). Kyrgyz was offered once to two students in 1994. This is the same year Turkmen, which was offered continuously through 2009, was added (Turkmen will be offered again in 2015). In 1997 Chechen was offered to three students. Tibetan was offered in 2001 and 2002 to a total of eight students. In 2003, Pashto, Tajik and Uyghur were added. Tajik was offered continuously through 2011, Pashto through 2012 and Uyghur through 2013. That last summer, it was only offered at the intermediate level to two students. Mongolian was added in 2007, Dari in 2010, Tatar and Turkish in 2011, Persian in 2012 and Urdu (in combination with Hindi) was offered for the first time in 2014 to six students. Dari was last offered, at two levels, in 2013. Mongolian, Persian, Tatar, Turkish, and Urdu continued to be offered.

Factors determining language offerings each year include national student demand; the research needs of IU faculty and graduate students; academic year enrollments in the IU Department of Central Asian Studies



(CEUS); prioritization by CEUS and IU's Inner Asian and Uralic National Resource Center (IAUNRC); funding from IAUNRC and other national centers with specific interest in supporting particular languages or groups of languages; and textbook development by other units at IU, most notably by the Center for the Languages of the Central Asian Region (CeLCAR).

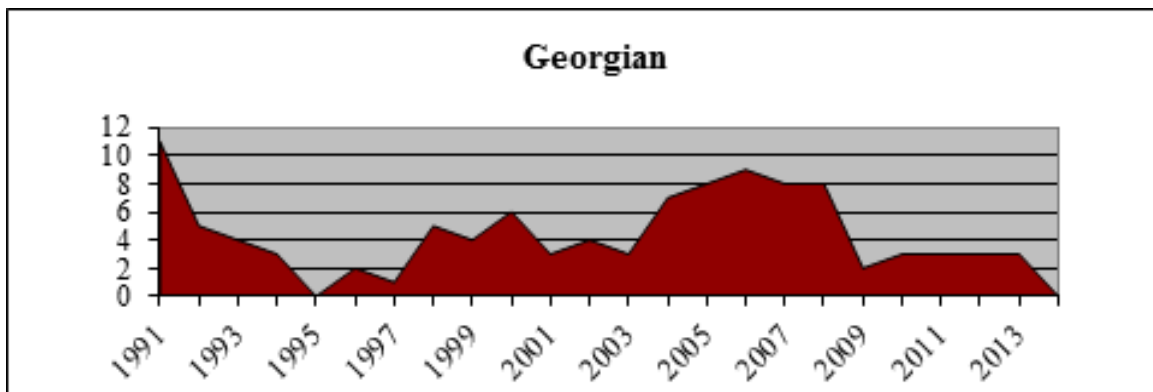
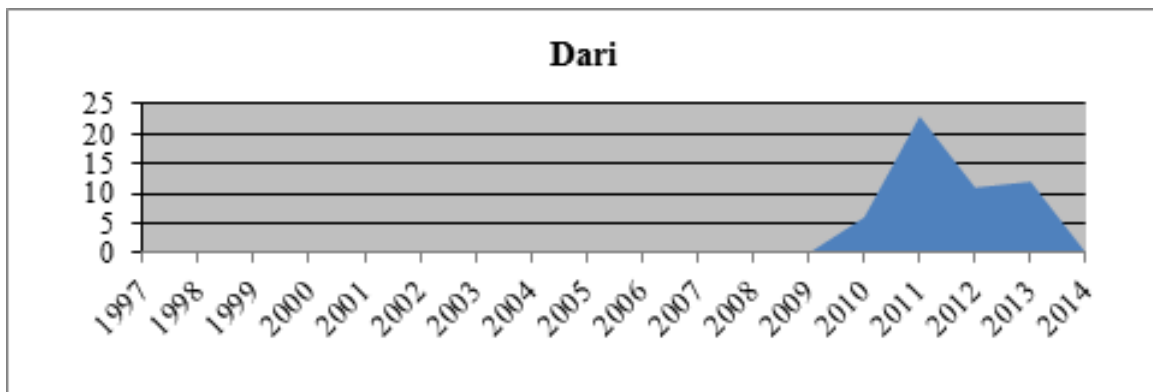
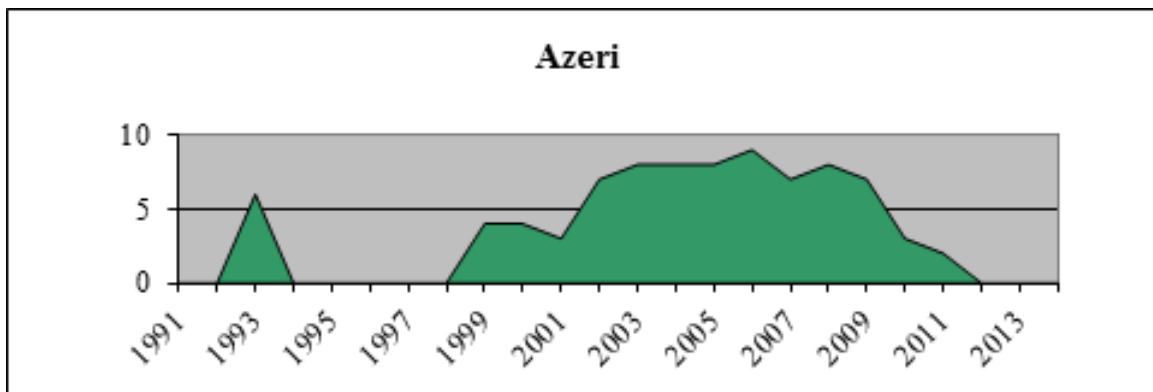
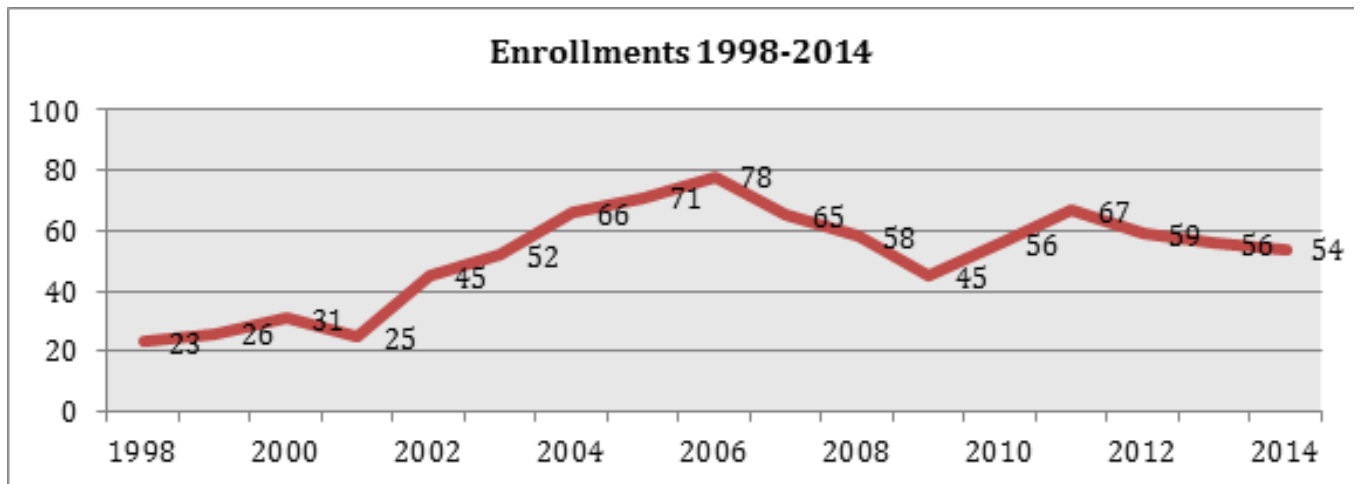
A guiding principle for the Workshop in deciding which courses to offer is promotion of and support for academic year language study, and the academic priorities for CEUS and its faculty as manifest in their research and teaching interests. Over the past 23 years, these academic-year priorities have increasingly been the primary motivator for offering various languages through the Workshop. Indeed, this academic mission impacts student demand (i.e. enrollments) as can be seen in the cases of Uyghur and Kazakh. As faculty focus has shifted away from the regions where these languages are spoken (because of political factors affecting access to the region for Uyghur and retirements for Kazakh), fewer students are able to pursue research in these areas at IU and student demand has decreased significantly.

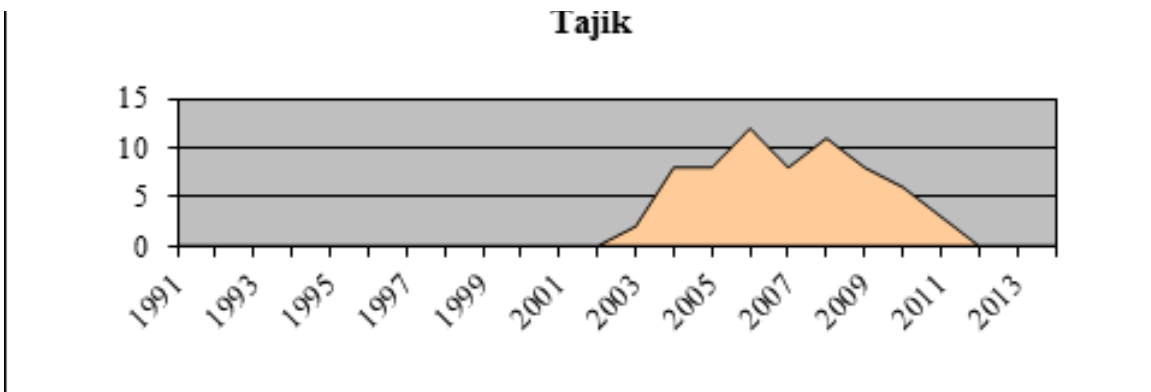
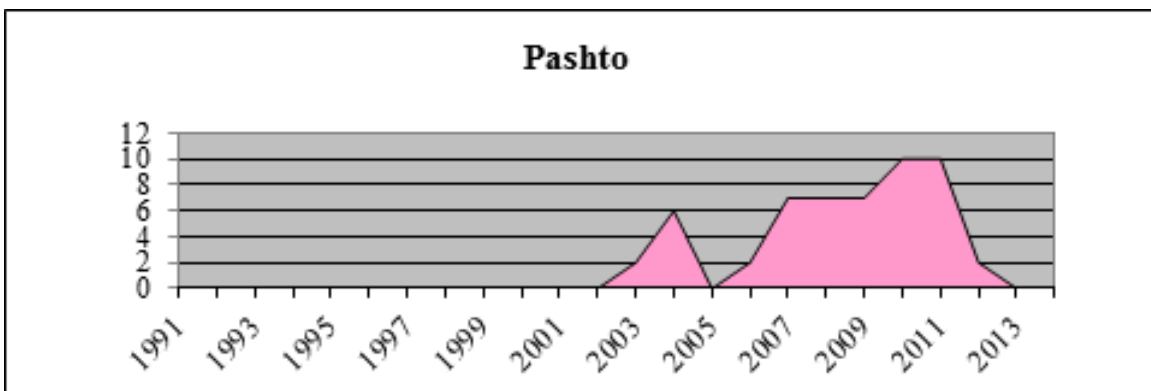
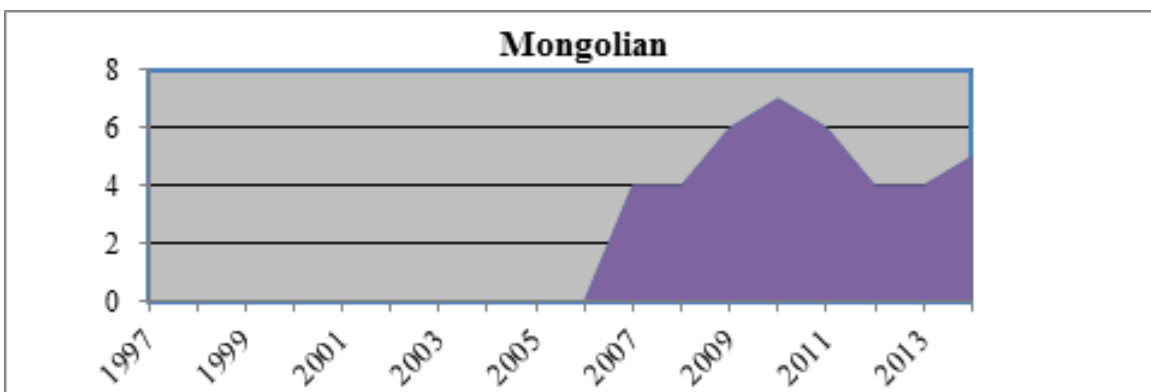
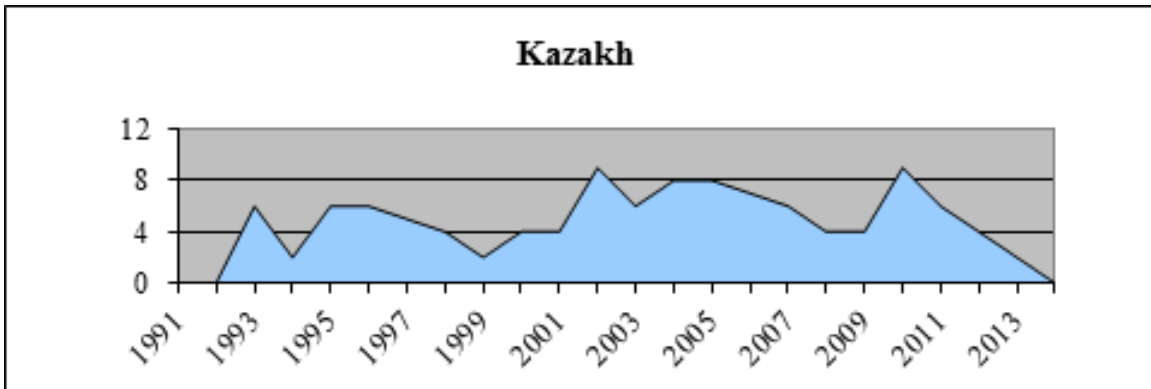
Another major academic mission at IU supported by the Workshop is textbook development in CeLCAR. The Workshop offers an ideal setting for textbook developers who seek to assess how well drafts of the materials work in the classroom. Students in the Workshop represent a national audience of learners at the graduate and undergraduate level from institutions across the US as well as professors and other professionals. Most students of Central Asian languages are experienced language learners, so are discerning about the methodologies and approaches employed in class and in their textbooks. They use the CeLCAR materials intensively for two months and complete anonymous surveys that provide CeLCAR developers with feedback at the midterm and final points of the program. Some of the developers also teach from their textbooks in the summer and can thus gauge the efficacy of their texts daily in practice. Thus every year, the Workshop leadership consults with CeLCAR about course offerings that would best help them in their work. Finally, Turkish has been added and expanded to two levels to support the work of the Turkish language Flagship, which works together with the Workshop director on curriculum and hiring for the Turkish courses.

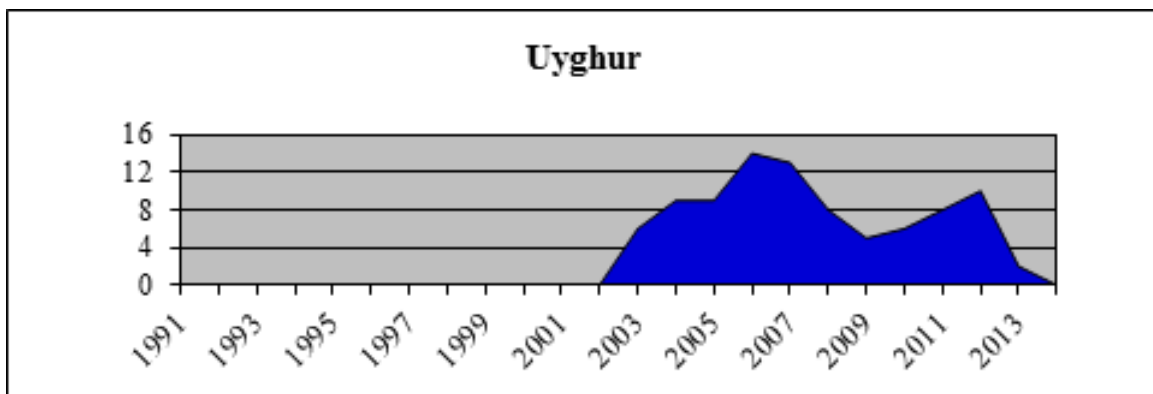
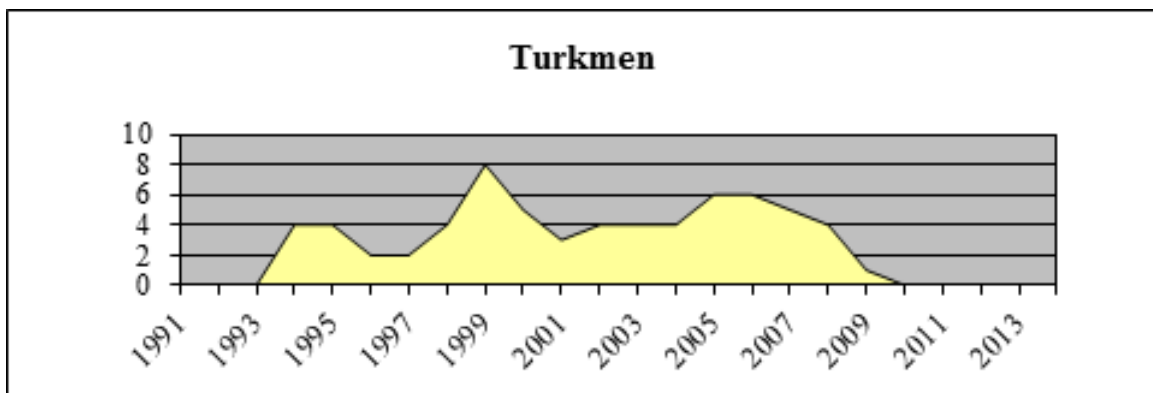
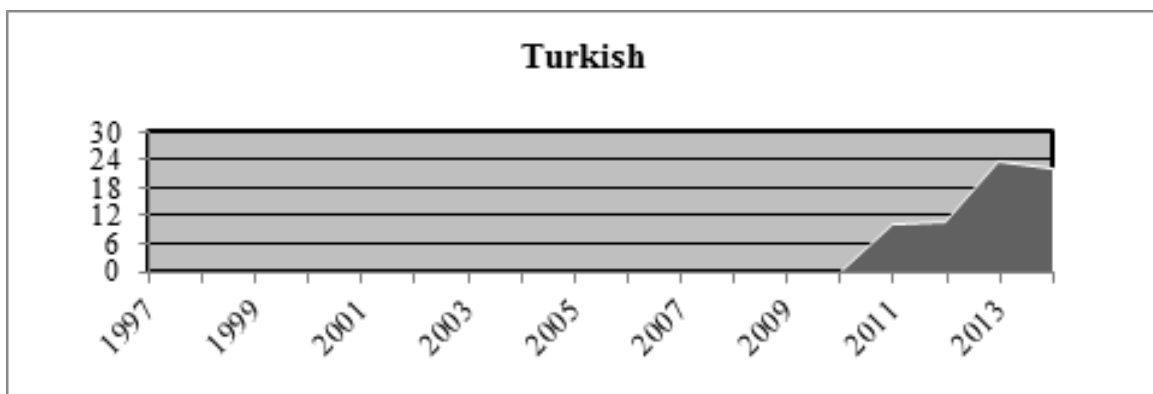
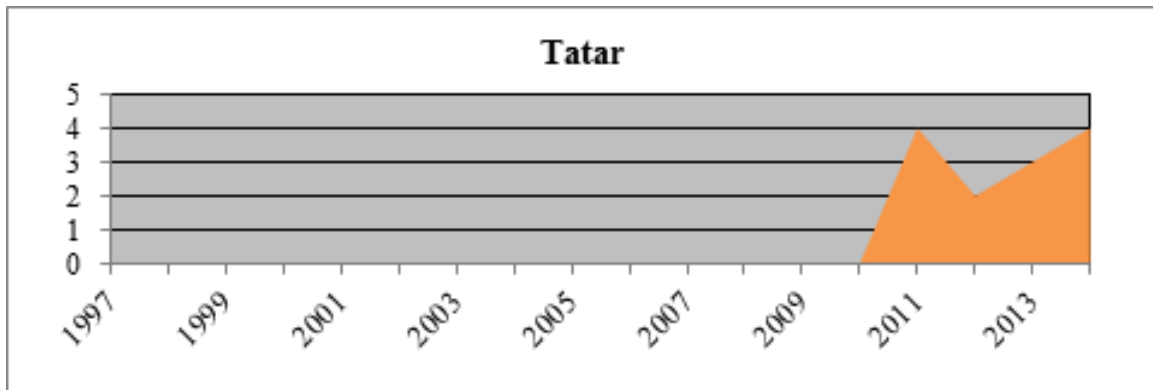
The principle of promoting academic year language study means aligning the curricula and syllabi in the Workshop with those in the academic year so that students can seamlessly enroll in the next level of the language after a summer of intensive study in the Workshop. In many cases, this dictates summer instructional hiring decisions, because it is ideal for CEUS faculty members teach the courses both in the summer and academic year. This principle also influences decisions about summer course offerings. The Workshop will not offer courses if they threaten enrollments in the academic year, which has increasingly led to decisions not to offer more than one or two levels of instruction in each language in the summer or not to offer courses at all.

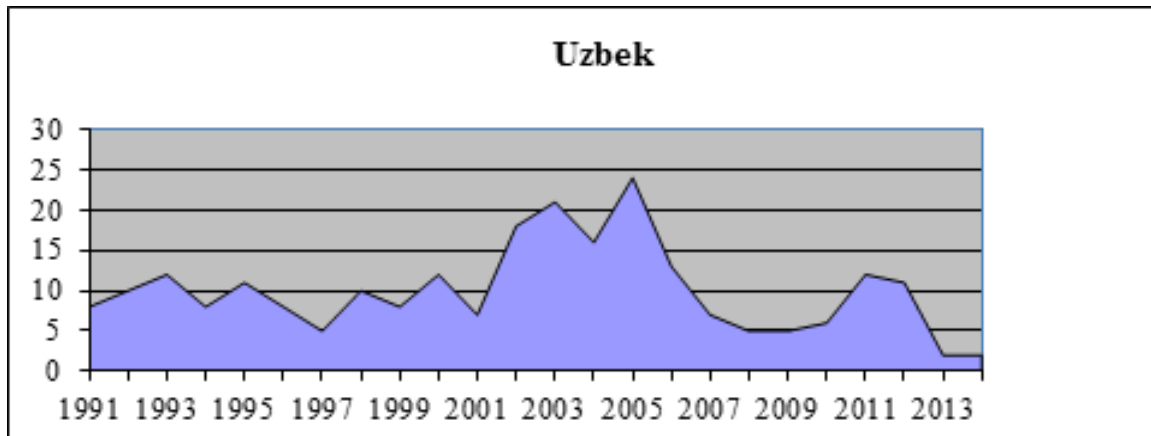
### **Enrollments: trends and factors affecting enrollments**

The following charts track the enrollments trends for all languages except Chechen, Urdu, and Tibetan (which have only been offered once or twice respectively and whose enrollments are detailed above). Factors that most impact enrollments include the research priorities of students and faculty as described above, funding, access to study abroad in regions where the languages are spoken, and world events. For example, when they were high priority for the US military, Afghan languages attracted many students, but there is much less demand for Dari and Pashto than there was even five years ago.









## Students

In the Workshop, over 75% of those pursuing Central Asian languages are graduate students. The majority of undergraduate students are in the Reserve Officer Training Corps. As noted above, students also include professionals working for the government and professors (since 2009 government agencies represented include the National Security Agency, and Departments of Defense, Energy, and State; faculty have come from IU, Northwestern, the University of Chicago, and other schools). Most students have funding for study abroad following their Workshop study and many spend time in country at some point. They study history, politics, religion, biology, anthropology, resource management, architecture, ethnomusicology, linguistics, etc. In 2011, a group of female students started the national Women's Central Eurasian Network in the Workshop, because it was a place where many of them met and could interact for the first time on a large scale. That same year, the Workshop initiated two years of graduate student roundtables for students to talk about their research, doing research in Central Asia, having opportunity to present on their research. Since that time, students present on their research or time overseas in other extracurricular programming.

## Funding

There are two main types of funding for summer intensive language programs: funding for instructional support and funding for student fellowships. Student funding includes Department of Education Title VI Foreign Language and Areas Studies Fellowships (FLAS); Title VIII funding that comes through flow-through agencies like the Social Science Research Council (SSRC) and directly from grants from the State Department to the Workshop; scholarships for ROTC undergraduate students through the national Project GO program sponsored by the Department of Defense's Defense Language and National Security Education Program (DLNSEO); and for military linguists and language specialists through the DLNSEO-sponsored Language Training Center grant the Workshop received for 2013. Funding for faculty hires has come from the Title VIII flow-through agencies and Title VI National Resource Centers (NRC) at IU and other institutions. IAUNRC annually works with a consortium of other NRCs that regularly contribute funds to support instruction for select groups of languages depending on the focus of each NRC. In the past decade, especially most recently, there have been serious fluctuations in these funding sources because of changes in agency priorities and also because of partners' priorities. For example, SSRC stopped receiving funding from Title VIII to support summer language programs in 2008 and in the 2013 fiscal year, the State Department did not appropriate any funding for the Title VIII program. Similarly, serious cuts to the Title VI National Resource Centers in 2011 and again in the 2014 Title VI competition curtailed funding for instructional salaries, which limits the Workshop's ability to offer courses with very low enrollments. In 2014 these affects were felt most seriously as fewer fellowships

were available to support graduate students and professionals and less funding was available to support languages with very low enrollments.

## **Faculty**

The first instructors in Central Asian Languages in the Workshop came from overseas and other institutions in the US. All were native speakers. The majority of these summer hires returned annually to teach and instruction is fairly consistent each year. There have been some minor fluctuations in Azerbaijani and Kazakh instructors. The majority of changes have been seen when a second instructor is needed for the rare summer when a second or third level of Dari, Mongolian, Pashto, or Uyghur is taught. The Georgian instructor was a faculty member from the IU Department of Slavic Languages and Literatures. Beginning in 1999, the senior lecturer in Persian from CEUS taught Azerbaijani for ten years with just one break in the summer of 2008. In 2006, a materials developer from CeLCAR, a graduate student in the IU Department of Linguistics who also worked as a CeLCAR materials developer, and two new CEUS faculty members began teaching regularly in the summers in Pashto, Uzbek (both levels), and Uyghur respectively. A new CEUS faculty hire in Mongolian language joined the Workshop faculty in 2007. Two more CeLCAR materials developers began teaching Dari and second-year Uyghur in 2010. In 2012, a doctoral student from the IU Department of Anthropology began teaching Persian and a Masters student from CEUS began teaching Turkish. These were only the second and third students hired to teach Central Asian languages. The CEUS student is the first and only non-native speaker hired to teach a Central Asian language in the Workshop. With the addition of intermediate Turkish in 2013, an academic year hire from CEUS and the Turkish Language Flagship joined the faculty.

All of those with academic year appointments at IU have taught each summer their languages were offered. There are two exceptions to this: in 2013 one of the CeLCAR developers finished her degree and could not teach because of restrictions on her visa. And in 2014, IU's senior lecturer in Uyghur taught Uzbek (she was raised and educated in Uzbekistan).

For many years IAUNRC, CeLCAR, and CEUS controlled the Central Asian hiring decisions, faculty oversight, and instructional budgets for the Workshop. When the Workshop underwent fundamental restructuring by the College of Arts and Sciences in 2011, instructional oversight, hiring, and budgeting processes for the Central Asian languages reverted to the Workshop director. Though this work is now centralized, the director continues to work closely with the directors of the IAUNRC, CeLCAR, and CEUS on all of these matters.

## **Instructional Support**

Central Asian language faculty have always participated in instructor orientations at the start of each summer, pedagogy workshop conducted by outside specialists, and have benefitted from anonymous midterm and final student course evaluations. Until the centralization of all faculty oversight in 2011, the CEUS Language Coordinator was responsible for otherwise supporting the curricula and faculty in the Central Asian languages. A graduate student in Linguistics and Second Language Studies was hired in this role in 2004. With excellent attention to detail and a fine grasp of best practices in second language acquisition, she began a series of regular faculty orientations that focused on aligning the various methodologies employed in the Central Asian language classrooms with a focus on communicative competencies, use of authentic materials and multimedia, and discussions about expectations in American classrooms. This last point was critical for many of the faculty who were trained and otherwise taught in Central Asia using primarily Soviet methods. To reinforce work done in the orientations, she also regularly observed classes in the summers, using a rubric for providing feedback and underscoring the most critical ideas and approaches introduced earlier in the summer. Simultaneous with the 2011 change in Central Asian language administration (including instructional oversight) in the Workshop, a new CEUS language coordinator was hired. However, the effects of the earlier language coordinator's efforts



are still very much in evidence in the high quality, student-centered, engaging CEUS language courses taught since 2011.

Indeed, the current Central Asian language instructors are some of the most seasoned teachers in the Workshop each summer. Their work is based in language acquisition research and national standards and this is reflected in excellent student evaluations, the director's observation reports, and feedback from outside specialists who visit Workshop classes for various purposes. Since 2012, CEUS faculty members have been leaders at meetings and presentations for all Workshop instructors and have generously shared their expertise with other Workshop colleagues.

## Language Materials

Of the 15 Central Asian languages that have been offered in the Workshop, only five have had published materials easily accessed by students in the US from the first summer when those languages have been taught: Georgian, Persian, Turkish, Urdu, and Uzbek. The first published textbooks used in the Workshop were Howard Aronson's *Georgian: A Reading Grammar* published by Slavica Publishers and Khayrulla Ismatulla's *Modern Literary Uzbek* from the IU Uralic and Altaic Series published by Routledge. These were the only materials published in the US used in the Workshop's Central Asian language classes until 2009. In other languages, course materials included books published in Central Asia that instructors brought for students; instructors' own materials produced for the courses; CeLCAR materials under development for Dari, Kazakh, Pashto, Tajik, Uzbek, and Uyghur; and a textbook published in Mongolia students purchase from the Mongolian society. In 2009, the Georgian instructor published her own textbook with Hippocrene, which replaced the Aronson volume that year. Georgetown University Press published CeLCAR's Tajik textbook in 2009 and its Uzbek book in 2010; these published versions have been used in the beginning Tajik and Uzbek courses since then.

## In lieu of references

The following tables detail Central Eurasian language offerings and enrollments in the Workshop from 1998-2014.

1998

Language	Levels offered	Combined Enrollments
Georgian	1	5
Kazakh	1	4
Turkmen	1	4
Uzbek	1 and 2	10
<b>Total</b>		<b>23</b>

1999

<b>Language</b>	<b>Levels offered</b>	<b>Combined Enrollments</b>
Azerbaijani	1	4
Georgian	1	4
Kazakh	1	2
Turkmen	1	8
Uzbek	1 and 2	8
<b>Total</b>		<b>26</b>

2000

<b>Language</b>	<b>Levels offered</b>	<b>Combined Enrollments</b>
Azerbaijani	1	4
Georgian	1	6
Kazakh	1	4
Tibetan	1	3
Turkmen	1	5
Uzbek	1 and 2	12
<b>Total</b>		<b>31</b>

2001

<b>Language</b>	<b>Levels offered</b>	<b>Combined Enrollments</b>
Azerbaijani	1	3
Georgian	1	3
Kazakh	1	4
Tibetan	1	5
Turkmen	1	3
Uzbek	1 and 2	7
<b>Total</b>		<b>25</b>

2002

<b>Language</b>	<b>Levels offered</b>	<b>Combined Enrollments</b>
Azerbaijani	1 and 2	7
Georgian	1	4
Kazakh	1	9
Tibetan	1	3
Turkmen	1 and 2	4
Uzbek	1 and 2	18
<b>Total</b>		<b>45</b>

2003

<b>Language</b>	<b>Levels offered</b>	<b>Combined Enrollments</b>
Azerbaijani	1 and 2	<b>8</b>
Georgian	1	<b>3</b>
Kazakh	1 and 2	<b>6</b>
Pashto	1	<b>2</b>
Tajik	1	<b>2</b>
Turkmen	1 and 2	4
Uyghur	1 and 2	6
Uzbek	1 and 2	21
<b>Total</b>		<b>52</b>

2004

<b>Language</b>	<b>Levels offered</b>	<b>Combined Enrollments</b>
Azerbaijani	1 and 2	<b>8</b>
Georgian	1	<b>7</b>
Kazakh	1 and 2	<b>8</b>
Pashto	1	<b>6</b>
Tajik	1	<b>8</b>
Turkmen	1 and 2	4
Uyghur	1	9
Uzbek	1 and 2	16
<b>Total</b>		<b>66</b>

2005

<b>Language</b>	<b>Levels offered</b>	<b>Combined Enrollments</b>
Azerbaijani	1 and 2	<b>8</b>
Georgian	1	<b>8</b>
Kazakh	1 and 2	<b>8</b>
Tajik	1	<b>8</b>
Turmen	1 and 2	6
Uyghur	1 and 2	9
Uzbek	1 and 2	<b>24</b>
<b>Total</b>		<b>71</b>

2006

<b>Language</b>	<b>Levels offered</b>	<b>Combined Enrollments</b>
Azerbaijani	1 and 2	9
Georgian	1	9
Kazakh	1 and 2	7
Pashto	1 and 2	2
Tajik	1 and 2	12
Turkmen	1 and 2	6
Uyghur	1 and 2	9
Uzbek	1 and 2	24
<b>Total</b>		<b>78</b>

2007

<b>Language</b>	<b>Levels offered</b>	<b>Combined Enrollments</b>
Azerbaijani	1 and 2	7
Georgian	1	8
Kazakh	1 and 2	6
Mongolian	1	4
Pashto	1	7
Tajik	1 and 2	8
Turkmen	1 and 2	5
Uyghur	1 and 2	13
Uzbek	1 and 2	7
<b>Total</b>		<b>65</b>

2008

<b>Language</b>	<b>Levels offered</b>	<b>Combined Enrollments</b>
Azerbaijani	1 and 2	8
Georgian	1	8
Kazakh	1 and 2	4
Mongolian	1	4
Pashto	1	6
Tajik	1 and 2	11
Turkmen	1 and 2	4
Uyghur	1 and 2	8
Uzbek	1 and 2	5
<b>Total</b>		<b>58</b>

2009

Language	Levels offered	Combined Enrollments
Azerbaijani	1 and 2	7
Georgian	1	2
Kazakh	1	4
Mongolian	1	6
Pashto	1	7
Tajik	1 and 2	8
Turkmen	1	1
Uyghur	1 and 2	5
Uzbek	1 and 2	5
<b>Total</b>		<b>45</b>

2010

Language	Levels offered	Combined Enrollments
Azerbaijani	1	3
Dari	1, 2, and 3	6
Georgian	1	3
Kazakh	1 and 2	9
Mongolian	1	7
Pashto	1 and 2	10
Tajik	1, 2, and 3	6
Uyghur	1, 2, and 3	6
Uzbek	1 and 2	6
<b>Total</b>		<b>56</b>

2011

Language	Levels offered	Combined Enrollments
Azerbaijani	1	2
Dari	1 and 2	3
Georgian	1	3
Kazakh	1	6
Mongolian	1 and 2	6
Pashto	1 and 2	10
Tajik	1	3
Tatar	1	4
Turkish	1	10
Uyghur	1	8
Uzbek	1 and 2	12
<b>Total</b>		<b>67</b>

2012

<b>Language</b>	<b>Levels offered</b>	<b>Combined Enrollments</b>
Dari	1 and 2	<b>11</b>
Georgian	1	<b>3</b>
Kazakh	1	<b>4</b>
Mongolian	1	<b>4</b>
Pashto	1	<b>2</b>
Persian	1	11
Tatar	1	2
Turkish	1	11
Uyghur	1 and 2	10
Uzbek	1 and 2	1
<b>Total</b>		<b>59</b>

2013

<b>Language</b>	<b>Levels offered</b>	<b>Combined Enrollments</b>
Dari	1 and 2	<b>12</b>
Georgian	1	<b>3</b>
Kazakh	1	<b>2</b>
Mongolian	1	<b>4</b>
Persian	1	<b>4</b>
Tatar	1	3
Turkish	1 and 2	24
Uyghur	2	2
Uzbek	1	2
<b>Total</b>		<b>56</b>

2014

<b>Language</b>	<b>Levels offered</b>	<b>Combined Enrollments</b>
Mongolian	1	<b>4</b>
Persian	1	<b>14</b>
Tatar	1	<b>4</b>
Turkish	1 and 2	<b>26</b>
Urdu	1	<b>6</b>
Uzbek	1	2
<b>Total</b>		<b>56</b>





