

UYGURCA ve TÜRKÇEDE GELECEK ZAMAN KIP ÖZELLİKLERİ

THE FUTURE TENSE PROPERTIES of UYGHUR and TURKISH

*Murat ORHUN**

Özet:

Makalede, Uygurcadan Türkçeye bilgisayarlı çeviri uygulaması geliştirmek üzere Uygurca ve Türkçede bulunan gelecek zaman kipleri detaylı örnekler ile analiz edilmiştir. Gerçi bu diller Ural-Altay diller kategorisinde aynı diller grubuna ait ve tümce yapısı yönünden benzer özelliklere sahip olsa da yine de göz ardı edilmeyecek kadar farklılık taşımaktadır. Gelecek zaman kipi, Uygurcada bulunan en karmaşık yapılardan biridir ve Türkçede bulunan benzeri yapılarla tüm ayrıntıları ile karşılaştırılmıştır. Uygurca gelecek zaman kipinin Türkçeye doğru aktarılması için bazı kurallar tanımlanmıştır.

Anahtar kelimeler: Uygurca, Uygur Biçim bilimi, Bilişimsel dil bilimi, Biçimbilimsel belirsizlik giderme, Bilgisayarlı çeviri.

Abstract:

In this article the future tense properties of Uyghur Language are analyzed with specific examples in order to implement a prospective machine translation system that translates Uyghur into Turkish. Even though, these languages belong to the same Ural-Altaic language family and have almost the same syntax structure, however there are some differences that cannot be ignored. Future tense is the most complex grammatical structure in Uyghur Language and it is compared to similar structures in Turkish with all of its aspects. Some rules have been defined in order to translated Uyghur future tense into Turkish correctly.

Keywords: Uyghur language, Uyghur morphology, computational linguistics, morphological disambiguation, machine translation.

Introduction

Machine translate is one of the hot topics of Natural Language Processing, which is a subfield of Artificial Intelligence. To translate a language into another, the most important thing is to keep the meaning of the original content of a source language and transfer it to the target language without losing any information. So far, some machine translation systems have been implemented to use for a special purpose. For example, Meteo that translates weather reports from English into French (Chandioux, 1976), and the File Title translation system translates from Japanese into English (Nagao,

* Dr. Öğr. Üyesi, İstanbul Bilgi Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi, Bilgisayar Mühendisliği Bölümü İstanbul /TÜRKİYE
murat.orhun@bilgi.edu.tr Orcid: 0000-0002-5013-1119.

1984). At the same time, there are some machine translation systems for general purpose such as Arabic-Chinese (Habash, Hu, 2009), French-Arabic (Alsharaf, Cardey, Greenfield, Shen, 2004), Russian-English (Huet, Manishina, Lefevre, 2003), and statistical machine translation systems (Philipp, Europarl, 2005). They are the most high-quality systems designed for most close languages (Hajic, 2000; Hajic, 1987). Though all Turkic languages have the same syntax (Hengirmen, 2002), yet different number of rules have been defined in order to design machine translation systems. For example, Azerbaijani to Turkish (Hamzaoğlu, 1993), Turkmen to Turkish (Tantuğ, 2007; Tantuğ, Adalı, Oflazer, 2007), Turkish to Crimean Tatar (Altıntaş, 2000) etc.

In the last decade, some computational researches have been done on Uyghur languages such as morphological analyzers (Orhun, Tantuğ, Adalı, 2009a; Orhun, Tantuğ, Adalı, 2009b), simple machine translation systems (Mahsut, 2004; Orhun, Tantuğ, Adalı, 2011; Yang, 2009) and some works about corpus constructions (Abaidulla, 2009; Aibaidulla, Lua, 2003). In Uyghur language, verbs are very complex and have a dominant effect on tenses. In most of the cases, the properties of a sentence could be defined by a verb. Some of the verbs have more than one meaning and irregular forms that could affect the entire sentence. Even though Uyghur verbs appear in three different tenses, but some verbs belong to more than one tense at the same time.

In this paper, some of proposals have been suggested in order to translate a future tense from Uyghur into Turkish. In Uyghur language, there are some words which belong to more than one class such as noun, adjective, adverb and verb etc. (Tömür, 1997; Osmanof, 1997, Öztürk, 1993; Kaşgarlı, 1992). Some words have more than one meaning and that meanings are completely different from each other (Osmanof, 1997). If these words are translated from one language into another, for example Uyghur to Turkish, some of the words are not changed orthographically. A reader can understand those words without considering other meanings. The reason is that both Uyghur and Turkish languages belong to the same Ural – Altaic languages family. Another reason is that the culture and mentality are very closely related to each other (Tantuğ, 2007). Uyghur is spoken in some of Central Asian Republics and mostly in Sin Kiang Autonomous region of China and different alphabets are used in different regions such as Cyrillic and Arabic scripts (Kaşgarlı, 1992). In this article, the standard Latin alphabet is used, which is standardized by the UKIJ to write Uyghur sentences (UKIJ, 2008; Duval, Janbaz, 2006).

To explain a sentence in Uyghur language and compare it with Turkish, English translation of Uyghur sentence is given in all examples.

Example 1:

Uyghur: Bu at nahayiti tiz yügüreydü.

Turkish: Bu at çok hızlı koşar.

English: This horse runs very fast.

Whenever a Uyghur or Turkish read this sentence, they can understand these sentences they are about a “horse”. Even though the word “at” has other different meanings such as “throw” and “name” etc., nobody doubts these sentences may have other meaning except for a “horse”. But the following Uyghur sentence has two different meaning at the same time and must be translated as two different sentences in Turkish even in English.

Example 2:

Uyghur: Men bu atni bek yaxshi körimen.

Turkish1: Ben bu atı çok severim.

Turkish2: Ben bu adı çok severim.

English1: I love this horse very much.

English2: I love this name very much.

In computer-based translation, all of the translations should be considered, but select the most correct one only. These kinds of situations are very common in Uyghur verbs, adjectives and adverbs etc. Especially some of the Uyghur verbs have very complicated character and more than one form (Belkiz, 2007; Belkiz, Ablajan, Dilmurat, 2007; Gülnar, 2007). But this paper describes some verbs and their effects on the tenses of a sentence.

This paper has been organized as follows; section 1 gives a simple overview about some words/verbs in Uyghur and their translations into Turkish; section 2 gives some information about simple future tense of Uyghur; section 3 describes the effects of some verbs on the future tense of Uyghur and their translations into Turkish; section 4 some explanations, suggestions and proposals about how to translate future tenses from Uyghur into Turkish and section 5 deals with the conclusions.

The Simple Future Tense in Uyghur

In Uyghur language, all of the verbs can appear in three different tenses such as future, present and past tenses. Even though both of the Turkish and Uyghur languages belong to the same language family, there are some

differences (Hengirmen, 2002; Tantuğ, 2007). In this paper the effects of the verbs on the simple future tenses are explained mainly. In Uyghur language, the simple future tense could be considered as a simple present tense even though it has been written without any grammatical difference. But in Turkish, the simple future tense and the simple present tense have different forms (Hengirmen, 2002). In modern Uyghur language, there are two different structures used to form future tense from a verb.

1. Adding the “maq+chi” and “mek+chi” suffixes to the end of verbs. Actually the “maq” and “mek” suffixes are two different forms of the single suffix “mAk”. The meta “A” character changes according to harmonization rule of Uyghur language (Orhun, Tantuğ, Adalı, 2009a; Orhun, Tantuğ, Adalı, 2009b).

Example 3:

Uyghur: Yaz+maq+chi

English: Will write

Uyghur: Yaz+maq+chi+men

English: I will write

Uyghur: Yaz+maq+chi+sen

English: you will write

Uyghur: Yaz+maq+chi+siler

English: you (plural) will write

Uyghur: Yaz+maq+chi+men+mish

English: I might write in the future

2. Adding the “i” + personal suffixes (men, sen, siz, biz, siler) to end of verbs.

Example 4:

Uyghur: Yaz+i+men

English: I will write

Uyghur: Yaz+i+sen

English: you will write

Uyghur: Yaz+i+siler

English: you (plural) will write

The main difference between these two types of future tenses is that the second type future tense may give a simple tense meaning while the first type of future tense gives explicitly future tense meaning.

Example 5:

Uyghur: Men uning’gha xet yazimen.

English: I will write him a letter.

In Example 5, these sentences describe an action that will happen in future, so we can consider this is a future tense.

Example 6:

Uyghur: Qushlar asmanda uchidu.

English1: Birds fly in the sky.

English2: Birds will fly in the sky

In Example 6, a habitual action is described. Therefore, this sentence could be considered as future tense or present tense. For Uyghur language, it does not matter if this sentence is considered as a simple tense or future tense. On the other hand for Turkish language, there are completely different. Because in Turkish the simple tense and the future tense have different forms.

In Turkish the future tense could be formed by adding the “cak/cek” suffixes while the simple present is formed with the “ar/er” suffixes at the end of a verb. Actually “cak” and “cek” also, “ar” and “er” are two different forms of “cAk” and “Ar” suffixes. The meta “A” character changes according to harmonization rule of Turkish language (Oflazer, 1995).

Example 7:

Turkish: Kuşlar gökyüzünde uçar.

English: Birds fly in the sky.

Turkish: Kuşlar gökyüzünde uçacak.

English: Birds will fly in the sky.

Turkish: Yarın babam gelecek.

English: My father will come tomorrow.

Turkish: Balıklar genelde sularda üzer.

English. In general, fishes swim in the water.

In Example 7, changes of the meta character “A” is shown with different verbs. In Uyghur language, there are two future tenses forms that can be created by adding the “dighan” and “ghusu” suffix to the end of a verb (Öztürk, 1993). This type of future tense has definitely future tense meaning.

Example 8:

Uyghur: Bu xetni men yazmaydighandimen

English: I will not write this mail

Uyghur: O mühendis bolghusu

English: he /she will be an engineer.

In order to translate simple future tense from Uygur into Turkish, it must be analyzed and should be decided if it is in a simple present tense or future tense. Being a natural language, the content of the text should be analyzed to get a correct result. Unfortunately, this article deals with rule-based solution.

Future Tenses of Uyghur and Turkish

In Uyghur language, though the simple future tense has the meaning of both future and present some verbs have to be translated into a simple tense directly without considering that it may be in future tense.

Example 9:

Uyghur: Men Uyghurche bilimen.

Turkish1: Ben Uygurca bilirim.

English1: I know Uyghur.

Turkish2: Ben Uygurca bileceğim (Wrong).

English2: I will know Uygur.

In example 9, Although a sentence is in the future tense form in Uyghur language, it does have the present simple meaning. So, it must be translated in the simple present tense form into Turkish. Therefore, the second type of translation is wrong.

Example 10:

Uyghur: Men bugün axşam kino körmen.

English: I will watch a movie in this evening.

Turkish1: Ben bu akşam film seyredeceğim.

Turkish2: Ben bu akşam film seyrederim. (Wrong).

In Example 9, the Uyghur sentence mentions an action that will happen in future, so it must be translated as future tense into Turkish. It cannot be translated as simple tense in Turkish.

Example 11:

Uyghur: Men bazargha barimen.

English1: I go to market.

English2: I will go to market.

Turkish1: Ben pazara giderim.

Turkish2: Ben pazara gideceğim.

In example 11, the Uyghur sentence could be translated into both of present and future tenses at the same time.

Example 12:

Uyghur: Men bashta öyge barimen, uningdin kiyin bazargha barimen.
English: I will go home first, after that I (will) go to market.
Turkish1: Ben önce eve gideceğim, daha sonra pazara gideceğim.
Turkish2: Ben önce eve giderim, daha sonra pazara giderim

In example 12, the verb “barimen” must be translated as a future tense. Therefore, second translation of Turkish is not valid.

As a result, the simple future tense in Uyghur language cannot be translated as a future tense into Turkish directly. In some cases, it should be considered as simple tense in Turkish as well. There are many factors related to this. Such as, category of the words, suffixes, number of words used in sentences, even the order of the words (poem) etc. In the following section, some cases have been explained with specific examples in order to distinguish future tense from simple present tense.

Proposal for Distinguish Future Tense from Present Tense

In order to translate simple future tenses from Uyghur languages into Turkish, there are some rules defined by some constraints:

1. If a sentence includes some words that describe habitual and frequent event, then this sentence may be translated as a simple present tense instead of a future tense.

Example 13:

Uyghur: Men her künü mektepke barimen.
English: I go to school every day.
Turkish: Ben her gün okula giderim.

2.If a sentence includes some words related to future time, (tomorrow, the day after tomorrow, day, time, year, week), then this sentence may be translated as a future tense.

Example 14:

Uyghur: Men ete mektepke barimen.
English: I will go to school tomorrow. Turkish: Ben yarın okula gideceğim.

3. If a sentence includes adjectives, then this sentence is considered as a simple tense.

Example 15:

Uyghur: Men seni yaxshi bilimen, shunga meni aldaymen dep awre bolma.

English: I know you very well, so don't try to cheat me.

Turkish: Ben seni iyi bilirim, dolayısıyla beni kandırırım diye boşuna uğraşma.

4. If a sentence includes both time (see the 2'nd rule of this section) and adjectives, then this sentence should be considered as a future tense.

Example 16:

Uyghur: Men ete yahshi bir kiyim alimen.

English: I will buy a nice cloth tomorrow.

Turkish: Ben yarın iyi bir giyisi alacağım.

5. If a sentence includes some words about adverbs, then this sentence may be considered as present tense.

Example 17:

Uyghur: Men paqachilap su üzimen.

English: I swim in break strake style.

Turkish: Ben kurbağlama yüzerim.

6. If a sentence includes both time (see the 2'nd rule of this section) and adverbs, then this sentence should be considered as a future tense.

Example 18:

Uyghur: Men ete musabiqide paqichilap su üzimen.

English: I will swim like a frog at the competition tomorrow.

Turkish: Ben yarın karşılaşmada kurbağlama yüzme yapacağım.

7. If a sentence includes some words about direction, then this sentence should be considered as future tense.

Example 19:

Uyghur: Men bazargha birip bir kitap alimen.

English: I will go to market and buy a book.

Turkish: Ben pazara gideceğim ve bir kitap alacağım.

If a sentence does not meet any of the rules defined above, then this sentence could be considered as a simple tense.

Example 20:

Uyghur: Men akam, singlim we anam bilen bille Istanbulda yashaymen.

English: I live with my brother, sister and mother in Istanbul.

Turkish: Ben abim, kız kardeşim ve annem ile beraber Istanbulda yaşarım.

Conclusion

In this article, there are some rules described to distinguish the simple future tense from the simple present tense in Uyghur language. As a natural language, there are many cases involved and it is not possible to clarify the rules that covers all the cases in a short time.

Final purpose of this research is to achieve a high quality, full automatic bilingual machine translation system between Uyghur and Turkish languages. Unfortunately, there are lot of adapted words in Uyghur and it is inefficient to give a definition of general rules that work on all words. Therefore, these words affect the rules explained in previous sections. To analyze some words, it is necessary to research their origin, such as Chagatai and other old Turkic languages (Abdulla, 2016). To scrutinize the verbs correctly, a well implemented morphological analyzer is necessary and a big scope of morphological analysis is under construction. This is an ongoing research with all of its aspects and hope to be accomplished soon.

References

- Abaidulla, Y., et al., (2009), "Progress on Construction Technology of Uyghur Knowledge Base", International Symposium on Intelligent Ubiquitous Computing and Education.
- Abdulla, A., (2016), "Chaghatay Language is The Bridge between Old Turkic and Modern Uyghur language", International Journal of Uyghur Studies, No.6, 2016.
- Aibaidulla, Y., Lua, K., (2003), "The development of Tagged Uyghur Corpus", Proceedings of PACLIC17, Sentosa, Singapore, pp: 228-234.
- Alsharaf, H., Cardey, S., Greenfield, P., Shen. Y., (2004), French to Arabic Machine Translation: The Specificity of Language Couples", Actes de l'International Conference on Informational Technology, ITCC 2004, Nevada, USA, IEEE Computer Society, Vol2, pp:293-297.
- Altıntaş, K., (2000), Turkish to Crimean Tatar Machine Translation System, Computer Engineering Department. Msc Thesis, Bilkent University, Ankara.
- Altintas K., Çiçekli İ., (2001), "A Morphological Analyzer for Crimean Tatar", Proceedings of the 10th Turkish Symposium on Artificial Intelligence and Neural Networks, pp: 180-189.
- Belikiz, (2007), The 3253 different word forms Uygur Verb "qil". Corpus Linguistics and Corpus Based Research, Dept of Linguistics, College of Anthropology Xinjiang Normal University. <http://www.xjcorpus.net>.
- Belikiz, Ablajan, Dilmurat, (2007), The 2107 different words forms of Uyghur verb "bol". Corpus Linguistics and Corpus Based Research, 01/07/2007. Dept of Linguistics, College of Anthropology Xinjiang Normal University. <http://www.xjcorpus.net>.
- Chandioux, J., (1976), Météo: Un système opérationnel pour la traduction automatique des bulletins météorologiques destinés au grand public, Meta, 21, pp:127-133.

- Duval J.R., Janbaz, W. A., (2006), "An Introduction to Latin Script Uyghur, Middle East & Central Asia Politics", Economics, and Society Conference. Sept 7-9, 2006, University of Utah, Salt Lake City, USA.
- Gülınar E., (2007), "Resistance to Borrowing of Uyghur Verbs". Annual Conference, University of Washington, 2007.
- Habash, N., Hu, J., (2009), "Improving Arabic-Chinese Statistical Machine Translation using English Pivot Language", Proceedings of the 4th EACL Workshop on Statistical Machine Translation, Athens, Greece, pp: 173-181.
- Hajic, J., (1987), "RUSLAN - An MT System Between Closely Related Languages", in Third Conference of the European Chapter of the Association for Computational Linguistics (EACL'87) Copenhagen, Denmark, pp: 113-117.
- Hajic, J., et al. (2000), "Machine translation of very close languages", in Proceedings of the sixth conference on Applied natural language processing, Morgan Kaufmann Publishers Inc., pp: 7-12.
- Hamzaoglu, İ., (1993), Machine translation from Turkish to other Turkic languages and an implementation for the Azeri languages, in Institute for Graduate Studies in Science and Engineering. MSc Thesis, Bogazici University, Istanbul.
- Hengirmen, M., (2002), Türkçe Dilbilgisi, Engin yayın evi, Anakara, Türkiye.
- Huet, S., Manishina, E., Lefevre, F., (2003), "Factored Machine Translation Systems for Russian-English", 8th Workshop on Statistical Machine Translation, Sofia, Bulgaria.
- Kaşgarlı, S. M., (1992), Modern Uygur Türkçesi Grameri, İstanbul Orkun Yayınevi, 1992.
- Mahsut, M., et al., (2004), "An Experiment on Japanese-Uighur Machine Translation and Its Evaluation", Conference of the Association for Machine Translation in the Americas AMTA 2004: Machine Translation: From Real Users to Research, pp: 208-216.
- Nagao, M., (1984), "A Framework of a Mechanical Translation Between Japanese and English by Analogy Principle", Proc. Of the International NATO symposium on Artificial and Human Intelligence, Lyon, France, pp: 173-180.
- Oflazer, K., (1994), Two-Level Description of Turkish Morphology. Literary and Linguistic Computing, Vol. 9, No:2.
- Orhun, M., et al., (2009a), "Rule Based Analysis of the Uyghur Nouns", International Journal of Assian Language Processing, 19(1), pp: 33-43.
- Orhun, M., et al., (2009b), "Rule Based Tagging of the Uyghur Verbs", Fourth International Conference on Intelligent Computing and Information Systems, Faculty of Computer & Information Science, Ain Shams University Cairo, Egypt, pp: 811-816.
- Orhun, Murat et al., (2011), "Uygurcadan Türkçeye bilgisayarlı çeviri", İstanbul Teknik Üniversitesi, Mühendislik Dergisi, Cilt 10, Sayı 3, pp:3-14.
- Osmanof, M., (1997), Hazırkı Zaman Uyghur Edebiy Tilining İmla ve Teleppuz Lughiti. Xin Jiang Xeliq Neshiryatı. (in Uyghur).
- Öztürk, R., (1993), Yeni Uygur Türkçesi Grameri. Türk dil kurumu yayınları: 593, 1993.

- Philipp, K., (2005), Europarl: A parallel corpus for statistical machine translation. In Proceedings of MT Summit.
- Tantuđ, A.C., (2007), Akriba ve Bitiřken Diller Arasında Bilgisayarlı eviri İin Karma Bir Model. Bilgisayar Mühendisliđi Bölümü. Doktora Tezi, İstanbul Teknik Üniversitesi, İstanbul.
- Tantuđ, A. C., et al, (2007), “Machine Translation between Turkic Languages”, Proceedings of the ACL 2007 Demo and Poster Sessions, pp: 189–192.
- Tömür, H., (1997), Modern Uygur Grammar (Morphology). Yıldız Teknik Üniversitesi, Fen-Ed. Fak. T.D.E Bölümü. İstanbul.
- UKIJ, (2008), Uygur Komputer İلمي Jemiyeti . <http://www.ukij.org/>
- Yang, P., et al., (2009), “Chinese-Uyghur Machine Translation System for Phrase-Based Statistical Translation”, Journal of Computer Applications, 29(7).

Murat ORHUN
Uygurca ve Trkede Gelecek Zaman Kip zellikleri