Aminem Memtimin

Language Contact in Modern Uyghur

TURCOLOGICA

Herausgegeben von Lars Johanson

Band 108

2016

Harrassowitz Verlag · Wiesbaden

Aminem Memtimin

Language Contact in Modern Uyghur

2016

Harrassowitz Verlag · Wiesbaden

Contents

Preface

The first thing I feel I need to do in this preface is to explain the title of the book, 'Language Contact in Modern Uyghur': The use of the preposition 'in' may seem a bit odd, but I think it is the only possible one in view of what the work is trying to do. My research is based on present-day Uyghur language data and those beginning in the middle of the previous century, couched in the language use of modern Uyghur society: Therefore 'in' Uyghur. But these data have been used synchronically and diachronically, for exploring four quite different contact systems, a virtual language contact laboratory, as it were. There was not much physical speaker involvement for the contact with Persian and Arabic, two very different languages themselves engaged in an involved contact situation, which created the culture by which much of the Turkic population of Xinjiang has identified itself for already a thousand years. Although the Uyghurs at a quite early stage submitted to Mongol rule for a number of centuries, cultural influence went the opposite way: Mongol tribes submitted to Turkic language, culture and identity, flooding the Uyghur language with influence from the bottom up; the Mongolic influence, in many ways more significant than the others, has mostly gone unnoticed by scholars. The language used at that time was not yet 'Modern Uyghur'; Modern Uyghur became what it is during the Mongol domination. Russian influence opened up the western modern world to the Uyghurs, a huge and pretty sudden world-view revolution reflected in language. Uyghur cultural and linguistic contact with Chinese existed already in Tang times but became more socially relevant with the establishment of Qing rule in 1759; it is practically the only relevant contact nowadays, lending itself to wide-scale synchronous research. Although much of our subject matter is in the past, our present-day material lends itself to the contrast of five different types of language interaction, with five languages which are typologically quite different from each other and belong to four different genetic families.

This book is a revised and expanded version of my doctoral dissertation, defended in April 2014 in Frankfurt am Main, Germany. Recalling those 'fantastic' days, I would like to mention the kind persons who gave me their valuable time, ideas and encouragement and supported me with their inspiring suggestions. First, I would like to express my deep gratitude towards my 'Doktorvater', Professor Marcel Erdal, who supervised my whole study in Frankfurt; without his serious and scientific support, this endeavor could not have been carried out. Special thanks go to Professor Irina Nevskaya, who kindly led me in the area of contact linguistics with her profound knowledge and patient instruction. I also owe her much for her serious suggestions concerning the theoretical part and the Russian chapter. My equally great thanks are for Professor Muhämmätrehim Sayit, who supported my research from the beginning on. His important remarks on the Chinese part, as well as those of

Professor Huang Chenglong and researcher Wang Haibo, prevented some shortcomings.

Furthermore, I sincerely thank Professor Abdurishid Yakup, who inspired and helped me during his stay in Frankfurt and gave me very important suggestions during the revising process. Professor Arslan Abdulla from Xinjiang University made work more convenient for me by generously offering his research materials. I use the opportunity to also express my warm gratitude to Professor Bernt Brendemoen, who carefully read and commented on my whole work, and to Dr. Bayarma Khabtagaeva, who read the Mongolic part thoroughly; both kindly offered me their valuable suggestions, from which I profited much.

I am grateful to Frankfurt University, which offered me access to various valuable scientific materials and a rich library. I wish to thank all teachers and colleagues in Frankfurt University, from whom I absorbed new scientific ideas and methods, profiting from their knowledge and friendship. I am especially thankful to Professor Károly, Dr. Nugteren, Professor Ragagnin and Dr. Jügel for their advice in some scientific matters. Professor Lars Johanson kindly read my work and happily accepted it for publication; I again wish to express my thanks to him. Mr. Şahin Beygu greatly helped me with his serious editing work. I owe great thanks to the China Scholarship Council, which financed my study in Germany, and the Uyghur Language and Literature department at Minzu University of China which also supported my study. Last but not least, I wish to express my gratitude to my mother Hashirem Pazil and my father Memtimin Abdurehim, my sister Aynur and other relatives who have supported me for many years; without their material and spiritual support I would not have been able to carry out my research.

This work is obviously based on the valuable contributions of many previous investigators, some of which may not have been duly appreciated here for lack of space and time. I would here like to express my warm gratitude to all those researchers with my sincere apologies. All the insufficiencies and shortcomings are, of course, mine.

Language contact is everywhere, as shown already by the names of my family members. It is such a complex phenomenon that a study such as this one can surely not cover everything in breadth and depth even concerning one language. Still, I hope that this perspective on Uyghur contact scenarios will serve as a useful base for further research in language contact and typology and, to some degree, as a small contribution for understanding some aspects of Uyghur society! I wish my reader would enjoy the work as much as possible: Bitte schön!

Räxmät!

Beijing, March 2016   Aminem Memtimin

1
Introduction

1.1 Theoretical and methodological background

Similarities between languages can be linked to several factors: coincidence, universality, genealogy and language contact. The present study focuses on the fourth one among these factors.

Language contact triggers not only linguistic similarities among non-related languages, but also divergence among genetically related languages, in our case the Turkic languages. Its study is a necessary ingredient of the synchronic and historical research on language.

The goal of contact linguistics is to uncover the various factors, both linguistic and socio-cultural, that contribute to the linguistic consequences of contact between speakers of different language varieties (Winford 2003: 11). Weinreich's (1953: 44) emphasis on combining structural consideration, psychological reasons and socio-cultural factors must be the guiding framework in contact linguistic research.

In the present research, the contact linguistic scenarios which have contributed to shape Modern Uyghur will be discussed in both their linguistic and sociolinguistic aspects, based on empirical material; to elucidate social interactions, each chapter also features a quite detailed historical section. Comparing Uyghur grammar and lexicon with those of the languages it has been in contact with, coincidence does not seem too likely, as all the languages involved (except-to some extent Turkic–Mongolic) are typologically so different from each other and belong to different language familes: Semitic, Iranian, Mongolic, Slavic and Sino-Tibetan. From the point of typological similarity, one would expect the integration of Mongolic elements to be easier, the integration of Chinese elements to be much more difficult, since Chinese is typologically so different from the Turkic languages in general. Universal features and constraints need to be taken into consideration in all aspects of grammar, as I have tried to keep in mind. Genealogy is relevant only in the Turkic–Mongolic relationship; I have here put effort into excluding this topic, as it belongs to prehistory. This is both a diachronical study (in the case of Arabic, Persian, Mongolian and Russian) spanning over more than one thousand years, and also a synchronical one, as the Chinese impact is going on.

Why do languages borrow from one other? What can be borrowed? How are these borrowings selected? Why are some structures more attractive than others? Could one list the scale of borrowed elements according to their borrowability?

What are the mechanisms of contact-induced language change? What is the outcome of language contact? Such questions have become central matters of contact linguistics and scholars have produced remarkable contributions on these topics both in the empirical and the theoretical domain. The present work evokes most of these questions.

Various theoretical frameworks have been proposed in contact linguistics. There are different terms for related concepts such as borrowing, code copying, transfer, adoption, imposition, camouflaged borrowing, PM (phonetic matching) and PSM (phono-semantic matching), calque (semantic borrowing), guestword (unassimilated borrowing), foreignism (phonetic adaptation), loanword and others. Some of the terms overlap with others, and were proposed by different authors in different times or within different frameworks.

Haugen (1950: 211 ff.) considers 'borrowing' not to be a proper term for linguistics, "since the borrowing takes place without the lender's consent or even awareness, and the borrower is under no obligation to repay the loan", as he says. He suggests calling the process 'adoption', but he still used the term 'borrowing', considering that it does not cause misunderstanding among linguists, since no appropriate term had yet been proposed. He defines borrowing as "attempted reproduction in one language of patterns previously found in another". The original 'pattern' is the 'model' language. He also distinguished sharply between the borrowing process and its results. 'Importation' and 'substitution' are related to the process: 'importation' for cases when the result in the recipient language is practically indistinguishable from the source, 'substitution' for cases of inadequate reproduction of the model. The degree of morphemic substitution decides the result of loans as 'loanword', 'hybrid' or 'loanshift'. Typological similarity has a decisive role in borrowability. 'Hybrid creation' is one of his useful terms; this is how he called something created secondarily in the borrowing language. His attitude towards borrowability was that all linguistic features can be borrowed, but they are distributed along a scale of 'adoptability' which is in some way related to structural organization. His method of combining synchronic and diachronic comparison is of use in identifying borrowings.

Johanson 2002 (1992) proposed the Code Copying Model, which distinguishes between global copies, selective copies and mixed copies. This model has been successfully applied during the last years in much work on Turkic language contacts and beyond. Here, the apter term 'copying' has replaced borrowing. In global copying, a B pattern is copied into an A basic-code clause in its entirety, i.e., as a block of materal, combinational, semantic and frequential structural properties. In selective copying, the model consists of selected structural properties of a B block, e.g., characterics of a material, combinational, semantic and frequential kind. Mixed copying contains at least one global copy, such as lexically mixed copies (Johanson 2002:18). In Johanson's sense, adoption signifies that the socially dominant code B influences the socially dominated A code, while in imposition, speakers of A carry their features into their B varieties.

Van Coetsem's distinction between two types of transfer is important. He calls one type 'borrowing' (pull transfer); here the agent of the transfer is the speaker of the recipient language. In the second type, called 'imposition' (push transfer), the agent is the speaker of the source language: In his terminology, 'transfer' covers all kinds of cross-linguistic influence. RL (recipient language) agentivity affects the less stable components of the RL, while SL (source language) agentivity affects the more stable ones (2000: 66). He also distuinguishes two types of operations as imitation / adaptation and imposition. In RL agentivity, the former functions, while imposition is at work in the SL agentivity.

Zuckermann 2003, on the other hand, developed the ideas of Haugen and applied the term 'camouflaged borrowings' under the framework of FEN (Folk Etymological Nativization); 'camouflaged borrowing' means covert borrowing, in which a SL lexical item is replaced by a target language (TL) item which relates semantically (calque / loan translation), phonetically or phono-semantically (PS, SPM and PSM) to the SL. He illustrated these types of borrowing with samples from different languages, and described the PSM (phono-semantic matching) model as an ideal perspective on the creation of new terms. PM (phonetic matching) signifies that TL material is originally similar to the source language lexical item phonetically but not semantically. In PSM, the target language material is originally similar to the SL lexical item both phonetically and semantically. He made the classification of borrowings based on whether they use SL or TL items as the basic material for the neologization. If it is the former, then 'guestword', 'foreign words' or 'loanwords' are expected; in the case of the latter, PM, SPM or PSM. Apparently, Turkish applied the second type of camouflaged (MSN) borrowing very effectively during the most active years of Turkish language reform, while Uyghurs have, in recent years with Chinese contact, been applying the second type, calque.

The contact linguistic handbooks of Thomason 2001 and Winford 2003 have been very useful; they are directly based on research, illustrate topics with a wide range of examples and analyze phenomena with a combination of social, psychological and structural frameworks.

Concerning Uyghur language contact, Arabic, Persian and Mongolic are old contact languages: They had contact with Chaghatay, the Muslim written language shared in many varieties by Central Asian Turkic peoples between the early 15th and the early 20th century, which is the predecessor of Modern Uyghur. But looking at Modern Uyghur, it seems that Mongolic influence on it went beyond what we witness in Classical Chaghatay both in the lexical and in the grammatical domain. Russian is a relatively recent contact language, whose influence reached its height in the first half of the 20th century but has since become marginal as contact language for Uyghur. These all contrast with Chinese, which is now in intimate contact with Uyghur. Hence, the description of Arabic, Persian, Mongolic and Russian linguistic contact is quite different from what we can observe in the ongoing contact with Chinese, which has been supplying Uyghur with borrowings and calques and is interacting with Uyghur oral communication in the form of code switching. It should

be clear that there is no real way to reconstruct oral contact situations of history as they took place: although classical Chaghatay authors are likely to have had good Persian communicative skills, what I took into consideration in the older contact languages was only their written forms; these were taken as the base of the comparison with Modern Uyghur. The grammar and lexicon of Modern Uyghur are in any case the result not only of these contacts but also of internal development.

Why is vocabulary more easily borrowed than grammar? The vocabulary has a weaker form of structuredness; it consists of open lists including an unlimited number of items. Grammar, on the other hand, has a strong structuredness and closed lists including a limited number of items (e.g. flexives). The speaker uses vocabulary items less frequently and thus less automatically and is more conscious of them. These items are in general less stable. Grammatical items, however, are more stable; they are used more frequently and therefore more automatically by the speaker, so he is less conscious of them (van Coetsem 2000: 107). The Uyghur dictionary has thousands of loans, but many of them are rarely used in normal communication; this question will be better answered through statistics of text frequency and token frequency. Of course, innovations of the vocabulary are necessitated by the adoption of a new religion and a new cultural paradigm, as was the case with Arabic and Persian; later, the replacement of social structures and the the introduction of technical and scientific reforms, as is the case with a large part of both the Russian and the Chinese vocabulary.

In fact, it is certainly "not just words that get borrowed: all aspects of language structure are subject to transfer from one language to another, given the right mix of social and linguistic circumstances" (Thomason 2001: 10–11). The hierarchy proposed by Thomason and Kaufman (1988: 50) to show the types of linguistic outcomes is of value in describing the typology of the states of contact languages. Van Coetsem (2000) proposed the stability gradient which takes both types of agentivity into account: inherent stability which is based on structuredness, and subsidiary stability which is determined by circumstantial factors such as speaker's attitude or affinity between languages. In RL agentivity, less stable items are borrowed primarily, while in SL agentivity, the most stable items will be transferred to the RL by imposition.

According to the borrowing scale of Thomason & Kaufman (1988: 74–76), the contact results on Uyghur in some cases reached the level of "very strong cultural pressure", causing "heavy structural borrowing"; e.g. in that prefixation was added to the structure of Uyghur adjectives and pronouns (see below).

In the Uyghur contact situation, I have attempted to discover what can be borrowed, how such borrowings are incorporated into the target language and the motivations for triggering contact induced change.

In contact linguistic theory, structural and social factors are distinguished as motivations for borrowing. Among structural factors, universal markedness, the degree to which features are integrated into the linguistic system, typological distance and transparency are important factors which promote or impede language interference,

while need, prestige, speakers' attitude, awareness, presence vs. absence of imperfect learning and the intensity of contact are social factors linked to the contact issue. Language attitude can either restrict influence or be an enhancing factor and lack of consciousness or automatic behavior strengthens stability. (see van Coetsem 2000, Thomason 2001, Johanson 2002). Van Coetsem (2000: 107) points out the relationship of high frequency with more stability and low frequency with less stability. Attractive features determining partly the likelihood of the copying of certain structures are illustrated well by Johanson (2002: 44–48) with Turkic language examples: transparentness, agglutinative structure, relatively invariant allomorphy, the syllabicity of affixes offering the possibility to pronounce them in isolation, peripheral position, more concrete semantic elements and less 'marked' structures are analysed as being more attractive than others.

Concerning Uyghur contact situations, it is evident that both social and structural motivations took part in promoting or impeding borrowings.

In the more specific studies on contact linguistics dealing with code-switching, Carol Myers-Scotton, who has done so much on this topic, should be mentioned as a pioneer with her Matrix Language Framework model; her work together with that of Penelope Gardner-Chloros have served me as base when considering Uyghur–Chinese code switching.

My attitude has been to view the language contact facts through the concepts of the different comprehensive frameworks, combining structural and sociolinguistic empirical material on the base of Haugen's ideas, the concepts developed by Johanson under the Code Copying Model, and the transfer theory in the sense of van Coetsem as well as the FEN model of Zuckermann.

1.2 The organization of the work

This study consists of an introduction, five chapters, the conclusion, one illustration (the photograph of the menu of an Uyghur restaurant, accompanied by transliterations and translations) as appendix, and the bibliography.

The introduction has three parts: First there is an account of the methodological and theoretical background and of previous research on the topic, then information on the database and on dictionary sources. Secondly there are lists of glossing term abbreviations, languages and sources and the transcriptions of the different languages referred to in the work. The third part of the introduction gives general information on the Uyghurs, on the modern Uyghur language, on the present contact situation, on the Turkic and Uyghur varieties in Xinjiang (where most of the Uyghurs live) and on contact with Chinese.

The first and second chapter are about Arabic–Persian contact, the third on contact with Mongolic, the fourth on contact with Russian and the fifth chapter on contact with Chinese. The second, third and fourth chapters start with quite detailed accounts of history, as we need history to explain the sociolinguistic facts: How

Islam and the Persians reached Central Asia and the Tarim Basin, how the Mongolians came to rule Xinjiang for so many centuries, how it came about that Russian had such a significant lexical influence on Uyghur. I did not add any such section for Chinese as the chapter on this language deals with the present and the previous history sections have some information also on that. After general history, each chapter has a section on the contact situation between the various languages and Uyghur. There follow detailed accounts on the various aspects of phonetic and phonological changes which words borrowed from those languages underwent, on how Uyghur dealt with the morphological elements in those languages, how it added its own derivational and inflectional segments to the borrowings, what parts of speech it borrowed and what parts of speech the borrowed words formed in Uyghur. Then there are sections on complex verbs derived from foreign verbal nouns and the like, the creation of verb and noun phrases involving foreign elements, the borrowing of noun phrases and other expressions and the syntactic aspects of contact with the various languages. The semantic sections give detailed lists of the types of lexemes borrowed, offer semantic classifications, mention semantic changes linked to the borrowing process, and discuss loan translations. In the chapter on Mongolic the various possible historical stages and Mongolic languages from which the borrowings might have come are discussed, in the chapter on Chinese the Xinjiang dialect and Standard Mandarin as alternative possible sources. The Chinese chapter has many more sections as appropriate for the lively contact situation.

1.3 Previous research

Much material has been brought together on the contacts of different Turkic languages such as Yakut, Turkish, Azeri, Chuvash, Chalkan, Kazakh, Kirghiz, Shor or Tuvan within the frameworks of Turcology and Altaic studies. There is some work dealing also with the contact phenomena around Chaghatay, which in many ways is the predecessor of Modern Uyghur. A. Yakup 2005: 170–190 deals in detail with all aspects of foreign lexical influence on the Turpan dialect, in work which is highly relevant for the whole Modern Uyghur domain. Y. Qurban 1988, R. Haimudula 1990, A. Abdulla 1996, P. Jilan 2002, Ä. Äxmät 2009, Niyu Dečiŋ 2002, A. Abudurexiti 2008 and A. Sayim 2006 are general descriptions of Uyghur borrowing from other languages; Rentzsch 2008 discusses the lexicon as well as various aspects of grammar. Ma Deyuan 2012 described in detail foreign language influence on the temporal expression of Uyghur. Li Pu & Gao Jinglian 2013 discussed some borrowed morphemes in Uyghur. Most other articles on this topic mainly showed some Chinese lexemes which were copied into Uyghur.

There are some important lexical material collections for Modern Uyghur. The explanatory dictionary of loanwords in Uyghur by H. Abduraxman et al. 2001 is practical for consulting, though it did not avoid some superficial judgment about the source of some loanwords; it never, e.g., mentions Mongolic as a source. A few

articles discussed the ultimate sources of borrowings; although such work may be interesting from the point of view of the dominance of some languages (such as Classical Greek) in universal coining, they are not very useful from the language contact point of view. Some articles about the Arabic–Persian, Russian and Chinese loan words in Uyghur have also been published; among these, the ones dedicated to Uyghur–Chinese contact are more numerous than the others.

Concerning Arabic–Persian loanwords in Uyghur there is Y. Ghojäxmät 1984, R. Hämdulla 1987, T. Rozi 1999, T. Raxman 1999, A. Ähät 2001, M. Maimaituerxun 2012, Gh. Ghäyurani & R. Ghänizat 2007 and P. Amuti 2009. There are several master theses on Arabic loanwords in Chaghatay, such as the one of A. Aihemaiti (2014). The following points may characterize these publications: Borrowed lexemes were classified into semantic domains and some semantic changes of Persian lexemes were analyzed; some criteria for distinguishing loan words from native words were also proposed, such as vowel harmony, consonant clusters and the distribution of some sounds. We find that the direction of some borrowings was not clarified and often judged superficially; e.g. Uyghur *očaq* 'oven' was wrongly analyzed as coming from Persian *oja:q*, positing a consonant change *j > č*. In fact, this is clearly an Oghuz form borrowed from there into Persian: It originally consisted of the stem *o:t* 'fire' and denominal suffix *-čAk* (see Erdal 1991: 108); voiceless consonants became voiced in Oghuz Turkic if preceded by a Proto-Turkic long vowel (e.g. *at* 'name' with 3rd person possessive suffix becomes *adı* 'her / his name' in Turkish).

Mongolic–Uyghur contact is the least investigated domain in Uyghur contact research, but it turns out to have been highly important both in grammar and in lexicon. A number of articles and books were devoted to the study of Mongolic influence on other Turkic languages, such as Poppe 1991, Csáki 2006, Khabtagaeva 2009, Sertkaya 1992, Ölmez 2007, Abdullayev 1992, Kincses 1997, Schönig 2000; the papers of Sertkaya, Kincses and Ölmez deal with Mongolic loanwords in Chaghatay. It is especially important to investigate the numerous Mongolian loanwords in Uyghur dialects. A. Abdulla & P. Mamut (1998) discussed some archaic words in the Qumul dialect and mentioned some "Altaic common words" which exist in this dialect as well as in Mongolian and Daghur Mongolic. Fu Maoji et al. (2000) deal with Mongolic loans in the Lopnor dialect; A. Yakup (2005) collected and classified the Mongolic loans in the Turpan dialect. The M.A. thesis of Qiqige 2005 compared the lexicon of the Lopnor dialect with Oirat Mongolian. In all, the Mongolic loans in Uyghur dialects are far from having been systematically investigated; there appear to be numerous dialect borrowings still waiting for investigation. We find that some reborrowings were not dealt with care; the origins of words were not clarified according to the phonological rules and the historical development of the language. E.g. Old Turkic *qatır* was copied into Mongolic, where it became *qačir* because of the Mongolic cancelling of the /ı : i/ opposition and the rule *ti > či*; its appearance in Modern Uyghur as *qečir* 'mule' is due to reborrowing. The existence of *qatır* in

Turkic and the phonological rules of Mongolic and Modern Uyghur explain its origin effectively.

Among publications about the Russian influence on Modern Uyghur we can mention R. Räxim 2001, M. Abdulla 2002, Yuan Shengwu 1996, Ä. Jappar 1989, K. Tahir 2002, A. Abudula, 2010, A. Abudurexiti 2008 and 2011 and T. Baiheti 2015. Here, Russian borrowings were again classified into semantic domains and some phonetic adaptations were discussed. The spelling of Russian loan words was another topic dealt with.

Russian and English share the general European coining system heavily dependent on Greek and Latin, in both languages sometimes received through French. It is therefore not surprising that they have many common lexemes. However, phonetics often show that the direct source for Uyghur was not English but Russian. One such word is Uyghur *maltoza* 'maltose', which has the same shape as in Russian; English *maltose*, on the other hand, does not end with a vowel in English. The first vowel of 'microphone' is a diphthong in English but not in Russian and Uyghur. Some even suggested correcting such borrowings according to English pronunciation. Words like *kapitalizm* or *materiyalizm* appeared in Uyghur at a time when contacts with Russia were strong but contact with English-speakers practically nonexistent. That Uyghur *karton* 'cardboard' is not from English *cartoon* ('story expressed through drawing'), e.g., but from Russian *karton* (ultimately from Italian via French) is shown by their semantics. Concerning borrowings from English we can mention the list of English loanwords in H. Abduraxman 2001 and the papers Sung Je 2005, Li Huixing 2003, Liu Geling 2007, W. Dawuti & G. Maituohuti 2010, Zheng Ranghong 2010, O. Dawut 2012 and A. Kawuzi 2012. These publications show us the semantic domains of English copies in Uyghur. Dawut dealt more systematically with the standardization of English loans in Uyghur.

Zhao Xiangru 1984, Zhang Shufang 1994, Ma Deyuan 1995, Zhong Jiafen 1995, X. Niyaz & M- Qasim 1998, Yuan Shengwu 1998, Zhang Yang 1998, 2009 and 2011, G. Ghopur 1999, Peng Yan 1999, Gao Liqin 2005a and 2005b, Wang Yang 2004, Zhong Jiafen 1995, Chen Shiming 2004, Ou Yangwei 2000, Yan Xinghong & Ou Yangwei 2000, Zhang Liping 2006, G. Aibai, 2007, M. Ajiaikepaer 2007, A. Taji 2008, Chen Huiyou & Wang Lijun 2010, Li Dehua 2011, A. Baki 2012, G. Abudula 2013, Z. Akebaierjiang 2014, R. Maimaiti 2014, Wangjuan 2015 and T. Kuerban et al. 2002 all deal with contact with Chinese, generally again with the semantic domains of Chinese borrowings. Hayasi 2009 looks at the topic of Uyghur–Chinese contact from the perspective of phonological nativization. Some papers discuss the matter of the standardization of Chinese spoken and written forms, e.g. suggesting that the spoken Chinese forms should be taken into consideration in standardizing loanwords in Modern Uyghur. In the book of Gao Liqin (2005b), statistics were offered of Chinese loanwords in different stages of Turkic from the earliest times onward. She found 65 Chinese words in the Uyghur part of the 18th

century pentaglot dictionary. Among the Modern Uyghur dictionaries, Šähidi's (1953) dictionary, she says, contains 128 Chinese loans among 12405 entries.[1] Zhao Xiangru 1984 discussed spelling rules of loanwords in Uyghur. Ma Deyuan 1995 explained some loan translations and idioms on the base of Chinese. The papers of G. Aibai 2007 and A. Baki 2012 are about the Uyghur influence on Chinese dialects. M. Ajiaikepaer 2007 discussed some Chinese loanwords in oral Uyghur dialects and positive or negative aspects of their influence on the development of Uyghur language. Zhang Yang 1997, 2009 and 2012, Liao Dongmei 2005 and Zhangyang & Tianyunhua 2014 discussed some structural innovations in local Chinese dialects under the influence of Uyghur. Further papers on the Uyghur impact on Xinjiang Chinese are not mentioned here as the topic is beyond the scope of the present study. A. Taji 2008 pointed out some structural innovations which took place in Uyghur under the influence of Chinese. R. Maimaiti 2014 wrote about the Chinese loanwords in the Turpan dialect; for this topic, cf. also especially A. Yakup 2005.

There is not much sociolinguistic work. The M.A. thesis of Wang Yang (2004), Wang Yang (2007) and Du Liangxia & Cui Youwei (2015) and some other papers describe the language attitudes of Uyghurs in Xinjiang, concentrating on the language attitude of middle school and college students and some adults in the Ürümqi, Turpan and Ghulja areas. A. Abudula 2014 observed language attitudes and its causes among the Uyghur intellectuals.

Work on the current Chinese influence on spoken Uyghur and on face-to-face Uyghur–Chinese communication is insufficient. Uyghur–Chinese code switching is a highly interesting phenomenon which has only recently attracted the attention of scholars, and studies in this domain have mainly dealt with its lexical aspect; its grammatical aspects are far from being systematically described. See A. Xialifu 2007 and Haifeng 2008 on this topic; the latter analyzed the causes of code switching in detail. The work of Cao Xianghong (2013) on this matter is important: On the base of materials collected in Xinjiang he described the state of language use, the occurrence of code switching and its functional and social aspects among Uyghurs from a sociolinguistic point of view. Further work on this topic is M. Simayi 2008, Zhang Caiyun 2009 and Xuelin 2009; the latter observed code switching occurrence in second language classrooms, their features and functions. Z. Abulimiti 2009 analyzed the motivation of code switching among Uyghur university students from a sociolinguistic point of view and pointed out that female students held a more positive attitude towards code switching than males; its frequency correlates with proficiency in Chinese. Another factor supporting code switching occurrence is the linguistic economy principle favoring Chinese over Uyghur in some expressions. Z. Abulimiti & Tang Yunfeng 2011 reached similar results. A.

---

1   It is not clear why translations such as *bäš žilliq plan* 'five years plan' or *juŋxua xälq jumhuri- yiti* 'People's Republic of China' are considered to be loan words. Nor should Uyghur deriva- tions like *säyxanä* 'vegetable house' or *säypuŋčiliq* 'tailor's profession' be put in the list of loans when their bases *säy* and *säypuŋ* are also mentioned as such.

Zaoreguli & Tang Yunfeng 2014 studied the motivation of code switching among Uyghur students in Ürümqi. Yang Fan 2012 carried out a survey on the situation of SMS usage and its motivation among Uyghur university students in Ürümqi. G. Akebaer 2008 investigated code switching among Uyghur university students and observed "that it more easily involves nouns than other word classes." This is certainly no longer true today for code switching in Ürümqi, as described by Cabras 2014a and Cabras 2014b. One error of many of the contributors to this discussion is not to distinguish between Chinese borrowings in use in spoken Uyghur and code switching; *diyänši*, e.g., is a common Spoken Uyghur copy of Chin. 电视 *dianshi* 'television' not reflected in the written language. Such established borrowings of the oral language have been mistakenly put into the category of code switching.

Generally speaking, the research on Uyghur mentioned above has supplied a base for further research and offered useful materials and suggestions. We find quite a large number of publications, but some serious weaknesses appear in some of them. One weakness is that elements copied into Uyghur were taken to have kept their original function also in the language in which they landed; this shows a basic misunderstanding of the phenomenon of language contact: The model and target language have different grammatical and semantic structures and the copied element can never do in the target language what it does in the model language. Take the copying of Persian prepositional phrases and compounds: These cannot be treated as Uyghur affixes if they did not acquire the function of forming new words in Modern Uyghur itself. If, on the other hand, they function freely within Uyghur structure, they can no longer be called borrowings. Such is the element *-xana*, suffixed to nouns to create names for places linked to the base noun, e.g. in *torxana* 'internet café' from *tor* 'net', created as a calque on Chinese *wǎngbā*, literally 'net café' from 网 *wǎng* 'net'. The Central Asian Turks borrowed *xana* from Persian a millenium ago, and it was inherited into Modern Uyghur: *Now* it is just as Uyghur as any originally Turkic suffix. When dealing with the impact of Chinese, the written and fixed varieties of the languages were mostly taken into consideration; the strong influence in oral contact was not much in the scope of previous research. Different source types such as spoken and written or standard and dialect varieties were not taken into consideration in dealing with Russian and Chinese copies; in particular, the phonetic rules of local Xinjiang Chinese dialect words were wrongly taken to be the result of adaptation to Uyghur (which they mostly are not, as discussed in section 6.3.2.). Most publications dealt only with lexical copies; grammatical copies were rarely touched upon. Another topic which needs further study, finally, is the numerous Chinese calques found in Uyghur (as the one mentioned a few lines before), echoing the semantic copies from Persian which flooded the earlier stages of the language.

1.4 The database and dictionary sources for the topic

The Uyghur lexical materials analyzed in my research partly come from the *Explanatory Dictionary of Uyghur* edited by the Millätlär Til-Yeziq Xizmiti Komiteti (Nationalities' Language and Writing Service Committee) in 1999, and its electronic version, prepared by several colleagues (including the present author) under the supervision of Prof. Muhäbbät Qasim in Xinjiang University during the year of 2006–2007. Colloquial forms, sentences and dialogues are taken from my recordings during 2006–2007 and 2009–2010 in Xinjiang; some sentences were extracted from Uyghur websites. Some other samples were taken from textbooks such as Tömür 1993. The Chaghatay dictionary of M. Bahawudun et al. (2002) is (together with Eckmann 1966 and Bodrogligeti 2001) used for Chaghatay examples; It also served me for bringing together the Modern Uyghur counterparts of the Chaghatay lexemes. The *Dialect and Subdialect Dictionary of Modern Uyghur* (Jilan et al. 2007) and the *Hazirqi Zaman Uyyur tili Diyalektliriniŋ Fonetika, Girammatika we Leksikisi Üstidä Selišturma Tätqiqat* (Comparative Studies on the phonetics, grammar and lexicon of the Modern Uyghur dialects) (M. Qasim 2014) were useful, among other things, for discovering Mongolian loans which are in use in the dialects though not in 'Standard' Uyghur.

    Concerning the contact languages I used the dictionaries and word collections of Steingass 1892 and Lambton 1961 for Persian, Hava 1982, Steingass 1972 and Wehr 1979 for Arabic, Ramstedt 1935, Lessing 1960, Doerfer 1963–1975, Mostaert 1941–44 and Nugteren 2011 for Mongolic, Wheeler 1972 for Russian and A. Äbäy et al. 2008 for Chinese; some Chinese examples were also extracted from internet websites.

1.5 Abbreviations and transcriptions

1.5.1 Linguistic glossing

1SG	1st person singular	CAUS	causative
1PL	1st person plural	COM	comparative
2SG	2st person singular	COND	conditional
2PL	2st person plural	CONJ	conjunction
3SG	3st person singular	COP	copula
3PL	3st person plural	CS	Code Switching
ABL	ablative	CVB	converb
ABSTR	abstract	DAT	dative
ACC	accusative	DER	derivational element
ACT	actionality	DISTR	distributive
AOR	aorist	EVD	evidentiality
C	consonant	EXCL	exclamation

FUT	future	POSS	possessive
GEN	genitive	PRES	present
IMP	imperative	PROG	progressive
IMPF	imperfect	PRTC	particle
INT	intentional	PST	past
LOC	locative	PURP.CVB	purpose converb
MOD	modality	Q	interrogative particle
NEG	negation	SG	singular
ORD	ordinal number suffix	SIM	simulative
PART	participle	V	vowel
PASS	passive	VN	verbal noun
PF	perfect	VOL	voluntative
PL	plural	VP	verb phrase
PN	proper name		

1.5.2 Languages and sources

Ar.	Arabic	Russ.	Russian
CC	Codex Cumanicus	SH	Secret History
Chag.	Chaghatay	S. Uyg.	Spoken Uyghur
Chin.	Chinese (Standard Mandarin)	Soy	Soyot
Hak	Hakas	Tel	Teleut
M. Chin.	Standard Mandarin Chinese (only used when opposed to X. Chin.; = Chin.)		
Mo.	Mongolian	Uyg.	Uyghur
Muq	Muqaddimatu 'l-Adab	W. Uyg.	Written Uyghur
Per.	Persian	WM	Written Mongolian
Kalm.	Kalmuck	X. Chin.	Xinjiang Chinese

1.5.3 Transcriptions for specific languages

For Uyghur sounds the following transcriptions are applied (the alphabetical order follows that of the Uyghur alphabet):

a	[a]	s	[s]	o	[o]
ä	[æ]	š	[ʃ]	u	[u]
b	[b]	ɣ	[ɣ, ʁ]	ö	[ø]
p	[pʰ]	f	[f]	ü	[y]
t	[tʰ]	q	[qʰ]	w	[w, v]
j	[dʒ]	k	[k]	e	[e]
č	[tʃ]	g	[g]	i	[i, ɯ]
x	[x]	ŋ	[ŋ]	y	[j]
d	[d]	l	[l]		
r	[r]	m	[m]		
z	[z]	n	[n]		
ž	[ʒ]	h	[h]		

For transliterating the Persian written forms, I principally followed Steingass 1875 with the following small changes:

ᶜ : [ʕ]   ṭ : [tˤ]   ẓ : [zˤ]   th : [θ]   ḍ : [dˤ]   dh : [ð]

The above transliteration also holds for Arabic, except that, when transliterating Arabic, I wrote *x* for [x]; for Persian I transliterate this letter as *kh* [x], as Steingass does.

Arabic and Persian transliteration:

Arabic-Persian alphabet	Arabic	Persian
ا	ʾ	ʾ
ب	b	b
پ	-	p
ت	t	t
ث	th	th
ج	j	j
چ	-	ch
ح	ḥ	ḥ
خ	x	kh
د	d	d
ذ	dh	dh
ر	r	r
ز	z	z
ژ	-	zh
ط	ṭ	ṭ
ظ	ẓ	ẓ
ك	k	k
گ	-	g
ل	l	l
م	m	m
ن	n	n
ص	ṣ	ṣ
ض	ḍ	ḍ
ع	ʿ	ʿ
غ	gh	gh
ف	f	f
ق	q	q
س	s	s
ش	sh	sh