Loan Adaptation in Zazaki

Mary Ann Walter
Massachusetts Institute of Technology

The lexicon of Zazaki has been heavily influenced by borrowing from other Middle Eastern languages. In this study I examine the repair of interdental segments and illicit consonant clusters as they are adapted online by a Zazaki speaker. Another corpus of loans from Turkish, the language that currently influences Zazaki most, is also discussed. Particular attention is paid to the treatment of Turkish vowels. I argue that a native allophonic process involving one of them facilitates the adoption of that one as a segment when borrowed, while a comparable one without allophonic status must be repaired.

1. Introduction

Zazaki is spoken at the nexus of three major linguistic areas. The Arabic, Turkic, and Indo-Iranian languages that surround it have permeated the vocabulary and the grammatical system of Zazaki over centuries of coexistence. Other regional languages with smaller footprints, such as Armenian, the Neo-Aramaic languages, and dialects of Kurdish, have also exerted an influence. With such a long shared history, a sprachbund has developed in which these languages share a great many phonemes and grammatical characteristics (not to mention lexical items), despite their membership in quite disparate language families. The following diagram shows some of the relationships between linguistic communities, and exemplifies the difficulty in untangling them (adapted from Kahn (1976)).

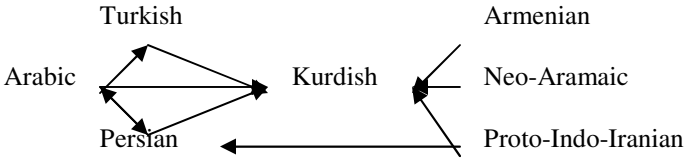

Figure 1: Linguistic Influences on Zazaki

Even the above picture is vastly oversimplified, as each category is an umbrella term. Thus "Arabic" includes the Classical, Levantine, and Iraqi/Eastern dialects, Turkish both Standard and Azeri, *et cetera*. Moreover, even an established Indo-Iranian etymology does not guarantee that a given form is native to Zazaki. Loans from Persian, in particular, are extremely common (Hasanpoor 1999).

It is clear that determining the origin of any given borrowing is quite difficult, as is tracing the path of transmission. For example, Zazaki is saturated with borrowings ultimately from Arabic, including much of the vocabulary relating to religion, civic institutions, and "high" culture. However, this is true of both Persian and Turkish as well. Consider the table given below, which details the origins of the word stems on a randomly-chosen page from Zazaki and Turkish dictionaries (Turgut 2001, Hony 1947).

Language	*Arabic*	*Turkish*	*European*	*Total*
Zazaki	7	3	0	22
Turkish	7	n/a	2	22

Table 1: Origins of Zazaki and Turkish vocabulary

The numbers involved are small, but based on this mini-sample it appears that roughly half the vocabulary is foreign. And while the source language is primarily Arabic, the presence of an equivalent number of Arabic loans in another language that neighbors and has had a strong influence on Zazaki means that their immediate provenance is unclear. A principled discussion of a form's adaptation depends upon accurate knowledge of the phonetic and phonological properties of the source word. Yet as the dictionary data demonstrates, one cannot be sure of the language via which a loanword passed into Zazaki, and therefore this knowledge is not at our disposal. While the path of transmission may sometimes be deduced by the form of the borrowing (those where Arabic /q/ corresponds to /k/, for example, did not come via Persian, for example, which preserves the segment, and is therefore probably from Turkish; similarly, a loan containing the glide /w/ is probably directly from Arabic, as both Persian and Turkish borrow the segment as /v/), this is not always the case.

To avoid such etymological confounds, I will primarily consider only two subgroups of Zazaki loanwords in this study. The first is a set of forms elicited via online adaptation from a native Zazaki speaker.

The second subgroup is from a recently published dictionary of Zazaki (Turgut 2001; the first and only one in existence so far as I am aware). Words which corresponded to Turkish forms were collected into a corpus (these were easily identifiable, as the dictionary is from Zazaki into Turkish). All those with Persian and/or Arabic counterparts were automatically eliminated. This left a set of forms which must be either Turkish forms borrowed into Zazaki, or vice versa. The semantic considerations detailed above as well as the relative status of the two languages militate for the former (Zazaki was banned from use until very recently, and all education and public life is carried out in Turkish). While the course's linguistic consultant criticized the dictionary for being too influenced by Turkish, this makes it ideal for the purpose of identifying and evaluating Turkish loans into Zazaki.

2. Online adaptation

Because of the sprachbund situation that obtains around the Zazaki region, disparities of phonotactics and phonological inventory have been generally leveled. Not many are present to be repaired over the course of adaptation from

a neighboring language into Zazaki (for more information about the native Zazaki system, see Adler & Kenstowicz, this volume). Thus stimuli were designed to elicit repairs of non-native segments, as well as of phonotactically ill-formed consonant clusters. Those described in 2.1 were elicited by the investigator producing a form orally and asking the consultant to repeat it back as it would sound in Zazaki. Those described in 2.2 were elicited by presenting the consultant with a form written in English orthography and asking her to say it out loud as it would be in Zazaki.

2.1 Segmental repairs

The inventory of Zazaki consonants is given below in Figure 2.

p p^h b	t t^h d		k k^h g	q
m	n		N	
f v	s z	Σ Z	x	
		tΣ dZ		
	l r P			
w	j			

Figure 2: Zazaki consonant inventory

As may be seen, the Zazaki surpasses English in its inclusion of a phonemic aspirated stop series, in addition to voiced unaspirated and unvoiced unaspirated ones. English stop consonants map straightforwardly onto this system, with allophonic aspiration status corresponding to the Zazaki phoneme with equivalent glottal specifications. Thus English voiced stops correspond to Zazaki voiced stops, English voiceless aspirates to Zazaki voiceless aspirates, and English voiceless unaspirated stops to Zazaki voiceless unaspirated ones. Examples of each correspondence are given in Table 2 below.

Zazaki lacks, however, the interdental consonants /T/ and /Δ/. The former is invariably borrowed as aspirated /t/. The latter typically becomes /d/. This is always the case when the segment is hyperarticulated (by the English speaker) and after repeated exposures. However, in some cases after single tokens it may surface as /v/ or /l/, according to the schema detailed in the following table:

Context	*Input*	*Output*
__ C [labial]	luΔbi	luvbi
	liΔbi	livbi
/i/ ___ C [front]	siΔtar	silta≅r
	p^hiΔkin	p^hilk^ji≅n

Table 2: Interdentals

A surprising instance of a *lack* of adaptation occurs with respect to the coda-devoicing rule active in the native Zazaki vocabulary. Words like English *sibling* and *dovetail* include a medial cluster of a voiced consonant followed by a voiceless one.. Complex onsets are not permitted in Zazaki, so the first consonant must be syllabified in a coda, and therefore subject to the devoicing

rule. Alternatively, vowel epenthesis between the two would lead to resyllabification and the possible preservation of the source's voicing specification (epenthesis is quite common in loan adaptation, as will be shown below).

However, our consultant chose neither of these options, and instead chose to preserve coda voicing without epenthesis, thereby violating the native phonotactics.

Finally, an outstanding issue in segmental adaptation is the treatment of English-like velar /k/. In addition to the velar series, the Zazaki inventory contains a velar series with the three members specified above (/g/, /k/, /k^h/). The velars are also subject to a palatalization rule conditioned by the following vowel. Finally, Zazaki also has two uvular consonants, the stop /q/ and fricative /®/.

English, on the other hand, is relatively impoverished with only /k/. It is subject to considerable fronting when preceding a front, non-low vowel. The lack of any *phonemic* contrast in place of articulation behind the palate, however, means that it is quite variable in its point of contact with the surface of the oral cavity (Stevens 1997). This variability means that English /k/ ranges over a greater acoustic space than Zazaki /k/. Particularly when produced before a back vowel, it may cross into the perceptual space of Zazaki /q/. Indeed, when English /k/ is borrowed into Zazaki, it appears as either /k/ or /q/ (very rarely, /g/).

The alternation is not obviously governed by the quality of the following vowel, however. The resulting Zazaki /q/ may appear not only before back vowels, but also before /i/, for example. Recall that in Zazaki such front, non-low vowels condition an allophonic palatalization contrast in preceding velars. Despite some degree of fronting, the coarticulatory effect of /i/ on English /k/ is comparatively small. Thus the consonant in an English /ki/ sequence would not be perceived by a Zazaki speaker as palatalized, and therefore not as an underlying /k/. It is then borrowed as /q/ instead.

It may be the case that the aspiration of the original segment has a bearing. Aspirated /k/ before back vowels tends to be borrowed as /q/, but before front vowels as aspirated and palatalized /k/. As uvulars generally have a longer voice-onset time than velars, English aspirated /k/ would correspond to /q/, and unaspirated to /k/. The allophonic aspiration of English /k/, then, would be less than that of the phonemic Zazaki phoneme, and more comparable to the VOT of the uvular. The /k/ before a front non-low vowel, however, could not be palatalized were it borrowed as /q/, and is therefore kept as aspirated /k/. The aspiration generalization is not without other exceptions – the second segment of initial /sk/ clusters, in particular, is often borrowed as /q/ despite its lack of aspiration in English.

In the event, the adaptation pattern of this segment is complex, and bears further investigation.

2.2 Phonotactic repairs

Like other languages in the region, Zazaki allows only simple onsets of only one segment, and complex codas of up to two consonants. Also like its neighbors, it has generally final stress (exceptions include when suffixed with the diminutive

ɪkE, which does not bear stress, and other final-vowel morphemes). As established borrowings, then, are unlikely to yield information concerning cluster repair, a set of words including them were presented to the language consultant for online adaptation. This set is an augmented version of the one used in Shademan (2002) for Tehrani Persian, a closely related language.

The resulting pattern of epenthesis closely matches that observed by Shademan (2002). As in Persian, the epenthetic vowel is inserted between the first and second consonant of a cluster. The vowel inserted is not the same in the two languages– in Zazaki it is the short, mid vowel *ɪ*, versus Persian *E*. But it is nonetheless the shortest vowel in the Zazaki inventory, as expected, and as Table 3 demonstrates. Duration values are averaged over four tokens per vowel.

Vowel	ɪ	ɪɪ, i, u, uu	e, a, aa, o, oo	ee
Duration (ms)	44	60-73	92-117	138

Table 3: Vowel duration

This is as expected according to the generalizations regarding epenthesis made by Lombardi (2002).

Also as in Tehrani Persian, there is vowel-harmonic epenthesis in some contexts, in addition to epenthesis of the default vowel *ɪ*. But while Persian permits the vowels /u/, /i/, and /o/ to be copied, Zazaki allows only /u/. This is always possible, and possibly required, in a sequence such as labial-sonorant-u (such as *pru*, *flu*, etc.). If the following vowel is /o/ rather than /u/, the epenthetic one can optionally be /u/ when the sonorant is /r/, but not when it is /l/. The potential for vowel-harmonic /u/ to surface epenthetically is variable when the next underlying vowel is /u/ but the first consonant is not labial. Variability also exists when the first consonant is not labial, but the other consonant and the vowel both are. It may be that only a palatal disallows /u/ in this case. These options are summarized in Table 4 below.

Context	*Input*	*Output*
+lab __ +son /u/	myut flut	muyut fulut
+lab __ +son /o/	brochure Freud Florida	bɪroΣir, buroΣir fɪrojd, *furojd fɪlorida, *fulorida
-lab __ +son /u/	drup zrunski	dɪrup, *durup zurunski
-lab __ +lab /u/	kvuski rmuspi Zvunti	kuvuski rumuspi Zɪvunti, *Zuvunti

Table 4: Initial cluster repair

Shademan (2002) finds that vowel copy may occur across the segments /r/, /l/, and the nasals. The segments are listed in descending order of duration, and are subject to increasingly stringent conditions respectively. While no durational data is available for Zazaki consonants, it is striking that roughly the same set

(sonorants) is susceptible to being copied across (with the addition of labial harmony that extends to obstruents).

In addition to the words tested by Shademan (2002), which all contain word-initial clusters, a supplementary word set was presented to the consultant with final clusters of three consonants. In final position, the situation is somewhat different. As before, the vowel is inserted before the first and second consonant of the cluster. But recall that Zazaki has generally final stress, (invariably for consonant-final words). It follows that the vowel inserted to break up a final cluster will always be stressed. This is unproblematic for the vowel *ɪ*, which bears stress in the diminutive suffix given above. That stress on epenthetic vowels is also permissible in Zazaki is independently verifiable. Consider the following form in Table 5. It results from the application of nickname truncation, in which the first maximal syllable of the name is preserved.

English	*Zazaki*	*Truncatum*
brΘd	bırEd	bır

Table 5: Nickname truncation

Despite the ability of epenthetic vowels to bear stress, however, the presence of stress has an effect on the choice of vowel to insert. Avoidance of stressing the (epenthetic) vowel *ɪ* seems the most likely explanation for the following alternation. While the default vowel *ɪ* is always an option, it alternates with others according to two conditions. When the epenthetic vowel is preceded by a labial consonant or the uvular /q/, it may become /o/. Otherwise, it may become /i/. This alternation is exemplified below in Table 6.

Input	*Output*
saklt	saqılt, saqolt
rutpk	rutpık, rutpok
kitnk	kitnık, kitnik
fepnt	fepnıt, fepnit
sitkp	sıtkıp, sıtkip
siktp	sıktıp, sıktip

Table 6: Final cluster repair

Note that a discrepancy exists between the back vowel chosen for the alternation word-initially (/u/) versus word-finally (/o/). The former is the more obvious choice, as it falls next on the durational continuum. That the latter ever appears epenthetically is puzzling.

3. Turkish loanwords

The linguistic situation of Zazaki is such that bilingualism is near-universal and has been for most of the last century. Due to it being a prohibited language until very recently, Turkish is the language of all public life. All education is conducted in Turkish, so that Zazaki speakers are subject to intense exposure to the language from the age of 5 years at the latest. Thus it is not surprising that

Turkish is the source, or at least the point of entry, of most recent loans. As stated above, however, the long-standing nature of the two languages' proximity means that not much is necessary in the way of adaptation.

One feature of Turkish that Zazaki does not display is phonemic gemination. In such cases, Zazaki straightforwardly degeminates, without any compensatory measures, as shown in Table 7 below.

Turkish	*Zazaki*	*Gloss*
dZEnnE≅t	dZEnE≅t	heaven (>Arabic)
dZEhEnnE≅m	dZEnE≅mE	hell (> Arabic)

Table 7: Degemination

Zazaki and Turkish also differ with respect to vowel inventory. In addition to the phonemes of Zazaki, Turkish also includes the high front rounded vowel /u_/, the mid front rounded vowel /o_/, and the high back unrounded vowel /∝/.

The third we will not consider here. The second is adapted as the vowel /o/, as in the following two cases:

Turkish	*Zazaki*	*Gloss*
o_rdE≅kE	ordE≅gE	duck (bird)
o_rnE≅k	orna≅k	sample, pattern

Table 8: /o_/ adaptation

This vowel shows up in the native system only in the verb *to sleep*, in which it follows a palatal aspirated /k/. One such form is *k^{h}jo_tbira* '*was sleeping*.' As palatalized velars seem to be conditioned by the following vowel, it appears that this must be a genuine, if unique, instance of /o_/, in order to trigger the change.

Adaptation in the case of the remaining vowel presents a far more complex case. While not phonemically present in Zazaki, /u_/ does surface as a fronted alternant of /u/ when following a coronal consonant. Examples follow in Table 8.

Zazaki	*Gloss*
Σu_anE≅	shepherd
Σu_≅ΣE	bottle
lu_lɪkE	pupil (of the eye)
Zu_	one
su≅kE	city

Table 8: /u/ fronting

Such fronting applies only variably (as shown by the last form above, an exception) but quite frequently. Moreover, it is fed by a process of word-internal hiatus-resolving raising. The following two examples demonstrate first the raising process, and then a case of it feeding fronting.

Input	*Output*	*Gloss*
dEwE≅ + E	dewi≅E	camel (fem.)
t^ho + o	t^hu_yo	you and

Table 9: Raising and fronting

The forms in the next table were borrowed into Zazaki via Turkish (originally from Arabic and French, respectively).

Zazaki	*Gloss*
du_nya≅, dinya≅, *dunya≅	world
u_nvE≅rstE	university

Table 10: Fronted borrowings

For both forms, the first vowel of the Turkish source form is /u_/. The first variant of the first borrowing, then, appears to be a straightforward instance of the vowel being borrowed as /u/, followed by application of the regular Zazaki fronting rule. That the /u_/ might also be borrowed as /i/, as in the second variant, is also unsurprising. Given the variability of fronting, however, it is mysterious why the third, starred variant is not a possible Zazaki rendition of the Turkish loan.

This particular form is quite an old borrowing. In fact, it is generally the case that Turkish /u_/ is no longer borrowed as /i/. Rather, it is kept as /u_/. And as with *world*, it is generally the case that such an instance of /u_/ cannot alternate with /u/, even though it generally would otherwise. Presumably this reflects the knowledge of the bilingual borrower that such an alternation does not occur in Turkish, from which the word is borrowed.

Moreover, in the second example front /u_/ appears without following a coronal consonant – an environment where it does not appear in the native system. Nevertheless, Zazaki speakers maintain the Turkish front vowel in their own language in this environment as well. Thus they have incorporated a borrowed segment in their inventory.

If this is possible for the high vowel, however, why does it not happened with the mid vowel /o_/? As we saw above, this vowel is always adapted to /o/ by Zazaki speakers. There are several ways in which this vowel differs in Zazaki from /u/ other than featurally. Recall from the duration data that it lasts approximately 50% longer than the high vowel. The fronting effect of a preceding coronal, therefore, may be hypothesized to affect a much larger proportion of the high vowel than the mid vowel.

To verify this discrepancy, recordings were made of the language consultant producing three tokens each of the following four syllables: *sho, shu, ho, hu*. Tokens were elicited with the presentation of the words visually to the subject. F2 values were then taken at 15 ms intervals for each token. Figure 2 shows the progression of F2 during the course of the vowels. Each line represents the averaged values of the three tokens.

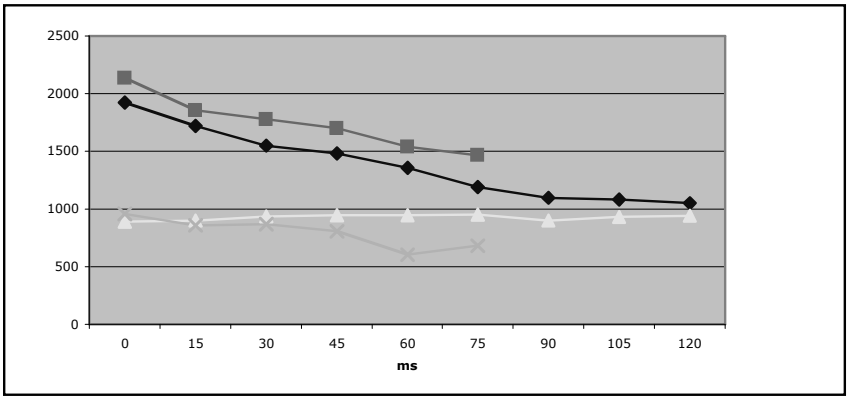

Figure 2: F2 in /u/ and /o/, post-coronal and noncoronal

First of all, the F2 values for *shu* are consistently higher than those for *sho*. This is contrary to expectation based on their featural specifications, but accords with what we expect to see from the influence of the preceding consonant.

Secondly, it is clear that at the endpoint of the average range for duration of /u/, 73 ms, F2 in *shu* is still on a downward trajectory that has not yet reached the value of the steady state in *hu*. At the end of the range for /o/, on the other hand, the F2 vowel following the coronal approximates that of its counterpart quite closely. For the /o/ vowel, then, the canonical F2 value is at least reached (even if it does not extend over much of the course of the segment).

Given the above, it is not surprising that a conditioned allophonic contrast might be established for the high vowel, but not the mid one. The coarticulatory effect of a preceding coronal, which is comparable for the two Zazaki vowels /o/ and /u/, results in distinct sorts of representations. The status of /u_/ as a conditioned allophone in Zazaki and the presence of a distinct representation for it, then, provides a beachhead for the segment in the language. When borrowed, it then becomes independent of the environmental restrictions placed upon it in Zazaki. For this to occur, Zazaki speakers must be aware of and able to perceive the difference.[1]

Bilingualism, finally, informs their behavior with respect to both languages, Turkish and Zazaki. The presence of the segment /u_/ in Turkish, and the resulting high degree of exposure of Zazaki speakers to it, may influence the allophony in Zazaki and speakers' sensitivity to its two allophones. This allophony, on the other hand, facilitates the borrowing of the segment as an independent segment into Zazaki words – as does not occur for the otherwise similar segment /o_/ -- in situations where the allophonic alternation does not apply.

[1] This has also been argued with respect to the English allophonic contrast in aspiration by Jones (2001), contra Whalen et al. (1997).

4. Conclusion

Due to its long history of coexistence with other speech communities, Zazaki displays a wealth of borrowings. Currently it appears to be in the process of incorporating a new tier of lexical items (in the sense proposed by Ito & Mester (1998)), with somewhat different phonotactics than the lexicon has had heretofore. Two classes of such items are online/English-like borrowings with voiced coda consonants, and Turkish borrowings involving the vowel /u_/. The latter case is an especially interesting example of the nature of language change and the role of bilingualism in it. Clearly Zazaki is continuing its tradition of lexical cosmopolitanism, and the language continues to change in interesting ways because of it.

References

Hasanpoor, J. (1999). *A Study of European, Persian and Arabic Loans in Standard Sorani*. Ph.D diss, Uppsala University.

Hony, H.C. (1947). *A Turkish-English dictionary*. Oxford: Clarendon.

Ito, J. & A. Mester. (1999). The Phonological Lexicon. *Handbook of Japanese Linguistics*. Ed. by N. Tsujimura. Malden: Blackwell.

Jones, A. (2001). *A Lexicon-Independent Phonological Well-Formedness Effect: Listeners' Sensitivity to Inappropriate Aspiration in Initial /st/ Clusters*. MA thesis, UCLA.

Kahn, M. (1976). *Borrowing and Variation in a Phonological Description of Kurdish*. Ph.D diss, University of Michigan.

Lombardi, L. (2002) Markedness and the typology of epenthetic vowels. *Proceedings of Linguistics and Phonetics.*

Shademan, S. (2002). *Epenthetic Vowel Harmony in Farsi*. MA thesis, UCLA.

Stevens, K. (1998). *Acoustic Phonetics*. MIT: Cambridge.

Turgut, H. (2001). *Zazaca-Türkçe sözlük*. Istanbul: Tij Yayınları.

Whalen, D. H., C.T. Best, and J. Irwin (1997). Lexical effects in the perception and production of American English /p/ allophones. *Journal of Phonetics 25,* 501-528.

Department of Linguistics and Philosophy
E39-245 MIT
77 Massachusetts Avenue
Cambridge, MA 02139
USA

walterma@mit.edu