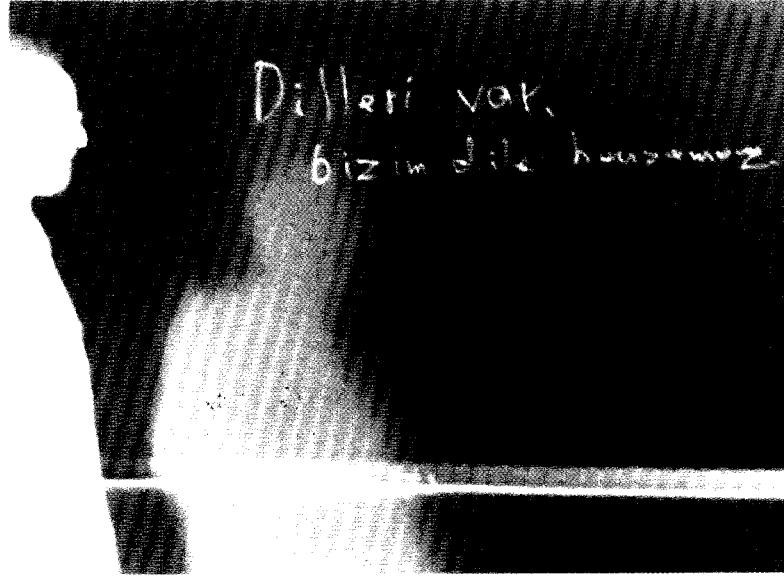


TURCOLOGICA 86

## Puzzles of Language

Essays in Honour of Karl Zimmer



Edited by  
Eser Erguvanlı Taylan and Bengisu Rona

Harrassowitz Verlag

# TURCOLOGICA

Herausgegeben von Lars Johanson

Band 86

2011

Harrassowitz Verlag · Wiesbaden

---

# Puzzles of Language

Essays in Honour of Karl Zimmer

Edited by

Eser Erguvanlı Taylan and Bengisu Rona

2011

Harrassowitz Verlag · Wiesbaden

---

# Why don't they meet face to face? On hiatus-preventing allomorphy in Turkish and its relatives

Lars Johanson

## 1 Introduction

Turkish and most of its Turkic relatives are known for their regular and transparent morphology. One of the unsolved problems relating to this domain is why vowels cannot occur next to each other across morpheme boundaries. Turkic disallows hiatus between two adjacent morphemes, i.e. sequences of a stem-final vowel and an initial vowel of a following bound marker. Allomorphy, alternation in complementary distribution, regulates the phonotactics. The shape of numerous bound markers, i.e. suffixes and clitics, varies according to whether the stem ends in a vowel or in a consonant. Vowel-final stems occur with consonant-initial markers, and consonant-final stems occur with vowel-initial markers. Thus the vowels do not come into adjacency, do not meet "face to face".

These variations, observable in inflection and derivation, are well-known. In synchronic descriptions, e.g. for didactic purposes, it is legitimate to deal with them as formal operations of the following kinds:

- (i) A vowel is dropped after stem-final vowels, and retained after stem-final consonants.
- (ii) A vowel is inserted before stem-final consonants, but not before stem-final vowels.

These operations are strictly synchronic. On the other hand, various assumptions about hiatus-preventing diachronic processes are found in the literature. One claim is based on the concept of "connective" vowels, which are thought to have been inserted between stem-final consonants and consonant-initial markers. Another claim concerns hiatus-preemptory "bridging" or "buffer" consonants, which are thought to have been inserted between stem-final vowels and vowel-initial markers.

What is the diachronic background of the observable synchronic variation? What circumstances gave rise to these kinds of allomorphy at actual real-time points in history? Hiatus resolution by means of the operations just mentioned seems diachronically implausible. Have there ever been any vowel+vowel junctures to resolve in the history of Turkic? There are at least no traces left of earlier developmental stages of the agglutinative morphology at which the phonotactic system would have allowed two adjacent vowels to be pronounced separately, e.g. cases of the type \**ya:ša-ar* ['to live'+present] 'lives'.

Suffix allomorphy in today's Turkish and its relatives relies on earlier morphological structures that have changed mainly through consonantal deletion and subsequent contractions including vowel assimilations. Fusion of various kinds has taken place at morpheme boundaries. The only illicit constellation is that vowels in two adjacent

---

morphemes are not allowed to occur next to each other, to come to stand “face to face”. The reconstructable history of Turkic phonotactics shows that they only have had two options:

- (i) to keep their distance, or
- (ii) to merge.

A possible exception will be mentioned in section 4.

## 2 Transcription and symbols

In what follows, Turkic data will in general be presented in a unified broad phonetic transcription. Turkish data will, however, be rendered in the official orthography.

Hyphens are used to indicate morpheme boundaries. The sign < means ‘has developed from’, and the sign > means ‘has developed into’. An arrow of the type ← means ‘is derived from’. Asterisks (\*) are used for unattested, reconstructed items. In descriptive contexts, they mark unacceptable forms and constructions.

Capital letters stand for morphophonemes, e.g. alternation due to assimilation processes such as sound harmony. *A* indicates an alternation of non-high vowels, and *I* an alternation of high vowels. If necessary, the number of alternants is indicated by raised numerical symbols, e.g.:

$A^2$  = alternation of non-high unrounded vowels (*a*, *ε*).

$I^2$  = alternation of high unrounded vowels (*i*, *ɨ*).

$I^4$  = alternation of high nonrounded vowels (*i*, *ɨ*, *ü*, *u*).

$U^2$  = alternation of high rounded vowels (*ü*, *u*).

$\tilde{I}^2$  = alternation of near-high lax vowels (*ĩ*, *ĩ̃*).

*C* indicates a consonant.

*V* indicates a vowel.

Vowel and consonant length is marked with a colon (:).

Bracketed elements occur in certain environments and are missing in others.

A bracketed initial consonant sign indicates the occurrence of the consonant after stem-final vowels and its absence after stem-final consonants, e.g. dative markers of the Turkish type  $-(y)A^2 = -a, -e, -ya, -ye$ .

A bracketed initial vowel sign indicates the occurrence of the vowel after stem-final consonants and its absence after stem-final vowels, e.g. possessive markers of the Turkish type  $-(I^4)m$ , i.e. *baba-m* ‘my father’, *gül-üm* ‘my rose’.

A bracketed final consonant sign indicates the occurrence of the consonant in front of case suffixes and its absence in other environments, e.g. third singular possessive markers of the Turkish type  $-(s)I^4(n)$ .

A bracketed initial zero sign indicates that the final vowel of the preceding stem drops when the marker is added, e.g. the Turkish present marker  $-(\emptyset)I^4yor$  or the Chuvash present marker  $-(\emptyset)A^2t$ .

A colon indicates that the addition of the marker to a stem-final vowel yields length of this vowel or of the last consonant of the stem. The Turkmen participle marker  $-A^4n/:n$  materializes as *gel-en* ‘having come’ ← *gel-* ‘to come’, *oķa-:n* ‘having read’ ← *oķa-* ‘to read’, etc.



## 4 Vowel hiatus in standard Turkish

A specific problem concerns a recent and limited sound change in one West Oghuz variety, namely modern standard Turkish, which is based on West Anatolian dialects. Here, a stem-final back velar stop *k* (and sometimes its front counterpart) changes to  $\emptyset$  in front of vowel-initial suffixes. This phenomenon is often called “the *k/ğ* alternation”, since, in the official orthography, the letter «*k*» is replaced by «*ğ*» before vowels. This «*ğ*» is, however, mostly not pronounced, which may give rise to hiatus-like structures.

Some non-standard Turkish varieties show forms with a preserved velar fricative  $\gamma$  in this position, e.g. *kuyruγ-u* ‘its tail’ ← *kuyruk* ‘tail’, *ayay-īm* ‘my foot’ ← *ayak*, *yatay-īm* ‘my bed’ ← *yataκ* ‘bed’. In standard Turkish, however, no consonant segment is pronounced in this position. Written word-forms such as *yatağ-ım* ‘my bed’ or *ayağ-ım* ‘my foot’ may be pronounced as [jata:m] and [aja:m]. This is the hiatus-preventing technique we find in South Siberian and some other Turkic languages, where the stem-final vowel merges with the initial suffix vowel into a long vowel, e.g. Tuvan *aya:m* ‘my foot’.

Pronunciations such as [jata:im] ~ [jataim] ‘my bed’ and [aja:im] ~ [ajaim] ‘my foot’ are, however, also observed in the speech of native speakers of standard Turkish. This articulation seems to represent hiatus as a result of consonant deletion and thus to be a counterexample of the rule that Turkic disallows vowel hiatus between two adjacent morphemes. Or is there a consonantal element left? As a matter of fact, the entity represented by «*ğ*», the so-called “soft *g*”, is still phonologically of non-vocalic nature. This is shown by the fact that stems ending in this entity take on postconsonantal suffix variants, e.g. [da:i] *dağ-ı* ‘its mountain’ ← [da:] *dağ* ‘mountain’, but never \*[da:sı], which would be obligatory after stem-final vowels.

## 5 Allomorphy of the type -(C)V

Another type of allomorphy is  $-(C)V$ , realized as  $-V$  after stem-final consonants and as  $-CV$  after stem-final vowels. It is claimed to be the most common type in Turkic inflectional morphology and is often interpreted as a process of insertion of a hiatus-bridging suffix-initial consonant, which is dropped in postconsonantal position. Operations of this kind are, however, implausible as actual diachronic processes.

### 5.1 $-(s)V$ and $-(š)V$

The bracketed suffix-initial consonants *s* and *š* in the patterns  $-(s)V$  and  $-(š)V$  are lacking after stem-final consonants. Though they are of unknown origin, they have certainly not emerged as inserted hiatus-bridging elements.

The third person possessive suffix has the shape  $-(s)I(n)$  in almost all old and modern Turkic languages, e.g. Turkish *kapı-sı* ‘its door’ ← *kapı* ‘door’.

The distributive numeral suffix is mostly found in the shape  $-(š)Ar$ , e.g. Turkish *dörd-er* ‘four each, four at a time’ ← *dört* ‘four’, *iki-şer* ‘two each, two at a time’ ← *iki* ‘two’. It has its own idiosyncratic historical developments. Thus, Karachay-Balkar uses  $-šAr$  also after numerals ending in consonants (except *bir* ‘one’). Chuvash uses  $-šAr$  much in the same

way, e.g. *tıvat-şar* 'four each', *ik-şer* 'two each', derived from *ik*, the 'clipped' form of *ikkı* ~ *iki* 'two'. The original shape of the marker is unknown.

## 5.2 -(n)V

The bracketed suffix-initial consonant *n* in *-(n)V* is absent after stem-final consonants. It is not likely to have emerged as an inserted hiatus-bridging element.

### 5.2.1 Genitive markers *-(n)I* and *-(n)In*

The oldest East Old Turkic genitive marker *-(n)İŋ* exhibits the alternation *-İŋ* versus *-nİŋ*. The corresponding genitive suffix in modern West Oghuz is *-(n)I<sup>h</sup>n*, where the bracketed *n* occurs after stem-final vowels and is absent after stem-final consonants, e.g. Azeri *at-in* ← *at* 'horse', *ata-nin* ← *ata* 'father', Turkish *taş-in* ← *taş* 'stone', *kapı-nın* ← *kapı* 'door'; cf. Ottoman Turkish *-(n)U<sup>h</sup>ŋ*.

The suffix-initial *n* has certainly not been inserted as a hiatus breaker that is deleted after stem-final consonants. It is a remnant of forms of the pronominal declension, which early influenced the nominal case markers. The oblique forms of the East Old Turkic pronoun *ol* 'this' derive from the stem *an-*, the genitive form being *aniŋ*. The nominal genitive marker may well go back to structures of the type *\*at aniŋ* and *\*ata aniŋ*.

The stem-final pronominal *n* is an archaic element that has also left other traces. In the nominal declension of most modern Turkic languages, a so-called pronominal *n* occurs in third person possessive suffixes before locative and ablative case markers, e.g. Turkish *kapı-sın-da* <door-3SG-LOC> 'at its door' or Chuvash *xul-in-çe* <town-3SG-LOC> 'in its town'. It has counterparts in Mongolic, Tungusic, and Japanese, which is an interesting fact for genealogical studies.

Turkmen, which is, as already noted, highly conservative with respect to vowel length, may give us a clue about the origins of the genitive marker discussed here. The suffix is realized as *-I<sup>h</sup>ŋ* after stem-final consonants, e.g. *at-in* ← *at* 'horse', *öy-ün* ← *öy* 'house', *diş-in* ← *diş* 'tooth'. After stem-final vowels we might expect *-nI<sup>h</sup>ŋ*, but here the stem-final vowels fuse with the suffix-initial vowels, yielding vowel lengthening, e.g. *gapi:nin* ← *gapi* 'door', *köçe:nin* ← *köçe* 'street', *ejε:nin* ← *ejε* 'mother', *kino:nin* ← *kino* 'movie'. The Turkmen genitive marker is thus *-I<sup>h</sup>ŋ/-:nI<sup>h</sup>ŋ*. Spoken Turkmen prefers the postvocalic form *-ŋ*, but the vowel length is preserved here as well.

The postvocalic forms point to contractions. The genitive suffix may ultimately go back to a combination with *\*aniŋ*, which may have eroded through contraction, though it is still reflected by vowel length after stem-final vowels. It has left no traces after stem-final consonants.

The corresponding Chuvash genitive marker is *-I<sup>h</sup>n* after stem-final consonants, e.g. *kil-ñn* ← *kil* 'house'. The allomorph *-nI<sup>h</sup>n* or *-n* occurs after stem-final *-A*, e.g. *laša-nñn* or *lašan* ← *laša* 'horse', *ine-nñn* or *ine-n* ← *ine* 'cow', *xula-nñn* ← *xula* 'city'. The allomorphy is not, however, adequately represented by the formula *-(n)I<sup>h</sup>n*. When *-I<sup>h</sup>n* is added to stems ending in *-İ*, the stem-final vowel is dropped, and the preceding consonant is almost always lengthened, e.g. *al:ñn* ← *alı* 'hand', *yit:ñn* ← *yiti* 'dog'. The formula for the Chuvash genitive marker is thus *-I<sup>h</sup>n/-nI<sup>h</sup>n ~ -n/-:I<sup>h</sup>n*. The consonant lengthening seems to be another



sign of contraction, in the sense of  $-C\check{I}^2+:\check{I}^2n>-C:\check{I}^2n$ . The distribution of the Chuvash genitive alternants displays further details, which are, however, not relevant for our topic.

Genitive markers of the type  $-(n)I\eta$  have been replaced by unified markers of the type  $-nI\eta$  in most Turkic languages. Already Old Uyghur texts display  $-nI\eta$  after stem-final consonants and vowels; Karakhanid sometimes shows  $-nIG$ . The shape is  $-nI\eta$  in most later languages, e.g. Chaghatay. This analogical extension produced a more regular pattern. The old allomorphy became a residual alternation that finally disappeared in more progressive languages.

In modern languages, the unified marker occurs in the Northwestern and Southeastern Turkic languages. Thus Crimean Tatar displays  $-nI^2\eta$ , e.g.  $at-ni\eta \leftarrow at$  'horse',  $ata-ni\eta \leftarrow ata$  'father', Tatar  $-nI^2\eta$ , e.g.  $at-ni\eta \leftarrow at$  'horse',  $eti-ni\eta \leftarrow eti$  'father', Noghay  $-D^3I\eta$ , e.g.  $bala-di\eta \leftarrow bala$  'child',  $terek-ti\eta \leftarrow terек$  'tree', and modern Uyghur  $-ni\eta$ , e.g.  $bali-ni\eta \leftarrow bala$  'child',  $köl-ni\eta \leftarrow köI$  'lake'.

### 5.2.2 Accusative marker $-(n)I$

Accusative markers of the type  $-(n)I$ , with morphological alternation  $-nI$  vs.  $-I$ , are found in some languages, e.g. Yakut  $eye-ni \leftarrow eye$  'peace',  $uot-u \leftarrow uot$  'fire'.

Turkmen shows  $-I^2$  after stem-final consonants, e.g.  $öy-i \leftarrow öy$  'house',  $at-i \leftarrow at$  'horse'. The postvocalic alternant is  $:-nI^2$  with vowel lengthening, e.g.  $köçe:ni \leftarrow köçe$  'street',  $eje:ni \leftarrow eje$  'mother',  $gapi:ni \leftarrow gapi$  'door'. Again, the lengthening seems to be a result of a contraction. The suffix may ultimately go back to a combination with the pronominal accusative form  $*ani$ . The latter has eroded, but is still reflected by vowel length after stem-final vowels, while it has left no traces after stem-final consonants.

The Chuvash dative-accusative marker is  $-A^2$  after stem-final consonants, e.g.  $kil-ε \leftarrow kil$  'house'. It is  $-nA^2$  after stem-final  $-A^2$ , e.g.  $laša-na \leftarrow laša$  'horse'.  $-A^2$  is also added to stems ending in  $-I$ , which drops, whereas the previous consonant is almost always lengthened, e.g.  $al:a \leftarrow alĩ$  'hand'. This lengthening, the same phenomenon as in the case of the genitive, is another sign of contraction, in the sense of  $-C\check{I}^2+-A > -C:A$ . The distribution of the alternants displays some further details without relevance for our topic.

As a parallel to the development of the old genitive marker, accusative markers of the type  $-(n)I$  have been replaced by unified markers of the type  $-nI$  in most Turkic languages. This analogical extension has produced a more regular pattern. The old allomorphy has finally disappeared in more progressive languages.

With the exception of Yakut, languages outside the Southwestern (Oghuz) branch exhibit suffix-initial  $-nI$  in all positions by analogy with the pronominal declension. Late East Old Turkic and Middle Turkic show  $at-ni$  instead of  $at-iy$ . Chaghatay has  $-nI^2$ ; cases of fusion of genitive and accusative suffixes are found early. Uzbek has the marker  $-ni$ , mostly coinciding with the genitive suffix. Crimean Tatar displays  $-nI^2$ , e.g.  $at-ni \leftarrow at$  'horse',  $ata-ni \leftarrow ata$  'father', Tatar  $-nI^2$ , e.g.  $at-ni \leftarrow at$  'horse',  $eti-ni \leftarrow eti$  'father', Noghay  $-D^2I^2$ , e.g.  $bala-di \leftarrow bala$  'child',  $tersk-ti \leftarrow terек$  'tree', modern Uyghur  $-ni$ , e.g.  $bali-ni \leftarrow bala$  'child',  $köl-ni \leftarrow köI$  'lake'. Postvocalic  $-nI$  is also found in many Southwestern varieties, e.g. Gagauz, Eastern Anatolian, Azeri and South Oghuz varieties (Songori, Aynallu, Khorasani, Afshar, etc.).

The unified accusative marker *-nI* is, however, not found in the conservative languages Ottoman Turkish, modern standard Turkish, and modern standard Azeri, all of which have retained the old accusative marker *-(y)I* (section 5.3.1).

### 5.3 *-(y)V*

One type of suffix morphemes begins with *-(y)V*, where the bracketed suffix-initial glide *y* occurs after stem-final vowels. Contrary to claims in the literature, this element is far from unstable, and it is not likely to have been originally inserted as a “buffer consonant”, a “hiatus-bridging” segment.

#### 5.3.1 *Accusative marker -(y)I*

An old type of accusative marker is *-(y)I*, e.g. Turkish *taş-t* ← *taş* ‘stone’, *oda-yı* ← *oda* ‘room’, Azeri *ev-i* ← *ev* ‘house’, *ata-yı* ← *ata* ‘father’. It goes back to *-(I)G*, e.g. East Old Turkic *at-ı̇y* ← *at* ‘horse’. *-(I)G* may also be the origin of the Khalaj accusative marker, whose postconsonantal form is *-i*, e.g. *hev-i* ← *hev* ‘house’. Its postvocal forms are *-y* ~ *-γ* ~ *-y(i)n*, e.g. *ba:ba-y* ← *ba:ba* ‘father’, which almost totally coincide with the genitive marker. It seems possible that the accusative and the genitive shared the same expression in Proto-Turkic.

#### 5.3.2 *Vowel-final converb markers*

Turkic languages possess a type of vowel-final converb markers. The choice of the vowel was originally the same as with the so-called aorist (section 5.3.3). The converb marker later developed differently from the aorist marker.

The East Old Turkic converb marker is *-V/-yU*. After stem-final consonants it appears as *-A* (normally with underived stems), *-U*, and *-I*, whereas the postvocalic allomorph is *-yU*. This situation is strikingly irregular compared to other suffix morphemes.

The Yakut marker is *-A/-I*. The postconsonantal allomorph is *-A*, e.g. *kül-ε* ← *kül-* ‘to laugh’, *üören-ε* ← *üören-* ‘to learn’. Stem-final vowels merge with the converb marker into *-I*, e.g. *üelēi* ← *üelē-* ‘to work’. South Siberian Turkic languages, e.g. Tuvan, show similar mergers of stem-final vowels and postvocalic markers. Similarly, the Chuvash marker *-A* merges with stem-final vowels; only two monosyllabic stems display a postvocalic marker *-yA*, namely *śi-yε* ‘eating’ ← *śi-* ‘to eat’, *tε-yε* ‘saying’ ← *tε-* ‘to say’.

In most other Turkic languages, the postconsonantal allomorphs have been unified to *-A*, whereas *-y* appears after stem-final vowels. The Turkmen marker is *-A<sup>4</sup>* after stem-final consonants, e.g. *ayd-a* ‘telling’, *bil-ε* ‘knowing’, *gör-ö* ‘seeing’, *tur-o* ‘standing’, while the postvocalic marker is *-y*, e.g. *oyno-y* ‘playing’.

In Turkic languages of the Northeastern and Southeastern branch, the converb *-A/-y* functions as the base of the finite intraterminal marker, e.g. *al-a* ‘takes’ ← *al-* ‘to take’, *aša-y* ‘eats’ ← *aša-* ‘to eat’.

The West Oghuz languages Turkish and Azeri display the marker *-(y)A<sup>2</sup>*. The glide *y* is not inserted as hiatus breaker, but continues the function of the old allomorph *-yU*. Turkish and Azeri are thus conservative in this respect.

The use of vowel-final converb markers is rather restricted in modern languages, where it mostly appears in combinations with descriptive verbs, e.g. Chuvash *-A pıl-* ‘to be able

to', -*A pušla-* 'to begin to', and in intensifying reduplications, e.g. Turkish *gül-e gül-e* 'constantly laughing', Chuvash *šir-a-šir-a* 'looking intensively'.

The Turkish converb occurs in combinations such as the ability marker  $-(y)A^2\text{-}bil-$  'to be able', the inability marker  $-(y)A^2\text{-}mA^2-$  'not to be able', actionality markers such as  $-(y)A^2\text{-}gel-$ , the expanded converb marker  $-(y)A^2\text{-}rA^2K$ , e.g. *al-arak* 'taking', *bekle-yerek* 'waiting', and the prospective marker  $-(y)A^2\text{-}jA^2G^2$ , e.g. *al-acak* 'will, shall take', *bekle-yecek* 'will, shall wait'. (On vowel raising see section 6.)

### 5.3.3 Consonant-final converb markers

Turkic languages possess another type of converb markers that mostly end in labial consonants, symbolized here by *-B*. Most languages have markers of the type  $-(V)B$ , e.g. Tatar *išle-p* 'having worked' ← *išle* 'to work'. East Old Turkic had the markers  $-(\tilde{I})b$  and  $-(\tilde{I})bAn$ .

The corresponding Turkish and Azeri markers, however, have the shape  $-(y)I^fB$ , e.g. Turkish *al-ıp* 'having taken' ← *al-* 'to take', *bekle-yip* 'having waited' ← *bekle-* 'to wait', *yaša-yıp* 'having lived' ← *yaša-* 'to live', Azeri *gör-üb* 'having seen' ← *gör-* 'to see', *oyna-yīb* 'having played' ← *oyna-* 'to play'.

It has been supposed that the glide *y* was inserted secondarily after stem-final vowels. The forms found in these two conservative languages are, however, likely to represent an earlier uncontracted stage. Modern Turkmen still displays  $-I^f p/A^2:p$ , i.e. with vowel lengthening, e.g. *ya:ša:-p* 'having lived' ← *ya:ša-* 'to live', *i:šle:-p* 'having worked' ← *i:šle-* 'to work', *oyno:-p* 'having played' ← *oyno-* 'to play'.

These forms suggest that the older shape of the marker was  $-(y)Vb$ , i.e.  $*ya:ša-yi-b$ ,  $*ba:šla-yu-b$  'having begun',  $*bekle-yi-b$  'having waited', etc. Uncontracted forms of this kind are actually found in older Turkmen texts (Baskakov et al. 1970:372-373).

Today, only conservative West Oghuz languages such as Turkish and Azeri display uncontracted forms of the corresponding converb markers.

### 5.3.4 Aorist markers

Turkic languages possess a so-called aorist, an old intraterminal category that was used to form presents and imperfects. The East Old Turkic aorist base marker was  $-Vr/-yUr$ . The vowels of the postconsonantal alternants were *A*, *I* and *U* (not predictable), the same vowels that occurred in the vowel-final converb marker (section 5.3.2).

The postvocalic allomorph was  $-yUr$ . Again, the initial *y* is not an element inserted as a hiatus breaker. It is likely that  $*-VyVr$  was the original shape of the marker. The short stem-final vowels were subsequently dropped, which led to contractions  $*-VyUr > -Vr$ . The forms with initial *y* survived longer after originally long stem-final vowels (Johanson 1975). The postvocalic allomorph early developed to  $-r$ , which occurs sporadically already in early sources. In Karakhanid of the 11th century, a few instances of  $-yur$  are found, but the normal postvocalic alternants are contracted forms such as *ba:šla-r* 'begins' ← *ba:šla-* 'to begin' (Brockelmann 1954:231).

In Khalaj, which has preserved numerous archaic features, the old postvocalic form  $-yUr$  has been retained, e.g. *bašla:-yur* 'begins' ← *bašla:-* 'to begin'. The form  $-yAr$  is found with a few stems, e.g. *sa:-yar* ← *sa:-* 'to comb cotton'. The marker  $-V(r)$  occurs after

stem-final consonants, e.g. *hat-ar* 'throws', *hat-a-m* 'I throw' ← *hat-* 'to throw' (Doerfer 1988:156-157). An overview of the Khalaj conjugation is found in Doerfer (1972).

Contracted forms in modern Siberian Turkic languages point to an older postvocalic form *\*-yVr*. Khakas has the postvocalic marker *-I:r*. Sayan Turkic varieties display *-V:r*, e.g. Dukhan *udu:r* 'sleeps' ← *udu-* 'to sleep', *sa:r* 'milks' ← *sa-* 'to milk', *ǰoru:r* 'goes' ← *ǰoru-* 'to go', *balıǰta:r* 'fishes' ← *balıǰta-* 'to fish' (Ragagnin in print). The Sayan variety Tuhan has the invariable marker *-i:r* after polysyllabic vowel-final stems, e.g. *aǰni:r* 'hunts' ← *aǰna-* 'to hunt' (Ragagnin 2009). For an overview of Sayan aorist vowels see Schönig (1989).

The Yakut aorist marker is *-A(r)/-I:(r)*. Stem-final vowels merge with the initial suffix vowel into a long high vowel *I:*, e.g. *ahi:r* 'eats' ← *aha:-* 'to eat'. The postvocalic markers evidently go back to *\*-yVr*, e.g. *basti:r* 'begins' ← *basta-* 'to begin' < *\*ba:šla:-yVr*.

The aorist has been claimed to be absent in Chuvash. As I have shown (Johanson 1975), however, it has actually survived as the so-called "indefinite future", whose marker is *-Ī*, e.g. *kil-Ī* 'will come' ← *kil-* 'to come', *šir-Ī-p* 'I will write' ← *šir-* 'to write'. Stem-final vowels are dropped, e.g. *vul-Ī-p* 'I will read' ← *vula-* 'to read', *vula-m-Ī-p* 'I will not read' ← *vula-ma-* 'not to read'.

The situation is similar in Turkmen, which is, as repeatedly noted above, highly conservative with respect to vowel length. The Turkmen aorist marker is realized as *-Ar* after stem-final consonants, e.g. *al-ar* 'takes, will take' ← *al-* 'to take', and as *-r* plus vowel lengthening after stem-final vowels, e.g. *ya:ša:-r* 'lives, will live' ← *ya:ša-* 'to live', *i:šle:-r* 'works, will work' ← *išle-* 'to work'. Also in this case, the postvocalic forms must be the result of *y* drop and contraction, since structures such as *\*ya:ša-ar* and *\*i:šle-er* are unlikely to have ever existed.

Azeri has developed the unified aorist base marker *-(y)Ar*, e.g. *al-ar* 'takes, will take' ← *al-* 'to take', *bil-er* 'knows, will know' ← *bil-* 'to know', *gel-er* 'comes, will come' ← *gel-* 'to come', *bašla-yar* 'begins' ← *bašla-* 'to begin'. Its distribution is thus predictable. *-(y)Ar* with its generalized low vowels has been created by analogy with the present base *-(y)Ir*, in which the high vowels have been generalized.

Most modern Turkic languages display the aorist marker *-Vr/-r*. Turkish has preserved the old non-predictable postconsonantal allomorphy, e.g. *bil-ir* 'knows, will know' ← *bil-* 'to know', *gel-ir* 'comes, will come' ← *gel-* 'to come'. The postvocalic marker is reduced to *-r*, e.g. *anla-r* 'understands, will understand' ← *anla-* 'to understand'.

### 5.3.5 Clitic personal markers

The modern Turkish clitic personal markers (non-past copulas) of the first person are *-(y)İ'm* 'I am' and *-(y)İ'z* 'we are', e.g. *burada-yım* 'I am here', *burada-yız* 'we are here'. The postvocalic glide is mostly taken to be an inserted hiatus-bridging segment. As I have suggested, the markers rather go back to earlier forms with a preceding element *i(y)-* < *\*er-ür* 'is', e.g. *\*iy-im*, *\*i-sin* (Johanson 2000).

### 5.3.6 Further markers of the type *-(y)V*

Turkish exhibits some further markers of the *-(y)V* type which are claimed to contain an inserted hiatus-bridging glide *y*.

The first person singular voluntative marker is  $-(y)A^2yI^2m$ , e.g. *yaşa-yayım* ‘let me live’ ← *yaşa-* ‘to live’. The first person plural voluntative marker is  $-(y)A^2II^2m$ , e.g. *yaşa-yalım* ‘let us live’. The second person imperative marker is  $-(y)I^4n(I^4z)$ , e.g. *yaşa-yın(ız)* ‘live’. The old necessitative-prospective marker is  $-(y)A^2sI^2$ , e.g. *yaşa-yası* ‘may s/he live’.

Not even these cases, however, verify the claim that the glide *y* has been inserted as a buffer consonant after vowel-final stems. Again, Turkmen displays counterparts with vowel lengthening that point to old contractions. The first person singular voluntative is  $-A^2yI^2n/$   $:-yIn$ , e.g. postvocalic *ya:ša:-yīn* ‘let me live’ ← *ya:ša-* ‘to live’. The first person plural voluntative is  $-A^2II^2η/:-II^2η$ , e.g. postvocalic *ya:ša:-liη*, the second person imperative marker is  $-I^2η/:-η$ , e.g. postvocalic *ya:ša:-η*, and the necessitative-prospective is  $-A^2sI^2/:-sI^2$ , e.g. postvocalic *ya:ša:-sī*.

### 5.3.7 $-(y)V < -GV$

The bracketed glide *y* is in many cases a reflex of an old morphophoneme *G*, which had front and back alternants, mostly *g* and *ɣ*. Its disappearance after consonant-final stems and preservation after vowel-final stems is a phonetic development typical of the Southwestern (Oghuz) branch of Turkic, particularly the West Oghuz languages Turkish and Azeri.

Also a number of Turkmen contracted forms have developed from suffix-initial glides that go back to a suffix-initial *G*. Astonishingly enough, Erdal sees no reason to assume that contraction of  $-GV$  passed through an intermediate stage involving  $-yV$  (2006:139).

The following examples represent retention of old velars in the shape of glides rather than insertion of glides in order to preempt hiatus.

The Turkish dative morpheme  $-(y)A^2$ , present in words such as *ev-e* ‘to the house’ and *oda-ya* ‘into the room’, goes back to  $-GA^2$ . The glide found in the postvocalic forms has thus not been inserted as a hiatus-breaker.

The corresponding Turkmen suffix is  $-A/A$ . The postconsonantal marker is  $-A$ , e.g. *at-a* ← *at* ‘horse’, *diš-ε* ← *diš* ‘tooth’, *toy-a* ← *toy* ‘party’, *öy-ε* ← *öy* ‘house’. The postvocalic marker is  $-A$ : i.e. with vowel lengthening due to consonant deletion, contraction, and fusion of the stem-final vowel with the suffix vowel, e.g. *köče:* ← *köčö* ‘street’, *da:ya:* ← *da:yī* ‘maternal uncle’, *düye:* ← *düyö* ‘camel’. After stem-final *a* and *o*, the marker consists of vowel length alone, e.g. *kino:* ← *kino* ‘to the cinema’, *ata:* ← *ata* ‘father’. Word forms such as *ğapa:* ← *ğapī* ‘door’, *köče:* ← *köčö* ‘street’, etc., have obviously developed from *\*kapi-ya*, *\*köče-ge*, etc., rather than from implausible forms such as *\*kapi-a* or *\*köče-ε*. The situation is similar in South Siberian Turkic languages, where the postvocalic dative markers display long vowels resulting from contraction.

The participle marker  $-(y)A^2n$ , present in Turkish *al-an* ‘taking, having taken’ ← *al-* ‘to take’, *bekli-yen* (written *bekleyen*) ‘waiting, having waited’ ← *bekle-* ‘to wait’, etc., goes back to forms in  $-GAN$ , e.g. *al-yan*, *bekle-gen*. The cognate Turkmen marker is  $-A^2n/:-n$ , i.e. with the postconsonantal allomorph  $-An$ , e.g. *gel-en* ‘having come’ ← *gel-* ‘to come’, and the postvocalic allomorph  $:-n$ , e.g. *ya:ša:-n* ‘having lived’ ← *ya:ša-* ‘to live’, *oka:-n* ‘having read’ ← *oka-* ‘to read’.

The optative marker  $-(y)A^2$ , present in Turkish *al-a-sın* ‘may you take’, *bekle-ye-sin* ‘may you wait’, etc., goes back to  $-GA(y)$ .

The converb marker  $-(y)A^2II^2$ , present in Turkish *al-alı* ‘since taking’, *bekle-yeli* ‘since waiting’, etc., goes back to  $-GAII$ .

The converb marker  $-(y)I^njA^2$ , present in Turkish *gel-ince* 'upon coming, *bekle-yince* 'upon waiting', etc., goes back to  $-GI^njA$ .

The agentive suffix  $-(y)I^f-jI^f$ , present in Turkish *al-ı-cı* 'receiver' ← *al-* 'to take', *dinle-yi-ci* 'listener' ← *dinle-* 'to listen', etc., goes back to  $-(I)G-JI$ . The corresponding Turkmen postvocalic marker is  $-y-jI$ , or, with lengthening of the stem-final vowel,  $-:jI$ , e.g. *dinle-y-jİ* 'listener' ← *dinle-* 'to listen', *okı:jİ* 'reader' ← *okı-* 'to read'. (On vowel raising see section 6.)

In all these cases, the bracketed glide *y* has evolved from a suffix-initial *G* and has thus not been inserted as a "buffer" consonant to prevent hiatus.

## 6 Allomorphy rules affecting stem-final vowels

Some allomorphy rules affecting stemfinal vowels have been mentioned above. A couple of additional rules will be touched upon here.

In Turkish, non-high stem-final vowels are raised before suffix-initial *-y* in certain verb forms. Thus the prospective marker  $-(\uparrow y)A^2jA^2K$  yields forms such as *arı-yajak* 'will/shall search' ← *ara-* 'to search' and *bekli-yecek* 'will/shall wait' ← *bekle-* 'to wait'. The first person plural volutative marker  $-(\uparrow y)A^2I^fm$  yields forms such as *arı-yalım* 'let us search' and *bekli-yelim* 'let us wait'. The converb marker  $-(\uparrow y)A^2$  yields forms such as *arı-ya* and *bekli-ye*. The current orthographic norm, however, prescribes harmonizing spellings such as *arayacak*, *arayalım*, *araya*, *bekleyecek*, *bekleyelim*, *bekleye*. Note that the bracketed initial *y* of the unaccentable imperative plural markers  $-(y)I^fn$  and  $-(y)I^fn-I^fz$  does not raise  $A^2$  to  $I^2$ , e.g. *ara-yİN*, *bekle-yİNİZ*.

Stem-final vowels undergo raising in certain Tatar verbal forms. The postvocalic present marker has the shape  $-(\uparrow y)$ , e.g. *qarı-y* 'looks' ← *qara-* 'to look', *işli-y* 'works' ← *işle-* 'to work'. The verbal noun marker is  $-(\uparrow U^2)w$ , e.g. *uqu-w* 'reading' ← *uqı-* 'to read', *biyü-w* 'dancing' ← *biyü-* 'to dance'.

Instances of stem-final vowel deletion at morpheme junctures were mentioned above. Morphophonemic changes of this kind affect many markers. The Chuvash present base marker is  $-(\emptyset)A^2t$ , e.g. *vul-at-ıp* 'I read' ← *vula-* 'to read'. The present participle is  $-(\emptyset)A^2kA^2n$ , and the prospective participle is  $-(\emptyset)A^2s$ .

The Turkish present base marker is generally considered to be  $-(I)yor$ . Since the preceding stem-final vowel is deleted, the formula for this marker is rather  $-(\emptyset)I^fyor$ . This formula allows verb forms to be construed in a predictable way according to normal vowel harmony rules.

Examples of postconsonantal forms: *al-ıyor* 'takes' ← *al-* 'to take', *ver-iyor* 'gives' ← *ver-* 'to give', *gül-üyor* 'laughs' ← *gül-* 'to laugh', *öl-üyor* 'dies' ← *öl-* 'to die', *sor-uyor* 'asks' ← *sor-* 'to ask', *kalk-ıyor* 'rises' ← *kalk-* 'to rise'.

Examples of postvocalic forms with deletion of the stem-final vowel: *başl-ıyor* 'begins' ← *başla-* 'to begin' (not *\*başla-yor*), *bekl-iyor* 'waits' ← *bekle-* 'to wait' (not *\*bekle-yor*), *yür-üyor* 'walks' ← *yürü-* 'to walk' (not *\*yürü-yor*), *ok-uyor* 'reads' ← *oku-* 'to read' (not *\*oku-yor*).

With negative stems in  $-mA^2$ , there is no need for a rule transforming the non-high vowel into a high one. We get forms such as *al-m-ıyor* 'does not take' (not *\*al-mı-yor*), *ver-*

*m-iyor* ‘does not give’ (not \**ver-mi-iyor*), *gül-m-üyor* ‘does not laugh’ (not \**gül-mü-iyor*), *öl-m-üyor* ‘does not die’ (not \**öl-mü-iyor*), *sor-m-uyor* ‘does not ask’ (not \**sor-mu-iyor*), *başla-m-iyor* ‘does not begin’ (not \**başla-mu-iyor*), *bekle-m-iyor* ‘does not wait’ (not \**bekle-mi-iyor*), *yürü-m-üyor* ‘does not walk’ (not \**yürü-mü-iyor*), *oku-m-uyor* ‘does not read’ (not \**oku-mu-iyor*), etc.

These allomorphy rules yield regular results and should be preferred to the complicated rules that are sometimes formulated for didactic purposes. For example, instead of the regular negation clitic  $-mA^2$ , Gerjan van Schaaik (2009) postulates a fourfold form  $-mI^2$  preceding  $-(I)yor$ , which is then realized as  $-yor$ .

The Turkish present base marker has developed from a sequence  $-V yori-r$ , a periphrastic construction consisting of a vowel-final converb plus \**yori-r*, the aorist of the auxiliary verb *yori-* ‘to move’, i.e. \**ara-yu yori-r* ‘moves searching’.

The above-mentioned corresponding Turkmen marker  $-yA:r$ , often contracted to  $-yA:$ , is traditionally said to have a similar origin. In my opinion, it rather goes back to a periphrasis consisting of a vowel-final converb + \**er-ür* ‘is’, the old aorist of *er-* ‘to be’. Contracted forms such as *gel-iyε:r* ‘comes, is coming’ and *yaz-iyε:r* ‘writes, is writing’ were still used as late as the 19th century. The modern marker has a simple structure without stem-final vowel deletion, e.g. *al-ya:r* ‘takes’, *bil-yε:r* ‘knows’, *gel-yε:r* ‘comes’, *bar-ya:r* ‘goes’, *işle-yε:r* ‘works’, *oқа-ya:r* ‘reads’, *al-ma-ya:r* ‘does not take’.

In the spoken language, the Turkish first person singular voluntative marker  $-(y)A^2yI^2m$ , tends toward the shape  $-(\emptyset)iyim$ , i.e. forms with stem-final vowel deletion and high front suffix vowels. These forms are even further contracted to  $-(\emptyset)i:m$ , e.g. *ariyim* ~ *ari:m* ‘let me search’ ← *ara-* ‘to search’, *bekliyim* ~ *bekli:m* ← *bekle-* ‘to wait’. For an overview of dialectal forms see Adamović (1985:244-245).

## 7 Conclusions

Some conclusions can be drawn with respect to hiatus-preventing allomorphy in Turkish and its relatives. The only constant rule is that vowels in two adjacent morphemes are not allowed to come into adjacency themselves, i.e. to stand “face to face”. They must keep their distance or merge. No Turkic languages have given up their resistance to vowel hiatus. Turkmen, Chuvash, Yakut and South Siberian Turkic have just applied their own methods, thus showing patterns different from those of other languages. They can no doubt tell us something about the diachronic background of the situation in other languages, e.g. Turkish and Azeri.

The popular myth about the insertion of “connective” vowels after consonant stems is highly implausible. The alleged insertion of “bridging” or “buffer” consonants in order to prevent vowel hiatus is unproven.

This is also true of the alleged hiatus-preemptory role of the glide *y*. The fact that the occurrence of morphemes of the structure  $-(y)V$  is more prominent in Turkish and Azeri than elsewhere in the Turkic-speaking world is due to the loss of the suffix-initial morphophoneme *G*, which is typical of these languages.

Cases of analogy cannot be totally excluded. The loss of *G* might have triggered the spread of  $-(y)V$  and the creation of analogical secondary forms that contributed to

morphological transparency. As a whole, however, there is no reason to suppose, as Marcel Erdal does, that the procedure of dropping the initial vowel was replaced in Turkish and Azeri by its “phonological opposite”, the addition of a palatal glide (2006:140). The author ascribes this allegedly fundamental change in morphophonemics to an areal influence of West Iranian languages. The Turkish verbal noun marker *-(y)lş* might even, as he supposes, be the result of contamination with a Persian suffix (1998).

The thesis regarding a decisive West Iranian influence on the alleged replacement of morphophonemic procedures is, however, unconvincing, especially in view of the situation obtaining in the non-Oghuz language Khalaj. The latter is spoken in central Iran, in a setting where the West Iranian influence ought to be particularly strong. Khalaj does not share the loss of the suffix-initial morphophoneme *G* and does not show any tendency to use *y* in the allegedly hiatus-preventing way observed in its Oghuz neighbours. This clearly contradicts Erdal's thesis. The author himself remarks that Khalaj “did not participate in the wholesale adoption of the palatal glide /.../ although its speakers live in the West Iranian heartland”. He adds the enigmatic comment that the fact that Khalaj “does not drop suffix initial /g/ may have meant that it lacked the base for doing so” (2006:141).