PROCEEDINGS

of

the 12th Seoul International Altaistic Conference

Multidimensional understandings
of languages and cultures of Altaic people

Dates: 16~19 July 2015

Venue: Shinyang Hall II, Seoul National Univ., Korea

The Altaic Society of Korea
Institute of Altaic Studies, SNU

Standard Turkic reduplications of
the extended *C*-type: the different types

Kamil Stachowski
Jagiellonian University, Poland

1. Introduction

Repetitions and reduplications are abundant in the Turkic languages. Some are structurally obvious, as Tksh. *birer birer* 'singly, one by one', some are more challenging. One specific kind among the latter are what could be called *partial interfixed* reduplications, i.e. words which have just a part of them repeated and appended to them, with some addition in between. Their semantics is essentially an intensification of the meaning of the original word.

By far the most numerously represented type among them is what I call the *C*-type, whose structure can be described so:

initial mora of the base + a single, lexically determined consonant + base

where *base* is the original word that undergoes reduplication, the single consonant appended to it is called the *closing consonant* (because without it the initial syllable of the final reduplication would usually end in a vowel; ℂ for short), there is no additional interfix beyond the ℂ, and a word is considered to have as many morae as many of the following it has: any vowel, a long vowel or a diphthong, consonantal auslaut (so e.g. Trkm. *gȫk* 'blue' is three morae long). Examples: Azer. *dopdolu* 'completely full' (AzRS), Kazakh *nypnyk* 'very tough, durable' (SKzkP), or Yak. *büsbütün* 'absolutely all, whole' (Pekarskij 1907–30). Below, the ℂ together with the interfix will be called the *closer*.

The *C*-type is also the most researched type, with several articles and three monographies devoted to it: Hatiboğlu 1973, Müller 2004, and Stachowski K. 2014. It is not limited to the Turkic languages. Besides several examples in Armenian, Tajik and other languages, that were clearly inspired by Turkish, and a few unclear words in Estonian and Finnish (see 2.4.1), it is also quite well represented in the Mongolic languages. These cases are being worked on (e.g. Kim 2014), but a comprehensive, exhaustive approach has not as yet been made. Nonetheless, the sheer number of Turkic examples of this type, and the state of research on them, is why this type is often referred to below.

This paper is considered with the other types of partial interfixed reduplications – or, at least, with formations which bear enough structural resemblence to them to be counted as such until their history and origins, and thus true nature, is investigated and revealed. It does not include full reduplications, interfixed or not, such as the above mentioned Tksh. *birer birer*, or Uzb. *karšyma-karšy* 'face to face, vis-à-vis' (UzbRS41). It is based on a certainly incomplete collection of more than three hundred examples from seventeen languages.

Section 2 presents a classification of the different types, using two criteria: the closing consonant, and the interfix. In lack of previous studies of the topic, the distinction between these two elements must be seen as provisional – as indeed should be the entire classification. It hopes to be practical in the same way that Linnæus's *Systema* remains relevant today, regardless of whether it is accurate in reflecting the actual genealogical lineage. Like in biology, a historically correct classification cannot be drawn until the origin of each type becomes known, and this certainly is not yet the case.

In the examples adduced below, two elements are often omitted: the source in which the base is attested, and the meaning of the reduplication. The former is not mentioned when it is the same as the source for the reduplication, and the latter when it is a simple intensification of the meaning of the base. All examples are given in a unified phonological transcription using the notation employed traditionally in Finno-Ugric studies; see Stachowski K. 2014: 30. Full stop delimits morphemes.

2. Types

It was mentioned above that the following classification is only apparently structural: until the history of each word becomes known, one can identify patterns but not decide whether they are meaningful or mere coincidences resulting from secondary evolution, contamination, or perhaps yet another process. I grouped all the reduplications by two properties: what appears to be the closer and what seems to be the interfix, and obtained a surprisingly high number of twenty-three different types. As it happens, even such a grouping leaves room for discussion, in particular regarding the nature of the closer ($\mathbb{C}$ or C_2).

The method is reflected in the titles of the subsections below. For clarity, similar types, or types which now appear separate but are relatively likely to prove ultimately not to be so, are grouped together in one subsection (two-digit numbers) and different subsubsections (three-digit numbers). The titles of subsubsections are the middle, variable part of the structure of the reduplications, so that each subusbsection presents formations whose built can be described using the formula

initial mora of the base + title of the subsubsection + base

where $\mathbb{C}$ = closing consonant (i.e. a single, lexically determined consonant that is unrelated to the base), C_2 = the first post-vocal consonant of the base (e.g. *r* in *kara*, but also in *ürdük*), A = low vowel, I = high vowel, the notation ‹$\mathbb{C}|C_2$› is used when it is not possible to determine whether the given closer is a $\mathbb{C}$ or a C_2, while ‹$\mathbb{C}$ or C_2› when examples with $\mathbb{C}$'s and C_2's are mixed within one subsubsection.

2.1. Closer = $\mathbb{C}$ or C_2; interfix = $\varnothing$

2.1.1. $\mathbb{C}|C_2 + \varnothing$

Kar. **birlej**[1] 'only, sole, single' ◊ E ***bi.r.birlej*** (KRPS, RKarS-Haf, RKarS-Lev, CKarED), SW ***bi.r.birłej*** (KRPS s.v. *бир-бирлэй* and *бирлэй*)

Yak. **buruolāχ** 'smoky' ◊ ***bu.r.buruolāχ*** (Pekarskij 1907–30)

Kar. **jumarlak** 'round' ◊ E ***ju.m.jumalak*** (KRPS, RKarS-Lev, CKarED), ***ju.m.jumarlak*** (KRPS, RKarS-Lev)

Tksh. **perišan** 'miserable, wretched' ◊ ***pe.r.perišan*** (Müller 2004: 109, Stachowski M. 2009: 118)

---

1 The base **birlej* does not seem to be attested in Eastern Karaim but cf. Kar.SW *birłej* id.

The above examples are unclear because what appears to be the closer in them, can be both a $\mathbb{C}$ or C_2. In Kar.E, both *m* and *r* are attested in *C*-type reduplications (e.g. *symsylak* ~ *syrsylak*), and *r* is used as a closer both in Turkish and Yakut (e.g. *tertemiz*, and *borbosχo*, respectively). It is only the southwestern Kar. *birbirłej* that suggests that we are dealing with C_2 here as I do not know of any Kar.SW reduplications closed with *r*. It is, nonetheless, quite possible that there exist such forms, only somehow they did not find their way into the dictionaries. I am not aware of any more words such as these.

2.1.2. $C_2 + \varnothing$

Kar. ***boklavat**[2] 'meanness, baseness, villainy' ◊ E ***bo.k.baklavat*** (RKarS-Haf), ***bo.k.boklavat*** (KRPS, CKarED)

Kar. **bošuna** 'in vain' ◊ E ***bo.š.bošuna*** (RKarS-Haf), ***bo.š.bošyna*** (KRPS, RKarS-Lev, CKarED)

Bshk. **ɣišrät** '1. feast; 2. dissolute life, debauchery' ◊ ***ɣi.š. ɣišrät*** (BškRS96)

Bshk. **minlek** 'self-esteem' ◊ ***mi.n.minlek*** 'conceit, pride' (BškRS58, BškRS96)

The above are the only certain examples that I am aware of (see above). At first sight, one might be tempted to reinterpret them as a simple reduplication of the entire initial syllable, rather than just the first mora, but *bošbošuna* and *bošbošyna* negate this possibility; likewise the less clear *perperišan* above. None of the examples that I am aware of, begins with a vowel, or has a long vowel or a diphthong in the initial syllable – which are

---

2 The base *boklavat does not seem to be attested. The connection with bok 'faeces' (KRPS) is quite apparent even if the form itself is not entirely clear.

the circumstances in which the reduplications presented here most frequently divert from the more common *C*-type.

The limited geographical spread suggests that this type may be a local innovation, but should the examples above prove to also historically belong to the same type, the one Yakut example among them would speak against this possibility. Given Yakut's prolificacy in the area of reduplication, a coincidental similarity should not be ruled out too hastily.

2.2. Closer = $\mathbb{C}$ or C_2; interfix = low vowel

2.2.1. $\mathbb{C}$ + low vowel (illabial)

Khal. **bälgälü̈.ɣ** 'completely clear' ◊ ***bä.s.ä.bälgälü̈.ɣ*** (Doerfer/Tezcan 1980)

Tksh. **düz** 'straight, smooth, even' ◊ ***dü.m.e.düz*** (Müller 2004: 119)

Bshk. **tiŋ** 'equal, similar' ◊ ***ti.p.ä.tiŋ*** (BškRS58, Juldašev 1981, BškRS96)

Ott. **sag** 'healthy, alive' ◊ ***sa.p.a.sag*** (Meninski 1680: II 2202 s.v. دب دري *dibdirī*: ‹صاغ صپا ʃapaʃagh›, Comidas de Carbognano 1794: 21: ‹صاغ صاپه sapa sagh›)

Kirg. **tūra** 'rightly, correctly' ◊ ***tu.p.a.tūra*** (KirgRS)

With thirty-two examples, this is the third most numerously represented type. The $\mathbb{C}$ is in the great majority of cases *p*, even though fourteen examples come from languages where other consonants are routinely also used as closers. The exceptional cases are: Khal. *bäsäbälgälü̈.ɣ* above, Tksh. *besebelli* 'completely certain, clear' (Kononov 1956: 157) and *dümedüz* above. Long vowels appear in the initial syllable in four cases, from Gagauz, Kirghiz, and Turkish, and are always shortened in the reduplication; diphthongs do not seem to be attested in this position. I am not aware of any reduplications of this type, of bases beginning with a vowel.

The geographical spread is fairly wide; I know of examples from Bashkir, Gagauz, Karakalpak, Kazakh, Khalaj, Kirghiz, Ottoman, and Turkish. This may suggest that the method is quite ancient. It is not possible to tell for certain what the origin of the interfixed vowel is, but one might venture a guess that it is there quite simply to increase the emphasis.

2.2.2. ℂ + low vowel (labial)

Kirg. **čogū** '(all) together; entirely' ◊ ***čo.p.o.čogū*** (KirgRS)
Dolg. **höp** *i.a.* 'appropriate, suitable' ◊ ***hö.r.ö.höp*** 'as if made for …' (DW)
Kirg. **kündüz** 'by day' ◊ ***kü.p.ö.kündüz*** 'in broad daylight' (KirgRS)
Kirg. **tono-** *i.a.* 'to rob' ◊ ***to.p.o.tonop*** 'having robbed clean' (KirgRS)

The above are all the examples that I know. I am not certain that they should be considered a separate type, but cf. Kirg. *tupatūra* in the group above, and *küpökündüz* here. However, it is also true that *tupatūra* is the only Kirghiz example of this type, which has an illabial interfix while all the remaining three have *o* or *ö*. Note should be also made of *tono-*, a rare case of a reduplicated verb.

The Dolg. *höröhöp* stands out in this group, structurally (*r* for the closer), and geographically. I know of no more examples of this type, either in Dolgan or in Yakut. It would be difficult to escape an association with Mongolic reduplications in *ra, ro,* and others, such as Bur. *boro borxigor* 'very homely (an old house)' or *tere tesxeger* 'very fat' (see Stachowski K. 2014: 271). Continuing this train of thought, I should like to mention Finnish where such examples can be found as *tipotiessään* 'completely lost', or *ypöyksin* 'completely alone'. No more, however, than structural similarity can be cited at the moment to connect them to Standard Turkic words; see 2.4.1.

2.2.3. $\mathbb{C}|C_2$ + low vowel

Gag. **parča** 'piece, part' ◊ ***pa.r.a.parča*** 'shattered, in pieces' (GagTS)

Bshk. **karšy** 'opposite, contrary' ◊ ***ka.r.a.karšy*** '1. opposing, located in opposite; 2. against each other, face to face' (BškRS58, BškRS96)

Tat. **karšy** 'opposite, contrary' ◊ ***ka.r.a.karšy*** (Laude-Cirtautas 1961: 34, TatRS)

Only the above three examples belong here, and all three are uncertain. Both bases, and especially *karšy*, appear in reduplications of many other types; *karakaršy* could be seen as an irregular, emphatic modifications of one of the more numerously attested, and thus more certain, type. Alternately, Laude-Cirtautas 1961: 34 sees in *karakaršy* an example of the usage of *kara* as an intensifier, similar to Ott. *kara johsul* 'poor and miserable', Sag., Shor *kara čagyš* 'completely (a)lone', &c. As for *paraparča*, it too is attested in reduplications of another type (see 2.6.1), and it should also be noted that *r* is used in Gagauz as a $\mathbb{C}$ in the *C*-type (*čyrčyplak*, GagTS). All this suggests that the overlap with C_2 may be purely coincidental

2.3.   Closer = $\mathbb{CC}$; interfix = low vowel

Kar. **ak** 'white' ◊ NW ***a.pp.a.ah*** (Józefowicz 2008 s.v. *bialuteńki* and *bieluteńki*)

Azer. **düz** 'smooth, even, straight' ◊ ***dü.bb.ä.düz***, ***dü.pbä.düz*** (AzRS)

Kmk. **boš** 'empty' ◊ ***bo.pp.o.boš*** (Dmitriev 1940: 71f, Doniyorova 2004: 19)

Uzb. **tik** 'vertical, steep, straight' ◊ ***ti.pp.a.tik*** (UzbRS41, Kononov 1960: 157, Bodrogligeti 2003: 352f)

Uzb. **tola** 'all' ◊ ***to.pp.a.tola*** (von Gabain 1945: 49)

In total, I know of thirty-four examples of this type. The $\mathbb{C}$ is always *p* or *b*, the latter only in Azeri, in the two reduplications of *düz* above, and in *sapbasaɣ* '1. completely healthy; 2. completely unscathed, sound' (AzRS). This is despite the fact that twenty-three of these words are from languages where other $\mathbb{C}$'s are also in use. The C_2 is never *p* or *b*, and there are no cases of long vowels or diphthongs in the initial syllable.

The geographical spread is moderate: Azeri, Karaim (NW), Karakalpak, Kumyk, and Uzbek (nineteen of the examples known to me). I suspect that the relation between this type and the type discussed in 2.2 ($\mathbb{C}$ + low vowel), is the same as between the *C*-type (*a.p.ak*) and the *C*-type with a doubled closer (*a.pp.ak*). If this were so, the doubling would likely only be there for added emphasis. In Stachowski K. 2014: 201f, I listed five ways in which forms such as *appak* could have arisen. This type, and the fact that it stands in nearly the same proportion to forms such as Gag. *düpedüz* 'openly, frankly' (GagTS) as *appak* does to *apak*, seems to reinforce the possibility that the doubling is a result of no more complex a process than emphatic lengthening; see therein and 207f for further commentary and examples. ("Nearly" because the type presented in 2.2 does sporadically employ other $\mathbb{C}$'s than *p*, which is not the case with the *C*-type with a doubled closer.)

2.4.   Closer = $\mathbb{C}$ or C_2; interfix = high vowel

2.4.1. $\mathbb{C}$ + high vowel

Gag. **düz** 'smooth, even, straight' ◊ ***dü.b.ü.düz*** '1. *intens.*; 2. everywhere, all over' (GagTS) ◊ ***dü.p.ü.düz*** (Pokrovskaja 1964)

Uzb. **ješil** 'green' ◊ ***je.m.i.ješil*** (von Gabain 1945: 49)

Yak. **lap** 'true, faithful, accurate' ◊ ***la.b.y.lap*** (Pekarskij 1907–30)

Yak. **öñös** 'stubborn, perverse' ◊ ***ö.r.ü.öñös*** (Pekarskij 1907–30)

Yak. **söp** 'worthy, fit, decent' ◊ ***sö.r.ü.söp*** 'just: fit, right, in time' (Pekarskij
    1907–30)

The two Gagauz and one Uzbek examples are the only three not from Yakut
where, including the two above, there are at least as many as twenty-six.
Intriguingly, only three of them have a ℂ other than *r*: *labyslap* above,
epiemeger 'having very small and tight lips', and *mayymaŋan* 'snow-white'
(all three Pekarskij 1907–30). The *γ* in this last form is in itself rather
unusual. It is also noteworthy that *öñös*, despite its vocalic anlaut, has *r* for
the closer. In the *C*-type, it is an almost exceptionless rule that such
reduplications are closed with *p* or *pp*; the only odd example is also Yakut,
and unclear in more than one way (*ūnutary*; Stachowski K. 2014: 215). One
might wonder whether the 23 Yakut cases with *r* should not be treated
independently from *epiemeger*, *labylap*, *mayymaŋan*, and the Gagauz and
Uzbek words. Two examples have a diphthong in the initial syllable, and it
has been shortened in both (*ču.ru.čuoχaγyr* 'with a thin neck', and
kü.rü.küökeger 'with a gracefully elongated neck', both from Pekarskij 1907–
30); long vowels do not seem to be attested in this position.

　　The geographical spread might be seen as suggesting that this type is an
ancient one. This may be possible, but a coincidence cannot be ruled out so
long as the origin of the interfix is not known. The vowel may be there
simply to add emphasis, and in such case the same method can be imagined
to have arisen independently in various languages.

　　Incidentally, two such languages might have been Estonian and Finnish,
where one finds *uhiuus* 'brand-new', *võhivõõras* 'total stranger' (Estonian), or
putipuhdas 'completely clean' and *supisuomalainen* 'purely Finnish' (Finnish).
See 2.2.2, and Stachowski K. 2014: 20f.

2.4.2. $\mathbb{C}|C_2$ + high vowel

Yak. **beriṅ́ńeχ** 'generous' ◊ ***be.r.i.beriṅ́ńeχ*** (Pekarskij 1907–30)
Yak. **moγojdōχ** 'full of snakes' ◊ ***mo.γ.u.moγojdōχ*** (Pekarskij 1907–30)
Yak. **murunnāχ** '(big-)nosed' ◊ ***mu.r.u.murunnāχ*** (Pekarskij 1907–30)
Yak. **ynaγar** 'paunchy, tubby' ◊ ***y.n.y.ynaγar*** (Pekarskij 1907–30)

I know of a dozen examples that belong to this group, all from Yakut. The $\mathbb{C}$'s are: *j* (1), *n* (3), *ŋ* (1), and *r* (7); all four are also attested in *C*-type reduplications (Stachowski K. 2014: 184). Noteworthy is the use of *n* for the closer in *ynyynaγar*, as is the resulting sequence of two adjoining vowels. As above, the domination of *r* suggests that these words might in fact belong to a separate type.

2.4.3. C_2 + high vowel

Yak. **balys** 'junior, younger' ◊ ***ba.l.y.balys*** (Pekarskij 1907–30)
Yak. **bolorχoj** 'cloudy, hazy' ◊ ***bo.l.u.bolorχoj*** (Pekarskij 1907–30)
Yak. **möltöχ** 'weak' ◊ ***mö.l.ü.möltöχ*** (Pekarskij 1907–30)
Yak. **ǯulaj** 'to fear, to be afraid' ◊ ***ǯu.l.u.ǯulaj*** (Pekarskij 1907–30)

Beside the above, I am only aware of three examples of this type, both from Yakut. All but one have *l* for C_2 (the exceptional one is *moγumoγojdōχ* 'full of snakes' (Pekarskij 1907–30)). I am not certain why this may be so. One possible explanation is that Yakut has as many as ten different $\mathbb{C}$'s attested for the *C*- and other types, an unusually high number, and this caused multiple cases to be misinterpreted here as having a $\mathbb{C}$ rather than C_2 for the closer. This could be seen as an argument in favour of treating the examples

with *r* as a separate type. Vocalic anlaut or long vowels or diphthongs in the initial syllable do not seem to occur among these cases.

2.5. Closer = ℂ or C_2; interfix = high vowel + *s* or *t*

2.5.1. ℂ + high vowel + *s*

Yak. **čepčeki** *i.a.* 'inexpensive, cheap' ◊ ***če.b.is.čepčeki*** (Pekarskij 1907–30)

Yak. **čuoɣur** 'motley' ◊ ***čuo.b.us.čuoɣur*** (Pekarskij 1907–30)

Gag. **jeni** 'new' ◊ ***je.p.iz.jeni*** (GagTS, Özkan 1996)

Yak. **kǖstēχ** 'strong, powerful' ◊ ***kü.b.üs.kǖstēχ*** (Pekarskij 1907–30)

Dolg. **sygyńńak** 'naked' ◊ ***hy.b.ys.sygyńńak*** (Ubrjatova 1985, DW, Stachowski M. 1997)

Yak. **ürdük** 'high, tall' ◊ ***ü.b.üs.ürdük*** (Pekarskij 1907–30)

With as many as eighty-six examples in total, this is the most well-represented type. The ℂ is almost always *b*; there is one word in *p* (Gag. *jepizjeni*), one in *n* (Yak. *tiiönüstüökün* 'big cheater' (Pekarskij 1907–30), and three in *r* (Yak. *borusbosχo* 'completely straight, upright', *bürüsbütün* 'all, whole' and, interestingly due to its vocalic anlaut, *irisitī* 'very hot, torrid' (all three Pekarskij 1907–30). The remaining fifteen bases that begin with a vowel all have *b*. Also unexpected is the fact that in four out of five cases, the diphthong in the initial syllable has been retained in the reduplication (as in *čuobusčuoɣur* above); long vowels, conversely, have been shortened in all four examples (as in *kübüskǖstēχ* above).

The geographical spread is rather peculiar: eighty examples are from Yakut, five from Dolgan, and one is from Gagauz. I am not aware of any similar forms in between. This may be seen as a suggestion that this type must be an ancient one, but perhaps it is a coincidence. The origin is unclear; if, however, it is simply a combination of two ℂ's with a vowel added to

facilitate the pronunciation, assembled to express an additional emphasis, then I should note that such a phenomenon can also occur spontaneously in spoken Turkish (H. Sofu – p.c.). But if this were the case, one would have to wonder why the second ℂ is always *s*, and the first almost always *b* (or *p*) when so many others are available (in Yakut, as many as ten are attested in the *C*-type). Surely, *p* – and *b* can be just an intervocalic realization of *p* – is by far the most common, and most probably the original ℂ, but the same cannot be said about *s*.

Were *b* indeed merely a result of intervocalic voicing, four words would need to be moved from here to the second group below: *čebisčepčeki* above, and *χobusχop* 'slander, calumny', *labyslap* 'absolutely true, faithful, accurate', *söbüsöp* 'very appropriate, convenient, good' (all three Pekarskij 1907–30).

2.5.2. ℂ + high vowel + *t*

Dolg. **čā̄lkā̄** 'white' ◊ ***čä.b.it.čā̄lkā̄*** (DW)

Yak. **sygyńaχ** 'naked, nude' ◊ ***sybyččygyńaχ***, ***sybytčygyńaχ*** (Pekarskij 1907–30)

Yak. **symyja** 'a lie' ◊ ***sybyččymyja*** 'an utter lie' (Pekarskij 1907–30)

The above three examples are all that I am aware of. I am not sure that they should be considered a separate type. The *t* might be merely a result of the pan-Siberian alternation between *s* and *t*. Very little is known about this phenomenon, and it is unclear whether it could have affected these particular words.

2.5.3. $\mathbb{C}|C_2$ + **high vowel** + *s*

Yak. **könö** 'smooth, straight, right, proper' ◊ ***kö.n.üs.könö*** 'very straight, righteous' (Pekarskij 1907–30)

Yak. **muŋnāχ** 'torturous' ◊ ***mu.ŋ.us.muŋnāχ*** (Pekarskij 1907–30)

Yak. **seber** 'clean, neat' ◊ ***se.b.is.seber*** (Pekarskij 1907–30)

Yak. **sür** 'fear, horror' ◊ ***sü.r.üs.sür*** 'dreadful' (Pekarskij 1907–30)

All the twenty-one known examples are from Yakut. The $\mathbb{C}$'s are *b* (2 examples), *n* (2), *ŋ* (2), *r* (11), *s* (2), and *t* (2). All of them are used as $\mathbb{C}$'s in Yakut *C*-type reduplications, and therefore it is impossible to tell whether here they have been used in the same capacity, or as a repetition of C_2. Interestingly, both words beginning with a vowel have *r* for the closer (*ürüsürdük* 'very high, tall', *yrysyrās* 'completely clean, pure, flawless, saint'); the only example with a long vowel or a diphthong in the first syllable is *lebislēbilēχ* 'very muddy' (all three Pekarskij 1907–30).

2.5.4. C_2 + **high vowel** + *s*

Yak. **aĥȳ** 'bitter' ◊ ***a.ĥ.ys.aĥȳ*** (Pekarskij 1907–30)

Yak. **saĥarχaj** '(reddish) yellow, rufous, bay, brown' ◊ ***sa.ĥ.ys.saĥarχaj*** 'very yellow' (Pekarskij 1907–30)

Yak. **čuoɣur** 'motley' ◊ ***čuo.ɣ.us.čuoɣur*** (Pekarskij 1907–30)

Yak. **ńulun** 'savourless, insipid' ◊ ***ńu.l.us.ńulun*** (Pekarskij 1907–30)

Together with the ones above, this group comprises twenty-one examples, all from Yakut. Three begin with a vowel, and in all three the closer does not happen to be *p* (it is twice *ĥ* and once *d*). Only one has a diphthong in the initial syllable, and it has not been shortened (*čuoɣusčuoɣur* above); no long vowels in this position.

2.6.   Closer = $\mathbb{C}$ or C_2; interfix = *Am, ama,* or *mA*

2.6.1. $\mathbb{C}$ or C_2 + *Am*

Gag. **parča** 'piece, part' ◊ *pa.r.am.parča* 'shattered, in pieces' (GagTS)

Tksh. **parča** 'piece, part' ◊ *pa.r.am.parča* 'shattered, in pieces' (Müller 2004: 102)

Azer. **täläsik** 'hastily, hurriedly' ◊ *tä.l.äm.täläsik* (AzRS)

I do not know of any more examples. In the case of *paramparča,* the closer might be a $\mathbb{C}$ or C_2 since *r* is used in both Gagauz and Turkish in this function; in the case of *tälämtälasik,* it should be considered C_2 as there appear to be no *C*-type or other reduplications in Azeri closed with *l.* Though these three examples certainly appear to be somehow linked to those listed below, the exact nature of this relation remains a mystery to me. See also *karymkatynas* in 2.7.5.

The rather condensed geographical distribution should be noted, as it might be a clue that this type arose – and, apparently, quickly died out – as a limited, local innovation in the Oghuz branch, probably not older than the 11th century. See also the group below.

2.6.2. $\mathbb{C}$ or C_2 + *ama*

Kklp. **karsy** 'opposite, contrary' ◊ *ka.r.ama.karsy* (KklpRS)

Kzk. **kajšy** 'opposite; contrary, contradictory' ◊ *ka.r.ama.kajšy* (KzkRS, SKzkP)

Kzk. **karsylyk** 'oppositeness; contrariness, contradiction' ◊ *ka.r.ama.karsylyk* (KzkRS, SKzkP)

Uzb. **karšylyk** 'opposition, contrast' ◊ *ka.r.ama.karšylyk* 'face to face, vis-à-vis' (UzbRS41, UzbRS59)

This group is strangely regular: the closer is always *-rama-*, and it comprises the reduplications of *karšy* and *karšylyk*, one each, in Karakalpak, Kazakh (in two phonetic variants), and Uzbek. Technically, the closer should be identified as C_2 in all the cases except Kzk. *kajšy* and *kajšylyk*. In these two, it is obviously not C_2, but it would be also quite surprising to consider it simply a $\mathbb{C}$ as *r* does not appear in this function in any of the 108 Kazakh examples for the C-type, all being closed by *p* or *pp* (Stachowski K. 2014: 90) – cf., however, *karymkatynas* in 2.7.5. I suspect that *karamakajšy* and *karamakajšylyk* must have been coined by analogy to *karamakarsy* and *karamakarsylyk*, possibly without an actual understanding of the process that had created the latter two in the first place.

Like the group above, it appears to be strictly limited to a single belt, this time across Central Asia, suggesting a local innovation. The fact, however, that it seems to be restricted to a single root may be an indication that these formations are in fact a product of some entirely different process, and the similarity to reduplications is purely accidental. See also the group below.

2.6.3. $\mathbb{C}$ + *mA*

Tksh. **dā̗ynyk** 'scattered, dispersed' ◊ ***da.r.ma.dā̗ynyk*** (Müller 2004: 100, Stachowski M. 2009: 120)

Azer. ****daɣyn*** ◊ ***da.r.ma.daɣyn***[3] (AzRS)

Bshk. **karšy** 'opposite, contrary' ◊ ***ka.p.ma.karšy*** '1. completely: opposite, contrary; 2. face to face, against each other' (BškRS58, BškRS96)

Kar. **tolu** 'full' ◊ E ***to.p.ma.tolu*** (RKarS-Haf s.v. *переполненный*)

---

3 Apparently, only attested in: ~ *etmäk* 'to smash, to break, to crush' and ~ *olmag* 'to collapse, to break down, to be defeated'. As for the base, *daɣyn does not seem to be attested as such but cf. *daɣylmag* '1. to collapse; 2. to break up', *daɣynyg* 'scattered' &c. (both AzRS), and cognate forms in other languages, such as Tksh. *dā̗ynyk* above.

This group contains eleven examples; seven are *karšy* and its derivatives in Bashkir, Tatar, and Turkmen; three are bases related to *dagy-* 'to scatter' in Azeri, Karaim, and Turkish; and only one is the Karaim *topmatolu* above. Considering this, I am not certain that this group really ought to be separated from the two below, but looking just at its structure, it should. The ℂ is always *p* with *karšy*, *r* with *dagy-*, and again *p* with *tolu*. There is one case of a long vowel or a diphthong in the initial syllable, in *darmadāᵧynyk* above.

The geographical spread is sickle-shaped: from Bashkir and Tatar, through Karaim, to Turkish and Azeri, and then on to Turkmen. Not much can be inferred from this fact alone, but I should note the similarity, perhaps insignificant, to the distribution of ℂ's other than *p* in the *C*-type, which is basically the same, only with the tips stretching out as far as to Dolgan and Yakut in the north, and to Uzbek in the south.

I need to mention two words here, Azer. *därmä-dešik* and *därmä-dešikli* 'all in wholes' (AzRS). Strictly speaking, they are not reduplications because the would-be reduplicated initial mora has in them a different vowel than the base. (Even if such words actually may form in specific conditions, e.g. Uigh. *japješil* 'intensely green', see Stachowski K. 2014: 212f.) But they are also most likely just modifications of *dälmä-dešik* id., a compound, not a reduplication (Stachowski M. 2009: 118). What makes them interesting here, is that they seem to be modifications fashioned after reduplications, and what is more, after reduplications of a rare and structurally unclear type. This apparent carelessness in the coining of new reduplication-like words may in fact be the disappointingly prosaic reason behind their typological richness.

2.6.4. $\mathbb{C}|C_2 + $ *ma*

Azer. **garyšyglyg** '1. entanglement, confusion; 2. disorder, unrest, stir' ◊ ***ga.r.ma.garyšyglyg*** 'chaos, randomness, mess' (AzRS) ◊ ***ga.t.ma.garyšyglyg*** 'chaos, randomness, mess' (AzRS)

Kar. **karyšyk** 'mixed' ◊ E ***ka.r.ma.karyšyk*** (KRPS, RKarS-Haf, CKarED), ***ka.r.ma.kiryšyk[4]*** (RKarS-Haf)

Ott. **karyšyk** 'mixed, mingled, confused' ◊ ***ka.r.ma.karyšyk*** (Meninski 1680: II 3579: ‹قارشق قارمه *karma karyśik*›, Jehlitschka 1895: 57: قارمه قاریشـیق *qarma qaryẓýq*)

There are ten examples in this group, and all are reduplications of *karyšyk* or one of its cognates. The sole reason for them being separated into another group is that the $\mathbb{C}$'s, *r* and thrice *t* in Azeri forms, are all attested in this function in *C*-type reduplications. In the case of *t*, however, it appears to only be used in just two *C*-type reduplications in Azeri, that of *garyš, and that of *garyšyg* (see Stachowski K. 2014: 51). It might be that this group should in fact be treated together with the one above, and the one below.

The geography is a segment of the sickle sketched for the group above. Here, examples come from Karaim, Gagauz, Ottoman, Turkish, and Azeri.

2.6.5. $C_2 + $ **ma**

Kklp. ***karyšyk** 'confusion, mess, chaos' ◊ ***ka.r.ma.karyšyk*** (KmkRS)

Uzb. **qaryšyq** 'confused, tangled ◊ ***qa.r.ma.qaryšyq*** (von Gabain 1945: 49)

Kar. **syk** '1. often; 2. thick ◊ E ***sy.k.ma.syk*** 'fill, overfilled' (KRPS, RKarS-Haf, RKarS-Lev, CKarED)

---

4 The base *kiryšyk does not seem to be attested independently. It is probably no more than a phonetic peculiarity; see Stachowski K. 2010: 152.

Azer. **täläsik** 'hastily, hurriedly ◊ **tä.l.mä.täläsik** (AzRS)
Kar. **tolu** 'full ◊ E **to.l.ma.tolu** (RKarS-Haf s.v. *книжный шкаф*)

The above are the only examples that I am aware of. The two reduplications of *karyšyk* are only listed here because neither Karakalpak nor Uzbek seem to employ *r* as a ℂ in the *C*- or other types. Perhaps they should be considered together with one of the two groups above, instead.

Also *sykmasyk* raises association with another type of reduplication. There exists in at least several languages, a fairly numerous group of formations composed of, most frequently, monosyllabic words, doubled, and with *ma* (or *ba, ha*) inserted in between them. Examples: Azer. *günbägün* 'everyday, day by day', *šarthašart* 'with a crash' (both AzRS), Bshk. *tinmätin* 'penny to penny' (BškRS96), Kzk. *žüzbežüz* 'face to face' (SKzkP), Trkm. *reŋbereŋ, reŋmereŋ* 'motley' (both TrkmRS), but also Kklp. *tikkemetikke* 'straight, direct' (KklpRS). Their semantics, however, tends to be as if the 'interfix' had a meaning similar to 'to' or dative. Certainly, a part of those cases is to be attributed to Pers. ﺑﻪ *be* 'to, into, onto', and therefore not partial interfixed reduplications, but *sykmasyk*, and maybe other examples that I do not know of, constitute a grey area. The Karaim word may be a full reduplication with a linking *ma* of unknown origin, like Kklp. *tikkemetikke*, but being monosyllabic, it may also be interpreted as a reduplication of the initial mora + C_2 + *ma*.

The geographical spread does not form a clear pattern. It is not so much a fragment of the above-mentioned sickle, as random points on it, without a clear connection.

2.7. Other

2.7.1. ℂ + *adan*

Kirg. **tūra** 'rightly, correctly' ◊ ***tu.p.adan.tūra*** (KirgRS)
Kzk. **žalɣyz** 'only, single, lonely' ◊ ***ža.p.adan.žalɣyz*** (KzkRS)

I am only aware of the above two examples. It is difficult to escape the association between what appears to be the interfix here, and the ablative suffix. One might wonder whether the reduplicated, closed, and interfixed initial mora of the base did not take on a life of its own in the minds of some speakers, and was reused here as a basis for comparison – or whether these two are not cases of a phenomenon similar to tmesis as in Engl. *how heinous ever it be* or, perhaps more in keeping with the emphatic nature of these formations, expletive interfixation as in Engl. *fan-bloody-tastic*. But these are just guesses, and the similarity might in the end prove purely coincidental.

2.7.2. ℂ + *ba, la,* or *ta*

Kar. **barabar** 'together' E ***ba.s.ta.barabar*** (KRPS, RKarS-Haf, CKarED)
Azer. ***daɣyn**[5] ◊ ***da.r.ba.daɣyn*** (AzRS)
Kmk. ***dagyn** ◊ ***da.r.ba.dagyn***[6] (KmkRS)
Tuv. **kara** 'black' ◊ ***ka.p.la.kara*** (TuvRS)

The four examples above are all that I know. I suspect that in reality only the two based on *dagy-* have anything in common with each other, and the

---

5 See fn. 3.
6 Apparently, only attested in: ~ *bolmak* 'to be completely destroyed, in shatters' and ~ *etmek* 'to destroy completely, to shatter'.

other two are both unrelated, isolate cases. I only put them together here for convenience, and because very little can be said about any of them. The Karaim word seems most like a genuine reduplication, but of a rather unusual type. The Azeri and Kumyk words may perhaps be purely phonetic, emphatic modifications of a better represented type with *ma* for interfix (see 2.6), only the appropriate form does not seem to be attested for Kumyk. Some guesses about *kaplakara* can be found in Stachowski K. 2014: 159.

2.7.3. $\mathbb{C}$ or C_2 + *gIn*

Yak. **ʒulagyr** 'bare, smooth' ◊ ***ʒu.l.gun.ʒulagyr*** (Pekarskij 1907–30)
Yak. **taraɣaj** 'bald, hairless' ◊ ***ta.r.gyn.taraɣaj*** (Pekarskij 1907–30)

I only know of these two examples, and am entirely at loss as to how to interpret them. It is not even clear whether the semantic similarity is of any significance.

2.7.4. $\mathbb{C}$ + *yl*

Tksh. **čyplak** 'naked, nude' ◊ ***čy.r.yl.čyplak*** (Kononov 1956: 157, Hatiboğlu 1973, Müller 2004: 87, Stachowski M. 2009: 119f)
Tksh. **syklam** 'wet ◊ ***sy.r.yl.syklam*** (Kononov 1956: 157, Hatiboğlu 1973, Müller 2004: 87, Stachowski M. 2009: 120)

Again, the above are all the examples known to me. The fact that the $\mathbb{C}$ is in both cases *r,* a fairly uncommon choice in the *C*-type in Turkish (8 out of 176 examples), suggests that it is perhaps a part of the interfix rather than a $\mathbb{C}$. If so, these would become the only examples known to me, of partial

interfixed reduplications in a Standard Turkic language, without any ℂ whatsoever. Either way, their structure remains entirely unclear to me.

2.7.5.   ℂ + *ym*

Kzk. **katynas**'relation(ship), connection' ◊ ***ka.r.ym.katynas*** (KzkRS, SKzkP)

Separating this one word, the only example of this kind that I am aware of, into its own group might be seen as rather generous. Possibly, it is merely a phonetic variant of the same type as *tälämtäläsik* in 2.6.1, but the difference in the height of the vowel, and the fact that the closer is clearly a ℂ here, cannot be overlooked. Perhaps there are more words like this one, only I do not know of them.

3. Closing thoughts

As mentioned before, the above classification must be considered provisional. But the material collected here is such that even a preliminary approach like this one brings to light several patterns and connections, and provokes more than one idea. It is not clear to me how the detailed histories of specific words can be traced, but I suspect that even without this knowledge, interesting and potentially fruitful observations can be made based on a more in depth look into the recurring phenomena (which I plan to do in the near future), and perhaps also based on a more complete collection of examples, one to which quantitative methods could be applied with a greater sense of security. Especially in light of the Dolgan and Yakut formations, I believe that it will be necessary rather sooner than later to thoroughly comb the Mongolic languages.

Abbreviations

AzRS = Tağıyev et al. 2006

Azer. = Azeri

Bshk. = Bashkir

BškRS58 = Ahmerov et al. 1958

BškRS96 = Uraksin 1996

CKarED = Aqtay/Jankowski 2015

DW = Stachowski M. 1993

Dolg. = Dolgan

Engl. = English

Gag. = Gagauz

GagTS = Baskakov et al. 1991

KRPS = Baskakov/Zajončkovskij/Šapšal 1974

Kar. = Karaim

Kar.E = Eastern (Crimean) Karaim

Kar.NW = Northwestern (Trakai) Karaim

Kar.SW = Southwestern (Halych/Lutsk) Karaim

Khal. = Khalaj

Kirg. = Kirghiz

KirgRS = Judahin 1985

Kklp. = Karakalpak

KklpRS = Baskakov 1958

Kmk. = Kumyk

KmkRS = Bammatov 1969

Kzk. = Kazakh

KzkRS = Bektaev 2001

Ott. = Ottoman

RKarS-Haf = Hafuz 1995

RKarS-Lev = Levi 1996

SKzkP = Aqtay/Jankowski 2011

Tat. = Tatar

TatRS = Golovkina/Osmanov/Denisova et al. 1966

Tksh. = Turkish

Trkm. = Turkmen

TrkmRS = Baskakov/Karryev/Hamzaev 1968

Tuv. = Tuvinian

TuvRS = Tenišev 1968

Uzb. = Uzbek

UzbRS41 = Kary-Nijazov/Borovkov 1941

UzbRS59 = Akobirov/Magrufov/Hodžahanov 1959

Yak. = Yakut

References

Ahmerov 1958 = Ахмеров К.З.. Краткий очерк грамматики башкирского языка – BškRS58: 743–802.

Akobirov/Magrufov/Hodžahanov 1959 = Акобиров С.Ф., Магруфов З.М., Ходжаханов А.Т. (ed.). *Узбекско-русский словарь. Ўзбекча-русча луғат*. Москва.

Aqtay G., Jankowski H. 2011. *Słownik kazachsko-polski*. Kraków.

Aqtay G., Jankowski H. 2015. *A Crimean Karaim-English dictionary*. Poznań.

Bammatov 1969 = Бамматов З.З. (ed.). *Кумыкско-русский словарь*. Москва.

Baskakov 1958 = Баскаков Н.А. (ed.). *Каракалпакско-русский словарь*. Москва.

Baskakov et al. 1991 = Baskakov N.A. (ed.), Gaydarci G.A., Koltsa E.K., Pokrovskaya L.A., Tukan B.P.. Gagauz Türkçesinin Sözlüğü. Ankara [transl. İ. Kaynak and A.M. Doğru].

Baskakov/Karryev/Hamzaev 1968 = Баскаков Н.А., Каррыев Б.А., Хамзаев М.Я. (ed.). *Туркменско-русский словарь*. Москва.

Baskakov/Zajončkovskij/Šapšal 1974 = Баскаков Н.А., Зайончковский А., Шапшал С.М.. *Караимско-русско-польский словарь*. Москва.

Bektaev 2001 = Бектаев Қ.. *Қазақша-орысша сөздік. Казахско-русский словарь*. Алматы.

Bodrogligeti A.J.E. 2003. *Academic reference grammar of modern literary Uzbek*. Los Angeles.

Comidas de Carbognano C. 1794. *Primi principj della gramatica turca ad uso dei missionari apostolici di Constantinopoli*. Roma.

Dmitriev 1940 = Дмитриев Н.К.. *Грамматика кумыкского языка*. Москва – Ленинград.

Doerfer G., Tezcan S. 1980. Wörterbuch des Chaladsch (Dialekt von Xarrab). Budapest.

Doniyorova S. 2008. *Parlons Koumyk*. Paris – Budapest – Torino.

Golovkina/Osmanov/Denisova et al. 1966 = Головкина О.В., Османов М.М., Денисова Н.Т. et al. (ed.). *Татарско-русский словарь*. Москва.

Hafuz 1995 = Хафуз М.Э.. *Русско-караимский словарь. Крымский диалект*. Москва.

Hatiboğlu V. 1973. *Pekiştirme ve kuralları* (= *Türk Dil Kurumu Yayınları* 372). Ankara.

Jehlitschka H. 1895. *Türkische Konversations-Grammatik*. Heidelberg.

Judahin 1985 = Юдахин К.К.. *Киргизско-русский словарь. Кыргызча-орусча сөздүк*. Фрунзе.

Juldašev 1981 = Юлдашев А.А. (ed.). *Грамматика современного башкирского литературного языка*. Москва.

Józefowicz G. 2008. *Słownik polsko-karaimski w dialekcie trockim*. Troki – Wilno – Warszawa – Wrocław – Gdańsk – Nashville.

Kary-Nijazov/Borovkov 1941 = Кары-Ниязов Т.Н., Боровков А.К.. *Узбекско-русский словарь*. Ташкент.

Kim H.-S. 2014. Intensifying Adverbial Particle or Emphatic Prefix? A Study of Intensive Adjectives in Mongolic Languages. – *Eonehag* 70, 305–328.

Kononov 1956 = Кононов А.Н.. *Грамматика современного турецкого языка*. Москва – Ленинград.

Laude-Cirtautas I. 1961. *Der Gebrauch der Farbbezeichnungen in den Türkdialekten* (= *Uralaltaische Bibliothek* 10). Wiesbaden.

Levi 1996 = Леви Б.З.. *Русско-караимский словарь. Крымский диалект. Урус-Ḳарай сӧзлӱк*. Одесса.

Meninski à Mesgnien F. 1680. *Thesaurus linguarum Orientalium Turcicæ Arabicæ Persicæ.linebreak Lexicon Turcico-Arabico-Persicum*. Vienna [2nd ed. by M. Ölmez, S. Stachowski, İstanbul 2000].

Müller H.-G. 2004. *Reduplikationen im Türkischen. Morphophonologische Untersuchungen* (= *Turcologica* 26). Wiesbaden.

Özkan N. 1996. *Gagavuz Türkçesi Grameri. Giriş – Ses Bilgisi – Şekil Bilgisi – Cümle – Sözlük – Metin Örnekleri* (= *Türk Dil Kurumu Yayınları* 657 = *Gramer Bilim ve Uygulama Yayınları* 18). Ankara.

Pekarskij 1907–30 = Пекарский Э.К.. *Словарь якутскаго языка*. Санктъ-Петербургъ – Петроградъ – Ленинград.

Pokrovskaja 1964 = Покровская Л.А.. *Грамматика гагаузского языка. Фонетика и морфология*. Москва.

Stachowski K. 2010. On First Syllable Reduplications in Karaim. – Mańczak-Wohlfeld E. / B. Podolak (ed.). *Studies on the Turkic World. A Festschrift for Professor St. Stachowski on the Occasion of His 80th Birthday*. Kraków. 139–156.

Stachowski K. 2014. *Standard Turkic C -Type Reduplications* (= *Studia Turcologica Cracoviensia* 14). Kraków.

Stachowski M. 1993. *Dolganischer Wortschatz*. Kraków.

Stachowski M. 1997. *Dolganische Wortbildung*. Kraków.

Stachowski M. 2009. Once again on the etymology of Turkish *çocuk* 'child'. – *Türkbilig* 17 (2009): 116–123.

Tağıyev M.T. et al. 2006. *Azərbaycanca–rusca lüğət*. Bakı.

Tenišev 1968 = Тенишев Э.Р. (ed.). *Тувинско-русский словарь*. Москва.

Ubrjatova 1985a = Убрятова Е.И.. *Историческая грамматика якутского языка*. Якутск.

Standard Turkic reduplications of the extended *C*-type: The different types

Uraksin 1996 = Ураксин З.Г. (ed.). *Башҡортса-русса һүҙлек. Башкирско-русский словарь.* Москва.

von Gabain A. 1945. *Özbekische Grammatik mit Bibliographie. Lesestücken und Wörterverzeichnis* (= *Porta Linguarum Orientalium* 25). Leipzig – Wien.