CHAPTER FIVE

CLASSIFICATION OF ALTAIC LANGUAGES AND DATING OF PROTO-ALTAIC

To demonstrate the genetic subclassification of Altaic we shall take the list of matches between Altaic subgroups in the realm of the basic vocabulary:

Item	Turkic	Mongolian	Tungus-Manchu	Korean	Japanese	PA
all	*büt-			*mòt-		*mŭtʿì
all	*Kop	*kow	*kupukte			*kŏpʿV
ashes		*hüne-sü	*pulńe-			*pʿoĺńe
bark	*Kāpuk			*kàph-	*kapa	*kʿáp̯ʿà
bark	*Kaŕ	*kajir(a)-				*kʿéŕà
bark	*Kɨrtɨĺ	*körü-sü	*xura-kta			*kʿi̯ŭrú
belly		*keweli	*kepel-			*kép̄ʿV
belly				*pă̄i	*pàrá	*pʿèjló
bite		*kem-			*kàm-	*kĕ̀ma
black	*Kara	*kara			*kùruà-	*kàru
blood		*či-su			*tí	*čʲúnu
bone		*ja-su		*s-pjə́	*pàniá	*pʿèjńé
breast		*čeɣeʒi	*ča(i)ʒan	*čjə́č	*tì(tí)	*čʿăjʒV
breast		*kökön	*kuku-n			*kŏ̀kʿè
burn	*jak-		*deg-ʒe-gi-	*thằ-	*dák-	*dĕ́kà
f.-nail		*kimul-su		*kom(h)		*kʿi̯uml[e]
f.-nail				*tòph	*túmá-i	*tʿi̯ŭpʿo
cloud				*kúrùm	*kùmua	*kʿòlmV
cold		*köji-ten	*xiŋ̈ü-			*kʿi̯ójŋo
cold	*sogɨ-k		*šig-			*šʲogo
come	*gẹl-				*kə̀-	*gĕ̀le
die		*bür-il-	*bu(r)-			*bŭri
dog	*ɨt		*ŋinda		*ìnú	*ŋĭ̀ndó
drink			*um(i)-	*mà-		*umV
dry	*Kūrɨ-	*kawra-			*káwá(ra)-k-	*kʿi̯óbarV
ear	*Kul-kak			*kúi		*kʿù̯jlu
earth		*siraɣu		*hằrk		*sʲắŕi

Item	Turkic	Mongolian	Tungus-Manchu	Korean	Japanese	PA
eat	*jē-		*ǯe-p-	*čā-		*ǯē
egg	*jumurtka	*ömdege	*umūkta			*úmu-tki
eye		*ni-dü	*ńia-sa	*nú-n	*mà-	*ni̯ȁ
fat	*jāg	*eɣü-kü				*i̯ā́gi
fat			*ximū-	*kìrɨ́m		*kʻi̯ǻŕme
feather	*jüg		*dekte-			*d[é]gì
feather		*hö-dün			*pánái	*pʻúńe
fire				*pɨ́r	*pə̀-i	*pʻȍre
fish		*ǯiɣa-			*(d)íwuá	*di̯ági
fly		*degde-	*deg-			*tēga
foot			*palga-n	*pár	*pànkì	*pȁlgà
full	*dōl-		*ǯalu-			*čā́lo
full			(*čak)	*čhǻ-		*čʻáko
full			*milte-		*mìt-	*mìltʻe
give	*bēr-		*bū-			*bi̯ṓr[é]
give				*tā-	*átá-pá-	*i̯ā́tá
go			*ŋene-	*nàń-	*ín-	*ŋḗni
good				*tjōh-	*də̀-	*di̯òge
green	*gȫk	*köke				
green		*nogo-ɣan	*ĺog-			*li̯ŏga
hair	*Kɨl(k)		*xińŋa-		*ká-i	*kʻíla
hair		*hü-sü	*puńe-			*pʻúńe
hand	*el		*ŋāla			*ŋȁli
head	*baĺč			*mərí		*mḗĺǯu
hear			*dōldī-	*tɨ̀d-		*tȕĺdi
heart			*miańam	*mǻńǻm		*mi̯óńù
heart	*jürek	*ǯirüke				*ǯȕr(V)kʻe
horn		*eber		*s-pɨ́r		*i̯opʻérV
I	*bẹ-	*bi	*bi		*bà-	*bȉ
I				*nà	*a-	*ŋa
know			*sā-		*sir-	*sā́ŕi
leaf	*japur-gak	*labči		*níph		*li̯àpʻ[à]
lie	*jạ-t-		*dē			*dȅ
lie		*keb-			*kəjə-	*kejbe
lie				*nū-b-	*ná-	*nḗ
liver	*biagɨr		*pākin			*pi̯ȁki
long			*ŋōli-		*nàn-kà-	*ŋȍla
long	*uŕɨ-n	*ur-tu				*i̯uŕo

Item	Turkic	Mongolian	Tungus-Manchu	Korean	Japanese	PA
louse		*sirke			*sìrám(u)i	*sɻə̀jrí
man	*ēr	*ere				*ā́ri
many				*mān(h)	*manai-	*mana
meat				*sā́rh	*sìsì	*šɻə̀lì
moon				*tā́r	*tùkùi	*tʿɻòlgu
mouth		*ama-n	*am-ŋa			*ā́mo
name		*nere		*(n)ìrh-		*nḗre
neck	*bōjn		*moŋa-n	*mjə-k	*nəmpV	*mṓjno
new			*nebi		*nípí-	*nébì
new	*jaŋɨ	*sine		*sái		*zèjńa
night			*dolba		*duà	*dŭ̀le
nose		*ka[m]ar	*xoŋa-	*kóh		*kʿɻŏŋa
not			*ā(n)-	*àn-	*nà-	*ànì
not		*e-se	*e-			*e
one	*bir				*pitə	*bɻuri
rain	*jag-			*pí		*pʿɻə̀ge
red		*hula-ɣan	*pula-	*pɨ̀rk-		*puli
road		*mör-			*mítí	*mɻóri
root		*ündü-sü	*ŋū̃ŋte		*mətə	*ŋɻŭŋtʿè
root		*hiǯa-ɣur	(*puǯuri)			*pɻoǯi
root			*pule-	*pɨrhā́i		*pʿŭ́li
round			*murV		*márə́	*múra
round			*toŋal-	*toŋkor-		*tʿòŋké
round		*tob-			*tùmpú-ra	*tʿòpʿú
round	*deg-	*tögörig				*tegá
sand				*mòr'aʿi	*mana-n-kua	*mā́ro
say		*kele-		*kə̀ró-		*kʿɻə̆́li
seed	*urug	*hüre				*pʿŭ̀ri
sit		*saɣu-			*súwá-	*šɻábu
skin				*kàph-	*kapa	*kʿā́pʿà
skin	*jīn		*nansa			*nā́ne
sleep	*ū-		*ŋuja		*úi-	*ŋū́ju
small		*öčü-	*ŋüši-			*ŋṓjčʉ
small		*ǯiǯig	*nisi			*ńĭ̆ŋči
stand	*dur-				*tàt-	*čŭ̀ra
star	*jul-duŕ	*ho-dun		*pjīr	*pə́sí	*pʿɻṓĺo
stone	*diāĺ	*čila-ɣu	*ǯola	*tōrh	*(d)ísì	*tɻṓĺì
sun			*sigū-n	*hā́i		*sɻṏ̆gu

Item	Turkic	Mongolian	Tungus-Manchu	Korean	Japanese	PA
swim		*oji-mu-			*ə̀jə̀nk-	*òje
tail	*Kudruk		*xürgü	*s-kòrí		*kˊi̯údo
that	*Ti-	*te-re	*ta-	*tjə́		*tˊa
that		*ča-	*čā-			*čˊa
this		*e-ne	*e-			*é
this	*gə̄			*ki	*kə́-	*kō
thou	*sẹ-		*si		*si	*si
thou				*nə̀	*ná	*ná
tongue		*kele	*xilŋü			*kˊi̯ălі
tongue				*hjə́	*sìtá	*si̯ə́jri
tooth				*par	*pa	*pala
tooth	*sīĺ	*si-dü				*sīĺa
tree		*mo-du	*mō			*mű́ro
two		*ǯiw-rin	*ǯube	*tubu		*ti̯ubu
warm	*jɨli-g	*dula-ɣan				*di̯ù̆lu
warm				*tă̆-	*àtà-	*ōtˊa
water			*mū	*mír	*mí-	*mi̯ù̆ri
water	*sɨb	*u-su				*ši̯uba
we	*bi-ŕ	*ba	*bue	*ú-rí	*bà-	*bĭ̆-
what	*nV	*ja-ɣu-			*nV̀	*ŋi̯V
white	*siarɨg			*hắi-	*sìruà-	*si̯ă̆jri
white		*čaga-ɣan	*šāk-			*šằkˊa
who	*kem	*ken	*xia			*kˊa(j)
who			*ŋū̄	*nú-		*ŋi̯V
woman		*eme		*ámh	*mía	*ə̌me
far	*ɨra-				*pàrú-ka	*pˊìrá
heavy				*mì-	*ə́m(p)ó-	*ámbe
near	*jạgu-		*daga			*dắgá
near	*jAkɨ-				*tù̀kà-	*dă̆kˊì
salt	*dūŕ	*dabu-su				*čo̯obeŕV
short		*hokar	*poKa-			*pi̯ŭ̆kˊì
snake		*mogaji	*mū̄kū̄			*mi̯ūko
snake				*pắjàm	*pàim(p)V̄	*pˊ[ò]jamV
thin				*kắnắr-	*kə̀má-	*kˊə̌ńó
thin		*nari-n	*ner-	*jɨ̄r-p-		*nèra
thin		*nim-gen	*niambu-			*ni̯ombu
wind	*jẹl	*sal-ki				*zălі
worm	*Kūrt	*koro-kai				*kˊo̯ó́ro

Item	Turkic	Mongolian	Tungus-Manchu	Korean	Japanese	PA
year		*oj	*aṅŋa			*ā́ṅu
year	*jɨl	*ǯil			*tə̀sì	*dǐlo
year			*sē	*sɨ̄r		*zēra

This list is basically the same as given in Starostin 1991 (pp. 25-63, 85-104), but with some additions and corrections added during the years of work on the Altaic dictionary, which have consequently resulted in some calculational changes, albeit statistically insignificant. The average percent of matches revolves around 20, which gives us the date of split of Proto-Altaic at around the end of 6th millennium B.C. We see an increase up to about 25% between Turkic, Mongolian and TM, and an increase to 33% between Korean and Japanese, which would speak in favour of two basic subbranches of Altaic.

However, if we look at the figures in more detail and take into account the division between 35 more stable items and 65 less stable items, proposed by S. Y. Yakhontov, the picture appears to be somewhat more complicated.

Language pairs	Matches in the standard 100 w.-list	Matches in Yakhontov's modified 100 w.-list	Matches in Yakhontov's 35-wordlist	35/65 w.list ratio
TuMo	25	24	11	31 / 20 : 1.55
TuTM	25	22	10	29 / 18 : 1.61
TuKo	17	13	5	14 / 12 : 1.17
TuJap	19	19	7	20 /18 : 1.11
MoTM	29	30	11	31 / 29 : 1.07
MoKo	18	17	8	23 / 14 : 1.64
MoJap	22	17	9	26 /12 : 2.17
TMKor	23	23	9	26 / 22 : 1.18
TMJap	22	20	8	23 /18 : 1.28
KorJap	33	30	11	31 /29 : 1.07

This chart shows us that while the overall 35 / 65 wordlist ratio is > 1 in all cases (the situation which indicates genetic relationship, meaning that the rate of matches within the most stable 35 word range is higher than the rate of matches within the less stable 65 word range), in two cases - Mongolian-TM and Kor.-Jpn. - it is dangerously close to 1. Lexicostatistically this may indicate the borders of ancient dialect zones within Proto-Altaic, suggesting that Tungus-Manchu does not really

constitute a unity with Turkic-Mongolian, and throwing some doubts on the genetic unity of Korean-Japanese.

Additionally it provides an explanation why the genetic situation within Altaic looks somewhat different from that within, e.g., Indo-European - which was the main reason for the whole anti-Altaic criticism. Altaic appears to be different from Indo-European in two main respects:

a) it is somewhat older than Indo-European: while the split of Proto-Indo-European (or Proto-Indo-Hittite) can be dated to the 4th millennium B.C., the split of Proto-Altaic must have occurred at least a thousand years earlier;
b) whereas the subbranches of Indo-European are rather old (e.g., Balto-Slavic may be dated around the beginning of the 1st millennium B.C, and Indo-Iranian - by the verge of the 2d and 3d millennia B.C.), the subbranches of Altaic are considerably younger: Turkic - beginning of our era, Mongolian - around the 10th century A.D., Tungus-Manchu - around the 4th century B.C., Japanese - around the 5th century A.D., Korean - around the 11th century A.D. Furthermore, whereas the oldest texts in ancient Indo-European languages are attested quite early, the written monuments of Altaic languages date back not earlier than to the 8th century A.D. This all creates an impression of a much more distant genetic relationship between Altaic languages than that between Indo-European languages. Consequently - basically due to the absence of archaic attestation - Proto-Altaic is somewhat more difficult to reconstruct than Proto-Indo-European. It is also quite possible that some other branches of Altaic might have existed, but they could have been wiped out by later waves of migrations of Altaic (and non-Altaic) languages.

The overall distribution of lexical isoglosses confirms the above classification, but yields some additional information on the dialect distribution within Proto-Altaic.

The core of the common Altaic vocabulary is constituted by etyma reflected in Turko-Mongolian (at least in Turkic or Mongolian) and in Korean-Japanese (at least in Korean or Japanese) - with or without Tungus-Manchu parallels. The number of such roots within the present volume is 1841; most of them (1553) are also reflected in Tungus-Manchu.

There are, however, two other lexical groups:

a) Turkic-Mongolian roots with Tungus-Manchu parallels (615); without Tungus-Manchu parallels (57)

b) Korean-Japanese roots with Tungus-Manchu parallels (195); without Tungus-Manchu parallels (23)

We see thus that Tungus-Manchu occupies a specific position within the Altaic family, sharing a large number of isoglosses both with the Turkic-Mongolian and the Korean-Japanese branches. In the dictionary we call the former "Western isoglosses" and the latter, "Eastern isoglosses". Historically such a situation may be explained by a "central" position of Tungus-Manchu among the Proto-Altaic dialects. At the same time, one can also not exclude a scenario of later prehistorical borrowings (already after the split of Proto-Altaic, but before a wide geographical separation of Turko-Mongolian and Tungus-Manchu, on one hand, and Tungus-Manchu and Korean-Japanese, on the other).

The position of Tungus-Manchu within Altaic thus resembles the position of Greek within Indo-European: due to its original "central" geographic location, Greek shares a large number of isoglosses both with European languages (primarily Italo-Celtic) and with Indo-Iranian languages.

The Proto-Altaic nature of almost a thousand "Western" or "Eastern" isoglosses mentioned above is questionable; nevertheless we decided to include them into the dictionary because potentially any of them may turn out to be common Altaic - there is always a chance that, e.g., a Korean match for a Western isogloss exists but has not yet been recovered, or a chance that an archaic etymon reflected in Korean-Japanese has been lost in Turko-Mongolian. We must stress that phonetically and morphologically these isoglosses behave just like all other common Altaic roots.

To sum up: Proto-Altaic split into three branches, viz. Turco-Mongolian, Tungus-Manchu and Korean-Japanese, around the 6th millennium B.C. Tungus-Manchu must have occupied a central dialectal position, which explains its shared isoglosses both with Turko-Mongolian and Korean-Japanese.

Two subbranches - Turko-Mongolian and Korean-Japanese - in their turn, had split rather early, around the 4th millennium B.C. However, reconstructing Proto-Turko-Mongolian or Proto-Korean-Japanese would in fact be almost equivalent to reconstructing Proto-Altaic, because of small time span separating those units from the original protolanguage. There is still some doubt about the existence of common Korean-Japanese: the specific similarities between these two subbranches might be due to secondary dialectal interaction.

The next splits occurred already closer to our era: first the split of Tungus-Manchu, next the split of Turkic, Japanese and Korean dialects.