

Var mı, yok mu? ("Does it or doesn't it exist?"): The Altaic dilemma (or: Aru, nai?)

Timothy Ian Mills  
University of Calgary

## 0 INTRODUCTION: WHAT IS A LANGUAGE FAMILY?

This paper is an overview of arguments for the relatedness of the two languages demonstrated in the title - Turkish and Japanese - and the series of language groups between the two, including the rest of the Turkic languages, the Mongolian and Manchu-Tungus families, and Korean, the close sister of Japanese. Before diving into the specific arguments for and against relatedness of the languages in question, it may be beneficial to think about what we mean by a language family.

We as linguists have a fairly rigorous way to define the spoken language of a person in a systematic way. After that point, we tend to get confused. What is a language, as spoken by a community? How is it different from a dialect? A *language family* is generally defined as a group of one or more languages, all directly derived from a single common historical language by means of natural historical-linguistic changes. Out of this definition, we can build methods of determining members of a common linguistic heritage - a language family - even when the parent language is no longer spoken in its original form. The problems with the definition of language families are that, first, the definition of a *language* is problematic at best, and second, not every language is derived from a single previous language. Take the case of creoles: most people, Bickerton (1990), Lefebvre and Lumsden (1990) among them, agree that Haitian Creole was born of both Fon-Ewe *and* French. Whatever the details behind the birth of this creole, it indisputably arose from processes as natural as those that generate other languages, such as English, from previous languages, such as proto-Germanic. If you agree with Lefebvre and Lumsden that this creole is simply the result of an intense case of lexical borrowing (coupled with the phonological simplification characteristic of pidgins and creoles), where the words of the superstratum language are imposed on the syntax and morphology of the substratum language, then it isn't even remarkably different from English, in which there is a high proportion of French words superimposed on the Germanic core. With respect to the family discussed in this paper, Japanese shows evidence of being an ancient creole created by contact between an Altaic language and an Austronesian language (Shibatani 1990). Even if this were true, though, the Altaic components are meaningful to an analysis of the Altaic language family.

Even given the uncertainty of creoles, though, we are confident that, generally, members of a language family can be said to be derived from one *core* language. The process of divining which languages are members of a family is still confounded by the fact that nothing can be actually tested and proven (or disproven). We can only gather evidence from all the various traces that a past civilization leaves - biological traces in the genes of descendent peoples, archaeological traces of migrations from a common point, cultural traces such as religion, and linguistic traces in their speech. As linguists, we approach this problem with a focus on the speech of the modern descendants of the

ancient unity, but we must always keep an eye on the other aspects. To ignore a crucial piece of evidence just because it is not within one's field of specialty is to devalue the entire endeavour, which is, at best, already a matter of very educated guesswork.

With this in mind, I will introduce the topic of this paper. Among historical linguistic studies, there is a great range of theories - from the firmly and almost universally accepted Germanic family, which includes mainly the German languages, Scandinavian languages, and English, to the almost universally denied Nostratic, a language family purported to reach across Europe, Asia, and India. Somewhere between these two extremes of certainty is Altaic. The Altaic family is a hypothesized genetic unity including the subfamilies of Turkic, Mongolian, and Manchu-Tungus, as well as the fringe languages of Japanese and Korean. The great geographical expanse of these languages encourages scepticism until one considers that Indo-European is said to reach from Icelandic to Hindi<sup>1</sup>. The origin of the Altaic family has been established in the vicinity of the Altai Mountains, from which the family gets its name. Menges (1968) gives a more thorough description of this origin, and the issue of the name "Altaic" (see section 5 of this paper). The initial divergence of the Altaic languages can be traced back about four millennia, to the beginning of the second millennium B.C., so there have been plenty of time effects and language contacts to obscure the nature of the original Altaic language and make the investigation of that language all the more interesting.

Having had less attention in studies of historical linguistics than Indo-European, the existence of the Altaic family is still hotly debated among scholars. Individuals such as Sir Gerard Clauson and G. Doerfer deny its existence outright and strive to explain the correspondences among the given languages away in other ways. Sceptics like Robert Austerlitz and A. Róna-Tas withhold judgement on complaints that reliable evidence is far too scant to conclude anything yet. Others, such as Nicholas Nicholai Poppe and Roy Andrew Miller, have applied modern comparative methods to the problem and decided that a genetic unity does, indeed, exist, and all that remains to be done is to flesh out the nature of that unity and, eventually, reconstruct the proto-language that birthed the modern descendants.

With that, an overview of the issues involved in the Altaic debate is in order.

## 1 OVERVIEW

The languages of the Altaic family have been alternately proposed as members of various far-flung families. Shibatani (1990) lists a number of theories which have received varying levels of acceptance, all suggesting where Japanese came from. The list ranges widely: the Altaic connection, with which this paper is principally concerned; the Malayo-Polynesian connection; the Indo-European connection; and other ideas connecting Japanese with Sumerian, Greek, or others. The Altaic and Malayo-Polynesian (or Austro-Asiatic) hypotheses seem to be the most widely accepted, leading to two more possibilities that Shibatani mentions: one is that Japanese "consists of an Austronesian

<sup>1</sup> Miller (1991a & 1991b) makes much of the fact that the opponents of the Altaic hypothesis ignore parallels to the well-established and accepted Indo-European.

substratum and an Altaic superstratum"; and the other is that it is a hybrid between these two families - something like an ancient version of the creoles mentioned earlier. This is supported by Loveday (1996), who only regards two hypotheses as being very viable: 1) that Japanese emerged from a pidginization and creolization environment when proto-Austronesian met proto-Altaic; and 2) that Japanese has a strong Altaic base, with a large number of Austronesian borrowings early on. Shibatani (1990:105) lists a few isoglosses which prove that, at the very least, historical contact resulted in some degree of borrowing from Austronesian into Japanese.

Alternative explanations of the origin of the other Altaic languages - Turkic, Mongolian, Manchu-Tungus - are rare. Anti-Altaic consensus seems to be that cognates across these languages are borrowed from a central *proto-Turkic* (which appears, coincidentally, to have much in common with the reconstructed proto-Altaic, according to Miller's 1991 paper). This is problematic because of the great geographical range of the family - from Turkey to the exceptionally isolated islands of Japan. Evidence of peripheral retention also highlights the anti-Altaic arguments as unnecessarily complex. So for now, the only hypothesis (and, as Miller (1991a) points out, such historical reconstructions must always remain hypotheses) that exists to be seriously considered, at least with respect to Turkic, Mongol, and Tungus, is the Altaic hypothesis, and the details of reconstruction pertaining to that.

### 1.1 Phonology and the Lexicon

In order to figure out the system of phonological correspondences between members of a language family, a significant base of cognates must be known. To establish a significant base of cognates in such an old language family as Altaic, some pattern of phonological correspondences must be known. This "chicken-and-the-egg" problem can be daunting, and it has, indeed, left its share of casualties in the form of sceptics and doubters along the development of the Altaic question. However, with a careful combination of inductive and deductive reasoning, study of the Altaic languages has brought forth a reliable set of correspondences from which a larger common lexicon can be hunted down, which will lead to a more thorough understanding of the phonological relationships.

This problem of cognate scarcity is daunting enough, considering the geographical expanse that the family covers. But on top of that is the fact that many of these cognates can be demonstrated to be more recent borrowings between the languages. However, by a concentrated effort to focus only on the (admittedly problematic) core words - kinship terms, body parts, and so forth - this smokescreen of borrowings can be at least somewhat cleared up. **Section two** deals with the various phonological and lexical arguments as put forward by both the Altaicists and the anti-Altaicists.

### 1.2 Morphology

Though lexical items are relatively easy to borrow, affixes are not. For this reason, morphological correspondences can provide very firm evidence for genetic affinity. Since phonetic changes tend to warp affixes greatly, a good basis of phonological

changes (**section two**) is very helpful before diving into the quagmire of determining morphological cognates. There are a number of such cognates that have been demonstrated among the Altaic languages. These will be discussed in **section three**.

### 1.3 Syntax

Unfortunately for comparativists, not all areas of linguistic typology can give us reliable backing in proving or refuting the existence of a language family. Syntax is one of those questionable areas. Just because two languages show a head-final, SOV word order doesn't necessarily mean they are related. The fact that the Indo-European languages generally demonstrate SVO word-order is probably due more to the fact that they have had a lot of areal contact than the fact that they are historically related, according to Ritter (pers. com.). Similarly, though the Altaic languages tend to be SOV, this by itself does not support the hypothesis. **Section four** addresses this problem, but only briefly because the syntax is neither a good argument for nor against the relatedness hypothesis.

### 1.4 Non-linguistic arguments

No language develops in isolation from other factors. Speakers of descendent languages are almost always descendants of the speakers of the original language. Migrations are subject to geographical influence, and they leave their mark on the land as well as the cultures and languages they come in contact with. Evidence such as genetic comparisons between speakers of languages, archaeological proof of migration, and so forth help us to support or refute statements about the linguistic history of languages. For these reasons, linguists must be able to look beyond the language of a people to the whole picture, including all of these factors in the establishment of a viable hypothesis of language genesis. **Section five** will provide an overview of these issues and what they tell us about the origins of the Altaic languages.

## 2 PHONOLOGY AND LEXICON

### 2.1 Phonology against lexicon

When dealing with such an ancient linguistic unity as Altaic, lexical cognates can be quite obscured by phonological shifts. They have also become more scarce, due to large amounts of loanwords, from within and from without the Altaic sphere. The change can be seen, in fact, in the title to this paper. The Turkish phrase and the Japanese phrase are literal translations of each other, and happen to contain two cognate words - the existential affirmative and negative. Japanese has (at least in the informal register used here) lost the question particles, which may not be directly related to the Turkish ones after all, and other phonetic events have further twisted the similarities until, to the untrained, or even reasonably sceptical eye, the phrases bear little resemblance other than meaning. The Altaic roots of the "*var*"/"*aru*" and "*yok*"/"*nai*" isoglosses shown by Miller (1971) are given in the morphology section of this paper. In order for Miller and colleagues to derive those morphological tendencies, though, a system of phonetic

correspondences was required. In order for a system of phonetic correspondences to be found, a sufficient base of cognates had to be found. To find these cognates, some understanding of the phonetic correspondence was needed. Among the 'casualties' that this dilemma has caused is the eminent Sir Gerard Clauson, whose disillusionment is described in Miller (1991a) as being a result of his inability to read an ancient Mongol text with his thorough knowledge of modern Turkish. However, once individuals such as Ramstedt, Poppe, and later Miller (1971) had plowed through massive amounts of data and come up with some correspondences, all that truly remained to pin down were the details. Some of those details are described below.

## 2.2 Cognate scarcity

As a result of phonetic change, borrowings, and semantic shift muddying the waters of comparison, Altaicists must be especially careful in selecting their notes of evidence to prove the ancient unity. The Altaic sceptic, Róna-Tas, introduces his 1975 article with a list of reasons why a given word may be cognate between two languages:

1. Historical contacts, 2. Areal convergencies, 3. Typological parallelisms, 4. Convergencies of independent origin, 5. Chance, 6. Genetic relationship.
- Our approach to the correspondences due to genetic relationship can be scientifically justified only if we go ahead and remove the correspondences caused by the first five other reasons. (Róna-Tas 1975:201)

This list is quite similar to one proposed by Doerfer, an anti-Altaicist, which is reviewed in Miller's summary of the Altaic debate (1991a:300-301). Doerfer's reasons are essentially the same as those of Róna-Tas, except that he neglects to mention the typological parallelisms - such as the fact that agglutinating languages will tend to have vowel harmony, as is exemplified in the Altaic family. The claim by Róna-Tas that we must rule out all other possibilities before we can say for certain that a word is a legitimate historical cognate is valid; because nothing is absolutely certain in the field of historical reconstruction, it is important to temper the possibility for random mistakes by being as critical as is practical. The scarcity of cognates we have, due to the expanse of time over which the Altaic languages have been diverging, is enough of a problem - we do not want to pollute it with an incautious analysis of those few that do exist.

## 2.3 Issues of phonological correspondence

The practice of establishing cognates between languages is not simply a matter of finding words that sound similar and have similar meanings. In order to rigorously defend the relatedness of languages, a *regular* system of phonetic correspondences must be found among words from the different languages. Except for a few pseudo-linguists who throw out wild proposals on such newsgroups as R. F. Hahn's well-intentioned *AltaiNet*, most linguists will agree that this is necessary when reconstructing a proto-language. However, regular phonetic correspondences are not enough. Róna-Tas (1975) demonstrates this in a specific investigation of a word borrowed from Mongolian to

Chuvash since the divergence of the Altaic languages. It is daunting to try to approach the matter of comparison in the face of the possibility that any cognate may be a borrowed word. However, given some set of rules by which we can sort out borrowed words from true historical cognates, this smokescreen that is a favourite argument of anti-Altaicists can be cleared somewhat.

One of these rules is the idea of *core words* - words that are less susceptible in languages to borrowing than others. Among these words are terms for body parts, which Miller (1991a:296) assures us constitute one of the most reliable sources of cognates among the Altaic languages. These words are generally used much more frequently than other words, and also tend to be among the closed-set lexical categories: prepositions, pronouns, determiners, and such. Among Altaic languages, pronouns are especially handy. Not only do they demonstrate phonological correspondences, but they reflect a syntactic peculiarity of proto-Altaic, as demonstrated by Miller (1971:155-178), that further cements the argument of historical relatedness.

Another way of winnowing out borrowings from true Altaic cognates is to refer to the concept of *peripheral retention*, as Miller (1991a:305) does. The basis is that recent history has shown that peripheral forms of a language will tend to preserve older forms, while the inner languages will be more innovative. The argument for comparative reconstruction is that, if you see a common form at the geographical edges of a language family, and a different form - often more innovative than the other - exhibited in the inner languages, then it is more likely to be derived from a common ancestral form, later changed in the central region of the family, than to have been borrowed from one end straight to the other end of the geographical area. An example of peripheral retention in Altaic is the four-liquid system of proto-Altaic and its reflexes in the modern languages. This will be discussed in the next section.

#### 2.4 List of phonological correspondences

Miller (1971) has summarized the reconstruction work so far in the Altaic languages, especially with respect to Japanese. In his appendix, he lists the known information about the proto-Altaic phonetic system and the reflex forms in the descendent languages. He claims that proto-Altaic had a system of nine vowels with a two-way length contrast and eighteen consonants. This set of phonemes is shown in table 1.

Table 1: The phonetic set of proto-Altaic.<sup>2</sup> (adapted from Miller 1971:49-50)

p	t	tʃ	k
b	d	dʒ	g
	s		
m	n	ɲ	ŋ
	l <sub>1</sub> r <sub>1</sub>	j	
	l <sub>2</sub>	r <sub>2</sub>	

i	y	ɯ	u
i:	y:	ɯ:	u:
e	ø	ɛ	o
e:	ø:	ɛ:	o:
a			
a:			

Note: The precise phonetic nature of the l<sub>1</sub> and r<sub>1</sub> is uncertain.

Of these consonants, most are unremarkable, having undergone little change besides some standard processes of lenition, gliding, and palatalization in the various Altaic languages. The \*p in initial position has become a fricative in many languages, including Japanese and Turkish. It has remained as p in other languages - specifically the Tungus languages of Goldi, Olcha, and Orok; and Korean. The Japanese, Korean, Manchu, and Mongolian examples in (1) show their reflexes (Old Japanese *f*, Korean *p*, proto-Korean-Japanese \*p, Manchu *f*, Mongolian *ø* from proto-Altaic word-initial \*p):

- (1) pA \*p<sub>yl</sub> > Mo. *ylije-* 'blasen', Ma. *fulgije-* 'id.', K *pul-*, MK *pil-* 'blow', OJ either *Fuk-* 'blow' (pKJ \*p<sub>ol</sub>g-) or *Fur-* '(rain, snow) fall, come blowing' (pKJ \*p<sub>or-</sub>), or possibly even with OJ *Fir-* 'break wind'. (Miller 1971:52)

The two velar phonemes, \*k and \*g, have remained virtually unchanged, except that in Mongol and Turkish they have split into a fronted pair and a back pair of allophones, corresponding to the tongue position of the immediately following vowels (Miller 1971:52).

More interesting is the problem of the unique four-liquid contrast that has been established for proto-Altaic. Early research in this field revealed a correspondence between Chuvash, Mongol, and Manchu-Tungus forms with l and Turkic, Korean, and Japanese forms with š or s. Some of these reflexes are shown in examples (2) to (4).

- (2) pA \*dal<sub>2</sub>i- > Mo. *dalda* 'heimlich, verborgen', MMo. *dalda* 'Schirm, Schutz', Ma. *dali-* 'verdecken, die Sicht verdecken, verbergen, verheimlichen', Ev. *dal-* 'bedecken', OT *yafur-* 'verdecken, verheimlichen', OJ *yasi-ro* 'enclosure for worship of native deities; later, buildings erected within that enclosure'. (Miller 1971:116)
- (3) pA \*al<sub>2</sub> ~ \*pal<sub>2</sub> > Uig. *aſuq* 'to hurry, be quick', Tar. *aldıra* id., OJ *Fasir-u* 'run', K *pal* 'foot', pKJ \*vaſyi or \*(p)afyi (Miller 1975:166)

<sup>2</sup> To be notationally accurate, all of these phonemes would be asterisked as hypothetical. The asterisks have been omitted here as they would unnecessarily clutter the diagram.

- (4) pA *\*bal<sub>2</sub>-* > Tk. *baŋ* 'Kopf', Chu. *pus'* 'Kopf, Anfang', Go. *balča, balja* 'Kopf, Gesicht', K *məli* 'Kopf', OJ *Fasi-ra* 'main pillar, support of a building' (-*ra* in OJ is a locative suffix) (Miller 1971:118)

Likewise, the languages of Chuvash, Mongol, and Manchu-Tungus (the "inner languages" when discussing peripheral retention in the Altaic family) demonstrate an *r* where the others (the "outer languages") have *z* (or, in the case of Japanese, other alveolar consonants). Examples (5) and (6) exemplify this correspondence.

- (5) pA *\*omor<sub>2</sub>u* > Tarj, Turk *omuz* 'shoulder', Chu. *āmār* 'chest, breast', Khak. *omury* 'front part of the chest (of horses)', Mong. *omuruyu(n)* 'sternum, clavicle; breast', OJ *omotaka-* 'horse with head held high' (Miller 1975:162)
- (6) pA *\*nar<sub>2</sub>* > Mo. *nirai* 'fresh, new, newborn', Ma. *narxun* 'green, fresh, new', Chu. *s'ur* 'spring (season)', OT *jaz* 'spring (season)', OJ *natu* 'summer', K *yəlim*, MK *nyəlim, nyəlom* 'summer', pKJ *\*nyəlom* (Miller 1975:159)

Comparing the reflexes of *\*l<sub>2</sub>* and *\*r<sub>2</sub>* in the modern languages to the reflexes of *\*l<sub>1</sub>* and *\*r<sub>1</sub>*, we see that there must have been a contrast, even though the nature of that contrast (allophonic or phonemic) and the precise phonetic nature of *\*l<sub>2</sub>* and *\*r<sub>2</sub>* is not known. Example (1) above shows how *\*l<sub>1</sub>* (often, as here, transcribed simply as *\*l*) has descended unchanged into the modern languages, allowing for the fact that Japanese and Korean have each developed a system of only one liquid. Examples (7) and (8) show a parallel example for *\*r<sub>1</sub>*.

- (7) pA *er* > Mo., MMo. *ere* 'Mann', Chu. *ar* 'id.', OT *er* 'Mann, Gatte' J *aru, are* 'that one' (Miller 1971:122)
- (8) pA *\*āzur-* > pKJ *\*turxye*, MK *tulh* 'two', J *ture* 'companion' (Miller 1971:122)

This system of four liquids is summarized in table 2, along with the standard transcription symbols for the proto-Altaic forms. Please note that, though *\*l<sub>1</sub>* and *\*r<sub>1</sub>* are considered to be phonetically a standard liquid and rhotic, the phonetic nature of *\*l<sub>2</sub>* and *\*r<sub>2</sub>* are by no means certain - ideas range from phonetically parallel forms such as Doerfer's *\*ljε* and *\*rjε*, cited by Miller (1991a:315) to something like *\*ʎ* for *\*l<sub>2</sub>* and *\*ʂ* for *\*r<sub>2</sub>*.

Table 2: Altaic four-liquid reflexes. (from Miller 1971:122)

proto-Altaic	Tk.	Chu.	Mo.	Tg.	pKJ	K.	J.
<i>*r<sub>1</sub></i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>*r(x)</i>	<i>l(h)</i>	<i>r</i>
<i>*r<sub>2</sub></i>	<i>z</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>*r</i>	<i>l</i>	<i>r, t</i>
<i>*l<sub>1</sub></i>	<i>l</i>	<i>l</i>	<i>l</i>	<i>l</i>	<i>*r</i>	<i>l</i>	<i>r</i>
<i>*l<sub>2</sub></i>	<i>ʂ</i>	<i>l</i>	<i>l</i>	<i>l</i>	<i>*ʂ</i>	<i>l(s, h)</i>	<i>s(i)</i>



At first it was thought that perhaps the phonetic processes of *lambdacism* (changing from  $\xi$  or  $s$  to  $l$ ) and *rhotacism* (changing from  $z$  or similar sounds to  $r$ -like sounds) were taking place. Alternately, it was proposed that the reverse processes of *sigmatism* (the reverse of lambdacism) and *zetacism* (the reverse of rhotacism) may more accurately describe the alternations seen. Subsequent work, notably that of Talât Tekin (1975), has established that the latter is more firmly supported by the evidence. The precise relationship between the  $l$  and the  $r$  that became sigmatized (or in some instances, *shigmatized*) and zetacized and their counterparts which did not is still uncertain, as noted by Miller (1996). The symbols  $l_2$  and  $r_2$  are used on the basis that their phonetic nature is probably fairly closely related to that of their counterparts,  $l_1$  and  $r_1$ , whose reflexes in the descendent languages are pretty uniformly just  $l$  and  $r$ , respectively. However, beyond this somewhat general description that  $*l_2$  has something to do with  $l$  and  $*r_2$  had something to do with  $r$ , nothing concrete is known.

Despite the details remaining to be nailed down, though, the matter of zetacism and sigmatism remains a quite convincing argument for the existence of an earlier Altaic unity. Just the fact that two languages such as Turkish and Japanese both preserve evidence of a previous four-liquid contrast - a rare occurrence in the world's languages - is enough to make us take note. Adding to this the almost trivially straightforward correspondences that Miller (1971) and others have established for the other consonants, as well as other evidence referred to in this paper, it is difficult to believe otherwise than that these two languages, along with the numerous others between them, are part of an ancient linguistic unity.

## 2.5 Conclusion - validity of correspondences

Taking the example of the four-liquid system (whose precise phonetic details are uncertain but not necessary to demonstrate relatedness), we can show that peripheral retention has given us a secure morsel of evidence in support of the Altaic theory. If we accept that the languages are historically related, we can imagine a situation where, as the languages began to diverge geographically, carrying with them a phonological system which contrasted two  $l$ -like sounds and two  $r$ -like sounds, a shift began at the core of the language which neutralized the contrast, collapsing these contrasts into one  $l$  and one  $r$ . This innovation spread out, stopping at the edge of the historical Turkic languages (we can see that the Turkic fringe language of Chuvash experienced this innovation, whereas the other Turkic languages did not) and at the Manchu-Tungus languages at the other end. This left the periphery - the rest of the Turkic languages, and proto-Korean-Japanese - with the contrast, which then phonetically diverged in processes that have been labelled 'zetacism' and 'sigmatism', *preserving the contrast*. The alternative, which the anti-Altaicists espouse, is that the proto-Turkic language, from which all "Altaic" borrowings are said to originate, lent words into proto-Mongol, which then passed them on to proto-Manchu-Tungus, and from there proceeded to proto-Korean, and was somehow passed across to the islands to end up in proto-Japanese. Then, by some incredible coincidence, the intervening language families merged these four liquids into two, and the two ends - the source of the borrowings and the farthest-removed languages from it - developed parallel forms different from that. This is a perfect example of why

we invoke the idea of peripheral retention to defend the existence of a language family, where in its absence all manner of complex nonsense would be proposed to account for the forms. From these correspondences - listed in table 2 and exemplified in items (1) through (8) - we can define the periphery and core of Altaic with respect to the collapse of the liquid system as follows. The periphery - the languages preserving the four-way contrast - consists of the Turkic languages except Chuvash<sup>3</sup>, and Japanese. The core, or inner languages, are Chuvash, the Mongol languages, Manchu, Goldi (a Tungusic language), other Tungusic languages (not demonstrated here), and Korean. The fact that we can delineate the periphery with respect to this change - with allowances for the fact that the border between inner and outer languages may have had subsequent influence back and forth - is strong evidence for the relatedness of the entire range of languages here.

### 3 DISCUSSION OF MORPHOLOGICAL ISSUES

#### 3.1 Importance of morphology.

Morphology, as was mentioned above, can be a significant factor in telling whether or not languages are related. Of the six sources of lexical cognates listed by Róna-Tas (1997:201), chance and genetic affinity are the only two which are at all likely to apply to morphological cognates. There are exceptions to this, of course, but these are far more rare than the borrowing of lexical items. Miller (1991a:308-9) praises Ramstedt for his work on establishing a list of morphological correspondences between the Altaic languages. This does, indeed, bode well for the Altaic hypothesis.

#### 3.2 List of established morphological correspondences.

In his 1971 book examining the Altaic linguistic family in light of Japanese correspondences, Miller lists a number of significant morphological correspondences. The support this gives for the Altaic hypothesis is undeniable. Two examples are the morphemes of existence, shown in examples (9) and (10), both from Miller (1971:37,284):

- (9) pA \**er-* > OT *bar-*, J *ar-u* affirmative existential: 'to have, to be, to exist'  
(10) pA \**ja-k-* > OT *joq*, J *na-k-* negative existential: 'to not have, to not be, to not exist'

From the OT *bar*, we get the modern Turkish *var*, which in the title of this paper corresponds to the Japanese *aru*. The OT form *yoq* is the modern *yok*, which parallels the Japanese *nai*, derived from *na-k-*.

Miller also gives an involved description connecting Japanese interrogatives to the other Altaic languages. Miller's argument centers around the fact that both Turkish and

---

<sup>3</sup> Example (4) seems to be an exception to the evidence that Chuvash is a *core* language. Perhaps this is from influence from the other Turkic languages.

Japanese show reflexes of the proto-Altaic "what?" morpheme, *\*ja:(-n-)*. From this form, Old Turkish *ne, nen* (Turkish *ne*) and Japanese *nani, nan-* are descended, with the same meaning (Miller 1971:196-8). The incorporation of these and other question morphemes into words of the descendent languages is the subject of an extended discussion by Miller. He shows how the proto-Altaic *\*ja:* form, in various suffixed constructions, has come down into other Altaic languages on page 190 of his book: pA *\*ja:g-*, to Mongolian *jayun* "what?"; pA *\*ja:n-*, to Mo. *jambar* (Middle Mongolian *jan*) "what kind of, what sort of?"; pA *\*ja:m-*, to Evenki *e:ma* "what kind of?"; pA *\*ja:du*, to Evenki *e:du*, Lamut *ja:du* "why?", and Goldi *xaidu* "where?".

Other morphological cognates are demonstrated in Miller's work, such as the gerund form of verbs, whose Japanese and Old Turkish cognates are shown in example (11). The Turkish form is used in adverbs of manner. The Japanese is labelled *gerund*, but Miller says that this may be misleading - its usage is almost identical, semantically, to the Turkish form.

- (11) pA *\*-ti* > OT *-ti/-ti, -di/-di, J -te, -de* (Miller 1971:285-292)

Another correspondence, also shown by Miller, is that of the instrumental morpheme, shown in example (12).

- (12) pA *\*-n + te* > OT *-da, -de, -ta, -te, J ni te* (from intermediate *\*ni ta*), pTung *\*āzi* > Ma. *de* (Miller 1971:285-292)

Here, you can see that the Turkish and Manchu-Tungus forms have lost the initial proto-Altaic *\*-n*, but semantic arguments support that it was there in the proto-Altaic forms.

In his paper reviewing the work in and against the Altaic field, Miller refers the reader to Ramstedt and Poppe's work establishing the set of morphological correspondences among the Altaic languages (Miller 1991a:309). The correspondences given here add up to a convincing argument for the Altaic unity, especially considering that most of them show direct correspondences between the two most distant languages of the family - Turkish and Japanese.

### 3.3 Conclusion of morphological evidence

As has been shown, there are secure morphological correspondences among the Altaic languages. In the face of evidence such as these morphemes listed above, as well as other correspondences that have been established by Altaicists such as Ramstedt, Poppe, and Miller, it is very difficult to deny that these languages are genetically related. The fact that the phonetic differences among these morphemes correspond to the established phonetic correspondences between these languages only reinforces the continuity that has proceeded from the original proto-Altaic language down to the modern languages.

#### 4 DISCUSSION OF SYNTACTIC ISSUES

In his overview of Altaic linguistics, Miller (1991a:308 and 1991b:33) laments the fact that the syntactic issue has been virtually ignored by all sides of the Altaic debate. At most, it is often mentioned that the languages in question demonstrate SOV word-order. Miller goes beyond this to say that the Altaic languages are similar in a more significant sense than that - they are all characterized "by a highly specific variety of nominal predication". Though it is an interesting point that Miller raises, McLennan (1996) demonstrates that this sort of syntactic classification is more a reflection of the common universal grammar than a result of genetic correlations. So although the commonality among the Altaic languages on this point is interesting, it can carry little if any weight in an argument for or against genetic unity.

Although any analysis of the Altaic languages is bound to mention syntactic properties, since syntax is an integral part of any language, this mention cannot go beyond the level of description. Shibatani (1990:96) includes a list of properties that are said to be common among the proposed Ural-Altaic family. Since these properties are mainly syntactic and therefore typologically immaterial to discussions of monogenesis, and since a dismally small number of Ural-Altaic cognates, mostly questionable, exists, this family has received even less recognition than Altaic. The syntax argument just doesn't work.

#### 5 DISCUSSION OF NON-LINGUISTIC ARGUMENTS

The fact that the Altaic peoples have been nomadic throughout recorded history makes it difficult to establish any archaeological certainties about them. There were no great Altaic cities, no lasting sedentary agricultural societies . . . none of the things that distinguish sea-side civilizations and those that live on arable, productive land. Menges, in his 1968 study of the Turkic peoples, is forced frequently to define the Turks (and their ancestors, the "Altaijans") in terms of the civilizations that border them and how the Altaijans have affected these cultures. Also, as far as recorded history goes, these neighbouring people - the Chinese, the Indo-Europeans, the Dravidians, and so forth - are the ones we must rely on for written evidence of the "barbarians" who lived in the open expanses of Inner Asia. These accounts are generally vague and subjective, which is understandable but frustrating. We cannot use them to trace the various groups of Altaijans back to one original nation of people (or alternately, to distinct nations or communities, as the anti-Altaicists would expect).

The establishment of historical movements is perhaps best approached from the point of view of a linguist, as by Menges (1968). He explains how certain linguistic borrowings across the borders between Altaic and Indo-European in the west, and Uralic in the north, give some evidence to the times and extent of contact between these groups. Given that the languages are the largest bodies of data that seem to have survived from those ancient days of the Altaic unity, it is not very surprising that linguistic evidence constitutes the bulk of relatedness arguments. The internal structure of the Altaic family, for example, is clearly delineated into the Turkic, the Mongolian, the Manchu-Tungus, and the Japano-Korean groups. This would tend to imply an early divergence where the

proto-Altaiic-speakers drifted apart into roughly four linguistic communities. Miller (1971:44) diagrams this, suggesting that proto-Altaiic first divided into a western and an eastern dialect. The western one developed into the modern Turkic languages. The eastern grouping further divided, leaving proto-Mongol and another community which diverged into proto-Tungus and proto-Japano-Korean. These divisions are based almost entirely on linguistic evidence, though this series of divisions and subdivisions suggests a geographical expansion away from a homeland in the area of the Altai Mountains. The geographical origin of the Altaiic languages and peoples is supported by Menges:

For the times prior to the separation and differentiation from the primordial nucleus groups of Altajic, which were later to become the four<sup>4</sup> Altajic divisions mentioned above, a habitat must be assumed which probably comprised all of the Central Asiatic steppes, so that the term "Altajic" languages is actually justified, since it designates that group of languages spoken around the Altaj Mountains, in a wider sense of the term, in this case on the steppes extending to the south around the Altaj. . . . The term Altaj and Altajic is very handy if understood in the above-defined, larger, not too literal sense. (Menges 1968:57)

The internal relatedness of the various branches of Altaiic - Turkic, Mongol, Manchu-Tungus, and Japano-Korean - seems fairly well-established, and based as well on existing historical records and archaeological evidence. The Turkic languages in particular are the most researched, least disputed group. The Mongol languages, to the east of these, are also internally secure in their relationship. Beyond that, the Manchu-Tungus family is a unit. Martin (1991) uses his in-depth knowledge of the languages to further cement the almost universally accepted fact that Japanese and Korean form a loose sub-family. Japanese is problematic in this matter, since it has a number of isoglosses with the Malayo-Polynesian languages, as demonstrated by Martin (1991:105).

Using these four groups to develop intermediate proto-languages - proto-Turkic, proto-Mongol, proto-Tungus (also called proto-Manchu-Tungus), and proto-Japano-Korean - would seem to be a natural way to progress toward a reconstruction of the original Altaiic language. However, as Miller (1991a:299) complains, Altaicists are accused of using three jokers to win the hand (excluding proto-Japano-Korean). While it is true that with every stage of historical separation, the possibility of error increases, the result of such thorough comparative methods of reconstruction as have gone into the establishment of these three proto-languages (as well as the proto-Korean-Japanese that Martin's work helps to establish) can hardly be compared to drawing random wildcards.

At the east end of Altaiic, the position of Korean and especially of Japanese in the historical structure of the family is uncertain. Korean, being more closely tied, geographically, to the Asian continent and the other Altaiic languages, seems to

---

<sup>4</sup> Menges (1968:56-57) includes the Hunnic group as a fourth branch of Altaiic, in the south and southwest of the original Altaiic geographical area, but that is not dealt with here, because this group is not mentioned in any of the other literature. Also, he excludes the Japano-Korean languages with little more than a comment of dismissal.

demonstrate a more firm connection to the Asian languages than does Japanese. This may also be due to the fact that, as Shibatani (1990:103-114) explains, Japanese seems to demonstrate significant layers of isoglosses with other languages, among which are Dravidian, Austronesian, and Papuan. Knowing whether the Japanese language is, at its core (whatever that means), an Altaic language, or a language from one of these other sources, is not essential to our purposes here, since as Poppe points out in the introduction to Miller's 1971 book, the Altaic stratum of Japanese can still help us reconstruct this ancient language.

## 6 CONCLUSION

### 6.1 Summary of arguments.

Though plagued with layer on layer of borrowings, the lexical issue is addressed, and a firm base of common words is established, from which is built a system of phonetic etymologies to reconstruct the original proto-Altaic forms. With this grip of the phonetic nature of the proto-language, we can proceed to further flesh out our knowledge of the proto-language by examining morphological cognates, which are naturally less subject to borrowing than lexical items, though more susceptible to phonetic mutation over time. Though syntactic typology does little to help prove genetic relatedness or to reconstruct the proto-forms, it is reassuring to know that the Altaic languages do not differ significantly on this point. Beyond the language, we can use biological and socio-historical evidence as further support for the idea that the languages are members of the same family.

Taking all of this evidence together, it is clear that there was once a linguistic unity - a **proto-Altaic** - from which the modern language groups of Turkic, Mongolian, Manchu-Tungus, and Japano-Korean are descended. The precise nature of this proto-language is still being uncovered through tested and true methods of comparison and reconstruction, as the Indo-Europeanists have done with their currently less debated language family. The great time that has elapsed since this unity diverged confounds our efforts, so that there will always be some uncertainty over the nature of proto-Altaic. But that it existed, we can be fairly certain. It does not seem premature to predict that Altaic will soon take its place among the firmly established language families, such as Indo-European and Dravidian.

### 6.2 Discussion of approaches to language comparison.

In the study of the linguistic unity of these languages, and the reconstruction of the proto-language, it is important to pool our efforts. So perhaps this is the time to note some of the different approaches to the Altaic problem. There are linguists such as Sir Gerard Clauson, who according to Miller (1991a:294-5) seems to have been convinced of the unrelatedness of the Altaic languages because his knowledge of Turkish was of no avail in trying to read the *Secret History of the Mongols* (a text written in Middle Mongolian in the thirteenth or fourteenth century, later romanized). With him we may group Doerfer, whose lexical counterarguments to relatedness tend to involve more attention to

dictionary translations than to semantic and phonological similarity. These linguists demonstrate a dismaying disregard for established scientific methods, and should perhaps be paid the attention that such an approach merits.

Others, such as Robert Austerlitz and A. Róna-Tas, exercise a healthy, open-minded scepticism. This can be respected, since though they reserve judgement, they seem interested in the inductive method of gathering all the evidence possible, then evolving a hypothesis out of it; the others deduce, first fixing an idea in their minds, then gathering only the evidence they believe will support their idea.

The optimists - Roy Miller, Nicholas Poppe, G. J. Ramstedt, and so forth - form the other end of the spectrum, having accepted Altaic as proven on the basis of current evidence and proceeding from there to establish more detailed correspondences. It is from these people that the bulk of lexical, phonological, and morphological support for the Altaic hypothesis has come.

A balance between the logical sceptics and the logical optimists should keep the study of Altaic relatedness on an even keel, and produce a solid body of work that will, hopefully, establish a realistic and thorough view of one more aspect of the linguistic heritage of our species.

### 6.3 What is a language family? (revisited)

This question is returned to now because, despite all of the near-certainties and thoroughly scientific paths of inquiry that have formed the science of comparative linguistics, it still remains to be seen exactly how and why languages develop. A short list of questions regarding this problem is given by Austerlitz (1991:361):

- *Filiation*: How do dialects become languages, if this is how languages are born?
- *Meaning change*: What are its societal preconditions?
- *Panchronic rules for sound laws*.
- *Pathways of loans*. Which foreign words (*Fremdwörter*) become loanwords (*Lehnwörter*) and which do not?

All of these are relevant to Altaic, and indeed to all language families. First, how does a single language develop into separate, mutually unintelligible languages? Second, what motivation is needed for cognate words in these separate languages to shift in meaning? In what manner do sounds change over time? And finally, how do we dig through the layers of loanwords to reach the pure ore of proto-language cognates beneath? All of these questions are important, and not all of them have been sufficiently answered for us to be able to travel smoothly back along the history of languages. Until they are, languages such as proto-Altaic, which at second glance almost certainly existed, may remain the limit of our range of historical reconstruction.

## REFERENCES

- Austerlitz, Robert. 1991. 'Alternatives in long-range comparison'. In Lamb et al (eds), 353-364.
- Berta, Arpad, Bernt Brendemoen, & Claus Schonig (eds). 1996. *Symbolae Turcologicae*. (Swedish Research Institute in Istanbul, Transactions vol. 6) Stockholm: Almqvist & Wiksell Internat.
- Bickerton, Derek. 1990. 'Haitian demographics and creole genesis'. In Canadian Journal of Linguistics 35(3):217-219.
- Lamb, Sydney M. & E. Douglas Mitchell (eds). 1991. *Sprung from Some Common Source: Investigations into the Prehistory of Languages*. Stanford, California: Stanford University Press.
- Lefebvre, Claire & John Lumsden. 1990. 'Reply to Bickerton'. In Canadian Journal of Linguistics 35(3):221-224.
- Ligeti, Louis (ed). 1975. *Researches in Altaic Languages*. (Proceedings of 1971 PIAC) Akademiai Kiado, Budapest.
- Loveday, Leo J. 1996. *Language Contact in Japan: A Socio-linguistic History*. Oxford, New York: Oxford University Press.
- Martin, Samuel E. 1991. 'Recent research on the relationships of Japanese and Korean'. In Lamb et al (eds), 269-292.
- McLennan, M. Sean. 1996. *The Split VP Hypothesis, Verbal Nouns, and Japanese*. An Honours Thesis. University of Calgary Linguistics Department.
- Menges, Karl H. 1968. *The Turkic Languages and Peoples: An Introduction to Turkic Studies*. Wiesbaden: Otto Harrassowitz.
- Miller, Roy Andrew. 1971. *Japanese and the Other Altaic Languages*. Chicago, USA: The University of Chicago Press.
- Miller, Roy Andrew. 1975. 'Japanese-Altaic lexical evidence and the Proto-Turkic <<zetacism-sigmatism>>'. In Ligeti (ed), 157-172.
- Miller, Roy Andrew. 1991a. 'Genetic connections among the Altaic languages'. In Lamb et al (eds), 293-327.
- Miller, Roy Andrew. 1991b. 'Anti-Altaists Contra Altaists'. In Decsy, Gyula & A.J.E. Bodrogligeti (eds). Ural-Altaic Yearbook, Volume 63
- Miller, Roy Andrew. 1996. 'Trk. \*l<sub>2</sub>, \*r<sub>2</sub> and Korean'. In Berta et al (eds), 157-167.
- Poppe, Nicholas. 1975. 'Altaic linguistics - an overview'. In Jakobson, Roman & Shiro Hattori (eds), Sciences of Language. Tokyo Institute for Advanced Studies of Language.
- Ritter, Elizabeth. 1992. 'Cross-linguistics evidence for number phrase'. In Canadian Journal of Linguistics 37(2):197-218.
- Róna-Tas, A. 1975. 'The Altaic theory and the history of a Middle Mongolian loan word in Chuvash'. In Ligeti (ed.), 201-211.
- Róna-Tas, A. 1991. *An Introduction to Turkology*. (Studia Uralo-Altaica No. 33). Attila Jozsef University, Szeged.
- Shibatani, Masayoshi. 1990. *The Languages of Japan*. (Cambridge Language Surveys). Cambridge: Cambridge University Press.



Sinor, Denis (ed). 1990. *The Cambridge History of Early Inner Asia*. Cambridge: Cambridge University Press.

Tekin, Talât. 1975. 'Further Evidence for <<Zetacism>> and <<Sigmatism>>'. In Ligeti (ed.), 275-284.

Tim Mills

Department of Linguistics

University of Calgary

Calgary, Alberta, Canada T2N 1N4

timills@acs.ucalgary.ca

