

# Swadesh 100 on Japanese, Korean and Altaic <sup>1</sup>

Martine ROBBEETS

**Keywords:** Japanese, Korean, Altaic, basic vocabulary, genetic relationship

The Altaic origin of Japanese is among the most disputed questions of language history. Given this lack of consensus, an argument that is often advanced is: “there is no basic vocabulary relating Japanese to Korean and Altaic”. The present paper investigates this postulation from a methodological and data-oriented perspective. First it seeks to advance a more precise methodology involving the concept of basic vocabulary. Next it evaluates the etymological proposals made so far, relating Japanese basic vocabulary items to Korean and Altaic. Starting from 92 basic vocabulary items out of Swadesh 100 list for which an etymology relating the Japanese entry to Korean and Altaic has been presented in the past, 41 etymologies stand the selection criteria. The paper concludes that the similarities we find between Japanese, Korean and Altaic are more likely to be the result of common genetic inheritance than of borrowing.

## 1. The “yes-no” debate

Is Japanese an Altaic language, or not? The question is undoubtedly one of the hot spots of contemporary historical linguistics. For many linguists today the Altaic affinity of Japanese remains what it originally was: a question. The lack of a global state of the art leads to the dissemination of extreme answers to the Altaic question for Japanese, ranging from an absolute ‘no’ (Doerfer 1974; Unger 1990; Nichols 1992; Janhunen 1992, 1994; Kiyose 2002; Shogaito 2002; Vovin 2003; Kuribayashi 2004) to an exclamatory ‘yes!’ (von Siebold 1832; Boller 1857; Ramstedt 1924; Murayama 1958, 1966; Miller 1971; Menges 1975; Street-Miller 1975; Street 1977; Finch 1987; Starostin 1991; Vovin 1994 a, b; Wang 1999; Itabashi 2001; Starostin-Dybo-Mudrak 2003).

Given this “yes-no” debate on the Altaic affiliation question of Japanese, an argument that is often heard is “there is no basic vocabulary relating Japanese to Korean and Altaic”. As early as 1820 Abel-Rémusat expressed his doubts on the existence of an Altaic language family comprising Turkic, Mongolian and Tungusic, arguing that the languages in question did not share a sufficient amount of basic vocabulary. Since then the concept of shared basic vocabulary has been used widely in attempts to refute or to support the Altaic theory (Clauson 1956, 1969 a/b; Ligeti 1971 a/b; Poppe 1972; Doerfer 1974, 1988, 1995; Helimsky 1992; Ramer-Vovin-Sidwell 1997). Whether they present evidence in favor of the Altaic hypothesis or not, studies of shared basic vocabulary for a long time ignored Japanese and Korean etymological proposals. Recently, exceptions can be found in Starostin (1991); Wang and Ogura (1999) Itabashi (2001), Shogaito (2002), Vovin (1994 a/b, 2003), Starostin-Dybo-Mudrak (2003); Kuribayashi (2004).

---

<sup>1</sup> This paper is written with the support of the CANON Europe Foundation, creating the facilities for my postdoctoral research at the department of linguistics of the University of Tokyo in 2004.

The argument that “there is no basic vocabulary relating Japanese to Korean and Altaic” can be understood in two different ways. First it can mean that nobody has ever proposed basic vocabulary evidence relating Japanese to Korean and Altaic. And, second, it can also imply that nobody has ever proposed *reliable* basic vocabulary evidence relating Japanese to Korean and Altaic.

The first postulation can easily be dismissed. My doctoral dissertation (Robbeets 2003) aims at presenting a state of the art for the etymological evidence relating Japanese to Korean, and Altaic. My approach is empirical in the sense that the material, the body of the evidence that authors advocated in the past for a Japanese-Korean or Japanese-Altaic connection is at the core of my study. This material is gathered in an etymological index which contains 1806 Japanese lexical entries and 59 morphological entries along with the various etymological proposals that scholarship has suggested for the lexeme or morpheme in question. Just skimming through the etymological index of Japanese appended to my doctoral dissertation, it can be seen that for 92 Japanese entries belonging to Swadesh 100 list various etymologies have been proposed by linguistic scholarship in the past. In short, many scholars have proposed much basic vocabulary evidence relating Japanese to Korean and Altaic. So, the first postulation is clearly invalid.

The second postulation, however, is far more reasonable. Although many scholars have proposed much basic vocabulary evidence, the quality of this evidence is often unsatisfactory. A rather optimistic approach of Altaic basic vocabulary like Itabashi's (2001) has been sharply criticized by Shogaito (2002). In a paper recently presented at the Institute for Humanities and Nature in Kyoto, Kuribayashi (2004) objects to what he calls an overpermissive approach of Altaic basic vocabulary advanced in Starostin-Dybo-Mudrak 2003. Daring comparisons of Japanese basic vocabulary with Korean and Altaic, like the ones we find in Starostin 1991 and Vovin 1994 a/b, serve as a basis for numerical explorations in the origins of the Japanese language. Wang and Ogura (1999) apply statistical techniques like the Neighbor Joining Method and the Bootstrap on Vovin's basic vocabulary list. This list is at the very least controversial because Vovin at present (2003) rejects his former findings. Vovin (1994 a, b) is credited for the disproval of Benedict's Austro-Tai hypothesis for Japanese (1990). But his disproval is mainly based on the availability of better Altaic basic vocabulary evidence competing with the Austronesian etymologies for Japanese etyma proposed by Benedict. But today, in an article that recently appeared in the Nichibunken research report, Vovin (2003) denies the existence of any shared basic vocabulary relating Japanese to the so-called Altaic languages, undermining his own, former arguments against Japanese-Austro-Tai. What is going on here? If we want to prevent a further polarization of the “yes-no” debate, it is high time to turn to the core of the argument and let the data speak for themselves.

## **2. Basic vocabulary or basic word?**

Turning to the core of the argument involves two matters. First there is a methodological matter, concerning the concept of basic vocabulary. Second there is a material matter, relating to the quality of the data. Let us turn to the concept of basic vocabulary first. The core of a basic vocabulary argument is a probabilistic matter. A basic vocabulary argument is a relative argument, not an absolute one. If we find matches within the basic vocabulary, the probability that the matches are due to common ancestorship is higher

than the probability that we are dealing with borrowings. The strength of the concept of basic vocabulary is in its low borrowability and in its high resistance to replacement. But, we cannot turn the argument upside down. Shared basic vocabulary is not a *conditio sine qua non* for the demonstration of common ancestorship.

With low borrowability and high resistance to replacement, we refer to relative tendencies. Low borrowability refers to the fact that words corresponding to the semantic items on the list are less prone to borrowing in comparison with the overall lexicon. High resistance to replacement refers to the relatively low probability that native items on the list get replaced by synonymous loanwords. However, every single item on the basic vocabulary list can get borrowed. And every individual item on the basic vocabulary list is free to stay or to leave. As an illustration of how words for semantic items present on the basic vocabulary list can get borrowed, I refer to a number of Chinese loanwords in Japanese. For 11 items on Swadesh's 100 list, 11. one; 12. two; 16. woman; 17. man ; 27. bark; 29. flesh ; 30. blood; 33. grease; 52. heart; 53. liver; 72. sun, a Chinese loanword is attested in Japanese. Japanese has borrowed *iti* 'one'; *ni* 'two'; *zyosei* 'woman'; *dansei* 'man'; *zyuhi* 'bark'; *niku* 'flesh, meat'; *ketueki* 'blood'; *sibo@grease ,fat*; *sinzo@heart*, *kanzo@liver*, *taiyo@the sun* from Chinese. Except for one item, *niku* 'flesh, meat' for which we find OJ *sisi* 'flesh, meat', but no counterpart in contemporary Japanese, the Chinese loanword is in use side by side with its native synonym.

Contemporary Japanese *hito-* 'one', *huta-* 'two', *onna* 'woman', *otoko* 'man', *kawa* 'bark', *ti* 'blood', *abura* 'grease ,fat' *kokoro* 'heart', *kimo* 'liver', *hi* 'sun' have not been replaced by the Chinese loanwords, but are still present in the language in competing distribution. But 11 Chinese borrowings on 100 in the basic vocabulary is 11 % while the proportion of Chinese loan vocabulary in the entire Japanese lexicon exceeds 60 %. This fact illustrates that semantic items belonging the basic vocabulary list are less prone to borrowing. Moreover, only a single item, 29. Flesh, has been completely replaced by a Chinese loanword. Strictly speaking, only 1% of the basic vocabulary has been replaced by borrowing. This illustrates the tendency of basic vocabulary to resist replacement more successfully.

In short, the stability of basic vocabulary is not in its individual items, but in the low borrowability and resistance to replacement of the system as a whole. So we can say that there is no such thing as a basic word, but there is such a thing as basic vocabulary.

### 3. The rules of the game.

The application of a basic vocabulary test is a matter of fair play. There are a number of rules that should be kept in mind.

#### 3.1. Use a preset list.

A basic vocabulary list is a list with semantic items that are chosen with respect to their cross-linguistic tendency for stability. There are several such lists available in linguistic literature, such as the Swadesh 100 word list, the Swadesh 200 word list, or the shorter Yakhontov or Dolgopolsky lists. It does not matter which list one chooses. What matters is that the test items have been set in advance, for cross-linguistic use, not with the

specific point one wants to make in mind. Because it is the most widely used basic vocabulary list, I have chosen the Swadesh 100 list to start from.<sup>2</sup>

### **3.2. Consider all proposals.**

Every single etymological proposal within the basic vocabulary list deserves to be examined on its own merits. A basic vocabulary test always starts from a state of the art of proposals made by scholarship in the past. It is the Altaic hypothesis for Japanese in general that is at trial and not our personal interpretation of this hypothesis. The etymological index that gathers Korean and Altaic etymologies proposed for Japanese etyma so far (Robbeets 2003) is convenient in this context.

### **3.3. Set up sifting criteria**

Of course not all proposals that have ever been made for Japanese basic vocabulary items are valid ones. When we are looking for matches within the basic vocabulary, it is important to define what we mean with a match. A match is a set of words that correspond regularly in form and closely in meaning. In the following chapter (4) I intend to illustrate the formal and semantic criteria that I adopt for separating the stronger etymological proposals from the weaker ones.

### **3.4. Draw legitimate conclusions.**

A basic vocabulary test can never bluntly tell us whether the languages under consideration are related or not. Arguing for genetic relatedness consists in the establishment of a set of regular sound correspondences. But, it also consists in the demonstration that these correspondences are not the result of massive borrowing. That is where the basic vocabulary test comes in. Given a set of matches, the basic vocabulary test can only indicate whether the matches are likely to be the result of borrowing or whether they are more likely to be cognates. In fact, this brings us back to the concept of basic vocabulary as discussed in the previous chapter (2).

## **4. Sifting the evidence.**

Let us now turn to the quality of the data, relating Japanese basic vocabulary to Korean, Tungusic, Mongolian and Turkic forms. The sifting criteria that I adopt in order to separate the stronger etymological proposals from the weaker ones are the following.

---

<sup>2</sup> The following list of 100 meanings was proposed by the linguist Morris Swadesh in the 1950s and is still most widely used as a basic vocabulary list. 1. I 2. you 3. we 4. this 5. that 6. what 7. who 8. not 9. all 10. many 11. one 12. two 13. big 14. long 15. small 16. woman 17. man 18. person 19. fish 20. bird 21. dog 22. louse 23. tree 24. seed 25. leaf 26. root 27. bark 28. skin 29. flesh 30. blood 31. bone 32. egg 33. grease 34. horn 35. tail 36. feather 37. hair 38. head 39. ear 40. eye 41. nose 42. mouth 43. tooth 44. tongue 45. claw 46. foot 47. knee 48. hand 49. belly 50. neck 51. breast 52. heart 53. liver 54. drink 55. eat 56. bite 57. see 58. hear 59. know 60. sleep 61. die 62. kill 63. swim 64. fly 65. walk 66. come 67. lie 68. sit 69. stand 70. give 71. say 72. sun 73. moon 74. star 75. water 76. rain 77. stone 78. sand 79. earth 80. cloud 81. smoke 82. fire 83. ash 84. burn 85. path 86. mountain 87. red 88. green 89. yellow 90. white 91. black 92. night 93. hot 94. cold 95. say 96. good 97. new 98. round 99. dry 100. name

First I omit etymologies for Japanese basic vocabulary in which the internal analysis of the Japanese form contradicts the external comparison. Second, I omit similarities that could be the result of universal tendencies in the structuring of language. Third, I set up semantic constraints for the comparison of the meanings. Fourth, I omit etymologies that reflect irregular sound correspondences.

In what follows I will not go into the sifting details for every single etymological proposal for Japanese basic vocabulary that has been eliminated. Rather, I will restrict myself to the illustration of the sifting criteria with examples taken from basic vocabulary proposals relating Japanese to Korean and Altaic.<sup>3</sup> In the next chapter (5) the basic vocabulary evidence that stands the test will be presented.

#### 4.1. Conflicting internal analysis.

There are a good number of cases for which the internal analysis of the Japanese basic vocabulary item contradicts the external etymology for the word. Such cases are eliminated from the basic vocabulary core evidence. This sifting criterion can be illustrated with the etymologies proposed for Swadesh's 19. fish.

That we find *uo* 'fish' in contemporary Japanese for OJ *iwo* 'fish' is probably due to the assimilation of the initial vowel to the following labial glide. The problematic internal phonological analysis of OJ *iwo* 'fish' as proposed by Starostin (1997, 334) and Starostin-Dybo-Mudrak (2003, 477) casts doubt on the Altaic etymology that the authors advance for this word.

First the reconstruction of initial pJ *\*d-* in Starostin's reconstruction pJ *\*(d)iwuá* 'fish' is dubious. Based on an argumentation for which I refer to my doctoral dissertation (Robbeets 2003), I take the view that there was no phonological voice distinction in proto-Japanese. In cases where it is possible to distinguish between initial *V-* and *yV-* in Old Japanese, I reconstruct pJ *\*yV-* instead of pJ *\*dV-* for the latter alternant. Thus for *abu* 'horsefly, gadfly', I reconstruct pJ *\*anpu*, while for *yabu* 'refuse heap, garbage dump' I reconstruct pJ *\*yanpu*; for *oru* 'weave' I reconstruct pJ *\*ora-*, while for *yoru* 'twist, twine', I reconstruct pJ *\*yora-*; for *uru* 'sell, deal in' I reconstruct pJ *\*ura-*, while for *yuru* 'shake', I reconstruct pJ *\*yura-*. However, initial *i-* < *\*i-* and *i-* < *\*yi-* are completely neutralized in attested Japanese. The choice between the reconstruction of pJ *\*i-* and pJ *\*yi-* in words like OJ *iwo* 'fish' is completely dependent on the external cognates. If the remaining phonemes of this etymological proposal would represent a perfect fit, it would be methodologically safe to switch back and forth between internal and external evidence in this case. But this is not the case. The correspondence of the medial vowels and consonants in pTg *\*Z#gi* (~-j-) 'a kind of fish (salmon)', pMo *\*Z#a-su* 'fish' and pTk *\*ja#n* 'sheat-fish' (1997, 34) and in pK *\*thi* 'fish (suffix in fish names)' (Starostin-Dybo-Mudrak 2003, 477) are problematic enough to disregard Starostin's proposal.

Another word for 'fish' is *sakana*. Miller (1971, 98) compares this word to Na. *sogdata*, Ev. *sugZanna*, Lam. *huZanra* and reconstructs pTg *\*sugZ#sa* 'fish, trout'. But, in Old Japanese *iwo* is used for 'fish' and OJ *sakana* is used for 'any dish that is consumed when drinking sake'. Treating this three-syllabic word as 'a wine side dish', it is easily

---

3

assumed that we are dealing with a compound of *saka-* ‘sake, rice wine’ and *na* ‘greens, side-dish’. This internal analysis is in conflict with the external etymology, proposed by Miller.

#### 4.2. Universal tendencies in linguistic structuring.

Some similarities between languages cannot be attributed to a common ancestor. They are the result of universal tendencies in linguistic structuring. Observing that nursery terms like Eng. *tits* and J *titi* ‘breasts’, or Eng. *papa* and J *haha* ‘mother’ or sound symbolic words like Eng. *zigzag* and J *gizagiza* ‘notched’, or Eng. *knock knock* and J *kon kon* ‘knock knock’ are similar does not tell us anything about the common ancestor of English and Japanese. Being due to universals in language, mama-papa terms and sound symbolic words are generally regarded as poor diagnostics of genetic relatedness (Jakobson 1960). Universal tendencies in linguistic structuring are well studied as far as the lexicon is concerned, but universals might also be at work within morphological systems. An allomorph may either expand or retract in function, and evidently the universally preferred, unmarked, and otherwise favoured variants are most prone to expand and have a good chance of eventually ending up as the main or sole variant. (Nicholson 2003). The literature has also noted the central role of nasals in universally favoured variants (Jakobson 1960; Nicholson 2003; Campbell 1997). That is why I am disinclined to accept the etymology proposed for the Japanese entry under Swadesh number 8. not.

When turning to the Japanese negative auxiliaries *-(a)n-* and *-(a)z-*, I follow the majority view of linguistic scholarship today that the *mizenkei* is nothing but a surface stem that results from combining an auxiliary with initial *-a-* to the verb root.<sup>4</sup> Further, it is commonly agreed that the negative OJ *-az-* represents a fusion of OJ *-ani* and OJ *s-* ‘do, make’. The negative adjective *nai* and the preverbal negative *na* are related as well. The comparison with the Korean negative marker MK *ani* has been proposed by Whitman (1985, 244) and Unger (2000, 664) and a broader Altaic etymology, extending this binary comparison with the Tungusic negative markers Ewen *aḿ*, Oroch *ana*, Olch *ana*, Orok *ana*, Na. *anaḿ* ‘not, no’, pTg *\*ana* and the Chuvash negative imperative *an*, pTk *\*an* are provided by Miller (1971, 255; 1985, 38, 49-50), Menges (1975, 96-110), Starostin (1991, 253, 267, 277); Starostin-Dybo-Mudrak (2003, 300-01), Vovin (1998). The phonological fit is perfect, but maybe not as compelling as it seems. It can be observed that cross-linguistically nasals tend to be incorporated in negative markers. Perhaps nasals are common in negative markers throughout the world because they are relatively unmarked, universally favored sounds. Many languages have negatives on *m-*, e.g. Sanskrit *mā*, Modern Greek *mi(n)*, Somali *ma*, proto-North Caucasian *\*mV*, proto-Sino-Tibetan *\*ma*, putative Nostratic *\*ma*, or negatives on *n-*, e.g. English *no*, Dutch *neen*, Russian *net*, Lithuanian *ne*, French *non*, Latin *ne*, Sumerian *nu* (verbal negative), *na* (prohibitive), etc. Whether it is the individual negative form, or the mini *yes-no* paradigm that is soundsymbolically marked, we cannot exclude the possibility that a universal determinant of linguistic structure is at work here and that the Altaic negatives are the result of some natural phonological convergence.

---

<sup>4</sup> Ono 1953; Whitman 1985, 244; Takeuchi 1999, 91; Unger 2000, 664; Vovin 2003 b, 168; Robbeets 2003, 110-111.

### 4.3. Semantic overpermissiveness

There is no rule that tells us to keep the semantics of the compared forms identical in a basic vocabulary test. There is such a rule, however, when applying another method, the method of linguistic dating, called glottochronology or lexicostatistics. The effectiveness of the glottochronological method depends on the postulation that semantic drift over time is constant. In this quantitative treatment of basic vocabulary it is essential for the effectiveness of the method to keep the compared meanings constant. However, there is no rule that tells us to keep the semantics constant when applying a basic vocabulary test in a qualitative approach as the one we are adopting now. If we were to deal with the question whether Dutch is relatable to English and German or not, we would take Dutch *hoofd* ‘head’ and compare it to English *head* and German *Haupt* ‘chief, leader, main’, even if the exact semantics would require German *Kopf* ‘head’. We would take Dutch *boom* ‘tree’ and compare it to English *beam* and German *Baum* ‘tree’, even if the exact semantics would require English *tree*. We would take Dutch *vlees* ‘meat’ and compare it to English *flesh* and German *Fleisch* ‘meat’, even if the exact semantics would require English *meat*. Counting only exact semantic equivalents would be untenable. However, the greater the semantic latitude permitted in compared forms, the more likely that the phonological correspondences are purely coincidental. Therefore it is essential to set up certain constraints for semantic permissiveness. It is an art to decide whether a semantic association is acceptable or not, but in order to be on the safe side, semantic divergence will be tolerated only to a very limited extent. Once a certain semantic shift has already been established elsewhere in the world, it gets more confirmatory force. But, although elsewhere attested, the semantic latitude can still be too permissive to exclude coincidental similarity of the compared forms. For this reason I have set up rather severe semantic constraints. The central question here is: “Is Japanese relatable to Korean and Altaic?” That is why I start from etymologies for which the meaning of the Japanese form belongs to the Swadesh 100 list. Permitting a limited degree of semantic latitude, the meaning of the compared forms must not necessarily fall within the 100 list. The etymological proposals for the Japanese entry under Swadesh 18. Person is rejected on the basis of semantic overpermissiveness. The meaning of *na* in contemporary Japanese is ‘name’, but from a number of compounds in Old Japanese like OJ *womi<sub>1</sub>na* ‘woman’, OJ *oki<sub>1</sub>na* ‘old man’ and from contemporary *mina*, *minna* ‘all, everybody, everything’ and *otona* ‘adult’, an additional meaning ‘person’ can be reconstructed. If the semantic shift that occurred in this word answers to the unidirectional tendency of semantic shift from more concrete to more abstract, then ‘person’ can be reconstructed as the primary meaning and ‘name’, which is the abstract concept used in reference to the concrete person is the secondary development. From this viewpoint the external comparison of pJ *\*na* ‘person’ to the Korean word for ‘tale, story’ and the Tungusic words for ‘shamanize, fairy-tale’, the Mongolian words for ‘magic; legend’ and the Turkic words for ‘tale, legend; luck, omen; word; riddle’ following Starostin-Dybo-Mudrak (2003, 888-89) is semantically overpermissive. However Whitman’s (1985, 202) comparison of OJ *-na* ‘person’ to MK *\*nwom* ‘person’ makes a nice match, including the accentual conditions for *\*m* loss.

#### 4.4. Irregular sound correspondences

If the goal of a basic vocabulary test is to find out the number of matches within the basic vocabulary, then we must first explain what is meant by a match. Explaining the criteria for matches, we need a list of phonological correspondences on which our judgement can be based. The consonant and vowel correspondences that I accept relating proto-Japanese to proto-Korean, proto-Tungusic, proto-Mongolian and proto-Turkic are given in the footnotes.<sup>5</sup> These sound correspondences are based on the arrangement of a

| <sup>5</sup> pJ | pK           | pTg          | pMo          | pTk                |
|-----------------|--------------|--------------|--------------|--------------------|
| *p-             | *p-          | *p-          | *p-          | *b-                |
| *p- / *w-       | *p-          | *b-          | *b-          | *b-                |
| *-p-            | *-p-         | *-p-         | *-b- / -V-   | *-p-               |
| *-p- / *-w-     | *-p-         | *-b-         | *-b- / -V-   | *-b-               |
| *-np-           | *-Cp- / -pC- | *-CP- / -PC- | *-CP- / -PC- | *-(C)P-            |
| *t-             | *t-          | *t-          | *t-          | *t-                |
| *t- / *y-       | *t-          | *d-          | *d-          | *j-                |
| *-t-            | *-t-         | *-t-         | *-t-         | *-t-               |
| *-t-            | *-t-         | *-d-         | *-d- (Z#)    | *-d-               |
| *-nt-           | *-Ct-        | *-CT-        | *-CT-        | *-CT-              |
| *t-             | *c-          | *c#          | *c#          | *c#                |
| *-t-            | *-c-         | *-c#         | *-c#         | *-c#               |
| *t-             | *c-          | *Z#          | *d-          | *d-                |
| *y-             | *c-          | *Z#          | *Z#          | *j-                |
| *-y-            | *-l-         | *-d- (-Z#)   | *-d- (-Z#)   | *-d-               |
| *-y-            | *-l-         | *-j-         | *-j-         | *-j-               |
| *-nt-           | *-c-         | *-(C)Cε      | *-(C)c#      | *-(C)c#            |
| *k-             | *k-, h-      | *k-, x-      | *k-          | *k-                |
| *k-             | *k-          | *g-          | *g-          | *g-                |
| *-k-            | *-k- (-h-)   | *-k-         | *-k-         | *-k-               |
| *-k-            | *-k- (-h-)   | *-g-         | *-g- / -V-   | *-g-               |
| *-nk-           | *-Ck-        | *-N-         | *-N-         | *-N-               |
| *-nk-           | *-Ck-        | *-CK- / -KC- | *-CK-        | *-CK-              |
| *s-             | *s-          | *s-          | *s-          | *s-                |
| *-s-            | *-s-         | *-s-         | *-s-         | *-s-               |
| *m-             | *m-          | *m-          | *m-          | *b-                |
| *-m-            | *-m-         | *-m-         | *-m-         | *-m-               |
| *n-             | *n-          | *n-          | *n-          | *j-                |
| *n-             | *n-          | *l-          | *n-          | *j-                |
| *-n-            | *-n-         | *-n-         | *-n-         | *-n-               |
| *n-             | *n-          | *n/          | *Z#          | *j-                |
| *-r- / *-t-     | *-l-         | *-r-         | *-r-         | *-r-               |
| *-r-            | *-l-         | *-r-         | *-r-         | *-r <sub>2</sub> - |
| *-r-            | *-l-         | *-l-         | *-l-         | *-l-               |
| *-s-            | *-l(h)-      | *-l(C)       | *-l(C)-      | *s# *l(C)          |
| *-ns-           | *-l(h)-      | *-l(C)-      | *-l(C)-      | *-s#               |
| *a-             | *a-          | *a-          | *a-          | *a-                |
| *-a-            | *-a-         | *-a-         | *-a-         | *-a-               |
| *o-             | *e-          | *e-          | *e-          | *e-                |
| *-o-            | *-e-         | *-e-         | *-e-         | *-e-               |
| *-a-            | *-e-         | *-e-         | *-e-         | *-e-               |
| *-u-            | *-wo-        | *-e-         | *-e-         | *-e-               |
| *o-             | *wo-         | *o-          | *a-          | *a-                |
| *-o-            | *-wo-        | *-o-         | *-o-         | *-o- / -ˆ-         |
| *o-             | *O(<? *u-)   | *u-          | *ü- / *ö-    | *ü- / *ö-          |
| *-o-            | *u-          | *u- (-/gü)   | *-ö- / -ü-   | *-ö- / -ü-         |
| *u-             | *wu-         | *u-          | *ö- / ü-     | *ö- / ü-           |
| *-u-            | *-wu-        | *-u- (-/gü-) | *-ü- / -ö-   | *-ü- / -ö-         |



body of 350 relatively strong etymologies in matrices, for which I refer to my doctoral dissertation (Robbeets 2003). Again it must be stressed that a basic vocabulary test is not a tool for establishing regular sound correspondences. It can only serve as a tool to indicate the borrowability of the underlying evidence, given a situation in which regular sound correspondences have already been established.

An etymology that relates a Japanese form to Korean or Altaic will be accepted on condition that it reflects regular sound correspondences for the initial consonant, the medial consonant and the medial vowel of the Japanese proto-form.

For *o*, OJ *wo* ‘tail’ we can reconstruct pJ *\*wo*, so we expect the correspondence (pJ *\*w-*, pK *\*p-*, pTg *\*b-*, pMo *\*b-*, pTk *\*b-*) for the initial glide. Starostin-Dybo-Mudrak (2003, 1037) compare *o*, OJ *wo*, pJ *\*wo* ‘tail, tag’ with pTg *\*Nujelse* ‘hair under the neck; tie, scarf’, pMo *\*öjekej* ‘lower part of animal's belly’ and pTk *\*öjek* ‘part of animal's skin under the neck or between legs’. In addition to being semantically weak, there is no regular phonological correspondence.

Supported by numerous compounds such as *mabuta* ‘eyelid’, *manako* ‘eyeball’ etc. and by the quality of the vowel in *me*, OJ *me*<sub>2</sub> ‘eye’, we can safely reconstruct pJ *\*ma* ‘eye’. Attempts to reconstruct pJ *\*na* ‘eye’ on the basis of a disputable analysis of *namida*, OJ *nami<sub>1</sub>ta* ‘tears’ as *\*na* ‘eye’ followed by *\*mi<sub>1</sub>* ‘water’ and the formant *\*-ta* that possibly indicates an undefined quantity are too speculative (Starostin-Dybo-Mudrak 2003, 981-82; Robbeets 2003, 211). Whitman (1985, 128, 160-61, 237) proposes the Korean cognates MK *ma·li*, *me·li* ‘head’, but the semantic latitude permitted here is considerably large. Moreover, Martin (1992, 56; 1996, 20) reconstructs pK *\*mati*, supported by a phonograph reading in the Kyeyrim Yusa *MA- . TTYEY* which is a disturbing phonological correspondence. Other etymological proposals (Starostin 1991, 252, 265, 275; 1997, 339; Starostin-Dybo-Mudrak 2003, 981-82; Vovin 1994 a, 109; 1994 b, 380) for *me* ‘eye’ all compare the word to various forms with an initial dental nasal like pK *\*nwun* ‘eye’, pTg *\*n~~h~~* ‘eye’, pMo *nidin* ‘eye’ which is not the expected reflex according to the correspondence (pJ *\*m-*, pK *\*m-*, pTg *\*m-*, pMo *\*m-*, pTk *\*b*). Starostin further proposes pTk *\*ja~~h~~* ‘tear’, but apart from the problematic initial correspondence, the semantics of this comparison are overpermissive. Thus, although the etymology relating the Korean, Tungusic and Mongolian words for ‘eye’ is conceivable as far as the initial consonant is concerned, it is clear that a Japanese cognate is lacking.<sup>6</sup>

### 5. What survives the sifting.

What is the result of this sifting process? I started from 92 basic vocabulary items out of Swadesh 100 list for which an etymology relating the Japanese entry to Korean and Altaic has been presented in the past. After sifting the evidence the following 41 etymologies stand the test.

---

|        |             |      |      |            |
|--------|-------------|------|------|------------|
| *u-    | *Ø- (<?*o-) | *u-  | *u-  | *u-        |
| *-u-   | *-o-        | *-u- | *-u- | *-u- / -˘- |
| *-aCa- | *-oCo-      | *-u- | *-u- | *-u-       |
|        | *-oCi-      |      |      |            |
| *i-    | *i-         | *i-  | *i-  | *i-        |
| *-i-   | *-i-        | *-i- | *-i- | *-i- / -˘- |

<sup>6</sup> Doerfer 1995, 252; 2001, 198-99; Georg 2003, 442-43 completely reject the remaining Altaic etymology.

**1. I** (1)

(Miller 1971, 161; Itabashi 1998, 129)

OJ *wa* 'I', pJ \**wa*

WMo *ba* first person plural exclusive, pMo \**ba*

**9. ALL** (2)

(Whitman 1985, 127, 194, 239; Starostin 1991 256, 269, 279; Starostin-Dybo-Mudrak 2003, 906)

J *moro-* 'all, every, many, together', pJ \**moro-*

MK *mulus* 'all, in general', pK \**mulu-* 'all'

**10. MANY** (3)

(Martin 1966, 239; Starostin-Dybo-Mudrak 2003, 514)

J *ooi* 'many, copious, abundant', pJ \**opo-*

K *opwus (ha-)* / *opus (ha-)* 'substantial, sufficient, moderately wealthy'

WMo. *abarga*, Khal. *avraga*, pMo \**aba-* 'huge'

Karakhanide Tk. *ap / ep*, (14th C)Tk. *ap, apac`k* 'quite, completely', Kirg. *apej* emphatic interjection, pTk \**ap / ep* emphatic strengthening particle

**11. ONE** (4)

(Martin 1966, 238; Whitman 1985; Starostin 1991, 253, 267, 277; 1997, 329; Starostin-Dybo-Mudrak 2003, 364; Blazhenk 1999, 132)

J *hitotu*, pJ \**pito-* 'one'

MK *pi`lus ~ pilos* 'beginning', pK \**pilu, \*pilo* 'beginning'

OTk. *bir*, Turk. *bir*, Tat. *ber*, Uig. *bir*, South.-Yugh. *b`r*, Az. *bir*, Tkm. *bir*, Chuv. *pe#e*, Yak. *bi#* Kirg. *bir*, Kaz. *bir*, Bash. *ber*, Sal. *byr*, all meaning 'one', pTk \**bir* 'one'

**12. TWO** (5)

(Martin 1966, 249; Vovin 1994, 243; Starostin 1991, 109, 256, 268, 278; Starostin-Dybo-Mudrak 2003, 1107; Blazhenk 1999, 132-33)

J *hutatu*, pJ \**puta-* 'two'

MK *pcak* 'pair', pK \**p(o)cak*

OTk. *buc#k*, Tk. *buZ#k*, Tat. *poc#k*, pTk \**buc#k* 'half'

**13. BIG** (6)

(Whitman 1985, 212; Miller 1985, 151; Starostin 1991, 256, 270, 280; 1997, 329)

J *hutoi* 'big, fat', OJ *puto<sub>1</sub>-*, pJ \**putuo-* 'big, fat'

MK *pwuT-* 'swell, increase', pK \**pwu`to-*

WMo. *büdügün*, Khal. *büdüñ*, Kalm. *büdüñ, bödüñ*, Dong. *biedun*, Bao. *beidoñ*, Dag. *budEn*, Yogh. *budüün*, Mgr. *budin*, ZM *beidun* (18-3a), Mgr. *beidün*, pMo \**büdü-Vün* 'thick'

OTk. *bedük*, Tk. *büyük*, Tat. *biji#*, Uig. *büyük*, South-Yugh. *bez`k*, Az. *böyük*, Tkm. *bejik*, Yak. *bödöM*, Kirg. *bijik*, Kaz. *bijik*, Bash. *bejek*, pTk *\*bedü-k* ‘big, high’<sup>7</sup>

### 15. SMALL (7)

(Martin 1966, 249; Starostin-Dybo-Mudrak 2003, 448)

J *tiisai*, pJ *\*tipi-* ‘small, little’

K *cop-*, MK *cwop-*, pK *\*cwop-* ‘narrow’<sup>8</sup>

Ev. *c#ikte*, *c#ileme*, pTg *\*c#-*, *\*c#u-* ‘narrow’

### 17. MAN (8)

(Starostin 1997, 327; Starostin-Dybo-Mudrak 2003, 335)

OJ *wo-no warapa* ‘boy’, *wo-no ko<sub>1</sub>* ‘boy’, *woka* ‘buck, male deer’, pJ *\*wo-* ‘male’

Neg. *beje*, Na. *beje*, Ev. *beje*, pTg *\*beje* ‘person, man’

WMo. *beje*, Khal. *bije*, Kalm. *bi@bij@*, Dong. *beije*, Dag. *bej*, Yogh. *bEi*, Mgr. *bi@*, pMo *\*beje* ‘body, person, self’

### 18. PERSON (9)

(Whitman 1985, 202)

OJ *womi<sub>1</sub>na* ‘woman’, OJ *oki<sub>1</sub>na* ‘old man’ and from contemporary *mina*, *minna* ‘all, everybody, everything’ and *otona* ‘adult’, pJ *\*na* ‘person’

MK *·nwom*, pK *\*nwom* ‘person’<sup>9</sup>

### 23. TREE (10)

(Whitman 1985, 138-39, 226)

J *ki*, OJ *ki<sub>2</sub>*, pJ *\*ko* ‘tree’

MK *kuluh*, pK *\*kul* ‘stump, counter for trees’ (Old Koguryo *\*kʷl* ‘tree’)

### 27. BARK (11)

(Martin 1966, 225; Whitman 1985, 171, 221; Starostin 1991, 15, 38-39, 67, 92, 251, 264, 27)

J *kawa*, pJ *\*kapa* ‘bark, skin, shell’

K *kkaptaygi* ‘shell, skin’ < *kaph* + *taygi* ‘thingy’, K *kkaphwul* ‘skin, outer layer, film’,

MK *ka . phol* ‘case, sheath’, pK *\*kaph-* ‘bark, skin, outer layer’

OTk. *qab`q*, Tk. *kabuk*, Tat. *kab`k*, Uigh. *qobuq*, Az. *Gab`G*, Tkm. *Ga@q*, Chuv. *xoba#*

Kirg. *kab`k*, Kaz. *qab`q*, Bash. *kab`k*, pTk *\*Ka@uk* ‘bark, shell’

### 28. SKIN (12)

(Miller 1985, 151; 1996, 151; Starostin 1991, 74; Starostin-Dybo-Mudrak 2003, 782-83).

<sup>7</sup> Although the Turkic vowel correspondence pTk *\*-e-* is problematic, the Turkic forms for ‘big, high’ are likely to be related.

<sup>8</sup> Although the Korean vowel correspondence pK *\*-wo-* is problematic, the Korean word for ‘narrow’ is likely to be related.

<sup>9</sup> Although the Korean vowel correspondence pK *\*-wo-* is problematic, the Korean word for ‘person’ is likely to be related.

J *kara*, pJ \**kara* ‘skin, husk, hull, nutshell’

Neg. *ejekte*, Na. *xerekte*, Olc#*xerekte*, Orok *xerekte*, Ev. *erekte*, pTg \**xere-kte* ‘skin’  
OTk. *qaz*, *qað*˚, Tat. *kajr*˚, S.-Yugh. *kazd*˚k, Chuv. *xoja*#, Yak. *qat*˚r˚k, pTk \**kar*<sub>2</sub> ‘bark, scales’

### 30. BLOOD (13)

(Hattori 1959, 203-33; Murayama 1966, 154; 1974, 135, 136, 168, 171-2; Ozawa 1968, 120-1; Street-Miller 1975, 140; Vovin 2003, 24; Starostin-Dybo-Mudrak 2003, 401).

J *ti*, OJ *ti* ‘blood’, OJ *ti* ‘spirit, force’, pJ \**ti* ‘blood’

WMo. *c#u*, Khal. *cus*, Kalm. *cusn*¢Dong. *c#sun*, Bao. *c#oN*, Dag. *c#s*, Yogh. *c#sun*, Monguor. *cEZ#*, ZM *c#sun*, Mogh. *c#sun*, pMo \**c#* \**ti* ‘blood’

OTk. *t˚n*, Chuv. *c##*, Yak. *t˚n*, pTk \**t˚n* ‘spirit, breath’

### 31. BONE (14)

(Martin 1966, 226; Vovin 2003 a, 24; 2003 b, 22; Starostin 1991, 12, 13, 39-40, 67, 92, 251, 264, 274; Starostin-Dybo-Mudrak 2003, 1131-32)

J *hone*, pJ \**poni*(C)a ‘bone, skeleton, rib’

K *ppy*e, MK *˚spye*, pK \**peCi* ‘bone’

Neg. *xenNen*, Na. *feiNge*(n), Olc#*penE*, Orok *pene*(n), Ev. *xenNen*, pTg \**pen!* ‘knee’

### 41. NOSE (15)

(Starostin-Dybo-Mudrak 2003, 1185-86)

J *hana*, pJ \**pana* ‘nose’

Neg. *xun*, Lit. Ma. *fun*, Na. *pu*¢Olc#*pu*(@), Orok *pu*(@), *pu*@Ve, Ev. *hunNukte-*, pTg \**pun!* ‘smell, to smell’

WMo. *ünür*, *ünir*, Khal. *üner*, Kalm. *ünr*¢Dong. *funi*, Bao. *hun-de-*, Dag. *xu*@, Yogh. *honEr*, Monguor. *funir*, *funis-*, pMo \**pünir* ‘smell, to smell’

### 42. MOUTH (16)

(Martin 1966, 237; Whitman 1985, 135, 227; Miller 1985, 142; 1996, 172)

J *kuti* ‘mouth, aperture, opening’, pJ \**kutu*

MK *˚kwut*, pK \**kwut* ‘hollow, pit, cave’

Oroc#*gudE-* ‘break, tear’, Orok, Na. *gudE*@ hole’, pTg \**gude* ‘a hole’

### 43. TOOTH (17)

(Starostin 1991, 256, 268, 278; Starostin-Dybo-Mudrak 2003, 1075)

J *ha*’, pJ \**pa* ‘tooth’

Na. *paloa*, Olc#*pali*¢pTg \**palV* ‘molar’

### 46. FOOT (18)

(Starostin 1991, 275; Starostin-Dybo-Mudrak 2003, 292-93)

J *asi* ‘foot, leg’, pJ \**asi*

Neg. *alc*#@Na, Na. *alc*#a, Olc#*alc*#¢(n), pTg \**alc*#kan ‘knuckle-bone’

WMo. *alc*#, *alc*#, Khal. *alcan*, Kalm. *alc*n¢alcE, Dag. *alc*#¢pMo \**alc*# ‘knuckle-bone’

OTk. *as#q*, Tk. *as#*, Az. *as#G*, Tkm. *as#j*, Kirg. *as#*, pTk *\*as##/ \*alc##* ‘knuckle-bone’

**49. BELLY** (19)

(Martin 1966, 243; 1996, 47; Miller 1971, 153; Whitman 1985, 187, 193, 211; Starostin 1991, 251, 264, 274; Starostin-Dybo-Mudrak 2003, 1131; Vovin 2003 a, 24; 2003 b, 22)  
*J hara*, pJ *\*para* ‘stomach, abdomen’  
*K pay*, MK *\*poy*, pK *\*po(l)i* ‘belly’

**53. LIVER** (20)

(Martin 1966, 249; Starostin-Dybo-Mudrak 2003, 775)  
*J kimo*, OJ *kimo<sub>2</sub>* ‘liver, courage’, pJ *\*kimuo*,  
*K him* ‘strength’, MK *him* ‘strength, sinew’, MK *kim* ‘breath’, pK *\*kim*  
 WMo. *kim*, Khal. *xim*, pMo *\*kim* ‘sausage, offal’

**55. EAT** (21)

(Starostin 1997, 330; Starostin-Dybo-Mudrak 2003, 667)  
*kuu*, pJ *\*kup-* ‘eat, consume’  
 WMo. *kebi-*, Kalm. *kew-*, Dag. *keme-*, Mgr. *kef@* pMo *\*kebi-* ‘chew’  
 OTk. *käv-*, Tk. *gevele-*, Az. *gävälä-*, Tkm. *gävü-s#me-*, Chuv. *kavle-*, Yak. *kebi@* Kirg. *küj-s#*, pTk *\*kēb-* ‘chew’

**56. BITE** (22)

(Starostin-Dybo-Mudrak 2003, 662)  
*J kamu* ‘bite, gnaw, chew, masticate’, OJ *kam-* ‘brew’, *kamosu* ‘brew’, pJ *\*kamo-* ‘bite, gnaw, chew, masticate’  
 Lit. Ma. *kem-ki-* ‘bite’, pTg *\*kem-*  
 WMo. *kemeli-*, *kemile-*, Khal. *ximle-*, Kalm. *keml@*, Bao. *kamel-*, Dag. *keme-*, Yogh. *kemle-*, pMo *\*kemeli-*, *\*kemi-* ‘gnaw, bite’  
 OTk. *kemür-*, Tk. *gemir-*, Tat. *kimer-*, Az. *g"mir-*, Tkm. *gemir-*, Kirg. *kemir-*, pTk *\*kemür-* ‘gnaw, bite’

**60. SLEEP** (23)

(Starostin 1991, 72, 94, 253, 266, 276; Starostin-Dybo-Mudrak 2003, 964)  
*J neru*, pJ *\*na-(C)i-* ‘lie down, sleep’  
 Neg. *ne@Na. ne@Olc#ne@Ev. ne@pTg \*ne@* put’

**62. KILL** (24)

(Starostin-Dybo-Mudrak 2003, 671)  
*J korosu* ‘kill, murder’, pJ *\*koro-*, *\*kora-*, pJ *\*koros-*  
 OTk. *keris# küres#* Tk. *güres#* Az. *güläs#* Tkm. *göres#* Kirg. *keris#* pTk *\*kerüs#*  
 ‘quarrel, fight’

**63. SWIM** (25)

(Starostin 1991, 79, 99, 277; Starostin-Dybo-Mudrak 2003, 1043)  
*J oyogu* ‘swim’, pJ *\*oyonk-*

Neg. *oji* ~~an-~~, Na. *onban-*, Orok *onno* ~~@~~ Ev. *uju-,uju-hta-*, pTg *\*ojV* 'swim'.  
Mmo. *ojna-*, WMo. *ojimu-*, Khal. *ojmo*, Kalm. *öm-*, pMo *\*oji-mu-* 'swim'<sup>10</sup>

**65. WALK** (26)

(Starostin 1997, 334; Starostin-Dybo-Mudrak 2003, 277-78)

J *ayumu* 'walk, step', pJ *\*ayum-*

Lit. Ma. *aja-* 'run quickly', Na. *a@* 'step', Orok *aja-mu* ~~z#~~ 'swift', pTg *\*aj@* 'run, step',

WMo. *ajan*, *ajan*, Dag. *ajan*, pMo *\*aja-* 'journey'.

**66. COME** (27)

(Starostin 1991, 254, 265, 274; Starostin-Dybo-Mudrak 2003, 538)

J *kuru* 'come', pJ *\*ko-*

Orok *gilin-*, Ev. *gel-*, pTg *\*gel-* 'get hardly on one's way'

WMo. *gelgüri-*, *gelderi-*, Khal. *geldre-*, Kalm. *geldr* ~~ç~~, pMo *\*gel-* 'walk slowly'

OTk. *kel-*, Tk. *gel-*, Tat. *kil-*, Uigh. *käl-*, S.-Yugh. *gel-*, Az. *gäl-*, Tkm. *gel-*, Chuv. *kil-*,

Yak. *kel-*, Kirg. *kel-*, Kaz. *kel-*, Bash. *kil-*, Sal. *gel-*, *gej-*, pTk *\*ge* ~~ç~~ 'come'.

**70. GIVE** (28)

(Miller 1981, 857; Starostin-Dybo-Mudrak 2003, 1398-99)

J *tamau* 'grant, give, bestow on', *tamawaru* 'be given', *tame* 'benefit, sake', OJ *tabar-* 'humbly receive, give (to me)', OJ *tab-* 'deign to give', *taberu* 'eat' < 'present (politely)',

pJ *\*tama-*

Neg. *tama-*, Na. *tamã* 'price', Olc ~~#~~ *tama-*, Orok *tama-*, Ev. *tama-*, pTg *\*tama-* 'pay'.

**71. SAY** (29)

(Martin 1966, 249; Unger 1977, 87, 131; Whitman 1985, 246; Starostin 1991, 111, 253, 267, 277; Starostin-Dybo-Mudrak 2003, 589-90)

J *iu* 'say, tell, talk, speak', *iwau* 'congratulate', pJ *\*ipa-* 'say'

K *ip* 'mouth', MK *iph-* 'recite', pK *\*ip*, *\*iph-*

Ev. *ipku-*, *ipke-*, pTg *\*ipke-* 'order, tell'

**72. SUN** (30)

(Whitman 1985, 212; Martin 1996, 29)

J *hi* 'sun', pJ *\*pi*

MK *pyeth* 'sunshine, sun'; MK *pich* 'light', pK *\*picuk*, *\*picok*

**74. STAR** (31)

(Martin 1966, 243; Whitman 1985, 184, 212; Starostin 1991, 13, 36, 90, 254, 267, 277; Starostin-Dybo-Mudrak 2003, 1155-56; Martin 1996, 39; Vovin 2003 b, 22)

*hosi* 'star', pJ *\*posi*,

MK *pyel* 'star', pK *\*peli* (hyangka 12, 7. semantogram + *li*)

---

<sup>10</sup> Although the Mongolian initial vowel correspondence pMo *\*o-* is problematic, the Mongolian verb for 'swim' is likely to be related.

**79. EARTH** (32)

(Yi 1977, 38; Ramstedt 1949, 161)

OJ *na*, pJ \**na* ‘earth, ground, plain’K *nara*, MK *nàráh*, (Koguryo OK \**nua*), pK \**nara* ‘country’Neg. *na*, Ma. *naa*, Lit. Ma. *na*, Na. *na*, Olc#*na*, Orok *na*, Jur. *na*, pTg \**na* ‘earth’,Bash. *jalan*, pTk \**jalaN* ‘field’**82. FIRE** (33)

(Martin 1966, 232; 1996, 47; Whitman 1985, 132, 212; Vovin 2003 a, 25; 2003 b 22; Starostin-Dybo-Mudrak 2003, 1172)

J *hi*, OJ *pi*<sub>2</sub> ‘fire’, pJ \**po(r)i*MK \**pul* ‘fire’, pK \**pul*Neg. *xi-wun*, Lit. Ma. *fárlngá-*, Ev. *hiriw-*, *huri-*, pTg \**puri-* / \**piri-* ‘to dry (over fire)’**83. ASH** (34)

(Starostin-Dybo-Mudrak 2003, 1085)

J *hai*, OJ *papi* ‘ashes’, pJ \**papi*,K *pap*, pK \**pap* ‘refuse, scraps’WMo. *baVasu*, Khal. *ba*, Kalm. *ba*, *ba* ‘to defecate’, Dong. *basun*, *ba-* ‘to defecate’,Monguor. *ba*, pMo \**baVa-su* ‘dirt, excrements’**84. BURN** (35)

(Martin 1966, 227; Miller 1971, 90-94; Miller 1980, 157; Whitman 1985, 214; Starostin 1991, 251, 269, 279; 1997, 334; Starostin-Dybo-Mudrak 2003, 469)

J *taku* ‘burn (something)’ ~ *yaku* ‘set fire, burn’, pJ \**tak-*MK \**tho-* ‘burn’, MK *tahi-* ‘set fire, heat’, pK \**toho-*, \**tahi-*,OTk. *jaq-*, Tk. *jak-*, Tat. *jag-*, Uigh. *jaq-*, Az. *jax-*, Tkm. *jaq-*, Yak. *saq-*, Kirg. *z#q-*, Kaz.*z#q-*, Bash. *jaq-*, pTk \**jak-* ‘burn (something)’**87. RED** (36)

(Whitman 1985, 146, 183, 222)

J *hani* ‘red clay’, *beni* ‘rouge’, *hena* ‘red clay’ pJ \**pa-* ‘red’MK *polka* / *pulka-* ‘red’, MK *pulk-* ‘red’, *polk-* ‘bright’, pK \**pol(o)k-*Neg. *xolajin*, Ma. *fElEgiaN*, Lit. Ma. *fulg’an*, Na. *folg’ã(n)*, Jur. *fulagian*, Ev. *xula-ma-*,  
*rin*, pTg \**pula-* ‘red’WMo. *ulaVan*, Khal. *ula*, Kalm. *ula*, Dong. *xulan*, Bao. *felaN*, Dag. *xula*, Yogh. *z#*Monguor. *fula*, ZM *ula* (13-8), Mogh. *ulo*, pMo \**pula-Van* ‘red’**89. YELLOW** (37)

(Whitman 1985, 133-35, 226)

J *ki*, OJ *ki*<sub>2</sub> ‘yellow’, OJ *kugane* < *ku* + *kane* ‘metal’, pJ \**ku* ‘yellow’MK *kwùli* ‘copper’, (Koguryo OK \**kul* ‘yellow’), pK \**kwuli* ‘yellow, copper’Neg. *ojat*, Na. *Xo*, Olc#*Xoj-pü(n)*, Ev. *uri-m*, pTg \**xuri-* ‘grey’WMo. *küreN*, *kürin*, Khal. *xüren*, Kalm. *kürN*, Dag. *kurel*, Mgr. *kure*, pMo \**küre-* (*küri-*)  
‘(dark) brown’

**90. WHITE** (38)

(Street-Miller 1975, 131-39; Whitman 1985, 168, 234, 236; Starostin 1991, 73, 86, 255, 268, 278; Starostin-Dybo-Mudrak 2003, 1264-65)

J *siro*, OJ *siro*<sub>1</sub> ‘white’, *siraha* ‘white feather’, *sirahama* ‘white beach’, *siraga* ‘white hair’, pJ \**siruo*, \**sira*

MK *hoy-* ‘white’, MK · *huy-* ‘white’, MK *syey-* ‘whiten (of hair, of face)’, pK \**silV-* \**siCo* (i-breaking) > \**syo-* > \**hyo-* - (metathesis) > MK *hoy-* ‘white’  
 \**siCu*(i-breaking) > \**syu-* > \**hyu-* - (metathesis) > MK · *huy-* ‘white’  
 \**siCu-i-* (i-breaking) > \**syu-i-* - (lenition block) > MK *syey-* ‘whiten’

WMo. *sira*, Khal. *s#r*, Kalm. *s#r-E*, Dong. *sEra*, Bao. *s#a*, Dag. *s#ra*, Yogh. *sEra*, Monguor *sTra*, Mogh. *s#a*, pMo \**sira* ‘yellow’

OTk. *sar*<sup>V</sup>, Tk. *sar*<sup>ˆ</sup>, Tat. *sar*<sup>ˆ</sup>, Uigh. *seriq*, S.-Yugh. *sar*<sup>V</sup>, Az. *sar*<sup>ˆ</sup>, Tkm. *sa@*, Chuv. *s#ra#* Yak. *araGas*, Kirg. *sar*<sup>ˆ</sup>, Kaz. *sar*<sup>ˆ</sup>, Bash. *har*<sup>ˆ</sup>, Sal. *sa(:)r*<sup>ˆ</sup>, , pTk \**sa@g* ‘yellow’ (< pTk \**siar*<sup>ˆ</sup>-g ‘yellow’ < pTk \**sira-g*)

**93. HOT** (39)

(Martin 1966, 234; Miller 1971, 84, 87)

*atui* ‘hot’ ~ *yu* ‘hot water’, pJ \**yu-* ~ \**tu-* ‘hot’,

Ev. *dul@*, pTg \**dul* ‘to warm (of sun)’

WMo. *dulaVan*, Khal. *dula@*, Kalm. *dula@*, Dag. *dula@*, Yogh. *dulaan*, pMo \**dula-Van* ‘warm’

OTk. *j’lV*, Tk. *l’k*, Tat. *Z#*, Uigh. *ilman*, S.-Yugh. *ilV*, Az. *j’l’G*, Tkm. *j’l*, Yak. *s’la@*, Kirg. *Z#uu*, Kaz. *Z#*, Bash. *j’l*, Sal. *jili*, pTk \**j’l-g* ‘warm’

**97. NEW** (40)

(Martin 1966, 225)

*ara* ‘new’, *arai* ‘rough, natural, crude’ (Yo.) ‘new’, *atarasii* ‘new’, pJ \**ara-* ‘new, bare’ K *al-* ‘naked, bare’, pK \**al*

**98. ROUND** (41)

(Finch 1987, 51; Starostin 1991, 93, 256, 269, 278; Starostin-Dybo-Mudrak 2003, 955-5) *marui* ‘round’, pJ \**maru-*, \**maro-*

Neg. *mejel*, Lit. Ma. *murgēn*, Na. *murgl*, Olc#*murū-muru*, Orok *morolime*, Ev. *murume*, pTg \**murV-* ‘round’

WMo. *murui*, Khal. *muruj*, Kalm. *mur’u@*, Dag. *morc#ui*, Mgr. *muri@*, Mo \**murui* ‘curve’ Tkm. *bur-*, Chuv. *pʃW rʃWm-*, pTk \**bura-* ‘to twist, wind round’

**6. Conclusion.**

Back to the initial postulation: “Nobody has ever proposed *reliable* basic vocabulary evidence relating Japanese to Korean and Altaic”. Starting from 92 proposals of basic vocabulary evidence relating Japanese to Korean and Altaic, I eliminated more than half of the proposals (ca. 55%) on the basis of methodologically severe sifting criteria. For the remaining 41 etymologies out of Swadesh’s 100 list we know that the reconstruction of



the Japanese proto-form is legitimate, that universal tendencies in linguistic structuring are excluded, that the semantic latitude in the comparisons is permissible and that the sound correspondences are regular.

The distribution of the 41 etymologies over the different branches is as following.

|                               |    |                        |    |
|-------------------------------|----|------------------------|----|
| number of binary comparisons: | 12 | number of K cognates:  | 24 |
| number of 3 branches:         | 18 | number of Tg cognates: | 21 |
| number of 4 branches:         | 11 | number of Mo cognates: | 19 |
| number of 5 branches:         | 0  | number of Tk cognates: | 17 |

What does this tell us? It tells us that the postulation that there is no reliable basic vocabulary relating Japanese to Korean and Altaic is invalid. A basic vocabulary test can tell us whether it is safer to attribute the similarities to genetic inheritance than to attribute them to borrowing. On the basis of the data available in literature, a set of rather severe sifting criteria and the methodology of historical comparative linguistics, I cannot but conclude that the similarities that we find between Japanese, Korean and Altaic are more likely to be the result of genetic inheritance than of borrowing.

## References

- Abél-Rémusat, J. P (1820), “Recherches sur les langues tartares, ou mémoires sur différents points de la grammaire et de la littérature des Mandchous, des Mongols, des Ouigours et des Tibétains” Paris: Imprimerie Royale.
- Benedict, Paul K. (1990), “Japanese/ Austro-Tai”, Ann Arbor: Karoma.
- Blazek, Václav (1999), “Numerals. Comparative Etymological Analyses of Numeral Systems and their Implications”. Brne: Masarykova.
- Boller Anton (1857), “Nachweis, daß das Japanische zum ural-altaischen Stämme gehört” *Sitzungsberichte der philosophisch-historischen Classe der Kaiserlichen Akademie der Wissenschaften* 33.393-481. Vienna.
- Campbell, Lyle, (1997), “Amerind personal pronouns: a second opinion”, *Language* 73.339-51.
- Clauson, Gerard (1956),. “The Case Against the Altaic Theory”, *CAJ* 2.181-87.  
 ----- (1969 a),. “A Lexicostatistical Appraisal of the Altaic Theory” *CAJ* 13.1-23.  
 ----- (1969 b), “Leksikostatističeskaja ocenka altajskoj teorii” *Voprosy jazykoznanija* 6.22-41.
- Doerfer, Gerhard (1974), “Ist das Japanische mit den altaischen Sprachen verwandt?” *ZDMG* 124. 103-142.  
 ----- (1988), “Grundwort und Sprachmischung. Eine Untersuchung an Hand von Körperteilbezeichnungen” Wiesbaden: Franz Steiner Verlag.  
 ----- (1995), “The Recent Development of Nostratism”, *Indogermanische Forschungen* 100.252-267.  
 ----- (2001), “Altaica gottingensia”, *CAJ* 45.181-229.
- Finch, Roger (1987), “Verb Classes in the Altaic Languages”, *Sophia Linguistica* 26.41-61.

- Georg, Stefan (2003), "Japanese, the Altaic Theory, and the Limits of Language Classification", *Perspectives on the Origins of the Japanese Language*. 429-49. Kyoto: International Research Center for Japanese Studies.
- Hattori, Shiro (1959), "Nihongo no keito" Tokyo: Iwanami shoten.
- Helimskij, Eugene (1992), "Review of Doerfer 1988", *Word* 43.322-330.
- Itabashi Yoshizo (1998), "The Old Japanese Personal Pronouns as an Etymological Problem", *Eurasian Studies Yearbook* 70.123-154.
- (2001), "Arutaishogengo (gancho) to nihongo", *Gengo Kenkyu* 120.106-116.
- Jakobson, Roman (1960), "Why 'mama' and 'papa'?", *Perspectives in Psychological Theory* 21-9.
- Janhunen, Juha (1992), "Das Japanische in vergleichender Sicht", *JSFO* 84.145-161.
- (1994), "Additional notes on Japanese and Altaic", *JSFO* 85.236-40.
- Kiyose, Gisaburo & Norikura (2002), "The Relationship among the Languages of the Korean Three Kingdoms, Proto-Japanese and Tungusic, based on historical sources", *Transactions of the 47th International Conference on Eastern Studies*. Tokyo.
- Kuribayashi, Hitoshi (2004), "Arutai gozoku ni tsuite", *Paper presented at the Symposium "Yushia gengoshi no genzai". Sogohikyukankyokaku kenkyu* 2004.07.04.
- Ligeti, Louis (1971a), "La théorie altaïque et la lexicostatistique", *Researches in Altaic Languages*. 99-115.
- (1971 b), "Altajskaja teorija i leksikostatistika", *Voprosy jazykoznanija* 8.21-33.
- Martin, Samuel Elmo (1966), "Lexical Evidence Relating Korean to Japanese", *Language* 42.185-251.
- (1996), "Consonant lenition in Korean and the Macro-Altaic Question", Honolulu: University of Hawaii Press.
- Menges, Karl Heinrich (1975), "Japanisch und Altajisch", *AKML* 41.3.
- Miller, Roy Andrew (1971), "Japanese and the Other Altaic Languages", Chicago: The University of Chicago Press.
- (1980), "Origins of the Japanese Language", Seattle: University of Washington Press.
- (1981), "Altaic Origins of the Japanese Verb Classes", *BHD* 16.815-880.
- (1985), "Externalizing Internal Rules. Lyman's Law in Japanese and Altaic", *Diachronica* 2.137-65.
- (1996), "Languages and History. Japanese, Korean, and Altaic", Bangkok: White Orchid Press.
- Murayama Shichiro (1958), "Einige Formen der Stammverkürzung in den altaischen Sprachen", *Oriens* 11.224-30.
- (1966), "Mongolisch und Japanisch, ein Versuch zum lexikalischen Vergleich.", *Collectanea Mongolica, Festschrift für Professor Dr. Rintchen, zum 60. Geburtstag. Asiatische Forschungen* 17. 153-56.
- (1974), "Nihongo no kenkyu" Tokyo: Kōbun-dō.
- Ōno, Susumu (1953), "Nihongo no doji no katsuyokai no kigen ni tsuite", *Kokugo to*

- kokubungaku*  
350.47- 56.
- Ozawa, Shigeo (1968), “Kodai nihongo to chūsei mongorugo sono jakkan no tango no hikaku kenkyū”, Tokyo: Fuzanbō.
- Poppe, Nicholas (1972), “A New Symposium on the Altaic Theory”, *CAJ* 16.37-58.
- Ramer, Alexis Manaster; Vovin, Alexander; Sidwell, Paul (1997), “On Body Part terms as Evidence in Favor of the Altaic Hypothesis”, *UJNF* 15.116-138.
- Ramstedt, Gustaf John (1924), “A Comparison of the Altaic Languages with Japanese”, *Transactions of the Asiatic Society of Japan. 2nd ser.* 1.41-54.  
----- (1949), “Studies in Korean Etymology”, *MSFO* 95.
- Robbeets, Martine Irma (2003), “Part I: Is Japanese related to the Altaic languages? Part II: An Etymological Index of Japanese”, Leiden: Doctoral Dissertation.
- Shogaito, Masahiro (2002), “Itabashi Yoshizo@Arutaishogengo (gancho@ngo) to nihongo.” ni taisuru komento”, *Gengo Kenkyu* 21.112-121.
- Starostin, Sergej Anatol’evic (1991), “Altajskaja problema i proisxozhdenie japonskogo jazyka”, Moscow: Nauk.  
----- (1997.) “On the “consonant splits” in Japanese”, *JIES. Monograph series* 22.326-341.
- Starostin Sergei; Dybo Anna; Mudrak Oleg (2003), “Etymological Dictionary of the Altaic Languages”, Leiden: Brill.
- Street, John (1974), “On the Lexicon of Proto-Altaic: A Partial Index to Reconstructions”, Madison: the Author.  
----- (1977), “Altaic elements in Old Japanese. Part 2”, Madison: the Author.
- Street, John and Miller, Roy Andrew (1975) *Altaic Elements in Old Japanese. Part 1.* Madison: The authors.
- Swadesh, Morris (1955), “Towards Greater Accuracy in Lexicostatistic Dating”, *International Journal of American Linguistics* 21.121-137.  
----- (1971), “The Origin and Diversification of Language”, Chicago: Aldine Atherton.
- Takeuchi, Lone (1999), “The Structure and History of Japanese”, New York: Longman.
- Unger, Jim Marshall (1977), “Studies in Early Japanese Morphophonemics”, *Yale University Dissertation.* Indiana: Indiana University Linguistics Club.  
----- (1990), “Japanese and What Other Altaic Languages?”, *Baldi* 45.547-561.
- von Siebold, Philipp Franz Balthazar (1832), “Verhandeling over de afkomst der Japanners”, *Verhandelingen des Bataviaasch Genootschaps* 13. 185-275.
- Vovin, Alexander (1994 a) “Long-Distance Relationships, Reconstruction Methodology, and the Origins of Japanese”, *Diachronica* 11.95-114.  
----- (1994 b), “Is Japanese related to Austronesian?”, *Oceanic Linguistics* 33.387-390.  
----- (1998), “Altaic so far?”, *Migracijske Teme* 15.155-213.  
----- (2003 a), “Nihongo keitoron no genzai: kochira kara doko e?” *Perspectives on the Origins of the Japanese Language.* 15-39. Kyoto: International Research Center for Japanese Studies.  
----- (2003 b), “A Reference Grammar of Classical Japanese Prose”, Richmond: Curzon.

Martine ROBBEETS

- Wang, William S-Y. (1999), “Explorations in the Origins of the Japanese Language”,  
*Omoto* 1999.309- 334.
- Whitman, John Bradford (1985), “The Phonological Basis for the Comparison of  
Japanese and Korean”, Cambridge: Doctoral dissertation.
- Yi, Ki-Mun (1977), “Geschichte der Koreanischen Sprache”, Wiesbaden: Dr. Ludwig  
Reichert Verlag.

**日本語、韓国語、アルタイ語に関するスワデシュ 100表**  
マルチン ロッベーツ

日本語の系統問題は国際的な比較言語学のホットスポットです。アルタイ語説をめぐる論争によく述べてくる主張は日本語と韓国語といわゆるアルタイ語に共通基礎語彙がないことです。この論文は方法論的、実証的な見地から共通基礎語彙がない仮定を検討している。