

# Borrowing for a Universal Language

**Young-hee Chung**

*Sejong University*

## Abstract

This paper is a proposal to develop a universal list of vocabulary made up of borrowings from major languages of the world. This proposal depends on the observation (that vocabulary is relatively vulnerable to foreign influence.

In other words, universal vocabulary is to be made up of borrowing from several different languages, just like vocabulary of natural languages is in part made up of loanwords from various languages. Such cosmopolitan vocabulary will have several advantages including linguistic neutrality.

## 1. Introduction

Whoever has traveled abroad would have noticed that the signs for travellers in international airports are written in several different languages. This is just one trivial example of the costliness of the polyglot world. To cite more examples, the expenditure on language services such as simultaneous and consecutive interpretation in international organizations is extremely high, so is the expenditure related to foreign language learning in countries such as Korea. No one, therefore, will deny that the polyglot world costs us dearly, though language diversity could be an indicator of human creativity, and an asset of human beings. The costliness of the polyglot world would become more remarkable in the world to come where the

importance of information processing and communication increases. Economy is the ultimate goal in every human behavior, and the linguistic area cannot be exceptional. It is natural, thus, for human beings to seek the least costly way to communicate, and a common international language would be the answer to it.

There have appeared several schemes for one common international language and Esperanto was the most promising one. But even Esperanto seems to have ended up in failure. Even though it is often argued that the success or failure of any constructed language ultimately is unlikely to be determined primarily on linguistic issues, it is too hurried and early to give up devising a new common linguistic medium for international communication, since we have not tried every possibility.

Here, I will approach the issue a bit differently from the previous ones, and propose to develop, as a first step toward a common linguistic medium, a universal list of vocabulary made up of borrowings from major languages of the world. Our proposal is based on the observation that vocabulary is, compared to other linguistic areas such as syntax and phonology, relatively vulnerable to foreign influence.

In the following sections, universality of the phenomenon of lexical borrowing will be discussed, as well as advantages of the present proposal. A detailed discussion of how to construct a universal vocabulary will appear next.

## **2. Borrowing: a New Strategy for a Universal Language**

### **2.1. Lexical Borrowing**

Borrowing is a word to “describe the adoption into a language of a linguistic feature previously used in another”. (Linguistic terminology Dictionary) Borrowing occurs at various levels of grammar-syntax, phonology, and vocabulary. It is observed by

linguists, however, that borrowing normally begins at the lexical level. When borrowing comes to have a major effect on lexical semantics, there is often a great deal of cross-linguistic syntactic influence as well. Phonology is most unlikely to be affected by borrowing. (Odlin 1989) In other words, vocabulary is the most vulnerable to foreign influence, and so least resistant to borrowing. Thus, every language comprises loanwords, every speech community borrows lexical items from others whenever there is a need.

English is an excellent example to show how extensive lexical borrowing could be in a language. More than 10,000 words were borrowed in the first 150 years of Modern English (from 1,500 to present), and approximately ten thousand words during the period of Middle English (from 1,100 to 1,500) (see Baugh & Cable 1978). The number of loanwords in English is, of course, much more than that: it is estimated that more than half of its vocabulary is derived from Latin (Baugh & Cable 1978). Compared to such extensive lexical borrowing, other areas of the language such as syntax and phonology exhibit relatively weak influence of borrowing.

Since vocabulary is the most vulnerable and least resistant to borrowing, it will be the first and the most promising area where universality could be attained. Therefore, a scheme toward a common international linguistic medium had better aim at, first of all, universality of vocabulary throughout the world. This is where our scheme departs from previous schemes where a grammar including vocabulary, syntax and phonology as a whole was forced upon the learners. Universality in grammar, I argue, has to be preceded by universality of vocabulary.

## **2.2. Cosmopolitan Vocabulary and Linguistic Nationalism**

In most of the previous schemes, words were created arbitrarily. Thus, in Esperanto, for example, the selection of roots is so arbitrary that speakers of some languages might be misled. The Esperanto word “forest”, for example, does not mean "forest" ('forêt' in French)

but 'absence', misleading the English speakers. In addition to it, arbitrarily constructed vocabulary might present an unfamiliar appearance to every speech community.

Given that borrowing of words between languages is natural and common, and loanwords of most languages are of mixed character, borrowed from many other languages, we rather propose to construct vocabulary which is made up of borrowings from various languages of the world. The mixed character of vocabulary is nothing new and rare. English presents a superb example of cosmopolitan vocabulary. English in origin is classified as a Germanic language. However, more than half vocabulary is derived from Latin. Some of these borrowings have direct, a great many through French, some through the other Romance languages. As a result, English shares a great number of words with those languages of Europe which are derived from Latin, notably French, Italian, Spanish, and Portuguese (Baugh & Cable 1978). Not just from European languages has English borrowed, but also from various non-European languages though to a lesser extent- such as Hebrew, Arabic, Hungarian, Hindi-Urdu, Bengali, Malay, Chinese, the language of Java, Australia, Tahiti, Polynesia, West Africa and so on. The following list of 26 Modern English words constitutes an ABC of imports:

<u>Word</u>	<u>Source</u>
amen	Hebrew
bandanna	Hindustani
canyon	Spanish
dengue	Swahili
emotion	French
focus	Latin
gimmick	German
henna	Arabic
igloo	Eskimo
jute	Sanskrit
kamikaze	Japanese
lilac	Persian

mahogany	West Indian
nougat	Provençal
opossum	North American Indian
pickaninny	Portuguese
quirt	Mexican Spanish
rocket	Italian
salt	Indo-European
they	Old Norse
ukulele	Hawaiian
vassal	Celtic
welcome	Anglo-Saxon
xylophone	Greek
yen	Chinese
zebra	Bantu

(John Nist 1966)

The mixed character of its vocabulary is considered as one of the assets of the English language: one thing is that it has contributed to enrichment of the English vocabulary. Another advantage is that English is likely to present a somewhat familiar appearance to anyone who speaks one of the lending languages, particularly either a Germanic or a Romance language. There are parts of the language which he feels he does not have to learn, or learns with little effort.

Thus, vocabulary made up of borrowings from many different languages might have more advantages than disadvantages. First, by comprising terms unique to each culture can vocabulary be enriched. Second, by containing words of many different languages, it is likely to appear familiar to the speakers of those lending languages. Learners will find it comparatively easy to learn if the vocabulary shares lexical similarities with their native tongue. Third, by getting many different languages involved, we can build universal vocabulary in a true sense, not just reflecting one or two particular language(s). This last sense is an important issue in a scheme for a universal language. Remember that most previous schemes were criticized for not being universal: their grammar and vocabulary

mostly resemble those of European languages. Non-European languages have totally been ignored. The chance is that a scheme biased toward a particular language or language group will hardly succeed in the world to come, where linguistic nationalism will be more intensified than ever.

Linguistic nationalism, the belief that one's native tongue is better than any other language, seems to be growing these days. We see several instances of it throughout the world. For example, Swahili, an African native language, is getting unprecedentedly much attention among Africans and so is spreading over a vast area of Africa. The role of English in that area, in contrast, is diminishing gradually. In Canada, the movement of Quebec to get independence might be due to the conflict between the English language and the French language to some extent. The list can go on and on. It is of no doubt that linguistic nationalism will be more intensified in the world to come.

Of course, I am not proposing to design vocabulary in the way to reflect all languages equally; it would hardly be possible to reflect all languages spoken on the earth. Rather I propose a vocabulary made up of borrowings from major languages. Now, the question is how is the majorness to be measured.

### **2.3. Major Languages and their Proportions in Universal Vocabulary**

As the criteria to measure the majorness of a language, I suggest, among others, the number of native speakers, geographical diffusion, cultural or scientific importance, and functional expansion (learned as second language or foreign language). The following chart lists the top twelve languages in the number of speakers, as of 1975.

Mandarin Chinese	450-500 million
English	275-300 million
Spanish	140-150 million
Russian	130-140 million

German	100-105 million
Japanese	95-100 million
Arabic	85-90 million
Portuguese	80-85 million
Bengali	75-80 million
Urdu	75-80 million
French	65-70 million
Hindi	65-70 million

(John Nist 1966)

Among the languages listed above, we have excluded Portuguese from the group of major languages due to its linguistic similarities to Spanish and relatively limited geographic diffusion. Bengali has also been excluded since it is used geographically in a restricted area, and is not important culturally and politically, and has few secondary speakers. Hindi is chosen as one of the major languages over Urdu which is very similar to Hindi, since Hindi is an official language in India. Despite relatively small number of the native speakers, French is counted as a major language due to its considerable functional expansion and importance in culture and science as well as wider geographic expansion. To the list above are added Swahili and Korean. Swahili, whose native speakers are more than two million, is spreading rapidly as a lingua Franca in North East Africa and so selected as a representative of Africa. Korean spoken in the Korean peninsula by more than 80 million speakers deserves to be one of the 12 major languages due to its great number of speakers. Now we come up with eleven major languages: English, Spanish, Chinese (the Mandarin dialect only), German, French, Hindi, Japanese, Korean, Arabic, Swahili, and Russian.

It cannot be denied that these 11 major languages will be ranked differently in significance in each category. So, for example, though Chinese (Mandarin Chinese) outranks the other languages in the number of the native speakers, it ranks low in other categories. In contrast, English will probably outrank the other languages in other categories than the number of the native speakers. Different degrees

of significance of the languages cannot not be ignored when we construct common vocabulary: the more significant a language is, the more words will be borrowed from it. To calculate the exact proportion of each language is an issue to be settled on the basis of a world-wide consensus.

#### **2.4. Criteria for Selection of Words**

For each set of words, we have 11 candidates borrowed from the 11 major languages. Out of the 11 candidates we select the best word. How do we determine which one is best? I propose several criteria for the selection. First, a word with unmarked sounds will be preferred. Markedness is measured by the cross-linguistic frequency of the sounds and unmarked sounds are sounds of high frequency across languages. Languages tend to have a mix of sounds, some found in many languages and some not so commonly found. For instance, in a sample of 317 languages, the vowels /i/, /u/, and /a/ all appear in the inventories of over 250 languages; similarly, the bilabial nasal /m/ appeared in almost 300 languages, and the bilabial stop /b/ in almost 200 languages. In contrast, some sounds are rarer: for example, the German /x/ appeared in 76 languages and the German /ts/ in 46. (Odlin 1989) The significance of such cross-linguistic facts is that there seems to be a rough correlation between the frequency of a sound and the difficulty acquiring it. Thus, unmarked sounds are relatively easier to acquire. Second, a shorter word in length will be preferred, given that that shorter ones will be easier to memorize. Using short words will not pose any problem. For example, in Korean nouns mostly have less than four syllables. Third, a word with a CV syllable type (a consonant followed by a vowel) is preferred. Cross-linguistically it has been observed that CV is the most widespread syllable type. In addition to it, the errors documented by Tarone suggest that speakers of all languages may be predisposed to using CV syllable in a second language. (Odlin 1989) When no candidate is of CV, then the next observed syllable



type, CVC will be chosen. Consonant clusters will be avoided. Fourth, if one of the candidates is already used world widely, that will be selected (i. e. radio, computer, etc.).

The chosen words will be transcribed in Roman alphabets according to the principle of one-to-one correspondence between symbol and sound. In the future, speakers of different language will contact each other more through internet than any other way. Since writing plays a dominant role in internet, it is important to follow the principle of one-to-one correspondence. Writing also plays a significant role in tourism, which will be one of the major ways in the future through which people around the world contact each other.

### **3. Conclusion**

The cost of the polyglot world is higher than ever as information processing and communication increase. A common linguistic medium will provide an economical way for international communication. The present proposal differs from the previous schemes in that priority is given to attaining universality of vocabulary on the basis of the observation that vocabulary is the most vulnerable and least resistant to foreign influence. Universal vocabulary is to be made up of borrowings from several different languages, just like vocabulary of natural languages is in part made up of loanwords from various languages. Such cosmopolitan vocabulary will have several advantages including linguistic neutrality.

The dream of a universal language will not come true easily and in a short time. For universal vocabulary to penetrate each community, it will take not only time but also a world-wide effort and consensus. Now what we can do is to search for better schemes for a universal language.

### **References**

Baugh, A & Thomas Cable. 1978. *A History of the English Language*. London: Routledge & Kegan Paul.

Bright, William, ed. 1992. *International Encyclopedia of Linguistics*. New York: Oxford University Press.

Coulmas, Florian. 1992. *Language and Economy*. Oxford: Blackwell.

Large, Andrew. 1985. *The Artificial Language Movement*. Oxford: Basil Blackwell.

Naisbitt, John and Patricia Aburdene. 1990. *Megatrends 2000*. William Morrow and Co., Inc.

Nist, John. 1996. *A Structural History of English*. New York: ST. Marlin's Press.

Odlin, Terence. 1989. *Language Transfer*. Cambridge: Cambridge University.