

Journal of Universal Language 4
March 2003, 17-31

Comparisons between Pidgins and ‘Unish’

Eun-Joo Kwak

Sejong University

Abstract

In this paper, I review major linguistic features of pidgins and compare them with those of a new artificial language named ‘Unish’. I will show that although Unish is a constructed language, it is easy and natural by modeling pidgins. Furthermore, I argue that Unish is a proper form of language to carry the role of universal language in the globalization era.

Keywords: pidgins, artificial language, lexicon, linguistic features

1. Introduction

One of special categories of human languages is the one-called ‘pidgins’, makeshift languages which arise very suddenly by different language users to carry out basic communicative functions in a situation like trade. Until recently, these makeshift languages have been regarded as inferior and crude languages and thus not seriously considered as linguistic topics. Although many pidgins developed in limited situations have disappeared through the history, some pidg-

ins have been survived and elaborated to 'creoles'. Moreover, a large number of native speakers of pidgins have contributed to the change of social status of pidgins. Thus, pidgins are currently accepted as a special category of natural languages, and the major characteristics of pidgins, i.e., 'simplification' and 'hybridization', are also considered as important features of human languages.

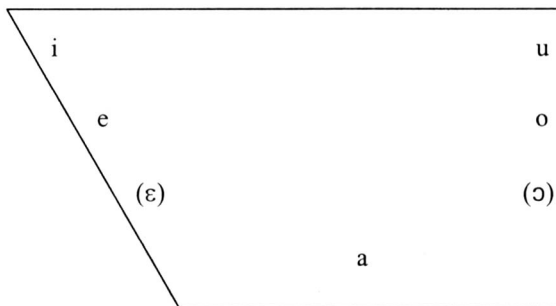
In this paper, I will review major features of pidgins and compare them with a newly constructed language named 'Unish'. Although Unish is artificially invented for general purpose, the structures of the lexicon and grammar are much similar to those of pidgins. Based on this, I will show that Unish is the easiest and simplest language for human communication, adopting the major features of pidgins, the simplest human languages.

2. Characteristics of Pidgins

2.1. Phonological Features

One of the distinctive features of pidgins is that these languages usually have a very simple inventory of phonological sounds. Most pidgins adopt a five vowel system including [i, e, a, o, u] or a seven

Figure 1. Pidgin vowel system



vowel system of [i, e, ε, a, ɔ, o, u].

These five or seven vowel systems are considered as easiest to pronounce and most prominent in cognition. Moreover, distinctive vowel length tends to be lost in pidgins, and thus there is no distinction between short and long vowels in many pidgins.

Along with the simple vowel system, part of the reasons of simple phonological structures of pidgins is the elimination of many marked sounds and the reduction in the number of phonological contrasts. For example, unique consonants in African languages such as clicks, implosives, and ejectives are usually not included in the phonological inventory of their pidgin languages. Hence, three click sounds in Zulu are replaced by [k] in their pidgin languages. Likewise, similar consonants tend to collapse to a more easily pronounced consonant. For example, in Tok Pisin, three phonological sounds of [s], [ʃ], and [tʃ] are reduced to [s] as shown in (1).

- (1) a. san ‘sun’
 b. sem ‘shame’
 c. sok ‘chalk’

Similarly, the contrasts between [θ, ð] and [t, d] are not sustained in Cameroonian pidgin.

- (2) a. dis ‘these’
 b. tink ‘think’

Any factors that are considered as complex or difficult to acquire by users of different languages tend to disappear in pidginization. For example, tones as occurring in many African languages, Chinese and Vietnamese are regarded as difficult phonological factors by speakers of a non-tonal language, and thus pidgins with tonal distinction are derived only when speakers of a tone language are in contact. Hence, the three punctual tones of Ngbandi-Sango-

Nyakoma are preserved in Sango since this pidgin is exclusively used by speakers of tone languages; however, it is not easy to observe in other cases.

The easiest and simplest phonological structures are open syllables of CVCV structure. Hence this bisyllabic structure is the most favored one in many pidgins, and double obstruent clusters are rarely encountered. When lexical items in question do not follow the simple phonological structure, they tend to be modified to fit with the favored one in pidgins.

2.2. Lexical Features

One of the prominent features of pidgin lexicon is a small inventory of lexical items. Compared with 25,000-30,000 lexical items of ordinary human languages, many pidgins just have about 1,500 lexical items. To overcome the shortage of lexical items, pidgin speakers make expressions by combining existing lexical items or depend on contexts to figure out concrete meanings. For example, a pidgin language based on Standard Fijian has only one lexical item *kato*, which corresponds to four lexical items in the standard language.

(3)	<u>Meaning</u>	<u>Standard Fijian</u>	<u>Pidgin Fijian</u>
a.	case, box, basket	kato	kato
b.	fishing basket	noke	kato
c.	coconut leaf basket	sū	kato
d.	woven leaf tray	I lalakai	kato

As the general domain of *kato* covers at least four different fields, speakers depend on a concrete context to pin down its meaning. Similarly, Kamtok’s *hia*, which is derived from *hear* in English, could mean ‘hear’, ‘sense’, and ‘understand’.

In addition to diverse meanings on a given lexical item, many words of pidgins may have many functions. For example, Kamtok’s

bad may function as an adjective, a noun, an adverb or an adjective-verb.

- | | | |
|--------|--------------------------|--|
| (4) a. | tu bad pikin | ‘two bad children’ |
| b. | Wi no laik dis kain bad. | ‘We do not like this kind of badness.’ |
| c. | A laikam bad. | ‘I like it very much.’ |
| d. | Di pikin bad. | ‘The child is bad.’ |

The multi-functional nature of pidgin words is the second strategy to compensate a small inventory of lexical items.

Another property of pidgins is the widely occurring reduplication. Given the small inventory of lexical items, pidgin speakers make a new word by duplicating existing words. For example, in Kamtok, a new word meaning ‘enormous’ is simply derived from the duplication of *big* ‘big’, and a word meaning ‘stare at’ is the duplication of *luk*, which means ‘see’.

- | | | | | |
|--------|-----|-------|---------|------------|
| (5) a. | big | ‘big’ | big big | ‘enormous’ |
| b. | luk | ‘see’ | luk luk | ‘stare at’ |

Similarly, existing words may produce a new word by compounding. In Kamtok, words corresponding to ‘greedy’ and ‘brave’ are derived from compounding of existing words.

- | | | |
|--------|---------|-------------------------------------|
| (6) a. | big ai | (<‘big’+‘eye’) = ‘greed’, ‘greedy’ |
| b. | drai ai | (<‘dry’+‘eye’) = ‘bravery’, ‘brave’ |

Thus, it is concluded that dynamic morphological derivation by compounding, which includes duplication, is another source to make pidgins elaborated languages.

2.3. Grammatical Features

In many pidgins, word order is rigid, and the pattern is usually SVO. Although OSV as in Hiri Motu or SOV as in Eskimo Trade Jargon show, SVO is not the only word order allowed for pidgins. However, many pidgins tend to follow SVO pattern regardless of the word order of their base languages. Moreover, many pidgins do not change word order of sentences to deliver interrogative meanings of imperative meanings.

Inflection is another grammatical factor to be considered as difficult to acquire. Hence, many pidgins do not have any inflection and agreement phenomena; if they have, pidgin speakers try to make it minimal.

(7)	<u>Kamtok</u>	<u>Tok Pisin</u>	<u>English</u>
a.	wan man	wanpela man	‘one man’
b.	tu man	tupela man	‘two men’
c.	A go.	Mi go.	‘I go.’
d.	I go.	Em i go.	‘He/She/It goes.’

As shown in (7), in both of the pidgins, Kamtok and Tok Pisin, differences of number or person do not affect the form of a noun *man* or a verb *go*. Mühlhäusler (1997) argues that Part of the reasons behind the strategy of this minimal inflection is for ease of decoding; maintaining one meaning for one form without inflection, language users easily acquire new pidgins. Similarly, a gender system is another casualty in pidginization. For example, whereas Standard German has a three gender system, Rabaul Creole German has only an invariant definite article *de*.

Pidgins often have a small number of prepositions; some pidgins have as few as two, one to indicate location, and another to indicate possession. In Tok Pisin, for example, the two most frequently used prepositions are *long* ‘along’ and *bilong* ‘belong’.

3. Linguistic Structures of Unish

The rapid globalization of the world requires successful standardization of many fields. Part of the important consensus that should be made in advance will be a language problem, i.e., to agree on a common language in a global era. Basically as candidates of the common language, two types of languages could be imagined. One is to select one of the existing natural languages and assign it the status of lingua franca, i.e., a designated communication method for different language speakers. The other is to choose a lingua franca from constructed or artificial languages, which are artificially made by an individual or a group of individuals for a certain purpose. Artificial languages take advantages over natural languages in that they are relatively easy to acquire and neutral to different language speakers in every aspect.

Given the merits of artificial languages, I will review a constructed language ‘Unish’, which is developed by the research team at Sejong University. The term ‘Unish’ means universal language targeted to a lingua franca in the coming globalized era and based on fifteen representative languages.¹ Basic strategies for developing this new language are ‘easiness’ and ‘commonness’; in other words, Unish is designed to be easy to any language speakers and common to many natural languages.

3.1. Phonological Structures

As reviewed in section 2, pidgins are considered as the simplest forms of human languages. Thus, the phonological structure of Unish is based on the general phonological features of pidgins.

¹ The fifteen languages include fourteen natural languages such as English, Spanish, Portuguese, Italian, French, German, Russian, Korean, Chinese, Japanese, Arabic, Hindi, Greek, Latin, and one artificial language of Esperanto.

First, the vowel structure of Unish consists of five vowels that are most commonly used in pidgins as well as in natural languages. As listed in section 2.1, pidgins accept either a five vowel system of [i, e, a, o, u] or a seven-vowel system of [i, e, ε, a, ɔ, o, u]. Unish takes the five-vowel system since this is simpler and more general. In addition to these five vowels, Unish also accepts semi-vowels of [y, w], which occur in many natural languages to make elaborate sounds. Hence, the correspondence between sounds of vowels/semi-vowels and spellings are as shown in Figure 2.

Figure 2. Unish vowels/semi-vowels and spelling correspondence

spelling	pronunciation
i	[i]
e	[e]
a	[a]
o	[o]
u	[u]
y	[y]
w	[w]

Second, Unish accepts consonants that are easy to pronounce and common to natural languages. For example, the dental sounds of [θ, ð] are not easy to acquire and pronounce and thus considered as representative sounds that are deleted in pidginization. Hence, these sounds are not included in the consonant system of Unish. Moreover, the correspondence between sounds and spellings are as straightforward as listed in figure 3. The list of figure 3 is very simple and straightforward, and thus anybody who has basic knowledge about the roman alphabets can easily understand the relations between consonants and spellings.

Figure 3. Unish consonants and spelling correspondence

spelling	pronunciation	spelling	pronunciation
b	[b]	n	[n]
c	[s]	ng	[ŋ]
ch	[tʃ]	p	[p]
d	[d]	q	[kw]
f	[f]	r	[r]
g	[g]	s	[s]
h	[h]	sh	[ʃ]
j	[dʒ]	t	[t]
k	[k]	v	[v]
l	[l]	x	[ks]
m	[m]	z	[z]

3.2. Selectional Principles of Unish Lexicon

As Unish is designed to be a full-grown language that is suited for any purpose of communication, its lexicon is not directly comparable to those of pidgins. First of all, the number of developed lexical items has already exceeded 8,500, and thus the lexical inventory of Unish covers more than those needed for everyday conversation. Moreover, Unish lexical items will be developed continuously to meet all sorts of topics for both speaking and writing. Hence, although Unish models pidgins in many aspects, the lexical structure is rather distinguished from those of pidgins. Since pidgins have a limited number of lexical items, the lexical structures of pidgins are characterized by multi-function and reduplication. However, Unish does not need to depend on these strategies as it has a sufficient number of distinct lexical items.

One of the important criteria in selecting Unish words is commonness of the origin. As discussed above, Unish is based on fifteen representative languages. Among the distinct fifteen words for a

given meaning, a word having the largest origin is preferred to other candidates. For example, the Unish word for ‘film’ is *film* since words having this origin occur in ten different languages. Similarly, the Unish word selected for ‘café’ is *kafe*, which has origins in eleven languages.

- (8) a. film ‘film’
 b. kafe ‘café’

The philosophy behind this strategy is that the most common word will be the easiest one to many people because people consider their mother tongue as the easiest language.

Another important criterion is ‘simplicity’ of words. To put it differently, the shortest word among candidates is preferably chosen. For example, fourteen different words are reviewed to deliver the prepositional meaning of ‘around’ as listed in (9b). Among these, the shortest one is *um* of German origin. Hence, the Unish word for ‘around’ is fixed to *um* according to the simplicity principle.

- (9) a. um ‘around’
 b. around, alrededor, em redor de, intorno a, autour, um, krugom, juwie, nomawarini, zài~zhouwéi, hawla, ccarongor, guro apo, cirkau

A more dynamic interpretation of this principle is to make words as short as possible by trimming unnecessary affixes. For instance, Unish word for ‘abbreviate’ is selected out of four different forms occurring in five languages as in (10b). Although these languages share the same origin for the word, their suffixes are diverse and have information only for the syntactic category.

- (10) a. abbrev ‘abbreviate’
 b. abbreviate, abreviar, abbreviazione, abrèger

Therefore, the Unish word for ‘abbreviate’ is trimmed as *abrev* by deleting suffixes and the duplicated consonant *b*.

The simplicity principle is not just understood as selection of words having short spellings. Phonological structures of words are also regarded as important factors. As discussed in the phonological structures of pidgins, the simplest and easiest phonological structure is open syllables of CVCV structure. Hence, words of CVCV structure are given high value in selection. For example, the Unish word for ‘appear’ is *apar*, which is derived from six different candidates from seven languages as listed in (11b). Out these words, Unish accepts *apar*, which perfectly suits the CVCV structure.

- (11) a. *apar* ‘appear’
 b. appear, aparecer, apparire, apparaître, apparere, aperas

As discussed above, suffixes are naturally deleted in consideration and the duplicated consonant *p* is also trimmed out. Finally, a single vowel *a* is preferred to *ea* in the second syllable structure. The result is the simple word of *apar*. Although not every word of Unish can follow the CVCV structure, this structure is considered as the ideal phonological one and adopted when it is possible.

Although simplicity both in spelling and pronunciation has high value in the lexical structure of Unish, words that are too simple to carry designated meanings are avoided for ‘distinctivenss’ sake. For instance, the shortest word meaning ‘goose’ is *e* in Chinese, and thus *e* should be selected for ‘goose’ according to the simplicity principle. However, *e* is considered too short to be used distinctively; moreover, *e* could cause confusion to different language speakers as it occurs in their languages with different meanings. Therefore, the Unish word for ‘goose’ is *gus*, originated in both English and Russian.

In addition to commonness and simplicity, diversity of origins is also considered in selecting words. When Unish words cannot be

selected according to the principles of commonness and simplicity, words from minority languages in Unish are given more value than those from languages taking large percentage in Unish. Given the diversity principle, the Unish words for ‘tree’ and ‘bird’ are *namu* from Korean and *tori* from Japanese, respectively.

- (12) a. *namu* ‘tree’
 b. *tori* ‘bird’

Of course, in addition to the minority of these languages, the CVCV structure of the words also contributes to the selection.

Finally, Unish words are also produced by dynamic compounding of existing words. For instance, instead of coining totally new words for either ‘dictionary’ or ‘Monday’, previously selected words are fully utilized to deliver the meanings. Hence, the Unish word for ‘disctionary’ is *motbuk*, which is the compounding of *mot* ‘word’ and *buk* ‘book’.

- (13) a. *mot* ‘word’ + *buk* ‘book’ \Rightarrow *motbuk* ‘dictionary’
 b. *luna* ‘moon’ + *di* ‘day’ \Rightarrow *lunadi* ‘Monday’

Likewise, *Monday* is a day of moon and thus the Unish word for this is determined as *lunadi*, i.e., ‘moon day’. Compounding provides an easy way to coin new words systematically and contributes the enlargement of the lexical inventory of Unish.

3.3. Grammatical Structures of Unish

The grammatical structures of Unish are very similar to those of pidgins since the latter structures are regarded as very simple and easy in language acquisition. First, the word order of Unish follows SVO structure, which occur in many pidgins. (c.f., Lee 2002) Moreover, word order is quite fixed in Unish, and thus much information

for syntactic category of lexical items is inferred from their positions in the sentence structures.

- (14) a. De ver-ed tori.
 you see-Past bird
 ‘You saw a bird.’
 b. De ver-ed wat?
 you see-Past what
 ‘What did you see?’

In both (14a), a declarative sentence, and (14b), an interrogative one, the sentences start with a subject that is followed by a verb and an object. As the word order is fixed, distinction between declaratives and interrogatives are made by punctuation, i.e., the question marker, in writing and by intonation in speaking. Interrogatives are assigned rising intonation pattern in speaking.

Second, pronominal structures of both personal and relative pronouns are made as simple as possible. Hence no distinction is made for the gender differences for personal pronouns, and only one relative pronoun is assigned, namely *dat*, regardless and animacy and clause structures.

As listed in figure 4, number distinction among pronouns is made by the suffix *-s* for plurals, and a single form of pronoun *le* is used for ‘he’, ‘she’ and ‘it’.

Figure 4. Person pronouns in Unish

person	singular	plural
1 st person	me	mes
2 nd person	de	des
3 rd person	le	les

Finally, inflection is allowed minimally. Gender distinction is not

made through suffix, and number agreement is not part of Unish grammar. However, information for tense is considered as crucial in sentence meanings, and thus *-ed* is used as suffix for past tense and *-il* is accepted as a suffix for future tense.

- (15) a. Me skrib buk.
 I write book
 ‘I write a book.’
- b. Le skrib buk.
 he/she write book
 ‘He/she writes a book.’
- c. Me scrib-ed buk.
 I write-Past book
 ‘I wrote a book.’
- d. Me skrib-il buk.
 I write-Future book
 ‘I will write a book.’

As shown in (15a) and (15b), the same form of verb *skrib* is used for the subjects of both the first and the third person pronoun. However, a past sentence of (15c) and a future one of (15d) are distinguished from (15a)-(15b) by the suffixes.

4. Concluding Remark

Just with the growing number of native speakers of pidgins and their elaborated descendants of creoles, the social and academic status of pidgins has also been enhanced. Considering that pidgins are the simplest and easiest forms of languages, linguistic features of pidgins may provide guidance for a proper form of language for the globalization era. To be used by different language speakers, a universal language for the new era should be easy enough to be ac-

quired with little effort.

In this study, I have reviewed major linguistic properties of pidgins and introduce a new artificial language name Unish. I have shown that Unish is constructed to model pidgins and thus has a strong basis to be accepted as a common and easy language by speakers of diverse language origins.

References

- Choo, M. 2001. The Need for “Unish”, a Universal Language and the Principles of its Development. *Journal of Universal Language* 2, 3-14.
- Chung, Y. 1996. An International Language for the World to Come. *Journal of Universal Language* 1, 56-70.
- Chung, Y. 2001. Borrowing for a Universal Language. *Journal of Universal Language* 2, 24-33.
- Comrie, B. 1996. Natural and Artificial International Languages: A Typologist’s Assessment. *Journal of Universal Language* 1, 35-55.
- Diamond, J. 1999. Sejong’s Achievement as Viewed by an American. *Journal of Universal Language* 1, 11-19.
- Holm, J. 2000. *An Introduction to Pidgins and Creoles*. New York: Cambridge University Press.
- Huber, M. 1999. *Ghanaian Pidgin English in its West African Context: A Sociohistorical and Structural Analysis*. Amsterdam: Benjamins.
- Large, A. 1996. The Prospects for an International Language. *Journal of Universal Language* 1, 20-34.
- Large, A. 1996. The New Babel: Language Barriers on the World Wide Web. *Journal of Universal Language* 3, 77-95.
- McWhorter, J. 1999. *Language and Language Contact in Pidgins and Creoles*. Amsterdam: Benjamins.
- Mühlhäusler, P. 1997. *Pidgin and Creole Linguistics*. London: University of Westminster Press.
- Romaine, S. 1990. *Pidgin and Creole Languages*. London: Longman.
- Sebba, M. 1997. *Contact Languages: Pidgins and Creoles*. New York: St. Martin’s Press.
- Singh, I. 2000. *Pidgins and Creoles: An Introduction*. London: Arnold.
- Todd, L. 1990. *Pidgins and Creoles*. London: Routledge.