

TO: Unicode Technical Committee  
FROM: S. and S. Oh|  
SUBJECT: Follow-up to L2/17-125 "Preliminary Proposal for Encoding New Hangul"  
and L2/17-126 "Comments on L2/17-125"  
DATE: 24 July 2017

Dear Committee Members,

Thank you very much for taking time to read my letter and to review the documents. I am also very thankful to God that I am having this precious opportunity to explain the issue, and express an opinion on the Korean addition in Unicode with Dr. Deborah Anderson. I realized that the members may not clearly understand Hangul distinction, even though this documents was focused mostly on the distinction. Therefore, I am adding the distinction of implementing Hangul software in this letter. Back in 1991, nobody expected that we will have this much change in human society. The Committee has lead the change because Committee is the most scientific organization. I believe that all of the members including former members are contributing their knowledge and skill to progress the communication and information technology. I know you have been done many research and you have devoted your research to facilitate human beings, I am extremely grateful to you as one of the beneficiaries

The Committee have worked independently, and exclusively without any political influence and also, any selfish capitalist influence. The Committee has achieved great progress with working in the fact and truth of science. I am expecting you will consider this project with the fact, and I hope you will make the positive decision for this long desired of Korean wishes. Thank you.

This letter includes the debate and refutation of the Jaemin Chung report to UTC in May meeting, and also it includes an opinion over the issue that discussed in the meeting. And it also has added an appeal for this project.

A. Introduction. B. Issue. C. Project D. Debate. E. Appeal.

A. Introduction.

Since 1996, no additional Hangul Unicode syllable was added (no Korean addition), even though there were multiple attempts. I am assuming that if the UTC had to

reserve certain amount left over code spaces, and therefore, the Committee was not able to approve any previous Hangul-proposals, then it should have to be understandable.

After reviewed Jaemin Chung' report to UTC (L2/17-126), I realized that it is very necessary to pinpoint the issue of this project. I believe that the Committee members will review it, if it is pinpointed, and I hope they will consider it.

The issue will be focused on

- a. Whether the characters are "in common use",
- b. Whether the characters are "invention". And also
- c. Whether this project will "hurt other's interest".

And I realized that those issues will be the "denial reasons" of this project.

The encoded single Chinese character can be implemented in their computing system. Recent decades, China and Taipei are using different version of Chinese characters. Chinese created a lot of modified format version of Chinese characters for easy using, and are popularly using now. Even though those modified characters might be looked "invented", but the "common use" must be a good guideline for encoding.

If an encoded character cannot be implementing in software, then it will be useless. Hangul is different with other writing system. Nevertheless, current Hangul software is working perfect in the computing system. Therefore, Hangul Unicode was properly encoded. and its distinction comes into the current Hangul Unicode.

A syllable is a combination of components, and is the basic unit that will be implementing in Hangul writing. The gathered and connected syllables are representing the words. The syllables could be compared to Alphabet by itself. The components (=jamo) are being the input commanding words which will be implementing the syllables. Therefore, the components and the syllables both are needed for Hangul software. But, even though as many components are added in Hangul Unicode, it will not be implementing in Hangul software for writing. Only additional syllables will bring the additional implementing in Hangul software. The syllables are the combinations of the initial and the medial {initial x media l}, or the initial and the medial and the final-components {initial x medial x final }.

It is important to understand the implementing to judge the issue. Therefore, I will suppose a miniature model. I am going to suppose the "B,C" are Hangul consonantal elements and the "1,2" are vowel elements. The consonants can be combined to make additional consonants suchlike "BB, CC, BC" or even "BCB (not allowed)". The vowel are also combined and suchlike "12" for additional vowel.

The consonants are using for the beginning or ending of a syllable (initial, final). Both of them are looked in same figure, but it has to be distinguished in the computing system, and Unicode have classified into the initial and final. I will call that "implementing component". The initial component (Ja) will be "B/C/BB/CC/BC", and the final be "b/c/bb/cc/bc". The medial will be "1/2/12". Among the 10 implementing Ja, I am going to suppose that "BC" was never been used in writing system and will be excluded for implementing components set. Then the rest will be 9 components; 4 initial "B/C/BB/CC", and 5 final "b/c/bb/cc/bc". The implementing Mo will be 3 medial "1/2/12". The altogether 12 Jamo components will be the "SET of utilizing components" for the mathematic combination which are "syllables".

The combination will be "4 initial x 3 medial = 12 syllables (B1/B2/B12/ C1/C2/C12/ BB1/BB2/BB12/CC1/CC2/CC12), and "4 initial x 3 medial x 5 final = 60 syllables. (B1b/B1c/B1bb/B1cc/B1bc/B2b/B2c/B2bb/B2cc/B2bc and so on).

I am going to suppose that among the 72 syllables, 4 syllables "BB1c/CC2bb/CC12cc/CC" are never been used in current language, and 3 syllables "BB1bc/ CC12bc" are not pronounced. But if the 7 characters are not included in the encoding, because "Not in use" or "Invented", then the software of 4 characters "B.C.1.2" will not be working properly

Accordingly, a single Hangul syllable is just one of the combinations of the input components set, regardless whether it is in use or not.

It will be like the finished products of the input materials in automated chemical plant. And the understanding and acknowledgment of this distinction was applied for the encoding of Hangul Unicode back in 1996.

Among 254 Jamo components, 19 initials, 21 medial vowels and 27 finals were applied for the "utilizing components set" in 1996, and all of the combinations (=syllables) were granted and approved to encoding.

The number of encoded syllables was  $19 \times 21 \times (1+27) = 11,172$ .

That is why so many "not in use" syllables are also in Hangul-Unicode.

The syllable character was not the object to discuss "IN USE or NOT". It was just the "LOGICAL RESULT".

And a lot of the "never been used syllables" were not being questioned "whether it is the invention or not" by the Committee at that time.

Moreover, the partial encoding will cause a lot of the disordered syllables in the software. {If "없" is not encoded: 없서 >> 업써,어브써,어빠서 }.

## B. Issue

### b1) Invention, in common use. .

If a new component is encoded, then it will become the Hangul Jamo components, and it can be in a set of implementing Jamo group.

Then the combinations of this group containing this new components will be resulted logically.

The new syllables are not the object to discuss the issue of "IN USE,or INVENTION", as long as the components are included in Hangul Jamo.

Therefore, the focus have to move to the components.

### b2) The components.

Hangul was invented the characters of 14 consonantal element and 10 vowel elements of the sounds. {Among 10 vowels, 4 double projection vowels of "ㅑ ㅓ ㅕ ㅗ" are not the basic vowels. It is the combined sounds like diphthong.}

And so many additional consonantal characters can be composed by combining with other consonants, and the vowel will do the same.

There are 254 Hangul Jamo components in Unicode currently. Jamo means the consonants and the vowels. Among them, 188 characters are the consonantal components (Ja) and 66 characters are the vowel components (Mo). Therefore, beside 24 first invented Jamo components, all rest of the components are the combined components.

The consonantal components must have its role in the syllables, either the initial or the final. That is why you can see some same characters encoded twice in Unicode (initial, final).

If a new consonant is included in Hangul component, then the additional consonants can be composed by combining with other consonants together. It should also have its roles in the syllable for the initial or the final.

All of the proposed components are the consonant characters, and the 6 final consonants are the combined characters of "R" with "ㄴ, ㄹ, ㅁ, ㅂ, ㅅ, ㅈ".

Accordingly, if the "R" is being included into Hangul components, then the "R ㄴ, R ㄹ, R ㅁ, R ㅂ, R ㅅ, R ㅈ" will not be the "invention", because the combining is the nature distinction of Hangul. Therefore, if the four Alphabet consonants are being included in Hangul Jamo, then the combined components and the Alpha-Hangul syllables are not the object to discuss "invention" or "in use".

Therefore, the main issue should be whether the four Alphabet characters into Hangul will be the inventing or the borrowing, and the Committee will allow the four Alphabet characters into Hangul, or not.

### b3) Borrowing.

Alphabet is originated from the scripts used by the Phoenician merchants in BC 16century. It has spread around the world, suchlike, Greek, Roman, French, German, Russian, Mongolian, Indonesian, and Vietnamese. They imported them from other country, and sometimes they modified them partially to make their own. Roman is most widely used around the world suchlike English, Spanish, Turkish, Indonesian and Philippine.

Unlike with other borrowing (importing), this project will be the partial borrowing from English. Any of above Alphabet is not suitable for Korean language, and this borrowing is just for indication of loan words only. Therefore, total Alphabet borrowing is not considerable.

Some countries in history had forced to use their own language to other people, and even they tried to eliminate other's language. American language policies is not to force to use only English nor to eliminate others language. Bi-language using is granted and small group's languages are supported and assisted, SEI activity is one of the role. Nevertheless American soft language policy, English is being attractive language to learn, and every schools in the world including the hostile country like North Korea are teaching English voluntary. Recent decades, English are creating most new words that indicate new technology, new invention, new idea and concepts. Loaning English words is prevailed, because the borrowing will be easier than to find or to create a corresponding Korean words.

Borrowing is not restricted nor prohibited by neither Korean nor American (or British) authority. Therefore the individual Koreans use more often English words loaned as for their own expression.

The Koreans are getting frustrated when they have to use "Fast" for the sound of "Past" in Korean writing. That is why Koreans eager to have the additions in Korean. (Homographs will be mentioned later)

The four characters are the characters of English phonetic sounds which are already existing. It cannot be "the invention", it is just "borrowing" from English.

Three of them are the same design of Roman alphabet "R/V/Z", and one of them is the modified "F" for F sound.

Allowing to borrow four Alphabet phonetic sound into Hangul will be greatly devoted to the individual Korean people.

I think the individual should have the choice of the proper expression as they have the choice of the borrowing English words freely.

Please do not betray the Korean's necessity of the addition, by repeatedly denying the proposals.

b4). Borrowing words from English.

b4-1). Borrowing English words in Alphabet:

In order to use English properly, it is essential to know the correct spelling and other grammar subject. Every S Korean are learning English in school, but they are not good enough to use English properly, because their grammar is very different with English. And English will not be properly implemented with Korean software.

Therefore, the borrowing English written in Alphabet will not be the supplanting.

b4-2). Borrowing English words into Hangul.

The Hangul standard called "한글맞춤법", was firstly enacted in early 1930', it rules mostly proper Hangul spelling, and other grammatical issue. An amendment was made in early 1980', in order to rule the indication of foreign loan words, because each individual uses different expressions what they believe it close to that sounds. There were not many English loan words at that time, and they did not expect many English words will be loaned.

They concluded that the confusion problem regard to the homographs. The amendment included the rule of "used only Hangul for the foreign loan words".

These Hangul standards are educational purpose and it does not have the enforcement power to prohibit the "indication against standards", even though that is used in the TV caption.

Recent years, the traditional communication such as the letters and telephone are less preferred, and they are increasingly borrow English words on online communication. The problem is getting greater, and it cannot be regarded as homographs anymore, because more than 20% of English words are containing the sounds of this project.

The confusion and ambiguity are getting unbearable,

b5) Hurt other' interest

The first word of "Terms of use" in Unicode chart is "anyone can use it freely". The encoding cannot be the preference for only certain person or company.

The advantage of encoding will belong to both the individual software user and the supplier. The progress is always made when they meet the demands. The denying or objection will not make the progress. Please do positive consideration of this matter.

C. Project.

The reason I attached the captured TV caption was not to prove the "in use", but to show the Alpha-Hangul mixed syllables are being accepted by the people.

c1) Project implementing

I am going to suppose that the TV caption syllable { th ㅏ ㅓ } is encoded, then it should able to implement in Hangul software. The "ㅏ" and "ㅓ" are already in Unicode. The "/th/" have to be include in Jamo initial component. The "/t/" and "/h/" have to be also included in initial Jamo, in order to be "/t/" + "/h/" = "/th/". These three initials {"/t/, /h/, /th/"} will be able to combine with the medial {up to 21}, and the combination will be up to  $3 \times 21 = 63$ .

These 63 syllables will also be combined with the final {up to 16 or 27} consonants, the combinations will be up to  $63 \times 16 = 1008$ . This TV caption character will be the one among the 1071 encoded syllables (=combination), in order to be implemented properly in the software. Nevertheless, we will have the "disruption" immediately.

Because, we already have the characters for "/t/" and "/h/", even if we don't have the character for "/th/". Accordingly, if any components that duplicates its sound with existing Jamo, then it should not be included in Hangul Jamo.

c2) Utilizing

This project is to express only foreign loan words. Therefore, it has minimized the input components number for the SET. Any unnecessary medial and final components are excluded. Any syllables not related with foreign loan words that will be disordered, when it is implemented.

The biggest feature of this project is utilizable. Any Korean who know some English will be able to read immediately.

c3) Reuse the spare components.

The only 67 components out of 254 Hangul Jamo Unicode are using currently. The rest components were sort through whether it can be the utilizing for this project.

c4) The "breakage" and Korean national body.

Korean standards are approved by the authorized Korean organization that is a sub-organization of ISO. They should not approve any standards against ISO approval. The mentioned "break" will be happened when it lacks the ISO approval. The UTC is the very special organization. It have been worked out very independent, exclusive from any other institute of organization, even it is the member of ISO. Especially, unlikely with other standards, KATS had nothing to do with Hangul Unicode last twenty five years. If the authorized Korean body's support is required, then I will follow up.

D. Debate

d1) The vowel "ㅡ".

In order to indicate the "speed" in Hangul, the "스피드" cannot be used, it won't be pronounced, it have to be "스피드 or 스피". The reason for mentioning "ㅡ" in the document is to express the necessity of the "ㅡ", along with the corresponding Hangul vowel for "AEIOU".

d2) Other language;

This project is only linked for Korean language users. It will be used to avoid or lessening the confusion and ambiguity of the foreign loan words and mostly from English. It is useless for other language users.

#### d3) Precedent

Korean had borrowing Chinese words in a long time. The borrowing happened, when the majority people knew the Chinese characters. Most Koreans are learning English, and the big volume of borrowing English is happening now. That is why Korean are need to borrow the English phonetic sounds also. The big volume borrowing from Arabic, Indian, Greek or other language won't apparently be happen, and it will not be necessary to borrow their phonetic sounds into Hangul, even in the future.

When a few Korean knew English, the English words suchlike "car, train, air plane" was substituted for the words of Chinese characters. Now most Korean know English, and more often English words are loaned as it sound, suchlike "smart phone, online communication, SNS and internet banking".

Most people in the world are Alphabet users. They might not having this problem when they loan words from English, even though, their Alphabet is not Roman. This project is only for the Korean to indicating loaning words. The mentioning of "Turkish, French or any other language" in her letter will not be related to the issue. It will be useless to other language users.

There are a lot of Chinese origin words in Korean present time, if this project is approved, then it is expected that a lot of English origin words will be included in Korean language along with Chinese, because the individuals can express better, and will use more English words.

#### d4) “z” and “ㄷ”

The traditional handwriting is seldom using these days. They will use this characters mostly on computing system. The "z" and "ㄷ" will be distinguished.

#### E. appeal.

Since after the May UTC meeting, I have learned that any further attempt to get the addition of Korean in Unicode will be denied, or will be approved for only non-implementable characters. I found the reason behind is that the guidelines for encoding is not suitable for implementing Hangul software. As it mentioned, in order

to the Hangul software be implemented, all of the combination of components set are needed, regardless whether the combination (=syllable) are in use or not. It can be comprehensible if it is compared with a misspelled word still can be implemented in computer, even though that is not using.

If the "common use" was the requirement for the encoding, and was applied to the encoding Hangul Unicode back in 1996, then the Korean people will not able to use computing system properly, till present time.

The Committee have working out with gathering the member's opinion, and then, made very proper and fair decision. I am grateful for the proper decision that was made in 1996 for Hangul Unicode.

This project is the borrowing four Alphabet characters into Hangul Jamo components, because there are no Hangul components representing those sounds. I have convinced that there not be any alternative proposal for this proposal, because any other proposal will be overlapped with current Hangul components.

The "A ㅏ" is occasionally seen on TV these days, but "A" cannot be borrowed into Hangul components, because 1) the corresponding Hangul vowel for the "A" is already existing in Hangul components, 2) the borrowed vowel cannot be combined with current Hangul vowels. Suchlike, the "ㅏ" and "ㅏ" will be "ㅏ+ㅏ=ㅏㅏ", but "ㅏ" and "A" cannot be combined together in its figure and its distinction. The "t" and "h" for the "th" will also be interfered with current Hangul components.

Hangul is very different with Chinese characters, even though both are in CJK Unicode category. The Chinese characters are indicating the meanings, and it is important to have the proper characters to express correct meaning. Meanwhile, Hangul indicate the sounds, and it is necessary to indicate correct sounds to make a proper expression.

It might be bothersome to research the distinction and aware the implementing of Hangul, but please review the document again, and consider it.

The judgement standards of sports game cannot be fair to every ones, but the encoding guideline should not be the sports game judgements. I believe that every Committee members are working hard to research and consider the fact to make fair and proper conclusion. I also strongly believed that the Committee had made proper decision for Hangul Unicode about quarter century ago, therefore, the Korean people

are able to use proper computing system today. Once again, quarter century past, the Korean are now desperately needed the addition of Korean. The demand is the borrowing four Alphabet into Hangul. All of the components and the syllables in this project are just the consequence of borrowing the four Alphabet.

The decision in 1996 was very proper and harmful to nobody. It will be very fair, and cannot be harm to anybody, even if the Committee allow the borrowing. I realized that only 10% of Unicode spaces are allocated as of today, please do re-consider it. And please make a positive conclusion.

Best regards.

Dear, Ms. Jaemin Chung.

Thank you very much for your deep concern over my project. I was very impressed at your vast and deep knowledge about the various language. I am not trying to make just refutation on your report to UTC. But I like to politely ask you to focus some issue in different angle. Firstly, about your mention of the "It will be a bad precedent". This is not an event that happens often. The motivation of this project is that Koreans are borrowing a lot of English words these days. The Koreans had borrowed from Chinese words long time. The borrowings are happen when a lot of the people know their letters. Now most Koreans are learning English, and the borrowing English occurs. But apparently the borrowing from other language suchlike Arabic, India, Greek, Vietnamese won't be happen, because most Koreans are not knowing nor learning their language. And there will not be any necessity of indicating their language. This is a project of borrowing some Alphabet from English. None of the characters are overlapped with current Hangul characters, and it will not be disrupted with other indication. Meanwhile, there are the Jamo components representing the sounds for "t", "h" and "H". The "A" cannot be the Mo component in Hangul.

It might not happen within several hundred years that a lot of Koreans are learning certain same language, and borrow a lot of words from them.

Even though you and the Committee allows the encoding for this project, it will not be any precedent, because no further characters not be overlapped.

I am doing this task because the Korean need this, and somebody have to do it. And I believe it is worthy to do. I should feel obligation if I mess up this precious opportunity with Dr. Anderson, which any Korean may not have it again. Please do positive consideration.

Thank you.