# Types and Patterns of English L1 Students' Misspellings in Korean

**Author:**
Shin, Seong-Chul

# Types and Patterns of English L1 Students' Misspellings in Korean

## Seong-Chul Shin

*School of Languages and Linguistics, Faculty of Arts and Social Sciences,*
*University of New South Wales, Sydney, 2052. s.shin@unsw.edu.au*

신성철. 2007. 영어권 한국어 학습자의 철자 오류의 유형과 패턴. *한국어 교육* 18-3: 99-122. 본 연구는 영어가 제 1 언어인 한국어 학습자의 작문에 나타난 철자 오류들의 빈도수를 측정하여, 그 오류들의 유형과 패턴 및 주요 오류 원인들을 알아보는데 있다. 이를 위해 3 개 대학에서 공부하는 한국어 학습자가 작성한 정규 시험지 작문에서 635 개의 유효 철자 오류를 추출하여 분석, 분류하였다. 그 결과 40 개 이상의 철자 오류 유형을 관찰할 수 있었으며 그 가운데 3 쌍 (6 개)의 오류 유형이 전체 오류의 31%를 차지할 만큼 고빈도 오류를 나타냈다. 이들 오류 유형들을 음운별로 나누어 실례와 함께 제시하고, 오류 유형의 패턴을 5 개로 분류하여 제시하였다. 오류의 원인은 영어와 한국어 사이의 차이나 전이로 인한 언어간의 요인과 한국어의 특정 음질들의 유사성으로 인한 언어내적인 요인과 그외 로마자 사용, '시각적' 혼동과 같은 일반적인 요인들을 논하였다. 교육적 측면에서 철자, 소리, 의미가 모두 결합된 교육 원리의 적용과 특히 의미의 유무 여부를 강조하여 교육할 것을 제안한다. (뉴사우스웨일즈대 UNSW)

주제어: 철자오류 (spelling errors), 오류 유형과 패턴 (error types and patterns),

## 1. Introduction

Orthography or spelling is uncompromising, and there are only two options to evaluate: correct or incorrect. James, Scholfield, Garrett and Griffiths (1993: 287-8) point out that the norms of grammar and pronunciation can vary in a foreign language situation, as long as communication is not hindered. Influential studies of learner language such as Selinker (1972), where the intermediate, transitional and approximate nature of the learner's language 'system' is emphasised might have led to generous perceptions towards learners' errors in second language learning and teaching. Unfortunately but understandably, the same tolerance or acceptability does not apply to orthography or spelling. No approximations are allowed in spelling. This is the starting point for one who wants to venture into this field.

This study aims to identify the orthographic[1] features that present particular difficulties to English L1-KFL learners, to classify them in terms of their type, frequency and pattern. A sound understanding of the nature and pattern of the errors should form the basis for adequate linguistic explanations and prepare the way for pedagogical measures to deal with the problematic areas. For this purpose, the study has 1) quantified and documented spelling errors from data sources; 2) stated the frequency of each type of error by a quantitative method; 3) linguistically classified errors by types and patterns; and 4) attempted to interpret descriptive statistical results and explain possible causes for some main errors.

A good number of EA studies in Korean have been carried out over the past few years, particularly from the mid 1990s to recent years.[2] The most recent studies include M-O. Kim (2003), E-J. Hong (2004), Y. Kim (2006), I-J. Cho (2006) and J-S. Kim (2007). Despite the complexity of 'sound-symbol relationship' in Korean due to its rich phonology, however, there are very few studies in Korean spelling error analysis (EA).[3] M-O. Kim (2001: 334-335) points out the fact that studies of spellings in KFL are extremely rare despite the fact that a considerable proportion of errors produced by KFL learners are in spelling. M-O Kim (2002: 504) further emphasizes the necessity of enhancing the teaching of spellings for English L1 speakers saying that the English-speaking learners of Korean produced the highest rate of spelling errors among the five language groups in question throughout the language levels.

The types of spelling errors and their causes have been examined in some studies. In James *et al.* (1993), for example, the types of errors are categorised into three, namely, L1 interference, non-interference and dual (or multiple) origin. James *et al*. (ibid.) claim that there is a strong L1 influence on L2 production (i.e. spelling errors). Similar claims are made by H-M. Sohn (1986) who examined composition errors produced by American university students learning Korean. He claims that many of the error types are attributable to different types of confusion, such as confusion between the allophonic Korean consonants, and concludes that phonological interference from L1 (English) is the strongest factor in this confusion. M-O. Kim (2001) examines spelling errors produced by Korean-American students learning Korean in Korea. The most frequent error types found in the study include consonant confusions between *ss* and *s*, *n*

---

[1] The term 'orthography' or 'orthographic' used in this study refers to spelling and thus orthographic errors refer to misspellings or spelling errors rather than misformed or incorrectly spaced characters. The question of spacing between words is not examined in this study.
[2] Kang (2003) reports that among the 37 Korean EA studies carried out from 1980s to the end of 2002, 21 studies (57%) was undertaken for the 200-2002 period.
[3] Studies of spelling errors in Korean that are currently available include H-M. Sohn (1986), M-O. Kim (2001) and S-C. Shin (2001).

and *ng*, *k* and *kk*, and between *c* and *ch* or *cc*, while, in vowel, confusions between *ey* and *ay*, *e* and *o* and between *o* and *wu* were most frequent.

The primary difference between the present study and the previous ones is found in the subjects and data. While previous studies were based on a single institution or a single group of KFL students mostly with mixed language backgrounds, this study is based on six groups of English L1-KFL learners at three different institutions. EA studies on multi-groups with a common language background at different institutions have the significance of being able to draw common errors by groups with a specific language background and generalize the outcome on more valid grounds.

## 2. Data and Data Processing

The data used in this study come from written examination papers administered at the three universities.[4] To maintain consistency, the data have been obtained only from regular formal examinations such as end-of-semester examination papers and do not include homework-type data such as worksheets and take-home essays. The study particularly focuses on two classes of textual data: a *free composition* type of data and a *reformulation* type[5] of data (Corder 1981). With this combination of two types of data, it is expected that the constraints of textual data can be complemented, thereby increasing the validity of the findings.

The data were collected and sorted out in the following way: 1) collect regular examination papers completed by 2nd and 3rd Year students of Korean; 2) select the papers which were completed only by students who were identified as English L1 speakers; and 3) select a free composition or a reformulated type of text (e.g. translation, completion) in the examination papers. Out of 167 examination papers, a total of 635 valid spelling errors were identified for analysis, before being classified by the researcher and his assistant.

## 3. Subjects

The subjects selected in this study are 167 young adult students who are enrolled in Level 2 and Level 3 (or 2nd and 3rd Year) classes at three universities. Prior learning experiences varied according to individual learners but I was informed that an average student in Level 2 must have studied a minimum of 150 - 200 hours and a student in Level 3 a minimum of 300 - 400 hours in a formal setting, or have reached its standard equivalent, before enrolling in these levels. The reason for choosing students at these levels was because I believed that these groups would

---

[4] One Korean and two Australian universities.
[5] It is a type that reconstructs the ideas and intentions of other people, e.g. translations.

provide me with a more appropriate ground to observe the students' tenacious learning difficulties than with the lower or higher levels of students.[6]

Another important aspect was the language background of the subjects. The subjects whose writings were selected in this study were L1 speakers of English or were believed to have English as their first language. For the purpose of this study, Korean-heritage or background students were excluded from the subject group (or at least it was intended not to include any of them). This was confirmed by student information (e.g. country of origin and names) obtained from the source institutions or by consultation with its relevant teaching staff. Out of 167 students in total, 85 (50.9%) were third-year level students, and 82 (49.1%) second-year level students.

## 4. Procedure

Identifying spelling errors requires a decision-making process which distinguishes between misspellings and other levels of errors, i.e. lexical, morphological and syntactic. In this study, the focus will be on grapheme-related errors that concern vowel and consonant misspellings. I adapt Finch's definition of a grapheme as an "individual letter as concept" (Finch 2000: 47). It is an individual letter representing a phoneme, such as *c* in *ca.ta* 'sleep', *ch* in *cha.ta* 'be cold' and *cc* in *cca.ta* 'be salty'. This way, spelling errors are distinguished from morphology-related errors, which are related to word forms (e.g. *chin.kwu.kwa* for *chin.kwu.wa* 'with a friend'), lexical errors related to 'meaning' (e.g. *cak.ta* for *cek.ta* 'be small in size' for 'be small in number'), and syntactic errors related to the functions of grammatical items (e.g. *hwa.lul.nayss.un.tey.to* for *hwa.lul.nayss.nun.tey.to* 'even if (I) got angry'). This study deals with misspelled words and does not include errors that are considered to be lexical errors. Also, errors related to morphological and syntactic usage are not treated as spelling errors in this study.

After identifying spelling errors, legible misspelled words were examined and listed along with correct spellings, then transliterated into Roman letters. Each misspelling was classified according to the graphemic features and coded by error type (e.g. *ay* for *ey*). If a word contained more than one type of error (e.g. *sam.pho* for *san.po* 'taking a walk'), the errors were classified

---

[6] In this study, the lower level of students generally refer to those who were in the 1st year of their study of Korean, while the higher level of students refer to those who have done four or more years of formal study in Korean.

into two or more types (i.e. *m* for *n* and *ph* for *p*). Numbers of each error type were computed and calculations were made of the distribution and relative frequencies of the error types.[7]

## 5. Presentation of the Results

Out of a total of 635 orthographic errors were analysed, I have found more than 40 error types that occurred three or more times. In the following the types of misspellings are presented in order of their frequency. According to the frequency rate they are grouped into three: high frequency errors (approximately 10% or more), mid-frequency errors (approximately 2% – 9%) and low-frequency errors (less than 2%).

### 5.1 Descriptive Statistical Information

The single most frequent error type was the substitution of *ch* by *c* (6.0%), followed by the substitution of *ay* by *ey* (5.5%). Table 1 below shows the overall frequencies of 40 spelling error types.

Table 1: Types of misspellings: Overall frequency

| Error Type | No. | % | | Error Type | No. | % |
|---|---|---|---|---|---|---|
| 1. *c for ch* | 38 | 6.0 | | 21. *l. for l.l* | 6 | 0.9 |
| 2. *ey for ay* | 35 | 5.5 | | 22. *o for wu* | 6 | 0.9 |
| 3. *o for e* | 33 | 5.2 | | 23. *ss for s* | 6 | 0.9 |
| 4. *ch for c* | 33 | 5.2 | | 24. *kk for k* | 6 | 0.9 |
| 5. *e for o* | 32 | 5.0 | | 25. *u for wu* | 5 | 0.8 |
| 6. *ay for ey* | 26 | 4.1 | | 26. *l.l for l.* | 5 | 0.8 |
| 7. *t for tt* | 16 | 2.5 | | 27. *.l for l.* | 5 | 0.8 |
| 8. *p for ph* | 13 | 2.1 | | 28. *n for ng* | 5 | 0.8 |
| 9. *ng for n* | 11 | 1.7 | | 29. *ng for ø* | 5 | 0.8 |
| 10. *wu for o* | 11 | 1.7 | | 30. *ph for p* | 5 | 0.8 |
| 11. *t for th* | 10 | 1.6 | | 31. *wu for u* | 5 | 0.8 |
| 12. *ø for h* | 10 | 1.6 | | 32. *uy for oy* | 5 | 0.8 |
| 13. *a for e* | 9 | 1.4 | | 33. *ø for ng* | 5 | 0.8 |
| 14. *way for oy* | 9 | 1.4 | | 34. *ay for a* | 4 | 0.6 |
| 15. *k for kk* | 9 | 1.4 | | 35. *l.l for .l* | 4 | 0.6 |
| 16. *ey for yey* | 7 | 1.1 | | 36. *l. for .l* | 4 | 0.6 |
| 17. *n.n for .n* | 7 | 1.1 | | 37. *.l for l.l* | 4 | 0.6 |
| 18. *s for ss* | 7 | 1.1 | | 38. *ph for pp* | 4 | 0.6 |
| 19. *e for u* | 7 | 1.1 | | 39. *tt for t* | 4 | 0.6 |
| 20. *e for a* | 6 | 0.9 | | 40. *th for t* | 3 | 0.5 |

---

[7] This study does not include errors made in transliterations of loan words into Hankul, as they are not regarded as spelling errors in a strict sense.

In terms of frequency, there were three groups of the most frequent error types where one type of error occurred more than 60 times ranging 10%: (a) confusion of *c* and *ch* (11.2%); (b) confusion of *e* and *o* (10.2%); and (c) confusion of *ay* and *ey* (9.6%), as shown in Table 2 below.

Table 2: Three most frequent types of misspellings (groups)

| Error Type Group | No of Errors | Percentage |
|---|---|---|
| *c* for *ch* and *ch* for *c* | 71 | 11.2% |
| *e* for *o* and *o* for *e* | 65 | 10.2% |
| *ay* for *ey* and *ey* for *ay* | 61 | 9.6% |
| Total | 197 | 31.0% |

In addition to the high frequency error types, there were 12 other groups of considerably frequent error types occurring 10 or more times each, and these include (in order of frequency): (a) *t-th-tt* substitution; (b) *l*-related; and (c) *p-ph-pp* substitution. Apart from these mid-frequency errors, there were some miscellaneous but still potentially active types of errors occurring 5 to 10 times, such as *ey* for *yey*, *n.n* for *.n*, *uy* for *oy*, and *ay* for *a*.

Among the 15 frequent error types, there were slightly more consonant error types (8 types) than vowel error types (7 types), but in terms of percentage value, there were slightly more vowel misspellings (24.3%) than consonant misspellings (22.1%). In all 40 types presented above, however, there were more misspellings of consonants (35.3%) than those of vowels (31.3%).

The subsection 5.2 below presents the frequent error types with some examples. Misspellings are indicated with * and their correct spellings are given in parentheses. To save space, only two examples are given for each type, with their primary meaning. Dots (.) are used to indicate a syllable boundary and the symbol > indicates that the erroneous spelling must be replaced with the correct spelling in that direction, i.e. 'a > b' means 'a' should be 'b'.

## 5.2. Descriptions of Error Types

The most frequent spelling errors came from the *c* and *ch* substitutions, where the *c for ch* type (6.0%) was more frequent than the *ch for c* type (5.2%). Observe the examples shown in (1) below.

(1) Palatal stops (11.2%): *c > ch* - *a.cim (a.chim) 'morning', *cayk (chayk) 'book'; *ch > c* - *cha.ka.yong (ca.ka.yong) 'family car', *chass.e.yo (cass.e.yo) '*sb* slept'.[8]

The most frequent error type in vowels was *ey*-[e] *for ay*-[ɛ] (5.5%) and this is compared with its opposite type, *ay for ey*, which was considerably less frequent (4.1%). The total substitutions between them were slightly fewer than the *o*-[o] and *e*-[ə] substitutions, where there was not much difference in frequency between the *o for e* type (5.2%) and the *e for o* type (5.0%). Examples of the most frequent vowel errors are shown in (2) and (3) below.

(2) Mid-back vowels (10.2%): *o > e* - *pon.ho (pen.ho) 'number', *kon.kang (ken.kang) 'health'; *e > o* - *pyel.le (pyel.lo) 'not really/particularly', *il.pen (il.pon) Japan.

(3) Mid- and low-front vowels (9.6%): *ey > ay* - *tam.pey (tam.pay) 'cigarette', *ney.il (nay.il) 'tomorrow'; *ay > ey* - *ka.kay (ka.key) 'shop', *swuk.cay (swuk.cey) 'homework'

Other frequent consonant errors came from the substitutions involving the lenis, aspirated and tensed series of stops (i.e. labial, alveolar, palatal and velar stops) and *s*-series, with varying degrees of frequency. Frequent three-series consonant errors were of the following types.

(4) Alveolar stops (5.4%): *t > tt* - *ta.la.ka.yo (tta.la.ka.yo) 'follow', *tay (ttay) 'occasion', 'the time'; *t > th* - *po.tong (po.thong) 'usually', *to.yo.il (tho.yo.il) Saturday; *tt > t* - *cel.ttay (cel.tay) 'absolutely', *ttak.a.yo (takk.a.yo) 'clean'; *th > t* - *thong.an (tong.an) 'during', 'duration', *thong.ku.la.mi (tong.ku.la.mi) 'circle'; *th > tt* - *thok.thok.ha.ko (ttok.ttok.ha.ko) 'be clever and'.

(5) Labial stops (4.3%): *p > ph* - *pi.wu.myen (phi.wu.myen) 'if one smokes', *pi.kon.hay.se (phi.kon.hay.se) 'because (sb) is tired'; *ph > p* - *pok.caph.hay.se (pok.cap.hay.se) 'because (sth) is complex', *phyeng (pyeng) 'disease/sickness'; *p > pp* - *na.pa.yo (na.ppa.yo) '(sb/sth) is bad.', *na.pu.ta.ci.yo? (na.ppu.ta.ci.yo?) 'They say it's bad, don't they?'; *pp > ph* - *a.ppass.sum.ni.ta (a.phass.sum.ni.ta) 'was sick'; *ph.ph > ph.* - *aph.phu.lo (aph.u.lo) 'in the future'; *p. > .pp* - *nap.a.yo (na.ppa.yo) '(sth) is bad'.

---

[8] '(sb)' refers to 'someboday', while 'sth' 'something'.

(6) Velar stops (2.4%) [9]: *k > kk* - *e.me.nim.key (e.me.nim.kke) 'to mother', *cam.kan.man (cam.kkan.man) 'for a moment'; *kk > k* - *ce.kkum (co.kum) 'a little/some'.

(7) Aveolar fricatives (2.1%): *s > ss* - *nal.si (nal.ssi) 'weather', *sep.ni.ta (ssup.ni.ta) '(sb) spends', '(sb) uses'; *ss > s* - *5 ssi (5 si) 5 o'clock, *ca.ssik.tung.i (ca.sik.tul.i) 'chidren (subj)'

Another group of consonant errors resulted from the substitution, addition or omission of nasal *n* or *ng* and lateral *l*. The nasal and lateral errors were of the following types.

(8) Nasals (5.2%): *ng > n* - *kong.weng (kong.wen) 'park', *ceng.kong (cen.kong) 'major study'; *n.n > .n* - cen.nyek (ce.nyek) 'evening', *sin.nay (si.nay) 'city'; *n > ng* - *kwan.ko (kwang.ko) 'advertisement', *cen.mal (ceng.mal) 'really'; *ng > ø* - *yeng.hayng (ye.hayng) 'travel', *thong.hayss.e.yo (tho.hayss.eo.yo) '(sb) vomited'; *ø > ng* - *ye.hwa (yeng.hwa) 'movie', *yo.ton (yong.ton) 'pocket money'.

(9) Lateral *l*-related (4.3%): *.l > l.l* - *ta.la.yo (tal.la.yo) '(sb/sth) is different', *ppa.li (ppal.li) 'quickly'; *l.l > .l* - *il.lum (i.lum) 'name', *mol.lu.ci.man (mo.lu.ci.man) 'don't know but'; *.l > l.* - *a.la.yo (al.a.yo) '(sb) knows', *ke.lo (kel.e) 'walk'; *l. > .l* - *han.kwuk.chal.ul (han.kwuk.cha.lul) 'Korean car (obj.)', *yel.e (ye.le) 'various'; *l.l > l.* - *tul.le.ol swu (tul.e.ol swu) 'can enter', *sal.lass.ul ttay (sal.ass.ul ttay) 'when one lived'; *l. > l.l* - *kel.yess.sum.ni.ta (kel.lyess.sum.ni.ta) '(sb) was caught', *ol.a.kal ttay (ol.la.kal ttay) 'when going up'.

Other frequent errors in vowels and diphthongs (glides) are substitutions between *wu*-[u] and *o*, *a* and *e*, *way*-[wɛ] and *oy*-[we], and *u*-[ɨ] and *wu*. See the examples in (10), (11), (12) and (13) below.

(10) High- and mid-back vowels (rounded) (2.7%): *wu > o* - *kwa.mwuk (kwa.mok) 'subject' , *sey.mwu (sye.mo) 'triangle'; *o > wu* - *a.i.po.the (a.i.pwu.the) 'from children', *el.kol (el.kwul) 'face'.

---

[9] In the current data, there was no misspelling related to *kh*. This does not necessarily mean that the KFL learners did not confuse *kh* and *k* or *kk*, but it is more appropriate to see in such way that it is reflecting very few *kh*-words in Korean. The *Korean Learner Dictionary*/KLD (Seo et al. 2004), whose entries were based on 26 KFL textbooks, lists only 42 vocabulary items that begin with *kh*, and approximately a half of them are English loan words such as *kha.mey.la* 'camera', *khaym.phe.su* 'campus' and *khem.phyu.the* 'computer'.

(11) Mid- and low-back vowels (unrounded) (2.4%): *a > e* - \*kyeng.ham (kyeng.hem) 'experience', \*can.hwa (cen.hwa) 'telephone'; *e > a* - \*te.um.cu (ta.um.cu) 'next week', \*cin.chel (cin.chal) 'medical examination'.

(12) Mid-front vowel (rounded) and diphthong *way* (2.1%): *way > oy* - \*an.tway.nun (an toy.nun) 'not allowed', \*way.kwuk (oy.kwuk) 'foreign country'; *oy > way* - \*an.toy.se (an tway.se) '(sth) wasn't done/working', \*koyn.chanh.ko (gwayn.chanh.ko) 'it's OK and'.

(13) High-back vowels (1.6%): *u > wu* - \*el.kul (el.kwul) 'face', \*wu.ul.hayss.e.yo (wu.wul.hayss.e.yo) '(I) was sad'; *wu > u* - \*mwu.swun (mwu.sun) 'what sort of', \*twu.sey.yo (tu.sey.yo) 'Please eat'

Lastly, the *h*-related addition or omission constituted another group of error types as shown in (14).

(14) Glottal fricative (2.5%): ø > *h* - \*pa.kkwu.ci.an.ass (pa.kkwi.ci.anh.ass) 'have not changed', \*ta.hayng.i (ta.hayng.hi) 'luckily'; *h* > ø - \*sel.hak.san (sel.ak.san) Mt Selak, \*coh.ha.hay.ten (coh.a.hayss.ten) 'used to like'.

In addition to these types, there were some other error types where one individual error type occurs frequently enough to be noteworthy but the overall occurrence in the pairs was less than 10 times. For example, the types such as *ey for yey* and *e for u* occurred 7 times each but their counterparts, namely, *yey for ey* and *u for e*, occurred only once each.[10] There are a number of miscellaneous errors in the data, but these errors are not considered to have much importance in this study as they occurred rarely and sporadically.

## 6. Patterns of Error Types

---

[10] I call this phenomenon 'leaning' in a separate paper to refer to the phenomenon where in the substitution errors involving two competing phonemes/graphemes one type of substitution occurred significantly more frequently than the counterpart, indicating that the perceived confusion or problem is one directional.

It is apparent that the error types presented in 5.2 above have certain patterns, which can be broadly categorised into five: 1) mismatch in *three-series* stops; 2) mismatch in vowel sounds; 3) misuse of nasals and lateral; 4) omission and addition of 'h'; and 5) miscellaneous errors including transliterations. In the following, the error types are presented in the five categories.

### 6.1 Mismatch in three-series consonants

Among the frequent error types presented above, 29.6% were misspellings of the three-series consonants.   Pairs of the lenis, aspirated and tensed consonant sounds are mismatched, particularly between (a) *c* and *ch*; (b) *t* and *tt*; (c) *t* and *th*; (d) *p* and *ph*; (e) *k* and *kk*; and (f) *s* and *ss*. The mismatch between the lenis palatal *c* and the aspirated palatal stop *ch* produced the single largest number of errors, and the comparative frequency of the pairs was quite similar, with the type *c for ch* was slightly more frequent than *ch for c*. Among the alveolar stops, the type *t for tt* or *t for th* was much more frequent than the opposite type *tt for t* or *th for t*. This shows that there was a clear mismatch between lenis *t* and tensed *tt* or aspirated *th*. This kind of trend in mismatch is also reflected in such pairs as *p* and *ph*, *p* and *pp*, and *k* and *kk*, where the lenis stops were much more frequently used than their aspirated or tensed sounds, while the mismatch between *s* and *ss* occurred with nearly the same rate.

### 6.2 Mismatch in vowel sounds

Six groups of considerably frequent error types (28.5%) constituted vowel misspellings. In these and other vowel error types, vowels that appear to be often perceived as 'similar' sounds by learners were mismatched, particularly between (a) *e*-[ə] and *o*-[o]; (b) *ay*-[ɛ] and *ey*-[e]; (c) *o* and *wu*-[u]; (d) *a* and *e*; (e) *way*-[wɛ] and *oy*-[we], and (f) *u*-[ɨ] and *wu*-[u]. Other pairs of vowel mismatch that are worth noting include *e* and *u*, and *ey* and *yey*. The mismatch between *e* and *o* produced the second largest number of errors overall, followed by the *ey* and *ay* pair.   In terms of the comparative frequency of each pair, error types such as *o for e* and *e for o* was virtually the same in occurrence, while such pairs as *ey* and *ay*, *o* and *wu*, *e* and *u*, *ey* and *yey*, and *way* and *oy* showed a clear difference in which the error types of *ey for ay*, *wu for o*, *e for u*, *ey for yey*, and *wey for oy* were much more frequent than their respective opposite types. Still some other error types such as *u for wu* and *a for e* presented considerable mismatch with the same or slightly higher frequency than their opposite cases.

## 6.3 Misuse of nasals and lateral

Out of the total 635 spelling errors in analysis, there were 49 (7.7%) nasal-related and 38 (6.0%) lateral-related misspellings. In this category of error type, a nasal or lateral sound has an effect upon the neighbouring sound, and as a result, the nasal grapheme was spelled doubly or switched with another nasal and the lateral *l* was also spelled singly or doubly across a syllable boundary. This kind of error occurrence was made particularly between (a) *ng* and *n* (e.g. *ceng.kong* for *cen.kong* 'major study', *kwan.ko* for *kwang.ko* 'advertisement'); and (b) when using *l* between vowels (e.g. *il.lum* for *i.lum* 'name', *ta.la.yo* for *tal.la.yo* 'be different'), and less frequently between (c) *m* and *n* (e.g. *tan.pay* for *tam.pay* 'cigarette'); (d) when using *n* between vowels (e.g. *a.cu.men.ni* for *a.cu.me.ni* 'middle-aged woman') and (e) by adding *ng* to the end of a neighbouring syllable when followed by a syllable which contains *ng* (e.g. *yeng.hayng* for *ye.hayng* 'travel'). Duplications of *l*, *n* and *ng* were apparent in this category.

The most frequent misuse of nasals was the use of *ng* for *n*, which was as frequent as the use of *n* for *ng* at a syllabic final position. The error frequency among the nasal pair *m* and *n* was apparently low, but it is notable that there were four errors in the *n for m* error type while there was none in the opposite type, *m for n*. Another significant misuse of nasal *n* was made by spelling it double, i.e. *n.n for .n*, through the duplication of the same nasal *n* that immediately follows in the next syllable. In the meantime, lateral *l*-related errors produced a number of different error types where the occurrence of each individual error type was relatively low but still the aggregated frequency (28 times) was significant. As in the nasal letters, lateral *l* was spelled single or double under the influence of another lateral, or its syllabic position was switched between vowels. In the current data, however, there was no case where lateral *l* was deleted at the beginning of a syllable, though occasionally observed in KFL learners' compositions, e.g. *tal.yek* for *tal.lyek* 'calendar'.

## 6.4 Omission and addition of 'h'

There were a considerable number of omissions and additions of both consonants and vowels, but a patterned omission and addition that is notable was the intersonorant 'h'. The grapheme 'h' was omitted or added, particularly when it came after, before or between nasals (e.g. *ta.hayng.i* for *ta.hayng.hi* 'luckily', *mek.ci.an.a.se* for *mek.ci.anh.a.se* 'didn't' eat, so…'), or when it came at a syllabic initial, as observed in examples such as *so.wa* for *so.hwa* 'digestion', *sel.hak.san* for *sel.ak.san* Mt Selak, and *chin.ku.hwa* for *chin ku.wa* 'with friends'. All these omissions and additions are related to such phonological phenomena as intersonorant

*h*-deletion or *h*-aspiration. Due to the 'weak' nature of the sound[11], there were many more omissions than additions, and it seems some omissions are predictable when the 'h' sound comes with certain expressions, such as *coh.a.yo* 'It's good' and *manh.a.yo* 'There's a lot' where the 'h' is not pronounced (i.e. [co.a.yo], [ma.na.yo]).   Thus, many 'h' omissions are related to the *h*-weakening between voiced sounds (vowels and consonants *m, n, l, o*) and *h*-deletion at the end of a verb stem in front of a verb sound, while some omissions are influenced by *h*-aspiration where 'h' is 'absorbed into the neighbouring sound, causing aspiration' (Choo and O'Grady 2003: 71) when it occurs before and after a lenis consonant, as in *coh.ta* [co.tha] 'be good/like', *ku.leh.ci.man* [ku.le.chi.man] 'but' and *kup.hi* [ku.phi] 'hurriedly'.

### 6.5 Miscellaneous errors

Some errors did not fit into any of the above types and it was difficult to characterize them since errors in this category were small in number and sporadic. It seems that some miscellaneous errors were related to a wrong graphic association or a wrong pronunciation, rather than simple omissions or additions, but it was difficult to find a particular pattern. Another type of sporadic error that might be taken into account was transliteration of loan words. Though the unorthodox transliterations may not carry as much linguistic significance as the other types of error, the current pedagogical reality does not allow these ill-formed transliterations in the orthography.[12]

### 7. Discussion and Explanation

Thus far I have presented some key areas of perceived difficulty in Korean spellings and establish certain patterns of error types. In the following, I will briefly discuss some distinctive features of the major error types and patterns (involving *c* and *ch*, *ay* and *ey*, and *e* and *o*) and explain the possible causes of these confusions.

---

[11] Kim-Renaud (1997: 166) claims that 'there is enough phonological evidence to regard the glottal approximant /h/ as one of the weakest among consonants.  It often is "invisible" in a phonological environment.'

[12] Some non-standard transliterations had certain features that were found in other error types, but it was difficult to characterize them as a pattern and I did not consider this to be important.   It is worth noting, however, that some transliterations made by Australian KFL learners were perhaps influenced by Australian pronunciations ('Orthodox' transliterations of loan words are largely based on American English, resulting in the transliteration of such loan words as 'bus' as *ba.su* instead of *be.su* and 'dollar' as *dal.la* instead of *dal.le*.). Collins (1989: 4) evaluates Australian English that there exists "a diverse range of varieties – migrant and Aboriginal pidgins, Aboriginal creoles, 'stabilised transference' varieties, and so on – forming a continuum between the different vernaculars and Standard Australian English".

(1) Three-series Consonants: The three-series consonant spellings involve differentiating the sound value of lenis [-tense, -aspirate], aspirated [+tense, +aspirate], and tensed [+tense, -aspirate] consonants, and the failure to do so is largely attributable to the confusion between their sound features.   In Korean, there is a clear distinction in sound quality between *p-, t-, k-, c-* and *s*-series: *p-ph-pp*; *t-th-tt*; *k-kh-kk*; *c-ch-cc* and *s-ss*, since they are all phonemic, producing lenis [-tense, - aspirate], aspirated [+ tense, + aspirated] and tense [+ tense, - aspirated] sounds, respectively. These lenis, aspirated and tense stops all appear at the beginning of a syllable in Korean and the meaning of a word changes according to the initial sound as in *cata* 'sleep', *chata* 'be cold' and *ccata* 'be salt' for *c-ch-cc* series. In English, however, the distinction is based on the feature [voice], i.e. in a manner of voiced or voiceless articulation and the two sound features [aspirate] and [tense] are not so significant, merely playing a "functional" role for allophonic variations (Sohn 1986: 497), while the sound quality of [lenis] is "not shared by English" (Sohn 1999: 154). The patterns of misspellings in the three series of Korean consonants, therefore, are attributable to the confusion among the three phonemic sounds in Korean, and this appears to be caused by phonological mismatches between Korean and English.

It is interesting to observe that there was a much higher rate of confusion between *c* and *ch* (11.2%) than any other pairs. The fact that the error occurrence in both *c for ch* and *ch for c* types was high without much difference between the two types suggests that there is a greater confusion between these types than any other combination among all the error types. It is obvious that confusion arises in differentiating whether a sound is aspirated or not, and this confusion was clearly revealed in the Korean aspirated *ch* and the 'very lightly' or 'slightly' aspirated *c*, resulting in the production of the largest number of errors.[13] It is because there is a certain degree of aspiration in some sounds that are defined as unaspirated in phonology (Kim-Renaud 1997: 164-65). In other words, the matter may be not whether aspirated but the degree of aspiration—the aspirateness of the lenis stop *c*-[c]—and this seems to be particularly the case in *c* and *ch*.[14] In fact, the unaspirated stops, *p*-[p], *t*-[t], *c*-[c] and *k*-[k], are all considered to carry a certain degree of the aspirate feature, which makes it extremely difficult for English L1 speakers to distinguish between the aspirated stops and the lenis stops. The distinction between the slightly aspirated stop *c*-[c] and the aspirated stop *ch*-[ch] can be regarded as more

---

[13] The result coincides with the findings in J-H. Lee (2003) where confusion among *c, ch* and *cc* produced the largest number of errors over beginner, intermediate and advanced levels, and it supports M-O. Kim (2001) where the series was one of the main spelling errors.

[14] Choo and O'Grady (2003: 41) distinguish [c] and [ch] by stating that [ch] has a 'heavy aspiration' and [c] has a 'very light aspiration', and Kim-Renaud (1997:165) makes distinction between the pairs by using terms 'slightly aspirated' and 'heavily aspirated'.

problematic as the descriptive statistical information about the error frequency has revealed. English does not have the lenis stop *c*-[c] in the phonetic quality and there are few or no words with the tensed stop *cc*-[c'] in English but there are thousands of English words that contain the equivalent sound quality of the aspirated stops *ch*-[ch] as in *cheese*, *church* and *choice*. On the other hand, the Korean three-series of *c*-[c], *ch*-[ch] and *cc*-[c'] are distinctive as each of them represents its own phonemic sound.

The main issue in question, therefore, is [aspirate] rather than the [tense] sound, and thus, the focus is on the perceived difficulty in distinction between *c*-[c] and *ch*-[ch] instead of the distinction between *c*-[c] and *cc*-[c'] or between *ch*-[ch] and *cc*-[c']. This is not because the English L1- KFL learners have little or no problem with the distinction between them but because there are much more frequent and substantial problems in the distinction between lenis and aspirated stops and the key element for the problem lies in the perceived aspiration of *c*-[c].

(2) Vowels: A very high frequency of errors was observed in two pairs of vowel spellings: *ay* and *ey* (9.6%), and *e* and *o* (10.2%). These confusions are primarily attributable to the "phonetic closeness" between them (Sohn 1986: 498) and the mismatch in sound features between English and Korean approximate sounds, but the habitual tendency of both Korean L1 and L2 speakers to follow easiness in pronunciation is also related to the frequent erratic productions in written form.[15]

The individual sound quality of *ey*-[e] and *ay*-[ɛ] is distinctive in Korean, and confusions between both sounds are caused generally by the phonetic closeness. Although the distinction is still maintained in initial syllables, especially in careful speech, as in *key* 'crab' versus *kay* 'dog', the distinction is so subtle that even Korean L1 speakers have trouble perceiving it when the *ey*/*ay* contrast is not emphatically made to distinguish one from the other and when there is no context (Choo and O'Grady 2003: 11-12). In this respect, the high frequency mismatch between the spellings *ey* and *ay* is largely due to the perceived similarity or closeness between the pair of vowels. The similarity implies both the perceived similarity in sound and the similarity in form.

At the same time, it is worth noting the differences in sound quality between the Korean and English equivalents as another source of the cause. The sound quality of Korean *ay* does not precisely match with that of English [æ]. The Korean *ey* sound is similar to the vowel in English words such as *bet* and *set*, but the quality of Korean *ay* is "somewhere between the qualities of English *e* in *end* and *a* in *ant*" (Sohn 1986: 497), normally ending with [e] quality. Thus, there is

---

[15] The *ey*-[e] and *ay*-[ɛ] sounds are often merged in casual speech to *ey* or closely to *ey*, even by Korean L1 speakers, and in fact, the contrast between these two sounds disappears in contemporary Korean, though the distinction is still maintained in careful speech by some speakers (Choo & O'Gray 2003).

a strong possibility of interference from English in perceiving and producing Korean *ay*. Such a mismatch in sound quality between Korean /ay/ and an English approximate sound, and a subsequent interference from English would contribute to the cause of confusions and lead to the written production of errors since the learner would have a wrong or confused graphic association of two sound features in his/her mind.

In addition to the phonological aspect, there is also a pragmatic aspect that needs to be addressed. Most Korean L1 speakers, particularly among young adults, pronounce *ay* as *ey* sound almost identically, as a sound closer to *ey*-[e]. For example Korean L1 speakers often pronounce *ay*-[ɛ] as ey-[e] in *nam.may* 'brother and sister', *tam.pay* 'cigarette' and *si.nay* 'city, town'. Although this type of error is not often found in Korean L1 speakers' compositions, it is quite normally used and accepted in their speech. Both speakers and listeners feel little or no 'irritation' on those utterances and are probably more 'comfortable' with such an alteration than the pronunciation based on spellings. Such usual production and wide acceptance is the result of casualisation of speech based on the relative ease of pronunciation. The comparative ease of *ey* is based on its phonemic location as the unrounded mid vowel and it is perceived as 'easier' to pronounce than the low vowel *ay*, thus saving the speaker effort. When Korean L1 speakers frequently or usually pronounce *ay* as [e] on the basis of comparative ease, it is quite possible that the pronunciation by KFL learners is affected and that the written production also reflects such an utterance.

As for the opposite case, i.e. the *ay for ey* type of error, I do not believe that ease of pronunciation has made a great influence on the production of such an erratic substitution, since the mid-front vowel *ey*-[e] is theoretically easier to pronounce than the low vowel *ay*-[ɛ]. And in practice, Korean L1 speakers hardly pronounce *ey* as [ɛ]-*ay* in actual speech and rarely produce *ey* as *ay* in written form, and thus the Korean L1 speaker factor is much less influential than the case of *ey for ay* type of error. In this respect, it could be due to confusions between the approximate sounds/symbols or due to graphic mismatch where the learner knows how to pronounce and yet uses approximate, i.e. wrong spellings.

Another vowel error found in this study was the mismatch between *o*-[o] and *e*-[ə]. The English L1-KFL learners presented a substantial difficulty in differentiating the two vowels (10.2%), with the error frequency rate of the type *o for e* (5.2%) and vice versa (5.0%) being almost the same. The confusions between Korean *o* and *e* are largely due to the mismatch in sound qualities between Korean and English. It seems that the learners perceived the short and monotonous Korean *o* as *e*-[ə], without knowing the difference in sound quality between English *o* and Korean *o*. English *o* carries [w] in its ending sound quality, while the quality of

Korean *o* is simply [o]. The interference from English *o* is further evidenced by a reasonably high proportion of error production in the pair of *wu*-[u] and *o*-[o], where *o* was perceived and transcribed as *wu* twice as many times as the opposite case.

If we look at the other side, however, it is noted that the learners perceived *e* as Korean *o* very frequently (5.2%) as well. First, this confusion seems to be related to phonetic closeness as mid-back vowels and/or phonetic variations wrongly perceived by the learners as if there is a variation in Korean *e* and *o*. There are two mid-back vowels in Korean (i.e. *e*-[ə] and *o*-[o]), while their English approximate sounds are four: two mid-central vowels and two mid-back vowels. In addition, Korean *e* or *o* has little variation in sound value, while its English approximate sound carries phonetic variations between the mid vowels (e.g. between [ʌ] and [ə] as in *just* and *but*). Thus, Korean *e*-[ə] erroneously perceived and transcribed as *o* seems to be due to at least two reasons. KFL learners tend to match the sound value of *e*-[ə] to its perceived variation that they believe is the most approximate sound—which is *o*-[o] articulated in the same location (i.e. mid-back). They also tend to wrongly treat the quality of the monotonous Korean *o* like that of English *o*, which is followed by [w] quality. The effect of this misinterpretation can lead the learners to perceive Korean *o* as closer to *e*-[ə] or *wu*-[u] than other mid-back vowels, *u*-[ɨ] and *a*-[a].

(3) General aspects: In addition to the above explanations, there are some general points that are worth noting in relation to KFL misspellings. Briefly, these include romanization, which is introduced at an early stage in some textbooks, the 'visual' aspects (i.e. 'visual' similarities) as opposed to 'oral' aspects, conjugations of irregular verbs, mora (i.e. the length of a syllable) and sound shift phenomena such as resyllabification. Some of these aspects are closely related to the results of the study and others are generally observed in KFL spelling.

(4) Implications: Pedagogical implications may be drawn from the error types and possible causes discussed thus far. Generally, it is useful for both KFL instructors and learners to have access to such information in that it can be used to raise their awareness of the types of errors and the kinds of (mis)perceptions. However, there is an issue with the KFL spelling process based on phonological interference or pronunciation, and that is, the KFL learners have trouble with identifying the exact match between the target sound in Korean and the perceived sound in their mind. It is often daunting for the KFL learners to distinctly perceive and produce a sound/spelling on the basis of the degree of aspiration, the tenseness or the sound qualities when the sound of the concerned Korean vowel or consonant does not exactly match with any of the English vowels or consonants. A more effective way of teaching and learning may be drawn from the fact that orthography or spelling is not just a matter of spelling or pronunciation but

also a matter of meaning. As a word consists of sound, meaning and symbol, it is necessary to apply these three concepts to spellings. Learners need to be given not only orthographic symbols and phonological explanations about sound features but also a semantic reinforcement which highlights the difference in meaning made by spellings in question. Learners need to be aware whether a syllable they transcribe has a meaning ('sensical') or not ('non-sensical'). In conventional KFL classes, however, a trend accompanying a high production of spelling errors is an unbalanced attention to the three aspects, especially relatively low attention to the semantic aspect.

## 8. Conclusion

The findings of this study indicates the common problem areas in spelling where both instructors and learners need to pay more attention. At the same time, the findings affirm that the mismatch in sound quality and patterns between the two languages was a significant cause that made the English L1-KFL learners become confused with the three-series consonants (especially *c* and *ch*) and various vowel pairs (especially *ay* and *ey*, *e* and *o*). Along with such interlingual factors, the main errors were also caused by intralingual features such as the perceived phonetic closeness of the two sounds and their variations in casual speech. The study reinforces the necessity of in-depth and more comprehensive explanations regarding the key areas in spelling which can provide both instructors and linguists with more concrete information about common trends in perception and production of given error items, which will in turn provide useful information for the future improvement of the materials and teaching methods for spellings in KFL. Whether it is for ordinary teaching or remedial teaching it is necessary to devise pedagogically effective learning and teaching strategies encompassing all three elements: sound, symbol and meaning.

## References

강승혜 Kang, S-H. (2003), "한국어 교육의 학문적 정체성 정립을 위한 한국어 교육 연구 동향 분석", 한국어 교육 14-1, 1-27 쪽.

김미옥 Kim, M-O. (2001), "맞춤법 오류에 관한 연구", 외국어로서의 한국어 교육 25-26, 331-359 쪽.

김미옥 Kim, M-O. (2002), "학습 단계에 따른 한국어 학습자 오류의 통계적 분석", 외국어로서의 한국어 교육 27, 495-541 쪽.

김미옥 Kim, M-O. (2003), "한국어 학습자의 단계별 언어권별 어휘 오류의 통계적 분석", 한국어 교육 14-3, 31-52 쪽.

김영주 Kim, Y. (2006), "영어권 고급학습자의 조사 및 동사 어미 사용 – 오류 분석", 외국어로서의 한국어 교육 31, 129-152 쪽.

김지숙 Kim, J-S. (2007), "베트남어권 학습자의 오류 연구", 한국어 교육 18-1, 49-71 쪽.

이정희 Lee, J-H. (2003), 한국어 학습자의 오류 연구. 서울: 박이정.

조인정 Cho, I-J. (2006), "영어권 한국어 학습자의 주격과 목적격 조사 대치 오류에 대한 비교언어학적 분석", 한국어 교육 17-3, 282-299 쪽.

홍은진, Hong, E-J. (2004). "영어권 한국어 학습자의 어휘 오류 분석과 교육 방안", 이중언어학 25, 331-358 쪽.

Choo, M. and W. O'Grady (2003). *The sounds of Korean: a pronunciation guide*. University of Hawaii Press, Honolulu.

Collins, P. (1989). Sociolinguistics in Australia: a survey. In Peter Collins and David Blair (eds.) *Australian English: The Language of a New Society*. St Lucia: University of Queensland Press, pp. 3-20.

Corder, S. P. (1981). *Error analysis and interlanguage*. Oxford: Oxford University Press.

Finch, G. (2000). *Linguistic terms and concepts*. Macmillan Press.

James, C. (1998). *Errors in language learning and use*. London: Longman.

James, C., Scholfield, P., Garrett, P. and Griffiths, Y. (1993). Welsh bilinguals' English spelling: an error analysis. *Journal of Multilingual and Multicultural Development* Vol. 14 No.4: 287-306.

Kim-Renaud, Y-K. (ed.) (1997). *The Korean alphabet: its history and structure*. University of Hawaii Press, Honolulu.

Richards, J. C. (eds.). (1974). *Error analysis: Perspectives on second language acquisition*. London: Longman.

Selinker, L. (1972). Interlanguage. In J. C. Richards (eds.), *Error analysis: Perspectives on second language acquisition*. London: Longman. 1974.

Seo, S-K. et al (2004). *Korean Learner's Dictionary*. Seoul: Ministry of Culture and Tourism.

Shin, S-C. (2001). An error analysis: Orthographic errors produced by Australian KFL learners. *Teaching Korean as a foreign language*, Vol. 25-26. pp 361-387. KLI Language Research and Education Institute, Yonsei University.

Sohn, H-M. (1986). Patterns of American students' composition errors in Korean. *Linguistic expeditions*. Seoul: Hanshin Publishing Co. 487-528.

Sohn, H-M. (1999). *The Korean language*. Cambridge University Press.

Spillner, B. (1991). *Error analysis: A comprehensive bibliography*. John Benjamins, Amsterdam.