Analysis of English Spelling Errors in a Word-Typing Game

Ryuichi Tachibana, Mamoru Komachi
Tokyo Metropolitan University
6-6 Asahigaoka Hino-city, Tokyo, Japan
tachibana-ryuichi@ed.tmu.ac.jp, komachi@tmu.ac.jp

Abstract

The emergence of the web has necessitated the need to detect and correct noisy consumer-generated texts. Most of the previous studies on English spelling-error extraction collected English spelling errors from web services such as Twitter by using the edit distance or from input logs utilizing crowdsourcing. However, in the former approach, it is not clear which word corresponds to the spelling error, and the latter approach requires an annotation cost for the crowdsourcing. One notable exception is Rodrigues and Rytting (2012), who proposed to extract English spelling errors by using a word-typing game. Their approach saves the cost of crowdsourcing, and guarantees an exact alignment between the word and the spelling error. However, they did not assert whether the extracted spelling error corpora reflect the usual writing process such as writing a document. Therefore, we propose a new correctable word-typing game that is more similar to the actual writing process. Experimental results showed that we can regard typing-game logs as a source of spelling errors.

Keywords: Gamification, Spelling Correction, Information Extraction

1. Introduction

People mainly operate computers through a keyboard interface. However, they often make spelling errors when using a keyboard; e.g., hitting an adjacent key, or using phonologically similar letters. In our study, we define *spelling errors* as (1) a typo from an incorrect keyboard operation, and (2) a spelling confusion from an incorrect identification.

Baba and Suzuki (2012) introduced two types of spelling errors: (1) spelling errors that people do not notice while typing (hereafter called *uncorrected spelling errors*), and (2) spelling errors that people correct while typing (hereafter called *corrected spelling errors*). The former can be a great source of online spelling corrections, while the latter have only been studied recently.

Detection and correction of spelling errors have been studied over the last few decades. In particular, the emergence of the web has created a large amount of consumer-generated texts, which include countless numbers of spelling errors. To overcome the problem of noisy texts, researchers have extracted spelling errors from web chatter by crawling through Twitter data (Aramaki et al., 2010) and from input logs by utilizing Amazon's Mechanical Turk (Baba and Suzuki, 2012). However, the former study does not guarantee the correctness of the extracted spelling errors, and the latter study includes the cost of crowdsourcing.

On the other hand, studies using gamification have become popular in the NLP literature (Deterding et al., 2011; Kumaran et al., 2014; Vannella et al., 2014; Rodrigues and Rytting, 2012). Gamification allows us to obtain resources without paying users by carefully designing a game suitable for information extraction.

In particular, Rodrigues and Rytting (2012) proposed to extract English spelling errors by using a word-typing game. This game includes the intended word and does not require the cost of crowdsourcing. However, the writing process of a word-typing game may differ from the usual writing process (e.g., writing a document).

Thus, one of the purposes of this study was to show that we can regard typing-game logs as a source of spelling errors. We implemented a word-typing game in two different configurations and compared the extracted spelling errors with those of Baba and Suzuki (2012).

The main contributions of this paper are as follows.

- We show that corrected spelling errors extracted from a typing game are similar to the spelling errors extracted from the usual writing process (Baba and Suzuki, 2012).
- Because the typing game allows us to extract correct pairs of intended word and actual input, we can extract more exact uncorrected spelling errors than what previous works did.

2. Related Work

First, we review related work on spelling-error correction. Damerau (1964) showed that over 80 percent of spelling errors are due to a single edit operation. Kernighan et al. (1990) performed spelling-error correction by modeling the probability of correction candidates using the noisy-channel model. They collected spelling errors based on the string-edit distance. Brill and Moore (2000) improved the precision of spelling error correction by extending the edit operation to multiple characters under the noisy-channel framework. Ahmad and Kondrak (2005) created a spelling-error correction model automatically, using web-search query logs. Their model learned the weighted edit distance without labeled data using the EM algorithm. Aramaki et al. (2010) analyzed the cause of spelling errors using Twitter crawled data and constructed a supervised classifier for detecting spelling errors. They collected spelling error candidates with an edit distance of 1. Both their studies and our study analyzed spelling errors; however, we propose a method of extracting error-correction candidate pairs without relying on the edit distance. In addition, their approach cannot extract corrected spelling errors.

Baba and Suzuki (2012) extracted error-correction data by logging users' keystrokes through Amazon's Mechanical Turk. They analyzed spelling errors by extracting pairs of error-correction candidates within an edit distance of 2. Both their study and our study could extract corrected and uncorrected spelling errors; however, we used a gamification approach that does not include a crowdsourcing cost.

Previous researchers have extracted language resources by gamification. Kumaran et al. (2014) extracted phrase equivalents through multiple users playing an online picture-drawing game. Vannella et al. (2014) constructed video games with the purpose of validating and extending knowledge bases. They demonstrated that video game based annotations consistently generated higher-quality annotations than crowdsourcing. Venhuizen et al. (2013) showed how to acquire resources for word sense disambiguation by using a multiplayer game of multiple-choice questions on word senses. Our study and these studies acquired language resources; however, our purpose was to acquire spelling errors, while these studies acquired other language resources. Furthermore, the games in these studies were designed as multi-player games to let users check other users' input because players sometimes cheat. Our typing game is a single-player game since we know the correct word from a dictionary.

Rodrigues and Rytting (2012) utilized typing race games to create spelling error corpora. Both their study and our study adopted typing games as a method to extract spelling errors; however, they did not analyze the difference between corrected and uncorrected spelling errors.

3. Extraction of English Spelling Errors Using a Word-Typing Game

We extracted English spelling errors from the typing-game logs. We acquired spelling errors using dynamic programming to compare the typing logs with the answers.

We implemented two word-typing games. The first one is a common word-typing game similar to that used by Rodrigues and Rytting (2012), and the second one is a new word-typing game in which users can correct their input before hitting the Enter key.

The first word-typing game implemented in this work is similar to that used by Rodrigues and Rytting (2012). Every time a user hits the Enter key, we send to the server the user's input, the correct answer, the time it took for the user to input one word, and his/her name. We used English sets based on Basic English (Ogden, 1930)[1]. We used 842 English words except for one phrase, "according to." Since we designed the typing game to help English-as-a-second-language learners, a translation of the typing word is displayed on the screen.

The second word-typing game uses the same resource and code but allows users to correct their input. Figure 1 shows a screenshot of the correctable word-typing game. The user inputs the English word shown in black. When the user presses a key, the typing game begins. The user's input is displayed in gray. If the user notices a spelling error

Edit type	Correct word	User's input	Spelling error
Deletion	edge	ege	$d \rightarrow \phi$
Insertion	fruit	furuit	$\phi \rightarrow u$
Substitution	able	anle	$b \rightarrow n$
Transposition	flat	falt	$a \leftrightarrow l$

Table 1: Spelling error examples in a word-typing game. (ϕ denotes the empty string.)

Figure 1: Screenshot of the typing game.

while typing, he/she can correct it on the fly by using the backspace key. When the user presses the enter key, the typing game judges whether the user's input is correct. If it is not correct, the user's input includes the uncorrected spelling error. Moreover, we judge whether the error was corrected by the backspace key. If it was indeed corrected, the user's input includes the corrected spelling error. Once the user inputs 50 words, the typing game ends.

We reward users when they type words correctly. If the input is correct, 100 points are added to the score. If not, 10 points are taken from the score. In addition, we facilitate competition and participation to make the game challenging. The faster the user's typing speed is, the higher the user's score. The highest score, the user who achieved it, and the current score are always displayed so that the user is aware of them.

4. Experiments

4.1. Settings

Common Word-Typing Game. Seven university students of computer science played the common word-typing game. We collected 712 uncorrected spelling errors from 4,724 English words.

Correctable Word-Typing Game. Nineteen university students of computer science played the correctable word-typing game. We collected 21,468 English words in three days. We extracted 3,883 corrected spelling errors and 1,334 uncorrected spelling errors. We manually checked the pairs with an edit distance of greater than 2 (Table 2).

[1]http://www.catch.jp/wiki/index.php?english\%2F800_Basic_English

Error type	Correct word	User's input	Token
Corrected	process	prosecc s	22
Accidental	support	f	4
Give up	different	edidere	10
Others	carriage	fjkdsaljfsadf	1

Table 2: Examples of uncorrected spelling errors with an edit distance of ≥ 3.

Edit distance	1	2	3	4 $\geq$
Corrected spelling errors	3,566	148	7	0
Uncorrected spelling errors	1,099	95	15	0

Table 3: Number of strings with spelling errors for each edit distance.

Two pairs with an edit distance of greater than 2 were introduced by computer bugs.

4.2. Results

In the following subsections, we analyze the difference between corrected and uncorrected spelling errors using the correctable word-typing game. We defer the comparison of the two typing games in the next section.

Corrected Spelling Errors. Table 3 shows that most of the corrected spelling errors are within an edit distance of 2, with a few exceptions an edit distance of 3. Although the number of corrected spelling errors with a large edit distance was not large (1% of all the corrected spelling errors), we can see that more than half of them were useful (corrected spelling errors). By augmenting the data by gamification, we can use the data as a source of online spelling correction for complex errors. These instances cannot be extracted by previous approaches like that of Aramaki et al. (2010).

Figure 2 shows that, of the corrected errors, substitution errors were the most frequent, deletion and insertion errors were almost the same, and transposition errors were few. Baba and Suzuki (2012) showed that corrected spelling errors were dominated by substitution, while deletion errors were the most common. They claimed that substitution mistakes were easy to find, while deletion mistakes tended to escape our notice. Because our results are the same as their findings, it follows that we can utilize typing-game logs as a source of spelling errors.

Uncorrected Spelling Errors. Table 3 confirms that most uncorrected spelling errors had an edit distance of 1 (Damerau, 1964). In addition, the ratio of corrected spelling errors to uncorrected spelling errors in our experiment was 3.00, while that of Baba and Suzuki (2012) was 2.94. We regard this as another piece of evidence that typing-game logs contain exact spelling errors.

Figure 2 shows that, in the uncorrected spelling errors, substitution and deletion errors were frequent, but insertion and transposition errors were few. The ratio of uncorrected spelling errors in our experiment differs from that of Baba and Suzuki (2012). We suppose it was because they extracted uncorrected spelling errors using a set of common

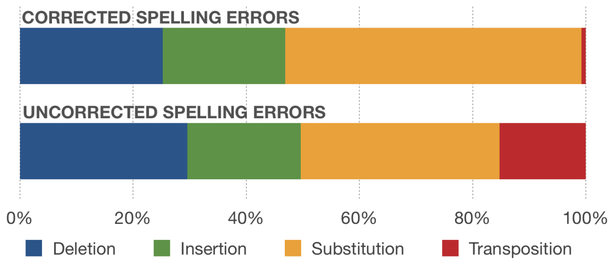

Figure 2: Ratio of spelling errors.

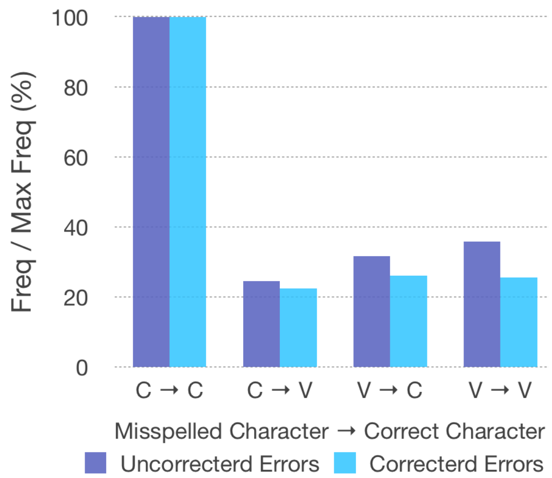

Figure 3: Substitution for consonants and vowels.

spelling errors from Wikipedia[2] and SpellGood[3] while uncorrected spelling errors in our experiment came from actual user input.

5. Discussion

5.1. Common Typing Game vs. Correctable Typing Game

We implemented a correctable typing game to extract corrected spelling errors. However, the difference in the two games may affect the spelling errors. Therefore, we analyzed spelling errors by comparing the common typing game and the correctable typing game.

Tables 4 and 5 show spelling errors with their context in the common typing game and correctable typing game. To investigate the influence of settings for spelling errors, we computed the Pearson correlation coefficients between target errors sorted by frequency in the (1) common typing game, (2) correctable typing game and (3) Baba and Suzuki (2012) dataset. The correlation between (1) and (2) was 0.918, while the correlation between (2) and (3) was 0.886. This shows that the difference in the two games has little effect on the spelling errors. Because our results are similar to Baba and Suzuki (2012)'s findings, it follows that we can utilize typing-game logs as a source of spelling errors.

[2] https://en.wikipedia.org/wiki/Wikipedia:Lists_of_common_misspellings/For_machines
[3] http://www.spellgood.net/

Target error	Frequency	Previous context	Frequency	Correct character	Frequency	Following context	Frequency
e	93	r	14	r	17	e	36
r	73	e	12	e	19	r	26
s	67	_	21	c	17	e	15
i	60	_	9	o	18	i	30
o	59	_	19	p	12	o	24
n	56	i	24	o	22	n	40
a	54	_	14	e	10	a	22
t	54	g	9	h, r	8	t	21
u	35	o	8	y	11	u	14
d	34	_	7	s	10	_	8
g	34	n	11	t	8	_	17
l	31	m	6	p	9	l	21
h	25	_	7	f	5	h, o	5
v	25	_	9	b	12	e	11
p	24	l, t	5	o	16	p	5
y	23	a	4	t	9	y	9
c	20	e, i	4	v	6	c, e	6
k	20	a	8	l	14	_	9
w	20	_	8	e	14	w	5
b	15	i	4	v	9	e	9
m	14	_	4	n	4	m	3
f	11	_	5	r	3	_	3
x	3	e, n, o	1	c	2	e	2
j	3	c	2	k	2	_	2
z	2	e, _	1	a, v	1	i, m	1
q	1	p	1	e	1	c	1

Table 4: Spelling errors in the common typing game (_ in Previous context shows the beginning of a word and _ in Following context shows the end of a word.).

5.2. User's Skill vs. User's Personality

We computed the Pearson correlation coefficients between the times for one keystroke, the ratio of corrected spelling errors, and the ratio of uncorrected spelling errors for each user. The correlation between the time for one keystroke and the ratio of corrected spelling errors was 0.356, that between the time for one keystroke and the ratio of uncorrected spelling errors was 0.329, and that between the ratios of corrected and uncorrected spelling errors was 0.093. The first two correlations show that the faster the usertypes, the more often the user tends to make mistakes. In addition, since there seems to be no correlation between the ratios of corrected and uncorrected spelling errors for each user, the user's personality appears to determine whether he/she corrected his/her errors.

5.3. Phonological Factors vs. Typos

We analyzed the substitution spelling errors. Tables 6 and 7 show confusion matrices for corrected and uncorrected spelling errors, respectively. The most frequent value for each spelling error is shown in bold. By looking at the most frequent value for corrected and uncorrected spelling errors, one can pair 23 and 21 spelling errors with the adjacent correct letters. These facts suggest that spelling errors due to incorrect keyboard operation are more frequent than spelling errors due to phonological causes in a typing game. We also investigated substitution errors for consonants and vowels, as in Baba and Suzuki (2012). Figure 3 shows the ratio of substitution errors for consonants and vowels normalized by the total number of consonants. Baba and Suzuki (2012) showed that the ratios of vowel-to-vowel errors were quite high in the uncorrected spelling errors; however, in our experiments there was little difference between corrected and uncorrected spelling errors. Moreover, the most frequent errors were not the vowel-to-vowel (V→V) errors but the consonant-to-consonant (C→C) errors in corrected spelling errors. Since consonants are more common than vowels, it also seems that spelling errors caused by incorrect keyboard operation dominate the spelling errors in a typing game.

6. Conclusion

We extracted English spelling errors using a word-typing game as in Rodrigues and Ritting (2012). We implemented two types of word-typing games that allow analysis of corrected and uncorrected spelling errors. Experimental results showed that typing-game logs can be regarded as a source of spelling errors. Our typing game logs are available at GitHub[4]. In the future, we would like to extract spelling errors using a language-learning application/service like Duolingo[5].

7. Acknowledgment

We appreciate the 19 people who played the game. This research was (partly) supported by the Grant-in-Aid for Re-

[4] https://github.com/hiyuricu/open_resource/blob/master/typing_game_logs.txt

[5] https://www.duolingo.com/

Target error	Frequency	Previous context	Frequency	Correct character	Frequency	Following context	Frequency
e	738	_	88	r	126	e	303
a	387	_	90	e	68	a	200
o	359	_	71	i	69	o	195
i	322	_	53	t	49	i	169
t	314	a	62	r	63	_	122
r	308	_	55	e	86	r	129
n	262	i	74	m	35	n	121
s	260	_	102	c	52	s	43
d	239	_	51	s	50	_	71
_	213	e	47	t	53	_	208
l	197	_	38	a	26	l	110
c	185	_	44	v	47	e	54
u	182	o	45	r	26	u	62
h	160	_	41	c	36	h	86
g	154	_	40	d	29	_	51
w	141	_	47	e	49	w	42
p	138	_	23	o	64	p	38
k	129	o	33	l	66	k	47
f	100	_	40	g	32	_	17
y	99	a	17	r	23	y	58
v	95	_	36	c	42	e	34
b	85	_	28	v	32	e	34
m	73	_	13	n	28	m	25
x	35	e	18	c	19	e	8
j	18	_	7	k	5	_	5
q	13	_	7	w	6	a	3
z	11	e	3	x	6	_	6

Table 5: Spelling errors in the correctable typing game (_ in Target error shows that users forget to input trailing strings, _ in Previous context shows the beginning of a word and _ in Following context shows the end of a word.).

search on Priority Areas, Tokyo Metropolitan University, “ Research on Social Big Data. ”

8. References

Ahmad, F. and Kondrak, G. (2005). Learning a spelling error model from search query logs. In *Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing*, pages 955–962.

Aramaki, E., Uno, R., and Oka, M. (2010). Typo writer: How humans conduct spelling errors? In *Proceedings of 16th Annual Meeting of Association for Natural Language Processing*, pages 966–969.

Baba, Y. and Suzuki, H. (2012). How are spelling errors generated and corrected? A study of corrected and uncorrected spelling errors using keystroke logs. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 373–377.

Brill, E. and Moore, R. C. (2000). An improved error model for noisy channel spelling correction. In *Proceedings of the 38th Annual Meeting of the Association for Computational Linguistics*, pages 286–293.

Damerau, F. J. (1964). A technique for computer detection and correction of spelling errors. *Communications of the ACM*, pages 171–176.

Deterding, S., Dixon, D., Khaled, R., and Nacke, L. (2011). From game design elements to gamefulness: Defining “gamification”. In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, MindTrek’11, pages 9–15.

Kernighan, M. D., Church, K. W., and Gale, W. A. (1990). A spelling correction program based on a noisy channel model. In *Proceedings of the 13th Conference on Computational Linguistics-Volume 2*, pages 205–210.

Kumaran, A., Densmore, M., and Kumar, S. (2014). Online gaming for crowd-sourcing phrase-equivalents. In *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*, pages 1238–1247.

Ogden, C. K. (1930). *Basic English: A general introduction with rules and grammar*. London: Paul Treber & Co., Ltd.

Rodrigues, P. and Rytting, C. A. (2012). Typing race games as a method to create spelling error corpora. In *Proceedings of the 10th International Conference on Language Resources and Evaluation*, pages 3019–3024.

Vannella, D., Jurgens, D., Scarfini, D., Toscani, D., and Navigli, R. (2014). Validating and extending semantic knowledge bases using video games with a purpose. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, pages 1294–1304.

Venhuizen, N., Basile, V., Evang, K., and Bos, J. (2013). Gamification for word sense labeling. In *Proceedings of the 10th International Conference on Computational Semantics (IWCS-2013)*, pages 397–403.

Spelling error	Frequency of correct letter																									
	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
a	0	0	5	0	**28**	0	2	0	4	0	2	2	1	0	17	0	5	5	27	3	6	0	2	1	0	0
b	0	0	4	3	0	0	1	7	1	0	1	0	2	7	0	1	0	0	0	0	0	**21**	0	0	0	0
c	1	4	0	5	1	5	5	0	0	0	2	0	1	3	1	0	0	2	25	4	0	**32**	0	9	0	1
d	1	4	22	0	6	27	7	1	0	0	6	2	1	1	0	0	0	3	**44**	8	0	2	1	1	1	0
e	**62**	2	2	3	0	0	0	1	19	0	1	1	1	2	19	0	0	48	11	16	9	2	40	0	3	0
f	0	1	4	17	0	0	**27**	3	0	2	3	0	0	0	1	1	0	5	1	2	0	0	0	0	1	0
g	0	7	9	18	1	**21**	0	13	0	1	7	0	0	0	2	0	0	1	0	12	0	3	1	0	0	0
h	1	7	1	0	1	7	7	0	3	3	7	0	0	1	0	0	0	2	0	0	**9**	0	0	0	3	0
i	5	0	0	2	14	0	2	1	0	0	0	4	1	3	**23**	1	0	2	2	6	15	1	1	0	3	0
j	0	0	1	0	1	**2**	0	**2**	0	0	1	**2**	0	1	0	0	0	0	0	0	0	0	0	0	0	0
k	1	0	1	4	0	1	0	1	1	0	0	**49**	2	0	3	0	0	0	0	3	0	0	0	0	0	0
l	0	0	4	3	2	2	1	0	4	1	**13**	0	4	1	6	6	0	1	3	5	0	1	0	0	0	0
m	0	0	4	0	3	0	0	0	0	0	1	0	0	**18**	1	0	0	0	0	0	0	0	3	0	0	0
n	1	12	10	0	4	0	9	3	1	0	2	2	**24**	0	2	1	0	1	9	2	3	5	0	1	1	0
o	16	0	3	1	6	1	0	2	**29**	0	0	5	2	3	0	23	0	1	3	1	4	1	1	0	2	0
p	1	1	0	1	2	4	0	0	0	0	0	3	0	3	**43**	0	1	4	1	2	0	1	1	0	0	0
q	1	0	0	0	1	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	**6**	0	0	0
r	1	0	2	3	**38**	6	0	2	1	0	0	4	1	6	5	3	0	0	6	23	1	2	7	0	1	0
s	26	1	**40**	20	9	5	4	0	2	0	2	4	1	3	0	0	0	5	0	12	0	0	7	2	1	2
t	2	0	5	21	5	5	8	1	1	0	0	5	3	0	2	3	0	**30**	7	0	0	0	2	1	15	0
u	4	0	1	0	0	0	1	0	**14**	0	0	2	2	4	4	1	0	5	0	3	0	1	7	0	11	0
v	1	29	**37**	0	3	1	1	0	0	0	0	0	2	2	0	0	0	1	5	0	0	0	0	0	0	0
w	1	0	2	0	**33**	0	0	0	0	0	0	2	0	0	0	2	14	10	1	3	2	1	0	0	0	0
x	0	0	**12**	2	0	0	0	0	1	0	1	1	0	0	0	0	0	1	2	1	0	0	0	0	0	3
y	0	0	1	0	4	0	0	3	0	0	0	0	0	0	0	0	0	2	0	**11**	**11**	0	0	0	0	0

Table 6: Confusion matrix for corrected spelling errors in Substitution.

Spelling error	Frequency of correct letter																									
	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
a	0	0	0	0	9	0	0	1	1	0	0	1	0	3	8	1	0	0	**13**	0	0	0	0	0	0	0
b	1	0	0	0	0	0	1	2	0	0	0	0	0	5	0	0	0	0	0	0	1	**9**	0	0	0	0
c	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	7	1	0	**10**	0	1	0	0
d	1	2	1	0	5	5	4	0	0	0	4	1	0	0	1	0	0	2	2	**8**	0	0	0	0	0	0
e	**19**	0	2	1	0	0	0	2	7	0	0	1	0	0	4	1	0	5	2	4	3	0	11	0	1	0
f	0	0	1	0	0	0	**4**	0	0	0	0	0	1	0	0	0	0	0	1	1	0	0	0	0	1	0
g	0	1	0	4	1	4	0	1	1	0	0	0	0	0	0	0	0	0	0	**8**	0	0	0	0	0	0
h	0	0	0	0	0	1	1	0	2	0	1	0	0	1	0	0	0	2	0	0	**3**	0	0	0	2	0
i	1	0	1	0	2	0	0	0	0	0	0	0	0	0	3	0	0	0	1	0	**4**	0	0	0	0	0
j	0	0	0	0	0	0	0	1	0	0	**2**	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
k	0	0	0	1	0	0	0	0	0	0	0	**5**	0	0	1	0	0	0	0	0	0	0	0	0	0	0
l	1	0	0	0	**3**	0	0	0	0	0	**3**	0	0	0	2	0	0	1	0	1	0	0	0	0	0	0
m	0	1	0	0	1	0	0	0	0	0	0	0	0	**7**	0	0	0	0	0	0	0	0	0	0	0	0
n	0	2	2	1	2	1	0	2	1	0	0	2	**4**	0	0	0	0	0	0	0	0	0	0	0	0	0
o	4	1	0	0	3	0	1	0	**9**	0	0	1	0	0	0	7	0	0	0	0	0	1	0	0	0	0
p	0	1	0	0	0	0	0	0	0	0	0	0	0	0	**6**	0	0	0	0	0	0	0	0	0	0	0
q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
r	0	0	0	0	5	0	0	0	0	0	0	2	0	0	0	1	0	0	0	**6**	2	0	1	0	0	0
s	5	0	**6**	3	1	0	0	0	0	0	0	1	0	1	1	0	0	1	0	0	0	0	3	0	0	0
t	1	0	2	6	5	3	4	2	0	0	0	1	0	0	0	1	0	**13**	1	0	0	0	0	0	11	0
u	0	0	0	0	1	0	0	0	**6**	0	0	0	0	**6**	1	0	0	0	0	0	0	0	3	0	4	0
v	0	3	**4**	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
w	0	0	0	0	**4**	0	0	0	0	0	0	0	0	0	0	1	0	1	1	2	0	0	0	0	0	0
x	0	0	**6**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
y	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	**3**	0	2	0	0	0	0	0	0
z	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0

Table 7: Confusion matrix for uncorrected spelling errors in Substitution.