Grotjahn, Rüdiger. (Hrsg.). (1996). Der C-Test. Theoretische Grundlagen und praktische Anwendungen (Bd. 3, S. 139-157). Bochum: Brockmeyer

Undine Roos

The reconstructability of Japanese characters: some new evidence*

In adapting the C-Test principle to the Japanese language we are confronted with various problems, such as how to define a word or what the second part of a character is. Research has shown that for a Japanese C-Test the lower part of every second character should be deleted.

In this context the following problems arise: (1) To what extent can testees reconstruct characters presented without context? (2) Is *kana* easier to reconstruct than *kanji*, which is usually thought to be more complex?

In the present paper it was possible to demonstrate that 38.6% to 74.7% of the items – depending on the deletion pattern – could be reconstructed properly when presented context-free and that k*anji* is indeed more difficult to reconstruct than *kana*.

1. Introduction

The idea for the research described in this paper arose when dealing with the problem of adapting the C-Test principle to the Japanese language. The C-Test and the underlying principle will not be discussed in detail here; for more information cf. e.g. Grotjahn (1992, 1994, 1995, 1996a, b), Grotjahn, Klein-Braley & Raatz (in press) and Klein-Braley (1994).

Roos (1990, 1994, 1995) proposed that a Japanese C-Test should be constructed as follows:[1]

> "... beginning with the second character of the second sentence the lower part of every second character is deleted." (Roos, 1995, p. 124)

* I am very grateful to Prof. Dr. Katsuo Tamaoka and Barbara Menzel from Matsuyama University (Japan), and Prof. Dr. Takashi Seki from Ôsaka University (Japan) who provided informants for the research described.

[1] There are rather few articles dealing with the C-Test principle in Japanese or in Chinese, where similar problems arise. Apart from the works just mentioned there is: Noma (1985), Shin (1990) and Tamaoka & Tanaka (in prep.) for Japanese; Arras & Grotjahn (1994) and Zhang (1985) for Chinese.

In applying this principle there are at least two problems to be solved:

1. Is it possible that testees reconstruct a character correctly without taking the context into consideration, i.e. do testees complete a character properly when only a part of it is presented context-free?[2]
2. Are the complex *kanji* more difficult to reconstruct than the less complex *kana*?

Before dealing with these problems I shall briefly describe the Japanese writing system.

2. The Japanese writing system[3]

The Japanese writing system is composed of *kana* (*hiragana* and *katakana*) and *kanji*. *Kanji* are Chinese characters, which generally represent the morphemes of the language, whereas *hiragana* are Japanese characters used mainly to express grammatical relations (postpositions, verb suffixes, etc.). *Katakana* are also Japanese characters; they are used to write all foreign words which were brought to Japan in the course of time. While the number of *kanji* characters in daily use is between 2.000 and 3.000, the number of *kana* characters is 104 for each system (*hiragana* and *katakana*).

The following example will illustrate the differences between *kanji*, *hiragana* and *katakana* (from: Steinberg & Yamada, 1978, p. 90):

The word

"karate" (空手)

consists of two morphemes:

"kara" (empty) and "te" (hand).

This word is generally written in *kanji* (空手).

However, it is also possible to write it in *hiragana*, for example if you do not know the correct characters. Then, it looks like this:

からて.

[2] When the solutions of the context-free versions are evaluated as "properly" or "correctly", what is meant is "exactly as in the complete text".

[3] This section is for the non-expert; those who are familiar with Japanese may skip it. For more detailed information cf. e.g. Coulmas (1980), Miller (1980), Roos (1995, Chapter 7) and Steinberg & Yamada (1978).

But you would never write it in *katakana*. An example of a *katakana*-word might be "terebi"

テレビ,

which is derived from the English word "television".

Kanji may have *hiragana* added to them, as for example in verbs and adjectives, i.e. they are mixed to form words. An example of this is

大きな ("large").

The *hiragana* and the *katakana* systems on the other hand are never mixed. Either *hiragana* or *katakana* is used to write a word. Thus, the word "terebi" is typically written in *katakana*, very rarely in *hiragana* but definitively never in a mixture of both.

These examples show that *hiragana* and *katakana* are less complex and have fewer strokes than *kanji* and are therefore easier to write.

3. The Japanese C-Test

Problems which arise when adapting the C-test principle to the Japanese language include the definition of the concept "word" and how to handle the postpositions. These will not be discussed here; this is done in detail in Roos (1994, 1995). In short one can say that there are theoretically at least 11 possibilities of constructing a Japanese C-Test. In Roos (1994, 1995) only 7 versions were taken into consideration; these are versions in which the *second* part of a character/word was always deleted, taking the order of writing into consideration. The versions 8 – 11 are added in the present paper; they are based on deletions of the *first* part of the characters/words.[4] The 11 versions are described in Figure 1. The unmutilated sentence runs as follows:

英語には、はるかな昔から伝えられていることわざが、何百とある。

(Translation: In English, there are hundreds of proverbs which came from ancient times.)

[4] Experiments with the deletion of the first half of a word are described in Boonsathorn (1987), Cleary (1988), Köberl & Sigott (1994) and Sigott & Köberl (1996).

Figure 1
The 11 possibilities of constructing a Japanese C-Test

Identification	Description
V1	The right half of every second character is deleted.
Example:	英語には、はるかな昔から伝えられている ことわざが、何百とある。
V2	The lower half of every second character is deleted.
Example:	英語には、はるかな昔から伝えられている ことわざが、何百とある。
V3	Every second character is emliminated.
Example:	英_に_、は_か_昔_ら_え_れ_い_ こ_わ_が、_百_あ_。
V4	Going by the order of strokes when writing a character and counting them the second half is deleted; characters consisting of only one stroke are not counted.[5]
Example:	英語には、はるかな昔から伝えられている ことわざが、何百とある。
V5	The number of strokes of Hiragana and Katakana is halved and for the Kanji only the root is left.
Example:	英語には、はるかな昔から伝えられている ことわざが、何百とある。
V6	One makes divisions into words, regarding postpositions as bound forms and deletes the second half of every second word.
Example:	英語には、はる___昔から伝えら_____ ことわざが、何___ある。

[5] Note that every Chinese and Japanese character must be written in a fixed order of strokes. The character は, e.g. must be written: (1) 丨; (2) 𠂆; (3) は.

Figure 1
continued

V7	One makes divisions into words, regarding postpositions as free markers, and deletes the second half of every second word.
Example:	英語には、は_かな昔か_伝えられている こと___が、何_とあ_。
V8	The upper half of every second character is deleted.
Example:	英語には、はるかな昔から伝えられている ことわざが、何百とある。
V9	The left half of every second character is deleted.
Example:	英語には、はるかな昔から伝えられている ことわざが、何百とある。
V10	One makes divisions into words, regarding postpositions as bound forms, and deletes the first half of every second word.
Example:	英語には、___かな昔から_____ている ことわざが、___とある。
V11	One makes divisions into words, regarding postpositions as free markers, and deletes the first half of every second word.
Example:	英語には、_るかな昔_ら伝えられている ___わざが、_百と_る。

Four Japanese texts were prepared in these 11 versions (the unmutilated texts are shown in Appendix 1) and presented to native speakers of Japanese at Ôsaka University (versions 1 to 7; 282 testees) and Matsuyama University (versions 8 to 11; 75 testees). The results for versions 1 to 7 can be found in Roos (1994, 1995), those for versions 8 to 11 – necessary for the solutions to the problems discussed here – are presented later in this paper.

In summary one can say that it is version 2 – where the lower part of every second character was deleted – which can be reconstructed best (i.e. with 99.7% accuracy) by native speakers of Japanese.

The C-Test principle underlying the "optimal" version 2 leads to the first problem: to what extent do testees take the context into consideration when reconstructing the characters?

4. Description of the experiment with context-free characters

4.1 Context-free versions

A context-free test was constructed in 5 different versions, which correspond to the C-Test versions 1, 2, 4, 8 and 9 in Figure 1. All these versions are based on characters. The versions 3 and 5, which are also based on characters, were skipped because in version 3 the whole character was eliminated and in version 5 only the root of the *kanji* was presented and it is senseless to present those items context-free.

The context-free test consisted of both *kanji* and *kana* items and can be described thus:

Versions 1, 2, 8, 9: 39 *kanji* and 30 *kana* (27 *hiragana* and 3 *katakana*), i.e. 69 items in total[6]

Version 4: 39 *kanji* and 23 *kana* (20 *hiragana* and 3 *katakana*), i.e. 62 items in total.

The 5 context-free versions are presented in Table 1 (the complete versions are shown in Appendix 2).

The different versions were randomly assigned to the testees who were native speakers of Japanese. They were asked to complete the characters without being given any information except whether the items were *kanji* or *kana*.[7]

[6] The Japanese C-Test investigated consists of 4 texts with 100 items altogether. Some characters, however, had to be reconstructed twice or more. As context-free characters they were only presented once. Thus, there were no items with two or more possible correct solutions. For example, in version 9 *kanji* item 7 _ and *kanji* item 8 _ were treated as two different characters (職 with the meanings "employment, job, work, office" and 識 with the meanings "knowledge, discrimination, writing"). For item 7 only 職 and for item 8 only 識 was scored as correct since we wanted to know whether the *expected* character would be supplied.

[7] The completions in the C-Tests based on deletions of characters showed that a *kanji* item was **never** completed as *kana* and vice versa. When items are presented in a running text, the testee can infer whether it is *kanji* or *kana* (cf. Section 2). When, however, parts of a character are presented context-free it is sometimes impossible to decide whether it is *kanji* or *kana*; e.g. the item _ can be completed among others as: 工 (*kana*) or 下 (*kanji*).

Table 1
Description of the 5 context-free versions

Version	Description	Example[8] Kanji	Hiragana	Katakana	No. of Testees
1	Deletion of the right part of the characters.	識	す	コ	22
2	Deletion of the lower part of the characters.	識	ふ	下	21
4	Deletion of the second half of the strokes in the order they are written. If a character consists of an odd number of strokes the deletion pattern is (n+1)/2.	識	力	一	23
8	Deletion of the upper part of the characters.	識	ゞ	土	22
9	Deletion of the left part of the characters.	識	ゔ	匚	22

4.2 Subjects

The context-free tests were given to 110 subjects in total (39 male and 71 female) all from Matsuyama University (Japan).

4.3 Scoring

The tests were scored as follows: 1 point for a correct answer and 0 points for a wrong answer. "Correct" were only solutions which corresponded exactly to the unmutilated character (cf. footnote 6).

[8] The characters in their unmutilated form are: 識 (*kanji*), 力 (*hiragana*) and 工 (*katakana*).

4.4 Results

4.4.1 General tendencies

Table 2 shows the means in percent (*Mean*) and standard deviations (*SD*) for *kanji*, *kana* and the total for the context-free tests as well as the means in percent (*Mean*) for the C-tests 1, 2, 4, 8, 9 (total in context).

Table 2
Means (*Mean*) and standard deviations (*SD*) for *kanji*, *kana* and the total test for the context-free versions and the means (*Mean*) for total in context

	Context-free						In context
	kanji		*kana*		Total		Total
Version	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*
1	82.98	5.28	57.11	9.33	70.08	6.07	98.8
2	68.27	6.36	69.64	4.93	68.98	3.21	99.7
4	47.70	7.48	29.48	8.49	38.62	6.45	95.2
8	52.79	11.18	62.85	4.62	57.84	6.92	96.8
9	70.95	6.61	78.78	10.13	74.73	6.98	98.5

Subtracting the mean total solution percentages for the context-free presentation from that of those with context we obtain:

Version 1: 98.8% – 70.08% = 28.7%
Version 2: 99.7% – 68.98% = 30.7%
Version 4: 95.2% – 38.62% = 56.6%
Version 8: 96.8% – 57.84% = 38.9%
Version 9: 98.8% – 74.73% = 24.1%

It can be seen that for the "optimal" version 2 the difference between "presentation in context" and "without context" is only 30.7%. In version 4, however, the difference is 56.6%. This means that obviously a large number of items in the "optimal" version used in Roos (1994, 1995) may have been solved correctly without taking the context into consideration, whereas for version 4 the impact of context appears

to be more distinct. In the light of this, it must be asked whether the C-Test principle underlying version 2 is really "optimal" for the construction of a Japanese C-Test.

In the next section it will be tested among other things whether the mean correct context-free solutions of these two and of the other versions differ significantly.

4.4.2 Kanji versus kana

In this section I shall try to determine whether there is a difference (a) between the total number of correct answers in the 5 context-free versions, and (b) between the number of correct solutions of *kanji* and *kana* within each version. This is a design for calculating a multivariate analysis of variance, but as several assumptions underlying this procedure (e.g. that of homogenity of the variance-covariance-matrices) are not fulfilled in the present data set and because of the exploratory nature of the research presented here the hypotheses were tested with the help of (matched) *t*-tests adjusting the significance-level according to Bonferroni-Holm (cf. Holm, 1979 and Stevens, 1990, p. 137).

In Table 3 (a) all 10 possible pairs of mean correct answers were tested. There is a significant difference between all totals of the context-free versions (except for versions 1 and 2) which means that for example the characters in version 4 are more difficult to reconstruct than those of version 2. We can thus conclude that when solving a C-Test in version 4 the context must definitively be taken more into consideration than when solving a C-Test in version 2.

Table 3
(a) *t*-tests for independent samples

	Version 1	Version 2	Version 4	Version 8	Version 9
Version 1	–	p = .466	p < .001 **	p < .001 **	p = .023 *
Version 2		–	p < .001 **	p < .001 **	p = .002 **
Version 4			–	p < .001 **	p < .001 **
Version 8				–	p < .001 **

* significant on α = .10 (adjusted)
** significant on α = .05 (adjusted)

Table 3 (continued)
(b) *t*-tests for paired samples

	Difference *kanji* – *kana*
Version 1	p < .001 **
Version 2	p = .513
Version 4	p < .001 **
Version 8	p < .001 **
Version 9	p = .001 **

** significant on α = .05 (adjusted)

In Table 3 (b) the results of the tests of the differences between *kanji* and *kana* items within each version are presented. Except for version 2 there is a significant difference between the mean correct answers of *kanji* and *kana*.

Going back to Table 2 and analyzing the means of *kanji* and *kana* we can see that in versions 8 and 9 (upper resp. left part deleted) the *kana* items could be more easily reconstructed than the *kanji* items. In versions 1 and 4 (right part deleted resp. deletion according the order of strokes) however, the more complex *kanji* items could be significantly better reconstructed than the *kana* items. But there is a simple explanation for this observation: For the *kana*-syllables:

"ka, ki, ku, ke, ko" (か、き、く、け、こ／カ、キ、ク、ケ、コ)[9]
"sa, shi, su, se, so" (さ、し、す、せ、そ／サ、シ、ス、セ、ソ)
"ta, chi, tsu, te, to" (た、ち、つ、て、と／タ、チ、ツ、テ、ト)
"ha, hi, fu, he, ho" (は、ひ、ふ、へ、ほ／ハ、ヒ、フ、ヘ、ホ)

voiced counterparts of the voiceless sounds can be obtained by adding a voicing-marker (").[10] For the k-sounds, e.g., we therefore get:

か、き、く、け、こ ⇒ が、ぎ、ぐ、げ、ご
(ka, ki, ku, ke, ko ⇒ ga, gi, gu, ge, go).

Deleting the right part of such characters as done in version 1, the voicing-marker always disappears, and as it is always written last, the voicing-marker also disappears when deleting the second part of a character based on the total number of strokes as done in version 4. When reconstructing a character which can have a

[9] The characters left from the slash are *hiragana*, those right are *katakana*.

[10] In the h-row, the voiced counterparts are "ba, bi, bu, be, bo"; the h-row is the only one where also a "half voicing-marker" (°) can be used; the results are the p-sounds.

voicing-marker the context must be taken into consideration. If this is not done, it is impossible to decide whether the character is voiced or voiceless. Skipping the 6 voiced *kana* items in versions 1 and 4, we get for version 1 a mean of 87.16% and for version 4 a mean of 49.49%, which reestablishes the relation.

However, with version 8 (deletion of the upper part) it is slightly different: In this version the voicing-markers are also deleted but nonetheless the *kana*-items could be reconstructed significantly better than the *kanji* items (62.85% *kana* and 52.79% *kanji*). This is due to the fact that we have in this version more correct answers in the unvoiced items than in versions 1 and 4. It thus seems that when the upper part of a character is deleted it can better be reconstructed than when the right half is deleted or when there is a deletion on the basis of the total number of strokes.

5. Summary and conclusion

The aim of the research described here was to solve the following problem: "Do native speakers of Japanese complete characters properly when they are presented context-free?" It was possible to demonstrate that the testees were able to reconstruct between 38.6% and 74.7% of the items correctly – depending on the deletion pattern. When the first half of strokes of characters was presented, only 38.6% could be reconstructed in the expected form, whereas when the right part of a character was presented, 74.7% could be reconstructed correctly. In the terms of *Gestalt*-Psychology (cf. e.g. Guss, 1977 & Katz, 1969) this means that the tendency to complete *parts of characters* is more distinct than completing *missing strokes*; in the last case the unity of the characters is more disturbed and therefore more difficult to reestablish.

On the basis of these results the deletion pattern for a Japanese C-Test must be reconsidered: On the one hand, we have version 2 – the deletion of the lower part of every second character – which could be solved to 99.7% correctly by native speakers of Japanese; but 68.9% could be restored without context. In an investigation with *learners* using this C-Test form there was a highly significant correlation with Level 2 of the *Japanese Language Proficiency Test*, thus providing an indication for criterion validity.

On the other hand, we have version 4 – the deletion of the second half of strokes according to their writing order. This could be solved to 95.2% by native speakers of Japanese and also fulfils the requirement of Raatz & Klein-Braley (1985) that

natives should score around 95%. This version has the advantage that only 40.9% of the items could be reestablished without context.

Certainly text difficulty is also an important factor: The *kanji* characters of the four texts had a mean difficulty of 3.83 (cf. Roos, 1995, p. 105), which means that they tend to belong to group A4 of the *Jôyô-kanji*[11] list and are comparatively easy, since they are learnt in primary school.

Further research with different learners and texts must be carried out to show which really is the "optimal" C-Test principle for the Japanese language.

References

Arras, Ulrike & Grotjahn, Rüdiger. (1994). Der C-Test im Chinesischen. In Grotjahn (1994), 1-60.

Boonsathorn, Somsak. (1987). *C-Tests, proficiency and reading strategies in ESL.* Ph.D. diss, Univerity of Alberta, Canada.

Cleary, Christopher. (1988). The C-Test in English: left-hand deletions. *RELC Journal, 19,* 26-37.

Coulmas, Florian. (1980). Schriftentwicklung, Schriftverarbeitung: Herkunft und Funktionsweise der japanischen Schrift. *Zeitschrift für Semiotik, 2,* 361-374.

Grotjahn, Rüdiger. (Hrsg.). (1992). *Der C-Test. Theoretische Grundlagen und praktische Anwendungen. Band 1.* Bochum: Brockmeyer.

Grotjahn, Rüdiger. (Hrsg.). (1994). *Der C-Test. Theoretische Grundlagen und praktische Anwendungen. Band 2.* Bochum: Brockmeyer.

Grotjahn, Rüdiger. (1995). Der C-Test: State of the Art. *Zeitschrift für Fremdsprachenforschung, 6*(2), 37-60.

Grotjahn, Rüdiger. (1996a). The C-Test bibliography: version December 1995. In Grotjahn (1996b).

Grotjahn, Rüdiger. (Hrsg.). (1996b). *Der C-Test. Theoretische Grundlagen und praktische Anwendungen. Band 3.* Bochum: Brockmeyer.

[11] *Jôyô-kanji* means: Characters for general use. The list, which can be found e.g. in Kenbô, Kindaichi, Kindaichi & Shibata (1982), is dated from 1981. It includes 1945 characters which – according to the Japanese Ministry for Education – are the most important and the most frequently used ones. There are the following categories:
A1-A6: The 996 characters to be learnt in the primary school;
中: The 812 characters to be learnt in the secondary school;
高: The 42 characters to be learnt in the high school;
Z: The last 95 characters to be learnt in everyday life.

Grotjahn, Rüdiger, Klein-Braley, Christine & Raatz, Ulrich. (in press). C-Tests: An overview. In James A. Coleman (Ed.), *University language testing and the C-Test. Proceedings of a conference held at the University of Portsmouth in April 1995.* Portsmouth: University of Portsmouth.

Guss, Kurt. (Hrsg.). (1977). *Gestalttheorie und Fachdidaktik.* Darmstadt: Steinkopff.

Holm, Sture. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6,* 65-70.

Katz, David. (1969). *Gestaltpsychologie.* Basel/Stuttgart: Schwabe.

Kenbô, Hidetoshi, Kindaichi, Kyosuki, Kindaichi, Haruhiko & Shibata, Takeshi. (Eds.). (1982). *Sanseido kokugojiten.* Tôkyô: Sanseido.

Klein-Braley, Christine. (1994). *Language testing with the C-Test. A linguistic and statistical investigation into the strategies used by C-Test takers, and the prediction of C-Test difficulty.* Unpublished Habilitationsschrift Universität Duisburg.

Köberl, Johann & Sigott, Günther. (1994). Adjusting C-Test difficulty in German. In Grotjahn (1994), 179-192.

Miller, Roy A. (1980). *The Japanese language.* Tôkyô: Tuttle.

Noma, Kazuko. (1985). *The C-Test in Japanese. A pilot study.* Centre for Applied Language Studies, University of Reading (England). (Ms.).

Raatz, Ulrich & Klein-Braley, Christine. (1985). How to develop a C-Test. In Christine Klein-Braley & Ulrich Raatz (Eds.), *Fremdsprachen und Hochschule 13/14: Thematischer Teil: C-Tests in der Praxis* (pp. 20-22). Bochum: AKS.

Roos, Undine. (1990). *Das C-Test-Prinzip und das japanische Schriftsystem. Eine empirische Untersuchung zur sprachspezifischen Anwendungsproblematik.* Unpublished MA thesis at the Faculty of Philology at the Ruhr-Universität Bochum.

Roos, Undine. (1994). The C-Test in Japanese. In Grotjahn (1994), 61-113.

Roos, Undine. (1995). *Ein C-Test für Lerner der japanischen Sprache. Entwicklung, Erprobung und Validierung.* Bochum: AKS.

Shin, Kimikei. (1990). Nihongo Cloze-Test-kara nihongo henkei C-Test-e. *Gengo bunka ronshû (Nagoya daigaku sôkei gengo sentâ) 11*(2), 213-225.

Sigott, Günther & Köberl, Johann. (1996). Deletion patterns and C-Test difficulty across languages. In Grotjahn (1996b).

Steinberg, Danny D. & Yamada, Jun. (1978). Are whole word kanji easier to learn than syllable kana? *Reading Research Quarterly, 14*(1), 88-99.

Stevens, James. (1990). *Intermediate statistics: A modern approach*. Hillsdale, NJ: Erlbaum.

Tamaoka, Katsuo & Tanaka, Tomoko. (in prep.). Nihon-no shakai – bunkateki-na shûkan-o hyôgen shita jukugo nincho-o kaishita bunka rikai-no sokutei. Universities of Matsuyama and Hiroshima (Japan). (Ms.).

Zhang, Lian Zhong. (1985). *The C-Test in Chinese*. Centre for Applied Language Studies, University of Reading (England). (Ms.).

Appendix 1: The four Japanese C-test-texts

Text 1

子供達は幼年学校に入学した後も、家庭において正確な知識のストックを増やし得るような場に直面する状況が続く。しかしながら、彼らの知識は未だ無秩序で取り留めがない。幼年学校の教師にとって最初の職務は、こうした知識をより秩序づけ、より正確なさせることであり、まだ子供達の語彙を増やし、言葉がより正確な意味をもつように手伝うことである。子供達のうち幾人かは、ほかの子供達が家庭環境を通じて得るような、知識的バックグラウンドを与え得るような経験を必要とするであろう。

After children entered the infant school they will continue to meet situations in their home lives which will increase their store of mathematical knowledge; but this knowledge will still be unordered and vague. The first function of the teacher in the infant school is to make this knowledge more ordered and more precise and to help children to build up a vocabulary in which the words have a greater precision or meaning. Some children will need to have experiences which will give to them that background knowledge which others have gained from their home environment.

Text 2

大変濫用される言葉である、‘リアリスム’には少くとも四つの基本的意味がある。しかし小説に関して用いられるときは、普通、日常生活の精密な模写を意味する。リアリズム小説と言えば、会話は口語体で読者が登場人物等を思い浮かべることが出来るものを言う。この意味では、現代小説の殆どすべてが過去の小説よりも、‘リアリスム’的である。何故なら、日常生活の措写と自然な調子の会話の構成と言うものは世代を重徐々に進歩する技術的な点にかかっているからである。

"Realism", a much abused word, has at least four current meanings, but when applied to novels it normally means a photographic imitation of everyday life. A "realistic" novel is one in which the dialogue is colloquial and physical objects are described in such a way that you can visualize them. In this sense almost all modern novels are more "realistic" than those of the past, because the describing of everyday scenes and the construction of natural-sounding dialogues are largely a matter of technical tricks which have passed on from one generation to another, gradually improving in the process.

Text 3

鼠たちは、何も好んで破壊主義者である訳ではない。彼らは、唇の上に突き出て、急速に伸びる門歯を削るために、かじり続ける必要があるのである。さもなくば、歯が肉に食い込むか、あるいはあごが開かなくなり、餓死してしまうであろう。それで、丁度口にあう電線をかじっては、戸口のチャイムを鳴らしたり、電気器具のヒューズを飛ばしたりするのである。更に悪いことには、不審火の多くは鼠が電線の絶縁体を食いちったことに端を発していると考えられている。飛行場管理者が機上における鼠の害に気を使うことは並大抵ではなく、もし鼠の糞がひとつでも見つかろうものなら、飛行機の運行を停止し、消毒するほどである。

Mice need not enjoy all that they destroy. They have to keep gnawing to file down the fast-growing incisor teeth that project beyond their lips. Otherwise, they would die of starvation, their teeth ingrown or jaws locked tight. Since electric wire is such a convenient jawful, mice amuse themselves by ringing doorbells and by fusing appliances. More seriously, many mysterious fires are thought to have been started by mice stripping insulation from wires. Aircraft managers are so alarmed at the damage mice might do on board that the sight of just one dropping can lead to an aircraft being taken out of service for fumigation.

Text 4

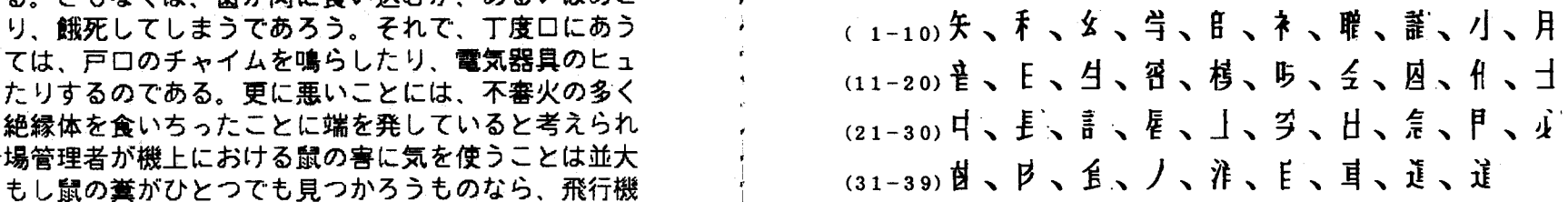
いつもは五人が消防自動車にのっている。一人は消防自動車を運転します。指導者はドライバーのよこに座っています。他の消防士は車の中に座っています。長い間は指導者はいつも消防署にいた。いろいろな種類の火をどのように処置するかしっている。そのまま消防士は火事場に着くときに、どんなふうに火と戦うかいつも指導者は決定します。何をするか消防士に話す。

There are usually five men in the crew of a fire engine. One of them drives the engine. The Leader sits besides the driver. The other firemen sit inside the cab on the fire engine. The leader has usually been in the Fire Service for many years. He will know how to fight different sorts of fires. So, when the firemen arrive at a fire it is always the leader who decides how to fight the fire. He tells each fireman what to do.

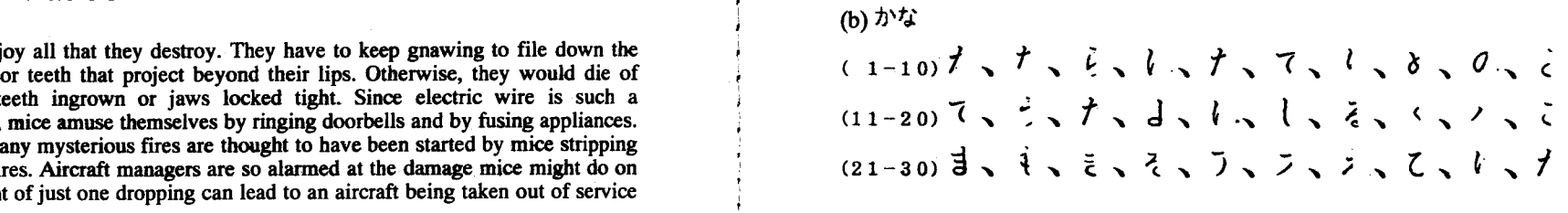
Appendix 2: The 5 context-free versions

Version 1: Deletion of the *right* half of the characters

(a) 漢字

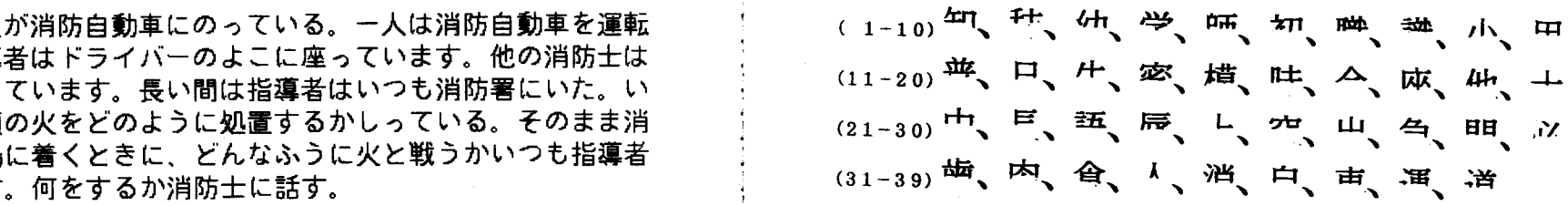


(b) かな

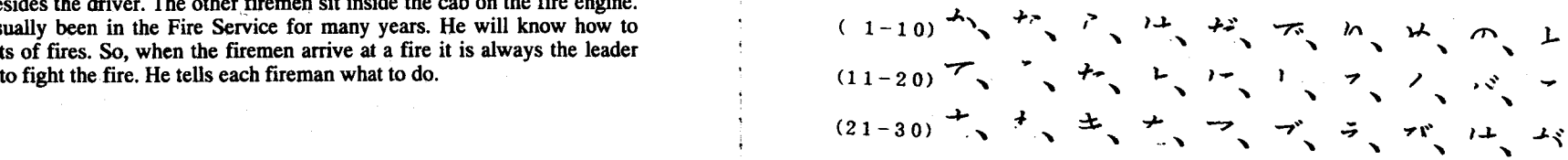


Version 2: Deletion of the *lower* half of the characters

(a) 漢字

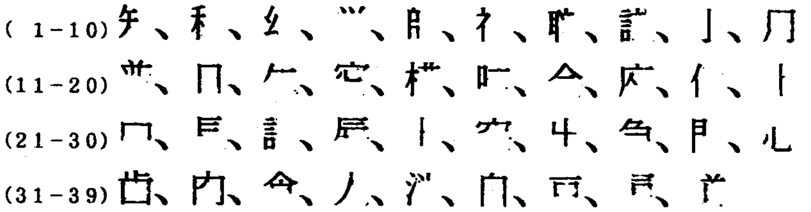


(b) かな

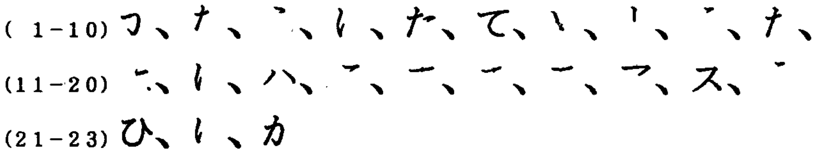


Version 4: Deletion of the *second half of the number of strokes*

(a) 漢字

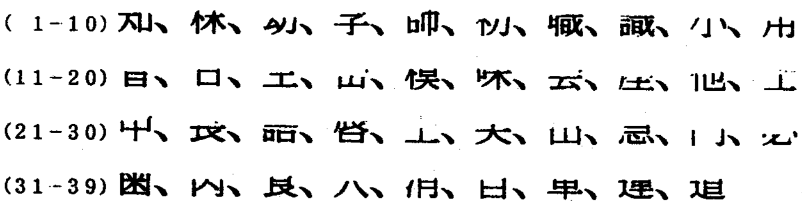


(b) かな

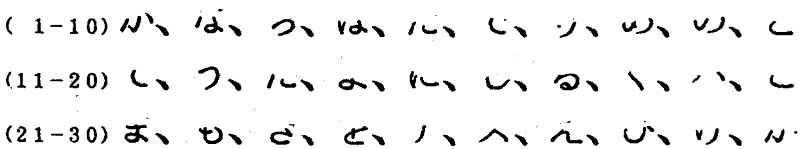


Version 8: Deletion of the *upper* half of the characters

(a) 漢字

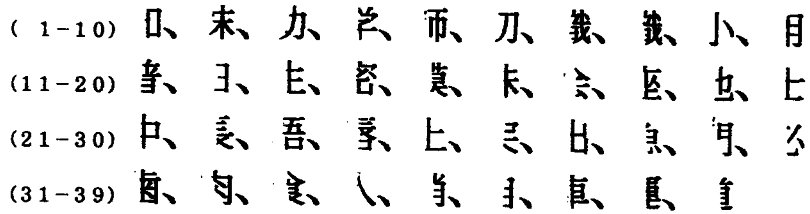


(b) かな

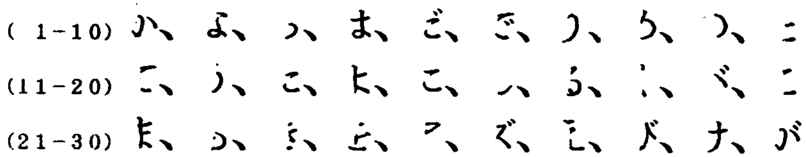


Version 9: Deletion of the *left* half of the characters

(a) 漢字



(b) かな



Solutions

(a) 漢字

(1-10) 知、秣、幼、学、師、初、職、識、小、用

(11-20) 普、日、生、密、模、味、会、座、他、士

(21-30) 中、長、語、唇、上、突、出、急、門、必

(31-39) 歯、肉、食、人、消、自、車、運、道

(b) かな (The underlined characters consist of only one stroke and were therefore skipped in version 4)

(1-10) か、な、ら、は、だ、で、り、め、の、と

(11-20) て、う、た、よ、に、し、る、く、バ、こ

(21-30) ま、も、き、を、ア、ズ、え、び、け、が