

# AN ANALYSIS OF SIMPLIFICATION STRATEGIES IN A READING TEXTBOOK OF JAPANESE AS A FOREIGN LANGUAGE

**Kristina HMEJAK SANGAWA**

University of Ljubljana, Slovenia

kristina.hmeljak@guest.arnes.si

## Abstract

Reading is one of the bases of second language learning, and it can be most effective when the linguistic difficulty of the text matches the reader's level of language proficiency. The present paper reviews previous research on the readability and simplification of Japanese texts, and presents an analysis of a collection of simplified texts for learners of Japanese as a foreign language. The simplified texts are compared to their original versions to uncover different strategies used to make the texts more accessible to learners. The list of strategies thus obtained can serve as useful guidelines for assessing, selecting, and devising texts for learners of Japanese as a foreign language.

**Keywords:** readability; simplification; Japanese as a foreign language; textbook analysis; reading

## Povzetek

Branje je eno od temeljev učenja drugega ali tujega jezika in je lahko posebej učinkovito, ko jezikovna težavnost besedila ustreza bralčevemu nivoju jezikovnega znanja. Članek nudi pregled dosedanjih raziskav na področju berljivosti in poenostavljanja japonskih besedil ter predstavlja analizo zbirke poenostavljenih besedil za učence japonščine kot tujega jezika. Iz primerjave poenostavljenih besedil z njihovo originalno različico izhaja seznam različnih strategij, ki so jih pisci uporabili, da izboljšajo dostopnost besedil za učence. Seznam teh strategij lahko služi kot iztočnica za ocenjevanje, izbiranje in sestavljanje besedil za učence japonščine kot tujega jezika.

**Ključne besede:** berljivost; poenostavljanje; japonščina kot tuj jezik; analiza učbenika; branje

## 1 Introduction

Reading is one of the bases of second language learning, and it can be most effective for the purpose of improving a reader's language skills when the text being read is not only appealing but also of the appropriate difficulty level for its reader. The development of reading skills through extensive reading can be supported on one hand by selecting appropriate material from existing texts and grading them according to objective or subjective readability criteria; and on the other hand by



adapting existing texts to the level of the intended audience, i.e. simplifying and abridging existing material. Both approaches have been extensively researched and implemented for major languages, especially English (see for example DuBay 2004 for an overview), and some research has been conducted on the readability and simplification of Japanese texts. However, factors affecting readability for learners of Japanese as a second language have not been thoroughly researched yet.

In the present paper, after reviewing previous research on the readability and simplification of Japanese texts, a collection of texts that have been simplified for L2 Japanese learners is analysed and compared to the original Japanese texts from which the simplified versions were adapted, in order to investigate text characteristics that could be considered important factors in determining the readability of a text for readers of Japanese as a foreign language.

## **2 Readability and simplification of Japanese texts: previous research**

Research in this area stems from different backgrounds and is targeted at different groups of weak readers. Some early work does not clearly specify in which context and by whom the measure is meant to be used, but most research is targeted either at young native speakers or at persons with disabilities.

Probably the earliest work on readability of Japanese texts (Morioka, 1952), inspired by Flesch's Reading Ease formula, reports on preliminary research at the National Institute for the Japanese Language to determine the criteria needed to develop a similar formula for Japanese. Regrettably, the project seems to have been discontinued.

Another early attempt at measuring the readability of Japanese texts was made by Sakamoto (1962), who manually analysed Japanese language textbooks for elementary school grades 1 to 6, using school grades as the scale of difficulty, and found that the ratio of frequent vocabulary, sentence length and the proportion of kanji characters in the text correlate with school grades.

A similar way of estimating the difficulty of written sentences is also proposed in a writing stylebook by Yasumoto (1983), who uses the average number of characters per sentence and the percentage of Chinese characters as indicators of text difficulty, but does not combine these two factors into a single formula.

Two decades after Sakamoto's research, when computers were already available for lengthier calculations, Tateishi et al. (1988a, 1989b) proposed the first readability formula for Japanese on the basis of four surface characteristics: the proportion of types of characters (Roman letters, hiragana, katakana and kanji); the length of continuous strings of the same type of character; the length of sentences; and the number of commas per sentence.

A more recent and very productive stream of research is work on readability formulae to predict the difficulty level of texts for Japanese school-children, to be

used in mother tongue education (Shibasaki & Tamaoka, 2010). Other formulae have also been developed by Sato et al. (2008, for young native speakers of Japanese), and Lee & Hasebe (2016, for learners of Japanese as a foreign language).

Another approach to the readability of Japanese, proposed by Sano and Maruyama (2008) is based on Halliday's concept of lexical density within the framework of Systemic Functional Grammar (Halliday, 1993). In this approach, lexical complexity is defined as the ratio of content words to ranking clauses in a text.

A second stream of research on Japanese readability is work on information accessibility and text simplification, aimed at facilitating communication with handicapped and elderly readers (Ichikawa, 2006), and paraphrase generation to assist handicapped readers with limited linguistic capabilities (Yamamoto et al., 2000, Inui and Yamamoto 2001, Inui and Fujita 2004, Nakano et al. 2005, Sato et al. 2004).

A third stream of research which bears on readability is work on computer-aided text revision, where readability criteria are used to highlight potentially incomprehensible passages and suggest more readable substitutions (Hayashi, 1992, Inui and Okada 2000, Ono et al. 2006, Oono and Inazumi 2007). These projects often use advice on clear writing from style manuals such as Kabashima 1979, Kinoshita 1981, Honda 1982, Mishima 1990 etc., which do not deal with numerical measurements of readability, but give hints on what factors can affect readability and should be considered in its measurement.

Linguistic factors which have been found to correlate with text readability in previous research can be divided according to the traditional levels of linguistic analysis: script (ratio of character type, punctuation, phonetic guides etc.), vocabulary, syntax (sentence length, clause length, ellipsis etc.), text and discourse (length, cohesion etc.). Statistical correlations between these factors and collections of graded texts have been described in previous research. However, the factors influencing the readability of a text for learners of Japanese have not been yet thoroughly researched. The following sections present a comparison between simplified texts and their originals and an analysis of the strategies used in this process.

### **3 Data**

The texts are reading passages in a textbook for intermediate learners of Japanese, the second in the set of textbooks developed by the International Student Center of Sanno University: 日本語を楽しく読む本・中級 *Enjoyable task reading in Japanese: Intermediate*, published by Bonjinsha in 1991 and reprinted multiple times, a popular textbook for reading instruction.

These texts were chosen because they are one of the very few available collections of pairs of authentic and simplified Japanese texts targeted at foreign learners of Japanese.

The reading passages included in the textbook were selected and simplified by the textbook authors, experienced teachers of Japanese as a foreign language. Selection criteria, as stated in the foreword, were: content (that should be interesting to adult learners of Japanese: worth reading, intellectually challenging, ), text type (as varied as possible, including narratives, expository and scientific writing, in order to offer learners the opportunity of practicing different reading strategies). Another criterion that is not stated in the foreword but was evidently applied, is length: not exceeding the length that can be read in a 90-minute lesson. The longest texts are approximately 1300 characters long, spanning one to two pages.

The foreword mentions that texts were rewritten for their target audience, learners of Japanese, while the afterword mentions that all textbook material was developed and used for two years in the Japanese course of Sanno University before being re-edited for publication in book form. Only vocabulary is mentioned in the foreword as a simplification criterion, but it is conceivable that strategies applied to text rewriting were based on the authors' experience as language teachers and empirically verified or found to be useful in their language classes.

The textbook has been used by the present author for Japanese language instruction in a class of 2nd year students of Japanese and received a positive response from the students, indicating that it is a good example of readable writing for students of Japanese. The pairs of texts were also given to read to a group of eight advanced learners of Japanese, who were asked to choose the easier in a pair of texts, the original and its simplified version. All participating students indicated the simplified versions as the easier to read.

In the forewords to each textbook in the series (*この本を使う先生へ To the teachers using this book*), the authors mention vocabulary as the main criterion used both when assessing the difficulty of texts included and also when rewriting texts for their intended readers. All texts in this textbook series are graded from one to three stars (one star indicating the easiest texts and three stars indicating the most difficult texts) according to the percentage of vocabulary included in the text but not present in the vocabulary lists used as a yardstick, and thus expected not to be known by readers at the given level. These percentages and vocabulary lists are shown in Table 1 on the next page.

**Table 1:** Vocabulary lists used as yardstick and percentage of new vocabulary

Textbook level	No. of expected known words	Vocabulary list used as yardstick	Words not covered by yardstick list in passages marked by stars:		
			☆	☆☆	☆☆☆
Pre-intermediate (1996)	2000	list of 2030 words in Nihongo kyōiku no tame no kihon goi chōsa (National Language Research Institute 1984)	5-6%	6-10%	10-20%
Intermediate (1991)	3500	unpublished vocabulary list compiled by the authors' research team	5%	5-10%	10-15%
Pre-advanced (1993)	6000	list of 2030 words in Nihongo kyōiku no tame no kihon goi chōsa (National Language Research Institute, 1984)	up to 30%	up to 30%	30% or more
		list of 6000 words in Nihongo kyōiku no tame no kihon goi chōsa (National Language Research Institute, 1984)	up to 15%	15% or more	15% or more

In the preface to the first volume the authors explicitly mention that vocabulary was overall the main criterion used in assessing text difficulty, while adding that the average length of sentences was also used as a secondary indicator of structural complexity, but no concrete data are given for these aspects of complexity. In the prefaces to the second and third volume in the series, only vocabulary is mentioned as the yardstick for assessing text difficulty.

Similarly, the level of proficiency which is expected from readers of each of the three volumes is defined in terms of hours (or months) of Japanese instruction received, which is supposed to reflect their vocabulary knowledge: readers who have studied Japanese for a certain period of time are expected to know a certain number of words, which should approximately correspond to the vocabulary prescribed for a certain level of the Japanese Language Proficiency Test (JF and AIEJ, 2004).

Table 2 shows the number of words which readers (learners of Japanese) are expected to know after different periods of study, as stated in the forewords.

**Table 2:** Expected proficiency of learners of Japanese using the textbook series *Enjoyable Task Reading in Japanese*

Volume	Expected Japanese instruction time	Expected No. of known words	Expected JLPT level of vocabulary knowledge
Pre-intermediate	7 months – 1 year or 400 - 600 hours	2000–3500	2nd–3rd level
Intermediate	9 months – 1 year or 450 - 600 hours	3500–5000	2nd–3rd level
Pre-advanced	1 year – 1 year and a half or 600 – 900 hours	5000–10000	1st level

As can be seen from the above descriptions, the authors of the textbooks have carefully controlled the vocabulary used in the reading passages, considering it the main factor of text difficulty.

Each reading passage is also preceded by lists of 10 to 20 keywords used in the text, with exercises to learn or reinforce vocabulary knowledge, including written form (Chinese characters), morphological, syntactic and collocational patterns, again emphasising the importance of depth and breadth of vocabulary knowledge for reading comprehension.

The textbook is divided into 9 chapters, each chapter containing one or two reading passages. All reading passages were analysed, except reading no. 4, where the original text was in English and only the simplified text used as reading material was in Japanese.

The following table presents the data used in this analysis: the titles of the simplified passages as they appear in the textbook, the title of their originals, the length of both (expressed in number of characters) and their sources.

**Table 3:** Analysed data: simplified texts in *Enjoyable Task Reading - Intermediate* and their originals (length expressed in number of characters)

Chap.	Title of simplified text	Length	Original title	Length	Source
1	禁酒	449	禁酒	352	『月刊アサヒ』1989年6月号、朝日新聞社 p. 342
2	いつもとちがう	682	いつもとちがう	527	『ポケットジョーク9トラベル』(植松黎編訳) 角川書店 p. 18

Chap.	Title of simplified text	Length	Original title	Length	Source
3	賢い農夫	1327	賢い百姓	1285	矢崎源九郎『世界の民話』（矢崎源九郎訳編）社会思想社 pp. 255~256
5	振り向き賃	526	振り向き賃	909	深田祐介『ちょっといい旅いい話』（深田祐介監修）旅行開発株式会社ジャルパック出版センターpp. 188-189
6	ママ、手が凍るんだよ 行康君、手すり変わるよ	707 829	ママ、手が凍るんだよ 行康君、手すり変わるよ	702 922	朝日新聞 1985年2月20日朝刊 朝日新聞 1985年3月1日朝刊
7	みんなって何人？	1277	みんなって何人？	2640	斉藤勇『人間関係の分解図』誠信書房 pp. 73~80.
8	握手 1 握手 2	1147 851	握手 1 握手 2	2408 1199	阿部謹也『逆光のなかの中世』日本エディタースクール出版部 pp. 97~100 樋口清之『日本の風俗の謎』大和書房 pp. 34~35
9	趣味	1219	趣味	1972	守屋毅『日本文明 77 の鍵』（梅悼忠夫編）創元社 pp. 136~138
	<b>Total:</b>	<b>9014</b>	<b>Total:</b>	<b>12916</b>	

#### 4 Procedure and results

All pairs of original and simplified texts were scanned, OCR-processed and the resulting files were manually checked to correct OCR errors. Pairs of files were then automatically compared using the document comparison software JDiff X (Matsumoto, 2010), all differences found were transcribed into a spreadsheet file and marked according to type, linguistic level and content of modification. Modifications within the same sentence or clause which stem from different rewriting strategies, or are carried out at different linguistic levels, were counted separately. For example, the following rewriting of one original sentence into shorter sentences involved multiple strategies at distinct linguistic levels.

Original sentence:

「亡くなった飲み友だちと約束してね、僕が飲みに行くときは、必ずオレの分も注文して飲んでくれという遺言を実践しているんだ」

Literally: *Having made a promise to a drinking pal who's dead, I'm executing his will that says, "always order my part too and drink it" when I go drinking.* – Quotation marks are not used in Japanese.

Equivalent modified sentences:

「先週、僕の親友が亡くなったんだが、彼が亡くなる前に約束してね」

Literally: *A good friend of mine died last week, and having made a promise before he died ...*

「はあ」

Literally: *Oh.*

「僕が飲みに行く時は、必ず、彼の分も注文して飲むということになったんだ。それで、その約束を実行しているってわけさ」

Literally: *... it was decided that when I go drinking, I would always order his part too and drink it.*

Firstly, one strategy was simplification at the syntactic level: both adnominal clauses in the first and last part of the original complex sentence (亡くなった飲み友だち *drinking pal who died*; ... オレの分も注文して飲んでくれという遺言... *the will that states always order my part too and drink it...*) were split into separate simple sentences, avoiding adnominal modification, a known hurdle for learners of Japanese.

Secondly, a simplification at the discourse level retained the same entity (僕, the first person narrator) as the subject of all clauses, avoiding the shift from the first person narrator (僕 *boku* - I) to the dead friend (using a more informal first person pronoun (オレ *ore* - I) within a short clause of reported speech (not marked by quotes: 必ずオレの分も注文して飲んでくれ *ore no bun mo chuumon shite nonde kure* - *always order my part too and drink it*) and then back to the first person narrator (遺言を実践している *yuigon o jissen shite iru* - *I am executing his will*), which could be confusing. This simplification also brought with it the omission of the very informal pronoun オレ (*ore* - I), thus resulting in a standardisation of register.

Thirdly, at the semantic level, two pieces of information were made more explicit: a concrete time setting (先週 *senshuu* - *last week*) was added, and the pronoun 僕の (*boku no* - *my*) was added to the noun 飲み友だち／親友 (*nomitomodachi / shin'yuu* - *drinking pal / good friend*).

Fourthly, at the level of vocabulary, three simplifications were made by substituting the less common word 飲み友だち (*drinking pal*) with a less specific but



more common one: 親友 (*good friend*), and the words 遺言を実践 (*yuigon o jissen - execute a will*) with 約束を実行 (*yakusoku o jikkô - keep a promise*).

Fifthly, two explicitations occurred at the script and punctuation level: the word と き (*toki - when*) written in hiragana was rewritten with its commonly used and unambiguous Chinese character 時, and a comma was added after the adverb 必ず (*kanarazu - always*).

One further boundary explicitation was carried out by inserting a back-channelling expression (はあ *Haa - Oh*) by the other participant in the conversation, in the middle of the longest remaining sentence.

In all such cases of multiple modifications, each modification was counted and transcribed separately, resulting in a list of 815 modification occurrences in the whole corpus. All transcribed modifications are reported in Appendix 1. Repeated modifications of the same item: e.g. the rewriting of わたし as 私 for three times in the same text was counted as 3 modifications; in such cases, the modification was transcribed once and the number of modifications was noted in the second column of the table in the appendix.

## 5 Analysis

Differences found between the simplified texts and their originals were grouped into three categories: simplification (including deletion), explicitation and standardisation (including visualisation). Strategies belonging to these categories were found to be used at different levels of linguistic analysis: from script, to vocabulary, morphology, syntax, semantics, to discourse and style. Let us consider each one in turn.

### 5.1 Simplification

Strategies of simplification were the most commonly used, amounting to 472 (of which 96 deletions) out of the 815 modifications found.

#### 5.1.1 Script simplification

Script simplification occurred a few times, where non-standard or low-frequent Chinese characters were rewritten with hiragana:

尋ねた。 → たずねた。; 年を取った → 年をとった; 全然 → ぜんぜん; 土産 → おみやげ; 嫌がります → いやがります; 出来た → できた; 皆様 → みなさん; 頑張った → がんばった; 頑張りなさい → がんばりなさい; 暖かく → あたたかくて; 間もなく → まもなく; 誰と誰 → だれとだれ; 石鹸 → せっけん; 挨拶 → あいさつ.

### 5.1.2 Vocabulary simplification

Vocabulary simplification was the most frequent of all changes, obtained by:

- substituting less common with more common **content words**:

誰しも → 多くの人; 珍奇な → 珍しい; 花をさかす → 花を作る; 法則が確立される → 法則が発見された; 奇妙な感慨にとらわれている → 不思議に思った; 育種学 → 植物学; 事柄 → こと; 実用についで → 使って; 食糧 → 食べ物; 増産する → 生産を増やす; 奇怪な → 珍しい; 熱中した → 夢中になった; 高度な → 難しい; 数式 → 数学の問題; 受容する → 受け入れる; しかるべき先生 → 専門の先生; 特殊な階層 → 特殊な人々; 介して → 通して; 大衆化し → 広がり; 国民的な趣味に拡張した → 国民的な趣味になった; その道の達人 → 一流の人たち; 生活のたし → 生活の助け; 先生連中 → 先生たち; 一覧表 → リスト; 現在 → 今; 盛況をきわめている → 流行している; カルチャー・センターのカリキュラム → カルチャー・センターの内容; かたっぱしから → 何でも; 都市大衆 → 庶民; 愛好した → 愛した; 絆 → 関係; 手をさしのべる → 手を差し出す; 仕来り → 習慣; 述べる → 言う; 口上 → 言葉; 大変巧みに → じょうずに; 握手にこたえて → 握手して; 帰宅した → 家に帰った; あわてて → 急いで; ぬぐって → ふいて; 古来の → 昔からの; etc.

- substituting less frequent **functional vocabulary** (which is usually learned later in courses of Japanese as a foreign language), with more basic expressions:

のみ → だけ; かならずしも...というわけではない → ...ばかりではない; といえは → は; しかも → そして; といったような → というような; のもとで → では; において → では; にもかかわらず → けれども; のおかげであった → があったからだ; にあたっており → である; ようにおもわれる → と思う; ...にせよ、...にせよ → 例えば、...でも、...でも; に対しても → にも; やたらと...したがる → すぐに...したがる; のだが → のに; 通っているかにみせねばならない → 通っているようにみせなければならない; ...すら → ...も; 当然と言えば当然でしょう → 当然かもしれません; 帰りがけ → 帰る時; ...握られっぱなしで → その間...握られていて; やたらと → すぐに; etc.

Difficult words are sometimes substituted with an explanation or definition: 古典 → 昔の文学作品; sometimes even substituted with words with a completely different meaning, if it does not change the overall gist of the text, such as 語学 instead of the less frequent word 手工芸 as an example of a hobby:

現代日本人も、あらたにみいだした文化を、テニスにせよ、手工芸にせよ、かたっぱしから「お稽古ごと」にしてしまう → 現代日本人も、現代

の文化を（例えば、テニスでも、語学でも）、何でも「お稽古ごと」にしてしまう。

Often, the substitution of a vocabulary item brings with it also syntactical modifications, such as:

- changes in part of speech, e.g. from noun to verb and vice-versa: の教授をして → を教えて; あらたにみいだした文化 → 現代の文化; ヨーロッパ伝来の → ヨーロッパから入って来た; 複数の相手のばあいには → 相手が何人かいる場合は; 婦人が優先される → 婦人たちから始める; 右手を失った人 → 右手のない人; or from verb to adverb to express a modal meaning: 相手はダメダメと首をふるに決まっています。 → 相手は、必ず、ダメダメと首を振ります; or from noun to adverb: 文化の実態が...であった → 文化が、実は、...であった;
- shortening from phrase to word, e.g.:  
まるでいいあわせたかのように → 必ず;
- argument distribution, resulting in the use of different particles:  
日本が西洋とあらためて接触した → 日本に西洋文化が入って来た; 日本人の間にもかなり浸透している → 日本人の間でもかなり一般的になっている; 間をもたせるのに苦労する → どうしたらいいか困る; 石鹸だらけであった → せっけんがたくさんついていた; 一体となる意識を目覚めさせ → 相手と一つだという意識を持つため; 四枚目ので女房を養います → 4枚目のは女房のために使います. The substitution of a difficult verb (ものがたる) with a basic verb (わかる) here results in a change in subject, from inanimate to animate: その文化の実態が趣味の「お稽古ごと」であったことを、この一覧表がものがたっている。 → このリストを見ると、江戸の庶民文化が、実は、「お稽古ごと」であったことがわかる。

Sometimes, vocabulary simplification is carried out by means of a paraphrase, where part of the original meaning is lost: 改めて見直し、考えさせられるチャンスに恵まれます。 → もう一度考えさせられます。; 私の中に、価値観の逆転が起きた → 私の考え方が変わった。

Vocabulary simplification at times also involves a change in cohesive devices, such as deictics instead of synonyms or paraphrases: 江戸時代に趣味として蓄積してきた高度な教養 → このような高度な教養; 店のおやじさん → 彼.

### 5.1.3 Morphological simplification

Morphological simplification could be seen in the use of more basic grammatical forms instead of markedly formal or markedly colloquial forms, such as:

- substituting the formal truncated connective form of affirmative predicates with the te-form: 会社を出 → 会社を出て; 冬暖かく、夏は気持ちいい → 冬はあたたかくて、夏は気持ちいい; and substituting the formal negative connective form ずに with the general ないで: 言わずに → 言わないで; 気づかずに → 気がつかないで; せず、 → しなかった。; and generally using basic instead of very formal forms: 同じく → 同じように; みせねばならない → みせなければなりません; なることでしょう。 → なると思います。

#### 5.1.4 Syntactic simplification

Syntactic simplification was also frequent, by means of:

- dividing sentences with coordinate clauses into separate, shorter sentences:

きっと私も、この子がいなかったら、気づかずに、そのまま生きて行ったんだろうな、と思い、最近、行康に「ありがとう」としみじみ言ったものでした。 → きっと私も、この子がいなかったら、気がつかないで、そのまま生きて行っただろう、と思います。行康には心から「ありがとう」と言いたい。

彼女はあわてて手の甲をぬぐって私に右手の甲をさし出し、お互いの右手の甲を合わせて握手の代わりとしたことがあった。 → 彼女は急いで手の甲をふいて右手の甲を差し出した。お互いの右手の甲を合わせて握手の代わりにしたことがあった。

握手は古来誓いや約定を固めるためになされた法行為であって、手の平の柔らかい部分を互いに合わせて一体となる意識を目覚めさせ、契約を公示する行為とされてきた。 → 握手は昔から契約の成立を確認するための行為であった。つまり、法律的な行為であった。手の平の柔らかい部分を合わせるのは、相手と一つだという意識を持つためだった。

封建制のもとで領主に臣従する家臣は両手を合わせ、領主がその手を自分の手をつつんだとき、臣従契約が成立した。 → 封建時代では、家臣になりたい人間が両手を合わせる。そして、主人がその手を自分の手で包むと、主従の契約が成立した。

- separating subordinate clauses and turning them into separate sentences, especially in the case of adnominal modifiers:

日本の伝統的な文化とされる茶道にせよ華道にせよ、いずれも、もともと特殊な階層の文化であったものが、こうした「お稽古ごと」を介して大衆化し、国民的な趣味に拡張したものである。 → 日本の伝統的な文化である茶道も、もともとは、特殊な人々の文化であった。が、こうした「お稽古ごと」を通して広がり、国民的な趣味になったものである。

The separation of complex sentences into shorter, simpler ones, at times resulted in (or was motivated by a desire of) bringing subject and predicate of the original complex sentence nearer to each other:

この実験は、人が集団の圧力を感じ、集団に従おう、あるいは従わざるをえないというような気持ちになるのは、三人以上であることを明らかにしています。 → この実験から、次のようなことがわかります。私たちが、集団の圧力を感じて他の人と同じ意見を言うようになるのは、3人以上の人が同じことを言った時であること。

### 5.1.5 Discourse simplification

Discourse simplification was obtained in three cases by maintaining topic continuity and avoiding topic-shift from one agent to another:

必ずオレの分も注文して飲んでくれという遺言を実践しているんだ → 必ず、彼の分も注文して飲むということになったんだ。それで、その約束を実行しているってわけさ;

四枚目ので女房を養いますが、女房はわたしに何の利益ももたらしませんから、これは捨てることとなります。 → 4枚目のは女房のために使いますが、これは利益にはなりませんから、捨てることとなります;

さて、大臣たちが集まったとき、王さまは → お城に帰ると、大臣たちを集めて、王さまは、こう言いました。

### 5.1.6 Deletion

Deletion was the most drastic form of simplification, used quite often: 95 instances of deletion were found, including deletion of:

- modal forms, e.g.:

どうも土産が足りない → おみやげが足りない;

- intensifiers, e.g.:

それよりかなり以前 → それ以前; 何の利益ももたらしません → 利益にはなりません; ただの一度でも → 一度でも

- aspectual forms, e.g.:

捨ててしまいます → 捨てます; 引き返して行って → 引き返して;

- extra-textual references, e.g.:

「ママ、手が凍るんだよ」(一二、一四ページ)を読んで → 「ママ、手が凍るんだよ」を読んで

- rhetorical devices, e.g.:  
誤解のないようにいっておくが → しかし
- semantically redundant, non-essential information or details, e.g.:  
林野庁業務部販売推進室長羽賀正雄さん → 林野庁業務部の羽賀正雄さん; 見知らぬ十二歳の坊やの姿に、八十歳を超した両親の光景を重ね合わせた → 80歳を越したいなかの両親のことを思い出した; 越後からやってくるたび、東京の街なかで → 東京に出てくると; 江戸時代の都市の文化 → 江戸時代の文化; 法則の基礎知識をもっていた → 法則を知っていた;
- redundant paraphrases, or information that can be inferred from the context, e.g.:  
「みんなではなく、たった三人じゃないか」 → 「3人だけ?」; その趣味の代表的な事例 → その例; ...に出版された...案内書 → ...の...案内書; 都市にすむ学者 → 都市の学者.

In some instances, whole paragraphs were omitted, such as the underlined part in the following example, which includes culturally-bound terms, where not only the words, but also the words' referents are probably not known to learners, and at the same time, being only exemplifications of the previous general statement, are not essential to convey the general meaning of the passage:

古典には、手を握り合って飛び上がったとか、手を取り合って喜んだといった文章がたくさん出てきます。『古事記』にもあれば、さらに軍記文学にも、家臣たちが成功して帰ってくると武将が手を取って喜んでくれる、という場面がよく見られます。手を取るとは、つまり握手です。

→ 昔の文学作品には、手を握り合って飛び上がったとか、手を取り合って喜んだといった表現がたくさん出てきます。手を取るとは、つまり握手です。

## 5.2 Explicitation

Strategies of explicitation were also very frequent: 203 instances of explicitation were found, encompassing all linguistic levels.

### 5.2.1 Semantic explicitation

Semantic explicitation was the most common, occurring in 110 cases, such as:

- adding concrete time or place settings, sometimes even with an extensive description:

ある男が、バーのカウンターにすわった。 → ある男が、バーのカウンターにすわった。はじめての男だった。; 思わず、言った。 → 大田区役所へ行った時、思わず、こう言った。; では二杯、ご用意しましょう → では、2杯いっしょに、ご用意いたしましょう; やっぱり後ろから → この時も、やっぱり後ろから; ニューヨークで小さな衣料品会社をやっているスタインバーグ氏は、もう二十年間も毎日同じレストランで昼食を食べている。 → 佐藤さんは、東京の銀座で、小さな出版会社を営んでいるが、なんでもきちょうめんである。仕事もきちんと片づけるし、食事と同じである。佐藤さんは、昼に、もう 20 年も同じ店で同じものを食べている。

- using a more specific word instead of a more general hypernym, e.g.:

とうとう、 → 次の日、; 連絡をとった → 電話をした; 店のおやじさん → 店の主人; こいつ → この客; やっている → 経営している

- using a hypernym and adding a definition:

... ためになされた法行為であって → ... ための行為であった。つまり、法律的な行為であった; or adding a hypernym to a word that readers might not know, instead of a definition: 水割りを → ウィスキー、水割りを;

- in one case even adding loan-word synonyms as furigana (here written in parentheses):

利己的遺伝子の乗り物 → 利己的遺伝子 [セルフッシュジーン] の乗り物 [ヴィークル]

Almost half of the semantic explicitations occurred at the **script** level (42 cases): words that were written in hiragana in the original text, but can be and usually are written with Chinese characters, were rewritten using these characters, which made them less ambiguous, both visually, in terms of word delimitation, and semantically, distinguishing between homophones:

いっている → 言っている; いままで → 今まで; とき → 時; ばあい → 場合; ひとり → 一人; もっている → 持っている; らくに暮らせる → 楽に暮らせる; わたし → 私; よんで → 呼んで, etc.

### 5.2.2 Boundary explicitation

Boundary explicitation was the next most common type of explicitation (70 cases), mostly by means of added punctuation:

- adding commas to separate ambiguous or just long strings of hiragana:

もう二十年間うちのスープを → もう 20 年も、うちのスープを; 男はそれからもときどき → 男は、それからも、ときどき; ところがある日 → ところが、ある日; あれお客さん → あれ、お客さん;

- adding commas to separate ambiguous strings of kanji:

実際手は → 実際、手は;

- adding commas to separate phrases, often topical phrases:

必ずオレの分も → 必ず、彼の分も; いったいどうしたというんですか? → いったい、どうなさったんですか。; わたしは飲まなかったって → 私は、飲まなくても; 大人でも似たようなもの → 大人でも、同じようなもの;

- adding commas between clauses:

時間をやりくりして大田区役所へ... → 時間を見つけて、その地下道へ...

- adding parentheses for emphasis or reported speech:

三人というのはみんななのです → 3人というのは「みんな」なのです; みんながゴルフをはじめたので.....というとき → 「みんながゴルフをはじめたので...」という時

- dividing one paragraph into two: ...]

と百姓は答えました。王さまは → ...」と農夫は答えました。<p> 王さまは; という話になる。ところが → という話になる。<p> ところが; 「さっきあなたは一、〇〇〇円とおっしゃったけれど、二、〇〇〇円でいかがでしょう」となる。振り向き賃が七、〇〇〇円についちまうわけです。→ 「さっきあなたは 1000 円とおっしゃったけれど、2000 円ではいかがでしょう」となる。<p> 振り向き賃が 7000 円になってしまうわけです。

- inserting back-channelling expressions in dialogues:

「亡くなった飲み友だちと約束してね、僕が飲みに行くときは、必ずオレの分も注文して飲んでくれという遺言を実践しているんだ」 → 「先週、僕の親友が亡くなったんだが、彼が亡くなる前に約束してね」<p> 「はあ」<p> 「僕が飲みに行く時は、必ず、彼の分も注文して飲むということになったんだ。それで、その約束を実行しているってわけさ」; 「そうですね。わかりました。飲みますよ。あれ、スプーンはどこですか?」 → 「そうですね、わかりました。飲みますよ。飲めばいいんですね」<p> 「ああ」<p> 「じゃあ、飲みます。あれ、スプーンはどこですか」

Boundary explicitation was also obtained by splitting long complex sentences into shorter ones, which also implies syntactic simplification, as mentioned in the previous sub-section, and by using Chinese characters instead of hiragana where possible, as



mentioned in the previous subsection, to mark the delimitation between words, which is not marked by blank spaces in Japanese standard script.

### 5.2.3 Syntactic explicitation

Syntactic explicitation was obtained by:

- adding an omitted argument to a predicate:  
交配をかさねて → 朝顔を交配させて; 手袋は、滑ってしまうから、つけてはいかれない。→ 手袋をすると滑ってしまうから、手袋をすることはできない。
- substituting an intransitive expression, where all agents are not immediately obvious, with a transitive one, where agent and patient are more clear (here the semantic value of the verb is also more specific):  
一杯目が終わったら → 一杯目をお飲みになったら;
- using polite or humble forms which disambiguate the subject of the predicate:  
一杯目が終わったら → 1杯目をお飲みになったら; ご用意しましょう → ご用意いたしましょう; いったいどうしたというんですか? → いったい、どうなさったんですか。
- adding particle の to split a compound noun into a noun with a nominal modifier:  
臣従契約 → 主従の契約; 現代日本 → 現代の日本; 遺伝法則 → 遺伝の法則。

Other, less common cases of **explicitation**, were cohesion and phonetic explicitation.

### 5.2.4 Cohesion

Cohesion explicitation, by adding cohesive elements, eg.:

- deictics (その, この, 上の etc., such as in その結果が図です → その結果が、上の図です) or
- temporal expressions (その時, そして etc.)

### 5.2.5 Phonetic explicitation

Phonetic explicitation by using hiragana instead of Chinese characters, where the pronunciation of the characters could be ambiguous: その後の事 → そのあとの

事、 宝物 → 宝もの; or by adding furigana on difficult Chinese characters: 疑惑 → 疑惑 [ぎわく]; 氷解 → 氷解 [ひょうかい]; 爽やか → 爽やか [さわやか] .

### 5.3 Standardisation

The third strategy used in the adaptation of texts for foreign language learners was the use of more standard or basic linguistic forms, i.e. forms which are usually learned at the beginning of Japanese language courses, instead of stylistically marked forms, which are usually learned later. 128 modifications were counted in this category, including the following means.

#### 5.3.1 Script standardisation

Script standardisation, using:

- standard punctuation instead of non-standard brackets, question marks and other punctuation:

技術や知識・芸能を → 技術や知識や芸能を; どうかいたしましたか? → どうかいたしましたか。; ですか? → ですか。; 《皆様の励ましは、...宝物とさせていただきます》 → 『みなさんの励ましは、...宝もの とさせていただきます』; 冬の朝。金属の手すりは、氷のように凍っています。 → 冬の朝、金属の手すりは、氷のように冷たくなっています。; ので..... → ので...; or using standard characters: 八十歳を超した → 80歳を越した.

#### 5.3.2 Tense levelling

Changing single predicates in non-past form to past form in texts which are otherwise written in the past form, to make the tense uniform throughout the text:

確立されるのは → 発見されたのは;

#### 5.3.3 Formality levelling

Formality levelling from de-aru to plain style:

であった → だった; or from plain to formal style: あった → ありました; だから → ですから; in texts which are otherwise written in this style, to make the style uniform throughout the text

### 5.3.4 Formality standardisation

Substituting colloquial or otherwise register-marked forms with standard (more polite) forms, which are generally learned earlier in Japanese language courses:

飲まなくたって → 飲まなくても; 話さん → 話さない; 飲んでみたまえ → 飲んでみてくれないか; 言え → 言わなければならない; 「... どうしてです」 → 「... どうしてですか」; ようがす。 → わかりました。

Exceptionally, in one case the opposite occurred: a colloquial, shortened form was used instead of a politer one: 振り向いてはいけない → 振り向いちゃいけない, probably for stylistic effect.

### 5.3.5 Visualisation:

Throughout the texts, numbers written in Chinese characters in the original texts, printed vertically, were replaced with arabic numerals in the textbook reading passages which are printed horizontally:

百枚 → 100 枚; 百度 → 100 回; 四枚 → 4 枚目; 四十五分 → 45 分; 十二時 → 12 時; 十七世紀 → 17 世紀 六、〇〇〇円と → 6000 円と; 五年生 → 5 年生; 五十四段 → 54 段; 二枚 → 2 枚; 二杯 → 2 杯; 二月二十日付の → 2 月 20 日付けの; 二十年間 → 20 年間; 二、五〇〇軒 → 2500 軒; 三人 → 3 人; 三つ → 3 つ; 一・三キロ → 1.3 キロ etc.

One modification was found that does not clearly belong to any of the four categories proposed, but could tentatively be categorised as standardisation, or could also be termed familiarisation or domestication. It was only used in one text: the setting of a story, originally happening in New York to a Mr. Steinberg, apparel vendor, was reset in Ginza, one of the most famous Tokyo districts, with Mr. Sato, publisher, as the main character:

ニューヨーク → 東京の銀座; スタインバーグ → 佐藤; 小さな衣料品会社をやっている → 小さな出版会社を営んでいる。

## 5.4 Other modifications

Some modifications were found for which no clear motive could be guessed: they may have been made for stylistic purposes, according to the rewriter's tastes, or may be the results of multiple modifications, where the original motive became blurred in

subsequent modifications. One substitution was probably just a spelling mistake, resulting in a colloquial *きてます* instead of *きます* in the sentence:

おやじは必ず後を追っかけてきます。 → おやじは必ず後を追っかけてきてます。

Other cases where the motive for the substitution were not clear were:

- one separation of a clause indicating reported speech into a separate paragraph, probably for dramatic effect:

「それ、みたまえ!」とスタインバーグ氏は言ったのである。 → 「わかったかね」 <p> そう佐藤さんは言ったのである。

- the substitution of a more standard full stop with a less standard comma after a polite predicate:

そうですか。わかりました。 → そうですか、わかりました。

- one substitution of a causal connective with a more polysemic and not easier connective:

身体を接触させず、ただ目で意志を相手に伝えることの方が多いからである。 → 相手の体に触れない。目で意志を相手に伝える方が多いのである。

In six cases, commas separating phrases in short sentences were deleted, which is slightly surprising, given that in other 39 cases, commas were added in such positions:

階段の途中で、息子は立ち止まってしまいます。 → 階段の途中で息子は何回も止まってしまいます。; だが、それは行康君にはひどくつらいのだった。 → だがそれは行康君にはひどくつらいのだった。; 行康君は、記者に、そう言った。 → 行康君は記者にそう言った。; と、区の担当者はいっている。 → と区の担当者は言っている。; 社会部に、行康君の母親の栗田敦子さん(四五)から次のような手紙が届いた。 → 社会部に、行康君の母親の栗田敦子さん(45)から次のような手紙が届いた。; たいがい、三、四人で終わりです。 → だいたい3、4人で終わりです。

## 6 Discussion

As could be seen in the previous section, multiple strategies were used when rewriting texts which were originally written for native speakers of Japanese, to be included in a reading textbook for intermediate learners of Japanese as a foreign language. A summary of all modifications, counting their number at different levels of linguistic analysis and by type of strategy, is given in the following tables.

**Table 4:** Number of modifications by level of linguistic analysis

Number of modifications by level of linguistic analysis	No. of occurrences
script	175
punctuation	56
vocabulary:	
content words	193
function words	44
vocabulary + syntax	167
morphology	20
modality	18
syntax	42
semantics	45
cohesion	10
discourse	32
formality	12
intertextuality	1
<b>Total</b>	<b>815</b>

**Table 5:** Number of modifications by strategy type

Number of modifications by strategy type	No. of occurrences	Percentage
simplification (of which 96 deletions (12% of total))	472	58%
explicitation	203	25%
standardisation (of which 80 visualisations)	128	16%
not categorised	12	1%
<b>Total</b>	<b>815</b>	<b>100%</b>

While the quantities of modifications on different linguistic levels cannot be objectively compared, since they refer to linguistic elements which occur in different scales of magnitude (the number of elements of vocabulary in a text is always larger than the number of phrases, clauses, sentences and paragraphs, thus making a comparison

impossible), it is still interesting to see how modifications were made on all levels, multiple times.

The modifications at different linguistic levels confirm the central role of vocabulary as declared in the foreword to the textbook and as could be inferred from the structure of the textbook containing many vocabulary exercises. It is not, however, the only level at which modifications were made: a considerable number of modifications was made at the script level, and all other linguistic levels were also touched by the modification process. Vocabulary modifications in many cases (167) brought with them also syntactic changes, and other aspects of the text were modified irrespective of the vocabulary used: many were structural modifications, touching syntax and discourse, to simplify syntactic and discourse structures or make them more explicit, cohesive devices were introduced, and stylistic changes (standardisations) were made at the level of formality.

As for type of strategy used, simplification was the most common strategy, accounting for more than half of the occurrences. It occurred at the level of vocabulary, where less frequent words were substituted with more common synonyms, explanation, definitions or paraphrases, at the level of morphology, where less common predicate forms were substituted with more basic ones, at the level of syntax, where long sentences with coordinate and subordinate clauses were split into smaller units, at the level of discourse, where roles were switched to maintain topic continuity, paragraphs of narrative were divided and shorter conversation turns introduced in dialogues.

However, alongside simplification, explicitation was another strategy that should not be overlooked, as it accounts for one quarter of the number of modifications, indicating that the authors of the rewritings considered it a useful device and found it useful when rewriting texts for their students.

It was used on the semantic level, adding information that could otherwise be inferred from the text, or cultural background that is not likely to be known to learners, or just inventing concrete settings to help readers reconstruct the narrative being told. Semantic disambiguation also occurred at the script level, where strings of hiragana were often rewritten in Chinese characters to disambiguate homophones.

Structural explicitation was also observed at different levels: boundaries between linguistic units were made clearer by the use of punctuation, script or layout; omitted predicate arguments were made explicit, polite forms were used to disambiguate the subject of the sentence, and some phonetic information was added by means of rewriting Chinese characters in hiragana or by adding furigana.

The third strategy used, standardisation, was used at the script level, to standardise punctuation and to make the text visually more familiar (using Arabic instead of Chinese numbers), and at the discourse level, both to uniform the use of the same tense or level of formality within one text (choosing one of two possibilities, such as past/non-past, or formal/informal, which are both known to learners), or to standardise the text as a whole by removing marked forms which are typical of less

standard registers (very colloquial, literary etc.) and not likely to be known by intermediate learners.

## 7 Conclusion and further work

Overall, it could be seen that the rewriters used some strategies which could be applied to the simplification of texts for most weak readers of Japanese, not only foreign learners of Japanese: short sentences and frequent vocabulary are two aspects of language that have been found to be easier to read in most research on readability in different languages.

However, it is also clear that the authors (rewriters) of these texts, teachers of Japanese as a foreign language, were conscious of the typical progression of formal Japanese language instruction, and tended to prefer vocabulary, morphology and syntactic structures that are learned earlier in language courses. These are very often also the most frequent forms in the Japanese language as a whole and learned earlier by Japanese children (especially in the case of content words), but some linguistic elements, such as standard polite language (as opposed to very colloquial or very formal speech) are typical of beginning language courses for foreigners, while colloquial language (including vocabulary and contracted or otherwise colloquial morphology), which is learned quite early by Japanese children, is learned later in formal language instruction and therefore relatively difficult for foreign learners of Japanese.

All the strategies which were highlighted in this analysis could be useful as guidelines when assessing the readability of texts for foreign learners of Japanese, both in an overall assessment of readability, and when devising methods and systems to pinpoint difficult aspects of particular texts as a first step to text simplification. Especially in the first case, when assessing overall readability, i.e. grading multiple text on one scale of readability, further and more extensive analysis of the weight of each of these aspects on overall readability is needed. In both cases, it would be useful to devise a system for automatic discovery and assessment of particular aspects of readability.

## References

- DuBay, W. H. (2004). *The principles of readability*. Costa Mesa, California: Impact Information.
- Halliday, M. A. K. (1993). Some grammatical problems in scientific English. In *Writing Science: Literacy and Discursive Power*, Pittsburgh, 1993 (pp. 69-85). University of Pittsburgh Press.

- Hayashi, Y. (1992). A Three-level Revision Model for Improving Japanese Bad-styled Expressions. In *COLING 1992 Volume 2: The 15th International Conference on Computational Linguistics* (pp. 665-671).
- Honda, K. [本田勝一] (1982). *Nihongo no sakubun gijutu* 日本語の作文技術 [*Japanese writing techniques*]. Tokyo: Asahi shimbun 朝日新聞.
- Ichikawa, A. [市川熹] (2006). Fukushi jouhou gaku towa nanika 福祉情報学とは何か. *Gekkan gengo* 月刊言語, 35(7), 26-34.
- Inui, K. [乾健太郎] & Fujita, A. [藤田篤] (2004). Iikae gijutsu ni kansuru kenkyuu doukou 言い換え技術に関する研究動向 [Paraphrase research trends]. *Shizen gengo shori* 自然言語処理, 11(5), 151-198.
- Inui, H. [乾裕子] & Okada, N. [岡田直之] (2000). Nagai bun wa tsune ni wakarunikui ka? Wakarunikusa no youin to sono izon kankei - Is a long sentence always incomprehensible? A structural analysis of readability factors 長い文は常にわかりにくいのか?: わかりにくさの要因とその依存関係. *NL SIG Technical reports* 情報処理学会研究報告自然言語処理, 2000(11), 63-70. [<http://id.nii.ac.jp/1001/00048682/>]
- Inui, K. & Satomi, Y. (2001). Corpus-based acquisition of sentence readability ranking models for deaf people. In *Natural Language Processing Pacific Rim Symposium Tokyo* (pp. 205-212).
- Japan Foundation [国際交流基金] & Association of International Education Japan [日本国際教育協会] (2004). *Nihongo nouryoku shiken shutsudai kijun - Japanese language proficiency test : test content specifications* 日本語能力試験出題基準. Tokyo: Bonjinsha 凡人社.
- Kabashima, S. [榊島忠雄] (1979). *Bunshou sahou jiten* 文章作法事典. Tokyo: 東京堂 Tokyodo.
- Kinoshita, K. [木下是雄] (1981). *Rikakei no sakubun gijutsu* 理科系の作文技術. Tokyo: Chuokoron shinsha 中央公論新社.
- Koide, K. [小出慶一] (1991). *Nihongo o manabu hitotachi no tame no Nihongo o tanoshiku yomu hon - Enjoyable Task Reading in Japanese* 日本語を学ぶ人たちのための日本語を楽しく読む本. Tokyo: Sanno University International Student Center 産能短期大学国際交流センター.
- Lee, J.-H. & Hasebe, Y. (2016). Readability measurement for Japanese text based on leveled corpora. In *Papers on Japanese Language from an Empirical Perspective*, Ljubljana: Academic Publishing Division of the Faculty of Arts, Univ. of Ljubljana.
- Matsumoto, S. (2010). *JDiff X Document Comparison Plug-in for Jedit X*. [[http://www.artman21.com/en/jdiff\\_x/](http://www.artman21.com/en/jdiff_x/)].
- Mishima, H. [三島浩] (1990). *Gijutsusha, gakusei no tame no technical writing* 技術者・学生のためのテクニカル・ライティング. Tokyo: Kyouristu shuppan 共立出版.
- Morioka, K. [森岡健二] (1952). Yomiyasusa no kiso kenkyuu 「読みやすさ」の基礎的研究. In *Kokuritsu kokugo kenkyuujo nenpou - Annual report of National Language Research Institute* 国立国語研究所年報 (pp. 91-108).



- Nakano, T. [中野智子], Endo, A. [遠藤淳], Sugawara Sh. [菅原昌平], Inui, K. [乾健太郎], & Fujita, A. [藤田篤] (2005). Lexical Paraphrasing for Improving Accessibility to the Web - Web サイトへのアクセシビリティ向上を目的とした難語の平易化. *IEICE technical report Welfare Information technology* 電子情報通信学会技術研究報告 WIT 福祉情報工学 25(25), 11-14.
- Ono, T. [小野貴博], Suganuma, A. [菅沼明], & Taniguchi, R. [谷口倫一郎] (2006). Nihongo bunshou suikou shien ni okeru kakariuke o gokai sareru bun no chuushutsu - Extraction of the sentences whose modification relation is misunderstood for a writing tool 日本語文章推敲支援における係り受けを誤解される文の抽出. *IPSJ SIG technical reports* 情報処理学会研究報告 FI 情報学基礎, 2006(94), 99-104.
- Oono, H. [大野博之] & Inazumi, H. [稲積宏誠] (2007). - Development of an education support tool for improvement of ability for sentence making 技術文章作成能力の育成を目指した教育支援ツールの開発. In 電子情報通信学会 第 18 回データ工学ワークショップ論文集 - DEWS2007.
- Sakamoto, I. [阪本一郎] (1962). Bunshou no goi hijuu no sateihou - Readability no kenkyuu no kokoromi 文章の語彙比重の査定法---Readability の研究の試み---. *Dokusho kagaku* 読書科学, 6(1), 37-44.
- Sano, M. & Maruyama, T. (2008). Lexical Density in Japanese Texts: classifying text samples in the Balanced Corpus of Contemporary Written Japanese (BCCWJ). In *Proceedings of ISFC 35: Voices Around the World*, Sydney, 2008 (pp. 359-364).
- Sato, S., Utsuro, T., Tsuchiya, M., Asaoka, M., & Matsuhoshi, S. (2004). Natural Language Processing Technologies to Enhance Readability. In *Proc. of International Conference on Informatics Research for Development of Knowledge Society Infrastructure* (pp. 46-53).
- Sato, S., Matsuyoshi, S., & Kondoh, Y. (2008). Automatic Assessment of Japanese Text Readability Based on a Textbook Corpus. In *Proceedings of the Sixth International Language Resources and Evaluation (LREC'08)*, Marrakech, Morocco, 2008.
- Shibasaki, H. [柴崎秀子] & Tamaoka, K. [玉岡賀津雄] (2010). 国語教科書を基にした小・中学校の文章難易学年判定式の構築 - Constructing a formula to predict school grades 1-9 based on Japanese language school textbooks. *Nihongo kyouiku kougakukai rombunshi - Japan journal of educational technology* 日本教育工学会論文誌, 33(4), 449-458.
- Tateisi, Y., Ono, Y., & Yamada, H. (1988). A computer readability formula of Japanese texts for machine scoring. In *Proceedings of the 12th conference on Computational linguistics* (pp. 649--654). Association for Computational Linguistics.
- Yamamoto, S. [山本聡美], Inui, K. [乾健太郎], Nogami, M. [野上優], Fujita, A. [藤田篤], & Inui, H. [乾裕子] (2000). 聾者向け文章読解支援のための文可読性基準の調査 - Exploring the Readability Criteria for Congenitally Deaf People: A Step toward Computer-Aided Text Reading. *IPSJ SIG technical reports* 情報処理学会研究報告 NL 自然言語処理, 135(17), 127-134.
- Yasumoto, B. [安本美典] (1983). *Settoku no bunshô gijutsu [Techniques of persuasive writing]* 説得の文章技術. Tokyo: Kôdansha 講談社.