IEICE TRANS. INF. & SYST., VOL.E91–D, NO.11 NOVEMBER 2008

PAPER *Special Section on Knowledge, Information and Creativity Support System*

Analysis of Eye Movements and Linguistic Boundaries in a Text for the Investigation of Japanese Reading Processes

Akemi TERA[†a)], ***Member***, **Kiyoaki SHIRAI**[††], ***Nonmember***, **Takaya YUIZONO**[†], ***and* Kozo SUGIYAMA**[†], ***Members***

SUMMARY In order to investigate reading processes of Japanese language learners, we have conducted an experiment to record eye movements during Japanese text reading using an eye-tracking system. We showed that Japanese native speakers use “forward and backward jumping eye movements” frequently [13], [14]. In this paper, we analyzed further the same eye tracking data. Our goal is to examine whether Japanese learners fix their eye movements at boundaries of linguistic units such as words, phrases or clauses when they start or end “backward jumping”. We consider conventional linguistic boundaries as well as boundaries empirically defined based on the entropy of the *N*-gram model. Another goal is to examine the relation between the entropy of the *N*-gram model and the depth of syntactic structures of sentences. Our analysis shows that (1) Japanese learners often fix their eyes at linguistic boundaries, (2) the average of the entropy is the greatest at the fifth depth of syntactic structures.

key words: *eye tracking, saccade, fixation, N-gram model, entropy*

1. Introduction

In the field of Japanese language education for non-native speakers, many Japanese language learners (Japanese learners) feel that a Japanese text is difficult to read. The reasons are generally summarized as follows: there is no space between words, a Japanese text consists of four different kinds of characters (*Hiragana*, *Katakana*, *Kanji and Roma-ji*), and it is difficult to look up words represented by Kanji in a dictionary since Kanji words have several possible readings. In order to support the learning of Japanese, an investigation of Japanese learners’ reading processes plays an important role.

Several studies aimed at analyzing reading processes of Japanese learners have been conducted. Tera et al. reported that “Na-adjectives”, “sahen” and complex verbs are often consulted by Japanese learners via a reading support system developed in [5]–[7], [9]*. These studies indicated that Japanese learners especially want to know information about predicates in the Japanese text.

Eye movements of language learners are also analyzed. Here we briefly introduce “fixation” and “saccade” in eye movements. When we read a text, we use our eyes to get information from the text. It is already known that people do not move their eyes smoothly or constantly while reading, but repeat fixation and saccade. “Fixation” means that the eyeballs stop moving for the eye to glance at the same point for a while in order to carefully read a text, while “saccade” means that the eyeballs move from a fixation point to the next fixation point. We call a saccade in the forward direction a “forward saccade”, while a saccade in the backward direction is a “backward saccade”.

So far, numerous attempts have been made to analyze fixation or saccade, as summarized in Rayner [8]. From the beginning of 1990, research on reading processes Japanese texts using an eye-tracking system were conducted, using an eye-tracking system. Osaka [2] reported differences in the locations where subjects fix their eyes according to the types of characters in the text. Shigematsu et al. [4] discussed differences of fixation of eyes between Japanese and Chinese students. Saida [3] proposed a new method that restricted learners’ field of view using cutoff frames in order to measure the range of field of view in the reading process. We also tried to investigate reading processes using an eye-tracking system [13], [14]. We found that (1) when reading text in native or familiar languages, fixation time tended to be short, (2) Japanese native speakers often showed “backward jumping eye movements” and “forward jumping eye movements”. It is well known that “backward jumping eye movements” occurs when a reader finds out some discrepancies while reading with respect to the meanings he holds, and is often used by skilled language learners [1].

Although many research studies investigated fixation or saccade in reading process, there has been no attempt to reveal the relation between saccade and the linguistic boundaries in a text in the reading process. Especially, it is still uncertain if learners of Japanese tend to start or end their backward saccade near linguistic boundaries. In this paper, we analyzed relation between backward jumping and linguistic boundaries with eye tracking data obtained in [13], [14]. In addition to conventional linguistic boundaries such as word or clause boundaries, we also consider the clear boundaries empirically detected based on the entropy of the *N*-gram model. Thus far, in the research to reveal reading processes, the eye fixation time or the fixation

Manuscript received April 4, 2008.
Manuscript revised July 4, 2008.

[†]The authors are with School of Knowledge Science, Japan Advanced Institute of Science and Technology, Nomi-shi, 923–1292 Japan.
[††]The author is with School of Information Science, Japan Advanced Institute of Science and Technology, Nomi-shi, 923–1292 Japan.
a) E-mail: tera@jaist.ac.jp

DOI: 10.1093/ietisy/e91–d.11.2560

*“Na-adjective” and “saahen” are word classes in Japanese. The Na-adjective is one kind of adjective, while sahen is a word which functions both as a noun and as a verb.

position is analyzed, but not the saccade time [4]. Fixation is the work required for getting information from characters, while saccade is the work required for searching the positions at which the characters exist. Thus we pay attention not only to eye fixation but also to saccade time.

The paper is organized as follows: our goal is explained in more detail in Sect. 2. Section 3 describes the detailed procedures of our analysis. Results of our analysis are reported in Sect. 4, and we discuss these results in Sect. 5. A preliminary survey of the relation between syntactic structures and entropy is presented in Sect. 6. We conclude the paper in Sect. 7.

2. Goal

The goal of this paper is to investigate the reading process of Japanese learners. More specifically, we investigate where backward saccades occur in the text. As described in Sect. 1, a backward saccade (backward jumping eye movement) occurs when a learner cannot understand the text smoothly, and wants to read previous sentences again to make sure that what he/she understood before is correct [1]. If this is correct, we suppose that Japanese learners would start or end a backward saccade not within a word but at linguistic boundaries. So we propose the following hypothesis:
[Hypothesis]
When reading a Japanese text, a backward saccade starts or ends around linguistic boundaries.

In this paper, we will empirically prove the above hypothesis through an analysis of real eye-tracking data.

In our analysis, linguistic boundaries in a text are defined in two ways. One is by borrow the concepts of conventional linguistic boundaries such as words, *bunsetsus*† or clauses. The other is empirically defining linguistic boundaries by measuring the entropy of the *N*-gram model trained from a large number of real texts. One of the advantages of the second approach is that we can consider the degree of intensity of linguistic boundaries. As we will introduce in Sect. 3.2, the entropy of the *N*-gram model reflects how likely the position in a text might be linguistic boundaries. That is, if the entropy of a certain position is great, it might be a clear linguistic boundary. We would like to analyze if backward saccade is more likely to occur at clear linguistic boundaries, which can be statistically detected.

Another goal is to examine the relation between the backward saccade and syntactic structures of sentences. We would like to know if Japanese learners start or end backward saccades at deep positions in syntactic structures. As a preliminary survey for this investigation, we examine the correlation between the entropy of the *N*-gram model and the depth of syntactic structures.

3. Methodology

In this section, procedures to verify our hypothesis are described. The experiment for recording eye-tracking data is briefly described in Sect. 3.1, while the verification of our hypothesis by use of the obtained eye-tracking data is discussed in Sect. 3.2.

Fig. 1 Experimental set-up.

3.1 Collecting Eye Tracking Data

The following equipment is used to obtain eye-tracking data of Japanese learners:

(1) An eye-mark recorder EMR-8 (NAC)
(2) A 44-degree lens
(3) An analysis system (NAC)
(4) A windows NT, 17-inch monitor
(5) A head stand

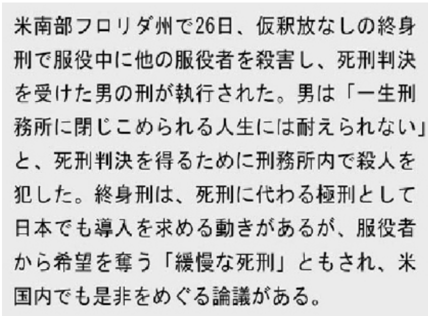
Since subjects are required to have eyesight over 1.0 to record eye movements with our eye-mark recorder, an eyesight check is performed before the experiment. An illustration of the recording of the eye tracking data with the above equipment is shown in Fig. 1.

The procedures for obtaining eye-tracking data consist of the following 7 steps:

(1) Choose intermediate learners as subjects according to the results of the Standard Japanese Ability Test and the English TOEFL/TOEIC.
(2) Explain an overview of the experiments and give instructions to the subjects, and then let them do preliminary experiments.
(3) Ask the subjects to read texts displayed randomly on a PC screen, one after another. Time for reading is unlimited.
(4) Do calibration for eye tracking.
(5) Record eye tracking data during reading.
(6) Ask the subjects which texts are most difficult and easiest to understand, and the reason why.
(7) Analyze the obtained eye tracking data as well as other data (Questionnaires (6), etc.).

The Japanese text used in the experiment is a newspaper article excerpted from the Chunichi Newspaper [12], shown in Fig. 2. It contains 178 characters. Figure 3 shows a schematic representation of the obtained eye tracking data overlapped with the text. In Fig. 3, circles indicate fixations,

†*Bunsetsu* is a linguistic unit widely accepted in Japanese: a chunk of words consisting of content words accompanied by some functional words.



米南部フロリダ州で26日、仮釈放なしの終身刑で服役中に他の服役者を殺害し、死刑判決を受けた男の刑が執行された。男は「一生刑務所に閉じこめられる人生には耐えられない」と、死刑判決を得るために刑務所内で殺人を犯した。終身刑は、死刑に代わる極刑として日本でも導入を求める動きがあるが、服役者から希望を奪う「緩慢な死刑」ともされ、米国内でも是非をめぐる論議がある。

Fig. 2 Japanese text used for the experiment.

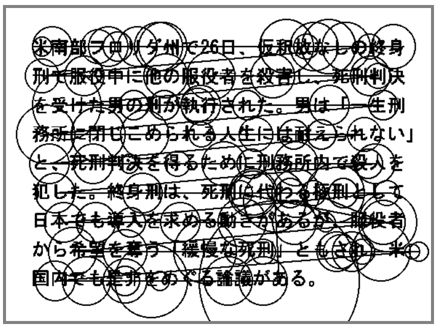


Fig. 3 Saccades and fixations on the text.

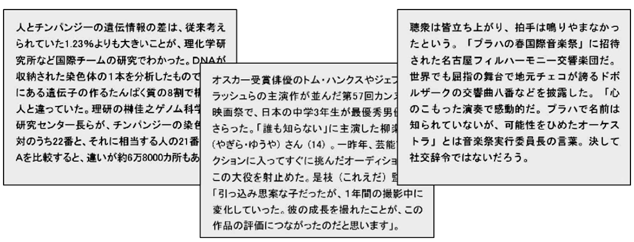


Fig. 4 Japanese texts used for the experiment.

while lines indicate saccades. The size of the circles represents the time of the fixations.

Furthermore, we use 3 more texts in this experiment (See Fig. 4). Using these texts, we analyze the difference in eye movements for the five different groups of learners that are mentioned below.

All of the subjects in this experiment are graduate students, 4 subjects are Japanese native speakers and others are Japanese language learners, from 22 to 28 years old. All of the foreign subjects except 3 get 60–80 scores for 100 full score Japanese ability test. For the investigation of reading processes, the categorization of subjects is a very important problem. Until recently, in the field of Japanese language education, subjects in experiments were usually categorized into 2 groups, “Kanji-area” and “Non-Kanji-area”. In 1999, Tera et al. divided subjects into 3 groups considering the familiarity of the subjects with Kanji. For example, Korean people use Kanji in their name or signboard, but they don’t use Kanji but Hangul in their writing. Hence, Tera et al. divided Kanji-area into 2 sub-groups, and then categorized all subject into 3 groups, which are “Kanji-area”, “Middle-area” and “Non-Kanji-area”.

Table 1 Nationality of subjects.

J	Japan (4)
K	China (5)
M	Korea (4)
NA	Nepal (1), Vietnam (1), Thailand (1)
NE	Belgium (1), Germany (1), Hungary (1), Spain (1)

In this experiment, we divide the Non-Kanji-area into 2 subgroups, “Non-kanji-Asia” and “Non-Kanji-Europe”, since their native languages, Asian languages and European languages, are rather different. That is, subjects are categorized into 5 groups.

We collected eye-tracking data for 20 subjects. They are classified into the following 5 groups according to their native language and familiarity with Kanji [9]:

- Japanese-Native-Speaker (J)
  Subjects who are Japanese native speakers.
- Kanji-area (K)
  Subjects who use Kanji in their native language, except Japanese.
- Middle-area (M)
  Subjects who know Kanji but do not usually use them.
- Non-Kanji-area-Asia (NA)
  Subjects who do not use Kanji and are from Asia.
- Non-Kanji-area-Europe (NE)
  Subjects who do not use Kanji and are from Europe.

The Numbers of subjects in each group as well as their nationalities are summarized in Table 1.

3.2 Analysis of Eye Tracking Data

The major goal in this paper is to examine if learners often start or end saccades near linguistic boundaries. First, we define three kinds of linguistic boundaries in a text: (1) word boundaries, (2) *bunsetsu* boundaries and (3) clause boundaries. A ‘clause boundary’ is a boundary of subordinate clauses or sentences. In our experiment, these boundaries in the text are manually annotated.

In addition to these linguistic boundaries, we try to define boundaries by the entropy of the N-gram model. The N-gram model is a well-known probabilistic language model widely used in the research field of natural language processing [10]. The N-gram model of characters is the probabilistic model predicting the appearance of a character (c_i) given its preceding $N-1$ characters ($c_{i-N+1}..c_{i-1}$). It is defined as follows:

$$P(c_i \mid c_{i-N+1} \cdots c_{i-1}) \tag{1}$$

The N-gram model can be automatically trained from a large amount of corpora [10]. Next, the entropy E of the N-gram model is given as below:

$$E = -\sum_{c_i} P(c_i \mid c_{i-N+1} \cdots c_{i-1}) \log P(c_i \mid c_{i-N+1} \cdots c_{i-1}) \tag{2}$$

In general, the entropy of the N-gram model is relevant to linguistic boundaries such as boundaries of words, phrases or clauses. For example, let us consider the cases where characters $c_{i-N+1} \ldots c_i$ are in a word. In such cases, only a limited number of possible characters would appear after the string $c_{i-N+1} \ldots c_{i-1}$. Thus the entropy would be low since the probability distribution $P(c_i \,|\, c_{i-N+1} \cdots c_{i-1})$ is not uniform, or is skewed. On the other hand, if there is a linguistic boundary between c_{i-1} and c_i, various characters could appear after the string $c_{i-N+1} \ldots c_{i-1}$. In such cases, the entropy would be high since $P(c_i \,|\, c_{i-N+1} \cdots c_{i-1})$ tends to be uniform. To summarize, if the entropy of the N-gram model is high, we can assume that there is a linguistic boundary at a position between characters c_{i-1} and c_i. More concretely, if the entropy at a position is greater than a certain threshold T, we regard that position as a 'clear' linguistic boundary.

When training the N-gram model of characters, we set N equal to 5, that is, we estimate a 5-gram model which predicts the probabilities of appearance of characters given 4 preceding characters. The model is trained by maximum likelihood estimation from newspaper articles published over 13 years. Then, the entropy of the 5-gram model is calculated at all positions in the text that subjects read in our experiment. The 'Position in the text' refers to a position between characters, and the entropy at a position between c_{i-1} and c_i is the entropy of the conditional probability distribution $P(c_i \,|\, c_{i-4}c_{i-3}c_{i-2}c_{i-1})$, where c_{i-4}, c_{i-3}, c_{i-2}, and c_{i-1} are the 4 characters appearing before that position. Unfortunately, the entropy cannot be calculated for all positions due to the data sparseness, since the entropy is not determined when the string $c_{i-4}c_{i-3}c_{i-2}c_{i-1}$ does not occur in the training corpus. Hereafter we call such positions "uncertain entropy positions".

Next, we extract fixations at the start and at the end of backward saccades from the eye tracking data. The positions of extracted fixations in the text area were identified, since our current goal is to examine if fixations happen at linguistic boundaries or at high entropy positions. However, it was rather difficult to decide the exact positions of fixations. Osaka [2] reported that readers saw between 2 and 5, mostly 3 and 4 characters when they fixed their eyes. This means that Japanese learners may see, not a point, but an area including several characters at fixations. Therefore, we suppose that subjects see 3 or 4 characters when they fix their eyes, and identify characters when they glance at fixations as shown in Fig. 5.

In this figure, a star shows the center of eye glance at fixation, while 'C' stands for characters supposed to be seen by subjects. That is, when the center of eye glance at fixation is on a character, we suppose that subjects see 3 characters, the character on eye glance and its previous and following characters (Fig. 5 (a)). Alternately, when the center of eye glance is between characters, we suppose that subjects see 4 characters around the center of eye glance (Fig. 5 (b)). Hereafter we call the area including characters glanced at by subjects "fixation area".

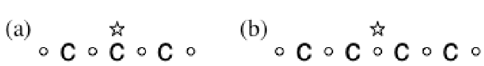


Fig. 5 Saccades and fixations on the text.

After identifying the fixation area for each fixation, we manually check if linguistic boundaries exist in the fixation area. More concretely, we look for linguistic boundaries among positions indicated by circles in Fig. 5. Then the linguistic boundary ratio *LBR*, defined by the formula below, is calculated.

$$LBR = \frac{\text{No. of fixations such that at least one linguistic boundary exists in the fixation area}}{\text{Total No. of fixations}}$$

LBR evaluates how likely fixations are to happen around linguistic boundaries. If the *LBR* is high enough, our hypothesis proposed in Sect. 2 is verified.

4. Results of the Analysis

4.1 Analysis of Eye Tracking Data

At first, we can have a look at the typical eye-fixation pattern of each group which we used in this experiment. Figure 6 shows the patterns of eye fixation when reading Japanese text, the size of each circle indicating the time of fixation. From this figure, we found that fixations of a Japanese subject occur for almost 1 in 10 characters. Fixations of a Kanji-area subject occur about every 3 or 4 characters, and most of the fixations occur on Kanji characters. Fixations of a Middle-area group occur also almost every 3 or 4 characters, like a Kanji-area group, but not only on Kanji characters, also on Hiragana characters. The Fixations of a Non-Kanji-Asia or a Non-Kanji-Europe group occur almost every 1 or 2 characters.

Figure 7 shows the averages of the saccade time and of the fixation time for each group†. Concerning the saccade time, Japanese native speaker group is longer. Also, the average saccade time of the Middle-area group is similar, although the standard deviation is much larger. Other groups' saccade times are shorter. Concerning the fixation time, for Japanese-native speaker group, it is the shortest, while for the other groups, it is longer.

In order to evaluate the difficulty of the 4 texts used in this experiment, we asked all subjects to choose the most difficult and the easiest text to read. Table 2 shows the numbers of subjects who chose each text as the most difficult or the easiest one. Each circle indicates the highest or higher numbers of objects. 9 subjects chose the Text-4 as the most difficult text. 9 subjects, most of who belong to the Kanji-area chose Text-1 as the easiest text. According to these observations, we roughly estimate the order of 4 texts as Text-1, Text-3, Text-2 and Text-4 from the easiest text, as shown in Table 2.

†Note that the saccade and fixation times shown here are rough estimations, since the sampling rate of EMR-8 is 30 Hz for binocular record.

Fig. 6 Typical pattern of eye fixation for each group.

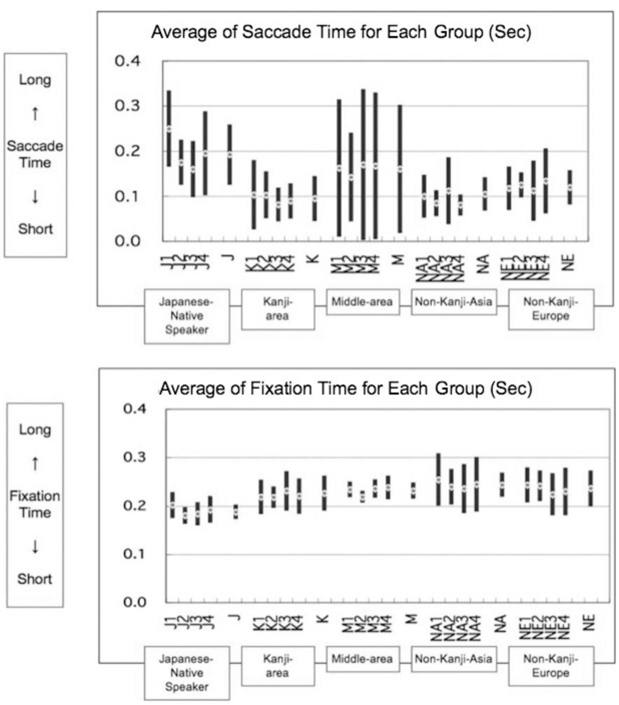


Fig. 7 Averages of saccade time and fixation time for each group.

Figure 8 shows the averages time of the saccade time and of the fixation time for each group (indicated in alphabetic order on the horizontal axis) and on each text (indicated its number). The saccade or the fixation time goes not seem to vary significantly from one text to another. Thus we cannot find any strong correlation between the saccade or the fixation time and the subjective difficulty of texts.

Table 2 Subjective difficulty of text.

Difficulty	ID	Text-1	Text-3	Text-2	Text-4	TOTAL
Most difficult	(J)	2	0	2	0	4
	(K)	0	0	2	3	5
	(M)	0	1	0	3	4
	(NA)	0	0	0	3	3
	(NE)	1	1	2	0	4
	TOTAL	3	2	6	9	20
	ID	Text-1	Text-3	Text-2	Text-4	TOTAL
Easiest	(J)	1	1	1	1	4
	(K)	5	0	0	0	5
	(M)	1	0	3	0	4
	(NA)	1	2	0	0	3
	(NE)	1	2	1	0	4
	TOTAL	9	5	5	1	20

◄····· Easy ······················ Difficult ···►

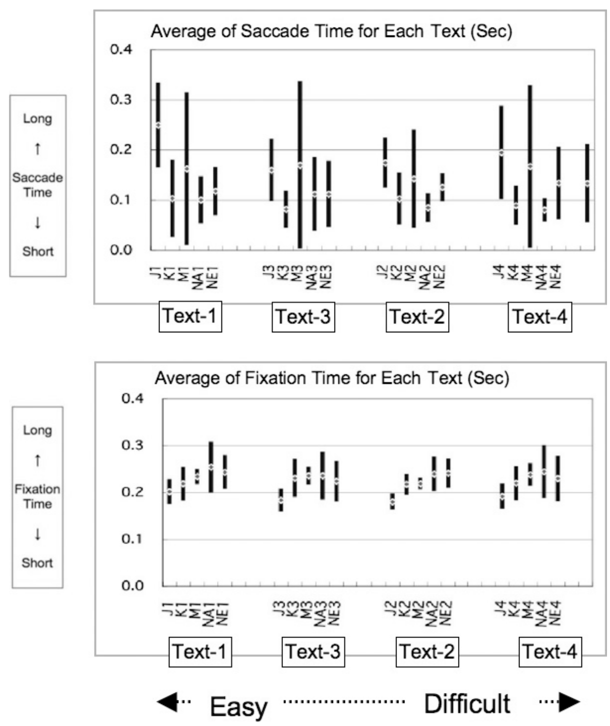


Fig. 8 Average of saccade time and fixation time for each text.

Table 3 Number of fixation times.

Fixation Times Group ID	Subjects	Text-1	Text-3	Text-2	Text-4	Total	Average
J-all	4	324	317	367	258	1266	317
K-all	5	644	753	993	694	3084	617
M-all	4	764	854	661	942	3221	805
NA-all	3	789	397	646	571	2403	801
NE-all	4	837	499	830	800	2966	742
Total	20	3358	2820	3497	3265	12940	647
(No. of Char.)		(178)	(221)	(202)	(172)	(773)	(193)

◄····· Easy ······················ Difficult ···►

To summarize, the types of eye movements, the saccade time and the fixation time differ according to the group of learners, namely their native languages. On the other hand, the difficulty of texts seems not to influence the eye movements of learners.

Table 3 shows the times of fixation. Each column indi-

Table 4 Analysis of backward saccade and entropy.

Start / Bunsetsu

	[certain]	(SD)
J-ave	0.950	(0.060)
K-ave	0.889	(0.089)
M-ave	0.933	(0.086)
NA-ave	0.913	(0.133)
NE-ave	0.921	(0.095)
(ALL-ave)	0.916	(0.085)
(E)	0.882	

End / Bunsetsu

	[certain]	(SD)
J-ave	0.900	(0.207)
K-ave	0.846	(0.092)
M-ave	0.911	(0.068)
NA-ave	0.754	(0.163)
NE-ave	0.868	(0.147)
(ALL-ave)	0.850	(0.140)
(E)	0.882	

Start / Clause

	[certain]	(SD)
J-ave	0.233	(0.122)
K-ave	0.154	(0.083)
M-ave	0.244	(0.158)
NA-ave	0.275	(0.095)
NE-ave	0.263	(0.396)
(ALL-ave)	0.223	(0.203)
(E)	0.214	

End / Clause

	[certain]	(SD)
J-ave	0.200	(0.064)
K-ave	0.239	(0.138)
M-ave	0.200	(0.153)
NA-ave	0.159	(0.111)
NE-ave	0.197	(0.109)
(ALL-ave)	0.204	(0.119)
(E)	0.214	

Start / Entropy 2.5

	[certain]	(SD)	[all]	(SD)
J-ave	0.950	(0.067)	0.900	(0.061)
K-ave	0.818	(0.107)	0.846	(0.083)
M-ave	0.893	(0.128)	0.800	(0.085)
NA-ave	0.830	(0.107)	0.754	(0.126)
NE-ave	0.933	(0.070)	0.868	(0.128)
(ALL-ave)	0.875	(0.100)	0.837	(0.101)
(E)	0.815		0.786	

End / Entropy 2.5

	[certain]	(SD)	[all]	(SD)
J-ave	0.833	(0.137)	0.783	(0.069)
K-ave	0.826	(0.177)	0.855	(0.144)
M-ave	0.857	(0.118)	0.844	(0.125)
NA-ave	0.786	(0.152)	0.739	(0.125)
NE-ave	0.868	(0.097)	0.829	(0.089)
(ALL-ave)	0.833	(0.139)	0.815	(0.115)
(E)	0.815		0.786	

Start / Entropy 3.0

	[certain]	(SD)	[all]	(SD)
J-ave	0.700	(0.138)	0.650	(0.130)
K-ave	0.727	(0.170)	0.675	(0.220)
M-ave	0.714	(0.318)	0.622	(0.239)
NA-ave	0.681	(0.099)	0.609	(0.159)
NE-ave	0.800	(0.112)	0.697	(0.164)
(ALL-ave)	0.730	(0.179)	0.657	(0.186)
(E)	0.669		0.611	

End / Entropy 3.0

	[certain]	(SD)	[all]	(SD)
J-ave	0.691	(0.122)	0.633	(0.060)
K-ave	0.598	(0.224)	0.564	(0.148)
M-ave	0.600	(0.245)	0.578	(0.242)
NA-ave	0.524	(0.086)	0.435	(0.184)
NE-ave	0.793	(0.149)	0.632	(0.120)
(ALL-ave)	0.640	(0.199)	0.567	(0.165)
(E)	0.669		0.611	

cates the text, while each row indicates the group of learners. The last row shows the number of characters in texts.

Table 3 shows that Japanese native speakers read texts with fewer fixations than other groups. The number of fixation times that of foreigners are about twice or 2.5 times that of Japanese. The number of fixation times is highest for Text-3, and lowest for Text-2. But we cannot find any significant differences among the 4 texts.

4.2 Analysis of Backward Saccade

In this section, we report the results of our analysis of backward saccades and linguistic boundaries. We found 367 backward saccades from the eye-tracking data of 20 subjects. The *LBR* is calculated for fixations at the start and at the end of these backward saccades. Results are summarized in Table 4. It shows the *LBR* for several cases, classified according to the following aspects:

- Start or end

  Fixations are distinguished if they are at the start or at the end of backward saccades. The *LBR* for start fixations are shown in the tables on the left, and the *LBR* for end fixations in the tables on the right.
- Types of linguistic boundaries

  In Table 4, we showed the results for *bunsetsu* and clause boundaries as well as two kinds of boundaries defined by entropy ('Entropy 2.5' and 'Entropy 3.0'). When we examined the relation between backward saccades and word boundaries, LBR was equal to 1 for almost all subjects. This is because we tried to find boundaries around 3 or 4 characters from the center of fixations (see Fig. 5). Therefore we omitted showing detail results for word boundaries.
- Threshold of entropy (T)

  When we define linguistic boundaries according to entropy, the number of linguistic boundaries in the text can be controlled by the threshold T. In this experiment, we set the threshold T to 2.5 and 3. Among 177 positions in the text†, 52 (29.4%) and 35 (19.8%) positions are regarded as linguistic boundaries when $T = 2.5$ and $T = 3.0$, respectively. We manually checked these 52 positions and found that 10 are clause boundaries, 29 *bunsetsu* boundaries and 50 word boundaries. While only two positions are wrongly identified as linguistic boundaries. We suppose that these are more clear boundaries than others.

- Groups of subjects

We separately calculate the *LBR* for fixations of subjects in the 5 different groups described in Sect. 3.1. In Table 4, the average and standard deviation (SD) of the *LBR* for each group is shown in each row. The same data for all subjects is also shown in the 'All-ave' row.

- Uncertain entropy position

When we define linguistic boundaries according to the entropy, the entropy at some positions cannot be calculated due to data sparseness as described in Sect. 3.2. If such uncertain entropy positions are located in the fixation area, the judgment of whether a linguistic boundary exists in the fixation area is also uncertain. So we calculated the *LBR* when the entropy at all positions in the fixation area can be calculated (in 'Certain' column), and the *LBR* for all fixations, including ones such that uncertain entropy positions exist in the fixation area ('All' column). In the latter case, the *LBR* is just an approximation. Note that it is underestimated since uncertain entropy positions may be real linguistic boundaries.

The expectation of the *LBR* is also shown in the 'E' row in Table 4. This is defined as the proportion of points such that linguistic boundaries exist in the neighborhood of all points in the text. Here, points in the text mean both centers of characters as indicated by the star in Fig. 5 (a), and positions between characters as indicated by the star in Fig. 5 (b). That is, the expectation of the *LBR* represents the probability that when a point in the text is randomly chosen as the center of fixation, one or more linguistic boundaries exist at that point.

5. Discussion

We here discuss the results shown in Table 4. First, the *LBR* for start points of backward saccades are greater than the expected ratio (E). On the other hand, the *LBR* for end points are lower than E excepting for the case of 'Entropy 2.5'. These results indicate that our hypothesis is valid for start points, but not for end points. That is, language learners tend to start their backward saccades at linguistic boundaries. This suggests that learners might recognize linguistic boundaries when they start to move their eyes backward. Hereafter, we focus on only start points of backward saccades.

Next, we will compare types of linguistic boundaries. We found that the difference between the average of the *LBR* and E for start points is 0.034 for *bunsetsu*, 0.009 for clause, 0.042 for entropy ($T = 2.5$) and 0.061 for entropy ($T = 3.0$), respectively. Thus, the differences for boundaries defined by entropy are greater than those for *bunsetsu* or clause boundaries. It may indicates that start points of backward saccades more likely agree with clear linguistic boundaries empirically estimated from a large corpus than *bunsetsu* or clause boundaries.

Finally, we will discuss the differences of groups of subjects. The LBR of group J (Japanese-Native-Speaker) for start fixations is relatively high: it is the greatest among all groups for *bunsetsu* and entropy ($T = 2.5$). We guess that this is because Japanese native speakers have rich knowledge about Japanese, and can recognize linguistic boundaries correctly. On the other hand, the LBR of group K (Kanji-area) is the lowest for all types of linguistic boundaries except for entropy ($T = 3.0$). Surprisingly, the *LBR* of group NE (Non-Kanji-area-Europe) is relatively high for all cases. Although Europeans might not be familiar with Kanji characters, they often start their backward saccades around linguistic boundaries. The above observations may suggest that the reading process of Japanese learners is not strongly related to their familiarity with Japanese or Kanji.

6. Syntactic Structure and Entropy

Another question about backward saccades is whether Japanese learners tend to start or end their backward eye movements at deep positions in syntactic structures of sentences. As a preliminary survey, we investigated the relationship between the depth of syntactic structures and the entropy of the N-gram model. First, we obtained syntactic structures of sentences in the text used in our experiment, using the Japanese CALL system *Asunaro*††. For each character in sentences, we obtained the entropy of the 5-gram model at the position after the characters and its depth in a syntactic tree, where the depth is the distance from the root to the character. We obtained the statistics shown in Table 5.

The correlation between the depth and the average of entropy is −0.0941. This means that there is no relationship between the two variables.

Table 5 Averages of saccade time and fixation time for each group.

Depth	Average of Entropy	No. of Characters
3	1.9314	15
4	1.3876	5
5	2.4844	20
6	1.9822	19
7	1.3822	49
8	1.8715	30
9	1.6572	29
10	1.0735	11
All	1.8599	178

†Note that the text used in this experiment contains 178 characters.

††http://hinoki.ryu.titech.ac.jp/

7. Conclusion

In this paper, we analyzed eye-tracking data to investigate the reading process of Japanese learners. We proposed the hypothesis that learners tend to start and end their backward saccades around linguistic boundaries. The analysis on the real eye tracking data showed that our hypothesis is valid to some degree.

In future research, large-scale experiments are required. The size of the experiment reported here is not sufficient in terms of both the number of texts and the number of subjects. It is necessary to conduct a larger experiment and to analyze the results further in order to reveal the reading process of Japanese learners. We have already obtained eye-tracking data for 3 other texts with the same subjects. The analysis of these 3 texts will be performed shortly. A detailed comparison between the nationalities and languages of Japanese learners will also be made. Another problem of the current experiment is that the sampling rate of eye tracking system is not so high, 30 Hz for binocular record. Analysis on the eye tracking data obtained by a more sophisticated recorder is desired. Furthermore, we plan to empirically investigate the relation between backward saccades and the depth in syntactic structures.

References

[1] M. Ford, J. Bresnan, and R.M. Kaplan, "A competence-based theory of syntactic closure," in The Mental Representation of Grammatical Relations, ed. J. Bresnan, pp.727–796, MIT Press, 1982.
[2] N. Osaka, "Size of saccade and fixation duration of eye movements during reading," Psychophysics of Japanese Text Processing, Journal of the Optical Society of America, vol.A, no.9, pp.5–13, 1992.
[3] S. Saida, Reading and Eye Movement, Experimental Psychology of Eye Movement, pp.167–197, Nagoya University Press, Nagoya, Japan, 1993.
[4] J. Shigematsu and T. Konosu, "The research on the reading processes by non-native speakers using an eye-camera," Spring Proc. KNG, pp.31–42, 1993.
[5] A. Tera, T. Kitamura, and K. Ochimizu, "Japanese reading support system "dictlinker"," Autumn Proc. KNG, pp.43–48, 1996.
[6] A. Tera, "Extensive reading support system for learning Kanji," Nihongo-Gaku, Nihongo Kyouiku, vol.16, no.6, pp.101–108, 1997.
[7] A. Tera, T. Kitamura, K. Ochimizu, T. Graham, and A. Lavin, "Japanese reading support system (DL) evaluation — Results from MIT user survey," JLEM, vol.4, no.1, pp.26–27, 1997.
[8] K. Rayner, "Eye movements in reading and information processing: 20 years of research," Psychological Bulletin, vol.124, pp.372–422, 1998.
[9] A. Tera, T. Kitamura, and M. Okumura, "The verification of Japanese reading support system "DL" — The research of the process of Japanese learner's reading," JLEM, vol.6, no.1, pp.20–21, 1999.
[10] C. Manning and H. Schutze, Foundations of Statistical Natural Language Processing, The MIT Press, 1999.
[11] A. Tera, H. Motizuki, and A. Shimazu, "An analysis of the words that Japanese language learners need during reading," Autumn Proc. 27th JSISE Conference, pp.259–260, 2002.
[12] Chunichi Newspaper, Chunichi Newspaper Company, 2004.
[13] A. Tera and K. Sugiyama, "Eye-tracking analyses of Japanese reading processes (IV) — Forward and backward jumping eye movements," Proc. IEICE Gen. Conf. 2006, TL2005-35-48, pp.43–48, 2006.
[14] A. Tera and K. Sugiyama, "Eye-tracking analyses of Japanese reading processes," Proc. KICSS, pp.108–114, 2006.

Akemi Tera received the M.S. degree in Knowledge Science from School of Knowledge Science, Japan Advanced Institute of Science and Technology. Her main field is the Japanese language education and Japanese language learning support system. She is now a PhD student of Japan Advanced Institute of Science and Technology.

Kiyoaki Shirai received the B.S (1993), M.S (1995) and D.S (1998) degrees in Engineering from Tokyo Institute of Technology. He was a research assistant at Tokyo Institute of Technology from 1998 to 2001. He has been an associate professor at School of Information Science, Japan Advanced Institute of Science and Technology since 2001. His research interests lie in the fields of natural language processing.

Takaya Yuizono received the B.E., M.E., and Dr. of Engineering from Kagoshima University, in 1994, 1996, 1999, respectively. He was a research associate in Kagoshima University, a lecturer and an associate professor in Shimane University, respectively. He has been an associate professor at School of Knowledge Science, Japan Advanced Institute of Science and Technology since 2005. His research interests include in groupware, computer-supported cooperative work and knowledge medium. He is a member of ACM and IEEE.

Kozo Sugiyama received the M.S. and D.S degrees in Earth Sciences from Nagoya University in 1971 and 1974, respectively. During 1974–1997, he worked for Fujitsu Ltd and Fujitsu Laboratories Ltd to study system engineering, information visualization, and creativity support. He stayed in International Institute of Applied Systems Analysis, Austria in 1982–1983. He is now a professor of School of Knowledge Science, Japan Advanced Institute of Science and Technology.