10th International Conference on Speech Prosody 2020
25-28 May 2020, Tokyo, Japan

Silent and oral sentence reading in Estonian: investigating the effect of phonetic quantity on eye movements

Pärtel Lippus, Kaidi Lõo

Institute of Estonian and General Linguistics, University of Tartu, Estonia
partel.lippus@ut.ee, kaidi.loo@ut.ee

Abstract

This paper studies the processing of prosodic information in silent and oral reading of Estonian. In Estonian, that has a three-way quantity opposition, the long and overlong quantity degrees are not distinguished in orthography. In a frequent morphological word type, genitive vs. partitive forms are marked with quantity alteration, resulting in ambiguous homographs where the correct word form can only be determined from the context. Such words are expected to require more processing compared to those with no quantity alteration and unambiguous orthographic form.

An eyetracking experiment was carried out with 24 native Estonian speakers performing a reading task both in silent reading and read-aloud mode. The target words were embedded in the central position of compound sentences, and the placement of the target word context was altered. The results show that the words with quantity alteration required more processing compared to the words with no quantity alteration, when the participants were reading aloud. In the case of silent reading the quantity effect did not appear, but the processing was faster when the context preceded the target word.

Index Terms: eye movements, quantity, oral reading, silent reading, Estonian

1. Introduction

A widely researched topic in psycholinguistics has been whether and how we use phonological information while reading. A common finding is that reading aloud is slower than silent reading as articulatory processes are involved [1].

However, it has also been argued that in silent reading we use inner speech (see e.g., [2, 3, 4]). For instance, readers are often confronted with phonological information that is not directly expressed by orthography. One such example is lexical stress in English. A word such as *suspect* is a noun when the stress is on the first syllable i.e., [ˈsʌspekt] or a verb when the stress is on the second syllable i.e., [səsˈpekt]. In support of the claim that prosodic information is used while reading, studies have reported that reading words with a more complex stress pattern takes longer than reading words with a simple stress pattern [5, 6, 7]. Similarly, vowel length has been shown to affect silent reading. For example, a German word such as *Weg* 'path' is pronounced with a long [eː], but a word such as *weg* 'gone' with a short [ɛ], although the written orthography for both is a single *e*. Reading words with long vowels has been shown to take longer than reading words with short vowels [8, 9].

Evidence on phonological effects on reading in natural sentence context has been relatively sparse. A few recent studies have, however, suggested that phonological effects also emerge in the context of silent sentence reading (see e.g., [10, 11, 12, 13]).

The current study investigates phonological effects in the processing of Estonian compound sentences including inflected target nouns in alternating quantities. Previously, Estonian nominal processing has been studied on a single word level by Lõo et al. [14] and [15]. These studies concentrated on frequency and paradigmatic effects of inflected forms. The analyses of the word naming task, however, indicated that homographic inflected forms (e.g., *kassi* 'cat, genitive/partitive' can be pronounced either with long or overlong [s]) were overall produced longer compared to non-homographic nouns, indicating processing difficulties possibly due to quantity ambiguities of the written form [16].

The focus of the current paper is on Estonian three-way phonetic quantity system, which is not always expressed in orthography. Estonian phonetic quantity has not been investigated from the functional processing perspective, but rather from the perspective of phonetic description, e.g. [17, 18].

The three-way quantity system in Estonian is not directly linked to morphology, that is, the minimal pairs and triplets of quantity are not necessarily grammatical forms of the same lexical word. In fact the most frequently used examples of minimal triplets are not related at all: short (Q1) [sɑtɑ] 'hundred' - long (Q2) [sɑːtɑ] 'send sg2imp' - overlong (Q3) [sɑːːtɑ] 'to get' or Q1 [kɑlɑ] 'fish' - Q2 [kɑllɑ] 'Calla, Bog Arum' - Q3 [kɑlːlɑ] 'pour, sg2imp'.

In terms of the Estonian morphological system, there are productive word types where the quantity alteration is integrated in the conjugation system. For some word types, quantity is the only marker of the basic cases: genitive forms are in Q2 ([seppɑ] 'smith', [koːlʲi] 'school', [kolʲli] 'bogey') and partitive forms in Q3 ([sepːpɑ], [koːːlʲi], [kolʲːli]). Yet there are other word types in Estonian where the cases are marked with an ending without changing the quantity degree, and within this group of words, there are unique segmental combinations that do not have minimal pairs or triplets in terms of quantity, e.g. Q2 [hærrɑ] 'mister', Q3 [pɤrːku] 'hell'; i.e. the same phoneme sequence with different quantity is not a meaningful word in Estonian.

Estonian contemporary orthography was developed in the 19th century and it was largely copied from the Finnish orthography, cf. [19]. Finnish has a binary quantity system and in the orthography short sounds are marked with one letter while long sounds use double letters. In the Estonian orthographic system the distinction between Q2 and Q3 is not marked on regular bases. Q1 is written with a single letter and for both Q2 and Q3 a double letter is used, so that in a written text the quantity and the grammatical case of the words like *kooli* or *kolli* can be identified only from the context. In the orthography, there is an exception for stop consonants: since there is no voiced vs. voiceless opposition of stops in Estonian, the letters *g, b, d* are used for Q1, while single letters *k, p, t* mark Q2 (e.g. *sepa*), and double letters mark Q3 (*seppa*). For learners this exception ac-

 10.21437/SpeechProsody.2020-62

tually complicates the systematic understanding of the Estonian phonology, but allows us to test the prosodic processing effect in a context-dependant reading task.

The current paper addresses the following questions:

1) Does quantity alteration require more processing compared to the words that do not have quantity alteration? We expect longer processing time and more need to look back at the target word in the case of words that use quantity alteration for case marking as compared to words that do not have quantity alteration.
2) Does orthographic ambiguity in the marking of quantity alteration require more processing? We expect the words where the quantity is clear from orthography to be read faster.
3) Is the prosodic processing activated for quantity alteration also in the silent reading mode? Based on previous findings form other languages we expect the words with quantity alteration to require more processing also in the silent reading mode.

2. Experiment

The experiment was carried out at the Institute of Psychology, University of Tartu, Estonia. Twenty four native speakers (16 female, 8 male; mean age=29.5 years) took part in the study on voluntary basis. They all had normal or corrected-to-normal vision.

2.1. Materials

The stimuli were 320 compound sentences. The target word was positioned sentence-medially followed by a dependent clause (a filler), and the context of the target word was placed in the end of the sentence. The length of the test sentences ranged from 76 to 87 characters (mean 83). The target word was a disyllabic CVC.CV word with primary stress, including only short vowels and an intervocalic geminate or consonant cluster. Thus the target words were either in long (Q2) or overlong (Q3) quantity degree. Target words were matched in terms of frequency (based on the 15 million token Balanced Corpus of Estonian) across the sentence types. Three types of target words were combined with two sentence structures into four types of stimuli (see Table 1 for examples):

A: No quantity alteration. The segmental sequence only exists in one quantity degree and there is no meaningful word in the language that has the same sequence in other quantity degree. The target word is preceded and followed by a filler phrase and the context of the target word is in the end of the sentence (same sentence structure is used for types B and C).

B: Ambiguous quantity alteration. Quantity alteration is used for case marking (usually genitive form is in Q2 and partitive form in Q3) but the quantity is orthographically ambiguous and in written text determined only by the following context.

C: Unambiguous quantity alteration. Quantity alteration is used for case marking (as in type B), but the quantity is marked by orthography (stop consonants, that are written with single letter in Q2 and double letter in Q3).

D: Context before target word. Additionally there was a fourth set of sentences where the target words are orthographically ambiguous (as in type B) but the context that determines the case and the quantity of the target word comes already in the beginning of the sentence, and the following clause until the end of the sentence is a filler.

The stimuli were arranged into two lists of 160 sentences (2 lists x 4 types x 2 quantites x 20 sentences = 320 stimuli in total).

2.2. Apparatus

The eye movements were collected with a tower mounted monocular EyeLink 1000 eye tracker by SR Research Ltd. with a resolution of 1000 Hz, ca 3.0 ms delay, and the average spatial accuracy of approximately 0.5 degrees of arc. For silent reading, a stabilizing chin and forehead rest was used, for reading aloud task only the forehead rest was used. Participants' dominant eye was calibrated on a nine-point grid (for five it was the left eye, for 19 the right eye). The stimuli were presented on a 21-inch Acer computer screen in lower case black 12-point Courier New font using the ExperimentBuilder software by SR Research Ltd.

2.3. Procedure

Participants completed a two-part self-paced reading task where they were asked to read out sentences both silently and aloud for comprehension. It was counterbalanced whether they started with the reading aloud or reading silently task. During the reading aloud task also the speech signal was recorded (to be analysed in future papers). In both tasks, each trial started with a drift correction on the left of the screen, after which the sentence appeared on a single line on the screen. Participants were instructed to press a key on the keyboard after they had finished reading (aloud) the sentence. Each task started with four practice trials, which were followed by 160 experimental trials. About half of the sentences were followed by a yes/no-comprehension question. Every 80th trial was followed by a short break. There was a longer break between the two parts. After each break, the participant was calibrated again. Before the start of the experiment, participants filled out a short language background questionnaire. The whole procedure lasted approximately 90 minutes.

3. Analysis and Results

Trials with incorrect answers to comprehension questions were removed from the analysis (2.7% of the data). All the experimental sentences were divided into four segments (interest areas). The current analysis focuses only on the second segment, i.e., the segment where the target word was presented.

Trials with segment reading times below 50 ms and above 1700 ms (2.5 standard deviations of the mean target word reading time) were also removed (9.7 % of the data). The data was analyzed using Generalized Mixed Model regression (the R-package *mgcv*; [20]). We opted for mixed modeling to account for item- and subject-level variability.

The main response variables were: 1) total reading time (i.e., the summed duration of all fixations in milliseconds; log-normalised) of the target word. 2) look-back at the target word (0 or 1, i.e., whether the current interest area received at least one regression from a later interest area, 1 if the interest area was entered from a later segment; 0 if not).

The main variables of interest were: 1) sentence type (A, B, C or D; sentence type B was chosen as the reference level as the main interest was the comparisons of ambiguous quantity with

Table 1: *Examples of the stimulus sentences. The target word is highlighted in bold and the context phrase is underlined.*

Type	Quantity	Sentence
A	Q2	Kokasaates üritati väita, et **pasta**, millel on palju lisandeid, valmib poole tunniga. *In the cooking show it was claimed that the* ***pasta*** *that has many sides will be ready in half an hour.*
A	Q3	Tänapäeval tuleb jahu poest ja **veski**, mis on küll muinsuskaitse all, muudkui laguneb. *Nowadays flour comes from the shop and the* ***mill*** *that is under heritage protection is gradually falling apart.*
B	Q2	Wikipediast võib lugeda, et **linna**, kus asub maakonnakeskus, asutas Napoleon. *It can be read from Wikipedia that the* ***town*** *where the county centre is, was founded by Napoleon.*
B	Q3	Wikipediast võib lugeda, et **linna**, kus asub maakonnakeskus, juhib reformierakond. *It can be read from Wikipedia that the* ***town*** *where the county centre is, is lead by the Reform party.*
C	Q2	Uus brauser on küll kiirem aga **nupu**, kust saab seadeid muuta, on arendaja ära peitnud. *The new browser is faster but the* ***button*** *for changing the settings has been hidden by the developer.*
C	Q3	Uus brauser on küll kiirem, aga **nuppu**, kust saab seadeid muuta, on raske üles leida. *The new browser is faster but the* ***button*** *for changing the settings is difficult to find.*
D	Q2	Ärimees ehitas sellise **lossi**, mis oli Talvepalee täpne koopia, ainult natuke väiksem. *The businessman built a* ***castle*** *that was an exact copy of the Winter Palace, just a little bit smaller.*
D	Q3	Ärimees tahtis endale sellist **lossi**, mis oli Talvepalee koopia, ainult natuke väiksem. *The businessman wanted a* ***castle*** *that was an exact copy of the Winter Palace, just a little bit smaller.*

other types), and 2) reading mode (silent vs. aloud; reading silently was chosen as the reference level).

3.1. Total reading time of the target word

The final model fitted to the total reading time of the target word (represented in Figure 1 and Table 2) showed a main effect of the sentence type. Target words in type D, where the context was given before the target words, were read faster than target words in type B, where the context was given at the end of the sentence (β=-0.23, t=-3.45, p= 0.001). There was no significant difference between type B and type A, and type B and type C.

There was a main effect of reading mode: sentences were read aloud slower than silently (β = 0.31, t=11.48, $p < 0.0001$). There was also a small interaction between the sentence type and the reading mode. Sentences in type A (where there is no quantity alternation possible) were read aloud faster than in sentence type B (β = -0.08, t=-2.21, p=0.03). Participants read overall faster when conducting the second part as opposed to the first part of the experiment (β = -0.06, t=-4.58, p=0.0001). This variable did not significantly interact with the sentence type or reading type. Finally, random intercept for target word and by-subject random smooths for trial were added to the model as random effects.

3.2. Looking back at the target word

The logistic regression model fitted to the looking back data (represented in Figure 2 and Table 3) was overall in line with the total reading time model. The likelihood of participants looking back at the target word later was smaller in sentence type D (β=-0.48; z= -2.54; p=0.01) compared to sentence type B.

There was no difference between sentence type B and sentence type A and between sentence type B and sentence type C. Participants were more likely to look back at the target word when reading aloud compared to reading silently (β=0.52, z=4.73, p=0.0001). There was also a significant interaction between the reading mode and the sentence type. In reading-aloud-mode, the target word in both sentence type A (β=-0.47, z=-2.21, p=0.003) and in sentence type D (β=-0.48, z=-2.87, p=0.005) was less likely to receive look-backs to the target compared to the target word in sentence type B. There was no significant difference, however, between sentence type B and sentence type C. Participants were less likely to look back at the target word when conducting the second part as opposed to the first part of the experiment (β=-0.19, z=-4.58, p=0.0001). Random intercepts for target word and by-subject random smooths for trial were added to the model as random effects.

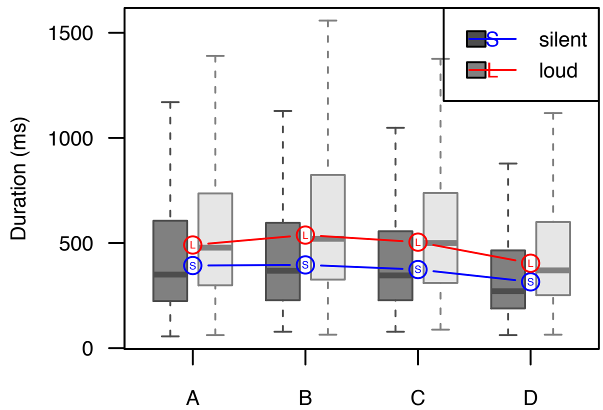

Figure 1: *The total reading time of the target word. The boxplot shows the distribution of the data and the points represent the model estimates.*

4. Discussion and Conclusions

Firstly, we analysed the total reading time of the target word (illustrated in Figure 1). In the case of silent reading, there was only an effect of sentence structure and the prosodic processing effect did not appear: reading time was significantly shorter with type D sentences, where the context was presented before the target word, but there was no difference between the other stimuli types. The results did not support our hypothesis whereby also in silent reading mode the prosodic processing would slow down the reading of type B sentences where the quantity degree of the target word was ambiguous.

In the reading aloud task, there also was a small difference

Table 2: *Summary of the partial effects in GAMM fitted to the log-transformed total reading time of the target word.*

A. parametric	**Estimate**	**St.E.**	**t-value**	**p-value**
(Intercept)	5.98	0.06	93.42	< 0.0001
typeA	-0.01	0.06	-0.98	0.90
typeC	-0.06	0.06	-4.10	0.33
typeD	-0.23	0.07	-3.45	0.001
readAloud	0.31	0.03	11.48	< 0.0001
secondPart	-0.06	0.01	-4.58	0.0001
typeA:readAloud	-0.08	0.04	-2.21	0.03
typeC:readAloud	-0.01	0.04	-0.20	0.84
typeD:readAloud	-0.06	0.04	-1.58	0.12
B. smooth terms	**edf**	**Ref.df**	**F-value**	**p-value**
s(subject,trial)	72.78	215	4.88	< 0.0001
s(target)	201.03	234	5.79	< 0.0001

Table 3: *Summary of the partial effects in logistic GAMM fitted to the probability of looking back at the target word.*

A. parametric	**Estimate**	**St.E.**	**t-value**	**p-value**
(Intercept)	-0.81	0.17	-4.65	0.0001
typeA	-0.06	0.17	-0.27	0.71
typeC	-0.16	0.17	-0.91	0.36
typeD	-0.48	0.19	-2.54	0.01
readAloud	0.52	0.11	4.73	0.0001
secondPart	-0.19	0.01	-4.58	0.0001
typeA:readAloud	-0.47	0.16	-2.21	0.003
typeC:readAloud	-0.28	0.16	-0.20	0.07
typeD:readAloud	-0.48	0.17	-2.87	0.005
B. smooth terms	**edf**	**Ref.df**	**Chi.sq**	**p-value**
s(subject,trial)	154.33	234	451.1	< 0.0001
s(target)	47.86	215	392.6	< 0.0001

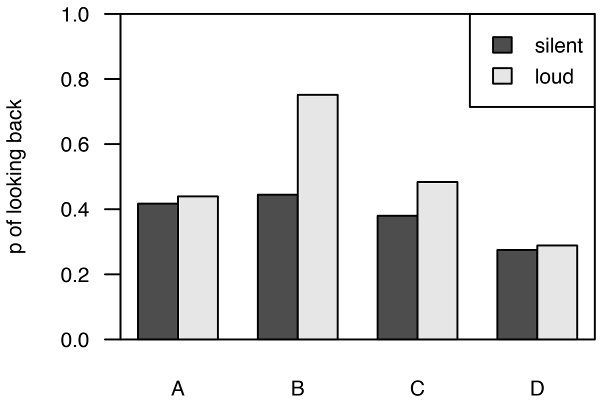

Figure 2: *The probability of looking back at the target word after reaching the end of the sentence estimated from the model.*

between type B and type A, but not between type B and type C sentences, type A being read slightly faster than type B and C target words. This result suggests that it is not only orthographic ambiguity that slows down the reading but more generally it depends on whether there is quantity alteration used for case marking or not. Type A target words that only exist in the language in one quantity degree are read faster. However, this effect was rather minute.

Secondly, we looked at the probability of looking back at the target word once having already reached a following interest area. From Figure 2 we can see that in the silent reading mode there is only a significant effect of the sentence structure: in the case of type D where the context was given already before the target word, the probability of looking back at the target word was significantly lower than in the case of other types.

In regard of looking back to the target word in the reading aloud task, there is a clear effect of prosodic processing. The model shows a significantly higher probability of looking back to the words with quantity alteration. For type B where the quantity marking is orthographically ambiguous the probability to look back is the highest, but also for type C the level is higher than in the silent reading mode.

These results are partly in line with the previous research. We replicate the well-established finding that sentences are read slower aloud than silently [1]. However, contrary to our expectations no effects of prosodic processing were detected in silent reading. This might have several reasons. First, most previous research has investigated phonological effects in a single word, but not in a sentence context [2, 3, 4]. Second, our current analysis concentrated only on the effects on the target word. However, previous eyetracking studies have detected that phonological effects might arise very early already in the parafoveal view, i.e., before the fixation on the actual target word [21, 22]. Investigating eye movements of the previous and following context will be subject to future research. Future research will also include looking in-depth into the data from the speech signal collected during the reading aloud task as well as conducting direct comparisons between the speech and the eye movement data.

5. Acknowledgements

We are very grateful to the volunteer participants. We would also like to thank Nele Põldver, Richard Naar and Tiit Mogom from the Institute of Psychology, Univesrsity of Tartu for their assistance with the data collection and Eva Liina Asu for proofreading this paper. This research was in part supported by Estonian Research Council Mobilitas Pluss postdoctoral researcher grant (no. MOBJD408) given to the second author. The paper is dedicated to the 102nd anniversary of the Republic of Estonia.

6. References

[1] K. Rayner, "Eye movements in reading and information processing: 20 years of research." *Psychological bulletin*, vol. 124, no. 3, p. 372, 1998.

[2] I. Berent and C. A. Perfetti, "A rose is a reez: The two-cycles model of phonology assembly in reading english." *Psychological Review*, vol. 102, no. 1, p. 146, 1995.

[3] G. Lukatela and M. T. Turvey, "Visual lexical access is initially phonological: 2. evidence from phonological priming by homophones and pseudohomophones." *Journal of Experimental Psychology: General*, vol. 123, no. 4, p. 331, 1994.

[4] J. C. Ziegler, L. Ferrand, A. M. Jacobs, A. Rey, and J. Grainger, "Visual and phonological codes in letter and word recognition: Evidence from incremental priming," *The Quarterly Journal of Experimental Psychology Section A*, vol. 53, no. 3, pp. 671–692, 2000.

[5] J. Ashby and C. Clifton Jr, "The prosodic property of lexical stress affects eye movements during silent reading," *Cognition*, vol. 96, no. 3, pp. B89–B100, 2005.

[6] M. Breen and C. Clifton Jr, "Stress matters: Effects of anticipated lexical stress on silent reading," *Journal of memory and language*, vol. 64, no. 2, pp. 153–170, 2011.

[7] E. Reinisch, A. Jesse, and J. M. McQueen, "Early use of phonetic information in spoken word recognition: Lexical stress drives eye movements immediately," *The Quarterly Journal of Experimental Psychology*, vol. 63, no. 4, pp. 772–783, 2010.

[8] M. Abramson and S. D. Goldinger, "What the reader's eye tells the mind's ear: Silent reading activates inner speech," *Perception & Psychophysics*, vol. 59, no. 7, pp. 1059–1068, 1997.

[9] G. Lukatela, T. Eaton, L. Sabadini, and M. Turvey, "Vowel duration affects visual word identification: evidence that the mediating phonology is phonetically informed." *Journal of Experimental Psychology: Human Perception and Performance*, vol. 30, no. 1, p. 151, 2004.

[10] J. Ashby, R. Treiman, B. Kessler, and K. Rayner, "Vowel processing during silent reading: Evidence from eye movements." *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 32, no. 2, p. 416, 2006.

[11] L. Huestegge, "Effects of vowel length on gaze durations in silent and oral reading," *Journal of Eye Movement Research*, vol. 3, pp. 1–18, 2010.

[12] M. Yan, Y. Luo, and A. W. Inhoff, "Syllable articulation influences foveal and parafoveal processing of words during the silent reading of chinese sentences," *Journal of Memory and Language*, vol. 75, pp. 93–103, 2014.

[13] J. Pan, J. Laubrock, and M. Yan, "Parafoveal processing in silent and oral reading: Reading mode influences the relative weighting of phonological and semantic information in chinese." *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 42, no. 8, p. 1257, 2016.

[14] K. Lõo, J. Järvikivi, and R. H. Baayen, "Whole-word frequency and inflectional paradigm size facilitate Estonian case-inflected noun processing," *Cognition*, vol. 175, pp. 20–25, 2018.

[15] K. Lõo, J. Järvikivi, F. Tomaschek, B. V. Tucker, and R. H. Baayen, "Production of Estonian case-inflected nouns shows whole-word frequency and paradigmatic effects," *Morphology*, vol. 28, no. 1, pp. 71–97, 2018.

[16] K. Lõo, "Morphology in Estonian speech production," *General's Paper, Department of Linguistics, University of Alberta*, 2016.

[17] I. Lehiste, "Prosodic change in progress: From quantity language to accent language," in *Development in prosodic systems*, ser. Studies in Generative Grammar, P. Fikkert and H. Jakobs, Eds. Berlin, New York: Mouton de Gruyter, 2003, pp. 47–65.

[18] P. Lippus, E. L. Asu, P. Teras, and T. Tuisk, "Quantity-related variation of duration, pitch and vowel quality in spontaneous Estonian," *Journal of Phonetics*, vol. 41, no. 1, pp. 17–28, Jan. 2013. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0095447012000721

[19] H. Laanekask and T. Erelt, "Written Estonian," in *Estonian language*, ser. Linguistica Uralica, Supplementary Series, M. Erelt, Ed. Tallinn: Estonian Academy Publishers, 2003, no. 1, pp. 273–342.

[20] S. N. Wood, *Generalized Additive Models*. New York: Chapman & Hall/CRC, 2006.

[21] J. Ashby, J. Yang, K. H. Evans, and K. Rayner, "Eye movements and the perceptual span in silent and oral reading," *Attention, Perception, & Psychophysics*, vol. 74, no. 4, pp. 634–640, 2012.

[22] A. W. Inhoff and R. Radach, "Parafoveal preview benefits during silent and oral reading: Testing the parafoveal information extraction hypothesis," *Visual Cognition*, vol. 22, no. 3-4, pp. 354–376, 2014.