Prosodic and visual indicators for linguistic processing in reading aloud

Armanda Costa[1] Isabel Falé[1,2] Paula Luegi[1]

[1]Universidade de Lisboa – CLUL | [2]Universidade Aberta

armandacosta@letras.ulisboa.pt; ifale@campus.ul.pt; paulaluegi@letras.ulisboa.pt

22nd AMLaP Conference: Architectures and Mechanisms for Language Processing, Bilbao, September 1 – 3, 2016

How can eye movements and prosody conjointly inform us about linguistic processing?

Reading is firstly a visual process that feeds a complex bundle of cognitive operations leading to comprehension. The reader has to deal with multiple linguistic sources to decode, categorize, parse and interpret sequences of words.

Visual word recognition (VWR) can proceed more or less easily depending on lexical properties, such as length, frequency and familiarity, phonological structure including syllable type and stress. Isolated words need to be organized in structures by imposition of working memory and the need to build interpretable syntactic units. For that, grammar and intonation play a crucial role, not neglecting the help of punctuation (Perfetti 1999; Hirotani et al. 2006).

Reading aloud involves all the representations and processes intervening in silent reading, more the planning and production of speech. Therefore it is as an extremely informative communicative task about the cognitive processes it implies. (Ashby et al. 2012; Benjamin & Schwanenflugel 2010).

Eye movements and **reading speech** can provide important indicators to the study of written language processing. Scan paths reveal patterns of saccadic movements, several measures of fixation time in a word or region mirror lexical and syntactic operations on the print input (Rayner et al. 2005). In reading speech, prosodic strategies such as intonation, speech rate or hesitations point out the same cognitive processes than eye movements, however they are expressed in a delayed temporal line (eye-voice span) (Inhoff et al. 2011; Frazier et al. 2006; Clifton et al. 2002).

In a modular and sequential perspective, after VWR, building structure takes place. For that reason, the reader has to pay attention to lexical information and to graphic cues to construct syntactic units and to establish boundaries between these units, i.e. to parse. We can expect that the reader needs to spend more time in setting a word that holds the periphery of a XP than when it occupies the position in its core; in the same way, we can expect that the reader will spend more time in setting a word that holds the periphery of a clause or sentence. The supposed time increase could be justified because there is a progressive accumulation of information to be processed: from the word itself, to its integration in a syntactic unit (XP), and to their incorporation into a discursive structure. This must be visible in eye movements and speech, by themselves and in combination.

Hypotheses

H1 Boundaries at syntactic and discursive positions, as *loci* for structure building and information integration, should be marked by visible prosodic markers and longer eye fixations, when compared with no boundaries.

H2 Discursive boundaries, as *loci* for wrap-up effects in the context of a complex text and under the influence of punctuation, must trigger higher fixation time (gaze and regressive fixations), and strong prosodic indicators of a intonational phrase boundary with significant decrease of F0 in declarative sentences.

H3 At a head phrase, as a no boundary position, we do not expect important variation of prosodic or eye movements variables comparing with boundary positions.

Experiment

Participants

17 European Portuguese native female speakers, students, proficient readers.

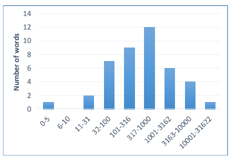
Experimental material for reading

Following previous studies on reading aloud and linguistic processing (Falé, Costa & Luegi, 2015, 2016), we assume that the high level of text complexity creates conditions leading to a less automatized behavior in reading for comprehension, triggering a greater reliance on structures, in a more likely bottom-up processing mode. So, in order to identify behavioral indicators of linguistic processing and information integration for reading aloud, we prepared a text with a specialized topic - *thermal-acoustic insulation* - adapted from an engineering journal.

A text of around 200 words was controlled considering phonological and lexical properties, namely: prevalence of words of 3 or more syllables, presence of words with complex syllable types (CVC, VC, CCV), low frequency words in the language (Graph 1). To control for these features we used FreP and CRPC, tools for computational analysis of EP databases.

Text Readability:

- Flesch-Kincaid Grade Level: 18.6;
- Flesch Reading Ease score: 3.1;
- Total of sentences: 9
- Average of words per sentence: 23
- Percentage of words with 3 or more syllables in text: 35%



Graph 1: Number of words in the text by frequency distribution.

We considered 3 experimental conditions with 7 occurrences each within the text:

- **Syntactic Boundary** (SB): the target word in the right periphery of an NP
- **Discursive Boundary** (DB): the target word in the right periphery of a sentence, followed by a punctuation mark
- **No Boundary** (NB): the target word is a head of a XP

A *resolução* [NB] deste problema, típico das actuais formas de vida urbana,

centra-se na existência de meios *técnicos* [SB] actuantes na oposição à propagação de *ruídos* [DB].

The resolution of this problem, typical of current forms of urban life,

focuses on the existence of technical resources actuating in the opposition the propagation of noise.

Procedure

Eye movements were recorded with a SMI IVIEW X™ HI-SPEED system, at a 1250Hz speed, and sound was recorded with a Logitech® Webcam Pro 9000.

For presentation, the text was divided in two blocks, font in size 22, Courier New, with two paragraphs spacing between rows, in a 17-inch screen.

Subjects were asked to read at their own pace trying to understand. After the reading aloud, participants answered a multiple-choice questionnaire, thus ensuring a reading comprehension task.

Independent variables

Position

- Syntactic boundary (SB)
- Discursive boundary (DB)
- No Boundary (NB)

Aims

- Identify processes of linguistic information integration undergoing in oral reading for understanding.
- Verify the effect of linguistic and discursive structure on speech and eye movements at specific *loci* in the text.
- Understand how prosody and eye movements can be related to reveal cognitive operations of structure building.

Dependent variables

Eye movements analysis

As sensitive measures to catch the processing of target structures, we selected 3 variables (Rayner et al. 2005):

- First fixation (FF) – average duration of the first fixation in a word; must reflect specific processes to visual word recognition, regardless the word context.
- First pass (FP) – which includes FF and other fixations before moving the eyes to right or left regions; could tap the processes involved in lexical access, required for their integration in a larger meaning or structural unit.
- Total time of word fixation (TTF) – including all fixations in a word; must reflect word integration in a semantic-discursive mental representation, and can reveal *wrap-up* effects.

Speech acoustic analysis

To identify prosodic boundaries we consider two acoustic parameters (Gussenhoven & Rietveld 1992):

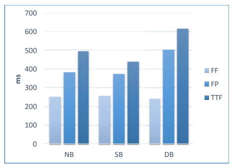
- Stressed vowel length (SVL) – as a marker indicating the proximity of a high level prosodic boundary: the longer the time vowel duration is, the higher the boundary is expected.
- Fundamental frequency of the stressed vowel (FO) – as an indicator of the syntactic position of the word: the more the word occupies the right periphery of a phrase or sentence, lower is F0 on the stressed vowel.

Results

Results for Syntactic and Discursive positions

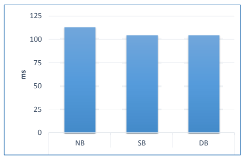
Eye Movements

- A Position effect in FP (F(2,16)= 7.729; p<0.001) is clear when comparing SN and DB (p=0.018), SB and DB (p=0.016), with longer reading times for DB. A Position effect is also registered in TTF (F(2,16)= 9.752; p<0.001), when comparing NB and DB (p=0.051) and SB and DB (p<0.001), with longer reading times for DB (Graph 2).
- The lack of Position effect in FF (F(2,16)=0.392; p=0.679) sustain the assumption that word targets in all positions are equivalent in terms of lexical properties.

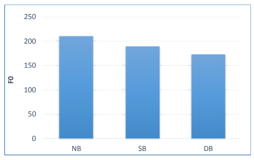


Graph 2: *Mean values (ms) for visual reading variables (FF, FP and TTF) by Position (NB, SB, DB).*

Reading Speech



Graph 3: *Mean values for "reading speech" variable SVL by Position (NB, SB, DB).*



Graph 4: *Mean values for "reading speech" variable F0 by Position (NB, SB, DB).*

- A Position effect in SVL (F(2,16)= 7.845; p=0.002) is evident when comparing NB and SB (p=0.002) and NB and DB (p=0.037), with lower values in DB.
- A Position effect in F0 (F(2,16)= 69.3; p>0.001) is evident when comparing NB and SB (p<0.001), with higher values in NB, and when comparing SB and DB (p<0.001), with higher values in SB.

Conclusions

Conjoined effects of eye movements and prosody are clear along the studied positions, however it is in Discursive Boundary that this effect is stronger:

- The increase of fixation time in both FP and TTF occurs simultaneously with a decrease of F0 values. This means that, at the *locus* where processes related with structural building end and prosodic phrasing occur, eyes take longer to complete the structure and possibly to resume prior information while speech indicates the sentence closure with lower F0.

Unexpectedly, the head of a XP, that we considered a no boundary position, seems to be in competition with a syntactic boundary, showing similar values in FP and even higher values in TTF.

- This result can be interpreted as a correlate of syntactic operations required by the projection of a bare syntactic category in a larger unit (X → XP).
- The higher duration of the stressed vowel in NB compared to boundary positions reinforce the hint that the nucleus of a syntactic phrase is time consuming for purpose of structure building.

Results reinforce the hypothesis that eyes and speech are working together closely unveiling indicators of the mental processes involved in structure building.

References

Ashby, J., Yang, J., Evans, K.H.C., Rayner, K. (2012). Eye movements and the perceptual span in silent and oral reading. *Atten Percept Psychophys*, 74: 634.

Benjamin, R., Schwanenflugel, P. 2010. Text complexity and oral reading prosody in young readers. *Reading Research Quarterly*, 45(4), 388-404.

Clifton, Jr., Ch., Carlson, K., Frazier, L. (2002). Informative Prosodic Boundaries. *Language and speech*, 45 (2), 87 – 114.

Falé, I., Costa, A., Luegi, P. (2015). Leitura em voz alta, movimentos oculares e prosódia. Integração de informação sintática e discursiva. In A. Moreno, F. Silva & J. Veloso (Eds.). *XXX Encontro Nacional da Associação Portuguesa de Linguística*. *Textos Selecionados*, Braga: APL, 261-275.

Falé, I., Costa, A., Luegi, P. (2016). Reading aloud: eye movements and prosody. *SP2016 Enhanced Proceedings*. (https://drive.google.com/file/d/0B5tX0NnTCf-ILVcxR18zYjI1MlE/view?pref=2&pli=1)

Frazier, L., Carlson, K., Clifton, C. Jr. (2006). Prosodic phrasing is central to language comprehension. *Trends in Cognitive Sciences*, 10(6), 244–249.

Gussenhoven, C., Rietveld, T. (1992). Intonation contours, prosodic structure, and preboundary lengthening. *Journal of Phonetics, vol. 20,* 283-303.

Hirotani, M., Frazier, L., Rayner, K. (2006). Punctuation and intonation effects on clause and sentence wrap-up: Evidence from eye movements. *Journal of Memory and Language, 54,* 425-443.

Inhoff, A., Solomon, M., Radach, R, Seymour, B. (2011). Temporal dynamics of the eye-voice span and eye movement control during oral reading. *Journal of Cognitive Psychology*, vol. 23, no. 5, 543-558.

Martins, F., M. Vigário, Frota, S. (2016). *FreP - Frequency in Portuguese*. Version V2016. Software in CD-ROM.

Multifunctional Computational Lexicon of Contemporary Portuguese, available at http://www.clul.ul.pt/en/resources/88-project-multifunctional-computational-lexicon-of-contemporary-portuguese

Perfetti, Ch. (1999). Comprehending written language: the blueprint of the reader. In Brown. C. & P. Hagoort (Eds.), *The neurocognition of language*. Oxford University Press, 167-208.

Rayner, K., Juhasz, B., Pollatsek, A. (2005). Eye movements during reading. In Snowling, M. & C. Hume (Eds.) *The Science of reading: a handbook*. Oxford: Blackwell Publishing Ltd, 79-97.

This work was financed by Fundação para a Ciência e a Tecnologia (UID/LIN/00214/2013 and SFRH/BPD/84138/2012)