

PROSODIC BOUNDARY INCONGRUENCES IN ORAL READING

LEITE, André Luiz de Faria^{1*}

LIMA, Aveliny Mantovan²

¹MSc in Linguistics, Postgraduate Program in Linguistics, University of Brasilia

²Professor of the Postgraduate Program in Linguistics and the Phonoaudiology Course, University of Brasilia – ORCID: <https://orcid.org/0000-0002-4305-4864>

Abstract: *Background: Characteristics of oral readings are well studied in school-aged children and teenagers, but not in educated adults. Objectives: Assess the prevalence of prosodic boundary incongruences in oral readings of adult, native, educated, Brazilian Portuguese speakers and analyze their correlations with specific linguist features. Design, settings, and participants: We studied an online video corpus of political speeches delivered by house members of the Brazilian parliament between 2017 and 2018, and their respective written texts. Measurements: We assessed a) prosodic boundary incongruences between oral readings and written texts, b) actor prototypicality of the subjects, c) thematic continuity of the sentences, and d) a variable called “sufficiency”, related to the concept of argumenthood, assorting each word according to its need for complementary words. The inter-rater reliability of the author’s perceptions of incongruences underwent Cohen’s Kappa test. Results: In 5 hours of oral readings, we found a median of 1.4 prosodic boundary incongruences per minute (interquartile range: 0.766 - 2.212). 80% of the incongruences were insertions of non-terminal or terminal boundaries. Prosodic boundary incongruency correlated positively with a) thematic continuity of the incongruent sentences (p-value = 0.0006345), b) the concept of “sufficiency” (p-value < 2.2e-16); and correlated negatively with c) first-person subjects (p-value = 0.0002584). Limitations: The assessment of the variables was subjective, and we did not control sentences for their lengths when analyzing variables “b” and “c”. Conclusions: Prosodic boundary incongruences were relatively common in our corpus. We introduced some hypotheses to explain the results.*

Keywords: Oral reading; Prosodic boundaries; Focus; Actor identification; Working memory.



^{1*} Corresponding author: andreleite68@gmail.com

1 Introduction

An old, recurrent metaphor represents writings as carcasses of speech and linguistic studies of written corpora as autopsies on oral language corpses (1). In case that is true, prosody would have probably been the first to perish.

Prosody, as stated in Barbosa and Raso (2), is "absent from writing in its acoustic manifestation, except for mere indications inferred from punctuation marks". In such manner, how can readers bring written texts back to life without making them look like dysprosodic zombies? Or can they? It applies to both silent and oral reading. In silent reading, readers move their eyes backward about ten percent of the time, and these backward saccades reveal some difficulties with the task of speech reconstitution (3). In oral reading (reading aloud), the struggle surfaces as incongruences between the written and the spoken text at the segmental (word) and the suprasegmental (prosodic) levels of speech.

Reading aloud is not equivalent to spontaneous speech. The end results of both processes may seem comparable, but their cognitive innings are not. The first noticeable difference is the material each task handles. Spontaneous speech handles "thoughts" (in a broad sense) while reading aloud handles graphic representations of oneself's or someone else's "thoughts". The graphic representations are not first-hand personifications of "thoughts," but their third-hand versions realized first as overt or inner speech and only then as a written text.

The second difference is that spontaneous speech predates reading by a long stretch of evolutionary time. Reading and writing are man-made artifacts; late cultural devices human brains are not hardwired to deal with. In order to learn how to read, human brains seem to "recycle" visual neurons in the left occipitotemporal region - "the brain's letterbox" - to recognize letters and words in both alphabetic and non-alphabetic (e.g., Chinese and Japanese characters) writing systems (4).

The third difference we chose to mention is that while spontaneous speech is closely knitted to the meanings of what is said, reading aloud may be disconnected from the semantics of the text. In reading, the pathways from the written input to the spoken output may, or may not, access the meanings of words, phrases, clauses, sentences, or discourse. When meanings are accessed, the access may be incomplete (only some words, or some clauses, and so on); and when meanings are not accessed, the cognitive route may go straight from orthographic to phonological units without even recruiting the lexicon. That has been called the dual-route model of reading (5).

Reading aloud performance has inspired some thinking and researching, and much of it is related to the acquisition and development of reading abilities in students (6, 7). Proficient readers, on the other hand, have received much less attention. Why is that? Is it because their oral readings are primarily good, fluent, and congruent, or just because they are less incongruent, less disfluent, and less bad? In any case, prosodic features, which are "absent from writing in its acoustic manifestation", and speech segmentation, which depends heavily on prosody (8), would be strong candidates for incongruences.

Speech, and oral reading, are not continuous flows of sound and information. Instead, prosodic boundaries segment them into noticeable units that may have syntactic, pragmatic, and cognitive purposes. Listeners can perceive and discriminate, with a high degree of inter-rater consistency, between two broad types of prosodic boundaries - terminal and non-terminal (9). The units that each boundary bounds can be interpreted at the discourse level as information modules (non-terminal boundaries) vs. utterances (terminal boundaries), corresponding, at the syntactic level, to phrases/clauses vs. sentences (8). As demonstrated in spoken corpora of many languages (10), utterances may be verbless in as much as one-third of their occurrences.

Although we do not expect to see as many (or any) verbless sentences in written corpora, they should not be proscribed as legitimate autonomous linguistic units in oral reading in the appropriate linguistic, pragmatic, and prosodic context.

If oral readers segment speech incongruently, they forge incongruent boundaries and units, and listeners with access to the written text should be able to identify the incongruences between the writing and the reading. The first question we ask is if listeners can reliably identify incongruent boundaries/units in oral reading. If they could, we would like to know why do readers produce those incongruent boundaries/units.

It is legitimate to guess that incongruent units result from some sort of language processing difficulty. For example, when a reader is reading a sentence, he may, at first, not know “what” or “whom” that sentence is about. As every sentence should ordinarily talk about “who is doing what”, that information constitutes relevant semantic knowledge that should translate into coherent units and prosodic boundaries. Bornkessel-Schlesewsky and Schlewewski (11) consider the "actor" in a sentence (the "who") to be a universal cardinal category "that provides an optimal and neurobiologically plausible solution to the demands of real-time information processing". According to their model, named "actor identification strategy", listeners/readers would search for prominence features associated with agenthood, identifying actors as prototypically human, animate, definite, first-person, nominative (in nominative-accusative languages), and positioned at the first argument position in sentences. Predicates - the "what" - would be inferred by exclusion.

Other source of semantic information that readers can access is the affiliation of the current sentence or segment with other sentences or segments in the discourse. Grosz and Sidner (12) proposed a well-known “computational theory of discourse structure” to analyze these affiliations, but it has some shortcomings when applied to oral reading. First of all, it includes every sentence and segment (current, previous, and forthcoming) in discourses, and readers, as a matter of fact, do not have access to forthcoming sentences or segments. Additionally, they suggested analyses that are too subjective and may yield biased and conflicting results between judges. A less subjective approach would be to consider the semantic affiliations of the current sentence with the sentence that came just before it only, as in the topic-focus (topic-comment, theme-rheme, etc.) approach (13). A sentence that talks about something already known could have a better chance of being appropriately segmented.

A third possibility is that oral readers construct prosodic boundaries responding to linguistic features that are local and narrowly focused. For example, since prosodic boundaries exist between words, the relation between adjacent words may have the upper hand when deciding to insert or not insert terminal or non-terminal boundaries. From this angle, any word in a sentence might call for a prosodic boundary or not, and the verdict for a boundary insertion would take into account its relations with previous words in the same sentence and, eventually, some additional information from the discourse context.

That brings us back to Izre'el's ideas (8) about speech segmentation. A read sentence would be segmented into information modules or, apart from its terminal boundary, would not be segmented at all. At each word, the oral reader would decide if he inserts a prosodic boundary. If he judges the current word as the last one in an information module, he could insert a boundary. Otherwise, he would not do it. Additionally, if he thinks the word is not the last one in an information module, he presumes something is missing in the current speech segment. When we speculate that the oral reader assumes the current word needs other words (or a single word) to fulfill its semantic or syntactic needs, our thinking becomes tangential to the concept of arguments vs. adjuncts.

According to Haspelmath (14), a verbal argument may be defined as "a phrase whose occurrence is made possible by a specific verb, and which therefore cannot occur with a generic verb". The author proposes a method to identify verbal arguments:

- a. I wrote a letter. > *I wrote, and I did a letter.
- b. I wrote with a pen. > I wrote, and I did it with a pen.

The sentence in (a) shows that "letter" cannot be moved into a neighboring clause with an anaphoric verb because it is an argument of the verb "to write". On the other hand, sentence (b) shows that "pen", an adjunct, can be freely moved away from the verb. This so-called "argument-adjunct dichotomy" has also been applied to nouns, adjectives, and even prepositions (15), but it may not be as unambiguous as it seems (16). As a matter of fact, many researchers have "abstracted away from this distinction, because identifying arguments and adjuncts is a notoriously difficult task, taxing many native speakers' intuitions" (17).

Furthermore, when applied to oral reading, the approach suggested by Haspelmath (14) suffers from the same ailments we pinpointed above in Grosz and Sidner's (12): it includes segments of the speech that are yet to come to the readers' eyes. For example, if "with a pen" in sentence (b) is an adjunct, it could be wrongly omitted when reading the sentence. The reader could be led to think that the sentence ended after the verb and insert a terminal boundary, which would be incongruous with the written text. In sentence (a), "a letter" is an argument, the verb "calls for it", but its fate in reading may be the same as any adjunct. In both situations, when the reader gets to "wrote", he has precisely the same information. To that end, all that matters to the reader is if the word he is reading at any moment needs additional word(s) to complete the current segment, regardless of the arbitrary category - information module or utterance, argument or adjunct - linguists attributed to them.

To better understand how oral readers segment their readings and why they do so, we established as our primary objectives to 1) assess the prevalence of prosodic boundary incongruences between written texts and their respective oral readings and 2) investigate the correlations between the incidence of prosodic boundary incongruences in oral reading and the following elements of the written texts: a) the prototypicality of the actor; b) the thematic continuity from one sentence to another; c) the need for other words to syntactically, semantically, or pragmatically complete the current speech segment. We hypothesize that sentences that a) are thematically continuous with the previous sentence or b) bear prototypical actors have a better chance to be prosodically congruent with the written text. Additionally, we conjecture that if the word or segment under reading does not need other words to be syntactically, semantically, or pragmatically fulfilled, the reader will have a bias to insert a prosodic boundary and incongruently segment his speech.

2 Methods

We performed a cross-sectional study on a corpus of oral readings, identifying prosodic boundary incongruences and some pre-determined linguistic features in incongruent and congruent sentences. Then, we investigated the statistical correlations between prosodic congruency and linguistic features.

2.1 Corpus and participants

The corpus consisted of a public online database of video recordings of speeches given on the Brazilian Senate floor from 2017 to 2018 by native Brazilian Portuguese (BP) speakers (18). The speakers are Brazilian senators, and the speeches are political. The files had MPEG-4

format with a mean bitrate of 500 kbps. Each speech was delivered by one individual speaker at one particular moment of a specific day. Most speeches consisted of oral read sentences and off-the-cuff non-read sentences, so we extracted speeches with at least one sentence that was read. Then, we selected one speech per speaker, choosing the speech with the largest number read sentences. After these steps, we ended up with a corpus of 39 speeches, delivered by 39 different speakers, with at least one oral read sentence per speech.

The speeches, as mentioned, are political, and have some specificities. They may talk about many issues but typically employ persuasive elements. Prosodically, they may use higher and more variable pitches (fundamental frequency) and tend to highlight emotions more often than not (19).

It must be noted that the written texts were drafted by professional speechwriters and specifically targeted to oral delivery (reading aloud). They were accessible to the study but are not available to the general public due to ethical reasons and institutional policies. Nonetheless, speech transcriptions by professional stenographers are publicly available under the label "notas taquigraficas" (18). As the transcriptions avoid reproducing reading errors that are evident to the stenographers, they happen to be close replicas of the written texts.

2.2 Variables

Along with demographic statistics like age, sex, educational attainment, and birthplace, we collected the following variables: a) incongruent prosodic boundaries (oral reading differs from written text), b) actor prototypicality, c) thematic continuity, and d) need for other words to syntactically, semantically, or pragmatically complete the current speech segment. Similar to a typical cross-sectional study, we could say that the presence of variable "a" determined the "cases", and its absence determined the "non-cases". Correspondingly, the presence of variables "b", "c", and "d" defined the "exposed," and their absence determined the "not exposed".

The exposition to the linguist features of variables "b" and "c" applied to the entire sentences where the variables appeared. In these situations, we had "exposed" and "not exposed" sentences that would have an incongruent prosodic boundary (variable "a") or not. On the other hand, the exposition to the linguistic features of variable "d" applied both to the sentences and to the exact word transitions where they appeared. Altogether, variables "b" and "c" looked at associations with incongruencies at the sentence level, and variable "d" also looked at associations at a local-word level. That is, with variable "d" there were also "exposed" and "not exposed" word transitions, in addition to "exposed" and "not exposed" sentences.

Sentences with incongruent prosodic boundaries had their inter-rater reliability between authors and a group of native BP speakers validated by Cohen's Kappa test.

2.3 Assessment of the variables.

We assessed each variable as follows.

2.3.1 Incongruent prosodic boundaries

The delimitation of sentences observed the punctuation (periods) defined by the written text. One of the authors read each sentence and then watched the video of its oral reading. Incongruences in speech segmentation between the oral readers and ourselves were annotated. Our measures for incongruences were very tolerant. We admitted as congruent all prosodic boundaries (or their absences) that could eventually be considered an acceptable speech segmentation by the intuitions of a native BP speaker. We only treated a boundary as incongruent when the speaker read the sentence without a "phrasing that was consistent with the author's syntax" (6), i.e., when the speech segmentation of the sentence as a whole, meaning the combination of its prosodic boundaries, was inconsistent with the proper delivery of the message intended by the written sentence. Formal syntactic criteria, per se, were not applied when evaluating prosodic boundary congruence.

Additionally, we categorized the boundaries as terminal or non-terminal. A congruent terminal boundary could only be inserted at the end of a written sentence and a non-terminal one at any point where, according to the indulgent definition we adopted, it would fit. Hence, four types of prosodic boundary incongruences were annotated: insertions of prosodic boundaries (1. terminal, 2. non-terminal) and deletions of prosodic boundaries (3. terminal, 4. non-terminal). Some annotated samples extracted from the corpus are provided in the Appendix, including audio files and English translations.

2.3.2 Actor prototypicality

We analyzed the syntactic subject of each sentence's clause along the lines followed by the actor identification strategy model (20). The classification included the following binary features: 1) person (first vs. other), 2) human (yes vs. no), 3) animacy (animate vs. inanimate), 4) position (before vs. after the verb), and 5) definiteness (definite vs. indefinite). We added the following features to the classification: 6) voice (active vs. passive), and 7) subject drop (yes vs. no).

About item 7, BP, in contrast with English or French, may not realize the subject. When the subject is not overtly present (subject drop), it can be figured out based on pragmatic or grammatical elements (i.e., agreement on the verb).

2.3.3 Thematic continuity

We labeled each sentence as thematic continuous or thematic discontinuous. Thematic continuity meant the current sentence topic (theme) had been a topic or comment (focus, rheme) in the previous sentence. In multi-clause sentences, we analyzed the main clause or the first coordinate clause. Additionally, we considered a first-person subject consistent with thematic continuity since the first person is always known and positioned at the center of any discourse.

2.3.4 The need for other words to syntactically, semantically, or pragmatically complete the current speech segment

The approach here was less orthodox and will be arbitrarily named "sufficiency". We tried to put ourselves in the readers' shoes, simulating an extreme situation where the reader would be completely blind to whatever would come after the word he was reading at any specific point in

time. At that moment, with only the syntactic, semantic, and pragmatic information gathered so far in the current speech and sentence, could that word be assumed to be the last word in an utterance or information module? For instance, in "I wrote a letter", "letter" could be the last word in the sentence/utterance, and the reader could be led to insert a terminal prosodic boundary. However, the sentence/utterance went on: "I wrote a letter with a pen". Again, it would be possible to think that "pen" was the last word in the sentence/utterance, and a terminal prosodic boundary could be inserted. But, again, the sentence could be much longer, offering many new opportunities for prosodic boundary incongruences: "I wrote a letter with a pen my father gave me as a birthday present last year when my mother came back home from abroad." In this case, we would say that, sequentially, the words "letter", "pen", "me", "present", "year", "home", and maybe "back" are "sufficient": they do not need other words to syntactically, semantically, or pragmatically complete the current speech segment.

We must not forget that our study deals with Brazilian Portuguese, which has some grammatical features of its own and, as such, behaves differently than English. For instance, adjectives in BP usually follow nouns, and that shapes opportunities for incongruent insertion of boundaries after nouns, as in "comprei um carro velho" (I bought an old car), where "velho" is the adjective (old) qualifying the noun "carro" (car). The reader may think the sentence goes as far as "carro", as in "I bought a car that is old", and insert an incongruent prosodic boundary after it. Nevertheless, we provide additional examples of putative sentences in English. 1) "Without spoken words, facial expression and gesture must carry the meaning." In this sentence, "facial expression and gesture" is the subject of the clause. However, the reader may think they are coordinated with "spoken words", as in "Without spoken words, facial expressions and gesture, [something else] must carry the meaning", and insert a non-terminal boundary after "gesture". 2) "To begin with, remember what a word is: a long-term memory linking of pieces of phonological, syntactic, and conceptual structures." The reader may read "memory" and think that "a long-term memory" is "what a word is", as in "a word is a long-term memory". He could insert a non-terminal or terminal prosodic boundary after memory, but that would be incongruent because "a word is: a long-term memory linking". 3) "Phrasing (also referred to as grouping) is associated with the segmentation of utterances into variable prosodic units and prosodic theory and phonological studies refer to several prosodic categories and units ranging from syllable to utterance." (21). Here, "prosodic theory" is part of the subject of the verb "refer" but the reader may think it is coordinated with "variable prosodic units", as in "the segmentation of utterances into variable prosodic units and prosodic theory". In this case, he could incongruently delete a non-terminal prosodic boundary after "prosodic units" and insert a non-terminal (or even a terminal) one after "prosodic theory". In a way, this reasoning could also apply to the so-called garden path sentences, as in the illustrious: "While Mary bathed the baby played in the crib." Here, the reader could insert an incongruent non-terminal boundary after "baby".

Taking all that into account, we labeled each word of every sentence according to its "sufficiency", as in the examples below, where "/" indicates "sufficiency" (of the previous word):

- a) *I ate/ my soup/ with a spoon/ my father gave me/ as a birthday present/ last year/ when my mother returned home/ from abroad.*
- b) *To begin with/, remember what a word is/: a long-term memory/ linking/ of pieces/ of phonological, syntactic, and conceptual structures.*

Sentence (a) labeling is straightforward, as every "/" marks a possible end to the utterance, but sentence (b) needs additional explanation. For example: "with" in "to begin with"

was considered “sufficient” because "to begin with" is a usual phrase, a fixed expression, and the reader does not need to look for other words to complement the segment (an "information module"), as he would need with "to" and "begin". The other marked words - "is", "memory", "linking", and "pieces" - follow the same pattern of the sentence (a) and could be the last word of an utterance.

Additionally, as can be seen (*in italics*) in sentences (a) and (b), we identified in every sentence the first segment that could be considered an utterance - the smallest linguistic unit with pragmatic autonomy and interpretability in isolation, “the counterpart to a speech act”, “the primary reference unit for the analysis of speech”, akin to the Language into Act Theory (22). For instance, "*I ate*" in sentence (a), and "*Remember what a word is*" in sentence (b) are the smallest (and first to appear) linguistic units that have interpretability in isolation and pragmatic autonomy. Thus, readers could eventually (and erroneously) interpret them as utterances. Bethink that this procedure is an adaptation of concepts originally applied to spoken speech; but, as we are dealing with written sentences destined to be read and acquire prosodic features, we take the liberty and run the risk of expanding its conventional applicability.

Summing up, we categorized each word along two axes: a) sufficiency (yes vs. no) and b) autonomy (belongs to the first segment in the sentence with pragmatic autonomy and interpretability in isolation: yes vs. no)

2.4 Validation of the variables

We compared the author's perception of incongruent prosodic boundaries with four other BP speakers' perceptions and assessed the inter-rater reliability with Cohen's Kappa test (23) performed in the software RStudio (24). Since participants had no formal linguistic education, they were shown, as a preparatory step, four sentences (from the corpus) with all types of incongruent prosodic boundaries. The sentences were annotated to represent the authors' specific perceptions of the incongruences. The annotations included only insertions ("/") and deletions ("*") of prosodic boundaries, regardless of the terminality or non-terminality of the boundary.

The protocol each participant individually followed at the preparatory phase was: 1) read the sentence, 2) listen to the audio of the actual oral reading of the sentence, 3) read the sentence with annotations showing the authors' perceptions of prosodic boundary incongruences, 4) repeat any of the steps at will, if needed.

After the preliminary phase, each participant received a set of audio files with the same 40 pairs of sentences from the corpus and their respective written texts, each pair containing one sentence with and one without prosodic boundary incongruence(s). The procedure they had to follow was: 1) read the sentence, 2) listen to the audio of the oral reading, 3) if one of the sentences of the pair is incongruent, mark it. They were asked to apply the procedure to at least ten pairs of sentences. The findings of the test are described in the results section.

2.5 Study size and sources of bias

Awareness of the value of the variable that defined non-cases and cases (prosodic boundary incongruence) could distort the assessment of exposition variables (variables b, c, and d). Thus, we tried to avoid information bias by evaluating the exposition variables without information about the value of the case-defining variable.

We analyzed many oral read sentences extracted from a database of speeches delivered by a few dozen speakers in a frame of time. There was a potential selection bias towards speakers that gave more speeches during that time. We avoided the selection bias by choosing only one speech per speaker. Even so, there was also a selection bias towards speakers that read

more material during their speeches. In order to have enough sentences to analyze, we did not try to eliminate that last bias.

We did not perform formal, a priori, sample size calculations to determine a proper size for our sample. However, we had an a priori estimate of the incidence of prosodic boundary incongruences and of the number of speeches speakers used to deliver. Therefore, we believed a two-year timeframe of the database would yield enough read sentences to analyze.

2.6 Statistical methods

We used descriptive statistics to summarize information about demographics and corpus characteristics. Then, as we deal with associations between categorical nominal variables, we applied chi-squared tests. Finally, when analyzing actor prototypicality, we tested the variables as a group (all characteristics lumped together) and, in order to discriminate between joint effect and individual effects, we tested each relevant subgroup.

3. Results

3.1 The speakers

From a universe of more than 81 potential speakers, we selected 39. The criteria for the selection were a) the speaker had read aloud any sentence of his speech, and b) the original written text upon which the oral reading was based was available to the authors. Table 2 shows the speakers' demographic characteristics. Note that “ages” reflect the moment speeches were delivered.

3.2 The corpus

The 39 speeches amounted to 8h28min of video recordings (individuals speaking with and without reading) with 5 hours of oral readings. The median oral reading time was 473 seconds (7.88 minutes). Table 6 shows reading times per speaker.

3.3 The inter-rater reliability

The Cohen's Kappa test included four participants, with the following demographic characteristics:

- 1) 22-year-old female Biotechnology undergraduate student
- 2) 20-year-old male Materials Engineering undergraduate student
- 3) 18-year-old male Computer Sciences undergraduate student
- 4) 16-year-old male High School student

Table 1 shows the Cohen's Kappa results and the number of pairs of sentences each participant evaluated. The Cohen's Kappa numerical results are classified according to Landis and Koch agreement categories (25).

Table 1: Cohen's Kappa Test: interrater reliability of PBI at the sentence level.

| Participant | Kappa | Agreement | Confidence intervals | Sentence pairs analyzed (n) |
|-------------|-------|-------------|----------------------|-----------------------------|
| 1 | 0,69 | substantial | 0,42-0,95 | 25 |
| 2 | 0,61 | substantial | 0,12-1,00 | 11 |
| 3 | 0,44 | moderate | 0,03-0,86 | 18 |
| 4 | 0,27 | fair | (-)0,19-0,73 | 13 |

Table 2: Demographic characteristics of speakers.

| Speaker | Age | Gender | *POB | **Educational attainment |
|---------|-----|--------|------|----------------------------|
| 1 | 74 | M | SP | College degree |
| 2 | 69 | M | PE | College degree |
| 3 | 75 | M | AL | College degree |
| 4 | 54 | M | PB | College degree |
| 5 | 50 | M | PI | College degree |
| 6 | 61 | M | SC | College degree |
| 7 | 54 | M | SE | College degree |
| 8 | 56 | M | PA | College degree |
| 9 | 54 | M | SP | College degree |
| 10 | 75 | M | CE | College degree |
| 11 | 65 | M | CE | College degree |
| 12 | 71 | M | RJ | College degree |
| 13 | 70 | M | RN | College degree |
| 14 | 50 | M | AL | College degree |
| 15 | 39 | M | AC | College degree |
| 16 | 58 | M | GO | College degree |
| 17 | 64 | F | GO | College degree |
| 18 | 81 | M | MA | College degree |
| 19 | 85 | M | PB | College degree |
| 20 | 48 | M | RN | College degree |
| 21 | 75 | M | RS | College degree |
| 22 | 61 | F | BA | College degree |
| 23 | 72 | F | GO | College degree |
| 24 | 43 | F | MS | College degree |
| 25 | 73 | F | SP | College degree |
| 26 | 59 | M | SP | College degree |
| 27 | 68 | M | RS | Less than secondary school |
| 28 | 67 | M | PA | Secondary school |
| 29 | 78 | M | MS | College degree |
| 30 | 75 | M | PB | College degree |
| 31 | 52 | M | RJ | College degree |
| 32 | 70 | F | MG | College degree |
| 33 | 55 | M | RR | College degree |
| 34 | 76 | M | MG | College degree |
| 35 | 60 | M | RR | College degree |
| 36 | 62 | M | SC | College degree |
| 37 | 56 | F | SC | College degree |
| 38 | 60 | M | MT | College degree |
| 39 | 49 | M | GO | College degree |
| Mean | 63 | .. | .. | .. |
| SD | 11 | .. | .. | .. |

*POB: place of birth by state; **Postgrad studies not considered

3.4 The prosodic boundary incongruences

Table 6 shows the prosodic boundary incongruences we identified, distributed by speaker. Speakers produced a median of 1.4 prosodic boundary incongruences (PBI) per minute across speeches. Most incongruences (80%) were prosodic boundary insertions, notably non-terminal boundaries (54% of all incongruences, 67% of boundary insertions). Figure 1 shows a histogram of the number of speakers in each category of PBI per minute, and Figure 2 depicts the median and interquartile range of PBI per minute in the corpus.

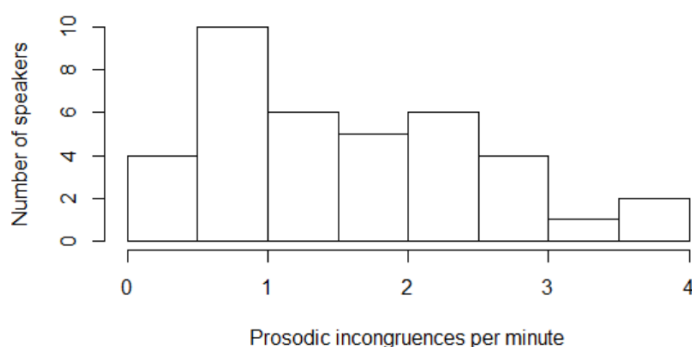


Figure 1: Histogram: frequency of prosodic boundary incongruences per minute per speaker

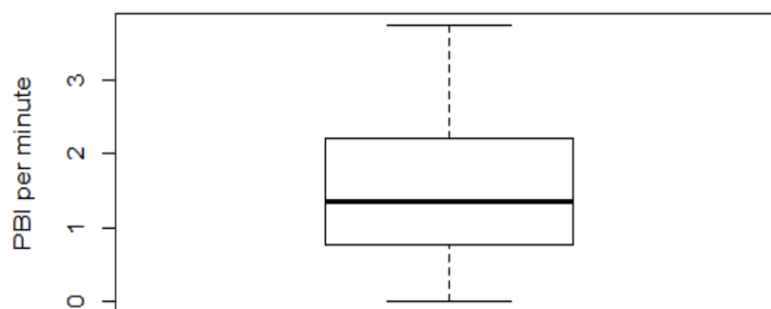


Figure 2: Boxplot: prosodic boundary incongruences per minute (median 1.4, interquartile range 0.766 - 2.212)

Another way to look at the data is to acknowledge the speech rate (in words per minute) and the PBI incidence per spoken word. The basal speech rate for each speaker was measured on stretches of fluent oral readings to avoid unwanted effects of incongruent prosodic boundaries and other dysfluencies on the calculations. The mean speech rate was 119.7 words per minute (interquartile range 109.5 - 127), and the incidence of PBI per spoken word was 1.39 PBI per 100 words (interquartile range 0.7-1.9). Pearson's product-moment correlation test showed no correlation between basal speech rate and incidence of PBI per word ($t=0.051908$, $p\text{-value}=0.9589$).

3.5 Prosodic boundary incongruences vs. linguistic features

3.5.1 Thematic continuity

Table 3 shows the number of incongruent and congruent sentences in each category of thematic continuity: continuous and discontinuous. There were a disproportionately high number of thematically continuous incongruent sentences. Pearson's Chi-squared test with Yates' continuity correction suggests an association between thematic continuity and prosodic boundary incongruence (p-value = 0.0006345).

Table 3: Thematic continuity in congruent and incongruent sentences.

| Thematic continuity | Sentences (n(%)) | | Total | PBI (events) |
|---------------------|------------------|-----------|-------|--------------|
| | Incongruent | Congruent | | |
| Yes | 223(73%) | 320(56%) | 543 | 264 |
| No | 81(27%) | 252(44%) | 333 | 184 |
| Total(%) | 304(100%) | 572(100%) | 876 | 448 |

Table 4 shows the number each type of prosodic boundary incongruence (PBI) in each category of thematic continuity. There seemed to be a disproportionately higher incidence of insertion of non-terminal boundaries (INT) associated with thematic continuity, but Pearson's Chi-squared test did not confirm it (p-value = 0.1426).

Table 4: Thematic continuity of the sentences of each PBI subtype

| Thematic continuity | PBI subtypes (number of events) | | | |
|---------------------|---------------------------------|----|-----|----|
| | INT | IT | DNT | DT |
| Yes | 151 | 62 | 38 | 13 |
| No | 89 | 55 | 32 | 8 |

Prosodic boundaries - INT: insertion non-terminal; IT: insertion terminal; DNT: deletion non-terminal; DT: deletion terminal

3.5.2 Actor prototypicality

We identified 38 different types of "actors" from the combination of the prototypicality features we choose to analyze. Table 5 shows the distribution of the most frequent types, bundling all the features in a six-position string of characters - XXXXXX -, corresponding, sequentially, to the following features: 1) Voice: active (A) vs. passive (P) vs. or subject drop (D); 2) Position: before (1) vs. after (0) the verb; 3) Person: first (1) vs. other (0); 4) Human: yes (1) vs. no (0); 5) Animate: yes (1) vs. no (0); 6) Definite: yes (1) vs. no (0). The last column of Table 5 (actor congruency ratio) shows the ratio of the counting of each type of actor in each type of sentence. Negative ratios indicate there were more in incongruent sentences with that type of actor.

Table 5: Distribution of types of actors in prosodic congruent and incongruent sentences

| Prototypicality features | Sentence congruency | | Actor congruency ratio |
|--------------------------|---------------------|-------------|------------------------|
| | Congruent | Incongruent | |
| A10001 | 326(46%) | 198(40%) | 1.65 |
| D11110 | 93(13%) | 48(10%) | 1.94 |
| D11111 | 65(9%) | 24(5%) | 2.71 |
| D10001 | 30(4%) | 33(7%) | -1.10 |
| A10000 | 27(4%) | 31(6%) | -1.15 |
| A00001 | 23(3%) | 16(3%) | 1.44 |
| A11111 | 22(3%) | 5(1%) | 4.40 |
| A10110 | 21(3%) | 17(3%) | 1.24 |
| A10111 | 20(3%) | 31(6%) | -1.55 |
| P10001 | 19(3%) | 17(3%) | 1.12 |
| A00000 | 10(1%) | 7(1%) | 1.43 |

Table 6: Oral readings and prosodic boundary incongruences (PBI)

| Speaker | *PBI per type | | | | PBI Total | Reading time (sec) | PBI per min |
|----------|---------------|----------|----------|--------|--------------|--------------------------|-------------------|
| | Insertion | | Deletion | | | | |
| | INT | IT | DNT | DT | | | |
| 3 | 19 | 5 | 7 | 0 | 31 | 496 | 3.8 |
| 5 | 4 | 5 | 4 | 0 | 13 | 210 | 3.7 |
| 39 | 7 | 4 | 0 | 0 | 11 | 204 | 3.2 |
| 33 | 12 | 3 | 1 | 0 | 16 | 339 | 2.8 |
| 19 | 11 | 3 | 2 | 0 | 16 | 340 | 2.8 |
| 17 | 5 | 6 | 5 | 1 | 17 | 378 | 2.7 |
| 16 | 8 | 5 | 1 | 4 | 18 | 430 | 2.5 |
| 29 | 16 | 7 | 1 | 2 | 26 | 678 | 2.3 |
| 26 | 9 | 3 | 2 | 1 | 15 | 396 | 2.3 |
| 8 | 9 | 9 | 6 | 0 | 24 | 644 | 2.2 |
| 6 | 10 | 3 | 2 | 5 | 20 | 549 | 2.2 |
| 36 | 10 | 0 | 7 | 1 | 18 | 495 | 2.2 |
| 18 | 11 | 3 | 4 | 0 | 18 | 502 | 2.2 |
| 27 | 23 | 5 | 8 | 1 | 37 | 1171 | 1.9 |
| 35 | 12 | 3 | 0 | 0 | 15 | 494 | 1.8 |
| 24 | 2 | 2 | 0 | 0 | 4 | 133 | 1.8 |
| 21 | 13 | 4 | 3 | 0 | 20 | 743 | 1.6 |
| 7 | 9 | 1 | 4 | 0 | 14 | 551 | 1.5 |
| 15 | 4 | 4 | 2 | 0 | 10 | 415 | 1.4 |
| 14 | 5 | 3 | 2 | 3 | 13 | 573 | 1.4 |
| 10 | 6 | 8 | 2 | 0 | 16 | 722 | 1.3 |
| 2 | 2 | 1 | 0 | 0 | 3 | 146 | 1.2 |
| 38 | 2 | 6 | 1 | 0 | 9 | 473 | 1.1 |
| 22 | 0 | 1 | 0 | 0 | 1 | 57 | 1.0 |
| 37 | 1 | 2 | 0 | 0 | 3 | 187 | 1.0 |
| 31 | 1 | 0 | 1 | 1 | 3 | 190 | 0.9 |
| 25 | 6 | 1 | 1 | 0 | 8 | 517 | 0.9 |
| 13 | 5 | 0 | 1 | 0 | 6 | 412 | 0.9 |
| 11 | 2 | 3 | 0 | 1 | 6 | 459 | 0.8 |
| 28 | 4 | 1 | 0 | 1 | 6 | 482 | 0.7 |
| 34 | 0 | 5 | 0 | 0 | 5 | 434 | 0.7 |
| 32 | 5 | 1 | 0 | 0 | 6 | 533 | 0.7 |
| 30 | 2 | 2 | 1 | 0 | 5 | 521 | 0.6 |
| 12 | 1 | 6 | 0 | 0 | 7 | 831 | 0.5 |
| 20 | 2 | 1 | 0 | 0 | 3 | 416 | 0.4 |
| 1 | 1 | 1 | 1 | 0 | 3 | 934 | 0.2 |
| 9 | 1 | 0 | 0 | 0 | 1 | 366 | 0.2 |
| 23 | 0 | 0 | 1 | 0 | 1 | 528 | 0.1 |
| 4 | 0 | 0 | 0 | 0 | 0 | 18 | 0.0 |
| Median | | | | | | 473 | 1.4 |
| Subtotal | 240(54%) | 117(26%) | 70(16%) | 21(5%) | 448(100%) | | |
| Total | 357(80%) | | 91(20%) | | 448(100%) | | |

*PBI: prosodic boundary incongruence
INT: insertion non-terminal; IT: insertion terminal; DNT: deletion non-terminal; DT: deletion terminal

The most frequent type of actor was "A10001": active voice, pre-verbal, non-first person, non-human, inanimate and definite. Subject-drop actors occupy the next three positions. Passive voice was relatively rare, representing only 3% of actors in any kind of sentence.

Pearson's Chi-squared test with simulated p-value applied to the numbers of actors in congruent and incongruent sentences in Table 5 yields a p-value of 0.0002332, suggesting an association between actor prototypicality and congruency. The actor congruency ratio (last column of the Table 5) indicates the most congruent type of actor was A11111: active voice, pre-verbal, first-person, human, animate, and definite. Conversely, the most incongruent actor was A10111, which differs from A11111 only at the first-person feature (it is non-first person).

In order to pinpoint the differences in congruency ratios between actors, we reclassified the actors into subgroups. Table 7 shows the first reclassification, considering only the voice and subject drop features. Pearson's Chi-squared test suggests no association between those features and congruency (p-value = 0.5115).

Table 7: Sentence congruency vs. active voice, passive voice and subject drop

| Prototypicality feature | Sentence congruency | |
|-------------------------|---------------------|-------------|
| | Congruent | Incongruent |
| Active voice | 468(66%) | 334(67%) |
| Passive voice | 44(6%) | 37(7%) |
| Subject drop | 198(28%) | 127(26%) |

The next reclassification segregated the results of each of the other features. Table 8 shows them as a stack of contingency tables.

Table 8: Sentence congruency vs. other prototypicality features

| Prototypicality feature | | Sentence congruency | |
|-------------------------|---|---------------------|-------------|
| | | Congruent | Incongruent |
| Position | 0 | 59 | 48 |
| | 1 | 651 | 451 |
| Person | 0 | 528 | 416 |
| | 1 | 182 | 83 |
| Human | 0 | 472 | 342 |
| | 1 | 238 | 157 |
| Animate | 0 | 472 | 342 |
| | 1 | 238 | 157 |
| Definite | 0 | 172 | 136 |
| | 1 | 538 | 363 |

We see, at first, that the results for “human” and “animate” are identical. It brings to light that our corpus didn't have animals or other credible non-human animated entities as actors. Next, we applied the Pearson's Chi-squared test with Yates' continuity correction to each contingency table, obtaining the following p-values: position (0.4925), first person (0.0002584), human (0.4909), and definite (0.2614). These results indicate there is an association between first-person subjects and prosodic boundary congruency. Although we did not analyze interactions between variables, the association we found between A11111 subjects and congruency may stem from the first-person feature alone.

The first-person feature had an additional characteristic that we must consider (and will discuss later): only 12% of them were overtly expressed. The other 88% were subject-dropped.

3.5.3 Sufficiency

By “sufficiency”, we mean the need for other words to syntactically, semantically, or pragmatically complete the current speech segment. Words that did not need other words were “sufficient”. We analyzed every word of the written texts to check if they were sufficient or not (in the context they appeared). Table 9 shows the number of words in each category of sufficiency (yes/no) in each type of sentence (with or without PBI). Pearson's Chi-squared test with Yates' continuity correction suggests an association between sufficiency and incongruency (p-value = 0.001381), which means that incongruent sentences had comparatively more sufficient words than congruent sentences.

Table 9: Sentence congruency vs. word sufficiency

| Sufficiency | Sentence congruency | | Total |
|-------------|---------------------|-------------|-------|
| | Congruent | Incongruent | |
| No | 6946 | 5573 | 12519 |
| Yes | 4209 | 3704 | 7913 |
| Total | 11155 | 9277 | 20432 |

Next, we analyzed only incongruent sentences, labeling every word as 1) PBI (yes/no) and 2) sufficiency (yes/no). Note that PBI here does not refer to sentences, but to each specific PBI we found in the readings. Table 10 shows the results, and Pearson's Chi-squared test with Yates' continuity correction suggests a strong association between word sufficiency and PBI (p-value < 2.2e-16).

Table 10: Prosodic boundary incongruence vs. word sufficiency

| Sufficiency | Prosodic boundary incongruence | |
|-------------|--------------------------------|-----|
| | No | Yes |
| No | 5510 | 30 |
| Yes | 3359 | 418 |

Table 10 included all types of PBI and maybe we should have excluded PBIs of the type “deletion” from the analysis. Prosodic boundaries can only be deleted where prosodic boundaries should exist, and they mustn't exist after words that need other words to complete their speech segments. Indeed, we found only one event of PBI of the type “deletion” after a word that was not sufficient. For this reason, we excluded PBI deletions from the counting and represented the results in Table 11. Pearson's Chi-squared test with Yates' continuity correction keeps sustaining a p-value < 2.2e-16, confirming the association between sufficiency and prosodic boundary incongruence of the type “insertion”.

Table 11: Prosodic boundary incongruence (only insertions) vs. word sufficiency

| Sufficiency | Prosodic boundary incongruence | |
|-------------|--------------------------------|-----|
| | No | Yes |
| No | 5510 | 29 |
| Yes | 3359 | 328 |

At last, Table 12 shows the association between PBI and the “autonomy” of the segment where the word is. Considering only words without “sufficiency” (words that need others), a word has “autonomy” if it belongs to a segment that can be interpreted as an utterance.

Table 12: Prosodic boundary incongruence vs. word autonomy

| Autonomy | PBI | |
|----------|-----|------|
| | Yes | No |
| Yes | 12 | 1412 |
| No | 18 | 3843 |

As we have already seen, there were not many PBI associated with words without sufficiency. From Table 12, we see that, for those words, “autonomy” also does not correlate with congruency: it did not matter if the word belonged to a potential autonomous utterance or a non-autonomous information module (Pearson's Chi-squared test with Yates' continuity correction, p -value = 0.1585).

4 Discussion

We aimed to assess the prevalence of prosodic boundary incongruences in oral readings and to investigate their associations with some linguistic features of the written texts.

The material we analyzed amounted to 5 hours of political oral readings, delivered by 39 native BP speakers from all regions of Brazil, most of them males, in their fifties or sixties, with at least a college degree. Results are briefly summarized below.

4.1 Key results

4.1.1 Prevalence of prosodic boundary incongruences

Prosodic boundary incongruences (PBI) were relatively common, arising more than once per minute of reading, with an interquartile range of 0.766 - 2.212, and only one speaker (who read only one sentence) performing a PBI-free reading. Incongruent prosodic boundary insertions accounted for 80% of all PBIs. Inter-rater reliability of prosodic incongruence of sentences measured by Cohen's Kappa Test with four participants showed an agreement with the author's judgments between fair and substantial.

4.1.2 Prosodic boundary incongruences vs. thematic continuity

We hypothesized that thematic-continuous sentences would have a better chance to be prosodically congruent with the written texts. Surprisingly, we saw the opposite: thematic-continuous sentences had more PBI than thematic-discontinuous sentences. In other words, familiarity correlated with incongruency.

4.1.2 Prosodic boundary incongruence vs. actor prototypicality

Our hypothesis was that prototypical actors would be associated with prosodic congruency, and, indeed, we found that active voice, pre-verbal, first-person, human, definite actors were more prevalent in congruent sentences. However, analyzing each of those features individually, we found that only the first-person feature had a statistically significant association with congruency.

4.1.3 Actor prototypicality vs. thematic continuity

Our thematic continuity assessment included first-person actors as a criterium for continuity. Since first-person actors were independently associated with congruency, as seen in the assessment of actor prototypicality, the first-person feature of thematic continuous sentences may have reduced their association with incongruency. Nonetheless, we still found a statistical-significant association between thematic continuity and incongruence. Furthermore, if we had not included the first person as a criterium for thematic continuity, we might find some association between thematic continuity and subtypes of PBI, as seen in Table 4.

4.1.4 Sufficiency

We hypothesized that prosodic boundary incongruency could be related to local word-level features and, then, “sufficiency”, as we defined it, could be associated with a readers' bias towards prosodic boundary incongruence. Our findings confirmed it, showing significantly more sufficient words in incongruent sentences and significantly more PBI after sufficient words. Along with that, we found that it did not matter if a non-sufficient word belonged to a potential utterance or an information module.

4.2 Limitations

In addition to the non-experimental design, our study has limitations we will try to diagnose and report.

4.2.1 Corpus and speakers

The speakers represent only a particular stratum of BP speakers: skewed to male, older, educated, upper class. Speakers who read more were overrepresented compared to those who read less or did not read at all. The speeches were also very specific, as they were political.

4.2.2 Variables

There was no hard science in the measuring of our variables. Firstly, even though we assessed the inter-rater reliability of prosodic boundary incongruence perceptions, they are still perceptions. Secondly, the concept and criteria we proposed to measure “sufficiency” are still fuzzy and need more clarification and inter-rater validation. Therefore, when it comes to the association between prosodic incongruency and word sufficiency, what we can say for sure is that our perception of prosodic boundary incongruence is strongly associated with our perception of word sufficiency. Thus, replications of this study would be reassuring.

4.2.3 Confounding

We analyzed variables associated with the congruency of entire sentences. However, sentences may have different extensions, and larger sentences may have more incongruences than shorter ones. Since we did not control sentences by their extensions, actor prototypicality and thematic continuity may be associated with sentence extension, and sentence extension may be the middleman between those variables and congruency. When it came to sufficiency, we also analyzed specific locations inside sentences, and thus, in that setting, extensions were not confounding factors.

4.3. Interpretation

4.3.1 Thematic continuity

Supposing the association between thematic continuity and prosodic segmentation incongruence is genuine, it would be interesting to speculate why. The speculations we put forward are consistent with the interpretation that written language is incongruous with spoken language, and mechanisms of language processing designed by nature to help the latter do not necessarily benefit the former.

One line of thought would be the idea of priming. Current sentences tend to be biased towards previous sentences' syntactic, semantic, or even prosodic (26) characteristics. Therefore, when the current sentence topic has been mentioned in the previous sentence, that may redeem latent primed features that do not help the current sentence's speech segmentation and, probably, may even disrupt it. Another possibility is that active mechanisms like predictive processing that "exploits multiple constraints in parallel across the different levels of linguistic representation" may play a role in misguiding the reader's interpretation (27). Finally, we could propose that a more semantic approach, like the concept of "preparedness", meaning "information that is given makes contact with linguistic material that came before, as well as with background knowledge, and integration of the input with preceding context and knowledge leads to the creation of a rich semantic representation", induce a misrepresentation of the forthcoming sentence (28).

4.3.2 Actor prototypicality

Actor prototypicality, the next variable we studied, showed an association between first-person actors and sentence congruence. A reanalysis of the data revealed that most first-person actors were not overtly expressed (subject drop, 88%). However, subject drop alone was not associated with congruency. Reviewing the data, 265 of 325 subject drops were first person, suggesting that maybe the 60 second and third-person subject-dropped actors were associated with incongruency.

In either way, it seems that the Actor Identification Strategy model (20) does not help prosodic speech segmentation in oral reading. What does seem to help speech segmentation is a higher degree of grammaticalization of the actor: BP marks first-person subjects on the verb, which means BP has morphosyntactic properties that unequivocally relate an argument (in this case, the first-person subject) to its clause, helping readers understand the sentence's grammatical relations and apply proper speech segmentation.

4.3.3 Sufficiency and autonomy

From the analyses of "autonomy" and "sufficiency", we learned that a) belonging to the smallest linguistic unit with interpretability in isolation and pragmatic autonomy (the minimal utterance extracted from a larger sentence) does not bring about prosodic congruency, and b) prosodic speech segmentation appears to be highly responsive to local word-level properties. Whenever a word in any written phrase or clause or sentence was interpretable as the last one in its respective segment, the odds were that an incongruent prosodic boundary (usually, a non-terminal prosodic boundary insertion) could surface.

4.3.5 Beyond the bounds

A guideline for reporting observational studies (29) advises researchers to give "a cautious overall interpretation of results". Nevertheless, even at the risk of lacking in moderation, we will put forward a particular interpretation of our results. As we have seen, readers inserted, overall,

more than one incongruent prosodic boundary per minute, revealing a bias toward incongruent prosodic boundary insertion and hyper-segmentation of their readings. But why did they segment more instead of less?

Chomsky (30) holds that boundless expressions are the "most basic property of human language", and recursive structures can yield sentences with infinite words. Christiansen and Chater (27) maintain that language processing works under pressure and that the fundamental constraint on language is the working memory. Cowan (31) clarifies that working memory has a mean capacity of 3.5 independent items, ranging from a minimum of 2 to a maximum of 6 items. So, how can sentences have infinite words if working memory has such a limited capacity?

One possibility is that sentences incorporate each new word to the bulk of previous words in a straightforward merge operation. Then, working-memory limits would never be under pressure because it would have only one (the bulk of words) or two (the bulk plus a new word, before the merge operation) of its slots occupied at any time. However, we know that sentences, as defended by Chomsky (30), are composed of "hierarchically structured expressions", and dependencies between words are not always contiguous.

As a matter of fact, hierarchical structures have been interpreted as a domain-general cognitive response to memory constraints. In language, they materialize as progressive merging of phonetic, phonological, word, phrase, clause, sentence, and discourse-level units, from lower to higher levels (32). Besides, these levels are not abstract or arbitrary, as they have tangible neurophysiological signatures in the brain (33). The most studied of those signatures is a centroparietal electroencephalographic positive wave that has been traced to domain-general cognitive phrasing, or segmenting, of any flow of sequential units that must be dealt with by the human brain (34).

In language, there is a correlation between neurophysiological markers of cognitive segmentation and prosodic boundaries. It is not a cause-effect relationship but an association: apparently, prosodic boundaries happen simultaneously with the closure of linguistic segments (35). Our finding that readers are biased to prosodically over segment their oral readings suggests that they are forming shorter linguistic segments in their working memories. It may emerge from a universal cognitive pressure to chunk words into units, transfer those units to long-term memory, and free space in the reader's working memory.

4.3.6 Generalizability

Our results have some characteristics that may hinder their universality. Firstly, we deal with a particular language, and, as we saw, grammatical features may be relevant to the readers' inclination to produce prosodic boundary incongruences. Then, other languages may have different prevalence of the phenomenon.

Secondly, we dealt with a particular group of BP speakers and a specific kind of oral readings: political speeches. Therefore, our results may not apply to other speakers in other circumstances.

However, as long as we proposed a universal mechanism behind the bias towards incongruent prosodic boundary insertion, it is fair to stipulate that it must be ubiquitous if it expects to have any merit.

5. Appendix

5.1 Foreword

Samples of sentences from the corpus are provided below. Each example includes a) the original written text in Brazilian Portuguese (in *italics*) with annotations showing the authors' perceptions of incongruent prosodic boundaries, b) an audio file with the corresponding oral reading, and c) an English translation. In order to give an idea of the original constituent arrangements, translations try to keep constituents in the same relative positions as in the original texts, neglecting better translations options.

The annotations on the written texts are as follows:

- Incongruent insertion of prosodic boundary: "/" non-terminal; "/" terminal.
- Incongruent deletion of prosodic boundary: "*" non-terminal; "*" terminal.

5.2 Examples

1) *Por mais suicida que possa aparentemente sugerir*, uma versão mais suave dessa proposta/ deveria ser, no mínimo, contemplada.* (audio example 1)

Even if it may look suicidal, a mild version of that proposal should be, at least, considered.

2) *Hoje, mais de 30 milhões de brasileiros estão atolados em dívidas impagáveis//, segundo o IDEC.*** (audio example 2)

Today, more than 30 million Brazilians are stuck in insurmountable debt, according to IDEC.

3) *Essa ampliação/ do atendimento é muito bem-vinda, e deve ser buscada e estimulada.* (audio example 3)

The increment of services is welcome and must be pursued and encouraged.

4) *Esses cálculos foram tabulados/ pelos auditores com base nos dados de empregos formais do Ministério do Trabalho.* (audio example 4)

The estimates were tabulated by the auditors, based on employment data from the Ministry of Work.

5) *É certo que, em 2015, não havia crise, já que os reservatórios estavam com 100% de suas capacidades.*** (audio example 5)

It is true that in 2015 there was no crisis since the dams were at their full capacities.

6) *O fluxo de trabalhadores nascidos no Brasil/ em direção ao Japão também é notável//, a partir da década de 1980.*** (audio example 6)

The migration of Brazilian workers towards Japan is also notable, starting in the '80s.

7) *Na opinião de alguns*, Estados economicamente mais fracos/ ocupam espaços de decisão para muito/ além daquilo que se poderia prever dentro de um quadro de normalidade federativa.* (audio example 7)

To some people, States that are economically weaker hold ruling powers that are beyond what should be expected in proper federations.

8) *É preciso, porém, avaliar se a grande maioria da população vem sendo prejudicada em razão da irresponsabilidade de alguns poucos perdulários.*** (audio example 8)

It is necessary to check if most people are not being harmed because of some profligates.

9) *No entanto, para assegurar o abastecimento de água/ aos brasilienses, são indispensáveis obras de infraestrutura// que deem conta de nosso aumento populacional.* (audio example 9)

However, to guarantee water supply to Brasilia's inhabitants, it is imperative to develop an infrastructure that matches our population growth.

10) *Prevê-se, ainda, que o aumento da renda real do trabalhador brasileiro/ neste ano chegue a 2,5%,* depois de dois anos consecutivos de perdas.* (audio example 10)

It is foreseeable that the increase in the earnings of Brazilian workers this year will get to 2.5% after two consecutive years of losses.

11) *Já no setor de prestação de serviços há bons indícios de recuperação, especialmente em transportes e logística, com o aumento do aluguel de galpões.*** (audio example 11)

Still, in the service sector, there are good signs of recovery, especially in transportation and logistics, with an upsurge in rentals of warehouses.

12) *Por tudo isso, entendo que o modelo de Zona Franca celebra seu jubileu/ de ouro com fôlego juvenil//, em razão da atualidade e pertinência do programa.* (audio example 12)

Because of all that, I think the model "Free-Trade Zone" celebrates its Golden Jubilee with the energy of a teenager because of the modernity and the usefulness of the program.

13) *Quando o PIB brasileiro ou sul-americano vai bem, as empresas da Zona Franca ampliam a geração de empregos e os investimentos*, fortalecendo o crescimento econômico no bioma amazônico.* (audio example 13)

When the Brazilian or south-American GDP is fine, companies of the "Free-Trade Zone" expand job openings and investments.

14) *Haverá menos violência – e a que houver será mais bem combatida –, haverá mais e melhores escolas, professores mais bem preparados e motivados, haverá mais médicos, enfermeiros, leitos e hospitais.*** (audio example 14)

There will be less violence - and whenever there is, it will be better countered -, there will be more and better schools, teachers that are well prepared and motivated, there will be more doctors, nurses, beds, and hospitals.

15) *Administração competéssima, muita pesquisa, muito desenvolvimento, muito trabalho//, em todos os patamares da organização*, da base até o topo — tudo convergindo para colocar a empresa na fronteira do conhecimento humano.* (audio example 15)

A competent administration, a lot of research, a lot of development, a lot of work, in all levels of the organization, from the bottom to the top - everything is converging to put the company on the brink of human knowledge.

16) *Conforme aponta o IBGE, praticamente todos esses empregos foram gerados no mercado informal, aquele em que o trabalhador permanece à margem de direitos trabalhistas*

constitucionais, como férias/ remuneradas, licença-maternidade, décimo terceiro/ salário, entre outros. (audio example 16)

As reported by the IBGE, virtually all jobs came from the informal market, where workers remain removed from constitutional labor's rights, like paid vacations, maternity leaves, end-of-year bonuses, among others.

17) *O planejamento das ações para os quatro anos de mandato/ que terei nesta Casa vem não apenas da experiência que adquiri nos cargos que ocupei, mas dos contatos que venho mantendo com a população e com as autoridades do Estado.* (audio example 17)

The planning of the actions to the four years term of office I will have in this House comes not only from the experience I bring from the positions I filled but from the connections I keep with people and authorities from my State.

18) *Cerca de 75% da água do nosso País está localizada nos rios/ da Bacia Amazônica, que é habitada por menos de 5% da população.* (audio example 18)

About 75% of our country's water is on rivers of the Amazon Basin, which is inhabited by less than 5% of our population.

19) *É a recompensa pelo trabalho que permite o consumo.* ** (audio example 19)

It is the earnings from work that allow consumption.

REFERENCES

1. O'Connell DC., Kowal S. *Fluency and Hesitation*. In: *Communicating with One Another. Cognition and Language: A Series in Psycholinguistics*. Springer, New York, NY; 2008. https://doi.org/10.1007/978-0-387-77632-3_3
2. Barbosa P, Raso, T. *A segmentação da fala espontânea: aspectos prosódicos, funcionais e aplicações para a tecnologia*. *Revista de Estudos da Linguagem*. 2018 Oct; 26(4):1397-1433. Available from: <http://www.periodicos.letras.ufmg.br/index.php/relin/article/view/14315>
3. Staub A, Rayner K. *Eye movements and on-line comprehension processes*. In: *The Oxford Handbook Of Psycholinguistics*. New York: Oxford University Press; 2007. p. 326-342.
4. Dehaene S. *Reading in the brain: the science and evolution of a human invention*. Penguin Books, New York, NY; 2009.
5. Coltheart M. *Modeling Reading: the dual-route approach*. In: *The Science of Reading: A Handbook*. Blackwell Publishing Ltd, Hoboken, NJ; 2005. p. 6-23. Available from: <http://dx.doi.org/10.1002/9780470757642.ch1>.
6. Washington, Dc: National Center For Education Statistics, Institute Of Education Sciences, U.S. Department Of Education, 2005. Daane, MC, Mary C, et al. *Fourth-Grade Students Reading Aloud: NAEP 2002 Special Study of Oral Reading*. [cited 2018 Oct 24] [72 p.]. Available from: <https://nces.ed.gov/nationsreportcard/pdf/studies/2006469.pdf>
7. Kuhn M et al. *Aligning Theory and Assessment of Reading Fluency: Automaticity, Prosody, and Definitions of Fluency*. *Reading Research Quarterly*. 2010; 45(2): 230-251.
8. Izre'el S. *Syntax, Prosody, Discourse and Information Structure: The case for Unipartite clauses — A View from Spoken Israeli Hebrew*. *Revista de Estudos da Linguagem*. 2018; 26(4): 1675-1726. Available from: <http://dx.doi.org/10.17851/2237-2083.26.4.1675-1726>.
9. Raso T, Teixeira, B, Barbosa P. *Modelling automatic detection of prosodic boundaries for brazilian portuguese spontaneous speech*. *Journal Of Speech Sciences*. 2020; 9: 105-128. Available from: <http://dx.doi.org/10.20396/joss.v9i00.14957>.

10. Landolfi A, Sammarco C, Voghera M. *Verbless clauses in Italian, Spanish and English*. JADT 2010: 10th International Conference on Statistical Analysis of Textual Data. 2010. p. 1187-1194. http://lexicometrica.univ-paris3.fr/jadt/jadt2010/allegati/JADT-2010-1187-1194_066-Landolfi.pdf
11. Bornkessel-Schlesewsky I, Schlesewsky M. *Competition in argument interpretation: evidence from the neurobiology of language*. In: Macwhinney B, Malchukov A, Moravcsik E, editors. *Competing Motivations in Grammar and Usage*. New York: Oxford University Press; 2014. p. 107-126.
12. Grosz B, Sidner CL. *Attention, intentions, and the structure of discourse*. Computational Linguistics, Cambridge, Ma. 1986: 12(3): 175-204.
13. Partee B. *Topic, Focus and Quantification*. Semantics And Linguistic Theory. 1991: 1: 159.
14. Haspelmath M. *Arguments and Adjuncts as Language-Particular Syntactic Categories and as Comparative Concepts*. Linguistic Discovery. 2014: 12(2): 3-11.
15. Levin B. *Argument Structure*. Oxford Bibliographies [internet]. New York: Oxford University Press; 2013 [updated 2018 apr 26]. Available from: <http://dx.doi.org/10.1093/obo/9780199772810-0099>.
16. Przepiorkowski, A. *How not to distinguish arguments from adjuncts in LFG*. In: Arnold D, Butt M, Crysmann B, King TH, Muller S, editors. *Proceedings of the Joint 2016 Conference on Head-driven Phrase Structure Grammar and Lexical Functional Grammar*. Stanford, CA: CSLI Publications; 2016. p: 560-580.
17. Merlo P, Ferrer EE. *The Notion of Argument in Prepositional Phrase Attachment*. Computational Linguistics. 2006: 32(3): 341-378.
18. Senado Multimídia. Brasil: Senado Federal; [date unknown]. Available from: <https://www12.senado.leg.br/multimidia>.
19. Lima AM, Constantini AC. *Prosódia e fonoaudiologia: do fonoestilo ao transtorno da linguagem*. In: *Prosódia da Fala*. Editora Edgard Blucher, [S.L.]; 2017. p. 133-144. Available from: <http://dx.doi.org/10.5151/9788580392593-08>.
20. Alday PM, Schlezewsky M, Bornkessel- Schlezewsky I. *Towards a Computational Model of Actor-Based Language Comprehension*. Neuroinformatics. 2013: 12(1): 143-179.
21. Botinis A, Granstrom B, Mobius B. *Developments and paradigms in intonation research*. Speech Communication. 2001: 33(4): p.263-296.
22. Cresti E, Gregori L, Moneglia M, Panunzi A. *The Language into Act Theory: A Pragmatic Approach to Speech in Real-Life*. In: Koiso H, Paggio P, editors. *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018), LB-ILR2018 and MMC2018 Joint Workshop: Language and Body in Real Life Multimodal Corpora 2018: Multimodal Data in the Online World*. ELRA, Paris; 2018. p: 20-25.
23. Alboukadel K. *Cohen's Kappa in R: For Two Categorical Variables* [Internet]. [Place unknown]: Datanovia, [Date unknown]. Available from: <https://www.datanovia.com/en/lessons/cohens-kappa-in-r-for-two-categorical-variables/#:~:text=Cohen%27s%20Kappa%20in%20R%3A%20For%20Two%20Categorical%20Variables,-20%20mins&text=This%20process%20of%20measuring%20the,is%20called%20inter%20drater%20reliability.&text=The%20Cohen%27s%20kappa%20is%20a,that%20removes%20this%20chance%20agreement>
24. RStudio: Integrated Development for R. Boston (MA): R RStudio Team; 2020. Available from: <http://www.rstudio.com/>.
25. Landis JR, Koch GG. *The Measurement of Observer Agreement for Categorical Data*. Biometrics. 1977: 33(1): p. 159.
26. Jun, SA, Bishop, J. *Priming Implicit Prosody: Prosodic Boundaries and Individual Differences*. Language and Speech. 2015: 58(4): p.459–473.
27. Christiansen M, Chater N. *The Now-or-Never bottleneck: A fundamental constraint on language*. Behavioral And Brain Sciences. 2015: 39: p.1-72.
28. Ferreira F, Chantavarin S. *Integration and Prediction in Language Processing: A Synthesis of Old and New*. Current Directions In Psychological Science. 2018: 27(6): p.443-448.

29. Vandenbroucke JP, von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, et al. *Strengthening the reporting of observational studies in epidemiology (STROBE): explanation and elaboration*. PLoS Med. 2007; 4(10): e297.
30. Chomsky N. *Minimal Recursion: exploring the prospects*. In: Roeper T, Speas M, editors. *Recursion: Complexity in Cognition*. [S.L.]: Springer International Publishing; 2014. p. 1-15.
31. Cowan N. *The magical number 4 in short-term memory: a reconsideration of mental storage capacity*. Behavioral And Brain Sciences. 2001; 24(1): p. 87-114.
32. Blumenthal-Dramé A, Malaia E. *Shared neural and cognitive mechanisms in action and language: The multiscale information transfer framework*. Cognitive Science. 2018; 10(2): p.1-18.
33. Ding N, Melloni, L, Zhang, H et al. *Cortical tracking of hierarchical linguistic structures in connected speech*. Nat Neurosci. 2016; 19: p. 158–164.
34. Hilton M et al. *Parallels in Processing Boundary Cues in Speech and Action*. Frontiers In Psychology. 2019; 10: p.1-12.
35. Friederici AD. *Language Functions in the Brain: From Auditory Input to Sentence Comprehension: Processing Prosodic Information*. Language in our brain: the origins of a uniquely human capacity. Cambridge, MA: The MIT Press; 2017: p. 71-81.