Understanding syntax structure of language after a head injury

Abstract

Cognitive-communication disorder is a type of language alteration commonly associated with a traumatic brain injury, but can also be a sequel from a disease. There is a latent possibility of establishing a connection between the ongoing epidemic language alteration sequelae with that language reorganization after a head trauma, that includes modifications in the syntax production process.

From a set of syntax indices, we propose a grammar-based analysis depicting triangle-segmented maps. The produced maps could serve to interpret how grammar rules progress, when further syntax complexity is required.

The central finding is that the suggested context-free grammar allows a suitable representation of the construction of syntax production in individuals after a traumatic brain injury in a post-recovery stage.

Introduction

A disorder in cognitive communication disrupts communication capability, a type of acquired language deficit that can follow a traumatic brain injury (TBI). This sequela affects different aspects of people lives, that can go from the loss of friends to the risk of an unstable psychological state. This implies depression, anxiety, delirium, and psychotic behaviors, that brings patients in an uninterrupted cycle of drop-backs.

TBI represents a worldwide leading cause of death and disability. Nowadays, some evidence suggests that these cases share some syntax deficits that can appear in language sequelae for oxygen deprivation caused by the SARCOV-II virus. Among such deficits is the essential day-to-day communication affected by altering syntax comprehension and production.

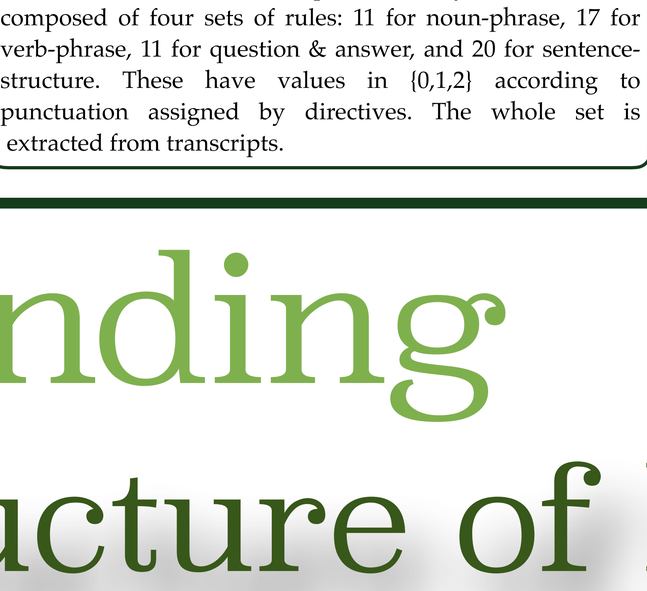
Syntax dictates the valid combinations of words that make sense in language. We look for an improved understanding of the information that comprises the complete set of characteristics describing syntax. For this, we propose and explore triangle-segmented maps, representing the indices of syntax production.

For that purpose, a context-free grammar (CFG) is defined to generate those maps. Our main finding is that the proposed grammar and its related graph representation seem to support the identification of evolving patterns within features assessing syntax production.

Objectives

1. The recognition of the behavior of syntax features in atypical language.
2. The understanding of the evolution of the restructuring of language.
3. A means to enable to identify patterns of the TBI-affected evolving language.

Related Work



Materials & Methods

Data. The TBI corpus is a product of a project between the University of Sydney and Carnegie Mellon University.

The retelling of Cinderella story was worked for different time points. However, given that the 24-months stage has less instances than the rest, it was extended with the recount of an-important-event samples.

Feature set. The index of productive syntax (IPSYN) is composed of four sets of rules: 11 for noun-phrase, 17 for verb-phrase, 11 for question & answer, and 20 for sentence-structure. These have values in {0,1,2} according to punctuation assigned by directives. The whole set is extracted from transcripts.

Approach

For a grammar-based analysis, a context-free grammar is associated with a triangle-segmented polygon. This is inspired by the association that Gross & Lenin did of a side-oriented triangle with the derivation of a given grammar. The generated map serves to contrast the progress of the syntax construction of a control sample versus instances of study when rules are more elaborated. In the next definition, observe the introduction of the symbol ∅2.

Definition.- Let set G = (V, Σ, S, P),

where Σ = {∅2, 0, 1, 2} (below the colors associated to the map)

∅2 → 1

0 → 2

is the finite terminal symbol set, ∅2 specifies the absence of the base to generate the next triangle. V = {S} is the non-terminal or variable set. The finite set P of productions is given by:

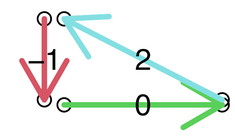
S → 2SS

S → ∅2SS

S → 0 | 1

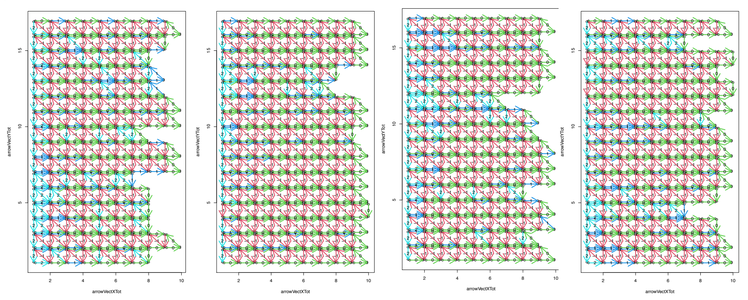
The unique symbol S ∈ V also denotes the start variable.

Example:

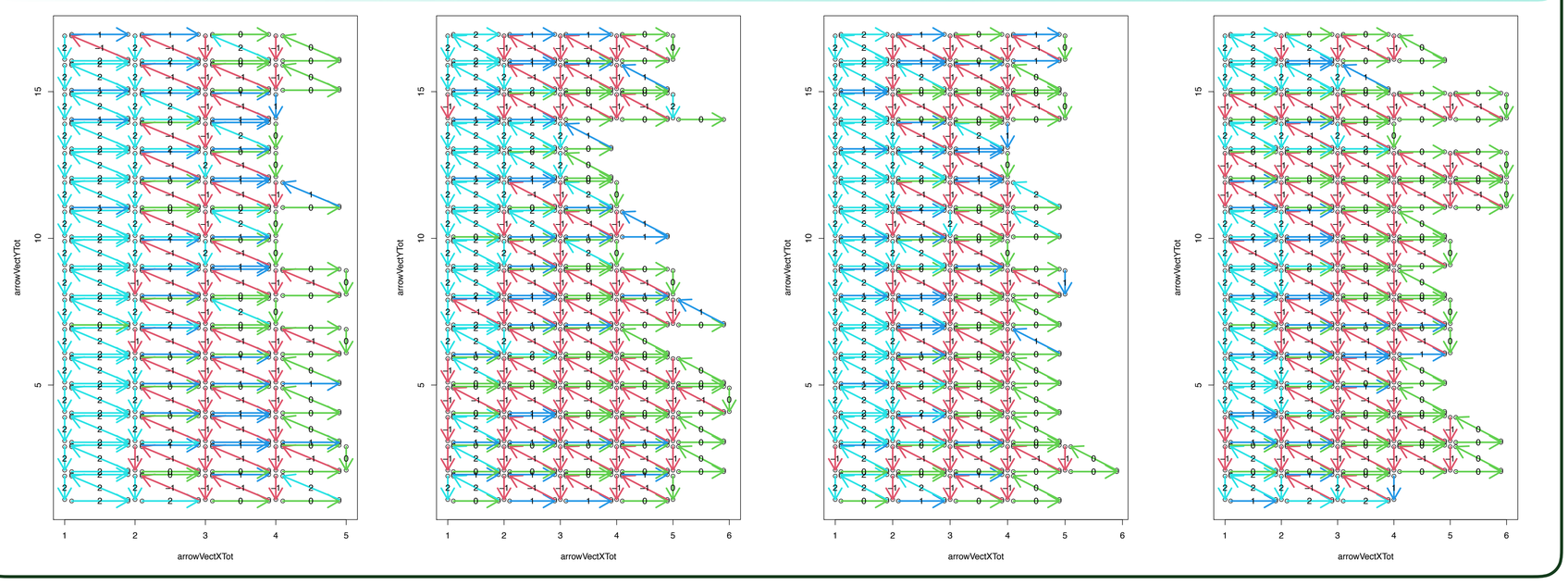


Rationale. Starting with the non-terminal S associated with a real side, the base of the triangle and of the next derivation, the structure of the map advance according to the grammar. When the base is absent, a ∅**2** is fed allowing more complex syntax composition.

Sentence-Structure maps



Results - Noun-Phrase Group maps



Discussion

Note that the complexity of the rules determining the IPSYN indices increases with their subsequent definitions, including some dependencies, the conditional rules. Thus, we hypothesized that posterior markers are built on the basis of those preceding. This justifies the connection of the suggested CFG with the previously described triangular polygon map representation. Considering the real side as the base on which the subsequent syntax structures are built and its absence to continue assembling the syntax could represent a signal of the intended restructuring of the examined language. For this reason, the ∅2 symbol was introduced in the specification of the grammar, which is associated with the −1 value (red arrow).

A three-bar pattern is evident in the figures, the noun-phrase group, though the negative group shows more dominant completeness of the syntax structure. Note, at the start of the feature responses production chain, the grammar rule S → 2SS is principally applied, in agreement with our hypothesis. This suggests that the basis directives of the grammar are satisfied, moreover with a notable favorable outcome for the control group compared to the response for the three and twenty-four months. On the second colored delimited block, the grammar rule S → $\varnothing_2$SS predominates in both groups. The last batch is extended to two columns for the three, twelve, and twenty-four months samples, whereas the negative set consists of only one, with more recurrent insertions of $\varnothing_2$ to our study group than for the negative set.

Conclusion & Further work

A triangle-segmented polygon diagram served to decipher the syntax construction of the analyzed groups. Although some cases that do not adhere to it need attention.

Counting with all transcriptions will enable us to replicate this analysis on the entire samples of the recovery stage. Comparing syntax production results against the current, looking for an understanding of the presented alterations around the TBI-affected language.

References

1. Ahmed, S., Venigalla, H., Mekala, H.M., Dar, S., Hassan, M., Ayub, S.: Traumatic brain injury and neuropsychiatric complications. Indian J. Psychol. Med. 39, 11– 121 (2017) .
2. Gross, M., Lentin, A.: Introduction to Formal Grammars. Springer, Heidelberg (1970).
3. Hegde, M.N., Freed, D.: Assessment of Communication Disorders in Adults Resources and Protocols. Plural Publishing, Inc. (2022).
4. MacWhinney, B.: The CHILDES Project: Tools for Analyzing Talk, 3rd edn. Lawrence Erlbaum Associates, Mahwah (2000).
5. Roldán-Palacios, M.: Closure under substitution to proper flexagons. Bachelor thesis, BUAP (2018).
6. Roldán-Palacios, M. and López-López, A., 2023. Feature analysis and classification of impaired language caused by brain injury. In *Natural Language Processing in Healthcare* (pp. 19-35). CRC Press.
7. Roldán-Palacios, M.: A multi-level analysis of language production: features compromised by traumatic brain injury. Master thesis, INAOE (2021).
8. Scarborough, H.S.: Index of productive syntax. Appl. Psycholinguist. 11(1), 1–22 (1990)
9. Stubbs,E.,etal.:Procedural discourse performance in adults with severe traumatic brain injury at 3 and 6 months post injury. Brain Inj. 32(2), 167–181 (2018).
10. TBIbank. https://tbi.talkbank.org/. Accessed Sept 2021

Authors

Roldán-Palacios Marisol, López-López Aurelio.

Instituto Nacional de Astrofísica Óptica y Electrónica (INAOE)

Acknowledgement

INAOE CONACYT