Interactive Poster: Displaying English Grammatical Structures

Pourang Irani

University of Manitoba
Department of Computer Science
irani@cs.umanitoba.ca

Yong Shi

University of Manitoba
Department of Computer Science
yongshi@cs.umanitoba.ca

ABSTRACT

This report describes ongoing work focused at designing a technique for visually representing English grammatical structures. A challenge in representing grammatical structures is to adequately display the linear as well as hierarchical nature of sentences. As our starting point we have adopted a radial space-filling technique based on Clark's etymological chart of the 19th century. Clark devised this chart for the purpose of instructing students English grammar. We have automated the chart with basic visual features and interaction techniques. We report the results of a preliminary evaluation that suggests that subjects are able to better identify parts of a sentence after minimal training with the interactive visualization system.

Keywords

Visualizing English sentences, language structure visualization, radial space-filling visualization.

1. INTRODUCTION

As part of the writing process, the writer needs to know how to recognize complete thoughts and accordingly vary sentence structures to reflect these. Understanding the structure and various relationships between components in a sentence facilitates coherent writing. Many grammarians and English instructors hold that analyzing a sentence and portraying its structure with a consistent visual scheme can be helpful—both for language beginners and for those trying to make sense of the language at any level [3]. This is especially true for language learners who tend to be visual-learning types. One approach to better learning and understanding grammatical structures is to use diagrams.

Several types of diagramming notations have been developed for capturing and representing structures in English grammar. Some of these are Clark's diagrams [2], syntactic trees [1], and Kellogg-Reed diagrams [4]. In Clark's diagrams, words, phrases, and sentences are classified according to their roles, and their relation to each other. Clark's diagrams are hierarchical in that the first stage decomposes the parts into the appropriate structural units (subject, verb, noun, etc.). At a lower level, each unit is broken down into it various components. The elements are visually depicted by showing each unit as an outlined shape oval, and connection between units as lines or appendices. Syntactic trees provide a hierarchical representation of sentence structures. At the most bottom level, leaf nodes contain each atomic unit of the sentence. Above each leaf node in the tree, the specific role played by each atomic unit in the sentence is presented. These could be nouns, pronouns, prepositional phrases, adverbs, etc. In a recursive fashion, the role of each unit (compound or atomic) is depicted as a node of the tree. The most widely used form of sentence visualization has been developed by Brainerd Kellogg and Alonzo Reed, and is known as Kellogg-Reed diagrams. In the Kellogg-Reed diagrams, a sentence is divided into its component

parts using solid and dashed lines. The most important cut being between the subject and the predicate. Horizontal lines are used for key structural elements, such as subject, verb, and direct object. Modifiers are placed on a diagonal bar and under the key elements they modify. Several hierarchies can also result from sentences that contain compound elements. Overall, these notations are weak in representing different types of relationships and semantics used in English grammatical structures. It is important to clearly reveal these relationships in order to allow the student to fully grasp the grammatical concepts. While these representations are complete, they are disjoint and do not provide a unified classification of the various types of possible sentence structures. As a result, they may not facilitate the learner who is particularly unaware of the range of sentence constructs in the language.
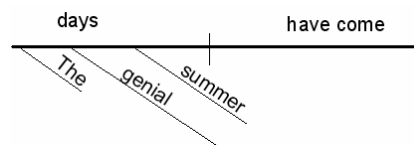

Figure 1. Kellogg-Reed diagram for the sentence "The genial summer days have come"

The inherent structure of these representations is either linear (as in the case of the Kellogg-Reed diagrams) or hierarchical (syntax trees). We hypothesized that adopting a representation that is at the same hierarchical and linear will facilitate analysis of sentences into their constituents.

2. CLARK'S ETYMOLOGICAL CHART

An alternative to providing separate and disjoint diagrams for the various forms and patterns of sentences is to create a compact representation. The representation would need to depict the linear as well as hierarchical construction of sentences in order to provide the learner a stronger view of the sentence. Such a compact representation has been proposed by Clark [2] in the 19th century and is entitled as Clark's etymological chart. While Clark's terminology is in certain places antiquated, the chart is compact and provides the learner with a concise representation of the various functional elements that could be part of a sentence. We have implemented this chart as a starting point over which all other visualization and interaction features are developed. A remarkable feature of Clark's representation is the compactness that allows the entire system of grammatical constituents of sentence patterns to be depicted. Figure 2 shows our implemented version of Clark's chart with the various elements of a sentence.

Clark's chart uses a radial display technique similar to that used by Sunburst [5]. While Sunburst is designed to display any form of hierarchy, Clark's chart imposes a strict ordering of the constituent nodes based on the sentence being represented. At the center of the chart is the root node representing the entire sentence. At the next level, the chart contains two nodes, one

representing the *principle parts* and the other the *adjuncts* or *qualifiers* of the elements in the sentence. The principle part is further decomposed into a node representing the *subject*, the *predicate* and the *object* of the sentence. The adjuncts are separated into primary and secondary, the former qualifying elements within the principle part of the sentence, while the latter qualifying elements within the primary adjunct. At deeper levels in the hierarchy the various functions that constituent elements represent are depicted. For example, a subject can be represented by a word, a phrase or another sentence. In turn a word can either be a noun or pronoun. A noun can either be a proper or common noun either being in the masculine or the feminine gender, and finally in the singular or plural form.

We have adapted Clark's chart as the base representation and have augmented it with perceptual and interactive elements (Figure 2.a). We use color as the primary perceptual feature for highlighting the various components of a sentence. A color highlights all constituent elements of a sentence part through the sub tree of the hierarchy. A common problem affecting radial displays is the layout of the text. To facilitate text readability we implemented automatic smooth zooming whereby the chart is rotated to position the node of interest in a vertical readable position (Figure 2.b).

To initially validate the effectiveness of the radial chart for language structures, we conducted a preliminary evaluation. Six computer science students from the University of Manitoba participated in the evaluation. None were familiar with any sentence diagramming methods. A pre-training evaluation was conducted to determine the students' ability for parsing sentences into their components. All six subjects demonstrated a low and equal performance rate. To perform the evaluation we included a range of simple and complex sentences in the tool. By selecting a particular sentence its visual representation would get highlighted in the chart. Students were given time to familiarize themselves with the tool by selecting the various sentences and viewing their structure in the chart (lasted 20 minutes). The experiment then consisted of displaying a sentence and presenting the subject with a range of possible structures to choose from within the chart. The subject was then asked to select the visual representation that best suited the sentence. All subjects scored higher in the post-training evaluation after using the tool. These results provide a hint at the potential benefits that the chart may afford.

3. FUTURE WORK AND CONCLUSION

In this poster we discuss the automation of Clark's etymological chart for the purposes of helping learners decipher sentence structure and their parts. The space-filling radial representation was evaluated and the results showed that subjects were able to break sentence constituents better using the visual aid.

An objective of a visual tool for depicting sentence structure would be to facilitate learning and self correction of grammatical errors. Self correcting tools exist in editors such as MS Word ™. However, the methods simply hint at possible sentence errors without giving much recourse to a possible solution. Our future work will consist of further developing the tool to aid learners in identifying and possibly self-correcting grammatical errors. We will additionally augment the tool with focus+context techniques such as those discussed in [5] that will allow users to manipulate the chart for extracting vital information to their tasks.
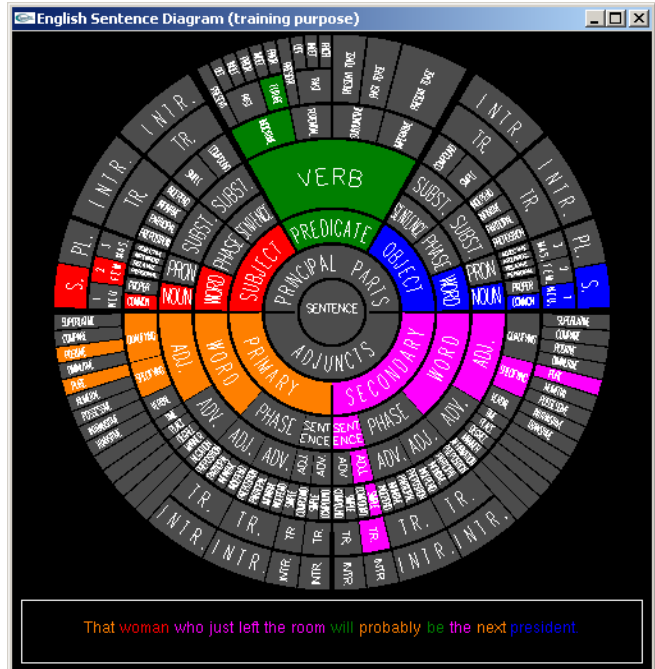
a) Augmenting Clark's etymological chart with visual features such as color to display sentence structures.

b) Automatic zooming that rotates the radial display to align the text with the user's node of interest. Here the user clicked on the Subject node to bring it into focus and then on the Predicate node.

Figure 2. Representation of Clark's etymological chart to highlight sentence structure using color and to facilitate interaction using automatic zooming

4. REFERENCES

[1] Chomsky, N. (1965). Aspects of a Theory of Syntax, Cambridge: M.I.T. Press.

[2] Clark, S.W. (1853). A practical Grammar: in Which Words, Phrases, and Sentences are Classified According to their Offices, and their Relations to Each Other, New York, A. S. Barnes & Co.

[3] Pinker, S. (1989). Learnability and Cognition: The Acquisition of Argument Structure, The MIT Press, Boston, MA.

[4] Reed, A and Kellogg, B. (1878). Elementary English Grammar, New York, Clark & Maynard.

[5] Stasko, J. and Zhang, E. (2000). Focus+Context Display and Navigation Techniques for Enhancing Radial, Space-Filling Hierarchy Visualizations, Proc. of the IEEE Symposium on Information Visualization 57-65.