

COGNITIVE TYPOLOGY AND DIGITAL METHODS: A CORPUS-BASED ANALYSIS OF CONCEPTUAL AND INFORMATION-PROCESSING DIFFERENCES BETWEEN ENGLISH AND AZERBAIJANI

Халилова Гульнар Байрам гызы
Университет Одлар Юрду,
преподаватель английского языка
на кафедре иностранных языков

Abstract. *This study applies techniques from cognitive typology and corpora analysis to contrast English and Azerbaijani. It shows how different linguistic structures reveal the workings of different types of information processing and that syntactic constructions can be studied scientifically.*

Keywords: *cognitive typology; corpus linguistics; information processing; English; Azerbaijani; digital methods*

Introduction

The paradigmatic change in the field of linguistics from the system-centered to the anthropocentric one has greatly influenced the research on the subject matter of linguistics. The language is perceived not only as a set of rules that determine the way in which the process of communication occurs but also as an illustration of the processes taking place within the human mind. The result of this paradigmatic change became the development of the cognitive typology approach.

This method is based on studying the differences in cognitive strategies and ways of information processing used in various cultures and their impact on the typological diversity of languages. Thus, the syntax can be regarded as a cognitive interface where the structure and meaning of thoughts are encoded and conveyed to the recipient. A comparison of two languages - English and Azerbaijani - provides a fascinating insight into the nature of cognitive approaches to information processing.

While the field of cognitive typology provides some rich theoretical insights, most of the existing studies in this domain tend to be qualitative in nature or at least limited by the lack of quantitative data. There is an increasing need for the development of empirical tests of cognitive ideas, particularly with the help of corpus approaches [McEnery, Hardie, 2012, p.23].

The purpose of this research is to link theory and data by integrating cognitive typological and computer-based corpus methodologies. In particular, the paper focuses on the way in which English and Azerbaijani conceptualize their world differently and process information.

Theoretical Framework: Cognitive Typology

Since a transition from the "system-structural" paradigm in linguistics to the "anthropocentric" paradigm takes place, one should conduct the research concerning not only syntax as a grammatical category but also the representation of ways of thinking and the models of the world. It is worth noting that when one conducts a comparative analysis between the inflectional-analytical structure of the English language and the agglutinative-synthetic structure of the Azerbaijani language, one comes across an entirely new phenomenon which can be called "cognitive typology". The point here is

not in the word order itself (SVO or SOV) but in different cognitive strategies of information encoding/decoding.

The English language is a subject language typologically, and it is also known for a fairly static syntax. The latter means that there is “positional determinism” operating in this language, whereby the function of a word in a sentence depends on its position rather than morphology.

Word order is the means of transforming lexical items in Germanic languages; thus, the physical placement of elements becomes significant. According to Fakhreddin Veysalli in his book on linguistics *Fundamentals of Linguistics*, “Here the lexicon is strengthened all the time. There is information coming from other parts as well. These cooperate, they control structurally.... The parts of generative grammar are systematically interrelated. Each part has its own function and structure.” [Veysalli, 2013, s. 317].

The above-mentioned cognitive approach suggests a linear processing approach. The processing approach involves sequential analysis and the building of expectations by the speaker, which is consistent with psycholinguistic approaches to processing, such as incremental processing theory [Levy, 2008, p.19].

Dependency length is one of the most significant characteristics of cognitive processing that can be measured. It refers to the number of items separating two syntactically related components in a sentence. Dependency length has frequently been applied as an indicator of the level of difficulty involved in cognitive processing. A large amount of evidence suggests that languages generally have an inclination toward minimizing the distance between dependencies to ensure effective communication [Futrell et al, 2015, p.37]. Dependency length minimization is seen as a universal property that influences the structure of syntactic constructions across languages.

However, with respect to English, the word order of which is quite rigid, the dependency distance between constituents tends to increase in some cases. In particular, the inflexible positions of basic constituents of the sentence imply that speakers do not have enough possibilities for altering their order in order to reduce the syntactic dependencies of processing. Consequently, prediction processes and incrementally organized analysis are applied to compensate for difficulties related to rigid syntax.

At the same time, Azerbaijani as an agglutinative language with flexible word order offers sufficient possibilities to speakers for manipulating syntactic elements and changing their order. The ability to modify the order of elements makes it possible to achieve shorter distances between syntactically related constituents and reduce the burden of processing. Therefore, it can be stated that Azerbaijani applies more adaptive information organization and employs syntactic variation to achieve the optimal performance of cognitive operations.

An especially interesting aspect of their dissimilarity is in terms of their approach to achieving communicative intent. In the theory of communicative dynamics, which was postulated by the Prague Linguistic School, the main purpose of a sentence is not only to transmit information but to do so in such a manner that the most informative component stands out [Firbas, 1992, p.65].

In English, emphasizing an element other than the subject, particularly the object, usually involves making grammatical changes to the sentence structure. One way of doing this is by creating passive structures, which help the speaker emphasize information:

“The letter was written by the manager.”

As one could observe, this change entails not only the rearrangement of syntactic components within the sentence but also morphological alterations, like the application of auxiliary verbs and participles.

Nonetheless, in the Azerbaijani language, it is possible to achieve a similar communicative goal without making any structural adjustments in the verbal predicate itself. Namely, in this case, word order becomes flexible enough to enable the required element to be shifted into a focal position. According to V.N. Yartseva it is also important to note that this principle of communicative economy, where information reshaping requires the least amount of morphological adjustments, is inherent in other languages, as well [Ярцева, 1975, с.119].

English uses cleft sentences (“It was ... who...”) to show focus, making syntax more complicated. On the other hand, Azerbaijani uses intonation and word order, specifically the pre-verbal focus position. These differences demonstrate two different mental processing techniques:

English: structure-based marking

Azerbaijani: position-based and prosodic marking

In order to support the theoretical assertions empirically, the current study utilizes a corpus-based approach. Frequency, concordance, and syntactic pattern analyses are performed using software programs, such as Sketch Engine and AntConc.

The empirical foundation of the current study is provided by a collection of comparable corpora consisting of English and Azerbaijani linguistic materials in similar communicative spheres. To achieve reliable and balanced results, the data sample includes literary and journalistic sources along with transcriptions of spoken language materials. It is expected that, owing to the selection of different genres, syntactic patterns can be observed within various language registers without being influenced by the stylistic peculiarities of each specific genre.

The analytical model of the research is based on a number of important parameters, which are directly linked with the process of information processing at the level of syntax. Firstly, word order is analyzed because it constitutes one of the most significant differences between English and Azerbaijani from a typological perspective, functioning as an effective way of information organization. Secondly, the frequency and distribution of passive constructions are studied, which is essential since this type of construction plays an important part in the focus reorganization in English sentences. Thirdly, topicalization processes are examined, meaning the techniques used to emphasize certain elements within a sentence as its focal point. Lastly, dependency length is viewed as a measure of cognitive effort required.

The findings of the corpus study show a number of statistically significant trends that support the hypotheses posed at the theoretical stage. Thus, English has a relatively high number of fixed SVO (subject-verb-object) patterns, which suggests that it tends to rely on strict syntactic rules as the only way to organize meaning and grammar. Besides, there are many passive constructions in English, meaning that there is a tendency to change the focal point of the sentence through structural transformation. Azerbaijan uses a different word order pattern: it shows a lot of variability, providing room for organizing sentence components depending on communicative goals.

Moreover, the study has revealed that dependency spans in Azerbaijani generally turn out to be shorter than in flexible structures. This indicates the optimization of

syntactic organization relative to the process of cognitive operations. The language minimizes the load on the working memory by placing syntactically related elements into proximity. As a result, the conclusions about the more flexible strategy of information processing used by Azerbaijani and the rigid approach typical for English can be drawn.

Findings of this research have confirmed the fact that cognitive typology differences should not be regarded as mere theoretical concepts since they can be empirically verified. In this regard, a comparison between the features of the two languages - English and Azerbaijani - proves that different types of principles of syntax organization are used depending on specific cognitive patterns in both languages. In particular, while English is characterized by predictability and linearity in syntax organization, Azerbaijani is more flexible in its structure, allowing a speaker to vary it according to communicative needs.

These differences point to fundamental differences in how the respective languages process information. English is based on a largely sequential type of information processing, whereby the data is analyzed sequentially and relies considerably on position, as well as structural coding of grammatical and semantic relationships. Azerbaijani, by contrast, involves a flexible type of processing, whereby the structure of the information is quite dynamic and relies considerably on position as a means of emphasizing meaning.

Incorporation of corpus-driven computerized methodologies in such an investigation will be highly instrumental in establishing an empirical base to the field of cognitive linguistics by proving such patterns quantitatively. It will also present new possibilities of research within the field through an interdisciplinary approach that will integrate the fields of linguistics, cognitive sciences, and computerized methodologies.

Conclusion

Thus, this paper proves that it is possible to successfully employ cognitive typology together with digital and corpus-based approaches for the acquisition of empirical results in studying the interrelation of language and cognition. The combination of the theoretical perspective and quantitative data indicates that differences in typology imply the existence of differences in the ways people perceive and store information.

Specifically, the contrastive analysis of English and Azerbaijani illustrates how differences in word order, which are rigid for one language and flexible for another one, correspond to various information-processing principles, namely linear/predictive and adaptive/contextual ones.

However, the results achieved show the crucial role played by digital methodology in modern linguistics. In other words, corpus-based approach makes it possible to analyze various patterns that cannot be identified solely theoretically.

In future, it would be reasonable to apply the proposed scheme to a wider variety of languages and adopt more complex computational means.

References:

1. Firbas, J.: Functional Sentence Perspective. Cambridge University Press. 1992, 114 p.
2. Futrell, R., Mahowald, K., & Gibson, E.: Large-scale evidence of dependency length minimization. PNAS. 2015, 78 p.
<https://doi.org/10.1073/pnas.1502134112>
3. Levy, R.: Expectation-based syntactic comprehension. Cognition. 2008, 117 p.
4. McEnery, T., & Hardie, A.: Corpus Linguistics. Cambridge University Press, 2012. 312 p.
5. Veysəlli, F.Y.: Dilçiliyin əsasları. Bakı: Mütərcim, 2013, 420 s.
6. Ярцева, В.Н.: Историко-типологическое изучение грамматических категорий. Москва: Наука, 1975, 217 с.