Contents

Cognitive Factors Affecting the Prevalence of SOV and SVO Word Orders

Manouchehr Kouhestani[1], Arsalan Golfam[2]

Abstract

It is a typological observation in more than 90% of languages where the basic word order is either subject–verb–object (SVO) or subject–object–verb (SOV). Functional typologists believe that the prevalence of these two orders; in which the subject precedes the two other elements, and the verb and the object are contiguous is due to the functions of language in the real world. Hence, the two principles of subject salience and verb-object contiguity have been proposed. The typological explanations put forward for these two principles hold that transitive sentences of a language have come into existence as a result of the encoding of the prototypical transitive action scenario. In such a scenario, subject salience is a result of the fact that the transitive action scenario is started by the doer of the activity. Also, because of the tight causal relationship between the activity and its receiver, the linguistic counterparts of these two elements, too, tend to be contiguous. Since functional pressures can only be manifested in language through human cognition, the present paper looks at the cognitive processes involved in the cross-linguistic prevalence of the afore-mentioned word orders.

Keywords: Word Order; SOV; SVO; Cognition; Analogy; Iconicity.

_____

[1] Assistant Professor, Department of Foreign Languages and Linguistics, Shiraz University, Shiraz. m.kouhestani@shirazu.ac.ir.
[2] Associate Professor, Department of Linguistics, Tarbiat Modares University, Tehran. golfamarsalan@gmail.com (Corresponding Author).

Introduction

Language typology and the study of linguistic universals have been approached from different angles, but there are very few studies that look into the findings of these disciplines from a cognitive science point of view (Auwera & Nuyts, 2007:1074-91). A considerable number of typological studies on the linearization of the three elements Subject (S), Object (O), and Verb (V) point out that in most languages of the world, the basic word order is either SOV or SVO (Dryer, 2013; Greenberg, 1966; Tomlin, 1986). Resorting to communicational needs of human species, functional typologists have tried to provide explanations for such observations (Evans and Green, 2006).

Functional explanations for linguistic phenomena are also cognitive necessarily. Hudson (2014:253) believes that what counts as cognitivism, in linguistic studies, is to posit that the findings of cognitive sciences, including cognitive psychology, can be applied to linguistics. Functional pressures are always of a cognitive nature, and the effects of cognition on language are always functional as well. Considering what has been argued so far, we strive to explain the reason for the fact that SOV and SVO word orders are so prevalent in languages of the world.

A closer look at these word orders, makes it clear that in both, the subject precedes the other two elements (subject salience); also, the verb and the object are close to each other (verb-object contiguity in the SVO and SOV word orders) Greenberg, 1966; Tomlin, 1986; Comrie, 1989; Kemmerer, 2012).

Here, we will try to investigate cognitive factors leading to the cross-linguistic prevalence of SOV and SVO word orders and to achieve this goal, we will benefit from notions like "gestalt perception," "perspective," "attention," "animacy," "analogy," and "sign." In the following sections, we will discuss some theoretical issues first.

Figure and Ground in Prototypical Transitive Action Scenario

When people look at the world around them, they view complex scenes made up of different things. These things are perceived as located on a common background. In fact, despite the fact that the visual input is comprised of many colored dots, the human mind understands things as organized in groups. Psychologists have long been interested in finding the reason. Gestalt psychologists are among those who have proposed explanations for this observation. Gestalt principles are an attempt to formulate regularities upon which the perceptual input is grouped in the form of Gestalts.

In visual perception, these unified forms are regions of the visual field, and parts of these unified forms are perceived as connected pieces which are separated from the rest of the visual field. According to the principles of Gestalt psychology, the human mind divides the scene being viewed into two parts. That part of the scene which attracts more attention is called the "figure" and the other standing in the background is labeled as the "ground" (Koffka, 1935).

In order to analyze the figure and the ground in the prototypical transitive action scenario, we first need to consider such a scenario. The doer of the action is an animate entity who does an activity as a result of which an inanimate receiver becomes affected and undergoes a change in its status. For example, a scene where a

human being slices a loaf of bread is a prototypical transitive action scenario.

The figure and the ground have to be specified in such a scene before anything else. The visual field is divided into two parts. One part of the visual field, the *ground*, has no internal organization and only serves as a background for the realization of the activity. In contrast, the part of the scene that attracts the most attention is called the *figure*. In the case of the prototypical transitive action scenario, the three elements *doer* of the activity*, activity* itself, and *receiver* of the activity work together to form the figure. In other words, these three elements together comprise the internal organization of the prototypical transitive action scenario. This scenario stands in the figure and looks more to the front, more salient, and worthier of attention in the eyes of the viewer.

Sentences as Iconic Signs

Saussure considers the sign as a double entity that comes into existence through the association of two terms (Saussure, 2011). One of these two terms is a *form* that does the signification, and the other is a *meaning* that is signified. These two are inseparable just like the two sides of a paper sheet. The *signifier* is a perceivable, material, and acoustic/visual signal leading to a mental image, and that mental image is the *signified*. Since the signifier and the signified have an associative link in the mind, the linguistic sign is a two-sided mental entity (Ribière, 2008).

Words are not the only linguistic signs. A linguistic sign is an abstract structure instances of which are at work in the language system. Each linguistic sign ought to have a formal component as well as a semantic one. Signs can be both atomic and

composed of other signs. With this in mind, it is obvious that the range of linguistic signs includes morphological structures, syntactic structures, and even whole discursive segments.

With regard to the relationship between the sign and its object, Peirce (1974) puts forward a three-way classification of signs: icons, indexes, and symbols. What concerns our purpose the most here is the icon. Peirce thus defines an icon: "An icon is a representation of what it represents." In simpler words, an icon is a pattern that physically resembles what it stands for. According to Haiman (1994), iconicity is the association of a formal property in the sign with a property in its referent.

There are different types of icons. Imagic icons are the prototypes of icons. An image is a simple sign that resembles its referent. These properties may be visual or auditory. Photographs, statues, etc. are some examples.

Metaphorical icons display a parallelism in something else and hence, distinguish a property of the sign's.

Diagrammatic icons are those that reflect the relations among the different parts of something else within themselves and among their parts. In other words, a diagram is a systematic alignment of signs that do not necessarily resemble their referents, but their relations with each other mirrors the relations holding among their referents. In particular, the configuration of the object and its referent are similar, but the individual signs and their referents need not be similar (Peirce, 1074).

Analogy as Factor behind Formation of Icons

Analogy can be understood in two different senses. It may be understood as a similarity

in which, identical relations hold between the source domain and the target domain. Analogy may also refer to an inference based on which if two things are similar in some respects, they are similar in other respects as well.

Gentner (1983) believes that analogy is important in cognitive science because of a number of reasons. Analogy makes possible transfer from one concept to another, from one situation to another, and from one domain to another and thus, helps explain new concepts. When analogy is acquired, it can be used as a mental model to understand a new domain (Halford, 1993).

What distinguishes analogy from other types of similarity is that for two situations to become analogous, they need to be similar in terms of their relational structures. Gentner (1974) proposes a theory titled *structural mapping*. According to this theory, analogical mapping requires the alignment of the two situations based on their similarities, especially their relational structures. It means that the elements of the two scenes can enter a one-to-one correspondence.

Analogy lets the human mind make conclusions about a target domain. In fact, one reason for resorting to analogy is to learn something about the target domain. This is done through utilizing the person's knowledge of the source domain. Also, if two relations are paired, their arguments should be paired too.

There are principles at work in analogy. One such principle is *consistency*. Everything in the source domain should be paired with one and only one thing in the target domain.

Analogies can be static or dynamic. It is important to distinguish *analogy as structure* from *analogy as process*. Itkonen

(2005) believes that analogy as structure is a static relation between two systems whereas, analogy as process is a dynamic relation that generates analogical structures. Static analogy refers to relations that hold between two pre-existing structures, but dynamic analogies are those that generate the target domain based on the source.

In the remainder of this paper, we will explain how the viewer looks at the prototypical transitive action scenario and perceives it as the figure against a ground. Also, we will explain that SVO and SOV word orders are iconic representations of the afore-mentioned scenario which are generated as the outcomes of two instances of dynamic analogy.

Role of Perspective and Selective Attention in Salience of Doer of Activity

Cognitive linguists have acknowledged the importance of attention as one of the most important cognitive processes. This is because attention, as a mental process, plays a very significant role in language. Langacker (1987) believes that attention is intrinsically related to the intensity or energy level of cognitive processes and leads to more importance or salience. From among the many cognitive processes leading to the richness of the mental experience, some receive added attention and become salient as the focus of attention.

Langacker believes that while watching a scene in the world, attention is unequally paid to one aspect of the scene. He claims that this is due to a number of focal settings. In fact, the human viewer imposes a unique construal upon the scene through a focal setting and linguistic organization.

The three parameters of focal setting are *selection*, *perspective*, and *abstraction*. We will pay more attention to perspective and

explain how the selection of a specific perspective will eventually lead to subject salience in SOV and SVO word orders.

In any scene, the perspective is of utmost importance in the salience of the participants relative to one another. The grammatical roles subject and object reflect the perspective to the scene and are rooted in human cognition. The fact that in SOV and SVO word orders, the subject appears in the beginning means that the referent of this grammatical element receives more attention than those of other elements.

Ungerer and Schmid (2013) believe rather than being a syntactic idea, perspective is more of a cognitive nature, and what underlies the selection of various perspectives is the cognitive ability to direct attention to things. They believe the perspective from which we look at a situation in the world depends on what attracts our attention.

The person who is watching the prototypical transitive action scenario selectively directs their attention to that element of the scene which bears more importance to them. As a result, other elements of the scene receive less attention. However, this does not mean that the human being consciously considers something as more important.

Animacy as Attention Attracting Feature

The feature-present/feature-absent effect plays a significant role in attracting attention. In other words, if some feature is present (positive information), it can attract more attention to the whole of which it is a part (Treisman, 1980; Treisman ey al., 1985; Treisman et al., 1988). Franconeri, Hollingworth (2005), Wolfe (2000) and (2001) claim that when people look for a feature, the feature in question attracts their

attention more easily. The question here is which feature of the doer of the activity makes it worthy of more attention in the prototypical transitive action scenario.

We argue that the most important feature that attracts the scene viewer's attention is animacy. In the prototypical transitive action scenario, the doer of the activity is animate.

In the course of human evolution, identifying animate entities in the visual field has been of great importance. One of the most significant features of animals is their ability to move around. The human ability to detect motion in animals has led to the identification of potential prey and predators and has helped the Homo sapiens survive up until this time.

Legerstee (1991) investigates the role of the person and the object in eliciting infants' imitation. A group of infants were exposed to two different conditions. In the first, the infants were exposed to tongue protrusions and mouth openings modeled by an adult. In the second though, the same infants were exposed to these gestures simulated by two objects. When these infants encountered human gestures, they imitated the same to a considerable degree, whereas when they encountered inanimate objects, there was no imitation by the infants. It can be concluded that the human being is able to differentiate humans (as animate entities) from objects from the beginning of their life.

Crichton and Lange-Küttner (1999) presented infants who were 16 to 20 weeks old with objects moving across a 60-cm distance. The researchers examined four conditions: induced movement while holding the object, induced movement while pushing the object, self-propelled mechanical movement in which the object moved by an internal clockwork, and self-

propelled biological movement where an animate object moved by internal impulse. The study found out that animate objects with self-propelled movement could easily be distinguished from inanimate ones. Again, we see that the human infant prefers to direct its attention to animate objects than to inanimate ones.

Pratt, Radulescu (2010) designed an experiment to study how differently human beings reacted to animate and inanimate motions. Inanimate motion is that which can be predicted such as motion as a result of two objects colliding. In contrast, the animate motion happens in an unpredictable way. The experiment showed that the subjects reacted more quickly to animate motion than to inanimate motion. The researchers conclude that this difference is due to the fact that animate motions are perceived as the movements of animate entities, and to the human species, such movements are much more important than the other type.

Another significant piece of evidence regarding the special attention humans pay to animate entities comes from comparing the human ability to detect changes in animate and inanimate entities. One such research was conducted by New, Cosmides (2007). They believe that mechanisms of visual attention select information about animals for processing whereas it is not so for inanimate entities. They conducted experiments to see whether human attention is more sensitive to changes in animals or to those in inanimate objects. The results showed that changes in animals are detected more easily and more accurately than those in cars, buildings, tools, and even plants. This is in spite of the fact that in the modern world, vehicles pose more numerous and more dangerous risks to humans than animal predators. This is due to the importance of identifying predators for the ancestors of humans and that today's humans have not had enough time to adapt to the modern world they themselves have created.

These studies demonstrate that from a very early age, the human being is able to distinguish between animate and inanimate entities and systematically pays more attention to animate entities. So, it might be safe to claim that animacy is the feature present in the doer of the activity that makes it so worthy of attention.

Role of Proximity Principle in Grouping of Activity and its Receiver

There are a number of principles in Gestalt psychology based on which different things in a scene are grouped together. As mentioned before, the human viewer divides the scene into two parts: the figure and the ground. In the prototypical transitive action scenario, the figure is composed of three elements: the doer of the activity, the activity, and the receiver of the activity.

As mentioned in the previous section, the doer of the activity which is the animate initiator of the scenario and attracts the most attention. As a result, it becomes more salient than the other two elements. In a prototypical transitive action scenario, the receiver of the activity undergoes a change as a result of the activity initiated by the doer. There is no doubt that the receiver of the activity is related to the doer through the activity. So, it is logical to say that the relationship between the activity and the receiver is a close one. However, this closeness is not of a physical nature. What leads to the grouping of the activity and its receiver is of a conceptual nature. We believe that the proximity principle in the

Gestalt grouping should be expanded to include conceptual proximity as well.

Role of Analogy in Conceptualization of Prototypical Transitive Action Scenario

In what can be considered a prelude to analogical mapping, the human viewer separates the figure from the ground then proceeds to identify its parts, i.e. the doer, the activity, and the receiver.

If the static analogy were at work, a structural alignment would take place – that is, similar relational structures in the source domain and the target domain would be detected and a one-to-one correspondence would be establish between their parts. But things are different here. The analogy is of a dynamic nature, and the non-existent target domain is constructed based on the already present source domain. Thus, the relational structure of the target domain is a reflection of that in the source domain.

The first step in the mapping phase of a dynamic analogy is to project the relational structure of the source domain onto the target domain to be constructed. In this phase, both the comprising parts of the source and relations among them are projected onto the target domain, and the target domain comes into existence consequently.

After the activity, its doer, and its receiver are identified, their relations to one another need to be recognized as well. Among these three elements that make up the figure, two elements (the doer and the receiver) are of a material nature and one (the activity) is energy. It is the activity that links the two material elements of the scene. Nevertheless, the two material elements located at the two ends of the activity are not of the same status. The doer is the initiator of the scenario. If it were absent, there would be no scenario in the first place. This leads the human viewer of the scene to assign a higher status to this entity. However, once the activity is initiated by the doer, there comes into existence a relationship between the activist and the entity receiving it. The change brought about in the receiver happens right after the activity is applied to it. The human mind understands this cause and effect relationship and groups these two elements together.

After the elements of the scene are identified and the relations among them are detected, analogy comes into play to create the target domain (the mental concept of the scene). Each relation in the source domain should correspond to one and only one relation in the target domain. As discussed before, the two material elements are linked to each other via the activity. Thus, a relation forms in the target domain in which two entities are linked to each other by means of another element. But at this stage, the elements in the source domain have no counterparts in the target domain.

The elements created in the mental concept of the scene based on the comprising elements of the figure are as follows: agent as the conceptual counterpart of the doer, action as the counterpart of the activity, and patient as the counterpart of the receiver.

Since the relations among the elements were mapped onto the target beforehand, at this stage each element finds its place in the relational structure. This is how the mapping of the relational structure of the real-world event onto the mental concept becomes possible.

When the analogical mapping is taking place, the perceptual salience of the doer of the activity reflects itself in the mental

concept of the scene. In the target domain of this analogy, the conceptual counterpart of the doer benefits from a status separate from the other two elements. It also situates itself before them.

The close relationship between the activity and its receiver too needs to be reflected in the mental concept. This is accomplished by grouping together the action and the patient in the resulting concept.

What we have, thus far, said about the salience of the agent and the proximity of the action and the patient is all in the realm of human mind. But subject salience and verb-object contiguity are linguistic phenomena and need explanations of their own.

Role of analogy in Formation of SOV and SVO Word Orders

Here, we strive to explain how analogy maps the mental concept of the scene onto the language and how this mapping leads to the formation of SOV and SVO word orders.

After the concept of the scene is formed, analogy comes into play again and converts the concept into language. This second analogy too is dynamic because again, the source domain (this time the mental concept) is there, but the target domain (which should be linguistic in nature) is not formed yet.

Once more, the elements of the source domain should be identified. Each relation in the source domain should correspond to one and only one relation in the target domain. Thus, a relation forms in the target domain in which two entities are linked to each other by means of another element. At this stage, the elements in the source domain have no counterparts in the target domain. The subject as the linguistic counterpart of

the agent, verb as the counterpart of the action, and the object as that of the patient are formed in the target domain – i.e. the linguistic structure.

In order to find out how the afore-mentioned mental concepts lead to the creation of SOV and SVO word orders, we should look more carefully at the analogical mapping of the agent, action, and the patient in the concept.

We explained that in the mental concept of the prototypical transitive action scenario, the agent is conceptualized before and apart from the other two elements to mirror the more attention directed to the doer of the activity. Furthermore, the cause and effect relationship between the activity and its receiver is reflected through a close relationship between the action and the patient in the mental concept of the scene.

The salience of the agent in the mental concept of the prototypical transitive action scenario leads to the initial positioning of the subject. Also, in order to reflect the conceptual closeness between the action and the patient in language, their linguistic counterparts (the verb and the object) appear close to each other. In the mental concept, it was the mere closeness of the action and the patient that mattered regardless of their order. Consequently, the verb and the object can appear in either possible order. The outcome of this mapping is subject salience as well as verb-object contiguity. This explains as why SOV and SVO word orders are so prevalent in the languages of the world. They are icons that signify events in real-world scenes.

Conclusion

In order to explain why SOV and SVO word orders are so prevalent cross-linguistically, we first decomposed these two orders to

their common characteristics. It has been argued that subject salience and verb-object contiguity can be observed in both of them. We explained that these orders are diagrammatic iconic signs that reflect the human's conceptualization of the prototypical transitive action scenario. This happens through the application of dynamic analogy at two stages. First, a concept of the scenario forms in the mind of the viewer, and in the second stage, the concept is converted to sentences in language. Because the sentences tend to be iconic, the subject comes first and the verb and the object remain contiguous in both word orders to reflect the importance ascribed to the doer activity and the causal closeness of the activity and its receiver.

References

[1] van der Auwera J, Nuyts J., (2007). Cognitive linguistics and linguistic typology. *Handbook of Cognitive Linguistics* Oxford University Press, Oxford. 1074-91.

[2] Dryer MS., (2013). *Order of Subject, Object and Verb Leipzig*: Max Planck Institute for Evolutionary Anthropology

[3] Greenberg J., (1966). Some universals of grammar with particular reference to the order of meaningful elements. In J Greenberg, ed, *Universals of Language* 73-113 Cambridge, MA.

[4] Tomlin RS., (1986). B*asic word order: Functional Principles*: Croom Helm London.

[5] Evans V, Green M., (2006). *Cognitive linguistics: An Introduction*: Edinburgh University Press.

[6] Hudson R., (2014). *Cognitive functionalism in language education. Theory and Practice in Functional-Cognitive Space*. 68:253.

[7] Comrie B., (1989). *Language Universals and Linguistic Typology: Syntax and Morphology*: University of Chicago Press.

[8] Kemmerer D., (2012). The Cross-Linguistic Prevalence of SOV and SVO Word Orders Reflects the Sequential and Hierarchical Representation of Action in Broca's Area. *Language and Linguistics Compass*. 6(1):50-66.

[9] Koffka K., (1935). *Principles of Gestalt Psychology, International Library of Psychology, Philosophy and Scientific Method*. Harcourt Brace, New York.

[10] Saussure FD., (2011). *Course in General Linguistics*. New York: Columbia University Press; 2011.

[11] Ribière M., (2008). B*arthes: Humanities-Ebooks*.

[12] Peirce CS., (1974). *Collected Papers of Charles Sanders Peirce*: Harvard University Press.

[13] Haiman J., (1994). Iconicity and Syntactic Change. *The Encyclopedia of Language and Linguistics*. 3.

[14] Gentner D., (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science* 7(2):155-70.

[15] Halford GS., (1993). *Children's Understanding: The Development of Mental Models*.

[16] Itkonen E., (2005). *Analogy as Structure and Process: Approaches in linguistics, Cognitive Psychology and Philosophy of Science*: John Benjamins Publishing; 2005.

[17] Langacker RW., (1987). *Foundations of Cognitive Grammar: Theoretical Prerequisites*: Stanford University Press; 1987.

[18] Ungerer F, Schmid HJ., (2013). *An Introduction to Cognitive Linguistics*: Routledge.

[19] Treisman AM, Gelade G., (1980). A Feature-integration Theory of Attention. *Cognitive Psychology* 12(1):97-136.

[20] Treisman A, Souther J., (1985). Search Asymmetry: a Diagnostic for Preattentive

Processing of Separable Features. *Journal of Experimental Psychology* 114(3):285.

[21] Treisman A, Gormican S., (1988). Feature Analysis in Early Vision: Evidence from Search Asymmetries. *Psychological Review* 95(1):15.

[22] Franconeri SL, Hollingworth A, Simons DJ., (2005). Do New Objects Capture Attention? *Psychological Science* 16(4):275-81.

[23] Wolfe JM., (2000). *Visual Attention. Seeing.* 2:335-86.

[24] Wolfe JM., (2001). Asymmetries in Sisual Search: An Introduction. *Perception & Psychophysics* 63(3):381-9.

[25] Legerstee M., (1991). The Role of Person and Object in Eliciting Early Imitation. *Journal of Experimental Child Psychology* 51(3):423-33.

[26] Crichton MT, Lange-Küttner C., (1999). Animacy and Propulsion in Infancy: Tracking, Waving and Reaching to Self-propelled and Induced Moving Objects. *Developmental Science* 2(3):318-24.

[27] Pratt J, Radulescu PV, Guo RM, Abrams RA.. (2010). It's Alive! Animate Motion Captures Visual Attention. P*sychological Science.* 21(11):1724-30.

[28] New J, Cosmides L, Tooby J.. (2007). "Category-specific attention for animals reflects ancestral priorities, not expertise". Proceedings of the National Academy of Sciences 104(42):16598-603.

عوامل شناختی مؤثر بر رواج ترتیب بنیادین سازه‌ها SOV یا SVO

منوچهر کوهستانی¹، ارسلان گلفام²

چکیده

یکی از مشاهدات رده‌شناختی این است که در بیش از نود درصد از زبان‌ها، ترتیب بنیادین سازه‌ها SOV یا SVO است. رده‌شناسان نقش‌گرا بر این باورند که رواج این دو آرایش سازه‌ای که در آنها فاعل پیش از دو عنصر دیگر قرار می‌گیرد و فعل و مفعول درکنار هم‌اند، به دلیل عملکردهای زبان در جهان واقع است. بنابراین دو اصل برجستگی فاعل و مجاورت فعل و مفعول را پیشنهاد داده‌اند. تبیین‌های نقش‌گرایانهٔ ارائه شده برای این دو اصل ادعا می‌کنند که جمله‌های متعدی زبان در نتیجهٔ رمزگذاری سناریوی کنش متعددی پدید آمده‌اند. در چنین سناریویی برجستگی فاعل به دلیل آن است که سناریوی کنش متعددی از سوی کنندهٔ فعالیت آغاز می‌شود. همچنین به‌دلیل رابطهٔ تنگاتنگ علی بین فعالیت و دریافت کنندهٔ آن متناظرهای زبانی این دو عنصر نیز گرایش دارند که در مجاورت یکدیگر قرار گیرند. ازآنجاکه فشارهای عملکردی فقط از طریق شناخت بشر می‌توانند در زبان تجلی یابند، این مقاله به فرآیندهای شناختی درگیر در رواج میان زبانی دو آرایش سازه‌ای پیش‌گفته می‌پردازد.

واژه‌های کلیدی: آرایش سازه‌ای، SVO، SOV، شناخت، قیاس، شمایل‌گونگی

¹. استادیار زبان‌شناسی، دانشگاه شیراز، شیراز، ایران m.kouhestani@shirazu.ac.ir

۲. دانشیار زبان‌شناسی، دانشگاه تربیت مدرس، تهران، ایران golfamarsalan@gmail.com (نویسندهٔ مسئول)