"OVS" – A misnomer for SVO languages with ergative alignment

Hubert Haider

Dept. of Linguistics & Centre for Cognitive Neuroscience, Univ. Salzburg

Abstract

Unequivocal OVS languages are not attested, contrary to claims in the literature. Languages filed as OVS in typological descriptions turn out to be SVO languages with ergative alignment that have been misclassified due to ill-chosen diagnostics for "S" and "O".

Introduction

Klingon[1] seems to be the only nominative-accusative language with uncontroversial object-verb-subject word order. This is how it has been designed in order to seem and sound alien. Natural languages with this property arguably do not exist. Languages that have been filed as OVS are ill-classified because of the inappropriate characterization of "S", as will be shown below. They are either *ergative SVO* languages or have been prematurely classified as OVS.

Greenberg's (1963) original sample of thirty languages contained only two languages that he classified as OVS (with VOS as alternative word order), namely Siuslaw and Coos (s. Greenberg's Appendix II). How did he arrive at this classification? He looked at simple sentences and the order of "meaningful elements", namely agent, action word and theme (aka 'patient') in simple transitive clauses with an agentive verb. He is very clear about his – preliminary – recourse to easily applicable criteria for the identification of the subject of a clause and that he is aware that his strategy is just a time-saving shortcut:

> "*I fully realize that in identifying such phenomena in languages of differing structure, one is basically employing semantic criteria. There are very probably formal similarities which permit us to equate such phenomena in different languages. However, to have concentrated on this task, important in itself, would have, because of its arduousness, prevented me from going forward to those specific hypotheses.*" Greenberg (1963:74).

A popular on-line encyclopaedia[2] characterizes "*Ergative-absolutive languages, sometimes called ergative languages*" as "*languages where the subject of an intransitive verb and the object of a transitive verb behave the same way in a sentence. Both behave differently to the subject of a transitive verb.*" Evidently, this way of describing the grammatical circumstances is biased towards the majority of languages with nom-acc alignment. The verbal argument that is "*the object of a transitive verb*" is an *object* only in a nominative-accusative setting. In an absolutive-ergative system it is the *subject* of the clause.

A more appropriate rendering would be this: In general, one of the two arguments of a transitive verb, viz. argument A and B, is aligned with the syntactic subject function and the other one with the grammatical function of the direct object. Consequently, this opens a system space for two systems. In one system, argument A is the subject, in the other system, argument B is the subject. In each system, the remaining argument will end up as the direct object.

---

[1] A language constructed by Marc Okrand for the Star-Trek movies as the language of the fictional Klingons. According to Okrent (2009: 273), meanwhile more than a dozen people claim to be fluent in this language.
[2] Wikipedia on "Ergative-absolutive language".

If A is the *agent* argument of an agentive verb and it is aligned with the subject function, the alignment system is called *nom-acc*. If B is the *non-agent* argument of an agentive verb and it is aligned with the subject function, the alignment is called *abs-erg*.

The syntactical subject is the morpho-syntactically privileged noun phrase aligned with an argument of the verbal predicate. 'Privileged' means amongst others that it agrees with the finite verb in languages with subject-verb agreement; it appears in a preverbal, structurally obligatory position in genuine [S[VO]] languages; it cannot be omitted without signalling this morpho-syntactically (i.e. passive in Nom-Acc languages, in direct correspondence to anti-passive languages with absolutive-ergative alignment).

If the alignment system is a system with structural cases, then a dependency relation holds between the assignment of subject case and the direct-object case. The object case is assigned only in the presence of the subject case (Haider 2000). If the primary subject candidate is syntactically unavailable, subject case is passed on and assigned to the object. The consequence is the familiar acc-to-nom (= object-to-subject) switch (1) or an ergative-to-absolutive (= non-subject-to-subject) switch (2), respectively, in passive.

(1) a. La mère a encouragé les filles   French
      the mother$_{Subj}$ encouraged the daughters$_{Obj}$
   b. Les filles$_{Subj}$ ont été encouragées (par la mère)
      the daughters have been encouraged (by the mother)

(2) a. ʔaaček-a   kimitʔ-ən   ne-nlʔetet-ən   Chukchi (Kozinsky et al. 1988: 652)
      youth-erg   load-abs   3pl.subj-carry-aor.3sg.obj
   b. ʔaaček-ət   ine-nlʔetet-gʔe-t   kimitʔ-e
      youth-abs   antip-carry-aor.3sg.subj-pl   load-instr
      'The young men carried away the/a load.'

Polinsky (2013) characterizes passive and antipassive in the common way that focuses on a circumstantial property, however: "*In the passive, the suppressed or demoted argument is the agent-like argument, in the antipassive, the patient-like argument*." The property that is syntactically relevant is not the thematic role of the demoted argument, it is the syntactic role, and this role is the grammatical function of a subject. Whenever the argument of the verb that would otherwise surface as a syntactic subject, viz. as nominative or absolutive, respectively, in a finite clause, is syntactically left out, this must be morpho-syntactically signalled. In French (1), the combination of a particular verb form (participle) and a particular auxiliary ('be'-type auxiliary instead of the 'have'-type one), blocks the primary subject argument. This is the way how Indo-European languages typically implement the syntactic elimination of the subject argument. Since French is a language with an [S[VO]] clause structure, argument that is direct object in (1a) surfaces in the position of the subject in (1b) and enters the typical agreement relation with the finite verb. In Chukchi (2), a prefix[3] of the verb is the morphological signal of the syntactically suppressed would-be subject argument. As a consequence, the ergative-marked noun phrase of (2a) switches its case to absolutive as the subject case.

---

[3] In Indo-European languages, affixal passive is suffixal, as for instance Latin ("-*ur*") or the Scandinavian S-passive with the suffix "-s", which is the continuation of the cliticized reflexive pronoun of a middle construction.

2. What matters is the syntactical function not the thematic quality

Content is easy to grasp; structure is hard to assess, but syntactically, structure matters more than content. Grammars define structures, and structures constrain the form of the presentation of content. Haspelmath (2014: 494) justly asks: "*Saying that Japanese generally has SOV order while English has SVO order is far more problematic, because it seems to presuppose that we can identify subjects, objects and verbs, i.e. abstract syntactic categories, in both languages. But on what basis*?" Regrettably, instead of insisting on the inevitability of providing structurally sound definitions as the basis of any comparative grammar, he suggests to stick to Greenberg's preliminary, and by now outdated, shortcut approach.

> "*The basic principle is* [...] *that languages can be readily compared only with respect to meanings and sounds/gestures, but not with respect to their categories, because only meanings and sounds, but not categories, are universal. Thus, instead of saying that English has SVO order, while Japanese has SOV order, we must say that English has agent-action-patient order, while Japanese has agent-patient-action order. This is not the normal notation.*" (Haspelmath 2014: 495).

Such a strategy is not promising. What matters is not "readily" but "correctly".[4] Languages can and in fact must be compared "*with respect to their categories*", but only after having ensured that one is comparing identical categories. This is exactly *not* what we do if we compare "agent-patient" order. Here we are at the core issue of this squib. If languages are compared "*only with respect to meaning*", the outcome is a confusing maze of patterns, since syntax – the result of evolutionary grammaticalization (Haider 2020) – overrules semantic distinctions.

The simplest case of a syntax-semantics incongruence is the categorization of the parts of speech. Many languages distinguish nouns, verb, and adjectives (3). Semanticists convince us that each form is basically a predicate. We know that the syntactic category determines the syntactic behaviour, and not so much the semantic content, since even the same content can be categorized differently, as for instance in (3). The verbal form behaves syntactically like a verb (1a), the nominal form in (3b) behaves syntactically like a noun, and the participial form like an adjective. The shared semantic content does not determine the syntactic behaviour.

(3) a. vom theoretischen Wert *abweich*-en$_{V°}$ German
    from-the theoretical value deviate-*Inf*
  b. ein [vom theoretischen Wert *abweichen*-d-er$_{A°}$]$_{AP}$ Wert
    a [from-the theoretical value deviate-*participle-Agr*] value
  c. das [*Abweich-en*$_{N°}$ des Wertes vom theoretischen Wert]$_{NP}$
    the [deviate-Inf the value$_{Gen}$ from-the theoretical value]

Analogously, the fact that a noun phrase represents the agent-argument does obviously not fully determine its syntactic status. In each of the following examples in (4), an agentive NP is present. However, nobody would claim that it is the syntactic subject of the clause in each case.

(4) a. *The president*/*He* messed up their lives.
  b. Their lives were messed up by *the president/him.*

---

[4] Comparative biology compares homologically. Analogous structures are seen as the result of convergent evolution. It is the structure that determines function (see Haider 2020).

    c. We shouldn't let *the president/him* mess up their lives.

Why are typologists convinced that the noun phrase that denotes the agent of the verb should universally be identified as subject in word order typologies? Almost sixty years after Greenberg (1963), who was unaware of different alignment systems, and in particular, of nom-acc versus abs-ergative systems of alignment, we know it better, and nevertheless languages are still classified following his preliminary and partially misleading way.

3. "OVS" languages are SVO with ergative alignment

Let's remember that the two "OVS" languages of Greenberg's original sample are ergative languages, that is, SVO languages with ergative alignment. Fifty years later, in WALS, Dryer & Haspelmath (2013) list the following languages as "OVS". *Four* of them have an ergative case-system (Kuikuro, Mangarrayi, Päri, Tuvalan). *Five* are caseless (i.e. 'neutral') but show ergative properties (Asurini, Hixkaryána,[5] Selknam, Tiriyo, Ungarinjin). The two remaining languages are Kxoe and Urarina. This means that structurally, nine of the eleven languages are "SVO" languages, modulo ergative alignment. As for the two alleged nominative-accusative OSV languages, the evidence is questionable.

For Urarina, Olawsky (2007: 45), who published a comprehensive grammar of this language, notes: "*The language has a nominative-accusative system but case is marked by constituent order only.*" How can one be sure that the system is nominative-accusative in Urarina if all we have is a semantically identified constituent order? Passive is inconclusive in this language since it is formed periphrastically through a nominalized verb functioning as a copular complement. The 'passivized' verb can take nominal morphology. However, there is an intransitivizer, viz "*ne-*" that produces O>S derivations of transitive verbs; see Olawsky (2006: 600), Muysken et als. (2016, Feature ARGEX8-1). In a Nom-Acc system, an intransitivizer is expected to produce SO>S, but not O>S. In ergative languages, an intransitivizer is expected to produce S,A>S, which in typological terminology is O>S. This seems to be exactly what happens in Urarina.

For Kxoe, Fehn's (2015:214) grammar of Grammar of Ts'ixa (Kalahari Kxoe) is very clear: "*There are three patterns available for transitive clauses: AOV, AVO and OAV, with the latter occurring less frequently than the other two. Although the dominant word order of the Khoe languages is thought to be AOV (cf. Heine 1976, Güldemann 2014), AVO is just as frequent.*" The type-assignment in WALS exclusively follows Köhler (1981). Kxoe is not a reliable testimony of OVS.

In a study on word order type and alignment, Siewierska (1996) lists four languages as "OVS" in her own sample of 237 languages, namely Makushi, Hixkaryana, Pari and Southern Barasano. Päri is an ergative language, according to Andersen (1988). Makushi is ergative according to Abbot (1991). Hixkaryana has been mentioned already above. As for Southern Barasano, Jones & Jones (1991) presented a syntax monograph which was reviewed by Dryer (1994). He points out a crucial weakness of their type assignment:

---

[5] According to Derbyshire (1979), an *object* receives the same morphology as an *intransitive* subject when verbs take on derivational morphology. This is an ergative feature, with separate morphology for objects and transitive subjects. As Birchall (2014:101) emphasizes, "*two commonly occurring verbal marking patterns in South American languages that are difficult to characterize as strictly ergative or accusative: hierarchical marking and split intransitivity.*" Kalin (2014) tries to motivate an SOV-based analysis with VP fronting.

"A count of all examples in the grammar shows both SV and VS order common, with SV slightly more common, though numbers of examples cited in a grammar is a poor source of data. But the frequency of SV examples both in the grammar and in the text examined does suggest that the claim that subjects tend to follow the verb is based on both noun and pronoun subjects rather than just noun subjects. If we interpret the notion of an OVS language as referring to clauses with a noun object and a noun subject (the standard usage in word order typology), it is not clear that Barasano qualifies." *"It is possible that it is best treated as indeterminately SOV/OVS, a word order type that appears to be quite common in the Amazon basin.* (Dryer 1994: 63).

In sum, there is no compelling evidence for OVS from non-ergative languages. For ergative languages, "OVS" means Abs-V-Ergative order, and this is *Subject-verb-object* order, under ergative alignment.

4. Properties shared by ergative languages

Siewierkska (1996:149) identifies and summarizes the following positions arrived at in the literature, based on a semantic definition of subject and object. Typologists agree that there is "*an association between ergative alignment and non-SVO order*" and "*an association between ergative alignment and object-before-subject order*". This would be surprising, given the fact that SVO is a major type.

In the literature Siewierska refers to, "SVO" and "OS" order is understood as "*Agent-Action-Patient*" and "*Patient-Agent*" order, respectively. What this neglects is the fact that 'ergative alignment' ought to be read as follows: The argument of a transitive verb that is a *direct object* in nominative-accusative alignment is the *syntactic subject* under ergative alignment, if 'subject' is construed *grammatically*. If we apply structural criteria, Siewierska's findings turn out as expected, straightforward, and cross-linguistically uniform properties of syntactic subjects across alignment systems.

First, an *ergative* language that would 'semantically' be identified as "SVO" is structurally an *OVS* language, with ergative-V-absolutive order. Structural OVS languages, however, are extremely rare if not inexistent.

Siewierska's *second* point, the "object-before-subject order" of ergative languages is in reality the Patient-before-Agent order, or Absolutive-before-ergative. Structurally, in ergative languages, this is subject-before-object, that is, the noun phrase with absolutive case precedes the noun phrase with ergative case. This – nominative before accusative – is the common serialization in Nominative-Accusative languages as well. Subjects precede objects. In sum, ergative languages pattern just like Nom-Acc-languages, modulo alignment, with SOV and SVO as the most frequent types. The allegedly non-existent "ergative SVO" do exist, as ergative languages that have been misidentified as OVS languages. The alleged "object-subject" order of ergative languages is in fact the cross-linguistically pervasive subject-object order, modulo ergative alignment. Hence, there is no reason for being surprised that an ergative "agent-V-patient" language, which would in fact structurally be an OVS language, has not been detected and presumably does not exist.

4. Conclusion

The *structural* identification of grammatical functions is the necessary, proper, and inevitable basis for cross-linguistic comparisons. 'Semantic' classifications of grammatical relations obviously lead astray. They rest on a hidden but wrong premise, namely, that universally, for verbs with an agent and a patient argument, the agent argument is the subject in a 'plain'[6] clause. This is true for Nom-Acc-languages, but crucially not for languages with ergative alignment. The equation of Agent with Subject works for Nom-Acc languages, but not for Abs-Erg languages. In these languages, the patient of a transitive verb is the grammatical subject. If one compares Agent-V-Patient patterns cross-linguistically, one compares the *subject* of Nom-Acc systems with a non-subject of Abs-Erg systems. It is not astonishing at all that such ill-defined "subjects" do not share relevant *grammatical* properties. If compared properly, that is, structurally, several puzzles disappear:

- Structural OVS languages are not only rare; they are virtually inexistent.
- The "OVS" languages listed in typological literature are SVO languages with ergative alignment.
- SOV and SVO is the most frequent type, both for nom-acc as well as for abs-erg alignment.

Bibliography

Abbott, Miriam. 1991. Macushi. In Derbyshire, Desmond C. and Pullum, Geoffrey K. (eds.), Handbook of Amazonian Languages 3. 23-160. Berlin: Mouton de Gruyter.

Andersen, Torben. 1988. Ergativity in Parï, a Nilotic OVS language. *Lingua* 75: 289-324.

Birchall, Joshua. 2014. *Argument marking patterns in South American languages.* LOT Dissertation series, number 352.

Derbyshire, Desmond C. 1979. *Hixkaryana*. (Lingua Descriptive Studies 1) Amsterdam: North-Holland.

Dryer, Matthew S. 1994. Review of Wendell Jones and Paul Jones, Barasano Syntax. *Word* 45(1): 62-65. DOI: 10.1080/00437956.1994.12098343

Fehn, Anne-Maria. 2015. *A Grammar of Ts'ixa* (*Kalahari Khoe*). Unpubl. doct. diss. Univ. Köln, Inst. f. Afrikanistik

Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In Joseph H. Greenberg (ed.). *Universals of Language*. 73-113. Cambridge, Mass.: MIT Press.

Haider, Hubert. 2000. The license to license. In Reuland, Eric (ed.) *Argument & Case: Explaining Burzio's Generalization*. 31-54. Amsterdam: Benjamins.

Haider, Hubert. 2020. Grammar change – a case of Darwinian cognitive evolution. Target paper for *ELT – Evolutionary linguistic theory*. [lingbuzz/005134]

Haspelmath, Martin 2014. Comparative syntax. In Carnie, Andrew, Yosuke Sato, and Dan Siddiqi (eds.) *The Routledge handbook of syntax*. 490-508. London: Routledge.

Jones, Wendell & Jones, Paul. 1991. *Barasano Syntax*. Studies in the languages of Colombia 2. Summer Institute of Linguistics & University of Texas at Arlington.

---

[6] 'Plain' means: non-passivized, non-middle, etc., or, in other words, non-derived form.

Kalin, Laura. 2014.The syntax of OVS word order in Hixkaryana. *Natural Language & Linguistic Theory* 32: 1089–1104. DOI 10.1007/s11049-014-9244-x

Köhler, Oswin. 1981. La langue kxoe. In Perrot, Jean (ed.) *Les langues dans le monde ancien et moderne*. 483-555. Paris: Centre National de la Recherche Scientifique.

Kozinsky, Isaac and Nedjalkov, Vladimir and Polinskaja, Maria. 1988. Antipassive in Chukchee. In Shibatani, Masayoshi (ed.), *Passive and Voice*, 651-706. Amsterdam: John Benjamins.

Mithun, Marianne. 2000. Ergativity and language contact on the Oregon Coast: Alsea, Siuslaw and Coos. In Andrew K. Simpson (ed.), *Proceedings of the Twenty-Sixth Meeting of the Berkeley Linguistics Society* (Special issue on the Syntax and Semantics of the Indigenous Languages of the Americas), 77-95. Berkeley: Berkeley Linguistics Society.

Muysken, Pieter, Harald Hammarström, Olga Krasnoukhova, Neele Müller, Joshua Birchall, Simon van de Kerke, Loretta O'Connor, Swintha Danielsen, Rik van Gijn & George Saad. 2016. *South American Indigenous Language Structures (SAILS) Online*. Jena: Max Planck Institute for the Science of Human History. (https://sails.clld.org).

Okrent, Arika. 2009. *In the land of invented languages: Esperanto rock stars, Klingon poets, Loglan lovers, and the mad dreamers who tried to build a perfect language*. New York: Spiegel & Grau.

Olawsky, Knut J. 2006. *A Grammar of Urarina.* (Mouton Grammar Library, 37) Berlin: Mouton de Gruyter.

Olawsky, Knut J. 2007. ObViouS OVS in Urarina syntax. In: Simpson, A. & P. Austin (eds.), *Endangered languages*. Special volume of *Linguistische Berichte*. 45-72. Hamburg: Buske Verlag.

Polinsky, Maria. 2013. Antipassive Constructions. In: Dryer, Matthew S. & Haspelmath, Martin (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info/chapter/108, Accessed on 2021-01-10).

Siewierska, Anna (1996). Word order type and alignment type. *STUF - Language Typology and Universals* 49(2): 149-176.