A comparison on orders of
S, O, V and R, T, V

Bingfu Lu (Shanghai Normal University)

Tianhua Luo (Nanchang University)

The result of a survey about the orders of R(ecipient), T(heme) and V(erb) reveals some interesting similarities and differences to that of S, O and V. The survey was based on 90 languages in China. The basic data are as follows:

Orders of S, O, V (Dryer 2005):

SOV	SVO	VSO	VOS	OVS	OSV
497	435	85	26	9	4

Our data of R, T and V:

RTV	RVT	VRT	VTR	TVR	TRV
35	3	19	6	0	0

If we classify the two sets into three groups with V as the reference point, the similarities and differences becomes clearer as follows.

V-final		V-initial		V-middle		lacking a basic order
RTV	TRV	VRT	VTR	RVT	TVR	
35	3	19	6	0	0	27

Table 1 (Lu & Luo 2007)

V-final		V-initial		V-middle		lacking a basic order
SOV	OSV	VSO	VOS	SVO	OVS	
497	4	85	26	435	9	172

Table 2 (Dryer 2005)

In the following, we will try to explain the similarities and differences of the two sets.

(1) V-final ordering

Data comparison:
SOV is overwhelmingly dominant over OSV. The ratio SOV/OSV is 497/4=124.25. Similarly, RTV is absolutely dominant over TRV. The corresponding ratio is 35/3=11.67. Both ratios are over 10.

Explanation:

The overwhelming dominance of SOV/RTV over OSV/TRV can be explained with the interaction of two principles of linear ordering: Linear Precedence (LP) and Semantic Proximity (SP).

LP claims that there is a hierarchy of precedence tendency among arguments. The hierarchy is motivated by several factors, including definiteness, animacy, given-new information, etc. They could constitute a natural class 'identifiability'. Specifically, LP expects that the more identifiable the referent of an argument is, the more strong is its tendency to precede.

SP claims that there is a hierarchy of proximity to the verb among arguments.

LP favors the orders [S…O] and [R…T]. SP favors the orders where O and T are closer to V than S and R are, i.e. {{{V}O}S} and { { {V} T} R} (where { }means proximity, not linear order). The two sets of word order distribution can be largely explained by the combination of the LP and SP motivations. Specifically, SOV/RTV satisfy both LP and SP, but OSV/TRV violate both LP and SP. Thus SOV/RTV are overwhelmingly dominant over OSV/RTV.

Furthermore, since S is the default topic (best candidate for topic), but R is not, the difference between S and O in terms of precedence tendency is enlarged, hence the difference of the two rations 124.25 and 11.67.

(2) V-initial ordering

Data comparison:
Both VSO and VOS are likely to occur. In fact, numerous languages belong to the mixed type of the two orders and should be just titled with a covering term 'V-initial languages'. Similarly, both VRT and VTR are frequently documented and in many languages it is hard to choose a basic order between the two. 13 SVO languages in our data belong to the mixed type of VRT and VTR. The ration of VRT: VTR is 19/6=3.17, very close to that of VSO:VOS, i.e. 85/26=3.27.

Explanation:
VSO/VRT satisfy LP but violate SP while VOS/VTR satisfy SP but violate LP. Thus both are likely to occur.
In addition, we may suppose that the power of LP is stronger than that of SP, thus the ratio about 3 of VSO/VRT over VOS/VTR.

(3) V-middle

Data comparison:
The two sets of V-middle orders differ obviously. SVO is overwhelmingly dominant over OVS. However, as basic word orders, neither RVT nor TVR language is documented in our data. Nevertheless, there is still a minor similarity between the two sets: In many VRT/VTR languages and few RTV/TRV languages, the RVT alternation is much more frequently used than the TVR, such as Bai and Bisu.

Explanation:
If we follow the vein of the explanations of V-final and -initial sets, we just cannot explain the difference between SVO:OVS and RVT:TVR. One may be tempted to say that SVO satisfies both LP and SP, based on the structure [S[VO]], while OVS satisfies SP in the assigned structure [[OV]S] but violates LP. However, this difference shouldn't result in such big difference of the numbers of SVO and OVS languages.

Consider the difference between VOS (26 lgs) and OVS (9 lgs).   Both satisfy SP but violate LP, but the difference in numbers of tested languages of each order is considerable.

There must be some other factors coming into play. A likely one is the surface order. Though it is possible to assign the structure [[OV] S] to OVS, superficially, O and S have equal proximity to V.

Thus, the claim that OVS satisfies SP is dubious. Hence, OVS just violates LP and is irrelevant to SP. Or in other words, the role of SP is neutral here. This explains the difference of VOS and OVS with respect to the numbers of tested languages.

If OVS does not satisfy SP, then, how about SVO? Superficially, S and O are also equally close to V. However, consider the following contrast between VO and OV orders with respect to X (oblique and adverbial), it is obvious that the combination of VO is much more compacted than that of OV, because X cannot insert into VO order, but can into OV.

VO order			OV order			more than one order with none dominant
VOX	XVO	VXO	XOV	OXV	OVX	
189	3	0	45	23	37	152

Table 3 (Dryer & Gensle 2005)

Therefore, OV is not a very compacted combination, but VO is. And hence [[OV]S] does not make much sense. In other words, the claim that OVS, being assigned with [[OV]S], satisfies SP is not solid enough.   To sum up, SVO satisfy both LP and SP while OVS violates LP and is irrelevant to SP. This difference in structure leads to that in the numbers of tested languages. Now, we are in the position to explain why both V-middle RVT and TVR cannot be the basic order. It seems that there is a strong universal tendency to put R and T on the same side of V. This assumption is supported by Table 3, given that one of R and T aligns with oblique. In Table 3, there are 257 languages having O and X on the same side of V, only 40 languages have O and X on the two sides of V.

The above analysis does not apply to the difference between SVO and OVS, due to the high topicness of S. the relation S and O is that of topic and part of predicate, while the relation of R and T is that of object and oblique, belong to the same category to a much greater extent than S and O.
One left issue waiting for an explanation in Table 3 is the considerable difference between 3 XVO languages and 37 OVX languages. Space limited, we will leave it for the future discussion.

RTV（35）	Qiang, Dongxiang, East Yugur, Daur, Primi, Kachin, Achang, Uyghur, Lisu, Hani, Jino, Lahu, Nasu, Kirgiz, Evenki, Uzbek, Bengni-Bogar, Drung, Langsu, Cangluo Monba; Daofu, Queyu, Lvsu, Nung; Cuona Monba; Motuo Monba; Lhoba, Darang Deng, Geman Deng; Rouruo, Anong, Bisu［SOV, 32］ **[RTV/trv]** Tibetan, Tuchia, Sibo［SOV, 3］	**RTV/TVR**（9） Zaiwa, Naxi, Salar, Tatar, West Yugur, Kazakh, Yidu, Zhaba, Muya
TRV（3）	Yi, Tuva, Gazhuo［SOV, 3］	
VRT（19）	Blang, Pouyei, Benglong, Hmong, Bunu, Kelao, Dong, Baheng, Buxing, Yanghuang, Bai［SVO, 11］ Amis, Bunun, Paiwan［VSO, 3］ **[VRT/vtr]** Jing, Li, Mulam, She, Lingao［SVO, 5］ **[RTV/vrt/rvt]** Bisu	**VRT/VTR**（3） Biao, Bugeng, Cun［3］
VTR（6）	Laka, Dai［SVO, 2］ **[VTR/vrt]** Sui, Mian; Huihui, Jiongnai［SVO, 4］	
RVT（0）	Bai (**VRT/rvt**), Bisu (**RTV/vrt/rvt**)	
TVR（0）		
Unknown/ unclear（15）	Nanai (Hezhen), Mongolian, Baoʾan, Tu, Tadzhik, Korean, Olunchun, Namuz, Shixing, Leqi, Pola, Xiandao［SOV, 12］ Va, Zhuang, Mulao［SVO, 3］	

Appendix: Ordering of R, T and V in 90 languages in China