Advances in Complex Systems
© World Scientific Publishing Company

Some limits of standard linguistic typology. The case of Cysouw's models for the frequencies of the six possible orderings of S, V and O.

Ramon Ferrer-i-Cancho

Departament de Llenguatges i Sistemes Informàtics, TALP Research Center, Universitat Politècnica de Catalunya, Campus Nord, Edifici Ω Jordi Girona Salgado 1-3. 08034 Barcelona, Spain.
ramon.ferrericancho@gmail.com

Received (received date)
Revised (revised date)

This article is a critical analysis Michael Cysouw's comment "Linear order as a predictor of word order regularities".

Keywords: reductionism, holism, inductivism, word order, linguistic universals, linguistic typology, human language.

1. Introduction

There is a long tradition in linguistic typology regarding word orders as a combination of more elementary word orders ([2] and refs. cited there). For instance, SOV and SVO can be seen as the the outcome of integrating the elementary orders SO and SV [2]. This standard reductionist view is in conflict with the holistic point of view of the mathematical framework put forward by Ramon Ferrer-i-Cancho, hereafter RF, for shedding light on certain word order biases. RF's mathematical framework is rooted in a long cognitive science tradition (see refs. in [5]) stressing the importance of syntactic arc length for word order. From a complexity perspective, the minimization of this length can be regarded as an emergent phenomenon in sufficiently long sentences. For instance, the order of a subject, a verb and an object does not matter in terms of arc length in pairs of these two elements (the distance between elements is one for any pair) but if one combines two of these pairs (with the constraint that the pairs have at least one common element, e.g., S in SO and SV) then distance begins to matter (given a sequential ordering of three words, the first and the third element are not consecutive). The models by RF and Michael Cysouw, hereafter MC, are linguistic examples of the holism versus reductionism debate in science. Another classic long standing problem in science is the degree of dependence models on real data. Solutions range from purely inductive models such as MC's to the more abstract approaches of RF or, at a higher level of abstraction, the Chomskyan tradition. The aim of this article is exemplifying the

Electronic version of an article published as: Ferrer-i-Cancho, R. Some limits of standard linguistic typology: the case of Cysouw's models for the frequencies of the six possible orderings of S, V and O. "Advances in complex systems", 2008, vol. 11, núm. 3, p. 421-432. DOI 10.1142/S0219525908001726.
© 2008 World Scientific Publishing Company. http://dx.doi.org/10.1142/S0219525908001726

limits of traditional linguistic typology using MC's model, not only as qualitative model for explaining the frequencies of the six possible orderings of S, V and O in world languages (as MC intends in [2]) but also as a quantitative model. The analysis of MC's models beyond what MC intended is pertinent to discuss the limits of radical reductionism and inductivism as well as for providing an scientific background on which future developments of reductionist and holistic approaches can be more objectively evaluated and integrated.

Before we proceed we have to warn the reader. First, the aim of this article is not arguing that the solution is excluding reductionism or induction but holism and abstraction cannot be obviated at least in some cases. Moreover, this article does not aim to show that RF's model is a better model than MC's for the frequency of the six possible orderings of S, V and O. We only argue that that RF's framework is grounded better and point out the limits of MC's more concrete proposal and the tradition it belongs to. Although MC as referee deemed it necessary, the original model in [5] was not aimed at explaining the frequencies of these orderings and [5] is essentially a vague speculation about it. RF's proposal does clearly less than MC's models.

The remainder of the article is divided into an analysis of MC's model as qualitative and quantitative model (Section 2) and a review of RF's approach (Section 3).

2. MC's model

MC argues that the frequency of the six possible orderings of S, V and O in world languages in free words could be easily explained by a combination of two strong preferences, one for SO over OS and another for SV over VS [2]. From the quantitative point of view, rigorous statistical analysis shows that (see Appendix B for the details):

(1) The assumption that SV and SO are independent is not supported by the actual data. MC never claims that they are independent but the accuracy of his model relies on statistical independency. Two orders x and y are statistically independent if and only if their joint probability equates the product of their respective marginal probabilities: $p_{xy} = p_x p_y$. If independence is not supported then $p_{xy} = p_x p_y$ ceases to be true and MC's proposals cannot be models in the strict sense.

(2) The assumption that SVO and SOV should be equally likely is not supported either. The same applies to VOS and OVS. MC is aware of this problem (Section 1 of [2]). However, assuming that the word order pairs above are equally likely is an arbitrary decision. There are more favourable and disfavourable alternatives (in terms of the fitting of the model) whose exclusion is not justified. This is a clear limit of reductionism: there are statistically significant difference in the frequency of the orderings of triples that cannot be explained.

(3) Only one of the six possible orderings has a frequency that is statistically con-

sistent with MC's model.

The same methodology used for the orderings of S, V and O in free words can be used to show the weakness of MC's proposal for verbal affixes (see Appendix C for the details).

In the next subsections we review various limits of MC's model from the qualitative point of view.

2.1. *A single dual "cause" for the occurrence of SVO and SVO in free words*

MC's model assumes that the frequencies of SVO and SOV are both the outcome of a universal preference for SV and SO. This assumption is problematic for various reasons:

- As shown in Section B.3, the difference in the frequencies of SOV and SVO cannot be explained by these two orders being equally likely. There is a bias towards SOV. This alone suggests that there is a at least one secondary factor not included in MC's model. That secondary factor cannot be arc length minimization, since SOV is a priori less economical than SVO in terms of Euclidean distances [5].
- Recent research in cognitive science suggests that SOV is natural order for non-verbal communication tasks, independently of the speaker's language a priori preference [7, 6]. These findings suggest that the secondary causes hypothesized above could actually be primary factors that are independent of the forces towards SVO. The absence or rarity of SVO in the experiments in [7, 6] can be easily justified by RF's approach: if 'sentences' are not sufficiently long (of three items as in [7], choosing an optimal word order (from the arc length minimization point of view) as SVO might not be enough rewarding. For instance, assuming $n = 3$, the sum of all arc lengths in SVO is 2 while it is 3 in SOV (see Section 5.1 of [5]). Of course, this argument relies on the assumption that the cost of an arc is a linear function of its length, which is questionable (if cost turned out to be an exponential function of distance then our argument would be invalidated). In contrast, MC's model cannot easily explain why SVO is absent or rare in [7]. Briefly, this absence provides support for word order as a multifactorial problem and suggests that holistic phenomena may play in an important role in word order organization (against the reductionism of traditional standard linguistic typology).
- MC reports on Table 2 of his comment [2] that SVO and OVS are the two most frequent word orders chosen for verbal affixes. SOV is not the most frequent word order and it is even very improbable that SVO and SOV are equally likely (as shown in Appendix B.3). Therefore, these results suggests that the factors favoring SOV in verbal affixes not only could be less intense than in free words under certain circumstances but also weaker than arc length minimiza-

tion. Rather than a single dual factor for SVO and SOV, empirical evidence suggests a combination of various forces.

2.2. *The origins of a universal preference for SV and SO (or SV and VO in verbal affixes)*

The hypothetical universal preference for SV and SO is an axiom of MC's model. One would expect that these biases are the result of some cognitive preference. This would make MC's model stronger. Instead of grounding his models on human cognition, MC constructs inductively different ad-hoc models for the ordering of free words and the ordering of verbal affixes. While the scenario changes, the human brain remains the same. For the reasons, RF's speculations make emphasis on the invariants of the word order problem. Arc length minimization is priori a general cognitively well-supported (recall [5]) factor to consider although its importance may depend on the context.

2.3. *The explanatory power of a preference for SV and SO in free words*

The preference for SV and SO that is valid for explaining free word order frequencies is not valid for verbal affix frequencies. MC is aware of this problem (Section 2 of [2]). A rigorous and detailed statistical analysis of MC's model for free words on verbal affixes confirms it (Appendix D). Traditional linguists would probably complain that affixes and free words are totally different entities so any extrapolation from free words to affixes is not possible. However, the ultimate goal of science is not the construction of ad-hoc explanations to concrete problems (as MC's models) but the construction of theories as general as possible with the smallest set of principles. Postulating a preference for SV and SO for free words and a preference for SV and VO in verbal affixes does not seem to be a step in this direction with our currently available knowledge.

A necessary step towards understanding the frequencies of word orders is understanding the adaptative landscape on which the evolution of word orders operates. The fact that SVO is on top of the frequency ranking in free words (2nd place) and affixes (1st place) suggests that arc length minimization plays a role in the evolution of languages and has some influence on actual word order frequencies. At present, we cannot probably say something more concrete than this based strictly on [5]. However, notice that this is not an ad-hoc statement since the general principle of economy or least effort [10] (instantiated in our case as a principle of Euclidean distance minimization) has been argued to underly many universal properties of language such as the shortening of words with frequency (the so-called law of brevity) [9], the shortening of syllables as word length grows [8] (the so-called Menzerath-Altmann law [1]), the exceptionality of crossings in syntactic dependency trees [4],... In contrast, the scope of a preference for SV and SO is narrower, so narrow that it fails for verbal affixes.

Notice that here we are criticizing MC's proposal from a higher level of abstraction, i.e the level of metamodelling. Instead of worrying about locally optimal models for separate problems (modelling level), we are putting emphasis on the common aspects of problems and reducing the differences between models (metamodelling level). The ultimate goal of metamodelling is unifying explanations.

2.4. *Radical induction*

As MC recalls, M. Dryer has argued very fervently against world wide word order correlations due to evolutionary dependencies among languages. Briefly, two languages sharing the same common ancestor are a priori more likely to share a certain feature than languages distantly related. The point raised by M. Dryer and revisited by MC suggests that these correlations are deceiving and ultimately useless for scientific research. Dryer's objections are critical for purely inductive approaches such as MC's models and the traditional way in which Greenbergian universals are obtained. RF's framework does not rely exclusively on induction as MC's models for two reasons. First, as reviewed above, economy is a principle of language. Second, a bias for the verb at the center is a mathematical truth in the simplified fashion in which it is presented in [5]. It does not depend on sampling effects or historical accidents in the evolution of language. We will provide various arguments that question Dryer's and MC's negative view specially with regard to well-grounded theories that do not rely exclusively on induction:

- It is obvious that evolutionary history plays a role in these correlations but MC's and Dryer's negative view implicitly assumes that the properties of a language are only determined by evolutionary history and that other factors do not play any role. A crucial issue is: if a language should adopt or adopts a certain feature from his predecessor, is this language or its successors able to get rid of this feature by means of evolution? What is the degree of freedom in language evolution? The diversity of languages only at the level of order suggests a high degree of flexibility.
- MC attempts to support Dryer objections showing that reducing the influence of evolution by using genera instead of languages for the counts decreases the statistical significance of the correlations. MC suggests that this decrease questions the validity of correlations obtained with languages. However, it is well-known that p-values tend to decrease as the sample size reduces. Therefore, one cannot unequivocally attribute the decrease (even above a certain significance level) to the the use of genera (or another higher unit) instead of languages. While there are 1481 languages on Table 3 of [2], the are 482 genera on Table 4 of [2].
- The correlations that one may discover neglecting such dependencies can be used to test the consistency of theories as in [5] or to check the degree of freedom from evolutionary history of well-grounded theories (theories that are not purely inductive).

What we have said above should not be interpreted as a free ticket for theories that do not care about statistical patterns of languages.

3. Reviewing RF's approach

A natural preference for SOV [7, 6] by factors that are independent from the optimality of SVO in terms of arc length minimization has implications for speculations on the evolution of language. RF's approach does not need to call for a neutral symmetry breaking between towards SVO and against its symmetric OVS to explain the rarity of OVS. Languages would be naturally attracted by SOV, which is adjacent to SVO but far from OVS in the permutation ring in Fig. 4 of [5].

Now we aim to check with the same methods if RF's speculations about the frequencies of the six possible orderings are well supported on free words and verbal affixes (notice that we can only use methods for which p-values below standard p-levels are possible). It can be shown that (Appendix E for the details).

- There is a tendency to put the verb at the center of the triple as expected from the arc length minimization principle) in both free words and verbal affixes.
- The only word orders that have a frequency greater than expected from random chance are SVO and SOV in free words and SVO in verbal affixes. The overabundance of SOV could be explained by factors other than arc length minimization, as argued throughout this article. The fact that OVS is not overabundant as its symmetric SVO could be explained by the interference of SOV discussed in this article.

In sum, our analysis suggests that arc length minimization has left a *statistical trace* in the abundances of word orders. This is a more modest conclusion than the qualitative predictions made by MC's models.

Acknowledgements

We are very grateful to M. Cysouw and P. Stadler as editor of ACS for the opportunity to debate controversial issues publicly and their support for new ideas. We thank P. Stadler, A. Hernández and S. Goldin-Meadow for helpful comments. This work was supported by the project FIS2006-13321-C02-01 from Spanish Ministry of Education and Science. It is strictly prohibited to use, investigate or develop, in a direct or indirect way, any of the scientific contributions of the author(s) contained in this work by any army or armed group in the world, for military purposes and for any other use which is against human rights or the environment, unless a written consent of all the persons in the world is obtained.

References

[1] G. Altmann. Prolegomena to Menzerath's law. *Glottometrika 2*, 2:1–10, 1980.
[2] M. Cysouw. Linear order as a predictor of word order regularities. *Advances in Complex Systems*, 2008.

[3] M. H. DeGroot. *Probability and statistics.* Addison-Wesley, Reading, MA, 1989. 2nd edition.
[4] R. Ferrer i Cancho. Why do syntactic links not cross? *Europhysics Letters*, 76(6):1228–1235, 2006.
[5] R. Ferrer-i-Cancho. Some word order biases from limited brain resources. A mathematical approach. *Advances in Complex Systems*, 2008.
[6] L. Gershkoff-Stowe and S. Goldin-Meadow. Is there a natural order for expressing semantic relations? *Cognitive Psychology*, 45:375–412, 2002.
[7] S. Goldin-Meadow, W. C. So, A. Özyürek, and C. Mylander. The natural order of events: how speakers of different languages represent events nonverbally. *Proceedings of the National Academy of Sciences*, 2008.
[8] R. Köhler. Das Menzerathsche Gesetz als Resultat des Sprachverarbeitungsmechanismus. In G. Altmann and M. H. Schwibbe, editors, *Das Menzerathsche Gesetz in informationsverarbeitenden Systemen*, pages 108–122. Olms, Hildesheim, 1989.
[9] G. K. Zipf. *The psycho-biology of language.* Houghton Mifflin, Boston, 1935.
[10] G. K. Zipf. *Human behaviour and the principle of least effort. An introduction to human ecology.* Hafner reprint, New York, 1972. 1st edition: Cambridge, MA: Addison-Wesley, 1949.

Appendix A. The binomial test

The probability of x successes in a series of y independent trials of probability π follows the binomial distribution if the probability of x successes is [3]

$$p(x) = \binom{y}{x} \pi^x \pi^{y-x}. \tag{A.1}$$

We say that x is binomially distributed (with parameters y and π) if its probability is governed by Eq. A.1. Given a certain empirical value z, we can test if it is consistent with a binomial distribution. More precisely, we can calculate

$$p(x \geq z) = \sum_{x=z}^{y} p(x), \tag{A.2}$$

the probability of obtaining a value of x greater or equal than z assuming that x is binomially distributed with parameters y and π. If $p(x \geq z)$ is below or equal a certain small value, e.g., 0.01 (the so-called significance level) we can conclude that it is unlikely that z comes from a binomial distribution with the parameters mentioned above. In this case, z is significantly larger than expected from a binomial distribution with these parameters. This procedure constitutes the binomial upper tailed test.

Similarly, we can test if z is smaller than expected from random chance using

$$p(x \leq z) = \sum_{x=0}^{z} p(x). \tag{A.3}$$

If $p(x \leq z)$ is below or equal the significance level we conclude that it is unlikely that z comes from a binomial distribution with the parameters mentioned above. In this case, z is significantly smaller than expected from a binomial distribution

x	N_x	p_x	$\hat{p}_x$
SOV	497	$497/1056 \approx 0.47$	$13221/30976 \approx 0.43$
SVO	435	$145/352 \approx 0.41$	$13221/30976 \approx 0.43$
VSO	85	$85/1056 \approx 0.08$	$1695/15488 \approx 0.11$
VOS	26	$13/528 \approx 0.025$	$5/30976 \approx 0.002$
OVS	9	$3/352 \approx 0.0085$	$5/30976 \approx 0.002$
OSV	4	$4/1056 = 1/264 \approx 0.0038$	$507/15488 \approx 0.033$

Table 1. Summary of the absolute frequency (N_x), relative frequency (p_x) and the predicted frequency according to MC's model ($\hat{p}_x$) of free word orders (x).

with these parameters. This second procedure constitutes the binomial lower tailed test.

Appendix B. MC's model for free words

Before we start, we need to introduce some notation. We use N_x to refer to the number of languages following order x. For instance, N_{SO} is the number of languages in which the subject S typically precedes the object O and N_{SVO} is the number of languages where typically the subject precedes the verb and the verb precedes the object. We define N_* as the total number languages considered. We define p_x as the proportion of languages following order x. p_x is an estimate of the probability that a language follows order x. We also define $\hat{p}_x$ as the predicted probability of order x according to MC's model. $\hat{p}_x$ is an estimate of p_x. The well-known statistics of word order frequencies and the prediction of MC's model are shown in Table 1 using our notation. In our sample we have $N_* = N_{SO} + N_{OS} = 1056$.

The essence of MC's model can be restated more formally by means of three points:

(1) SO and SV are independent. Therefore

$$p_{SOV} + p_{SVO} = p_{SV}p_{SO} \tag{B.1}$$

$$p_{VOS} + p_{OVS} = p_{VS}p_{OS} \tag{B.2}$$

$$p_{VSO} = p_{VS}p_{SO} = (1 - p_{SV})p_{SO} \tag{B.3}$$

$$p_{OSV} = p_{SV}p_{OS} = p_{SV}(1 - p_{SO}). \tag{B.4}$$

(2) $p_{SOV} = p_{SVO}$ and $p_{VOS} = p_{OVS}$. Therefore

$$p_{SOV} = p_{SVO} = \frac{p_{SV}p_{SO}}{2} \tag{B.5}$$

$$p_{VOS} = p_{OVS} = \frac{p_{VS}p_{OS}}{2}. \tag{B.6}$$

(B.7)

(3) The preferences for SV over VS and for SO over OS are axioms of the theory.

x	N_x
OS	$N_{VOS} + N_{OVS} + N_{OSV} = 39$
OV	$N_{SOV} + N_{OVS} + N_{OSV} = 510$
SO	$N_{SOV} + N_{SVO} + N_{VSO} = 1017$
SV	$N_{SOV} + N_{SVO} + N_{OSV} = 936$
VO	$N_{SVO} + N_{VSO} + N_{VOS} = 546$
VS	$N_{VSO} + N_{VOS} + N_{OVS} = 120$

Table 2. Summary of the absolute frequency (N_x) of free word pairs (x).

B.1. *The preference for SO over OS and for SV over VS*

MC does not provide any statistical evidence of these strong preferences. We need to check if the number SO languages and SV languages are, separately, greater than expected from chance using a rigorous statistical test. To this aim, we calculate the frequency of each possible ordered pair from the set O, S, V (Table 2).

The a priori chance that a random ordering of the triple O, S and V follows SO is 1/2 (there are 6 possible orderings of which 3 follow the order SO). It is easy to see that any possible ordered pair drawn from the set O, S, V has probability 1/2 of being obtained by random ordering of the triple. The number of languages that should follow a certain order pair from the set $\{O, S, V\}$ is binomially distributed with parameters $N_* = 1056$ and 1/2. The binomial upper-tailed test (Appendix A) shows that the actual number of languages following the order SV is greater than expected from random chance ($p - value < 10^{-10}$). The same applies to SO ($p - value < 10^{-10}$). In contrast, the binomial test indicates that the number of OV or VO is not either greater or smaller than expected from random chance ($p - value \geq 0.14$ in all cases). We conclude that MC's hypothesis of two preferences is statistically well-supported although this is not obvious from his proposal.

B.2. *The independence of SO from SV*

MC does not provide any evidence of the independence of SO from SV (or any absence of statistically significant dependence between SO and SV). Using the information in Table 1, we can construct a 2×2 contingency table for testing the correlation between SO and SV (Table 3). The right tailed Fisher exact test indicates that there is an statistically significant positive correlation between SO and SV ($p - value < 10^{-10}$), contradicting the assumption of independence of MC's model.

B.3. *The equally likelihood of certain pairs of word orders*

MC's model assumes that SOV and SVO are equally likely and thus N_{SOV} should be binomially distributed with parameters $N_{SOV} + N_{SVO}$ and 1/2. The upper tailed binomial test (Appendix A) shows that there are more SOV languages than ex-

	SV	VS
SO	$N_{SOV} + N_{SVO} = 932$	$N_{VSO} = 85$
OS	$N_{VOS} = 26$	$N_{VOS} + N_{OVS} = 35$

Table 3. Two times two contingency table of the languages following SV and SO and their respective reverses.

pected from that distribution ($p-value < 0.023$). Therefore, there is an unexplained preference for SOV over SVO. Similarly, MC's model assumes that VOS and OVS are equally likely and thus N_{VOS} should be binomially distributed with parameters $N_{VOS} + N_{OVS}$ and 1/2. The upper tailed binomial test shows that there are more VOS languages than expected from that distribution ($p - value < 0.003$). Therefore, there is an unexplained preference for VOS over OVS. In sum, the assumption of equally likelihood of certain word order pairs fails.

B.4. *The goodness of the predictions of MC's model*

Knowing that $p_{SV} = p_{SOV} + p_{SVO} + p_{OSV}$ and $p_{SO} = p_{SOV} + p_{SVO} + p_{VSO}$, we can estimate all the probabilities of the six possible orderings of S, V and O according to MC's model that are shown in Table 1 using Eqs. B.3, B.4, B.5 and B.6, and Table 2.

An indicator of MC's model goodness of fit is the relative error between $\hat{p}_x$ and p_x. For instance, the absolute difference between p_{OSV} and $\hat{p}_{OSV}$ is "small" ($|p_{OSV} - \hat{p}_{OSV}| \approx 0.028$) but the relative difference is of about two orders of magnitude larger ($|p_{OSV} - \hat{p}_{OSV}|/\hat{p}_{OSV} \approx 7.6$). Small absolute differences are misleading. The ultimate answer about the goodness of the model comes from testing if each of the six real frequencies is consistent with a binomial distribution. More precisely, the model assumes that N_x follows a binomial distribution with parameters N_* and $\hat{p}_x$. For instance, $N_{SOV} = 497$ should follow a binomial distribution with parameters $N_* = 1056$ and $\hat{p}_{SOV} = 13221/30976$. The right tailed binomial test (Appendix A) indicates that N_x is larger than expected from MC's model for $x \in SOV, VOS, OVS$ ($p - value < 0.003$ in all cases). The left tailed binomial test indicates that N_x is smaller than expected from MC's model for $x \in VSO, OSV$ ($p - value < 0.002$ in both cases). N_{SVO} is the only frequency consistent with MC's model.

Appendix C. MC's model for verbal affixes

Here we use the same methodology that we used to analyze MC's model on free word order frequencies in Appendix B to MC's model for the order of verbal affixes. We summarize the results:

- *Strong preference for SV over VS and VO over OV*
  It is shown in Appendix D that these preferences are statistically sound but there is also a preference for SO that is neglected by MC.

- *The independence of SV from VO*
  The Fisher exact test fails to detect a statistically significant correlation between SO and SV ($p - value > 0.47$) providing support for MC's assumption.
- *The equally likelihood of certain pairs of word orders*
  MC's model assumes that (a) SOV and OSV are equally likely and (b) VSO and VOS are also equally likely. The binomial test (Appendix A) fails to find evidence against the equally likelihood of SOV and OSV and against the equally likelihood of VSO and VOS ($p-value > 0.30$ in both cases), providing support for MC's assumption.
- *The goodness of the predictions of MC's model*
  The binomial test shows that there is no real word order frequency that is statistically consistent with the predicted frequencies by MC's model ($p - value < 3 \cdot 10^{-4}$ in all cases except for OVS ($p - value > 0.47$). There are less SOV, VOS, VSO and OSV and more SVO than expected from MC's model.

Appendix D. MC's model for free words on verbal affixes

Here we use the same methodology that we used to analyze MC's model on free word order frequencies in Appendix B to MC's model on the order of verbal affixes. We summarize the results:

- *Strong preference for SO over OS and SV over VS*
  The binomial test (Appendix A) shows not only a preference for SO and SV and VO over their respective reverses but also for VO over OV ($p - value < 10^{-3}$). In the case of verbal affixes, preferences do not reduce to just two.
- *The independence of SO from SV*
  The Fisher exact test indicates that there is a positive statistically significant positive correlation between SO and SV ($p - value < 10^{-10}$).
- *The equally likelihood of certain pairs of word orders*
  The binomial test shows that SOV and SVO are not equally likely ($p-value < 10^{-6}$). In contrast, the same test indicates that there is not enough statistical evidence against the equally likelihood of VOS and OVS ($p - value > 0.4$).
- *The goodness of the predictions of MC's model*
  The binomial test shows that there is no real word order frequency that is statistically consistent with the predicted frequencies by MC's model ($p - value < 0.01$ in all cases). There are less SOV, VSO and OSV and more SVO, VOS, OVS than expected from MC's model.

Appendix E. Review of RF's approach

First, we aim to support the hypothesis that there is a tendency for placing the verb at the center in free word orders. The probability that a random permutation of the triple has V at the center is $1/3$ (V is at the center in two out of six permutations). The number of languages with the verb at the center is $N_{*V*} = N_{SVO} + N_{OVS} =$

444. In the absence of any tendency, N_{*V*} is binomially distributed with parameters $N_* = 1056$ and $\pi = 1/3$. The binomial test (Appendix A) indicates that N_{*V*} is greater than expected from chance ($p - value < 10^{-8}$). Using the same approach on verbal affixes, it can also be shown that there is also a tendency to put the verb at the center ($p - value < 10^{-7}$). Second, we aim to show that SOV and SVO have a frequency greater than expected from chance and that the remaining orderings have a smaller frequency in free words. Consider that the probability that a random permutation gives a certain word order (e.g., SVO) is 1/6. In the absence of any tendency, N_x (where x is one of the six possible orderings of the triple) is binomially distributed with parameters $N_* = 1056$ and $\pi = 1/6$. The binomial test indicates that N_x is greater than expected from chance for SVO and SOV and smaller than expected from chance for the remaining orders ($p - value < 10^{-10}$ in all cases). Using the same approach on verbal affixes, it can also be shown that SVO is more frequent than expected from chance ($p - value < 10^-10$ whereas VSO and OSV are less frequent than expected ($p - value < 0.011$ and $p - value < 0.034$, respectively). There is not enough statistical evidence to reject the hypothesis that the remaining word orders have a frequency expected from chance ($p - value > 0.087$ in all cases).