

Disentangling word order and function assignment preferences in Modern Greek

Yanis da Cunha, Ioanna Chorai, Anne Abeillé

LLF, Université Paris Cité, France

<https://doi.org/10.36505/ExLing-2022/13/0013/000555>

Abstract

Construction alternations are influenced by several conceptual properties of referents, including animacy or definiteness, known as “prominence features” (Aissen 1999, Haspelmath 2020). However, it remains unclear whether these features influence construction alternations through word order or function assignment. To disentangle these effects, we conducted a quantitative corpus study in Greek, comparing a word order alternation (SOV/OVS) and a function assignment alternation (active/passive). We annotated a random sample of 600 sentences from the Corpus of Modern Greek (Arkhangelskiy, Kisilier 2018) and fitted a Bayesian categorical model to construction (SVO/OVS/passive). We show that animacy and definiteness features have an effect on function assignment, not on word order, but we found a short-before-long length effect.

Keywords: Modern Greek, construction alternations, corpus study, Bayesian modelling

Introduction

Construction alternations are influenced by several factors, including animacy, definiteness or length, as in dative alternation in English (Bresnan, Ford 2010). However, it remains unclear whether so-called “referential prominence features” (Aissen 1999, Haspelmath 2020) like animacy and definiteness affect construction choice through word order (eg. animate-first preference) or function assignment (eg. animate-subject preference) (Thuilier et al. 2021). Using sentence recall tasks, experimental studies both in English (Bock, Warren 1985, McDonald et al. 1993) and French (Thuilier et al. 2021) have tried to disentangle these preferences by comparing animacy effects on active/passive alternation (a function assignment alternation) and NP coordinations (a word order alternation). The authors report an animacy effect on active/passive alternation, but not on word order in coordinations, suggesting that animacy plays a role only for function assignment.

However, this conclusion is challenged by Feleki & Branigan (1999), who found an animate-first effect in Greek using the same sentence recall task with the SVO/OVS alternation. The authors thus claim that animacy only affects word order and that animate-subject preferences observed in English are just due to the tight connection between subject and first position in this language. But as Feleki and Branigan (1999) only took into account the SVO/OVS

alternation, word order and function assignment preferences have not yet been directly compared in Greek.

As for length, it is sometimes included among prominence features (Bresnan, Ford 2010), while McDonald et al. (1993) treat it separately, as it did not show any effect in sentence recall in English. Feleki and Branigan (1999) study did not include this factor in Greek.

Methodology

We aim at disentangling word order and function assignment preferences in Greek by conducting a quantitative corpus study. We used newspaper texts from the morphologically annotated Corpus of Modern Greek (Arkhangelskiy, Kisilier 2018). We randomly extracted a sample of 600 sentences, containing 200 active SVO (1), 200 active OVS (2) and 200 passive constructions with a preverbal subject and an expressed by-phrase agent (3). Verbs were all transitive and passivizable, and the arguments were only NPs, because clitic object pronouns are preverbal and subject pronouns mostly drop.

- (1) Oi ispanoi, prosperoun 1.500 eisitiria
the Spanish offer 1500 tickets
'The Spanish people offer 1500 tickets' (Makedonia, 2007/09/26)
- (2) Ipotheseis ekane o Oli Pen
hypotheses made the Oli Pen
'Oli Pen made assumptions'
- (3a) I synchroni techni empneetai apo ti vizantini (Makedonia, 2010/10/20)
the contemporary arts inspire.PASSIVE by the byzantine
'Contemporary art is inspired by Byzantine art' (To Vima, 2012/01/26)
- (3b) Tessera paidia travmatistikan apo tin puravliki epithesi
four children injure.PASSIVE by the missile attack
'Four children were injured by the missile attack' (To Vima, 2011/05/10)

Using Arg1 for active subject and passive by-phrase, and Arg2 for active object and passive subject, we annotated argument animacy, definiteness and length (number of words). With these three constructions (SVO/OVS/Passive), we can tease apart word order from function assignment preferences: in SVO/OVS, functions are the same but order differs, in OVS/Passive, there is the same Arg2-Arg1 order but function mapping varies (object Arg2 in active vs. subject Arg2 in passive).

For statistical analysis, we fitted a Bayesian categorical model to construction as a function of animacy, definiteness and length, using the *brms* package in R (Bürkner 2017). We used OVS as reference level for the outcome, as it minimally differs from both SVO and Passive. Verb lemma was used as a random variable. Length was included using a *Length difference* variable: $\log(\text{Arg1 Length}) - \log(\text{Arg2 Length})$ (Bresnan & Ford, 2010). We used weak

informative priors for regression coefficients (normal distribution with $\mu=0$, $\sigma=1$) and default brms priors for other parameters (Student's t-distribution with $\nu=3$, $\mu=0$, $\sigma=2.5$). Four sampling chains ran 4000 iterations with a warm-up period of 1000 iterations.

Results

Results from the Bayesian categorical model are reported in Figure 1. We examine 6 hypotheses: preferences for human-subject, human-first, definite-subject, definite-first, shorter-subject, shorter-first. In each case, we report mean parameter value (β_{Passive} for Passive/OVS contrast and β_{SVO} for SVO/OVS contrast), 95% credible interval and the probability P of β differing from 0. We find compelling evidence for human-subject preference for both Arg1 (1) ($\beta_{\text{Passive}}=-1.00$, $\text{CI}=[-1.48;-0.53]$, $P(\beta_{\text{Passive}}<0)=100\%$) and Arg2 (3b) ($\beta_{\text{Passive}}=1.84$, $\text{CI}=[1.13;2.58]$, $P(\beta_{\text{Passive}}>0)=100\%$). However, we do not find evidence for human-first preference for neither Arg1 (2) ($\beta_{\text{SVO}}=-0.20$, $\text{CI}=[-0.6; 0.21]$, $P(\beta_{\text{SVO}}>0)=21\%$) nor Arg2 ($\beta_{\text{SVO}}=0.56$, $\text{CI}=[-0.14;1.27]$, $P(\beta_{\text{SVO}}<0)=10\%$). Animacy thus has an effect on function assignment, not on word order. The same holds for definiteness: there is evidence for a definite-subject preference for both Arg1 (1) ($\beta_{\text{Passive}}=-1.85$, $\text{CI}=[-2.47;-1.24]$, $P(\beta_{\text{Passive}}<0)=100\%$) and Arg2 (3a) ($\beta_{\text{Passive}}=2.70$, $\text{CI}=[2.14;3.28]$, $P(\beta_{\text{Passive}}>0)=100\%$). But no evidence is provided for definite-first preference with Arg1 ($\beta_{\text{SVO}}=-0.15$, $\text{CI}=[-0.77;0.47]$, $P(\beta_{\text{SVO}}>0)=35\%$) or Arg2 (2)(3b) ($\beta_{\text{SVO}}=0.71$, $\text{CI}=[0.31;1.11]$, $P(\beta_{\text{SVO}}<0)=0\%$). Finally, we found a shorter-first preference ($\beta_{\text{SVO}}=-1.32$, $\text{CI}=[-1.57; -1.09]$, $P(\beta_{\text{SVO}}<0)=100\%$) but no shorter-subject preference ($\beta_{\text{Passive}}=-0.12$, $\text{CI}=[-0.36;0.12]$, $P(\beta_{\text{Passive}}>0)=20\%$).

Conclusion

Statistical modelling on Greek corpus data shows that prominence features (animacy and definiteness) only affects function assignment, which can be viewed as argument coding efficiency (Haspelmath 2020). By contrast, length only affects word order, which can be explained by dependency length minimization (Temperley, Gildea 2018). This difference brings evidence to distinguish the status of prominence features and length. While prominence features represent referential/conceptual information, length is more tied to phrasal/formal properties (McDonald et al., 1993). We failed to replicate an animate-first preference as found by Feleki, Branigan (1999) in Greek, which may be due to task effects: ecological production may not involve the same cognitive processes as sentence recall.

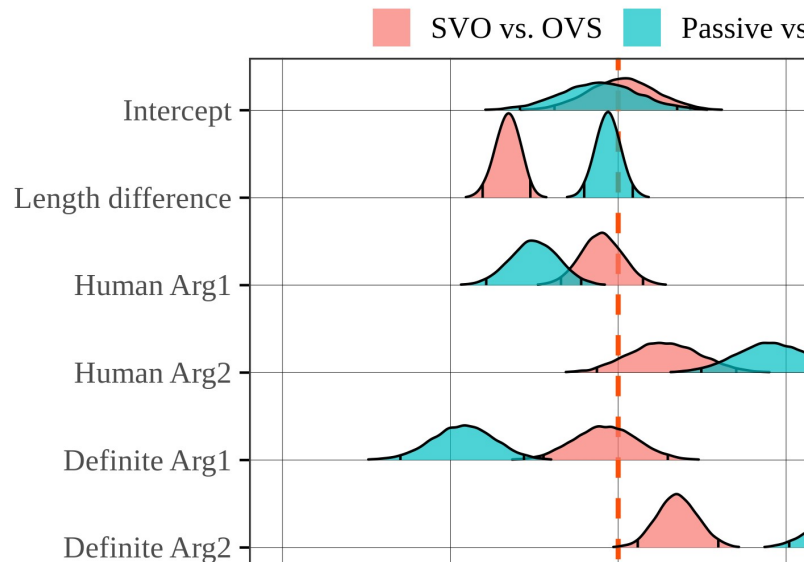


Figure 1. Fixed-effects posteriors in the categorical model. Vertical black lines represent 95% credible intervals.

References

- Aissen, J. 1999. Markedness and Subject Choice in Optimality Theory. *Natural Language & Linguistic Theory*, 17(4), 673-711.
- Arkhangelskiy, T., Kisilier, M. 2018. Corpora of Modern Greek: Achievements and goals. *Indo-European Linguistics and Classical Philology*, 5059.
- Bresnan, J., Ford, M. 2010. Predicting syntax: Processing dative constructions in American and Australian varieties of English. *Language*, 168-213.
- Bock, J.K., Warren, R.K. 1985. Conceptual accessibility and syntactic structure in sentence formulation. *Cognition*, 21(1), 47-67.
- Bürkner, P.-C. 2017. brms: An R package for Bayesian multilevel models using Stan. *Journal of statistical software*, 80, 1-28.
- Feleki, E., Branigan, H. 1999. Conceptual accessibility and serial order in Greek speech production. *Proceedings of the 21st Annual Cogsci Conference*, 961-101.
- Haspelmath, M. 2020. Role-reference associations and the explanation of argument coding splits. *Linguistics*, 45.
- McDonald, J.L., Bock, K., Kelly, M.H. 1993. Word and World Order: Semantic, Phonological, and Metrical Determinants of Serial Position. *Cognitive Psychology*, 25(2), 188-230.
- Temperley, D., Gildea, D. 2018. Minimizing syntactic dependency lengths: Typological/cognitive universal? *Annual Review of Linguistics*, 4, 67-80.
- Thuilier, J., Grant, M., Crabbé, B., Abeillé, A. 2021. Word order in French: The role of animacy. *Glossa: a journal of general linguistics*, 6(1), 55.