

Quantitative Analysis of Passives with Agent Phrase Based on Multilingual Parallel Data

Liubov Nesterenko¹[0000-0002-6872-5134]

National Research University Higher School of Economics, Moscow, Russia

Abstract. In this paper I discuss the advantages of using parallel data in linguistic research and demonstrate preliminary results of the study devoted to passives with agent phrase. For the study I used a parallel corpus of texts in nine European languages, the data set contained 983 fully aligned translation units. In my experiment I aimed to check whether the distribution of passives with agent phrase and related constructions used in translations depends on the semantic role of the participant that corresponds to the oblique argument of the passive.

Keywords: Passive Constructions, Parallel Corpora, Multilingual Data.

1 Introduction

In the last few years works presenting quantitative analysis on parallel data tend to appear more often, which seems to be a big step forward in linguistic research, some of the most recent examples of parallel corpora based studies are presented in [1, 4, 5], [17], [20]. Even though the benefits of parallel corpora were indicated more than a decade ago [6], we still could not say that investigations based on parallel data became mainstream among typologists.

At first sight, it might seem quite obvious that multilingual parallel corpora is a tool that suits the needs of typological researchers perfectly, but there are some issues regarding their usage that should be discussed in detail. As Natalia Levshina pointed out [15], there is a lack of data for the vast majority of the languages. Indeed, most of the corpora we have are based on Indo-European data, but that is not the only difficulty with parallel corpora. There are some other aspects of modern parallel corpora design that complicate their usage.

For example, the most popular and most rich in translations text is the Bible (over 100 languages, in corpus [2]), but the language used in the Bible has several differences from the modern language, which means it might not be useful for complex syntactic studies. Another aspect is the fullness of the corpus content. Some corpora contain lots of texts in different languages, but not each text has corresponding translations in all the languages represented in the corpus. It means one can hardly collect a fully aligned data set and has to decide if one is more interested in multiple translation units or if one is satisfied with a smaller number of units in order to involve a greater number of languages into the data set and have no gaps (missing translations) in the data. The tagging used in a parallel corpus also plays a big role, and thus, the most convenient option to avoid tagging mismatches between languages is to have a universal annotation

scheme. The corpora that have recently appeared are usually tagged with Universal dependencies guidelines[16]. Despite all these nuances the number of studies based on parallel data seems to increase in the last few years. One definitely can experiment with different approaches even using smaller data sets and get noteworthy results.

In this paper, I will show how linguists can benefit from a parallel corpus by investigating functional properties of passives with agent phrases (PAP). Also, I will present a preliminary analysis of PAP vs. related constructions, such as voice constructions. Since voice constructions is a phenomenon dependent on semantic and pragmatic factors, studies of grammatical voice or alternations involve a comparison of situations and their participants in a similar contextual environment. For a typological study of voice constructions parallel corpora give an opportunity to extract fully parallel translation units and, based on them, analyze situations with corresponding constructions cross-linguistically. And that is not the only advantage we obtain from parallel data for cross-linguistic analysis of PAP. As we will see, the parallel corpora approach helps to figure out language patterns that are beyond the reach of traditional studies of passives. Last but not least, parallel data with semantic annotation allows to train several statistical models, each corresponding to a certain language from the sample, and draw cross-linguistic analysis based on the resulting parameters of these models.

The rest of the paper is organized in the following way. Section 2 is devoted to the major characteristics of PAP; I compare this construction to agentless passives and discuss in what way it is specific and merits consideration. Then, in Section 3 I pay more attention to the benefits parallel data provide for the research of functional properties of voice constructions. In Sections 4 and 5 I describe the data, my experiments with PAP in European languages, and share my observations on the results.

2 Passives with Agent Phrase

Passive constructions have been a field of interest among linguists for a long time, and there are plenty of studies in this topic. The example below illustrates the difference between active and passive voice:

Example 1.

1. The paparazzi saw Zelda at the party.
2. Zelda was seen by the paparazzi at the party. [25]

Both sentences describe a situation with the same number of participants, in the first sentence “the paparazzi” is the subject and “Zelda” is the object, but in the second sentence “Zelda” takes the subject position while “paparazzi” is an oblique phrase. According to the book by F. Zúñiga and S. Kittilä [25] the prototypical passive voice possesses the following characteristics:

1. Syntactic valency is one less than in the active diathesis (e.g., the verb is monovalent when its active counterpart is bivalent).
2. Its subject corresponds to the non-subject P of the active voice.
3. Its peripheral, and optional, argument (typically marked by a non-core case or adposition) corresponds to the subject A of the active voice.

4. Passivization is formally coded on the predicate complex.

Similar considerations can be found in the works by K. Kazenin [9], L. Kulikov [13], M. Shibatani [21]. These features are quite basic and clearcut except for the third one. The syntactic status of the agent phrase is often formulated in such terms that the agent phrase is optional. It seems to be misleading because PAP can not be interchangeably used with agentless passives. Generally, one distinguishes between different types of passive constructions, and agentless passive can be classified as a certain construction type. From a typological perspective, we can say that there are languages that have only agentless passives, and there are also languages that have both agentless passives and PAP. However, in the language that has these two constructions, the functions of agentless passives and PAP are different since they appear in different contexts. The situations with fully demoted agents have other interpretations than PAP. Kiparsky also questioned the statement about the optionality of the agent phrase in his study [12], and he pointed out that situations encoded by agentless passives have a human agent that was demoted, while the agent phrase in PAP may correspond to different semantic roles. Kiparsky illustrates this by the following set of examples [12, p. 29]:

Example 2.

1. The castle is surrounded on all sides by water.
2. The castle is surrounded on all sides. [human surrounders only]

Example 3.

1. John was seen breaking into the house. [the seer is human]
2. John was seen breaking into the house by the dog.

Example 4.

1. The cave was entered. [the enterer is a person – not smoke, or an animal]
2. The peritoneal cavity was entered by a bullet.

Example 5.

1. It was expected that there would be food in the house. [can't be said of a raccoon]

Siewierska and Bakker also mention in their study [22] that the nature of the potential agent may be seen as performing a distinguishing role. In the case of agentless passives, the underlying agent is a human, but when it is expressed, it is not necessarily a human, it can be an animal or a natural force.

All this means that the question about the agent phrase should not be formulated in terms of optionality, since the semantic motivation for using agentless passives are quite clear. The thing which remains unclear is the distribution of agent phrases corresponding to different semantic roles and whether there are some lexical factors that influence the use of agent phrases.

In this paper, we focus on PAP construction more from a functional perspective. The observations about the functional properties of passives from literature can relate to passives as a phenomenon in the broad sense. In the book by F. Zúñiga and S. Kittilä [25], the authors list syntactic, semantic, and discourse-related motivations for using passive constructions instead of active. First, passive can be a tool for creating a syntactic pivot (see also [21]), like in the following example:

Example 6.

1. My friend[i] (S) arrived and Ø[i] (S) laughed.
2. My friend[i] (A) saved the boy[j] (P) and Ø[i/*j] (S) laughed.
3. The boy[j] (P) was saved by my friend[i] (A) and Øp*i/j] (S) laughed.

Sometimes due to syntactic restrictions, a speaker is forced to use passive instead of active, which is considered as a syntactic function of passives. Second, the semantic motivation for using passive lies in the need to express a lower transitivity value and stativization [8], [24]. The last one is probably the most prominent function of passives and the most debatable one, which is about P (patient) foregrounding and A (agent) demotion mechanisms. Some researchers pay more attention to the first issue [10], [19], and there is a point of view, found in [3], [21], that A demotion is a primary function of passives. But in some studies both P foregrounding and A demotion are regarded to be equally important, each one in its own way [24].

These functions are quite basic and we can not say that this list of functions is exhaustive. Can they be applied to every type of passives? Yes, since they are too general. Will they then fully characterize the peculiarities of the use of each passive construction type? Probably not, because it seems that different types of passives can reveal functions that are specific especially for them, and serve for very fine differences in construction usage. The issue I focus on is what can be the functional characteristics specific to PAPs; it seems that it goes beyond P's foregrounding, A's demotion, and stativization. We also try to find some evidence for Kiparsky's conclusion that lexical and semantic factors govern the distribution of agent phrases.

In the next section, I discuss the issues of using parallel data for voice constructions exploration. First, I briefly review Sansò's study [19] of (agentless) passives and impersonal constructions in functional perspective based on multilingual data, and then I move to my study of PAP.

3 Parallel Data and Functional Properties of Passives

As I pointed out previously, parallel data can be very useful for voice constructions exploration, especially for their functional properties. Several motivations for using passives mentioned in literature seem to be quite general and need to be studied in more detail. Traditional methods seem to be inappropriate for these purposes. Using parallel texts makes a big difference, one can focus on a particular situation and compare encodings (grammatical and lexical forms) used for its expression in different languages. Analysis of multiple parallel situations reveals distributions of related constructions and highlights the functions of the target construction.

The paper of Andrea Sansò [19] is an example of how parallel data can be used for a functional study of voice constructions. In his work, he demonstrates that passives and impersonals, which were previously considered to be similar, form a functional cline and actually have differences in their usage. Sansò's language sample includes Italian, Spanish, Polish, Danish, and Modern Greek. According to his work, there are at least three prototypical situation types, namely, "patient-oriented process", "bare happening"

and “agentless generic event”, that can be encoded by passives, middles, and impersonal constructions in the languages from his sample. For a patient-oriented process, the corresponding reason for defocusing the agent is that the agent is less discourse-central and individuated than the patient. Bare happening reflects conceptualization of the event depicted by the verb as a naked fact, e.g. “So the faith of the simple was mocked, the mysteries of God were eviscerated (or at least this was tried, fools they who tried), questions concerning the loftiest things were treated recklessly, the fathers were mocked” [19, p.241]. And the third situation type, agentless generic event, corresponds to situations with generic agent:

Example 7. German

Dieses Buch liest sich gut.
This book reads well. [11, p.147]

Example 8. French

Cela ne se dit pas.
This is not said/one doesn't usually say this. [19, p.243]

The situation types go hand in hand with the levels of agent defocusing or, to put it differently, the reasons for agent defocusing. For each language of the sample Sansò calculates a distribution of encoding strategies across situation types that appears to be statistically significant. Table 1 is an example of such distribution in Polish (similar calculations were also done for other languages).

Table 1. Passives/impersonal constructions and associated situation types in Polish, $\chi^2 = 473.49, p < 0.05$.

Situation type	Periphrastic passive	-no/-to construction	Middle construction
Patient-oriented process	87.39%	17.25%	2.61%
Bare happening	9.95%	79.31%	14.38%
Agentless generic event	2.76%	3.44%	83.01%

As a result Sansò elaborates a cline where certain constructions correspond to certain situation types, which proves his initial suggestion about the diversity of functions that passives and impersonal constructions have across languages.

Sansò's work demonstrates how parallel data can help reveal fine and not quite obvious differences between closely related constructions. And here I would like to emphasize some important features of multilingual parallel corpora for typological studies. One of the huge benefits we get from parallel data is the possibility of annotation transfer. In studies of voice and other grammatical alternations, semantic factors play an important role, which means one would probably need to make semantic annotation of the data. Even though nowadays there are many advanced annotation tools for different purposes, it is very unlikely that they will suit the needs of the researcher. Since in a study one would probably be interested in some very specific semantic features, like types of situations, or something that the researchers elaborate themselves. That means

the option of an automatic annotation is not available and one should annotate the data manually, which is quite an exhaustive task. Luckily, parallel texts allow us to annotate the situations of the translation units and then transfer this annotation to each translation. Another thing that parallel corpora make possible is various quantitative assessments, e.g., calculation of construction distributions across situation types, as in Sansò's study, as well as building statistical models with the evaluation of feature importances. Models built on parallel data allow to compare the resulting scores obtained for different languages. Fully aligned translation units that capture the pragmatic environment of situations and complex annotation create a space for advanced analysis with many research possibilities.

In my experiment, I try to figure out the special properties of PAP construction that have not been studied in detail before. As we see, agentless passives and impersonals can be matched to certain situation types. Taking into account the basic knowledge about passives functions, I propose some issues regarding PAP that will be discussed in detail further. First, PAP construction is usually opposed to active but is it always so or maybe there are other alternatives? Anticipating events I can say, that active is in fact not the only alternative to passive, which raises another question. What are the distributions of PAP and other constructions in the languages of the sample in use? And if there are differences, what do they tell us? Preliminary one can suppose, that if differences do take place then it indicates functional differences in PAP usage and there should be some kind of semantic motivation for that.

In the next section, I describe the corpus data I use and then in section 5 I move to the quantitative analysis of PAP and the results I obtained.

4 Corpus and Data

For this study I used a corpus of seven Harry Potter books in nine languages, English, German, Swedish, Italian, Spanish, French, Russian, Czech and Bulgarian. The amount of the text data per language is about 1 million tokens. All the texts were aligned with Gale&Church algorithm [7] at the sentence level and with Efmaraal toolkit [18] at word level. For morphological and syntactic annotation I used the UDPipe parser with Universal Dependencies models [16], [23]. Alignment and annotation was a primary processing stage, which allowed us to extract sentences with PAP constructions in each language from the sample and corresponding translations. The final data set included 983 translation units with a full set of translation equivalents for each unit.

5 Quantitative Analysis of Passives with Agent Phrase

In this section I provide some examples of PAP constructions with possible translation equivalents, describe the experiment and demonstrate the importance of the "agent role" factor.

5.1 The Distribution of Constructions in Translations

In grammars, PAP is usually opposed to transitive active sentences, but in my data, I found numerous cases, where PAP corresponds to some other construction than active.

There are several alternatives, it can be a passive with oblique agent phrase encoded differently than a typical passive agent, an existential or locative construction like “there is/are” or “X has Y”, an adjective and others, see the examples below.

Example 9. English “there is X on Y” vs. Russian PAP

English <...> her robes were ripped in several places and **there were numerous scratches on her face and arms.** (*J.K. Rowling, Harry Potter and the Order of the Phoenix, chapter 30*)

Russian <...> mantija porvana v neskol’kih mestah, a lico i ruki **ispeschreny carapinami.** (*Translation by V. Babkova, V. Golysheva, L. Motyleva*)

Example 10. English “X has Y” vs. Russian PAP

English He seems to **have sprouted little tentacles all over his face.** (*J.K. Rowling, Harry Potter and the Goblet of Fire, chapter 37*)

Russian Teper’ u etogo tipa vse lico **pokryto malenkimi schupalcami.** (*Translation by M.D. Litvinova*)

Example 11. English gerund vs. French PAP

English **Binding magical contract**, like Dumbledore said. (*J.K. Rowling, Harry Potter and the Goblet of Fire, chapter 17*)

French **Ils sont liés par un contrat magique**, comme l’a dit Dumbledore. (*Translation by Jean-François Ménard*)

For further analysis, I decided to merge the non-active alternatives into one type and determined three types of constructions used in translations: PAP, active, lexical variation.

Table 2. The distribution of construction types with respect to the semantic role of potential agent phrase, $\chi^2 = 569.58, p < 0.001$.

Semantic role	Active	PAP	Lex.Var.	Total
Agent	999	1097	532	2628
Non-agent	1311	2033	2875	6219
Total	2310	3130	3407	8847

There is one more notable thing regarding PAP, which is the semantic role of the participant the agent phrase corresponds to. In literature, one can find suggestions indicating that the semantic role factor probably is important. It was already mentioned in the paper of Shibatani [21], where he claimed agent demotion to be a salient feature of passives, that it is the agent that should be demoted and not the other participant even if can take the subject position in an active sentence. Similar thoughts about the agent role are found in Langaker’s work [14]. In addition to that, Siewierka [22] and Kiparsky [12] point out later that the distribution of agent phrases is probably governed by lexical and semantic factors. Taking that into account, I decided to check if there is any connection between the “agent role” factor and the distribution of constructions. I annotated each translation unit according to the feature that can be formulated as “Does the potential agent phrase correspond to the agent role?”. After that I was able to calculate

the distribution of construction types with respect to the semantics of the corresponding situations, results presented in Table 2.

This distribution shows that there are two semantically motivated sets of translations with different patterns of coding, and this distribution also appears to be statistically significant. The first group is a set of situations with semantic agents that generally correspond to PAP or active clauses, the second group, which appears to be almost of the same size, corresponds to the PAP vs. lexical variation cases.

At this point, I can claim that PAP definitely has some sort of semantic function other than just stativization of the verb. As I described the overall distribution, I will move to the individual distributions in languages of the sample.

5.2 PAP and the Agent Role Across Languages

As we have seen in the previous subsection the distribution of PAP and other constructions reveals a peculiar pattern. That distribution was calculated based on all the translations without any distinction made between the languages. Let us look closer at the distributions within each language in Table 3.

Table 3. The distribution of construction types in languages of the sample.

Language	PAP	Active	Lex.var.
English	0.48	0.12	0.40
German	0.28	0.27	0.45
Swedish	0.55	0.14	0.31
Russian	0.23	0.38	0.39
Czech	0.15	0.5	0.35
Bulgarian	0.36	0.23	0.41
Italian	0.52	0.14	0.33
French	0.45	0.19	0.35
Spanish	0.16	0.38	0.46
Translation units total: 983			

Obviously, some languages prefer PAP construction more than other languages which are prone to active and lexical variations. In order to get a more accurate picture of PAP behavior in the languages, I built binary logistic regression models with a single predictor which encodes, *if the potential agent phrase corresponds to the agent role*. The dependent variable has two values, namely, PAP vs. non-PAP. The summary of the models can be found in Table 4.

The p-values greater than 0.05 indicate that the results obtained in this model are not significant, which is true for Italian, Swedish and Spanish models. It means the predictor I used is not relevant for the distinction between PAP and non-PAP in these languages. The positive coefficients indicate that the predictor positively influences the choice of PAP construction; negative coefficients tell us the opposite. In English and German, using only this so-called “agent role” predictor, one can distinguish between PAP and

Table 4. Coefficients for the "agent role" predictor in logistic regression models by language.

Language	Accuracy	Coefficient	p-value	Intercept	p-value
English	0.69	1.93	<0.001	-0.62	<0.001
German	0.7	1.35	<0.001	-1.43	<0.001
Bulgarian	0.59	0.36	0.013	-0.7	<0.001
French	0.6	0.78	<0.001	-0.42	<0.001
Russian	0.47	-1.34	<0.001	-0.91	<0.001
Czech	0.42	-1.67	<0.001	-1.4	<0.001
Italian	0.52	0.31	0.026	-0.009	0.909*
Swedish	0.52	0.42	0.003	0.09	0.210*
Spanish	0.34	-0.15	0.439*	-1.63	<0.001

non-PAP with the accuracy of 0.7; the values for Bulgarian and French are lower. In Russian and Czech the predictor's coefficient is negative; this can be explained by the fact that PAP construction is non-frequent in these languages according to the data set.

The quantitative analysis helped us to detect major tendencies that languages have in the use of PAP and alternative constructions, but of course, the issues about PAP are not limited to those I have mentioned.

6 Conclusion

To sum up, I would like to emphasize one more time that multilingual parallel data takes comparative cross-linguistic analysis to the next level. The possibility to work with numerous observations and quantitatively assess them allows to analyze functional properties of constructions from a broader perspective.

The preliminary analysis of PAP has shown that the use of PAP is not limited to the discourse oriented mechanisms of participants promotion and demotion and is also semantically motivated. According to the data, there are languages that mostly use PAP as a counterpart of active construction (e.g., English, German) and those with predominant semantically motivated PAP cases (e.g., Russian, Czech). It seems that the group of observations with PAPs that have a non-agent participant in the agent phrase and the corresponding constructions classified as "lexical variations" definitely need a more thorough analysis, and I leave this for my future research.

References

1. Asgari, E., Schütze, H.: Past, Present, Future: A Computational Investigation of the Typology of Tense in 1000 Languages. In: Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing. pp. 113–124. Association for Computational Linguistics, Stroudsburg, PA, USA (2017). <https://doi.org/10.18653/v1/D17-1011>
2. Christodouloupoulos, C., Steedman, M.: A massively parallel corpus: the Bible in 100 languages. *Language Resources and Evaluation* **49**(2), 375–395 (jun 2015). <https://doi.org/10.1007/s10579-014-9287-y>

3. Comrie, B.: Passive and voice. p. 9 (1988). <https://doi.org/10.1075/tsl.16.04com>
4. Crible, L., Abuczki, Á., Burkšaitienė, N., Furkó, P., Nedoluzhko, A., Rackevičienė, S., Oleškevičienė, G.V., Zikánová, Š.: Functions and translations of discourse markers in TED Talks: A parallel corpus study of underspecification in five languages. *Journal of Pragmatics* **142**, 139–155 (mar 2019). <https://doi.org/10.1016/j.pragma.2019.01.012>
5. Cysouw, M.: Inducing semantic roles. pp. 23–68 (2014). <https://doi.org/10.1075/tsl.106.02cys>
6. Dahl, Ö.: From questionnaires to parallel corpora in typology. *Language Typology and Universals* **60**(2), 172–181 (jul 2007). <https://doi.org/10.1524/stuf.2007.60.2.172>
7. Gale, W.A., Church, K.W.: A Program for Aligning Sentences in Bilingual Corpora. *Computational Linguistics* **19**(1), 75–102 (1993), <https://www.aclweb.org/anthology/J93-1004>
8. Haspelmath, M.: The Grammaticization of Passive Morphology. *Studies in Language* **14**(1), 25–72 (jan 1990). <https://doi.org/10.1075/sl.14.1.03has>
9. Kazenin, K.I.: Passive voice. In: *Language Typology and Language Universals*, pp. 899–915 (2001)
10. Keenan, E.L., Dryer, M.S.: Passive in the world’s languages. In: Shopen, T. (ed.) *Language Typology and Syntactic Description*, pp. 325–361. Cambridge University Press, Cambridge. <https://doi.org/10.1017/CBO9780511619427.006>
11. Kemmer, S.: *The Middle Voice, Typological Studies in Language*, vol. 23. John Benjamins Publishing Company, Amsterdam (oct 1993). <https://doi.org/10.1075/tsl.23>
12. Kiparsky, P.: Towards a null theory of the passive. *Lingua* **125**(1), 7–33 (2013). <https://doi.org/10.1016/j.lingua.2012.09.003>
13. Kulikov, L.: *Voice Typology*. Oxford University Press (nov 2010). <https://doi.org/10.1093/oxfordhb/9780199281251.013.0019>
14. Langacker, R.W.: Dimensions of defocusing. In: *Voice and Grammatical relations*, pp. 115–137 (2006). <https://doi.org/10.1075/tsl.65.08lan>
15. Levshina, N.: Why we need a token-based typology: A case study of analytic and lexical causatives in fifteen European languages. *Folia Linguistica* **50**(2) (jan 2016). <https://doi.org/10.1515/flin-2016-0019>
16. Nivre, J., de Marneffe, M.C., Ginter, F., Goldberg, Y., Hajič, J., Manning, C.D., McDonald, R., Petrov, S., Pyysalo, S., Silveira, N., Tsarfaty, R., Zeman, D.: Universal dependencies v1: A multilingual treebank collection. In: *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC’16)*. pp. 1659–1666. European Language Resources Association (ELRA), Portorož, Slovenia (may)
17. Östling, R.: 6. Studying colexification through massively parallel corpora. In: *The Lexical Typology of Semantic Shifts*. De Gruyter, Berlin, Boston (2016). <https://doi.org/10.1515/9783110377675-006>
18. Östling, R., Tiedemann, J.: Efficient Word Alignment with Markov Chain Monte Carlo. *The Prague Bulletin of Mathematical Linguistics* **106**(1), 125–146 (oct 2016). <https://doi.org/10.1515/pralin-2016-0013>
19. Sansò, A.: ‘Agent defocusing’ revisited. In: *Passivization and typology*, pp. 232–273 (2006). <https://doi.org/10.1075/tsl.68.15san>
20. Schlund, K.: Active transitive impersonals in Slavic and beyond: a parallel corpus analysis. *Russian Linguistics* **44**(1), 39–58 (apr 2020). <https://doi.org/10.1007/s11185-020-09221-2>
21. Shibatani, M.: Passives and Related Constructions: A Prototype Analysis. *Language* **61**(4), 821 (dec 1985). <https://doi.org/10.2307/414491>
22. Siewierska, A., Bakker, D.: Passive agents: prototypical vs. canonical passives. In: Brown, D., Chumakina, M., Corbett, G.G. (eds.) *Canonical Morphology and Syntax*, chap. Passive ag (2013). <https://doi.org/10.1093/acprof>

23. Straka, M., Straková, J.: Tokenizing, POS Tagging, Lemmatizing and Parsing UD 2.0 with UDPipe. In: Proceedings of the CoNLL 2017 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies. pp. 88–99. Association for Computational Linguistics, Stroudsburg, PA, USA (2017). <https://doi.org/10.18653/v1/K17-3009>
24. Thompson, C.L.: Passive and inverse constructions. In: Voice and Inversion, p. 47 (1994). <https://doi.org/10.1075/tsl.28.05tho>
25. Zúñiga, F., Kittilä, S.: Grammatical Voice. Cambridge University Press (feb 2019). <https://doi.org/10.1017/9781316671399>