A multifactorial analysis on dative alternation in Russian

Yue Zou
yuezou@shisu.edu.cn
Hao Lin
linhao@shisu.edu.cn

Shanghai International Studies University

1 Introduction

Dative alternation in English refers to the alternative use of two near-synonymous constructions: the double object construction (give him a book) and the prepositional dative construction (give a book to him). Similarly, as illustrated in Examples (1-2), an alternation that is related to the relative order of theme and recipient constituents can also be observed in Russian.

(1) Ona dala Ivanu knigu.
She give-PAST Ivan-DAT book-ACC.
'She gave Ivan a book.'

(2) Ona dala knigu Ivanu.
She give-PAST book-ACC Ivan-DAT.
'She gave a book to Ivan.'

A large number of studies have been conducted to explore the factors that potentially have impact on the alternation in English (Gries, 2003; Bresnan et al., 2007; Bresnan & Ford, 2010; Kendall et al., 2011). Specifically for Russian, however, research on dative alternation is quite limited. On the one hand, researchers often had to rely on intuition to choose the prototypical verbs that can occur in the alternation. For instance, in Kizach (2012), the behavior of only one verb (dat') is investigated. On the other hand, the majority of studies on the alternation in Russian are monofactorial in nature (Kizach, 2012; Mykhaylyk et al., 2013). The relative importance of each independent variable compared to other ones and their interactions remain largely unknown.

As mentioned in Jaeger & Norcliffe (2009), similar constructions may be processed differently in different languages. In fact, it has been observed in previous studies (Bresnan & Hay, 2008; Gries & Bernaisch, 2016; Röthlisberger et al., 2017) that even in the same language (English), the effect sizes of some influencing factors vary in different dialects. Therefore, a detailed investigation of dative alternation in Russian is highly warranted.

The purpose of the present study is twofold: to expand the empirical base of Russian verbs that allow dative alternation, and to explore the factors that constrain the alternation in Russian. As will be shown below, our multifactorial analysis yields several interesting findings. In combination with previous research on the dative alternation in English and Chinese, our study can contribute to the discussion of identifying the interaction between language-specific characteristics and universal production principles.

2 Methods and results

2.1 Which verbs allow dative alternation in Russian?

In order to identify all the verbs that can occur in dative alternation, we parsed the offline version of the Russian National Corpus (RNC) with an NLP package `Stanza` (Qi et al., 2020) in Python. The offline version of the RNC contains texts in modern Russian and is of about one million words in size, where fiction, academic and journalistic texts, transcripts of oral speech and blogs are represented in roughly equal proportion.

After running the part-of-speech processor and the dependency parsing processor, we extracted all verbs that meet the following three requirements:

1. the verb is the syntactic head of an accusative noun, and the dependency relation between the two words is 'obj';
2. the verb is the syntactic head of a dative noun, and the dependency relation between the two words is 'iobj';
3. the verb can occur in both V-DAT-ACC and V-ACC-DAT patterns.

A total of 1668 verbs, which consist of 237 types, are identified. The top 10 most frequent verbs are

Proceedings of the Society for Computation in Linguistics (SCiL) 2024, pages 311-315.
Irvine, California, June 27-29, 2024

provided in Table 1. The complete list of verbs that allow dative alternation can serve as a reference for future research and is available for download at https://osf.io/3y8hj/.

As in English, dat' (give) is the most prototypical (frequent) verb in dative alternation in Russian. The result provides empirical justification (1) for choosing dat' as the only verb in previous research on the alternation in Russian and (2) for choosing dat' and its morphologically related verb otdat' (return) (together with their imperfective counterparts) for further detailed analysis in the present study.

Verb	Count
dat'	146
davat'	81
delat'	63
skazat'	49
pokazat'	47
otdat'	41
govorit'	37
pokazyvat'	25
darit'	23
otdavat'	22

Table 1: The top 10 most frequent verbs occurring in dative alternation.

2.2 What factors constrain dative alternation in Russian?

Instances of dative alternation containing the target verbs are extracted from the main RNC (URL: https://ruscorpora.ru/) and annotated with a number of variables. The sample contains 1299 instances from 226 writers. The coding process of each variable is briefly described below.

Chosen construction - the dependent variable is binary. A sentence is coded as an instantiation of the recipient-theme construction if the recipient precedes the theme, as in Example (1); it is coded as an instantiation of the theme-recipient construction if the recipient follows the theme, as in Example (2).

Pronominality of theme and **pronominality of recipient** - whether or not a constituent is pronominal is coded as a binary variable (pronoun and non-pronoun). The two variables and their interaction are included in our study to test the availability account (Ferreira & Dell, 2000), according to which the syntactic structure of utterances is sensitive to the accessibility of lexical information. The constituent that is in the pronominal form contains more accessible information, and is expected to occur before the non-pronominal one.

Length difference - length of theme and recipient is annotated by automatically counting the number of words in each constituent. The variable is recoded as five ordered levels because of its extreme skewed distribution: theme longer than recipient by more than 1 word (117 cases) > theme longer than recipient by 1 word (260) > equally long (729) > recipient longer than theme by 1 word (106) > recipient longer than theme by more than 1 word (87). It is included to test the hypothesis in Hawkins (2001, 2003, 2004) that "the human processor would prefer linear orders that minimize dependency processing effort". At the surface level, the hypothesis predicts that a construction where the shorter constituent occurs before the longer one should be preferred.

Animacy of recipient and **concreteness of theme** - for animacy, two levels are distinguished: animate (including human and animal) and inanimate. As in Bresnan et al. (2007), concreteness of theme (two levels: concrete or abstract) is included to compensate for the simplified binary distinction of animacy. Following previous literature (Bresnan et al., 2007; Zhang & Xu, 2023), we expect the theme-recipient construction to be preferred when the recipient is animate and the theme is concrete.

Verb and **aspect** - two pairs of verbs are included in our analysis: otdavat'-otdat' and davat'-dat'. The first verb in each pair is the imperfective, and the second is the perfective. VERB is included in the model to control for lexical idiosyncrasy. ASPECT is included to test if a change in aspect is accompanied by a shift in lexical meaning. An interaction is also considered, as changing from the perfective to the imperfective may either increase or decrease the verb's preference for one construction.

Additionally, the analysis includes a random intercept for each writer, which can be thought of as the individual adjustment to each writer's personal preference.

A mixed-effects regression analysis is implemented with the `lme4` package in R. Model selection was performed following the two-step strategy outlined by Zuur et al. (2009). R^2_{marginal}, $R^2_{\text{conditional}}$, and C-score of the final model are 0.738, 0.787 and 0.957, respectively. Classification accuracy of the final model is 0.895, which is significantly higher than the baseline (p_{binom} <

	Estimate	SE	z value	p value
(Intercept)	-0.970	0.386	-2.517	0.012
recipient animacy $_{\text{inanimate}}$	-0.583	0.258	-2.258	0.024
verb $_{\text{otdavat'-otdat'}}$	-0.584	0.344	-1.697	0.090
aspect $_{\text{perfective}}$	1.147	0.388	2.954	0.003
recipient pronominality $_{\text{pronoun}}$	3.674	0.320	11.491	< 0.001
theme pronominality $_{\text{pronoun}}$	-1.304	0.419	-3.108	0.002
length difference $_{\text{2-1}}$	-1.207	0.504	-2.394	0.017
length difference $_{\text{3-2}}$	-0.658	0.266	-2.476	0.013
length difference $_{\text{4-3}}$	-1.455	0.340	-4.280	< 0.001
length difference $_{\text{5-4}}$	-5.476	1.187	-4.611	< 0.001
verb $_{\text{otdavat'-otdat'}}$: aspect $_{\text{perfective}}$	-1.353	0.465	-2.913	0.004
recipient pronominality $_{\text{pronoun}}$: theme pronominality $_{\text{pronoun}}$	-1.980	0.595	-3.331	< 0.001

Table 2: Result summary: coefficient estimates, standard errors, z scores and p values for all main predictors and their interactions in the final model. Length difference is coded with the method of successive differences contrast. The first level is [theme longer than recipient by more than 1 word] and the last (fifth) level is [recipient longer than theme by more than 1 word].

0.001). No issue of multicolinearity or overdispersion is detected. A summary of results is provided in Table 2.

3 Discussion

Overall, the results in our study align remarkably well with previous findings in English and Chinese. As can be seen in Table 2, the probability of choosing the recipient-theme construction increases when the recipient is animate, pronominal, when the theme is nominal, and when the recipient is shorter than the theme. The observed effects of pronominality of theme, pronominality of recipient, animacy of recipient and length difference provide novel supporting evidence for the availability account and the dependency processing account.

Moreover, our analysis can shed light on some broader aspects of Russian grammar, and the effects of several other factors deserve further attention. In the following sub-sections, we discuss issues regarding default order, influence of verb semantics, and structural difference of dative alternation between Russian and English.

3.1 Is there a default order?

There exist mainly two views on default order for dative alternation in Russian. Within the generative framework, Junghanns & Zybatow (1997) and Dyakonova (2007) argued for treating the (DAT-ACC) order (recipient-theme construction) as the default, while Bailyn (1995, 2010) presented an opposite relationship. Potentially, a third view that is held in Gries (2003), where the alternating constructions do not form a single category (i.e. no default order), is also possible. Here we approach the issue by exploring the middle ground of the alternation.

	ACC-DAT	DAT-ACC
ACC-DAT	206	23
EITHER	182	168
DAT-ACC	33	687

Table 3: Confusion matrix for three-way classification. Each row indicates the number of the predicted construction, and each column indicates the number of the observed construction.

Based on the final model of our multifactorial analysis, we generated a 95% confidence interval of the predicted probability for each instance with the `bootMer` function in R. If the confidence interval includes 0.5, then the categorical prediction is changed to "either". The confusion matrix for the three-way classification is shown in Table 3.

We observed an asymmetrical relationship between the two constructions: the theme-recipient (ACC-DAT) construction in a sentence can often be replaced by a recipient-theme (DAT-ACC) construction (in 43.2% of cases), while the recipient-theme (DAT-ACC) can be replaced by a theme-recipient (ACC-DAT) construction only in a limited number of cases (19.1%).

The result is compatible with the first view and supports treating the recipient-theme construction as the default. On the one hand, the recipient-theme construction can have either a wide focus

(on the whole construction) or narrow focus (on the theme), which accounts for the fact that it can be used in more diverse contexts. On the other hand, the theme-recipient construction only has a narrow focus on the theme. Therefore, we argue that the recipient-theme construction is more "neutral", which is also consistent with a more general view of treating the DAT-ACC order as the default in Russian (Shvedova, 1980).

3.2 How does verb semantics influence choice of alternating constructions?

The effects of verb, aspect and their interaction indicate that, as in English and Chinese, choice of alternating constructions depend on verb semantics in Russian. More specifically, davat' and dat' have a higher preference for the recipient-theme construction. In addition, we find that the predicted probability of choosing the recipient-theme construction is significantly higher for the perfective dat' than the imperfective davat'.

The result regarding verb semantics is broadly consistent with previous findings in English and Chinese, where the verb for giving ("give" in English and "gei" in Chinese) is identified as the a distinctive collexeme for the ditransitive (recipient-theme) construction. Moreover, we argue that the difference in constructional preference between the perfective dat' and the imperfective davat' reflects their semantic nuances. For instance, it is mentioned in Podlesskaya (2006) that only the imperfective davat' can be used as a commercial transaction verb, which is less compatible with the ditransitive (recipient-theme) construction in English (Gries & Stefanowitsch, 2004). However, more verbs need to be included in future research in order to make a compelling argument about the semantic difference between the alternating constructions in Russian.

3.3 Why is the effect of concreteness insignificant?

Dative alternation in Russian is special in the sense that it is a "pure" constituent order alternation. Comparing our results with findings in other languages enables us to differentiate between factors that are related to the change in constituent order and factors that are linked to something else.

In English and Chinese, the alternation actually involves two types of change: a change in constituent order and a change in presence or absence of an overt preposition ("to" in English and "gei" in Chinese). The prepositional dative construction in English is derived from the caused-motion construction (Goldberg, 1995) and is associated with a movement meaning: "X causes Z to move to Y" (Gries & Stefanowitsch, 2004). An instantiation of the prepositional dative construction with an abstract theme is dispreferred in English (e.g. ? That movie gave the creeps to me.) as the theme and the construction are semantically incompatible with each other. In the same vein, the preposition "gei" in the prepositional dative construction in Chinese is also related to some kind of caused-motion (Zhang & Xu, 2023).

Unlike previous studies in English and Chinese, we did not detect a significant effect of concreteness of theme on dative alternation in Russian. Our analysis indicates that concrete and abstract themes are equally acceptable in the alternating constructions. Therefore, we argue that the effect of concreteness is associated with presence or absence of an overt preposition, rather than constituent order.

4 Conclusion

The present study investigates dative alternation in Russian to uncover factors that constrain the alternation with mixed-effects regression modeling. Overall, we find that the results in Russian align remarkably well with previous findings in English and Chinese. Choice of dative alternation is largely conditioned by the availability and dependency processing constrains in all three languages.

On the basis of our multifactorial analysis, a few broader issues are discussed. First, by exploring the middle ground of dative alternation, we identified the recipient-theme construction as the default and more "neutral" one in the alternation in Russian. Second, we found that verb semantics also has an influence on the alternation in Russian. The preference of verbs of giving for the recipient-theme construction may be cross-linguistic, although a future study in Russian that includes more verbs is needed to confirm the argument. Third, we attributed the effect of concreteness of theme to the use of an overt preposition, which accounts for the fact that it is statistically significant in English and Chinese, but not in Russian.

Acknowledgments

The research was sponsored by the Academic Mentorship Project (Project No. 2023DSYL010). We are grateful to the three anonymous reviewers for

their constructive comments that greatly improved the quality of the paper.

References

J. F. Bailyn. 1995. Underlying phrase structure and "short" verb movement in Russian. *Journal of Slavic linguistics*, 3(1):13–58.

J. F. Bailyn. 2010. What's inside VP? New (and old) evidence from Russian. In *Formal Approaches to Slavic Linguistics*, pages 21–37. Michigan Slavic Publications.

J. Bresnan, A. Cueni, T. Nikitina, and R. H. Baayen. 2007. Predicting the dative alternation. In *Cognitive foundations of interpretation*, pages 69–94. KNAW.

J. Bresnan and M. Ford. 2010. Predicting syntax: Processing dative constructions in American and Australian varieties of English. *Language*, 86(1):168–213.

J. Bresnan and J. Hay. 2008. Gradient grammar: An effect of animacy on the syntax of give in New Zealand and American English. *Lingua*, 118(2):245–259.

M. Dyakonova. 2007. Russian double object constructions. *ACLC Working Papers*, 2(1):3–30.

V. S. Ferreira and G. S. Dell. 2000. Effect of ambiguity and lexical availability on syntactic and lexical production. *Cognitive psychology*, 40(4):296–340.

A. E. Goldberg. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. The University of Chicago Press.

S. Th. Gries. 2003. Towards a corpus-based identification of prototypical instances of constructions. *Annual Review of Cognitive Linguistics*, 1(1):1–27.

S. Th. Gries and T. Bernaisch. 2016. Exploring epicentres empirically: focus on South Asian Englishes. *English World-Wide*, 37(1):1–25.

S. Th. Gries and A. Stefanowitsch. 2004. Extending collostructional analysis: A corpus-based perspective on alternations. *International journal of corpus linguistics*, 9(1):97–129.

T. F. Jaeger and E. J. Norcliffe. 2009. The cross-linguistic study of sentence production. *Language and Linguistics Compass*, 3(4):866–887.

U. Junghanns and G. Zybatow. 1997. Syntax and information structure of Russian clauses. In *Annual workshop on formal approaches to Slavic linguistics*, pages 289–319. Michigan Slavic Publications.

T. Kendall, J. Bresnan, and G. van Herk. 2011. The dative alternation in African American English: Researching syntactic variation and change across sociolinguistic datasets. *Corpus Linguistics and Linguistic Theory*, 7(2):229–244.

J. Kizach. 2012. Evidence for weight effects in Russian. *Russian linguistics*, 36(3):251–270.

R. Mykhaylyk, Y. Rodina, and M. Anderssen. 2013. Ditransitive constructions in Russian and Ukrainian: Effect of givenness on word order. *Lingua*, 137:271–289.

V. I. Podlesskaya. 2006. Auxiliation of 'give' verbs in Russian. In *Voice and Grammatical Relations. In Honor of Masayoshi Shibatani*, pages 271–298. John Benjamins.

P. Qi, Y. Zhang, Y. Zhang, J. Bolton, and C. D. Manning. 2020. Stanza: A Python natural language processing toolkit for many human languages. In *Proceedings of the 58th annual meeting of the association for computational linguistics: system demonstrations*, pages 101–108. Association for Computational Linguistics.

M. Röthlisberger, J. Grafmiller, and B. Szmrecsanyi. 2017. Cognitive indigenization effects in the English dative alternation. *Cognitive Linguistics*, 28(4):673–710.

N. Shvedova. 1980. *Russkaja Grammatika*. Izdatel'stvo Nauka.

S. Sokolova, O. Lyashevskaya, and L. A. Janda. 2012. The locative alternation and the Russian 'empty' prefixes: a case study of the verb gruzit' 'load'. In *Frequency effects in language representation*, pages 51–85. Michigan Slavic Publications.

D. Zhang and J Xu. 2023. Cognitive indigenization effects in the English dative alternation. *International Journal of Corpus Linguistics*, 28(4):559–585.

A. F. Zuur, E. N. Ieno, N. J. Walker, A. A. Saveliev, and G. M. Smith. 2009. *Mixed effects models and extensions in ecology with R*. Springer.