# When a cross-linguistic tendency marries incomplete acquisition: Preposition drop in Russian spoken in Daghestan

Anastasia Panova, Tatiana Philippova[1]

HSE University

[version accepted for publication in *International Journal of Bilingualism*][2]

## Abstract

**Aims and Objectives/Purpose/Research Questions.** The purpose of the study is to figure out what factors condition the phenomenon of preposition drop in locative, directional and temporal phrases. Specifically, we investigate what kind of phrases allow preposition drop in Russian spoken in Highland Daghestan and aim at understanding the rationale for this phenomenon.

**Design/Methodology/Approach.** We conduct a quantitative analysis of data extracted from the Corpus of Russian spoken in Daghestan, which includes interviews with 53 native speakers of 15 Daghestanian and Turkic languages, amounting to 228 thousand tokens.

**Data and Analysis.** Data from forty-seven (29 male; 18 female) L2 Russian speakers who produced a sufficient number of prepositional phrases (PPs) were included in the analysis. 50 PPs were collected from each speaker resulting in a dataset of 2350 PPs. Each PP was annotated for preposition drop and several sociolinguistic and linguistic parameters. We fitted a logistic mixed-effects regression model to determine which parameters are significant predictors for preposition drop.

**Findings/Conclusions.** We show that the probability of preposition drop depends on preposition type, phonetic context and the speaker's fluency in Russian. We propose that the prominence of preposition drop in the speech of Daghestanian highlanders results from an interplay of two factors: a typological tendency for certain spatial and temporal locations to be formally unmarked and incomplete acquisition of the Russian prepositional system.

**Originality.** This is the first detailed quantitative study of preposition drop, based on an inferential statistical analysis of data from a large number of L2 Russian speakers from Daghestan.

**Significance/Implications.** The results show that the apparently contact-induced phenomena like preposition drop may be explained both by typological tendencies and incomplete acquisition of L2. This paper is thus important both for the typological study of this phenomenon and for L2 acquisition research.

[1] Corresponding author: philippo@post.bgu.ac.il.

[2] The final version is published in the OnlineFirst section of the *International Journal of Bilingualism* website; https://doi.org/10.1177/1367006921990442

**Keywords**: prepositions; preposition drop; language contact; second language acquisition; typology; Russian

## 1. Introduction

Preposition drop (P-drop) is a cross-linguistic phenomenon. It is attested in standard languages, e.g. Modern Greek (Terzi, 2010), but has predominantly been discussed with respect to non-standard varieties. These include dialects (Bailey, 2018 and Myler, 2013 for British English dialects (1); Cattaneo, 2009 for Northern Italian dialects (2) a. o.) and contact-influenced varieties, for example, multiethnolects like the Berlin "Kiezdeutsch" variety of German (Wiese, 2009 a.o.), cf. (3).

(1) Northwest British English (Myler, 2013, p. 189)
*John came* [*to*] *the pub with me*.[3]

(2) Bellinzonese Italian (Cattaneo, 2009, p. 287)

| | | | |
|---|---|---|---|
| *te* | *ve* | [*a*] | *ginasctica* |
| 2SG | go | [to] | gymnastics |

'You go to gymnastics.'

(3) Berlin Kiezdeutsch German (Wiese, 2009, p. 792)

| | | | | |
|---|---|---|---|---|
| *morgen* | *ich* | *geh* | [*zum*] | *arbeitsamt* |
| tomorrow | 1SG | go | [to.the] | job.center |

'Tomorrow, I will go to the job center.'

A special case of contact-influenced varieties are creole languages, where P-drop has also been registered as a prominent feature (Holm, 2004, p. 232), cf. (4)-(5).

(4) Haitian Creole French (DeGraff, 2007, p. 122)

| | | | |
|---|---|---|---|
| *timoun* | *yo* | *al* | *Mache Pòspyewo* |
| children | DEF.PL | go | Market Post-Pierrot |

'The children have gone to the Post-Pierrot Market.'

(5) Sierra Leonean Creole English (Yillah & Corcoran, 2007, p. 194)

a. *I* *de* [*na*] *tɔng*
3SG COP [LOC] town
'(S)he is in town.'

b. *I* *go* [*na*] *tɔng*
3SG go [LOC] town
'(S)he went to town.'

[3] We use square brackets here and throughout to indicate a preposition that is optional in this environment and is dropped in the particular example. If it is not clear that there is an alternation between P-drop and an overt P, no preposition is specified.

As can be seen from (1)-(5), P-drop typically occurs when a PP denotes the goal of motion but is not limited to this context. To date, an extreme case seems to be reported for non-standard varieties of Russian: in Russian spoken by native speakers of minority indigenous languages, P-drop affects such contexts as spatial (6) and temporal (7) adverbial phrases, comitative phrases (8), etc.

(6) Erzya Russian (Shagal, 2016, p. 370)[4]

| | | |
|---|---|---|
| [*v*] | *Saranske* | *živët* |
| [in] | Saransk.LOC | live.PRS.3SG |

'(He/she) lives in Saransk.'

(7) Nanai Russian (Stoynova, 2019, p. 21, ex. 76)

| | | | | |
|---|---|---|---|---|
| *my* | *s* | *Amura* | *priexali* | *sjuda* |
| 1PL.NOM | from | Amur.GEN | come.PST.PL | here |
| [*v*] | *sem'desjat* | *vtorom* | *godu* | |
| [in] | seventy | second.LOC | year.LOC | |

'We came here from the Amur region in the year of 1972.'

(8) Archi Russian (Daniel, Knyazev, & Dobrushina, 2010, p. 75)

| | | | | |
|---|---|---|---|---|
| *ja* | *vot* | *podružilas',* | | |
| 1SG.NOM | PTCL | make_friends.PST.FSG.REFL | | |
| *podružilas'* | | [*s*] | *avarcami* | |
| make_friends.PST.FSG.REFL | | [with] | Avar_people.INS | |

'I made friends with the Avars.'

Contact varieties of Russian therefore provide a rich and intriguing material for studying the phenomenon of P-drop. This paper focuses on Russian used as a lingua franca in Daghestan, a region of the Russian Federation characterized by a remarkable language density. The empirical basis of our investigation is the Corpus of Russian spoken in Daghestan (DagRus), which provides ample data coming from Daghestanian highlanders whose native language (L1) is one of the local languages and who speak Russian as a second language (L2).
The paper addresses the following issues:

1) What factors condition and constrain the phenomenon of P-drop? What kind of prepositional phrases (PPs) allow P-drop and to what extent?
2) What can be a possible rationale for P-drop? Can any of the existing analyses of P-drop capture the pattern observed in the Russian speech of Daghestanian highlanders?

[4] We adopt the Leipzig Glossing Rules throughout. For brevity, we do not gloss gender, nominative case, and singular number on nouns and their modifiers, as well as aspect on verbs in the Russian examples, unless required for the analysis. The following abbreviations are used: 1, 2, 3 – 1st, 2nd, 3rd person; ACC – accusative; ATTR – attributive; CMPR – comparative suffix; CONJ – conjunction; COP – copula; DAT – dative; DEF – definite; DIST – distal demonstrative; ESS – essive form (static location in a spatial domain); F – feminine; GEN – genitive; HST – hesitation marker; INF – infinitive; INS – instrumental; IPFV – imperfective; LOC – locative; M – masculine; N – neuter; NEG – negation; NOM – nominative; OBL– oblique (nominal stem suffix); PFV – perfective; PL – plural; PRS – present; PST – past; PTCL – particle; PTCP – participle; REFL – reflexive; SBJV – subjunctive; SG – singular number/agreement; SUPER – spatial domain on the horizontal surface of the landmark.

The remainder of the paper is organized as follows. Section 2 provides an overview of the existing analyses of P-drop. Section 3 describes the data, methods of data collection and annotation. Section 4 presents a statistical analysis of the dataset that reveals predictors significant for P-drop. Section 5 interprets the results of the statistical analysis and discusses reasons underlying the observed P-drop pattern. Section 6 concludes and discusses P-drop from a cross-linguistic perspective.

## 2. Background

This section reviews previous treatments of the phenomenon of P-drop, grouped by the type of explanation they propose. This guides our annotation of PPs in Section 3. We relate our own analysis to the ones discussed here in Section 5.2.

### 2.1. Phonetic reasons

P-drop in contact varieties of Russian has been reported to mostly affect simple prepositions, especially *v* 'in(to)', which is sometimes accounted for by phonetic properties or phonotactic constraints of the speakers' native languages.

For instance, the Nanai, Ulch and Enets languages (spoken in the Russian Far East and Northern Siberia) are less tolerant to word-initial consonant clusters than Russian, and a quantitative study of Russian spoken by the Nanai, Ulch and Enets people (Khomchenkova, Pleshak, & Stoynova, 2017; Stoynova, 2019) has indeed shown that P-drop may be driven by cluster avoidance. Specifically, in these varieties the drop of *v* 'in(to)' is less likely before V-initial stems than before C-initial ones, with palatalized C-initial stems being an intermediate case. This proposal is supported by two additional facts: a) cluster simplifications are also occasionally observed at the word-level (*kusno* instead of *vkusno* 'tasty'); b) P-drop depends on the extent of phonetic interference from L1 exhibited by individual speakers (Stoynova, 2019, p. 21).

Phonetic factors are also mentioned in Daniel et al. (2010) that studied the Russian speech of Archi, Avar and Lak speakers from three Daghestanian villages. They observe omission of only three prepositions, namely *v* 'in(to)', *s* 'with; off, from' and *na* 'on(to)', noting that *ot* 'from', *u* 'at', *iz* 'from, of', *dlja* 'for' are never dropped and that the preposition *v* 'in(to)' is omitted most frequently. The latter finding is hypothesized to be partially a phonetic artefact of their annotation: *v* 'in(to)' is "sometimes realized as bilabial rather than labiodental and is hardly audible, so that, in many cases, it is not easy to decide whether it is a dropped [*v*] or its weak to zero realization." (Daniel et al., 2010, p. 75).

### 2.2. Morphosyntactic interference with other languages

P-drop has also been accounted for by appealing to morphosyntactic interference on the part of the minority language, specifically, in papers devoted to contact-influenced varieties of Russian.

Daniel & Dobrushina (2009, 2013), Daniel et al. (2010) propose that this process is conditioned by the morphosyntax of Daghestanian languages, characterized by rich nominal declension paradigms and the employment of postpositions rather than prepositions. Thus, the phrases corresponding to the Russian preposition + case-marked noun complex have the form of a case-marked noun or, less frequently, a case-marked noun + postposition in these languages, as illustrated in (9)-(10) for Mehweb Dargwa.

(9) a. Mehweb Dargwa (adapted from Chechuro, 2019, p. 64)
*ustuj-če-b*
table.OBL-SUPER-N(ESS)
'on the table'

b. Russian
*na* *stole*
on table.LOC
'on the table'

(10) a. Mehweb Dargwa (adapted from Lander, 2019, p. 321)
*hečʼ* *dubur-li-če* *aqu-r*
DIST mountain-OBL-SUPER up-NPL(ESS)
'over that mountain'

b. Russian
*nad* *toj* *goroj*
over DIST.INS mountain.INS
'over that mountain'

In addition, certain spatial locations (especially introduced by place names) may appear in an unmarked essive form in Daghestanian languages (see e.g. Daniel & Ganenkov, 2009). Based on the pattern of P-drop exhibited by Archi, Avar and Lak speakers, the authors conclude that "the preposition seems to drop only or primarily if the first language expresses the main meaning of the Russian preposition by morphological means" (Daniel et al., 2010, p. 77).

A range of papers, reporting P-drop in regional varieties of Russian, attribute it to morphosyntactic interference with the local Finno-Ugric and Turkic languages: Chuvash in Bajda (2018); Komi-Permyak and Tatar in Boronnikova (2014); Karelian, Vepsian, Mordvin, Tatar and Chuvash in Myznikova (2014); Erzya Mordvin in Shagal (2016). All these minority languages are head-final and express spatial and temporal meanings predominantly by case suffixes, which is what prompts their speakers to drop prepositions when speaking Russian, according to these authors.

Khomchenkova et al. (2017) and Stoynova (2019), investigating contact-influenced varieties of Russian spoken in the Russian Far East and Northern Siberia, also consider the morphosyntactic interference factor. While they propose that the drop of the preposition *na* 'on(to)' can only be accounted for by morphosyntactic influence of L1, most instances of P-drop involve monoconsonantal prepositions and are best accounted for by phonetic factors (see Section 2.1).

## 2.3. Markedness principle

One of the features of P-drop often cited in the literature is its tendency to occur in semantically unmarked contexts.

On the one hand, P-drop is discussed as being related to the semantics of the prepositions 'to' and/or 'at' which are cross-linguistically dropped most frequently. Gehrke & Lekakou (2013) and Bailey (2018), both referring to Zwarts (2008, 2010), notice that exactly these prepositions have the most neutral and the most basic spatial semantics. A similar point is found in Biggs (2015, p. 223): she quotes Caponigro & Pearl (2008) who assume that the prepositions *to* and *at* can be omitted while the preposition *from* cannot because of the difference in their degree of markedness. This is reminiscent of the so-called Goal/Source asymmetry — a cognitive bias that has been proposed to underlie various linguistic phenomena. For instance, goal path phrases are linguistically encoded more often than source path phrases (Lakusta & Landau, 2005; Lakusta et al., 2007; Georgakopoulos, 2018). More importantly, an adposition-like element marking the goal may be omitted while the one marking source typically may not (see Ihara & Fujita, 2000 for Japanese). On the other hand, Gehrke & Lekakou (2013) argue that restrictions on P-drop are explained by the semantics of the complement noun. According to their hypothesis, P-drop may only occur when the corresponding noun denotes a stereotypical location, e.g. *house, university*. Cattaneo (2009, p. 288), in turn, assumes that the possibility of P-drop is conditioned by the "familiarity" of particular locations to the speaker.

Similar argumentation is found in Comrie's (1986) discussion of locative constructions in Eastern Armenian. His key theoretical suggestion is that the formal markedness of a construction correlates with "the degree of markedness of the locational situation in the world being described" (Comrie, 1986, p. 87).[5] Specifically, a citation form of a noun can only be used in simple statements like "something is located somewhere". If the semantics of the predicate is different from "be located at", a case-marker is employed, sometimes accompanied by a postposition. Finally, case and postposition are both needed when one wants to specify the relation between the locatum and location (in, on, under, etc.).

Variation in the extent of formal marking in locative/directional constructions has also been described by Haspelmath (2019) in the context of "differential place marking", although he prefers to explain this variation in terms of frequency, expectedness and efficient coding (Haspelmath, 2019, p. 328).

## 2.4. Exceptional syntactic structure

P-drop in Indo-European languages has received extensive attention in the generative literature. All works we are aware of propose to account for the phenomenon by positing exceptional syntactic structure. As Bailey (2018, p. 56) points out, two major approaches have been taken to analyze P-drop constructions: assuming no PP projection or positing a null-headed PP. The former approach does not involve a PP at any level of syntactic representation and posits

[5] This is similar to the general idea of Aristar (1997) that "the most typical instance of a category is the most likely to be unmarked" (Aristar, 1997, p. 332).

pseudo-incorporation of the bare noun denoting location into a verb (Gehrke & Lekakou, 2013 for Modern Greek; Hall, 2018 for Multicultural London English). Analyses that do assume a PP projection in the contexts of P-drop differ with respect to how the null preposition is "licensed": it either incorporates into the verb (Ioannidou & Den Dikken, 2009 for Modern Greek; Bailey, 2018 and Myler, 2013 for British English dialects) or the noun denoting location undergoes movement into the PP (Longobardi, 2001 for Veneto dialects; Collins, 2007 for the English noun *home*; Cattaneo, 2009 for Bellinzonese Italian; Terzi, 2010 for Modern Greek). Biggs (2015) stands out, as she proposes the phonetically null element alternating with *at/to* in Liverpool English to belong to the category κ, rather than P. This κ is a semantically abstract head, having only a basic allative or stative meaning, depending on the context, whose function is to license inherent case on the noun phrase (NP).

While some of these formal analyses (Bailey, 2018; Biggs, 2015; Cattaneo, 2009; Gehrke & Lekakou, 2013) attempt to capture the intuition that P-drop happens in unmarked contexts (cf. Section 2.3), others propose a special syntactic structure as the sole explanation.

## 3. Data

We investigate the Russian speech of consultants from highland villages of Daghestan where Russian is used as a lingua franca. Daghestan is a republic within the Russian Federation located in the Northern Caucasus. Around fifty languages are spoken in the relatively small area of mountainous Daghestan (about 50,000 $km^2$). Most of them belong to the Nakh-Daghestanian (East Caucasian) language family, but speakers of three Turkic languages (Azerbaijani, Kumyk, Nogai) and one Iranian language (Tat) also live there.

Russian began to be taught at local schools that were established by the Soviet government in the 1930s. First, it was taught by the locals who had a decent command of Russian and then, in the 1950s, by Russian teachers who were sent by the government to teach in these villages (see Dobrushina, 2013, p. 382). Since then, most Daghestanians have acquired Russian through schooling. They are also exposed to Russian when watching TV and travelling to towns. Currently, Russian is widely and rather fluently spoken as L2 by Daghestanian highlanders (for details on the status of Russian in Daghestan see Dobrushina & Kultepina, 2020).

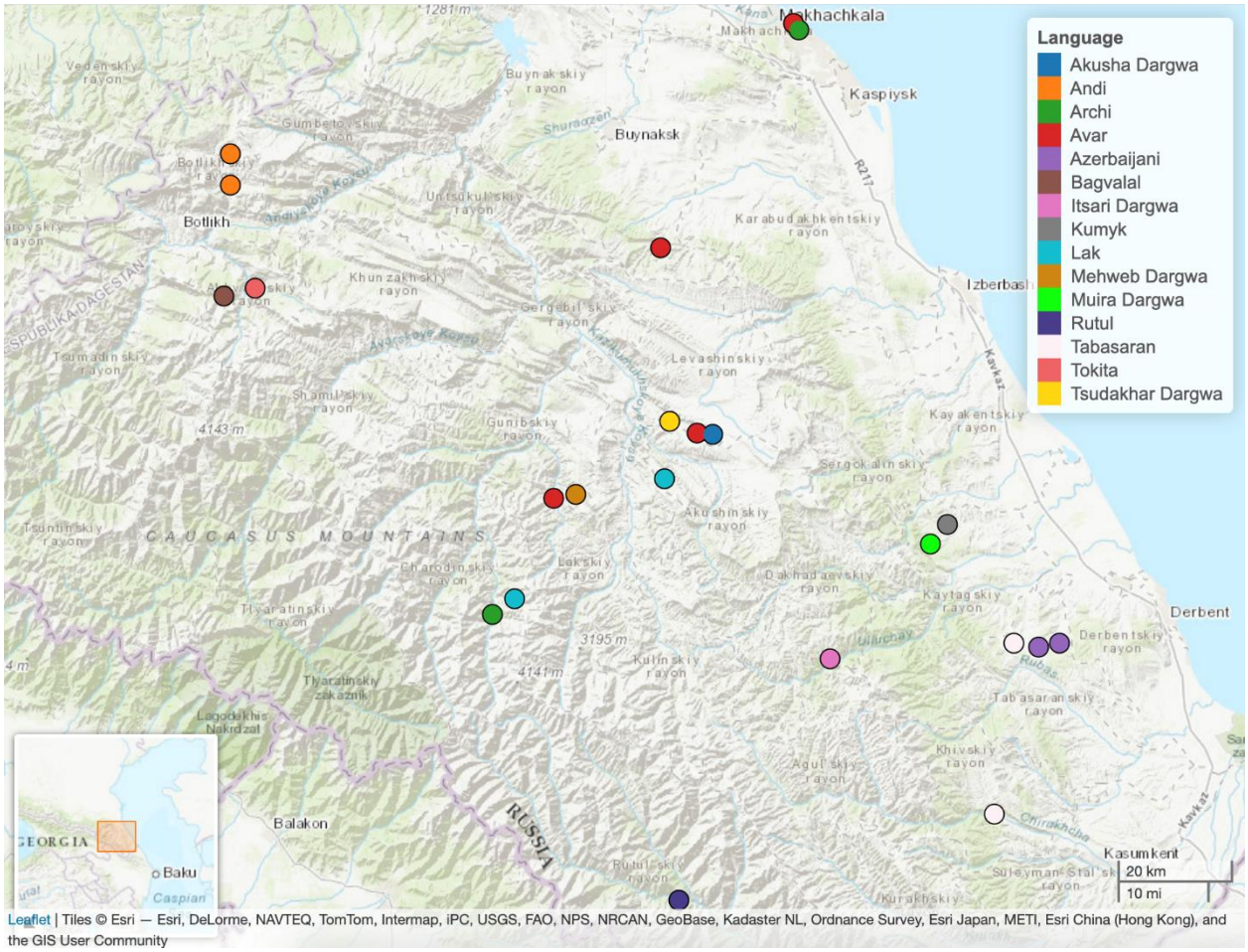
The interaction between ethnic Daghestanian languages and Russian has been studied in a number of papers, including Daniel & Dobrushina (2009, 2013), Daniel et al. (2010). These studies present an overview of specific linguistic features that are observed in the speech of Daghestanian highlanders. P-drop — the subject of this paper — is also discussed there in qualitative terms (cf. Section 2.2). We now turn to a detailed quantitative study of this phenomenon across a large number of speakers of different L1s.

### 3.1. Sampling

For our research we use data from the Corpus of Russian spoken in Daghestan (DagRus). The current version of the corpus comprises 50 sociolinguistic interviews with 55 consultants who are L1 speakers of 15 Daghestanian and Turkic languages. 46 interviews were recorded in 25

villages and 4 in the city of Makhachkala, the capital of the Republic, cf. Figure 1.[6] The total number of tokens produced by consultants is about 228 thousand.

Figure 1. Speakers' native languages and places where the interviews were recorded.[7]




In order to study P-drop we collected a dataset of PPs registered in the speech of all interviewees. Since the interviews are of varying length (10 to 95 min), we collected 50 PPs from the middle of each interview; six speakers who produced less than 50 PPs were excluded from the sample. The decision to avoid the beginning and the end of the interviews was guided by our strive to capture the most natural speech: bearing in mind the Observer's Paradox of Labov (1972), we expected that at the beginning the speakers would try their best to accommodate to the interviewers who speak Standard Russian; we also supposed that they would be too tired towards the end, possibly producing more non-standard features than they would when talking to their peers. On average, it took a speaker around 10 minutes to produce 50 PPs. Thus, in our sample, 50 consecutive PPs collected from a 60 min interview roughly came from the 25:00-35:00 fragment. In most cases, we went back to the original recordings

[6] Two speakers from Makhachkala who identify themselves as Laks speak Russian since their early childhood and their speech displays no deviations from L1 Russian (cf. Section 3.2), therefore we classify them as native speakers of Russian and do not include in our sample.
[7] The map was created with the "lingtypology" package for R (Moroz, 2017).

to double check whether a preposition was preserved or dropped. Whenever our perception clearly diverged from the corpus transcription, we went with the former, noting this discrepancy in our dataset.

### 3.2. Annotation

A bare NP was analyzed as involving P-drop whenever semantically warranted, based on comparison with L1 Russian. For example, the phrase *postupit' kursy* 'to enroll in a course' in DagRus was analyzed as involving omission of the preposition *na* 'on(to)' that one expects in L1 Russian (*postupit' na kursy* 'to enroll in a course').

Each occurrence of a PP (with or without P-drop) was annotated with a number of parameters. Sociolinguistic parameters included the speaker's ID, sex, year of birth, native language and education level. Apart from that, observations from previous studies of P-drop (reviewed in Section 2) led us to annotate the PPs with the following linguistic parameters:

(i) prepositional head;
(ii) initial phoneme of the prepositional complement (consonant/vowel);
(iii) complement type (toponym, temporal location, institution, other).
(iv) semantic type (Goal, Source, Location for spatial PPs)

Finally, we evaluated each speaker's fluency in Russian, based on two metrics, calculated for the examined interview fragment: their speech rate (average number of words per minute) and closeness to the L1 benchmark. The latter metric is the average number of deviations from L1 Russian per 100 words. We considered deviations at the morphological, syntactic and lexical level, excluding P-drop (see the Appendix for the full list of types of deviations that we counted, along with the examples and the variant expected in L1 Russian). To arrive at a more accurate and less subjective measure, each fragment was examined independently by both of the authors who are L1 speakers of Russian. Then, the initial annotator of the text fragment compiled a final, ''consensual'' list of deviations, including those that had been overlooked during the first round. Having obtained the data on speech rate and deviation ratio, we converted them into coefficients, with the 0.1 value attributed to the speaker exhibiting the lowest speech rate and the highest deviation ratio respectively. The coefficient values for the rest of the speakers were calculated using the following formula from Osborne (2013, p. 142): 1 − (native speaker average − individual score) / (native speaker average − lowest score) / 0.9).[8] The resulting fluency index is the average of the two coefficients.

## 4. Analysis

In this section we conduct a descriptive and inferential statistical analysis of the entire dataset to find out which linguistic and sociolinguistic parameters have an effect on P-drop. At the end

[8] The average native speaker values in our case came from the two L1 Russian speakers from Makhachkala that we excluded from the analyzed sample. The average speech rate for them was 152 words per minute and the average number of deviations per 100 words was equal to 0.

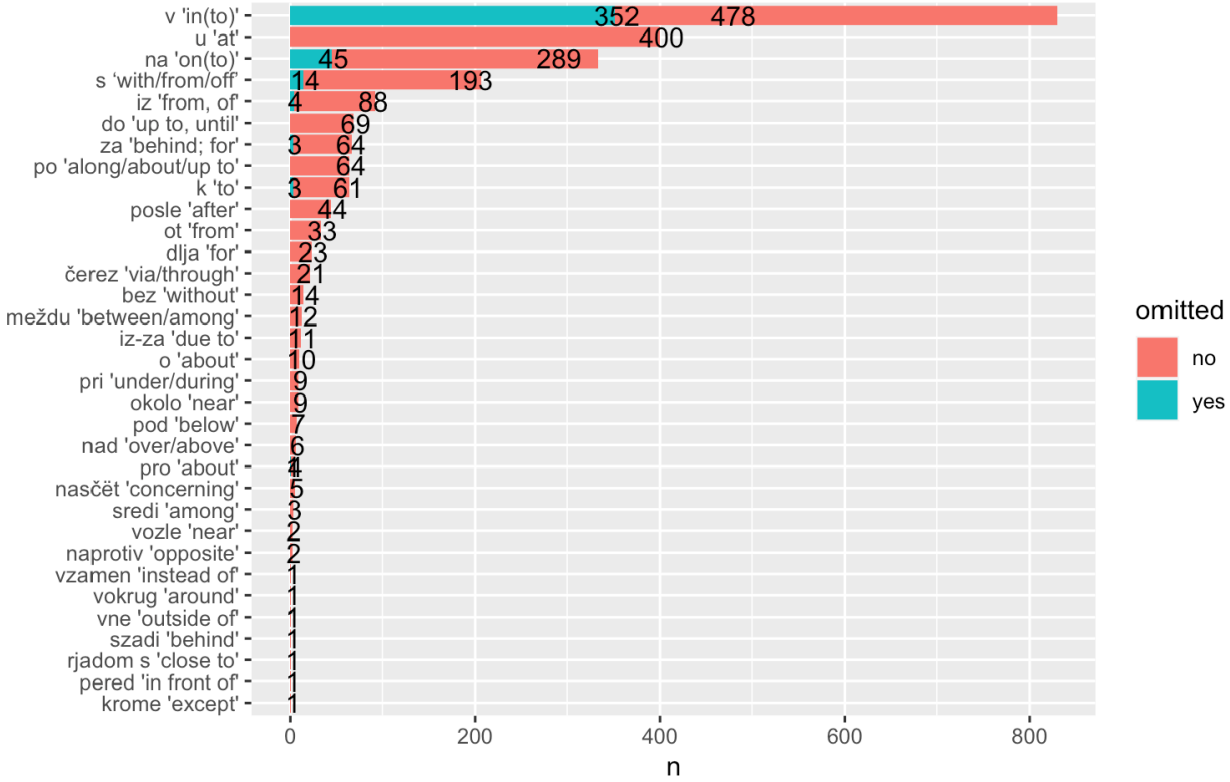
of this section we discuss the context type parameter for the relevant fraction of the data and show how it helps to reveal distinct P-drop patterns among speakers.

## 4.1. Descriptive statistics

The collected dataset consists of 2350 PPs (50 from each of the 47 speakers), 421 of which involve P-drop. 29 speakers in our sample are male, 18 speakers are female.

Let us look at how P-drop depends on the annotated linguistic parameters. As expected from previous research, different prepositions do not have equal propensity to be dropped. The bar plot in Fig. 2 shows that P-drop only ever occurs with 7 prepositions, namely *v* 'in(to)', *na* 'on(to)', *s* 'with/from/off', *iz* 'from', *za* 'for, behind', *k* 'to', *pro* 'about'.[9]

Figure 2. Number of omissions and productions per preposition.




The preposition *v* 'in(to)' exhibits an especially robust pattern: it is dropped in 42% of the PPs it heads; *na* 'on(to)' comes next with 13.5% of omissions. Minimal pairs illustrating the alternation between preposition omission and preposition retention in the speech of the same consultant are given for *v* 'in(to)' in (11) and for *na* 'on(to)' in (12).

[9] Preposition *pro* 'about', which is omitted in 1 out of 4 cases, is actually dropped in a context where L1 Russian marginally permits it as well, namely, in a fragment reprise question.

(11) *da, on* **[v]** ***Čuveke*** *rodil,*
yes, 3MSG.NOM [in] Chuvek.LOC give_birth.PST.MSG
*rodilsja* *on* ***v*** ***Čuveke***
be_born.PST.MSG.REFL 3MSG.NOM in Chuvek.LOC
'Yes, he was born [lit. gave birth] in Chuvek, he was born in Chuvek.'
[arhit.хив.42][10]

(12) a. *u nego vse zapisi*
at 3MSG.GEN all note.PL
**[*na*] *latinskom jazyke*** *byli*
[on] latin.ATTR.LOC language.LOC be.PST.PL
'he had all of his notes taken in latin script [lit. language]'
[yangikent.маллакент.40]
b. *vse zapisi vël* ***na*** ***latinskom jazyke***
all.ACC note.PL.ACC lead.PST.MSG on latin.ATTR.LOC language.LOC
'[he] took all notes in latin script [lit. language]' [yangikent.маллакент.40]

The data in Fig. 2 suggests that the P-drop pattern displays the Goal/Source asymmetry discussed in Section 2.3, since the prepositions *s* 'with/from/off', *iz* 'from' that may mark Source are dropped much more rarely than the typical Goal-marking prepositions *v* 'in(to)' and *na* 'on(to)'. This idea is supported by the data in Table 1, which shows that PPs encoding a Goal path display P-drop much more prominently than PPs encoding a Source path. Note, however, that PPs encoding Location allow P-drop roughly to the same extent as Goal-encoding PPs. This means that we cannot provide a uniform account for the observed P-drop pattern appealing to the Goal/Source asymmetry alone.

Table 1. Number of P-omissions and semantic type of PP.

| | | **omitted Ps** | **retained Ps** | **% omitted** |
|---|---|---|---|---|
| **Semantic type** | Goal | 81 | 165 | 32.9% |
| | Source | 7 | 127 | 5.2% |
| | Location | 161 | 275 | 36.9% |
| | Other | 172 | 1362 | 11.2% |

Since the prepositions *v* 'in(to)' and *na* 'on(to)' are the only ones that are systematically omitted and also constitute the bulk of goal and locative PPs (81.6% and 93.8% respectively), it is probably their inherent properties that are responsible for P-drop. We suggest that these prepositions are especially prone to omission because they are the ones used in most general locative and directional phrases, not necessarily specifying the relation between the locatum and the location.[11] According to Comrie (1986), these are precisely the environments that tend

[10] The information in the square brackets is the speaker's ID in the DagRus corpus.
[11] Incidentally, *v* 'in(to)' and *na* 'on(to)' are the two most frequent Russian prepositions (according to Lyashevskaya & Sharoff, 2009). This fact might well be a reflection of their multifunctionality and abstract nature.

to be least marked (cf. Section 2.3). Therefore, the prepositions *v* ‘in(to)’ and *na* ‘on(to)’ are grouped together in our further statistical analysis.

In Table 2 we can see that the frequency of P-drop does not seem to depend on the phonetic environment: prepositions are omitted more or less equally frequently before vowel-initial and consonant-initial complements.

Table 2. Number of P-omissions and initial phoneme of the P-complement.

| | | **omitted Ps** | **retained Ps** | **% omitted** |
|---|---|---|---|---|
| **Initial phoneme** | Vowel | 67 | 328 | 17.0% |
| | Consonant | 354 | 1601 | 18.1% |

Let us now examine the relation between P-drop and three sociolinguistic parameters: sex, native language family and education level, visualized in Fig. 3. Since some of the 15 native languages were represented by only one speaker, we merged the languages according to their genealogy. As a result, in Fig. 3 we have the (Nakh-)Daghestanian family (Andi, Archi, Avar, Bagvalal, Akusha Dargwa, Itsari Dargwa, Mehweb Dargwa, Muira Dargwa, Tsudakhar Dargwa, Lak, Rutul, Tabasaran, Tokita) and the Turkic family (Kumyk, Azerbaijani). A similar procedure was applied to the education level parameter. While the DagRus corpus distinguishes five levels of education (incomplete secondary, secondary, secondary specialized, incomplete higher, higher), we unify the former four into non-higher education and contrast it with higher education, since some of these levels characterize one or two speakers only. Building on what we found about the ability of various prepositions to drop, in Figs. 3-5 we plot the rate of omissions, only considering PPs that are headed by the seven Ps that are in principle omittable: this way we partially solve the problem of an uneven distribution of omittable and non-omittable prepositions across speakers.

As can be seen from Fig. 3, the difference between men and women with respect to the frequency of P-drop is more pronounced for the speakers of Turkic languages than for the speakers of Daghestanian. We can also see that the speakers of Turkic drop prepositions more frequently than the speakers of Daghestanian languages.[12] Finally, speakers with higher education tend to omit prepositions more rarely than those with lower education levels.

[12] Note that our sample contains 37 speakers of (Nakh-)Daghestanian and only 10 speakers of Turkic, so we can be more certain that the mean omission rate of our (Nakh-)Daghestanian sample comes close to the population mean rate.

Figure 3. Rate of omissions and sociolinguistic parameters (sex, native language, education level).

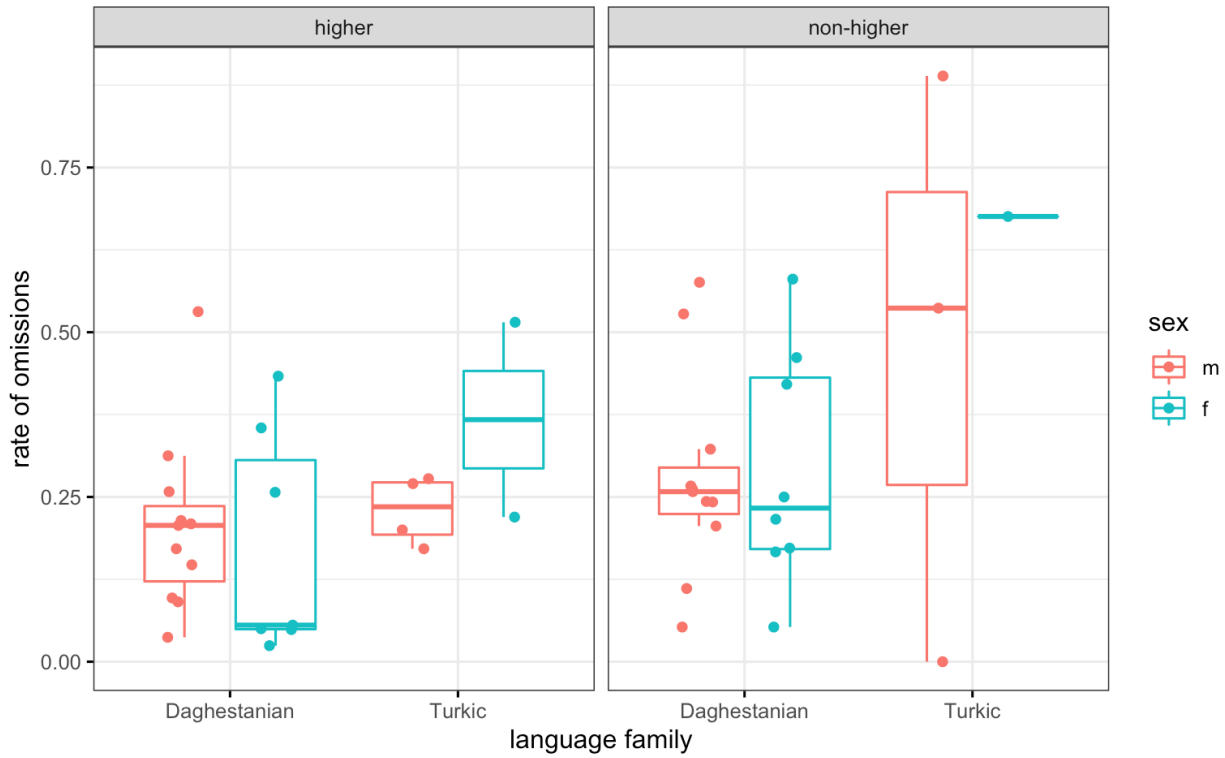



Figure 4 shows how the ratio of omissions to the number of produced omittable prepositions depends on the year a speaker was born in. Each point represents one speaker. While the distribution of the points does not allow us to make a definitive conclusion, the linear trend (with the confidence interval around it) shows that there is no significant correlation between the year of birth and the rate of omissions: the confidence interval is too wide for us to be sure that the average number of preposition omissions decreases along the horizontal axis.

Figure 4. Rate of omissions and year of birth.

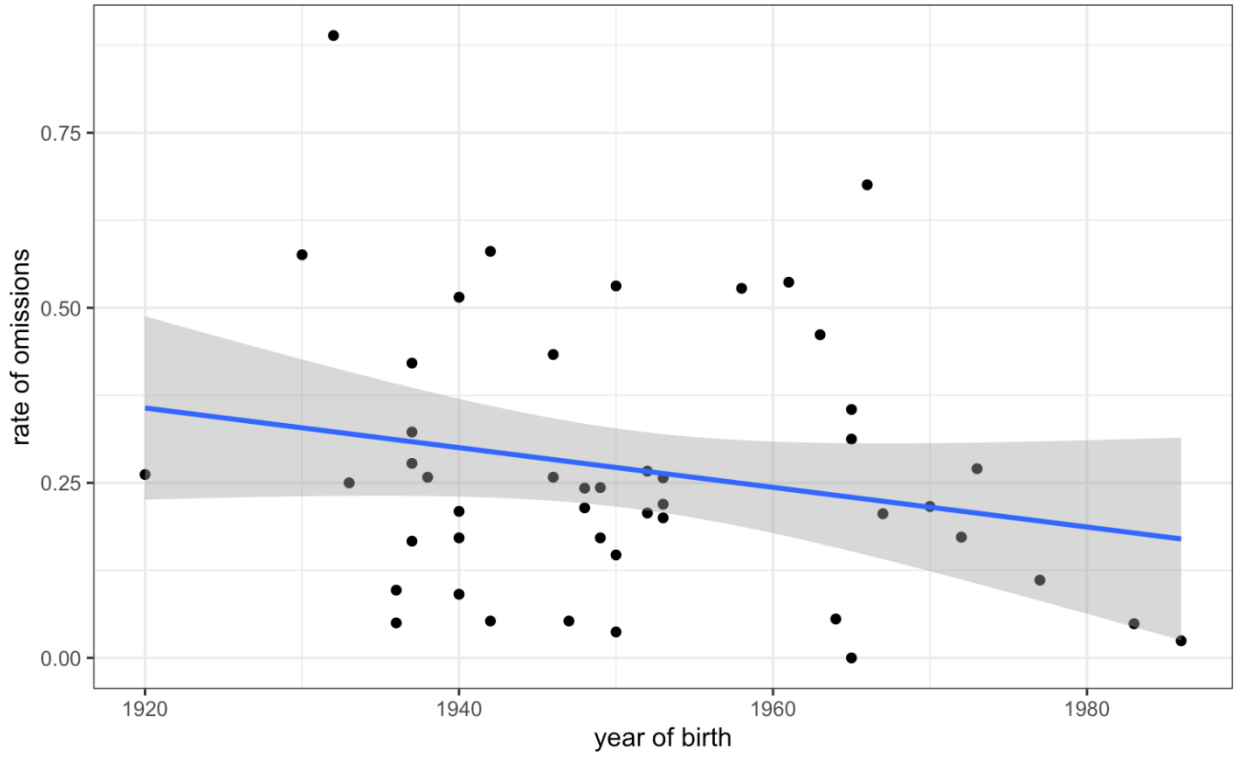
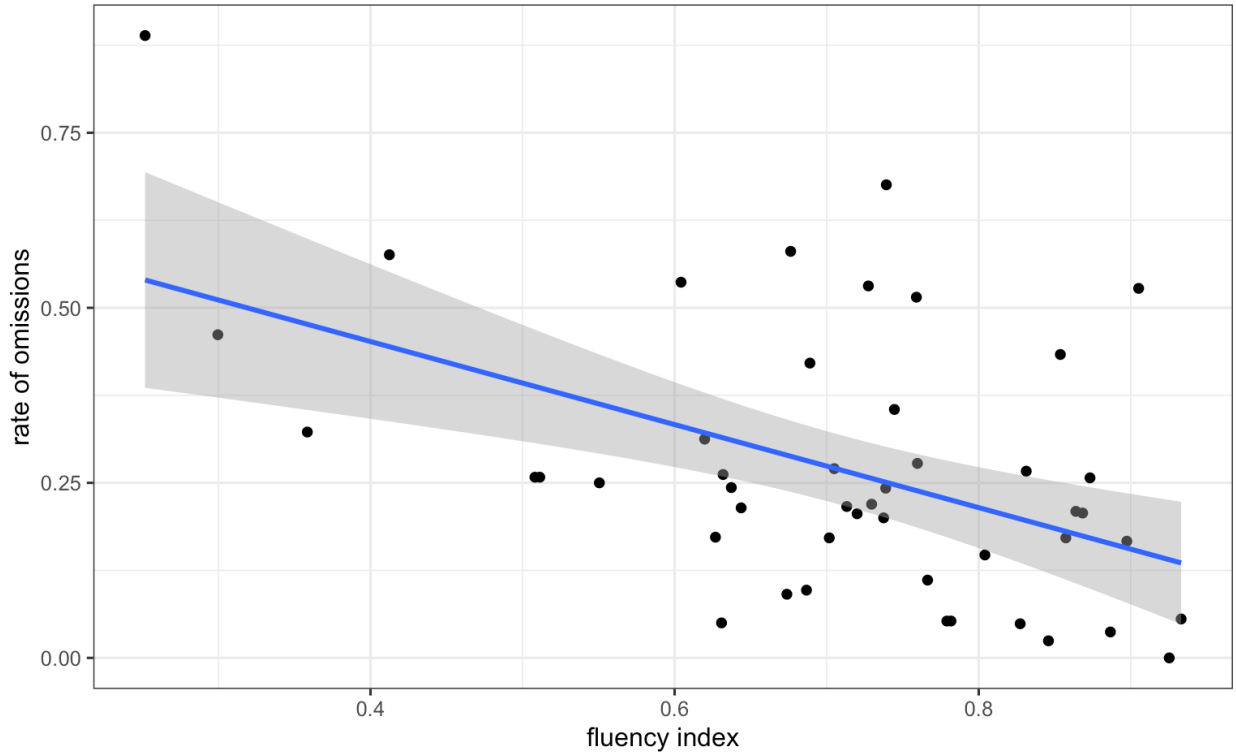



Figure 5 shows the relation between the rate of omissions and the speaker's fluency in Russian. The latter is represented by an index, combining speech rate and the ratio of deviations from L1 Russian (see Section 3.2 for details). Each point, again, corresponds to one speaker. The linear trend reveals that speakers who are more fluent in Russian tend to omit prepositions less frequently.

Figure 5. Rate of omissions and fluency in Russian.




In the following section we run a logistic regression analysis using the R software[13] to assess the significance of the factors discussed above.

4.2. Logistic regression

The linguistic and sociolinguistic factors described in Section 4.1 are summarized in Table 3.

Table 3. Variables and effect type.

| **Variable** | **Variable type (and values)** | **Effect type** |
|---|---|---|
| Preposition type | categorical (*v* 'in(to)'/*na* 'on(to)' vs. other) | Fixed effect |
| Initial phoneme | categorical (consonant/vowel) | Fixed effect |
| Speaker | categorical | Random effect |
| Year of birth | continuous | Fixed effect |
| Sex | categorical (male/female) | Fixed effect |
| Education level | categorical (higher/non-higher) | Fixed effect |
| Language family | categorical | Fixed effect |

[13] https://www.r-project.org/.

| | | |
|---|---|---|
| | (Daghestanian/Turkic) | |
| Fluency index | continuous | Fixed effect |

In order to see how significant each factor in Table 3 is, we employ a logistic mixed-effects model (Baayen, 2008, pp. 242-259; Gries, 2013, pp. 293-315; Levshina, 2015, pp. 254-266). A mixed-effects model is most fitting since it allows incorporating both fixed and random effects (Speaker in our case).[14] Aiming to arrive at an optimal model, we follow the backward stepwise variable selection procedure (Levshina, 2015, pp. 266-267), using the function drop1(). This function checks which predictor could be deleted to obtain a better fitted model (Gries, 2013, p. 266). We applied this function four times and, as a result, left out four parameters (sex, year of birth, education level and language family), based on the values of the Akaike Information Criterion (AIC). The results of the logistic regression for the three remaining parameters (fixed effects) are presented in Table 4.

Table 4. Fixed effects of the logistic regression.

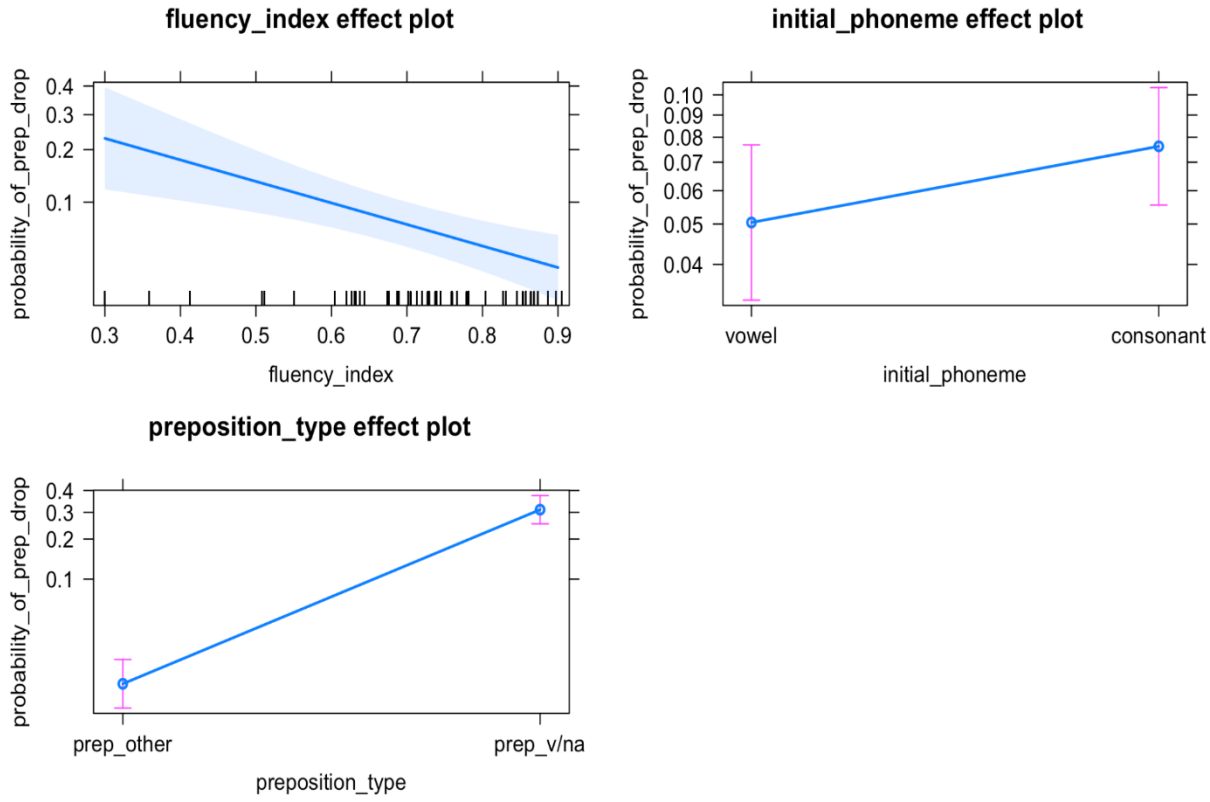
| | **Estimate** | **Std. Error** | **z value** | **Pr(>\|z\|)** | |
|---|---|---|---|---|---|
| **(Intercept)** | -1.8848 | 0.6764 | -2.787 | 0.005325 | ** |
| **Fluency index** | -3.3218 | 0.9195 | -3.613 | 0.000303 | *** |
| **Initial phoneme**<br>- vowel | -0.4406 | 0.1743 | -2.528 | 0.011459 | * |
| **Preposition type:**<br>- *v* 'in(to)'/*na* 'on(to)' | 3.5183 | 0.2232 | 15.765 | < 2e-16 | *** |

Let us discuss the columns with numerical values. The first column in Table 4 shows the estimates which specify the slopes of the regression line. Positive coefficient values in the Estimate show that the relevant predictor contributes to P-drop, while negative values mean that the predictor (or its particular value) and P-drop are negatively correlated. The second column displays the standard errors of estimated coefficients. The *p*-values in the fourth column are based on the z-statistics from the third column. They show how confident we can be in rejecting the null hypothesis that a parameter has no effect on P-drop. Asterisks mark predictors that are statistically significant (cannot be rejected as having no effect); their number reflects the degree of confidence.

We can see that the Intercept (corresponding to a situation when all continuous explanatory variables equal zero, and all categorical variables are at their reference levels) and all three predictors turn out to be significant: fluency index, initial phoneme and preposition type. Fig. 6 plots the effects of all three predictors.

[14] We used the function glmer() from the R package "lme4" (Bates et al., 2015)

Figure 6. Plots of the fixed effects of the logistic regression model.




These effects plots provide the predicted probability values of the outcome (P-drop) for given values of the predictors. These are obtained from "inserting" the value of a predictor into the model formula; the effect is calculated for one predictor at a time, while the other predictors are taken at mean values multiplied by their regression coefficients. A 95-percent pointwise confidence interval is drawn around the estimated effect of each predictor, based on standard errors computed from the covariance matrix of the fitted regression coefficients. The rug plot at the bottom of the uppermost left graph shows the location of the fluency index values. Fig. 6 shows that the most powerful predictors of P-drop are the speakers' fluency in Russian and preposition type.

### 4.3. Additional observations

During the annotation process, we noticed that the speakers seem to exhibit different patterns of P-drop, depending on the type of preposition and semantic context. In this section we classify the contexts into core and non-core and show how context type and preposition type reveal the existence of three groups of speakers in our sample.

#### 4.3.1. Contexts of P-drop

Recall that we annotated the collected PPs for the semantic type of the NP appearing with an overt or omitted preposition. In particular, we specified whether this NP denoted a toponym

(13), an exact temporal location (14) or an institution (15) — these are NPs that have been observed in the previous literature to be prone to less formal marking (see Section 2.3).

(13) *i* [*v*] *Čumljax* *rabotala* *ona* *učitelem*
and [in] Chumli.LOC work.PST.FSG 3FSG.NOM teacher.INS
'She worked as a teacher both in Chumli [and…].' [yangikent.янгикент.55]

(14) [*v*] *devjanosto* *sed'mom* *godu* *tam* *požar* *byl*
[in] ninety seventh.LOC year.LOC there fire be.PST.MSG
'In '97 there was a fire there.' [shangoda.мегеб.syn-139]

(15) *nu* *on* *mog* *by*
well 3MSG.NOM can.PST.MSG SBJV
[*v*] *institut* *postupit'* *pravil'no* *že?*
[in] institute.ACC enroll.INF right PTCL
*net,* [*v*] *texnikum* *pošël*
no [in] vocational_school.ACC go.PST.MSG
'Well, he could have gone to college, right? But no, he went to a vocational school instead.' [archib.арчиб.syn-114]

As can be seen from Table 5, the omission rate of the systematically dropped prepositions *v* 'in(to)' and *na* 'on(to)' turns out to be higher in these three contexts than elsewhere (Other). In addition, these contexts account for the majority (66%) of P-drop cases, so we refer to them as *core contexts*.

Table 5. Omission of prepositions *v* 'in(to)', *na* 'on(to)' and the semantic type of the complement.

| | | **omitted Ps** | **retained Ps** | **% omitted** |
|---|---|---|---|---|
| **Complement type** | Toponym | 133 | 131 | 50% |
| | Temporal location | 80 | 43 | 65% |
| | Institution | 49 | 85 | 37% |
| | Other | 134 | 509 | 21% |
| | Total | 396 | 768 | 34% |

4.3.2. Contexts of P-drop and inter-speaker variation

The speakers can be divided into three groups, according to contexts in which they drop prepositions, cf. Table 6.[15]

[15] One of the 47 speakers omitted no preposition and is not included into this classification.

Table 6. Groups of speakers according to their P-drop patterns.

| **group of speakers** | **number of speakers** |
|---|---|
| speakers who only omit prepositions *v* ‘in(to)’, *na* ‘on(to)’ and only in core contexts[16] | 8 |
| speakers who only omit prepositions *v* ‘in(to)’, *na* ‘on(to)’ in core and non-core contexts | 22 |
| speakers who omit prepositions *v* ‘in(to)’, *na* ‘on(to)’ in core and non-core contexts and also omit other prepositions | 16 |

Speakers of the first group only omit prepositions *v* ‘in(to)’ and *na* ‘on(to)’ and only in core contexts. The second group of speakers also omit only *v* ‘in(to)’ and *na* ‘on(to)’, but do it in other spatial, temporal and more abstract contexts as well, such as (16)-(17).

(16) *kogda* *Sovetskij* *Sojuz* *byl*
when soviet.ATTR union be.PST.MSG
**[*v*]** ***osnovnom*** *rabotali* [*v*] *sovxoze*
[in] general.LOC work.PST.PL [in] sovkhoz.LOC
‘In the times of the Soviet Union people here mainly worked in the sovkhoz’
[kina.кина.нд40]

(17) *oni* *tol’ko* [*na*] *tabasaranskom* *razgovarivali*
3PL.NOM only [on] Tabasaran.ATTR.LOC speak.PST.PL
‘they spoke only Tabasaran’ [arhit.хив.42]

The third group comprises speakers who omit *v* ‘in(to)’, *na* ‘on(to)’ and other prepositions in various contexts, cf. (18-19).

(18) [*za*] *pradedušku* *našego* *tože* *vyxodila* *ona*
[for] great-grandfather.ACC our.ACC too go_out.PST.FSG 3FSG.NOM
‘She was married to our great-grandfather as well.’ [a woman who married 12 times]
[archib.арчиб.syn-138]

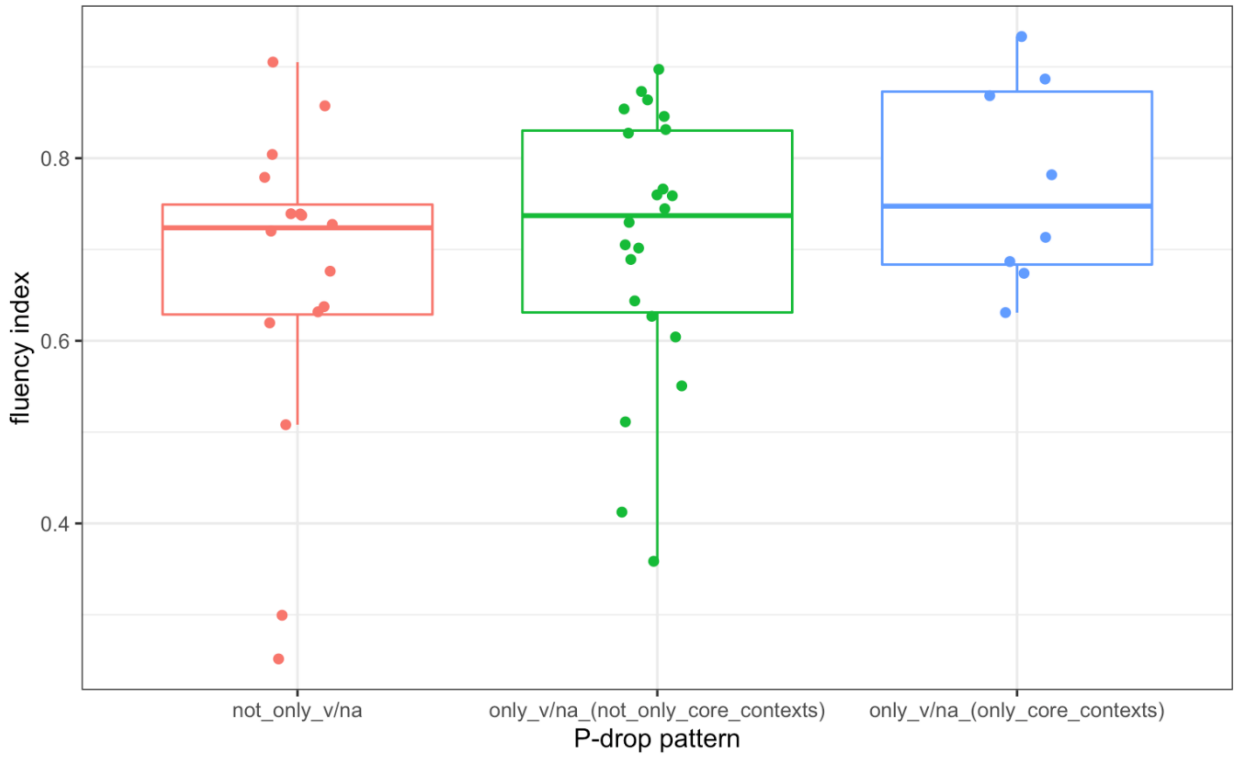
(19) *svobodno* *možno* *bylo*
Freely possible be.PST.NSG
[*s*] *rossijskim* *pasportom* *proexat’*
[with] Russian.INS passport.INS pass.INF
‘One was free to pass (the border) with a Russian passport’ [kina.кина.нд40]

A natural question to ask at this point is whether the observed P-drop patterns correlate with the speakers’ fluency in Russian. We can see from Fig. 7 that the higher the fluency index, the

[16] Importantly, all these speakers produced 6-15 PPs headed by *v* ‘in(to)’ and *na* ‘on(to)’ which do not belong to core contexts, so core contexts are indeed special for them.

narrower the range of environments with P-drop. However, the difference between the groups does not reach statistical significance ($p$ = 0.29, ANOVA test).

Figure 7. P-drop patterns and fluency in Russian.




Nevertheless, the fact that the speakers can be neatly divided into three well-defined groups suggests that there might be some psycholinguistic reality behind this classification: namely, the closer the consultant's speech is to L1 Russian, the more narrow is the scope of P-drop.

## 5. Discussion

### 5.1. Interpretation of the results

In this section we interpret the results of the statistical analysis and discuss what may underlie the significance of the three parameters — preposition type, fluency index and phonetic environment.

#### 5.1.1. Prepositions *v* 'in(to)' and *na* 'on(to)'

We have seen that the prepositions *v* 'in(to)' and *na* 'on(to)' are the ones that are systematically dropped by the speakers featured in DagRus. A possible motivation for the observed pattern is that these prepositions may have quite abstract, 'empty' semantics, in particular, in the core contexts for P-drop that involve specific time- and place-referring NPs.

In fact, in many prepositional languages a small group of nouns, such as town/city and street names may appear without a preposition when denoting a location (20)-(22). Stolz,

Lestrade, & Stolz (2014, p. 287), based on a sample of 147 languages, found that toponyms in the function of location, direction and source are zero-marked in 90% of the languages.[17] In fact, even Old Church Slavonic and Old Russian marked locative and directional phrases by case only, without prepositional 'support' (23).[18]

(20) French (Mel'čuk, 2018, p. 272)

| *on* | *s'est* | *vu* | *rue* | *de* | *Rivoli* |
|---|---|---|---|---|---|
| 3NSG | REFL.COP.PRS.3SG | see.PST.PTCP | street | of | Rivoli |

'We saw each other on Rivoli street.'

(21) Maltese (Stolz, Levkovych, & Urdze, 2017, p. 463)

| *jgħallem* | *Għawdex* |
|---|---|
| 3MSG.IPFV.teach | Gozo |

'He teaches on Gozo (an island).'

(22) Marshallese (Schlossberg, 2018, p. 139 via Haspelmath, 2019, p. 317)

| *ļe* | *e=j* | *pād* | *Lojkar* |
|---|---|---|---|
| man | 3SG=IPFV | be_located | Lojkar |

'He is at Lojkar.'

(23) Old Russian (PSRL, vol. 1, p. 27)

| *Svjatoslav''* | *bjaše* | *Perejaslavci* |
|---|---|---|
| Svjatoslav | be.3SG.PST | Perejaslavec.LOC |

'Svjatoslav was in [the town of] Pereyaslavl'.

In addition, Haspelmath (2019) distinguishes a group of what he calls topo-nouns "denot[ing] concepts which are commonly used as spatial landmarks" (Haspelmath, 2019, p. 322), cf. (24).

(24) Modern Greek (adapted from Terzi, 2010, p. 178)

| *pao/ime* | [*sto*] | *liman* |
|---|---|---|
| go.1SG/am | [to/at.DEF] | port |

'I go to/I am at the port.'

Turning now to temporal locations, it was observed in Haspelmath (1997, pp. 116-119) that a class of expressions denoting "various time periods combined with modifiers, especially demonstratives, the adjectives 'last' and 'next', and the universal determiner 'every'" systematically appear zero-marked in a number of languages (Haspelmath, 1997, pp. 116-119), cf. (25)-(26).

[17] In Stolz et al. (2014) zero-marking refers to the absence of both adposition and case marking.
[18] To the best of our knowledge, no monolingual variety of Russian displays this nowadays.

(25) English
a. *in the morning* vs. *this morning*
b. *on Friday* vs. *last Friday*

(26) Tagalog (Haspelmath, 1997, p. 117)
a. *sa* *Linggo*
at Sunday
'on Sunday'
b. *tuwing* *Linggo*
every Sunday
'every Sunday'

We observe a similar pattern in our data: for instance, in PPs meaning 'at this/that time', omission of *v* 'in(to)' is more frequent than retention: P-drop occurs in 13 out of 18 such examples.

The patterns of P-drop that we observe in DagRus thus accord well with the typological tendencies. In fact, they are even more pronounced in the varieties of Russian spoken in Daghestan than in the languages mentioned above. For instance, exact temporal locations (27) may be analyzed as modified time-denoting nouns and this is what expands the scope of contexts admitting less marking (P-drop).

(27) *i* *vot* *dobralis'*
and PTCL reach.PST.PL.REFL
[*v*] *tri* *časa* *uže* *na* *plato*
[in] three.ACC hour.GEN already on plateau.ACC
'And so, we reached the plateau already at three o'clock.'
[karata.тукита.нд14]

It might be the case that Russian spoken in Daghestan and other contact-influenced varieties exhibit P-drop in a wider range of contexts than typically discussed in the literature because prepositional complements are inflected for case in Russian. Moreover, the morphological case required in spatial PPs headed by prepositions *v* 'in(to)' and *na* 'on(to)' depends on whether the phrase encodes a goal of motion or a static location: *v dom* 'into the house(ACC)' vs. *v dome* 'in the house(LOC)'; *na kryšu* 'onto the roof(ACC)' vs. *na kryše* 'on the roof(LOC)'. Since case morphology encodes the essential semantic contrasts, in certain contexts the preposition becomes practically semantically 'empty', its contribution being very abstract. For example, in (28)-(29) the ambiguity between locational and directional meanings is resolved by case morphology, rendering the semantic contribution of the preposition rather redundant.

(28) a. *ezdili* [*v*] *Maxačkalu* *ljudi* *otsjuda*
go.PST.3PL [in] Makhachkala.ACC people from.here
'people from here used to go to Makhachkala' [chankurbe.дуранги.add.syn-1]

b. [*v*] *Maxačkale* *polučila* *pasport*
[in] Makhachkala.LOC get.PST.FSG passport.ACC
'[she] got a passport in Makhachkala' [archib.шалиб.syn-40]

(29) a. *ego* *otpravili* [*na*] *godičnye* *kursy*
3MSG.ACC send.PST.PL [on] one-year.ATTR.PL.ACC course.PL.ACC
'They sent him to take a one-year course.' [chuni.чуни.60ик]
b. *vot* [*na*] *takie* *kursax* *obučalsja*
PTCL[on] such.NOM/ACC course.PL.LOC study.PST.MSG
'So [I] took this kind of course.' [yangikent.янгикент.38]

5.1.2. Fluency in Russian

The index of fluency in Russian is another significant predictor for P-drop. That is, speakers who are more fluent in Russian tend to omit prepositions less frequently than those whose speech considerably deviates from the L1 benchmark.

Mastering prepositions is known to be a very challenging task in second language acquisition (see Celce-Murcia & Larsen-Freeman, 1983; Covitt, 1976 for English). In particular, Celce-Murcia & Larsen-Freeman (1983) observe that language learners make three types of mistakes: a) use a wrong preposition (30a); b) omit a required preposition (31a); c) use a superfluous preposition (32a). All these strategies are employed by speakers from our data sample, as evident from (30b, 31b, 32b).

(30) a. L2 English (Celce-Murcia & Larsen-Freeman, 1983, p. 261)
*My grandfather picked the name* ***on*** *me.* (instead of *for*)
b. DagRus Corpus [darvag.дюбек.нд15]
*udarila* ***na*** *spinu* (instead of ***v*** *spinu* / ***po*** *spine*)
hit.PST.FSG on back.ACC (in back / by back)
'[she] hit [me] on the back'

(31) a. L2 English (Celce-Murcia & Larsen-Freeman, 1983, p. 261)
*I served* [*in*] *the Army until 1964*.
b. DagRus Corpus [archib.арчиб.syn-138]
[*v*] *armii* *služil*
[in] army.LOC serve.PST.MSG
'[he] served in the Army'

(32) a. L2 English (Celce-Murcia & Larsen-Freeman, 1983, p. 261)
*I studied in Biology for three years.* (no preposition required)
b. DagRus Corpus [karata.тлибишо.нд32][19]
*tol'ko* *na* *odnoj* *kartoške* *pitalis'*
only on one.LOC potato.LOC nurture.PST.PL.REFL
'[we] only fed on potatoes.'

[19] In Standard Russian the verb *pitat'sja* 'feed on' requires an instrumental case-marked NP-complement, rather than a PP-complement. In fact, it is practically impossible to find instances equivalent to the English (32a),

Generalizing Celce-Murcia & Larsen-Freeman's (1983, p. 250) ideas about English, the following factors may cause difficulties in the acquisition of prepositions in L2:

a) information that is signaled by a preposition in L2 can be signaled by other means in L1: an inflection on a noun/article and/or a postposition;
b) L2 prepositions cannot be directly semantically mapped onto their functional equivalents in L1.

In light of the above we can try to explain why the predominant "mistake" made by the Daghestanian speakers of Russian in the realm of PPs is omission of the required preposition. This may have to do with the fact that the native languages of our speakers (Nakh-Daghestanian and Turkic) are head-final and, thus, postpositional. In these languages, location in space and time is encoded to the right of the nominal (predominantly by case suffixes, and sometimes also by postpositions), whereas in Russian the nominal may be marked both on the left (by prepositions) and on the right (by case suffixes). Therefore, we may expect that an individual who has not fully mastered Russian PPs will tend to omit, rather than replace or insert a preposition in case of uncertainty. This is most expected in contexts where preposition choice is idiosyncratic and/or where the preposition can be omitted in the target language. In L1 Russian this is most clearly seen in the domain of temporal expressions (33); we note that the variation illustrated in (33c, d) is apparently limited to phrases containing such modifiers as *pervyj* 'first' and *poslednij* 'last' and a restricted set of nouns in the accusative case.

(33) L1 Russian

a. *v* *pjat'* *časov*
in five.ACC hour.PL.GEN
'at five o'clock'

b. *na* *sledujuščej* *nedele*
on next.LOC week.LOC
'next week'

c. (*v*) *poslednee* *vremja*
in latest.ACC time.ACC
'recently'

d. (*v*) *pervyj* *raz*
in first.ACC occasion.ACC
'first time/on the first occasion'

This explanation can be extended to other contact-influenced varieties of Russian mentioned in Section 2, since in all those cases the minority language in contact with Russian is head-final. Of course, the most valid test for our hypothesis would be to take a head-initial, prepositional language in contact with Russian and see whether those speakers omit prepositions substantially less frequently; cf. Jarvis & Odlin (2000) who find that native

because employing a preposition when it is not needed in Russian typically leads to different case morphology as well.

speakers of Finnish (a postpositional language) omit required prepositions in their L2-English, while native speakers of Swedish (a prepositional language) do not. Possible candidates that are most similar to our case in terms of sociolinguistics would be speakers of Romani and German that have historically lived in Russia. Unfortunately, we do not have the type and amount of data that are needed for comparison with DagRus, so we leave this issue for future research.

### 5.1.3. Initial phoneme

The factor whose significance we are least certain about (based on the *p*-value in Table 4) is the initial phoneme of the P-complement. Recall that our statistical model returned the result that vowel-initial complements are less conducive to P-drop.

In principle, we could appeal to the properties of the phonological systems of Daghestanian and Turkic languages that our informants natively speak, namely, the absence of the [v] (labio-dental fricative) phoneme[20] and the ban on consonant clusters in the syllable onset (Kibrik & Kodzasov, 1990; Shiraliev & Sevortjan, 1971), as was done for other contact-influenced varieties of Russian (Khomchenkova et al., 2017; Stoynova, 2019). However, if we look at individual prepositions, we can see that it is the CV preposition *na* 'on(to)' that displays a striking contrast between V-initial and C-initial complements. It is dropped before vowels in 3,6% of cases and before consonants in 16,8% of cases. The monoconsonantal prepositions *v* 'in(to)' or *s* 'with; from, off', on the other hand, do not show a sharp contrast: *v* 'in(to)' is dropped before 39,7% of the V-initial complements and 42,9% of the C-initial complements; the respective percentages for *s* 'with; from, off' are 6.5% and 6.8%.[21] Although we do not know the reason underlying the unexpected pattern exhibited by *na* 'on(to)', it can definitely not be attributed to consonant cluster avoidance.

## 5.2. Relevance to previous research

In the Background section we reviewed four main groups of approaches to P-drop: a) phonetics/phonotactics-based account; b) accounts appealing to morphosyntactic interference with other languages; c) markedness-based analyses; d) formal syntactic accounts positing exceptional structure for P-drop constructions.

We have just discussed in Section 5.1.3. that phonetic and phonotactic reasons are rather unlikely to be definitive for P-drop exhibited by Daghestanian highlanders.

While it is not our aim here to provide a formal syntactic treatment of the phenomenon, if we were to do so, it would be clear that a pseudo-incorporation analysis would not be fitting or at least sufficient, as P-drop observed in DagRus is by no means restricted to bare, non-modified argument NPs. A formalization that appears to be most compatible with our data is the null κP proposal of Biggs (2015), since it is not predicated on PP adjacency to a verb or

[20] At the same time, all Daghestanian languages have a /w/ sonorant, realized as [β] in a subset of languages and contexts; Tabasaran and Rutul also have a /f/ phoneme. Azerbaijani and Kumyk (Turkic) have a /f/ phoneme and a bilabial /β/ phoneme.

[21] The percentages for the remaining four prepositions are not particularly informative, since there are too few instances of P-drop registered for them.

bareness of the complement NP; in addition, it can capture the fact that NPs in P-drop environments in DagRus typically bear proper case morphology.

Our explanation bears a certain affinity to the markedness and interference accounts. Specifically, our idea that the prepositions *v* 'in(to)' and *na* 'on(to)' are systematically dropped because they have a very abstract meaning in core contexts is close to the idea that these contexts are unmarked. Incomplete acquisition that we appeal to as another factor conditioning P-drop is related to morphosyntactic influence of the consultants' L1 on their Russian speech. However, this influence is more general than interference with case and postpositional systems of a particular L1: it is the absence of marking of the left-edge of the NP in the speakers' L1 that makes the acquisition of the Russian prepositional system a particularly challenging task (see Jarvis & Pavlenko, 2008, p. 94 for similar ideas concerning L2 English acquisition).

Thus, our account combines and complements insights from previous research, providing a more general, bipartite explanation.

## 6. Conclusion

In this paper we presented a quantitative corpus study of the phenomenon of P-drop in Russian spoken in highland Daghestan. Based on a dataset consisting of 2350 PPs, coming from sociolinguistic interviews with 47 speakers we found that three factors are significant predictors for P-drop. These are preposition type, fluency in Russian and phonetic context. We consider the former two factors and their synergistic effect to conceal deeper reasons underlying P-drop.

The preposition type factor is a manifestation of a cross-linguistic tendency toward less formal marking of certain spatial and temporal locations, such as toponyms and their like, and referential temporal expressions. The idea is that P-drop is particularly prominent precisely with those prepositions that are employed in the aforementioned contexts: *v* 'in(to)', *na* 'on(to)' in Russian and their correlates in other languages.

The fact that a lower fluency level corresponds to more extensive P-drop is reminiscent of one of the strategies that L2 speakers employ to cope with incomplete acquisition of the prepositional system — to avoid prepositional marking in case of uncertainty.

The aforementioned typological tendency can be more or less pronounced in languages of the world: for instance, several British English dialects exhibit P-drop in a slightly wider range of contexts than Standard English. When this tendency is coupled with incomplete acquisition, the probability of P-drop to be extended to more contexts becomes higher than at a chance level. Therefore, we find more P-drop in contact varieties, including pidgins and creoles.

We suggest that contact-influenced varieties of Russian investigated so far appear to exhibit more prominent P-drop than other languages for two reasons: a) it is in contact almost exclusively with head-final, postpositional languages; b) Russian prepositional complements are inflected for case, which encodes key semantic distinctions, such as location and direction and their metaphorical extensions. The first factor leads L2 speakers of Russian to omit rather than replace prepositions in case of uncertainty. Due to the second factor, prepositions with abstract locative semantics can be dropped without a significant loss of meaning.

To recap, the prominence of P-drop in the speech of Daghestanian highlanders results from an interplay of two factors: a cross-linguistic tendency for certain spatial and temporal locations to be formally unmarked and incomplete acquisition of the Russian prepositional system on the part of native speakers of postpositional languages.

**Acknowledgements**
We are grateful to all our colleagues who discussed previous versions of the manuscript with us, especially Peter Arkadiev, Ilya Chechuro, Michael Daniel, Nina Dobrushina, Ezequiel Koile, George Moroz, Maria Polinsky, Ilya Schurov, and Natalya Stoynova. We also thank the two anonymous reviewers and the editorial team of the International Journal of Bilingualism for providing critical feedback on the manuscript. Any remaining errors are our own.

**Funding**
The article was prepared within the framework of the HSE University Basic Research Program.

**References**

Aristar, A. R. (1997). Marking and hierarchy types and the grammaticalization of case-markers. *Studies in Language*, *21*(2), 313-368.

Baayen, R. H. (2008). *Analyzing linguistic data*. Cambridge, UK: Cambridge University Press.

Bailey, L. R. (2018). Some characteristics of Southeast English preposition dropping. *Iberia: An International Journal of Theoretical Linguistics*, *10*, 48-70.

Bajda, K. A. (2018, November 30 - December 1). *Russkij jazyk v Čuvašii: osobennosti rečevoj realizacii russkix predložno-padežnyx sočetanij v reči nostitelej čuvašskogo jazyka* [*Russian in Chuvashia: the peculiarities of the realization of the Russian preposition-case complexes in the speech of Chuvash native speakers*]. Paper presented at the Conference on Indigenous languages in contact with Russian: morphosyntactic and semantic interference, Moscow, Russia. https://drive.google.com/file/d/1Yy3pMBZEU0smO0Xi3vqoGGbaJv0ax5ow/view.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. doi: 10.18637/jss.v067.i01.

Biggs, A. (2015). A new Case for A-movement in Northwest British English. In U. Steindl, Th. Borer, H. Fang, A. Garcia Pardo, P. Guekguezian, B. Hsu, C. O'Hara & I. Chuoying Ouyang (Eds.), *Proceedings of the 32nd West Coast Conference on Formal Linguistics* (pp. 218-227). Somerville, MA: Cascadilla Proceedings Project.

Boronnikova, N. V. (2014). Ob interferencii v padežnoj sisteme (na materiale russkix spontannyx tekstov tatar i komi-permjakov) [About interference in case system (In Russian spontaneous texts of the Tatars and the Komi-Permyaks)]. *Socio- i psixolingvističeskie issledovanija*, *2*, 115-120.

Caponigro, I., & Pearl, L. (2008). Silent prepositions: Evidence from free relatives. In A. Asbury, J. Dotlačil, B. Gehrke & R. Nouwen (Eds.), *Syntax and Semantics of Spatial P* (pp. 365–385). Amsterdam: John Benjamins.

Cattaneo, A. (2009). *It is all about clitics: The case of a Northern Italian dialect like Bellinzonese* (Unpublished doctoral dissertation). New York University.

Celce-Murcia, M., & Larsen-Freeman, D. (1983). *The grammar book: An ESL*. Mexico: Harla.

Chechuro, I. (2019). Nominal morphology of Mehweb. In M. Daniel, N. Dobrushina & D. Ganenkov (Eds.), *The Mehweb language: Essays on phonology, morphology and syntax* (pp. 39–73). Berlin: Language Science Press.

Collins, C. (2007). Home sweet home. *NYU Working Papers in Linguistics*, *1*, 1–34.

Comrie, B. (1986). Markedness, grammar, people, and the world. In F. R. Eckman, E. A. Moravcsik & J. R. Wirth (Eds.), *Markedness* (pp. 85-106). Boston, MA: Springer.

Covitt, R. (1976). *Some Problematic Grammar Areas for ESL Teachers* (Unpublished Master's thesis). University of California, Los Angeles.

Daniel, M. A., & Dobrushina, N. R. (2009). Novye russkie [The new Russians]. In M. L. Remneva & S. V. Knyazev (Eds.), *Voprosy russkogo jazykoznanija, 13. Fonetika i grammatika: nastoyaščee, prošedšee, buduščee. K 50-letiju naučnoj dejatel'nosti S. K. Požarickoj* [*Issues in Russian linguistics, 13. Phonetics and grammar: present, past, future. In celebration of the 50th academic anniversary of S. K. Pozharitskaya*] (pp. 141-158). Moscow: Moscow State University.

Daniel, M. A., & Dobrushina, N. R. (2013). Russkij jazyk v Dagestane: problemy jazykovoj interferencii. [A corpus of Russian as L2: the case of Daghestan]. In V. P. Selegey (Ed.), *Komp'juternaja lingvistika i intellektual'nye texnologii: Po materialam ežegodnoj Meždunarodnoj konferencii Dialog. Vyp. 12. T. 1: Osnovnaja programma konferencii [Computational Linguistics and Intellectual Technologies: Papers from the Annual International Conference "Dialogue". Issue 12. Vol. 1: Main conference program]* (pp. 186-211). Moscow: Russian State University of the Humanities.

Daniel, M., & Ganenkov, D. (2009). Case marking in Daghestanian: Limits of elaboration. In A. Malchukov & A. Spencer (Eds.), *The Oxford Handbook of Case* (pp. 668-685). Oxford/New York: Oxford University Press.

Daniel, M., Knyazev, S., & Dobrushina, N. (2010). Highlander's Russian: Case Study in Bilingualism and Language Interference in Central Daghestan. In A. Mustajoki, E. Protassova & N. Vakhtin (Eds.), *Russian Language in the Multilingual World* (pp. 65-93). Slavica Helsingiensia, *40*. Helsinki: University of Helsinki.

DeGraff, M. (2007). Kreyòl Ayisyen, or Haitian Creole (Creole French). In P. L. Patrick & J. A. Holm (Eds.), *Comparative Creole Syntax: Parallel Outlines of 18 Creole Grammars* (pp. 101-126). London: Battlebridge.

Den Dikken, M. (2010). Directions from the GET-GO: On the syntax of manner-of-motion verbs in directional constructions. *Catalan Journal of Linguistics*, *9*, 23–53.

Den Dikken, M. (1995). *Particles: On the syntax of verb-particle, triadic, and causative constructions*. Oxford: Oxford University Press.

Dobrushina, N. (2013). How to study multilingualism of the past: Investigating traditional contact situations in Daghestan. *Journal of Sociolinguistics*, *17*(3), 376–393.

Dobrushina, N., Daniel, M., von Waldenfels, R., Maisak, T., & Panova, A. (2018). Corpus of Russian spoken in Daghestan. Moscow: Linguistic Convergence Laboratory, NRU HSE. (Available online at http://www.parasolcorpus.org/dagrus/, accessed on 07.04.2020.)

Dobrushina, N., & Kultepina, O. (2020). The rise of a lingua franca: The case of Russian in Dagestan. *International Journal of Bilingualism.* *https://doi.org/10.1177/1367006920959717*

Gehrke, B., & Lekakou, M. (2013). How to miss your preposition. *Studies in Greek Linguistics, 3*, 92–106.

Georgakopoulos, T. (2018). A frame-based approach to the source-goal asymmetry: Synchronic and diachronic evidence from Ancient Greek. *Constructions and Frames, 10*(1), 61-97.

Gries, S. T. (2013). *Statistics for linguistics with R: A practical introduction*. Berlin/Boston: Walter de Gruyter.

Hall, D. (2018, October 5-7). *PD-Drop and Pseudo-incorporation in London English*. Paper presented at the 49th Annual Meeting of the North East Linguistic Society (NELS 49), Cornell University, Ithaca, New York. http://cornell2018.nelsconference.org/files/abstracts/NELS%2049%20Abstract%20187.pdf

Haspelmath, M. (1997). *From space to time: Temporal Adverbials in the World's Languages*. Newcastle, UK: Lincom Europa.

Haspelmath, M. (2019). Differential place marking and differential object marking. *STUF - Language Typology and Universals, 72*(3), 313-334.

Holm, J. A. (2004). *Languages in contact: the partial restructuring of vernaculars*. Cambridge: Cambridge University Press.

Ihara, H., & Fujita, I. (2000). A cognitive approach to errors in case marking in Japanese agrammatism: The priority of the goal -ni over the source. In A. Foolen, & F. van der Leek (Eds.), Constructions in Cognitive Linguistics: Selected papers from the fifth International Cognitive Linguistics Conference (pp. 123–140). Amsterdam: John Benjamins.

Ioannidou, A., & Den Dikken, M. (2009). P-drop, D-drop, D-spread. In C. Halpert, J. Hartman & D. Hill (Eds.), *Proceedings of the 2007 Workshop in Greek Syntax and Semantics at MIT* (pp. 393–408). MIT Working Papers in Linguistics, *57*. Cambridge, MA: MIT Press.

Jarvis, S., & Odlin, T. (2000). Morphological type, spatial reference, and language transfer. *Studies in Second Language Acquisition, 22*(4), 535-556.

Jarvis, S., & Pavlenko, A. (2008). *Crosslinguistic Influence in Language and Cognition*. New York/London: Routledge.

Khomchenkova, I., Pleshak, P., & Stoynova, N. (2017, October 27-29). *Non-standard expression of spatial semantics in the contact influenced Russian Speech of Russian Far East and Northern Siberia*. Paper presented at the Language Contact in the Circumpolar World Conference, Moscow, Russia. http://web-corpora.net/tsakorpus_russian_nonst/pdf/IKh-PP-NS_spatial-semantics-abs.pdf

Kibrik, A. E., & Kodzasov, S. V. (1990). *Sopostavitel'noe izučenie dagestanskix jazykov. Imja. Fonetika* [A comparative study of Daghestanian languages. Nouns. Phonetics]. Moscow: Izdatel'stvo Moskovskogo Universiteta.

Labov, W. (1972). Some Principles of Linguistic Methodology. *Language in Society, 1*(1), 97-120.

Lakusta, L., & Landau, B. (2005). Starting at the end: The importance of goals in spatial language. *Cognition, 96*(1), 1-33.

Lakusta, L., Wagner, L., O'Hearn, K., & Landau, B. (2007). Conceptual foundations of spatial language: Evidence for a goal bias in infants. *Language Learning and Development, 3*, 179-197.

Lander, Y. (2019). The Mehweb "assertive" copula g$^{w}$a: a sketch of a portrait. In M. Daniel, N. Dobrushina & D. Ganenkov (Eds.), *The Mehweb language: Essays on phonology, morphology and syntax* (pp. 315–331). Berlin: Language Science Press.

Levshina, N. (2015). *How to do linguistics with R: Data exploration and statistical analysis*. Amsterdam/Philadelphia: John Benjamins.

Longobardi, G. (2001). Formal Syntax, Diachronic Minimalism, and Etymology: The History of French *Chez*. *Linguistic Inquiry*, *32*, 275–302.

Lyashevskaya, O. N., & Sharoff, S. A. (2009). *Častotnyj slovar' sovremennogo russkogo jazyka (na materialax Nacional'nogo korpusa russkogo jazyka)* [Frequency dictionary of Contemporary Russian (based on data from the Russian National Corpus)]. Moscow: Azbukovnik.

Mel'čuk, I. (2018). Les prépositions zéro en français. *Lingvisticæ Investigationes*, *41* (2), 269-284.

Moroz, G. (2017). _lingtypology: easy mapping for Linguistic Typology_ [R package]. https://CRAN.R-project.org/package=lingtypology

Myler, N. (2013). On coming the pub in the North West of England: Accusative unaccusatives, dependent case, and preposition incorporation. *The Journal of Comparative Germanic Linguistics*, *16*(2-3), 189-207.

Myznikova, Ja. V. (2014). Specifika funkcionirovanija predložno-padežnyx konstrukcij v govorax Priladož'ja i Srednego Povolž'ja [Peculiarities of preposition-case constructions in the Ladoga and Middle Volga dialects]. *Severnorusskie govory*, *13*, 149-157.

Osborne, J. (2013). Fluency, Complexity and Informativeness in Native and Non-native Speech. In G. Gilquin & S. De Cock (Eds.), *Errors and Disfluencies in Spoken Corpora* (pp. 139-161). Amsterdam/Philadelphia: John Benjamins.

PSRL — *Polnoe sobranie russkih letopisej* [A complete collection of the Russian Chronicles]. Vol. I. Saint Petersburg, 1846.

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Shagal, K. (2016). Contact-induced grammatical phenomena in the Russian of Erzya Speakers. In K. Shagal & H. Arjava (Eds.), *Mordvin languages in the field* (pp. 363-377). Uralica Helsingiensia, *10*. Helsinki: University of Helsinki.

Shiraliev, M. Sh., & Sevortjan, E. V. (Eds.). (1971). *Grammatika azerbaidžanskogo jazyka (fonetika, morfologija, sintaksis)*. [The grammar of Azerbaijani (phonetics, morphology, syntax)]. Baku: Elm.

Schlossberg, J. (2018). *Atolls, islands, and endless Suburbia: Spatial reference in Marshallese* (Unpublished doctoral dissertation). The University of Newcastle, Australia.

Stolz, T, Levkovych, N., & Urdze, A. (2017). When zero is just enough … In support of a Special Toponymic Grammar in Maltese. *Folia Linguistica, 51*(2), 453–482.

Stolz, T., Lestrade, S., & Stolz, C. (2014). *The crosslinguistics of zero-marking of spatial relations*. Berlin/Boston: De Gruyter Mouton.

Stoynova, N. (2019). Russian in contact with Southern Tungusic languages: Evidence from the Contact Russian Corpus of Northern Siberia and the Russian Far East. In A. Nikunlassi & E. Protassova (Eds.), *Russian Language in the Multilingual World* (pp. 9-36). Slavica Helsingiensia, *52*. Helsinki: University of Helsinki.

Terzi, A. (2010). On null spatial Ps and their arguments. *Catalan Journal of Linguistics, 9*, 167–187.

Wiese, H. (2009). Grammatical innovation in multiethnic urban Europe: New linguistic practices among adolescents. *Lingua, 119*, 782–806.

Yillah, S. M., & Corcoran, C. (2007). Krio (Creole English). In P. L. Patrick & J. A. Holm (Eds.), *Comparative Creole Syntax: Parallel Outlines of 18 Creole Grammars* (pp. 175-196). London: Battlebridge.

Zwarts, J. (2008). Priorities in the production of prepositions. In A. Asbury, J. Dotlačil, B. Gehrke & R. Nouwen (Eds.), *Syntax and Semantics of Spatial P* (pp. 85–102). Amsterdam: John Benjamins.

Zwarts, J. (2010). Semantic map geometry: Two approaches. *Linguistic Discovery, 8*, 377–395.

## Appendix

Types of deviations from L1 Russian counted for the purposes of assessing fluency level.

- agreement (noun-modifier; subject-predicate)

(A1) a. DagRus [arhit.хив.26]
*sovetskij vlast'* 'Soviet.**M** rule.F'

b. L1 benchmark
*sovetskaja vlast'* 'Soviet.**F** rule.F'

(A2) a. DagRus [balhar.балхар.нд22]
*togda naši proigral*
then our.**PL** lose.PST.M**SG**

b. L1 benchmark
*togda naši proigrali*
then our.**PL** lose.PST.**PL**
'Then we [the guys from our village] lost.'

- choice of the form of a pronominal expression

(A3) a. DagRus [chuni.чуни.31н]

| | | | |
|---|---|---|---|
| *doroga* | *byla,* | *ego* | *rasširili* |
| road(**F**SG) | be.PST.**F**SG | 3**M**SG.ACC | widen.PST.PL |

b. L1 benchmark

| | | | |
|---|---|---|---|
| *doroga* | *byla,* | *eë* | *rasširili* |
| road(**F**SG) | be.PST.**F**SG | 3**F**SG.ACC | widen.PST.PL |

'There was this road; it was widened.'

- government

(A4) a. DagRus [makhachkala.add-2]

| | | | |
|---|---|---|---|
| *tradicionnyj* | *islam* | *nas* | *učili* |
| traditional.**NOM/ACC** | Islam.**NOM/ACC** | 1PL.ACC | teach.PST.PL |

b. L1 benchmark

| | | | |
|---|---|---|---|
| *tradicionnomu* | *islamu* | *nas* | *učili* |
| traditional.**DAT** | Islam.**DAT** | 1PL.ACC | teach.PST.PL |

'We were taught traditional Islam.'

(A5) a. DagRus [yangikent.янгикент.41]

| | | |
|---|---|---|
| *Frukty* | *ne* | *bylo* |
| fruit.PL.**NOM** | **NEG** | be.PST.NSG |

b. L1 benchmark

| | | |
|---|---|---|
| *Fruktov* | *ne* | *bylo* |
| fruit.PL.**GEN** | **NEG** | be.PST.NSG |

'There were no fruit trees.'

- reflexive verbs without the reflexive affix or vice versa

(A6) a. DagRus [kina.кина.нд40]

| | | | | |
|---|---|---|---|---|
| *Vot* | *azerbajdžanskije* | *èto* | *otary* | *byvalis'* |
| PTCL | Azerbaijani.ATTR.PL | HST | flock.PL | be.PST.PL.**REFL** |

b. L1 benchmark

| | | | | |
|---|---|---|---|---|
| *Vot* | *azerbajdžanskije* | *èto* | *otary* | *byvali* |
| PTCL | Azerbaijani.ATTR.PL | HST | flock.PL | be.PST.PL |

'There used to be Azerbaijani flocks.'

(A7) a. DagRus [karata.тлибишо.нд22]

| | | | | |
|---|---|---|---|---|
| *Sčetovod* | *imel* | *togda* | *v* | *škole* |
| accountant | possess.PST.MSG | then | in | school.LOC |

b. L1 benchmark

| | | | | |
|---|---|---|---|---|
| *Sčetovod* | *imelsja* | *togda* | *v* | *škole* |
| accountant | possess.PST.MSG.**REFL** | then | in | school.LOC |

'There was an accountant in the school then.'

- aspect and tense form choice

(A8) a. DagRus [darvag.ерси.мд01]

| *Vot* | *èta* | *kniga,* | *ona* | *najdena* | *v* | *Drezdene* |
|---|---|---|---|---|---|---|
| PTCL | this | book | 3FSG.NOM | found.PTCP.FSG | in | Dresden.LOC |

b. L1 benchmark

| *Vot* | *èta* | *kniga,* | *ona* | *byla* |
|---|---|---|---|---|
| PTCL | this | book | 3FSG.NOM | **be.PST.FSG** |
| *najdena* | *v* | *Drezdene* | | |
| found.PTCP.FSG | in | Dresden.LOC | | |

'This book here, it was found in Dresden.'

(A9) a. DagRus [archib.арчиб.add.syn-1]

| *A* | *začem* | *mne* | *posmotret'* |
|---|---|---|---|
| CONJ | why | 1SG.DAT | found.**PFV**.INF |

b. L1 benchmark

| *A* | *začem* | *mne* | *smotret'* |
|---|---|---|---|
| CONJ | why | 1SG.DAT | found.**IPFV**.INF |

'And why should I look (at it)?'

- word order and headedness in relative clauses

(A10) a. DagRus [rikvani.зило.нд01]

| *pošël* | *k* | *kto* | *znaet* | *učiteljam* |
|---|---|---|---|---|
| go.PST.MSG | to | who.NOM | know.PRS.3SG | teacher.PL.DAT |

b. L1 benchmark

| *pošël* | *k* | *tem* | *kto* | *znaet...* |
|---|---|---|---|---|
| go.PST.MSG | to | DIST.PL.DAT | who.NOM | know.PRS.3SG |

'He went to those who know, to teachers'

(A11) a. DagRus [archib.арчиб.add.syn-2]

| *V Maxačkale* | *kotoryj* | *oni* | *postroili* | *dom* |
|---|---|---|---|---|
| in Makhachkala.LOC | which.ACC | 3PL.NOM | build.PST.PL | house |

b. L1 benchmark

| *dom* | *kotoryj* | *oni* | *postroili* | *v Maxačkale* |
|---|---|---|---|---|
| house | which.ACC | 3PL.NOM | build.PST.PL | in Makhachkala.LOC |

'The house that they built in Makhachkala'

- comparative constructions (positive rather than comparative form of the adjective/adverb; deviating word order)

(A12) a. DagRus [shangoda.мегеб.syn-153]

| *u* | *nas* | <...> | *reputacija* | *xorošaja* |
|---|---|---|---|---|
| at | 1PL.GEN | | reputation | good |
| *byla* | *čem* | *u* | *nix* | |
| be.PST.FSG | than | at | 3PL.GEN | |

b. L1 benchmark

| | | | | |
|---|---|---|---|---|
| *u* | *nas* | <...> | *reputacija* | *lučše* |
| at | 1PL.GEN | | reputation | good.**CMPR** |
| *byla* | *čem* | *u* | *nix* | |
| be.PST.FSG | than | at | 3PL.GEN | |

‘our reputation was better than theirs’

- lexical choice

(A13) a. DagRus [archib.арчиб.syn-138]

| | |
|---|---|
| *ručnyje* | *knigi* |
| **handmade**.PL | book.PL |

b. L1 benchmark

| | |
|---|---|
| *rukopisnyje* | *knigi* |
| **handwritten**.PL | book.PL |

‘handwritten books’

- other grammatical deviations

(A14) a. DagRus [archib.арчиб.add.syn-1]

| | | | | | |
|---|---|---|---|---|---|
| *A* | *papa* | *xočet* | *akkuratno* | *čtoby* | *delat’* |
| CONJ | dad | want.PRS.3SG | accurately | COMP.SBJV | do.INF |

b. L1 benchmark

| | | | | |
|---|---|---|---|---|
| *A* | *papa* | *xočet* | *akkuratno* | *delat’* |
| CONJ | dad | want.PRS.3SG | accurately | do.INF |

‘And dad wants to do it neatly.’

c. L1 benchmark

| | | | | | |
|---|---|---|---|---|---|
| *A* | *papa* | *xočet* | *akkuratno* | *čtoby* | *delali* |
| CONJ | dad | want.PRS.3SG | accurately | COMP.SBJV | do.PST.PL |

‘And dad wants (for us) to do it neatly.’