

“Nynešnie moi soparkniki, sokurortniki, sostakanniki” Productivity of the Russian Derivational Construction [so+X-nik] in Nonce Nouns

Valentina Benigni

Università degli Studi Roma Tre, Italia

Abstract This article investigates the productivity of the Russian derivational construction [so-X-nik] in the formation of nonce nouns. Adopting a constructionist perspective and a corpus-based approach, the study analyzes data from the Russian Trends corpus, a web-based monitoring corpus of journalistic texts, and the media subcorpus of the Russian National Corpus, focusing on the co-participative and deictic semantics of the pattern. The findings show that, beyond lexicalized forms (e.g. *sotrudnik* ‘collaborator’, *sovremennik* ‘contemporary’), the construction remains productive in creating low-frequency and occasional formations whose interpretation depends on contextual anchoring and shared knowledge. Particular attention is paid to plural marking in the formation of ad hoc categories and to the construction’s productive use in wordplay and expressive strategies in contemporary journalistic discourse.

Keywords Russian derivational morphology. Potential productivity. Nonce nouns. Co-participative meaning. Deictic morphology. Ad hoc categorization. Construction Grammar.

Summary 1 Introduction. – 2 Deictic Derivatives. – 3 Theoretical Background. – 4 Methodology and Data. – 5 Concluding Remarks.



Peer review

Submitted 2026-01-08
Accepted 2026-07-26
Published 2026-10-14



Open access

© 2026 Benigni | © 4.0

DOI 10.30687/BES/2785-3187/2026/01/002

1 Introduction

The Russian derivational pattern [so-X-nik] expresses co-participation through deixis, denoting an individual who is jointly involved, together with another participant (the reference point), in the event, activity, or situation encoded by the lexical base X. Canonical formations such as *so-vremen-nik* ‘contemporary’ and *so-trud-nik* ‘collaborator’ coexist with low-frequency and occasional nonce formations, illustrating the construction’s potential for creative extension. This is exemplified in (1), where a series of ad hoc derivatives (Benigni 2018) are used to denote context-bound groupings:

- (1) *Vpročem, požaluj, ved’ ne oni menja na nego segodnja vdochnovili, a nynešnie moi soparkniki, sokurortniki, sostakanniki.* (RNC,¹ N. Ju. Ustrjalov, *Dnevnik*, 1936)

‘However, it seems that it was not they who inspired me today, but rather my present co-park-goers, co-resort-dwellers, and drinking companions.’²

The nonce forms *so-park-nik-i* ‘people frequenting the same park’, *so-kurort-nik-i* ‘people staying at the same resort’, and *so-stakan-nik-i* ‘people drinking together’ instantiate the same derivational schema as entrenched lexemes, extending it beyond the domain of established vocabulary.

Traditional descriptions of Russian word formation classify [so-X-nik] as prefixal-suffixal derivation and focus primarily on established lexical items rather than on productive schematicity (Švedova 1980). From a constructionist perspective, by contrast, [so-X-nik] can be analyzed as a derivational construction – a schematic pairing of form and meaning whose productivity emerges from usage and is observable in both entrenched lexemes and context-dependent nonce formations (Goldberg 2006; Booij 2010). The availability of large electronic corpora now makes it possible to assess this productivity empirically.

Against this background, the paper aims (i) to model the deictic mechanism underlying referent identification in [so-X-nik] derivatives, (ii) to assess the construction’s productivity in contemporary Russian, and (iii) to identify recurrent tendencies in nonce noun formation.

The paper is organized as follows. Section 2 reviews the treatment of the pattern in Russian derivational morphology (§2.1), with particular attention to co-participative meaning and its deictic dimension (§2.2). Section 3 outlines the Construction Grammar framework. Section 4 presents a usage-based analysis based on the

¹ *Russian National Corpus.*

² Unless otherwise stated, all translations are by the Author.

Russian Trends (RT) corpus (§4.1) and the media subcorpus of the RNC (§4.2), with special attention to hapax legomena, contextual anchoring, and wordplay. Section 5 offers concluding remarks.

2 Deictic Derivatives

2.1 Previous Research

From a historical perspective, the model [*so-X-nik*] is often described as a calque from classical languages. For instance, the canonical form *sovremennik* ‘contemporary’ is glossed in Vasmer’s etymological dictionary (Vasmer, n.d.) as a calque of Latin *synchronus* or Ancient Greek *synchronos*.

Traditionally, formations of this type have been analyzed as a ‘prefixal-suffixal word-formation type’ (prefiksals’no-suffiksals’nyj sposob slovoobrazovanija). Šanskij, Tichonov (1987, 67) also introduced the notion of ‘circumfix’, emphasizing that the comitative prefix *so-* and the agentive suffix *-nik* are attached simultaneously to the lexical base, forming a single structural frame.

While Zemskaia (2009, 82) considers prefixal-suffixal noun formation only weakly productive in contemporary Russian, the reference grammar commonly known as *Grammatika-80* (Švedova 1980, 232, §524) nevertheless lists the circumfix *so...nik* among productive noun-forming patterns, citing items such as *sputnik* ‘companion, satellite’ and *sobutyl’nik* ‘drinking buddy’. Lexicographic sources further confirm the productivity of the model and its feminine counterpart *so...nica*, as illustrated by the gender pairs *soklassnik* (M) / *soklassnica* (F), both ‘classmate’, and *sokursnik* (M) / *sokursnica* (F), both ‘fellow student’ (Efremova 2000).

Recent developments are surveyed by Mineeva (2015), who analyzes prefixal-suffixal agentive formations from the past fifty years using neologism dictionaries, the RNC, and media sources. She identifies a range of co-participative derivatives with *so-*, distinguishing between noun-based formations (e.g. *so-klass-nik* ‘classmate’, *so-kafedr-ennik* ‘co-chair’ (academic), *so-kamern-ica* ‘female cellmate’, *so-postel’-ščik* ‘bedfellow’) and verb-based formations (e.g. *so-del-ec* ‘co-owner’, *so-duma-tel’* ‘co-thinker’, i.e. one who collaborates in thought processes).

2.2 Co-participative Function

Independently of the lexical base and the specific suffix attached to it, *so-* derivatives encode co-participation and joint involvement: an animate referent (in this case study marked by the suffix *-nik*)

is defined by sharing – via the comitative prefix *so-* – the event, place, activity, object, time span, or quality denoted by the base *X* with another participant. These formations are inherently deictic, since the reference point – the co-participant with respect to whom the referent is identified – is not morphologically expressed. The interpretation of the derivative therefore relies on relational anchoring to an implicit participant, recoverable through shared knowledge in well-established forms or through contextual cues in nonce formations.

So- co-participative derivatives belong to a broader class of deictic nouns that identify an animate referent in relation to another animate participant serving as a reference point. This deictic family also includes formations with other prefixes, such as *za-X-nik* and *pod-X-nik*, illustrated by *zapplečnik* (colloquial) ‘trainee who learns by observing another person at work’ (lit. ‘person positioned behind [another person’s] shoulder’) and *podkablučnik* (humorous) ‘henpecked man’ (lit. ‘man under [a woman’s] heel’).

From a frame-semantic perspective (Fillmore 1982), these deictic derivatives evoke relational frames in which the referent is profiled against an implicit but semantically salient reference point. In *so-*derivatives, this reference point corresponds to a co-participant in a shared event or situation, activating a co-participative frame. By contrast, in *zapplečnik* and *podkablučnik* the reference point is construed spatially as a figure of metaphorical proximity or dominance. In all cases, the reference point remains unexpressed but is crucial for the interpretation of the derivative, confirming the fundamentally relational and deictic nature of these constructions.

3 Theoretical Background

This study adopts Construction Grammar (CxG) as its primary theoretical framework, treating derivational patterns as conventionalized pairings of form and meaning that emerge from usage and are stored in speakers’ linguistic knowledge. From this perspective, formations following the [*so-X-nik*] pattern are not merely the mechanical outcome of concatenative derivation, but instantiate a constructional schema that licenses both entrenched lexical items and context-dependent nonce formations.

In CxG terms, the [*so-X-nik*] pattern can be represented as a partially specified morphological template with an open slot *X*. The construction combines the comitative prefix *so-* with the nominalizing suffix *-nik*, yielding nouns that denote animate human referents who stand in a relation of co-participation or shared reference with another participant with respect to *X*.

The construction can be schematized as follows, capturing the abstract pairing of form and meaning:

Form: [so-X-nik]_N

Meaning: <person who shares X with another participant (implicit reference point)>

In the data examined here, the X slot is overwhelmingly filled by nominal bases (with only marginal exceptions involving adjectival stems), most commonly denoting groups (e.g. *klass* > *so-klass-nik* ‘classmate’), activities (e.g. *trud* > *so-trud-nik* ‘collaborator’), places (e.g. *kamera* > *so-kamer-nik* ‘cellmate’), and abstract concepts (e.g. *vremja* ‘time’ > *so-vremen-nik* ‘contemporary’).

Metonymic extensions are also common, as in *so-butyl’-nik* ‘drinking buddy’, where the object *butylka* ‘bottle’ stands for co-participation in a drinking activity, and in *so-par-nik* ‘sauna buddy’, where the base *par* ‘steam’ evokes the shared sauna-bathing context.

Thus, at a lower level of abstraction, the construction can be specified by the following subschema:

Form: [so-N_[group/activity/place/object/other]-nik]_N

Meaning: <person who shares something (typically an activity, a place, or group membership) with another participant (implicit reference point)>

This subschema licenses both highly entrenched lexical items and low-frequency formations (e.g. *so-klan-nik* ‘clanmate’ < *klan* ‘clan’), and can extend to proper names (e.g. *so-zastekol’-nik* ‘fellow participant in the reality show *Za steklom*’).

The CxG approach also captures the deictic nature of [so+N-nik] derivatives. The construction presupposes an implicit reference point – another participant who shares N with the referent – thereby yielding a relational and indexical interpretation.

A salient property of the construction is its strong preference for plural marking, especially in nonce formations, as illustrated in (1). Pluralization of the pattern serves as a means of referring to non-lexicalized, occasional groups defined by the shared involvement of their participants in a salient situation or activity. At the same time, it constitutes a useful device for synthetic denomination, as it allows speakers to simplify the construction of referents by avoiding more complex denominations.

The plural pattern can thus be treated as a constructional extension of the matrix schema:

Form: [so-N_[group/activity/place/object/other]-nik-i]_{N,pl}

Meaning: <set of individuals grouped by shared, often occasional, co-participation in a group, activity, place, or situational frame involving N>

In sum, the CxG framework makes it possible to represent the [so-X-nik] construction as a network of hierarchically related sub-schemata within the broader family of deictic constructions, rather than as the output of a concatenative derivational rule.

More generally, a constructional approach allows for the integration of pragmatic and discourse-related factors into the analysis of the pattern, accounting for the relatively high frequency of [so+N-nik] occasionalisms in discourse genres that favor wordplay, creativity, and expressivity over strict lexical precision and terminological stability, such as journalistic writing and computer-mediated communication (CMC).

While this study does not explicitly adopt Relevance Theory (Sperber, Wilson 1995) as a formal framework, it shares its central assumption that context plays a crucial role in meaning construction. In particular, the interpretation of nonce [so-N-nik] formations depends on the interaction between schematic constructional meaning and contextual information, including shared frames, encyclopedic knowledge, pragmatic expectations, and the inferential identification of the implicit reference point. The referent is profiled against an unexpressed co-participant, whose identification is guided by discourse context, both linguistic and extra-linguistic, as well as by textual cues such as quotation marks, italics, or graphic highlighting, which frequently signal the non-lexicalized status of the derivative.

The nonce form *soplanetniki*, for instance, illustrates the inferential nature of this process. Although compositionally interpretable as ‘people living on the same planet’, the form is contextually interpreted as referring specifically to the Earth, yielding the enriched interpretation ‘fellow Earth-dwellers’. This interpretation is reinforced by its contrast with *inoplanetjane* ‘extraterrestrials’ in the article’s headline *Inoplanetiane ili soplanetniki?* (‘Extraterrestrials or fellow Earth-dwellers?’) (RNC, *Argumenty i fakty*, 1989.09.16). This example shows that identifying the intended referent depends on contextual inference rather than on constructional semantics alone.

4 Methodology and Data

The study adopts a corpus-based methodology within the framework of CxG, treating the [so-X-nik] pattern as a constructional schema whose productivity can be empirically assessed through usage data. The analysis targets both entrenched lexical items and low-frequency or occasional formations, which are central to a constructional view of derivational productivity.

Two datasets were compiled from two different corpora. The first was extracted via Sketch Engine from the RT corpus, a large monitoring corpus updated daily from RSS feeds of news articles and

other media sources, growing by approximately 3-4 million words per day. As a newsfeed-based corpus, RT is particularly sensitive to occasional and nonce formations, since media discourse is inherently oriented toward expressivity, evaluation, and stance-taking. These communicative goals favor the use of ad hoc lexical formations that create context-dependent labels and encode shared frames and attitudes with high semantic density. The results based on this dataset are discussed in §4.1.

The second dataset was extracted from the national media subcorpus of the RNC. It serves as a complementary resource, allowing the observation of [*so-X-nik*] formations in more traditional and editorially stabilized media discourse, and is especially useful for identifying entrenched items and comparing their distribution over time, given the corpus's longer temporal coverage.

Corpus queries targeted noun lemmas matching the formal pattern [*so-X-nik*]. All automatically retrieved items were manually filtered to exclude typographical errors, OCR noise, and formally similar but structurally unrelated forms that do not instantiate the construction under investigation.

To assess the productivity of the [*so-X-nik*] construction, the study follows quantitative approaches to morphological productivity developed in corpus linguistics (Baayen 2009). Productivity is evaluated not only in terms of overall type or token frequency, but primarily through the distribution of low-frequency types. Special attention is therefore paid to *hapax legomena*³ and low-frequency formations, which are taken to reflect speakers' ability to access the abstract constructional schema and to exploit it for the creation of new lexical items.

Following Baayen (2009, 902), the 'potential productivity' of a pattern can be estimated as the ratio between the number of hapax legomena and the total number of tokens in a corpus, capturing the growth rate of the lexical inventory associated with that pattern. From this perspective, a construction characterized by a small set of highly frequent, entrenched types alongside a large number of low-frequency or single-occurrence forms can be considered potentially productive, as it remains open to creative extension.

This quantitative approach is fully compatible with a CxG framework, where productivity is understood as a usage-based property of constructions. Low-frequency and occasional formations are therefore not treated as peripheral phenomena, but as crucial

3 While nonce nouns constitute a semantic category referring to context-dependent formations, *hapax legomena* are used here as a corpus-based proxy for such formations, providing empirical evidence of the construction's availability for occasional coinage.

evidence for the schematic abstraction of the [so-X-nik] pattern and for its role as a productive linguistic resource in discourse.

4.1 Corpus Analysis: Case Study 1 (*Russian Trends*)

The CQL query [lemma=“s.*nik” & tag=“N.*”] was conducted on 30 December 2025 to retrieve lexical realizations matching the formal pattern [so-X-nik] from the RT corpus, which at the time of the query contained approximately 3,3 billion tokens. Manual filtering excluded typos (for instance, *sobesednik* ‘interlocutor’ occurs in over 20 misspellings such as *sobesedenik*, *sobsednik*, *sobednik*), formally similar but structurally unrelated items (e.g. *sovetnik* ‘counsellor’ < *sovet* ‘advice’), and hyphenated compounds containing one element belonging to the pattern (e.g., *sotrudnik-kontraktinik* ‘contract employee’).

The final dataset comprises 33 relevant types and includes both highly frequent, entrenched forms,⁴ such as *sotrudnik* ‘collaborator’ (1.639.948 tokens, 496,88 instances per million - ipm), *sobesednik* ‘interlocutor’ (220.787 tokens, 66,89 ipm), and *sootečestvennik* ‘compatriot’ (51.860 tokens, 15,71 ipm), and low-frequency items with values below 0,002 ipm, including *solagernik* ‘fellow inmate of a labor camp’ (5 tokens), *somolitvennik* ‘prayer companion’ (4 tokens), and *sonaslednik* ‘co-heir’ (3 tokens).

In addition, the dataset contains 7 hapax legomena: *sopodel’nik* ‘accomplice (of a thief)’, *soslužebnik* ‘service colleague’, *sokelejnik* ‘monastic cellmate’, *sootrijadnik* ‘member of the same unit’, *sodel’nik* ‘accomplice’, *sorazmernik* ‘entity of the same size’, and *sosovetnik* ‘second adviser’. Two of these (*sootrijadnik* and *sodel’nik*) were attested very recently in the corpus, in June and November 2025 respectively. As noted by Baayen (2009, 902), the presence of hapax legomena provides direct evidence of the pattern’s potential productivity, showing that the [so-X-nik] construction remains available for creative formation of new lexical items beyond entrenched forms.

The following table summarizes the dataset distribution in terms of ipm (concordance size: 2.717.865):

⁴ Although highly entrenched nouns such as *sputnik* ‘companion; satellite’ and *sverstnik* ‘peer, person of the same age’ are diachronically related to patterns of co-participative derivation, the prefix *s-* in these forms is synchronically non-segmentable and non-productive. From a synchronic perspective, therefore, *s-* cannot be analyzed as an allomorph of the productive prefix *so-*. In line with a usage-based and synchronically oriented CxG approach, these forms do not instantiate the construction under investigation and are consequently excluded from both datasets created for the analysis of the [so-X-nik] pattern.

- $50 < \text{ipm} \leq 600$: *sotrudnik* ‘collaborator’ (496,88), *sobesednik* ‘interlocutor’ (66,89)
- $10 < \text{ipm} \leq 50$: *sootečestvennik* ‘compatriot’ (15,71)
- $5 < \text{ipm} \leq 10$: *soratinik* ‘comrade, ally’ (8,86), *soobščnik* ‘accomplice, conspirator’ (7,13), *součastnik* ‘co-participant’ (7,11), *sovremennik* ‘contemporary’ (5,41)
- $0,01 \leq \text{ipm} \leq 5$: *sobutyl’nik* ‘drinking buddy’ (1,43), *soplemennik* ‘tribesman’ (0,47), *sokursnik* ‘course mate’ (0,36), *sokamernik* ‘cellmate’ (0,09), *součenik* ‘classmate’ (0,04), *sokomandnik* ‘teammate’ (0,02), *sotrapeznik* ‘fellow diner’ (0,01)
- $\text{ipm} < 0,01$: *sosobstvennik* ‘co-owner’, *soplatformennik* ‘member of the same political platform’, *solagernik* ‘fellow camp inmate’, *somolitvennik* ‘fellow worshipper’, *sonaslednik* ‘co-heir’, *sorabotnik* ‘co-worker’, *soimennik* ‘namesake’, *soklassnik* ‘classmate’, *sopričastnik* ‘participant; person sharing involvement’, *soposrednik* ‘co-mediator’, *sovernik* ‘coreligionist; fellow believer’, *soklubnik* ‘fellow club member’, *sopodel’nik* ‘partner in crime’, *soslužebnik* ‘colleague; fellow employee’, *sokelejniki* ‘fellow cell-dweller (monastic cell)’, *sootrjadnik* ‘member of the same squad/detachment’, *sodel’nik* ‘accomplice’, *sorazmernik* ‘entity of equal measure’, *sosovetnik* ‘co-counsellor; fellow adviser’.

Forms in which the prefix *so-* combines directly with a base already suffixed with *-nik*, such as *so-rabotnik* ‘co-worker’ or *so-učastnik* ‘co-participant’, were retained. Although these formations do not constitute simultaneous prefix-suffix derivation from a strictly concatenative perspective, they fully conform to the co-participative construction from a CxG perspective.

These forms reveal that prefix and suffix are hierarchically organized, reflecting semantic and processing priorities. The nominalizing suffix *-nik*, which introduces an animate participant, occupies a higher rank than the co-participative prefix *so-*, whose function is to relate this participant – whether lexicalized or newly coined – to an implicit reference point.

This aligns with the psycholinguistic ‘complexity-based ordering model’ (Hay, Plag 2004; Baayen 2009, 915), which predicts that more productive and salient affixes tend to follow less productive ones. From this perspective, *-nik* functions as the core categorial operator, while *so-* operates deictically, anchoring the referent to its reference point.

Empirical support comes from the nonce formation *so-podel’nik*, attested in the corpus in November 2023. The base noun *podel’nik* already denotes an ‘accomplice’, i.e. a participant in a shared criminal act; the addition of *so-* is therefore referentially redundant and

constitutes a clear case of morphological ‘overcoding’,⁵ as it does not add new information. Nonetheless, its attestation demonstrates the ongoing productivity of the [so-X-nik] construction. The form appears in a polemical text predicting geopolitical collapse in Europe, evoking socio-political chaos that foregrounds criminal behavior:

- (2) *Odnako Borrelli smojutsja v svoë vremja na tëplye mestečki i sytnye pensii, a žiteli Evropy tam i ostanutsja ni s čem na razgrablennoj i razbomblennoj territorii. Gromče vsego na Vostoke – “Ubili, ubili!!!” – kričit **sopodel'nik** vora, kotoryj ili uže ukrad, ili vot-vot ukradët. V panike, smute, chaose i nerazberiche tolpy vor ili staja vorov čuvstvujut sebja, kak ryba v vode.* (RT, news2.ru)

‘Meanwhile, the Borrells will in due time slip away into comfortable positions and generous pensions, while the inhabitants of Europe will be left with nothing, on looted and bombed-out territory. Loudest in the East – “They’ve killed us, they’ve killed us!!!” – shouts the accomplice of a thief, who has either already stolen or is about to steal. In panic, turmoil, chaos, and mass confusion, a thief or a pack of thieves feels like a fish in water.’

Here, co-participation is already lexically encoded in the base noun and ensured by the context. The redundancy is therefore not accidental but pragmatically motivated. The surrounding text is saturated with elements activating a THEFT / CRIMINALITY frame: the reference point vor ‘thief’ is explicitly named, predicates such as ukrast’ ‘to steal’ and razgrabit’ ‘to plunder’ reinforce systematic wrongdoing, and nouns like panika ‘panic’, smuta ‘turmoil’, chaos ‘chaos’, and nerazbericha ‘confusion’ construct a background of disorder. Within this frame, sopodel’nik does not primarily serve an identificational function; rather, it is expressive and evaluative, foregrounding complicity as a morally salient, negatively charged property. In this sense, redundancy functions as a discourse strategy, whereby morphological ‘overcoding’ supports subjective evaluation rather than referential specification.

More generally, nonce derivatives rely heavily on contextual cues for interpretation. Their meaning is strongly context-driven, as illustrated by *sorazmernik* (< *razmer* ‘size’), used in automotive media discourse to denote a vehicle of comparable dimensions. Out of context, nothing in the form of the derivative signals the relational anchoring to the reference point.

⁵ The notion of ‘overcoding’ captures instances where different linguistic means encode more information than is strictly necessary for referential identification, often for pragmatic or evaluative effect. In the domain of discourse based word formation, pervasive overcoding has been observed in cases where multiple means of categorization co-occur to encode non-lexicalized concepts (e.g., in rhyming and imitative co-compounds in Russian; Alexandrova, Benigni 2021).

- (3) *Forthing T5 – èto “sorazmernik” Geely Atlas Pro. [...] Avtomobil’ vydelaetsja bolee tradicionnym dizajnom, v otličie ot rodstvennogo krossovera T5 Evo.* (RT, motorpage.ru)

‘The Forthing T5 is a size counterpart of the Geely Atlas Pro. [...] The car stands out for a more traditional design, unlike its related crossover, the T5 Evo.’

This is the only instance in the dataset where the derivative denotes a non-animate entity. It can be analyzed as a metonymic extension of the [so-X-nik] pattern, shifting from denoting <a person who shares X with another participant (implicit reference point)> to <an object that shares X with another object (implicit reference point)>. This illustrates the versatility of the construction, particularly for creating ad hoc, context-dependent concepts, a function also signaled graphically by the use of quotation marks.

4.2 Corpus Analysis: Case Study 2 (*National Media Subcorpus of RNC*)

A lexical query *so*nik* combined with the grammatical tag S (noun) was conducted on 30 December 2025 to retrieve lexical realizations of the [so-X-nik] pattern from the national media subcorpus of the RNC. As in the RT case study, the results were manually filtered to exclude typos, formally similar but structurally unrelated items, and hyphenated compounds containing an element belonging to the construction.

After filtering, the dataset comprises 67 types instantiating the [so-X-nik] construction. This result is noteworthy given that the RNC media subcorpus is substantially smaller than RT (approximately 818 million tokens, versus 3,3 billion). The larger number of types can plausibly be explained by the diachronic depth of the RNC media subcorpus: because it includes media texts published between 1983 and 2021, it is more likely to capture low-frequency or occasional formations that may not emerge in a synchronic monitoring corpus. The distribution of types by ipm frequency is summarized below:

- 50 < ipm ≤ 600: *sotrudnik* ‘co-worker; colleague’ (574,55), *sobesednik* ‘interlocutor’ (103,43)
- 10 < ipm ≤ 50: *sootečestvennik* ‘compatriot’ (34,44), *soratnik* ‘comrade; ally’ (24,18), *sovremennik* ‘contemporary’ (13,34), *soobščnik* ‘accomplice; conspirator’ (11,31)
- 5 < ipm ≤ 10: *součastnik* ‘co-participant’ (5,66)
- 0,01 ≤ ipm ≤ 5: *sokamernik* ‘cellmate’ (1,72), *soplemennik* ‘fellow tribesman’ (1,30), *sobutyl’nik* ‘drinking buddy’ (1,04), *sokursnik* ‘course mate’ (0,96), *součenik* ‘classmate’ (0,25), *sosobstvennik* ‘co-owner’ (0,13), *sotrapeznik* ‘fellow diner’

(0,07), sorabotnik ‘co-worker’ (0,03), solagernik ‘fellow camp inmate’ (0,03), soklubnik ‘clubmate; fellow club member’ (0,02), soplatformennik ‘member of the same political platform’ (0,02), soposrednik ‘co-mediator’ (0,02)

- ipm < 0,01: sogruppnik ‘member of the same group’, sokomandnik ‘teammate’, soimennik ‘namesake’, sovoprosnik ‘person participating in the same interview’, souznik ‘fellow prisoner’, sonaslednik ‘co-heir’, sopostnik ‘fellow faster’, sopričastnik ‘participant; one sharing involvement’, somolitvennik ‘co-worshipper’, sopalatnik ‘roommate; ward mate’, soplanetnik ‘fellow inhabitant of the same planet’, sočuvstvennik ‘sympathizer’, soprocessnik ‘person involved in the same legal proceedings’, soavtobusnik ‘fellow bus passenger’, sodel’nik ‘associate in the same undertaking’, soartnik ‘fellow artist’, sobločnik ‘fellow alliance member’, sovinovnik ‘person sharing the guilt’, sovospitannik ‘fellow pupil’, sodružebnik ‘person sharing a bond of friendship’, soduchovnik ‘spiritual companion’, sozastekol’nik ‘co-participant in the reality show Behind the Glass’, sokvartetnik ‘fellow quartet member’, soklannik ‘fellow clan member’, soknižnik ‘fellow contributor to the same book’, sokontinentnik ‘person from the same continent’, sokortnik ‘fellow tennis player’, somogil’nik ‘person sharing the same grave’, somodul’nik ‘person sharing the same module’, soobednik ‘person sharing a meal’, sootvetstvennik ‘person sharing responsibility’, soparnik ‘sauna buddy’, sopatentnik ‘co-patentee’, sopečal’nik ‘person sharing another’s sorrow’, sopodel’nik ‘person involved in the same affair’, sopokolennik ‘person of the same generation’, sopostel’nik ‘person sharing the same bed’, soputešestvennik ‘fellow traveller’, soraDnik ‘fellow supporter’ (playful spelling of soratnik), sorejsnik ‘fellow passenger on the same flight’, sotainnik ‘person sharing a divine mystery’, sotrampeznik ‘Trump supporter’ (playful blend based on sotrapeznik ‘fellow diner’), soumyšlennik ‘person sharing the same intent’, sofuršetnik ‘person attending the same reception’, soèkipažnik ‘crewmate’, soèlitnik ‘person belonging to the same elite’, soètažnik ‘person on the same floor’, sojazyčnik ‘person sharing the same language’.

As expected, the high-frequency and entrenched items largely overlap with those attested in the RT corpus. The main differences concern low-frequency forms and hapax legomena, which amount to 32 items in the RNC dataset and constitute the primary focus of the qualitative analysis in the following section.

4.2.1 Hapax Legomena: Contextual Anchoring and Wordplay

The qualitative analysis of single-occurrence forms confirms that the [so-X-nik] pattern is not merely a closed set of lexicalized items, but an active constructional schema that speakers exploit to create ad hoc denominations. In line with Baayen (2009, 902), single-occurrence forms provide direct evidence of speakers' access to abstract derivational schemata, making them particularly relevant for assessing the pattern's productive potential.

The low-frequency and hapax forms attested in the RNC cluster into coherent semantic classes according to the X component, which introduces the frame within which the co-participative relation is construed:

- member of the same group or institution: *sogruppnik* 'member of the same group' < *gruppa* 'group', *soklannik* 'member of the same clan' < *klan* 'clan', *soklubnik* 'clubmate' < *klub* 'club', *sokvartetnik* 'member of the same quartet' < *kvartet* 'quartet', *soèkipažnik* 'member of the same crew' < *èkipaž* 'crew', *soèlitnik* 'member of the same elite' < *èlit* 'elite';
- person co-located in the same place or spatial setting: *soètažnik* 'person sharing the same floor' < *ètaž* 'floor', *somogil'nik* 'person buried in the same grave' < *mogila* 'grave', *sopostel'nik* 'bedfellow' < *postel* 'bed', *sokortnik* 'fellow tennis player; person sharing the same tennis court' < *kort* 'tennis court';
- person participating in the same activity or event: *soobednik* 'person sharing a meal' < *obed* 'meal', *sofuršetnik* 'participant in the same buffet' < *furšet* 'buffet';
- person showing empathic, moral or ideological alignment: *sopečal'nik* 'co-sufferer, companion in sorrow' < *pečal* 'grief, sorrow', *soumyšlennik* 'person sharing the same intent' < *umysel* '(criminal) intent'.

Most of these formations are highly compositional and readily interpretable by analogy with entrenched members of the same semantic classes. Thus, *sotrapeznik* 'fellow diner' provides a model for nonce formations such as *sofuršetnik* 'buffet companion' or *soobednik* 'dining companion'.

More generally, however, interpretation is also supported by contextual anchoring. In (4), the surrounding cotext activates both the relevant frame (*trapeza* 'meal') and the constructional template (through the entrenched form *sobesednik* 'interlocutor'), while an explicit reformulation of *soobednik* (*to est' soedok različnyh jastv* 'that is, a person who shares a variety of dishes') removes any residual ambiguity.

- (4) *Parallel'no s ich razgovorom šla trapeza, v kotoroj Chottabyč vpolne učastvoval ne tol'ko kak **sobesednik**, no i kak **soobednik**, to est' soedok različnyh jastv.* (RNC, Chottabyč — II / Kommersant, 2000-08-13)

‘Parallel to their conversation, there was a meal in which Khotabych participated not only as a conversational partner but also as a dining companion, that is, a person who shares a variety of dishes’

Other nonce formations rely even more heavily on contextual anchoring and shared background knowledge. In (5), *soklannik* ‘fellow clan member’ is interpretable by analogy with more entrenched formations from the SOCIAL GROUP semantic class, such as *sokamernik* ‘cellmate’ or *soklassnik* ‘classmate’. At the same time, its interpretation is supported by the explicit mention of the reference point of the co-participative relation (‘a powerful mafia clan’) and by quotation marks signaling evaluative, occasional usage.

- (5) *Otec nekogda vchodil v moguščestvennyj mafioznyj klan, no potom sdal FBR svoich byvšich “**soklannikov**”.* (RNC, Knižnaja polka: pljažnoe čtivo / Komsomol'skaja pravda, 2005-07-07)

‘The father once belonged to a powerful mafia clan, but turned his former “clanmates” over to the FBI.’

In (6) the form *sozastekol'niki* ‘participants in the reality show *Za steklom* (‘Behind the glass’)’ presupposes familiarity with the title of the television show, which is nonetheless mentioned in the article’s headline. Once again, quotation marks highlight the occasional nature of the word formation:

- (6) *On poočeredno protanceval so vsemi. [...] Žanna očen' staralas', i vposledstvii Maks pri vystavlenii ocenok “**sozastekol'nikam**” osobo otmetil ee sposobnosti.* (RNC, “Za steklom”: Predskazaniya Margo / Izvestija, 2001-11-22)

‘He danced in turn with everyone. [...] Zhanna tried her best, and later, when Max graded the “Behind the Glass” participants, he especially highlighted her talent.’

A particularly salient subset of hapax legomena exploits graphic and morphological wordplay. In (7), the nonce formation *SoARTniki* (‘fellow ARTists’) blends the award name *Soratnik* (‘Comrade’) with the capitalized sequence ART, referring to the Art-Strelka cultural centre where the award ceremony takes place.

- (7) *na ArtStrelke v tretij raz vručili premiju “**Soratnik**” **SoARTniki** (to est' sami chudožniki) golosovali za naibolee jarko projavivšich sebja za god kolleg.* (RNC, Noč' v muzeje: Tret'jakovku brali šturmom / RIA Novosti, 2008-05-19)

‘at ArtStrelka, the “Soratnik” award was presented for the third time. The SoARTniki (i.e., the artists themselves) voted for colleagues who had distinguished themselves over the past year.’

The formation takes the event setting as its reference point for the co-participative interpretation while simultaneously evoking the award name *Soratnik*, allowing the identification of a complex ad hoc category, namely ‘artists affiliated with the Art-Strelka centre who participate in voting for the “Soratnik” award’. Its interpretation, however, depends on shared contextual knowledge, since its meaning is anchored to this specific discourse setting.

A similar mechanism underlies the nonce noun *sotrampeznik* in (8), a playful designation for a ‘Trump supporter’, which is formed through playful substitution of the segment *-trap-* with the proper name *Tramp* within the entrenched form *sotrapeznik* ‘fellow diner’. Here, the metalinguistic paraphrase (*storonnik Trampa* ‘supporter of Trump’) explicates the wordplay, while the associative link with the meaning ‘fellow diner’ ironically presents ideological support as a form of symbolic commensality – an implication that is largely lost in translation.

- (8) **Sotrampeznik – sotrapeznik** – *storonnik Trampa*. (RNC, *Trampotnja i trampizmy* / *Moskovskij komsomolec*, 2016-10-13)

‘A Trump supporter – a fellow diner – a supporter of Trump.’

A related but structurally revealing case is illustrated by the nonce formation *soraDniki* in (9), which exploits graphic wordplay to evoke the adjective *rad* ‘happy’ within the conventional form *soratniki* ‘comrades; fellow supporters’. The capitalized D invites this secondary reading, whose activation is reinforced by the surrounding cotext, notably the verbal form *raduet* ‘makes happy’. Rather than replacing the conventional reading, the graphic wordplay introduces an additional affective interpretation, yielding ‘supporters exchanging New Year wishes of happiness’, or more compactly, ‘happy fellow supporters’.

- (9) *Pust' raduet sem'ja i druž'ja, rabota tože (chotja by zarplatoj). [...] V obščem, pust' u vsech vezde i vseгда vse budet chorošo, s Novym godom soraDniki!* (RNC, *Čto zagadyvajut futbol'nye bolel'sčiki na Novyj god? Obzor gostevykh klubov rossijskoj Prem'er-ligi* / *Sovetskij sport*, 2012-01-02)

‘Let the family and friends be happy, and work too (at least with the salary). [...] In general, let everything be good for everyone everywhere and always. Happy New Year, happy fellow fans!’

In sum, these examples show that nonce formations instantiating the [so-X-nik] pattern are systematically interpretable through a

combination of constructional analogy and contextual anchoring. Their successful interpretation crucially depends on shared and encyclopedic knowledge, underscoring both the productivity of the pattern and its sensitivity to wordplay.

5 Concluding Remarks

This study has shown that the Russian construction [so-X-nik] constitutes a productive and structurally flexible word-formation pattern, capable of supporting fully lexicalized items, low-frequency derivatives, and nonce formations with strongly context-dependent meanings. Far from being a marginal or residual mechanism, the construction emerges as a robust resource for naming occasional social roles and participant configurations.

From a semantic perspective, the analysis has highlighted the centrality of deictic co-participative meaning, whereby the construction profiles a participant defined through shared involvement in an activity, affiliation, situation, or salient property with respect to an implicit reference point. This meaning proves sufficiently flexible to accommodate a wide range of fillers in the X-slot, including proper-name and phrasal bases, as in the case of *sozastekol'niki* ‘participants in the reality show *Za steklom*’, as well as forms involving graphic reanalysis and wordplay. The discussion of nonce formations has further shown that the construction is particularly sensitive to metonymy, analogy, and wordplay, allowing speakers to exploit formal similarity and contextual cues to generate novel yet interpretable meanings, including expressive ones.

The corpus analysis demonstrates the productive use of the [so-X-nik] construction in media discourse, where the pressure to rapidly and efficiently label new social configurations, ideological alignments, and participant roles is particularly strong. Importantly, the data challenge claims that prefix-suffix noun formation in contemporary Russian is only weakly productive. Instead, [so-X-nik] emerges as a living and creative pattern, especially visible in journalistic texts that allow for elaborated cotexts and pragmatic anchoring, in contrast to the more compressed environments of rapid newsfeeds. The comparison between the two corpora further reveals greater type diversity in the RNC dataset. This difference may plausibly be related both to the broader temporal coverage of the RNC corpus and to the availability of longer textual contexts, which provide richer discursive support for nonce formations than highly compressed newsfeed texts.

At the same time, the study opens several avenues for further research. These include investigating the relationship between the [so-X-nik] construction and alternative co-participative patterns (such as [odno-X-nik], derived through the prefixoid *odno-* ‘the same’), including pairs such as *sokursnik* ~ *odnokursnik*, both denoting ‘a person

studying on the same course’ and treated as synonymous in Russian lexicographic sources. Further research could also examine the role of textual and genre-specific factors in licensing nonce formations. Addressing these issues would further refine our understanding of how derivational morphology interacts with discourse strategies and cognitive operations of meaning construction.

Bibliography

- Alexandrova, A.; Benigni, V. (2021). “Chapter 12. Morphopragmatics of Rhyming and Imitative Co-Compounds in Russian”. Mauri, C.; Fiorentini, I.; Gorla, E. (eds), *Building Categories in Interaction: Linguistic Resources at Work*. Amsterdam; Philadelphia: John Benjamins Publishing Company, 317-54. <https://doi.org/10.1075/slcs.220.12ale>.
- Baayen, R.H. (2009). “Corpus Linguistics in Morphology: Morphological Productivity”. Lüdeling, A.; Kytö, M. (eds), *Corpus Linguistics. An International Handbook*. Berlin; New York: Mouton de Gruyter, 899-919. <https://doi.org/10.1515/9783110213881.2.899>.
- Benigni, V. (2018). “Ad hoc Categorization in Russian and Multifunctional General Extenders”. *Folia Linguistica*, 52(S1), 97-123. <https://doi.org/10.1515/flih-2018-0003>.
- Booij, G. (2010). *Construction Morphology*. Oxford: Oxford University Press.
- Efremova, T.F. (2000). *Novyj slovar' russkogo jazyka: Tolkovo-slovoobrazovatel'nyj* (New Dictionary of the Russian Language: Explanatory and Word-Formation Dictionary). Moskva: Russkij jazyk.
- Fillmore, C.J. (1982). “Frame Semantics”. Linguistic Society of Korea (ed.), *Linguistics in the Morning Calm. Selected Papers from SICOL-1981*. Seoul: Hanshin Pub. Co, 111-37.
- Goldberg, A. (2006). *Constructions at Work*. Oxford: Oxford University Press.
- Hay, J.; Plag, I. (2004). “What Constrains Possible Suffix Combinations? On the Interaction of Grammatical and Processing Restrictions in Derivational Morphology”. *Natural Language & Linguistic Theory*, 22, 565-96. <https://doi.org/10.1023/b:na1a.0000027679.63308.89>.
- Mineeva, Z.I. (2015). “Prefiksāl'no-suffiksāl'naja derivacija novych agentivov” (Prefix-Suffix Derivation of New Agent Nouns). *Vestnik Nižegorodskogo universiteta im. N.I. Lobačevskogo*, (5-6), 338-45.
- Sperber, D.; Wilson, D. (1995). *Relevance: Communication and Cognition*. 2nd ed. Oxford: Blackwell.
- Šanskij, N.M.; Tichonov, A.N. (1987). *Sovremennyj russkij jazyk: Slovoobrazovanie. Morfologija* (Contemporary Russian: Word Formation. Morphology). 2nd ed. Moskva: Prosveščenie.
- Švedova N.J. (ed.) (1980). *Russkaja grammatika. Tom 1: Fonetika. Fonologija. Udarenie. Intonacija. Slovoobrazovanie. Morfologija* (Russian Grammar. Volume 1: Phonetics. Phonology. Stress. Intonation. Word Formation. Morphology). Moskva: Nauka.
- Vasmer, M. (n.d.). *Ėtimologičeskij onlajn-slovar' russkogo jazyka* (Online Etymological Dictionary of the Russian Language). <https://lexicography.online/etymology/vasmer/>.
- Zemskaja, E.A. (2009). *Slovoobrazovanie kak dejatel'nost'* (Word Formation as Activity). 4th ed. Moskva: Librokom.

