Skolem Functions and Generalized Quantifiers for Negative Polarity Items Semantics

Ulyana Penzina[1] and Alexey Stukachev[1,2,3][0000−0002−4530−1961]

[1] Novosibirsk State University, Pirogova str. 1, Novosibirsk 630090, Russia
[2] Sobolev Institute of Mathematics, Koptyuga avenue 4, Novosibirsk 630090, Russia
[3] Novosibirsk State Technical University, Karl Marks av. 20, Novosibirsk 630073, Russia

u.penzina@g.nsu.ru, aistu@math.nsc.ru
http://www.old.math.nsc.ru/~stukachev

Abstract. We compare Skolem functions and generalized quantifiers as tools for representing formal semantics of natural language sentences. In particular, we discuss the case of Negative Polarity Items in languages with Negative Concord (for example, in Russian).

Keywords: Mathematical linguistics · Formal semantics · Generalized quantifiers · Skolem functions · Negative polarity items · Negative concord

1 Introduction

In mathematical linguistics, Negative Polarity Items (NPIs) and Negative Concord pose significant challenges in formal analysis of their semantics. In this paper we compare two approaches to the logical interpretation of NPIs and negative concord, based on generalized quantifiers and Skolem functions, respectively. Our analysis will explore sentences featuring both negative and indefinite pronouns using Skolem functions. The aim is to determine whether the presence of a negative particle in contexts with negative pronouns is obligatory for conveying their intended meaning, or serves merely as a syntactic element that does not impact the logical structure.

We indicate two main points of our analysis of generalized quantifiers and Skolem functions which show the latter are more suitable for the case of Negative Concord in Negative Polarity Items:

1) generalized quantifiers usually play a central role in the semantics of sentences (in particular, in monotonicity reasonings and entailments), while Skolem functions are technical in principle and allow other parts of a sentence (like verbs) to remain central;
2) functional types of Skolem functions $(e \to t) \to e$ are simpler than functional types of generalized quantifiers $(e \to t) \to ((e \to t) \to t)$ and have nothing on common with truth values and negation.

2 Polarity Items

Natural languages feature expressions that appear exclusively in "negative" contexts, known as negative polarity items (NPIs), as well as positive polarity items (PPIs), which exist solely in "positive" contexts.

Formal semantics has struggled to describe contexts hosting polarity items. Ladusaw [7] proposed that the environments licensing the use of NPIs share a common characteristic: NPIs are situated within the scope of "monotone decreasing functions" that permit transitions from sets to their subsets, but not the reverse.

In formal semantics, the basic domains are the sets of entities (type e) and of truth values (type t, with two ordered values $0 < 1$). For objects of type $(e \to t)$ (sets of entities, properties), "less or equal" is interpreted as "a subset of". In general, for inductively defined functional types $(a \to b)$, expressions of type $(a \to b)$ denote functions from objects of type a to objects of type b. The domain for expressions of type $(a \to b)$, $D_{(a \to b)}$ contains all functions from D_a to D_b. The ordering relation $\leq$ is defined for all functional types ending in t, incorporating the concept of material implication for sentences and extending the "subset of" relation for predicates. Monotonicity applies to functional types ending with t, meaning that for any f of type $(a \to b)$, f is monotone increasing if, for all x, y in D_a, $x \leq y$ implies $f(x) \leq f(y)$ or function f is monotone decreasing if for all x, y in D_a, $x \leq y$ implies $f(x) \geq f(y)$.

a. Positive Polarity Items (PPIs) are expressions licensed by monotone increasing functions;
b. Negative Polarity Items (NPIs) are expressions licensed by monotone decreasing functions.

The Ladusaw-Fauconnier Generalisation (LFG)[6] states: *occurrence within the argument of a decreasing function licenses negative polarity items, but occurrence within the argument of an increasing one does not.*

2.1 Classifications of NPI in Russian and English

There are many possible classifications of NPIs. For example, some researchers [14] classify negative pronouns as NPIs, whereas others, such as Paducheva [12] offers a nuanced view on the nature of Russian negative pronouns, arguing that they should not be classified as polarity items. She observes that NPIs operate independently without containing negation; instead, they are situated within or associated with a negative context (for instance, the use of *any*, see [11]).

Paducheva [13] categorizes negative polarity pronouns as a subset of indefinite pronouns in Russian, highlighting two distinct series: those ending in *-libo* (e.g., *kakoj-libo*, *chto-libo*) and those in the form of *by to ni bylo* (e.g., *kakoj by to ni bylo*, *chto by to ni bylo*).

Rossyaykin [15] elaborates on these ideas and points out the nuanced licensing requirements for negative concord items (hereinafter – NCIs), distinguishing the

NCIs from NPIs by the strict limitations of NCIs. Specifically, while the semantic content of a sentence influences the use of NPIs, NCIs necessitate proximity to a negation within the same clause for proper licensing.

In English, negative markers are represented by the so-called no-series: *no, nothing, no one/nobody, nowhere, never*. The class of negative polarity items include the *any*-series and the adverbs *ever, at all*. Negative pronouns belong to the class of negative markers because they function independently, which means that they do not require additional linguistic elements to express negation.

However, nowadays, a number of linguists, e.g., [10] claim that n-words themselves are not carriers of semantic negation. Rather, they should be regarded as semantically non-negative. That is why this paper adopts the stance that negative pronouns within negative concord contexts function as NPIs.

2.2 Negative Concord versus Double Negation

Negative concord is a phenomenon in which several formally negative units in a sentence express a single semantic negation. NC is observed in many languages; e.g. Romance, Slavic, Greek, Hungarian, Nonstandard English, West Flemish, Afrikaans, Lithuanian, Japanese [17].

Giannakidou distinguishes two classes of languages with negative concord: strict and non-strict [4]. In strict systems, the presence of negative words (n-words) necessarily requires an accompanying negative marker in the verb, regardless of their syntactic position in the sentence. This structure is characteristic of Slavic languages such as Russian and Polish. On the contrary, in languages with non-strict negative concord, such as many Romance languages, n-words can appear in front of a verb and licence negative concord with other n-words without the need for negation at the verb level. The analysis shows that even in languages with negative concord, NPI are often found, and it is sometimes possible to DN-read (and DN – double negation) some sentences. For example, in French. The use of the negative marker *pas* leads to a strictly definite DN interpretation in contexts corresponding to (1a), while in others, such as (1b), the use of DN-reading remains optional but pragmatically favoured:

French
(1) Ce n'est pas rien
(1a) It is not nothing;
(1b) = It is important;
(2) Personne n'aime personne [16]
(2a) No one loves anyone [NC];
(2b) = Everyone loves someone [DN].

The presence of strict and non-strict negative concord in such languages as Spanish, Italian, French emphasises the complexity and diversity of negation mechanisms in different linguistic systems, which require deeper study and understanding. In addition, [3] notes that even in non-standard English there are cases of negative concord: "Nobody said nothing to nobody" corresponds to

"Nobody said anything to anyone", but more often such usage is connected with dialectical peculiarities and is not normative.

The question of how negative concord can undergo compositional semantic analysis has been explored in both syntactic and semantic literature. According to the principle of compositionality, standard English is characterized as a DN-language, for which (1a) is equivalent to (1b), while this is not the case for other languages, e.g., Russian.

(1a) Nobody didn't sleep.
(1b) Everybody slept.

The presence of two negations within a predicate logic formula renders it truth-conditionally equivalent to a positive formulation in (2), considering the logical law of double negation, which posits that two logical negations cancel each other. Basic logical equivalences for formulas with negation are well-known:

1. For any formula Φ, $\neg\neg\Phi \Leftrightarrow \Phi$
2. For any variable x and any formula $\Phi, \neg\exists x\Phi \Leftrightarrow \forall x\neg\Phi$
3. For any formulas Φ and $\Psi, \neg(\Phi \wedge \Psi) \Leftrightarrow (\Phi \rightarrow \neg\Psi)$

(2) $\neg\exists x[person(x) \wedge \neg sleep(x)] \Leftrightarrow \forall x[person(x) \rightarrow sleep(x)]$

Also, issues arise when considering languages with strict NC. The synonymous sentences in (3a) and (3b) can be represented by the logical form (4a). However, the logical form (4b) can be wrongly assigned to (3b), by compositions of its parts, because in the Russian language, the combination of a negative pronoun with sentential negation does not yield a positive meaning:

(3a) Nobody slept
(3b) Nikto ne spal

(4a) $\lambda P.\neg\exists x[person(x) \wedge P(x)](\lambda v.sleep(v))$
(4b) $\lambda P.\neg\exists x[person(x) \wedge P(x)](\lambda v.\neg sleep(v))$

This conflict between the compositionality derived meaning and the actual interpretation of a sentence with NC illustrates the challenge that NC constructions pose for formal semantics.

3 Generalized Quantifiers

Montague's seminal paper [9] on PTQ presents an approach to formal semantics of quantifiers in natural language. It focuses on formal semantics of expressions like *every* and *some*. Montague's theoretical analysis is based on Frege Compositionality Principle: "The meaning of a complex expression derives from the meanings of its constituent parts and their syntactical arrangement". He proposes to consider a nominal group as a collection of properties, illustrating that the term *John* corresponds to a set of attributes uniquely ascribed to him, while *each person* refers to a set of qualities universally shared by people.

Based on Montague's groundwork, Barwise and Cooper [1] deal with the integration of generalized quantifiers theory with linguistic phenomena. A distinctive feature of this theory is a more convenient representation of quantifier expressions occurring in language.

$[\text{every student}]^e = \{P \subseteq D_e \mid [\text{student}] \subseteq P\}$
$[\text{most cats}]^e = \{P \subseteq D_e \mid |[\text{cat}] \cap P| > |[\text{cat}] \setminus P|\}$
$[\text{one in three cats}]^e = \{P \subseteq D_e \mid |[\text{cat}] \cap P| = 1/3|[\text{cat}]|\}$

3.1 Types of Quantifiers

Generalized quantifiers can be classified according to their types in lambda calculus [3]. For instance, quantifiers like "nobody" and "nikto" can be represented as functions that accept predicates and return truth values. In terms of types, the quantifier "nobody" can be represented as a function of type $(e \to t) \to t$, where e denotes the type of entities, and t denotes the type of truth values, see [8]. This means that the quantifier takes a function (or predicate) applicable to individual objects and returns a truth value. The types of main expressions are summarised in Table 1 for English and Table 2 for Russian, respectively.

Table 1. Categories and Types (English)

Category	Corresponding type	Basic expressions
Predicate negation	$(e \to t) \to (e \to t)$	not (He is **not** happy)
Sentential negation	$t \to t$	no, not (= It is **not** the case that)
Quantifier	$(e \to t) \to t$	somebody, anybody, nobody
Quantificational determiner	$(e \to t) \to ((e \to t) \to t)$	every, some, no
Skolem Functions	$(e \to t) \to e$	some, no (man)

Table 2. Categories and Types (Russian)

Category	Corresponding type	Basic expressions
Predicate negation	$(e \to t) \to (e \to t)$	ne (On **ne** spal iz-za shuma)
Sentential negation	$t \to t$	ne (= **Ne**verno chto)
Quantifier	$(e \to t) \to t$	kto-libo, nikto, nichego
Quantificational determiner	$(e \to t) \to ((e \to t) \to t)$	vsyakiy, kazhdij, kakoj-to
Skolem Functions	$(e \to t) \to e$	kakoj-libo, kakoj-to, nikakoj

3.2 Formalization of Negation

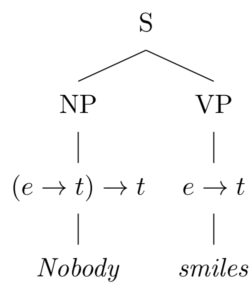
According to [5], the analysis of the English language identifies two principal forms of negation from a semantic perspective: predicate negation and sentential negation. It is important to note that in Russian linguistics, distinct sets of notions are used to distinguish semantic and syntactic forms of negation within a sentence, namely: general versus particular and predicate versus sentential [2].

Sentential negation is understood as general negation from a semantic standpoint. In contrast, predicate negation specifically targets the predicate within the sentence, as in the example "He is not kind." Predicate negation is expressed by the lambda expression $\lambda P.\lambda x.\neg P(x)$ with the type $(e \to t) \to (e \to t)$. An example of such negation is the phrase *It is not true that* (or *It is not the case that*), which can be analyzed as a function of type $t \to t$ that flips true to false and vice versa, $\lambda p.\neg p$. In English, the negation "not" can carry the meaning of both predicate and sentential negation, thus it has two lexical positions.

3.3 Generalized Quantifiers in Linguistics

Let's analyze the equivalent sentences "Nobody smiles" and "Nikto ne ulybaetsya" to identify one of the main challenges that arise when analyzing the types of expressions in sentences with negative concord. To do this, we will construct a formal representation of the English sentence using lambda functions and a syntactic tree with the indication of expression types, taking into account that "nobody" in this analysis is a generalized quantifier:

$$\lambda P.\neg\exists x.P(x)(\lambda x.smile(x))$$



When analyzing the corresponding sentence in Russian, a problem arises: despite the fact that the functional types of the expressions "nobody" and "nikto" are the same, the English sentence lacks a negative particle in the predicate.

Considering the functional types of expressions in the sentence "**Nikto ne** ulybaetsya" in Russian, the negation operator changes the truth value of the predicate that follows it. However, the type of the negative pronoun indicates a subsequent change in the truth value type. This would wrongly have implied that the sentence "Nikto ne ulybaetsya" should correspond to the sentence "Vse ulybayutsya (Everyone smiles)", which does not occur because, in negative concord, multiple formally negative particles in sequence express a single semantic negation.

However, according to the definition by Paducheva: "A negative pronoun is a non-referential indefinite pronoun *-nibud'* (someone) combined with negation, within whose scope it resides. For example, "nikto (nobody)" semantically equates to (not true that + -nibud'(someone))" [12]. Bearing this in mind, the following section will examine sentences with indefinite pronouns in Russian using Skolem functions, while considering that the sentential negation in these sentences corresponds to sentences with negative pronouns. This approach will be compared with the theory of generalized quantifiers.

4 Skolem Functions

Skolemization is a procedure by which existential quantifiers in a formula are eliminated. If an external quantifier in a formula is an existential quantifier, a Skolem constant is introduced in place of the variables bound by the existential quantifier. For example, $\exists x P(x)$ is converted to $P(c)$. If the existential quantifier is preceded by the universality quantifiers, a new k-local functional symbol f (a Skolem function that is not included in the formula) is introduced: $\forall y \exists x(P(y,x))$ is transformed into $\forall y(P(y, f(y)))$. In this paper, we shall use only the basic case of Skolem functions, with a single existential quantifier $\exists x P(x)$.

Skolem functions cannot be directly applied to negating existential statements; however, they are necessary for understanding how to interpret existential statements in affirmative formulations.

Let us then consider sentences with indefinite pronouns using Skolem functions, taking into account that the sentential negation of these sentences will correspond to sentences with negative pronouns. In first-order logic, the sentence "Someone smiles" corresponds to $\exists x(Smile(x))$ when using Skolem functions it corresponds to $Smile(c)$, involves replacing the existential quantifier with a specific function that selects the relevant entity without explicitly indicating it. The Skolem function acts as a witness to the truth of the existential statement; this function is independent of any other variables and selects an individual from the domain in such a way that the predicate "smiles" is true for this individual, where c is a Skolem constant.

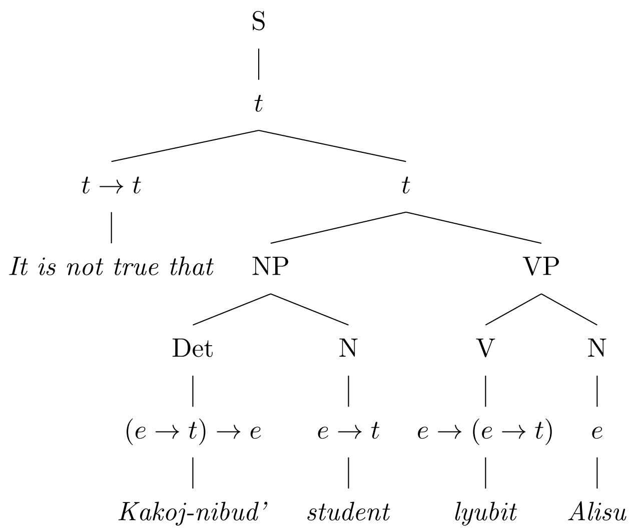
5 Generalized Quantifiers versus Skolem Functions

In order to identify the advantages of Skolem Functions in analysing sentences with NPI we will analyse a few examples comparing the functional types of generalized quantigiers and Skolem Functions.

Consider an example in English where a negative pronoun acts as a determiner. We can also construct a syntactic tree for it and specify the types of expressions, taking into account the functional types of GQ:

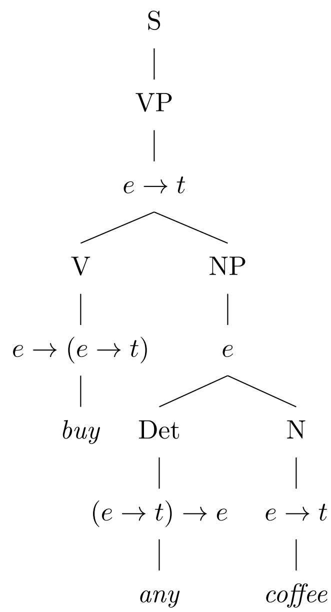
$$\lambda P.\neg\exists x.(Student(x) \wedge P(x))(\lambda x.love(x, Alice))$$

To consider the application of Skolem functions, we shall examine the sentence "Some student loves Alice". It corresponds to $\exists x(Student(x) \wedge love(x, a))$, so introduce a Skolem function $Student(f(a)) \wedge love(f(a), a)$, and examine how the types of expressions have changed. The main thing is that the type of the determiner has changed from $(e \to t) \to ((e \to t) \to t)$ to $(e \to t) \to e$. It is important to note that changing the functional type of the determiner allows shifting the focus from the generalized quantifier to the predicate. In this case, the Skolem function acts as an auxiliary tool in formalizing sentences with NPI. When a negative pronoun acts as a determiner in Russian, it is impossible to parse the types of expressions, as the truth type of the sentence “Nikakoj student ne lyubit Alisu” should change twice, but this does not occur due to the phenomenon of negative concord. Therefore, let us consider the sentential negation of the sentence “Kakoj-nibud’ student lyubit Alisu”. For this, we will construct a parse tree and note the types of expressions, where sentential negation is indicated by the phrase “it is not true that”, in the case of introducing a negative pronoun. In the sentence “Nikakoj student ne lyubit Alisu”, “kakoj-nibud’ (some)” is a Skolem function under negation, while “lyubit (loves)” changes to “ne lyubit (does not love)”, and “kakoj-nibud’” transforms into “nikakoj”. The Skolem function in the affirmative sentence guarantees the existence of a student who loves Alice, thus refining the semantics of the sentence.



6 Conclusions and Future Work

For future research in the application of Skolem functions to analyze natural language phenomena, it is crucial to mention that when analyzing natural language expressions using generalized quantifiers, several issues arise, one of which is the one-place predicate at the argument of the generalized quantifier. Noun phrases can be positioned as direct objects with a two-place predicate. In such cases, the verb and the quantifier expression cannot act as arguments to each other, leading to a breakdown in derivation. Pronouns from the *any-* series are categorized within the class of English NPIs. It is important to note that this study specifically examines the role of the determiner *any* when used as a direct object in negative sentences, as in these positions, *any-words* are translated by Russian negative pronouns. Consider the determiner *any*, noting that it possesses not a functional type $(e \to t) \to ((e \to t) \to t)$ but the type $(e \to t) \to e$ of a Skolem function. In this scenario, the noun phrase will have type e and can act as an argument for a two-place predicate.



Thus, in formalizing sentences containing a noun phrase with a pronominal determiner in the position of a direct object, the advantage of using Skolem functions over generalized quantifiers has been revealed: if the types of noun phrases (NPs) are considered using the functional types of generalized quantifiers, a problem known as the "quantifier in the object position" arises. However, if it is acknowledged that the construction has the functional type of a Skolem function, then the derivation will not be disrupted.

Acknowledgements The second author was supported by the IM SB RAS state agreement, project number FWNF-2022-012.

References

1. Barwise, J., Cooper, R.: Generalized quantifiers and natural language. In: Philosophy, language, and artificial intelligence: Resources for processing natural language, pp. 241–301. Springer (1981), https://doi.org/10.1007/978-94-009-2727-8_10
2. Coppock, E., Champollion, L.: Invitation to formal semantics. Manuscript, Boston University and New York University. eecoppock. info/semantics-boot-camp. pdf (2019)
3. Dowty, D.: The role of negative polarity and concord marking in natural language reasoning. In: Semantics and Linguistic Theory. pp. 114–144 (1994), https://doi.org/10.3765/salt.v4i0.2461

4. Giannakidou, A.: N-words and negative concord. The Blackwell companion to syntax pp. 327–391 (2006), https://doi.org/10.1002/9780470996591
5. Iordăchioaia, G.: Negative concord with negative quantifiers: A polyadic quantifier approach to Romanian negative concord. Ph.D. thesis, Tübingen, Univ., Diss. (2009)
6. Keenan, E.L., Westerståhl, D.: Generalized quantifiers in linguistics and logic. In: Handbook of logic and language, pp. 837–893. Elsevier (1997), https://doi.org/10.1016/B978-044481714-3/50020-5
7. Ladusaw, W.: On the notion affective in the analysis of negative-polarity items. Formal semantics: The essential readings pp. 457–470 (1980), http://dx.doi.org/10.1002/9780470758335.ch18
8. Loukanova, R.: Logic Operators and Quantifiers in Type-Theory of Algorithms. In: Bekki, D., Mineshima, K., McCready, E. (eds.) Logic and Engineering of Natural Language Semantics. LENLS 2022, Lecture Notes in Computer Science (LNCS), vol. 14213, pp. 173–198. Springer Nature Switzerland, Cham (2023), https://doi.org/10.1007/978-3-031-43977-3_11
9. Montague, R.: The proper treatment of quantification in ordinary english. In: Approaches to natural language: Proceedings of the 1970 Stanford workshop on grammar and semantics. pp. 221–242. Springer (1973), https://doi.org/10.1007/978-94-010-2506-5_10
10. Muchnová, D.: Negative concord in classical greek case study: Xenophon. Graeco-Latina Brunensia **24**(2), 163–176 (2019), https://doi.org/10.5817/GLB2019-2-11
11. Paducheva, E.: Negation. Materials for the project of corpus description of Russian grammar (http://rusgram.ru), on the rights of manuscript (2001)
12. Paducheva, E.: Implicit negation and pronouns with negative polarization. Voprosy Yazykoznaniya **1**, 3–18 (2011)
13. Paducheva, E.: Suspended assertion and nonveridicality: the case of russian negative polarity items. Russian Linguistics **39**, 129–162 (2015), https://doi.org/10.1007/s11185-015-9145-6
14. Pereltsvaig, A.: Polarity sensitivity as (non) veridical dependency by anastasia giannakidou. Language **76**(1), 219–220 (2000), https://doi.org/10.1353/lan.2000.0071
15. Rossyaykin, P.: Russian negative pronouns are licensed above negation (2021), https://doi.org/10.31862/2500-2953-2021-4-69-118
16. de Swart, H., Sag, I.A.: Negation and negative concord in romance. Linguistics and philosophy **25**, 373–417 (2002), https://doi.org/10.1023/A:1020823106639
17. Zeijlstra, H.: Sentential negation and negative concord. Ph.D. thesis, External Organizations (2004)