

Function and Support verb constructions in Slovenian naturalness theory

*Teodor Petrič**

Abstract

This study investigates Nominalization verb constructions (NVCs) in German, focusing on Functional verb constructions (FVCs) and Light/Support verb constructions (LVCs/SVCs). We evaluate their syntactic flexibility, semantic specificity, and processing complexity compared to base verbs (BVs) and free phrases (FPs) using Slovenian naturalness theory. Corpus analyses confirm that NVCs predominantly occur in less natural text types and grammatical environments. Experimental studies with Slovenian L2 German learners tested NVC processing. The results show that familiar FVCs facilitate processing, whereas LVCs cause processing delays due to deverbal noun complexity. These findings contradict the Underspecification Approach observed in German native speakers, and suggest that L2 learners primarily responded to the presence or absence of a resource-intensive nominalization rather than deeply processing these constructions in the same structured and dynamic way as native speakers. Our results contribute to understanding NVC representation in the mental lexicon and the cognitive mechanisms underlying L2 processing.

Keywords: Light verb construction, complexity, processing cost, naturalness, underspecification approach

Funkcijske in podporne glagolske zveze v slovenski teoriji naravnosti

V raziskavi se ukvarjamo s samostalniško-glagolskimi zgradbami (NVC) v nemščini, ki jih delimo na funkcijske glagolske zgradbe (FVC) in lahke/podporne glagolske zgradbe (LVC/SVC). Na osnovi slovenske teorije naravnosti ocenjujemo njihovo procesno zahtevnost v primerjavi z osnovnimi glagoli (BV) in prostimi besednimi zvezami (FP). Korpusne analize potrjujejo, da se NVC pogosteje pojavljajo v manj naravnih slovničnih okoljih in besedilnih vrstah kot BV. Eksperimenti s slovenskimi študenti nemščine kot drugega jezika (L2) kažejo, da dobro znana FVC pospešuje kognitivno obdelavo, kar podpira domnevo o reprezentaciji zgradbe v mentalnem slovarju, medtem ko LVC niso bile dosledno počasneje procesirane kot FP. Ti izsledki nasprotujejo hipotezi nedoločenosti, ki je bila podprta v eksperimentu z naravnimi govorniki nemščine, in nakazujejo, da so se študenti nemščine kot drugega jezika primarno

* Univerza v Mariboru, Filozofska fakulteta, teodor.petric@guest.um.si, teodor.petric@gmail.com

odzivali na prisotnost procesno zahtevne izglagolske izpeljanke. Naši rezultati prispevajo k razumevanju reprezentacije NVC v mentalnem slovarju in kognitivnih mehanizmov, ki so podlaga za procesiranje v drugem jeziku.

Ključne besede: Podporna/lahka glagolska zgradba, zapletenost, stroški kognitivne obdelave, naravnost, pristop nedoločenosti

1 Introduction

Nominalization verb constructions (NVCs) exist in many languages and typically consist of a semantically bleached verb (NV), a noun (NN) derived from a verb or adjective, and sometimes a preposition (e.g., German *in Erwägung ziehen*, *Bericht geben*; Slovenian *stopiti v veljavo*, *dati poljub*; English *do the cleaning*, *make a claim*). Compared to **free phrases** (FPs) (e.g., *einen Bericht lesen* / *read a report*), NVCs exhibit reduced referentiality, lower morphosyntactic flexibility, and more restricted combinatorial possibilities.

From a **diachronic** perspective, NVs evolved from full verbs through semantic bleaching, where they lost much of their original meaning and now primarily serve a grammatical function – providing verbal inflection for subject-verb agreement, tense, etc. – while the NN carries the core meaning and inherits the argument structure of its base verb.

German NVCs can be divided into **Function verb constructions** (FVCs) and **Support/Light verb constructions** (SVCs/LVCs) (Polenz 1988, 113–115). Certain NVs, such as *kommen* (*come*), *bringen* (*bring*), *stehen* (*stand*), and *finden* (*find*), have become highly grammaticalized, forming aspectually specific FVCs (e.g., *in Bewegung kommen* / *come into motion* vs. *sich bewegen* / *move*). Storrer's (2007) corpus study further highlighted how construction type and noun referentiality affect flexibility and modification – with FVCs showing reduced referentiality and stricter morphosyntactic constraints than SVCs, which take a direct object NN. Idiomatic NVCs are the most restricted. While NVCs are often semantically more specific than their base verbs, their use is governed by systematic factors rather than arbitrary choice.

Since NV and NN together form the predicate, this study compares NVCs to their base verbs (BVs) and examines their structural and functional similarities with free phrases (FPs). Accordingly, the article has the following **objectives**:

- To examine whether German NVCs are less natural than both their base verbs (BVs) and structurally equivalent free phrases (FPs) using Slovenian naturalness theory and to develop hypotheses for experimental and corpus-based studies.
- To analyse corpus data in German (L1) and experimental findings in German (L1 and L2), contextualizing them within our theoretical framework.

2 Theoretical background and hypotheses

Naturalness theory (see Dressler et al. 1987), utilized here in its Slovenian variant for hypothesis formulation, is a preference theory. The naturalness of a linguistic category is evaluated based on several universal, typological, and language-specific principles, such as uniformity, transparency, iconicity, structural constancy, and paradigmatic strength. **Slovenian Naturalness Theory (SNT)** evaluates the naturalness of (typically two) comparable syntactic variants based on principles like speaker/hearer parameters, cognitive effort, structural integration, frequency parameters, paradigmatic strength, processing effort, appropriateness, and typological distribution. The SNT predicts the contextual distribution of these variants (see Orešnik et al. 1990, Orešnik 2008, Cvetko Orešnik and Orešnik 2021).

Naturalness evaluations are represented using naturalness scales. The basic form of a naturalness scale is: $>\text{nat} (A, B) / \text{domain or principle}$, where the linguistic element or category A is considered more natural than B based on the specified principle or under the given circumstances (domain). Naturalness scales are essential for derivations in which SNT expresses its context-related predictions.

Compared to BVs, NVCs are strong syntactic variants, as they are syntactically more elaborate, less speaker-efficient, semantically more specific, and more listener-friendly than the corresponding BVs (Petrič 1993, 128–136, Petrič 1994, 190–194). The following naturalness scales support this evaluation:

- (a) A **less complex syntactic structure** (e.g., BVs) is easier for the speaker to compose and produce and is therefore, in line with the principle of **linguistic economy** (the “**drive for ease**”), more natural than a syntactically richer structure with a comparable function (e.g., NVCs). On the other hand, a **more elaborate syntactic structure** (e.g., NVCs) is **semantically more transparent**, making it advantageous for the listener, as it allows for **clearer decoding of the speaker’s**

intended meaning (the “drive for clarity”). This is reflected in the following naturalness scale: $>nat(BV, NVC)/\textit{principle of least effort}$, where NVCs are estimated to be less natural than BVCs.

- (b) According to the frequency principle, more frequently occurring linguistic categories (e.g., BVs) are more natural than less frequently occurring categories (e.g., NVCs), as more frequent categories can be retrieved from memory and produced more quickly by the speaker: $>nat(+, -)/\textit{token frequency}$.

The naturalness scales (a) and (b) can be aligned in parallel: if a syntactic construction V1 (e.g., an NVC) is cognitively more demanding than a functionally comparable construction V2 (e.g., a BV), we can expect that construction V1 (the NVC) will occur less frequently than V2 (the BV).

- (c) Strong syntactic variants are more likely to appear in **less natural grammatical environments**. Thus, NVCs are expected to occur more frequently in less natural contexts than in more natural ones (Petrič 1993, Petrič 1994) – for example, passive voice sentences are less natural than active voice sentences ($>nat(\textit{active}, \textit{passive})/\textit{voice}$). As such, NVCs are expected to occur more frequently than the BVs in passive voice sentences.
- (d) Strong syntactic variants are more likely to occur in **less natural text types**. Accordingly, NVCs – like other patterns characteristic of non-everyday communication – are expected to be more frequent in genres associated with distance communication conditions than in those associated with proximity communication conditions (Petrič 1998, 205–207, 219–221; cf. Koch and Oesterreicher 1990 for the proximity-distance model). For example, monologues are less natural than dialogues: $>nat(\textit{dialogue}, \textit{monologue}/\textit{situative parameters of communication})$. Thus, NVCs are expected to occur more frequently than V2 (the BVs) in monologic text types.
- (e) Strong syntactic variants require greater cognitive effort during retrieval, sentence encoding, and decoding. Since cognitive **processing is less automatic**, additional semantic processing during sentence assembly leads to increased processing costs, slowing down encoding and decoding ($>nat(\textit{automatic}, \textit{non-automatic})/\textit{cognitive processing}$). Thus, NVCs as strong variants are expected to slow down processing compared to BVs.

In contrast to **formally equivalent free phrases (FPs)**, the components of NVCs (FVCs, LVCs/SVCs) together form the predicate. This implies that NVCs are not fully compositional, as their meaning cannot be entirely inferred from the meanings of their components.

- (f) The semantically most defining component of NVCs is the NN, which largely inherits its argument structure from the base verb (BV). In contrast, prototypical nouns do not carry an argument structure. This distinction can be summarized by a naturalness scale, where a noun complement is more natural when it lacks an **argument structure**: *>nat(-,+)/argument structure of noun complement*. From this perspective, NVCs are considered less natural than FPs that contain nouns without argument structures inherited from verbs.
- (g) The **verbs in NVCs** are considered a variant of full verbs. NVs are typically morphologically simple, frequently occurring verbs with broad meanings. One of their semantic shades is the NV meaning, which results from semantic bleaching – typically less specific than in FP variants, making it open to interpretation and carrying only core semantic features (e.g., aspectual properties like terminativeness): e.g., *to give a summary* vs. *to give a gift*; *to come to the conclusion* vs. *to come to the conference*. From the speaker's perspective, it is convenient to use a single verb in different shades of meaning – functioning as both a full verb and a light verb depending on the noun complement. This evaluation can be summarized by the following naturalness scale: *>nat(light, full)/verb*.

Since the complex NN and its argument structure define the entire NVC, we generally expect NVCs to increase processing costs, leading to delayed encoding and decoding compared to FPs.

In psycholinguistics, two contrasting approaches explain LVC processing costs: the Separate Entry Approach and the Underspecification Approach. The Separate Entry Approach posits that LVCs are stored as distinct lexical units, allowing for faster processing due to direct retrieval and frequency effects. In contrast, the Underspecification Approach suggests that light verbs are stored without a fixed meaning, which is determined during sentence formation based on the noun complement. If the noun lacks an argument structure (e.g., *take an orange*), the verb remains non-light; if the noun has an argument structure (*take a walk*), the verb adopts a light meaning, integrating with the noun. Initially, the underspecified verb is retrieved without cost, but semantic integration in a second phase slows LVC processing.

- (h) **Idiomatic and highly familiar constructions** are assumed to be stored in the mental lexicon. According to the principle of least effort, constructions that can be readily retrieved from the mental lexicon involve more automatic processing, reducing cognitive load and making them more natural than those requiring additional resources during sentence assembly. Thus, idiomatic and familiar NVCs (especially FVCs) are expected to speed up processing compared to corresponding free phrases due to their greater automaticity.

3 Evidence from corpus studies

The assumptions formulated above were tested in two corpus studies, Petrič (1993) and Petrič (1998). Below some of the statistically significant results of Petrič (1993, 136–144) are summarized that support the above hypothesis that NVCs (FVCs and SVCs) behave like strong syntactic variants, i.e. as syntactic constructions with less natural properties occurring in **less natural grammatical environments**:

- NN-base verbs (e.g., *entscheid-*) and NNs (e.g., *Entscheidung*) were more morphologically complex, less frequent, semantically more specific than other verbs and nouns, respectively, in both FVCs and SVCs.
- Both FVCs and SVCs frequently appeared in dependent clauses, FVCs were particularly common in past tense clauses.
- In both FVCs and SVCs, NNs served more often as part of the informational focus (rheme) than other noun phrases.

Petrič (1998, 206–207) supports the assumption that NVCs typically occur in less natural text types, particularly in distance communication, which relies on formal linguistic strategies. The study showed that FVCs and SVCs are rare in proximity communication (i.e., spoken dialogues) but more frequent in distance communication (i.e., written monologic texts), aligning with this hypothesis. Corpus data from So (1991) further supports this claim. In 12 spoken dialogues from the Freiburg corpus (Institut für Deutsche Sprache 1974), FVCs and SVCs appeared in 1.7% and 6.3% of predicate expressions, respectively. In contrast, in five written monologues from So (1991, 160–199), these proportions increased to 2.7% and 7.4%. The difference between spoken and written contexts was statistically significant at $p < 0.05$ (Index of difference $DI = 4.8$ and $DI = 3.3$, respectively; cf. Pavlić 1971 for calculation of the non-parametric test).

4 Evidence from experimental studies

Two experiments were conducted to study the complexity of NVCs in German when learned as a foreign language (L2). Since the first experiment has already been published, only the second experiment will be described in more detail.

The **first experiment** by Basarić and Petrič (2015) demonstrated that familiar FVCs are retrieved holistically and may be represented as superlemmas (Sprenger et al. 2006) in the mental lexicon. Since FVCs occur more frequently than free phrases, they can be accessed more readily, resulting in lower processing costs during comprehension. Consequently, familiar FVCs were predicted to be processed faster than structurally similar free phrases, aligning with the least effort principle (see argumentation in (a) and (h)). This prediction was tested with 52 advanced Slovenian learners of L2 German using an adapted experimental design from Sprenger et al. (2006, 165–168). The participants were tasked to evaluate the meaning of sentences containing familiar FVCs, similarly familiar free phrases, and fillers. By using identical and unrelated primes, the experiment assessed whether the response times for FVCs were shorter when preceded by identical primes rather than corresponding free phrases. The findings supported our hypothesis: selected FVCs were activated more quickly after identical primes than structurally comparable free phrases. Consistent with Sprenger et al. (2006), we conclude that FVCs can be represented as superlemmas in the mental lexicon of German learners, provided they occur frequently enough in texts.

The **second experiment** by Basarić and Petrič aimed to demonstrate that the semantic complexity of LVCs slows down recipients' response times, in line with the principle of least effort (see argumentation in (e), (f) and (g)). Our experiment involved 46 Slovenian L1 speakers studying German as a foreign language at the Faculty of Arts in Maribor, all of whom also participated in the study by Basarić and Petrič (2015). We adapted the design and language materials of Wittenberg and Piñango (2011), using a cross-modal lexical decision task. Participants' CEFR proficiency ranged from B1 to B2, with a few at C1. Participants were divided into two lists: List A (21 participants) and List B (25 participants). In List A, reaction times were measured at verb offset (LC17), while in List B, they were measured 300 ms later (LC317). At both licensing points, a word probe appeared for ~400 ms, after which the participants decided whether it was a German word while listening to test sentences via headphones. Response time – the measure of processing load – was recorded based on their decision speed. Some example sentences from Wittenberg and Piñango (2011, 6) are shown below:

- (1) Light Verb Construction: Weil der Student seiner Kommilitonin vor dem Seminar eine *Zusammenfassung gab*, spendierte sie ihm letzte Woche einen Kaffee. (“give a summary”)
- (2) Same Noun Construction: Weil der Student seiner Kommilitonin vor dem Seminar eine *Zusammenfassung koptierte*, spendierte sie ihm letzte Woche einen Kaffee. (“copy a summary”)
- (3) Same Verb Construction: Weil der Student seiner Kommilitonin vor dem Seminar einen *Kugelschreiber gab*, spendierte sie ihm letzte Woche einen Kaffee. (“give a pen”)

We conducted a linear mixed-effects regression (Crawley 2013, 681–714) to examine the effect of construction type (LVC, Same Noun, Same Verb) on log-transformed response times (RTs), while controlling for probe word frequency and length. The model included random intercepts for Subject. At verb offset (LC17), linear mixed-effects models revealed a significant difference in RTs between LVCs and Same Verb constructions ($b=0.0863$, $SE = 0.0192$, $t=4.498$, $p<.0001$), indicating longer processing times for LVCs. A significant difference was also found between Same Noun and Same Verb constructions ($b=0.0733$, $SE = 0.0195$, $t=3.755$, $p=0.0005$). However, no significant difference was observed between LVCs and Same



Picture 1: Response Time at verb offset

Noun constructions. The results indicate that LVCs took approximately 9% longer to process than Same Verb constructions, while Same Noun constructions were processed ~7.6% more slowly than Same Verb constructions. At licensing point 300 ms after verb offset (LC317), linear mixed-effects models revealed a significant difference in RTs between LVCs and Same Verb constructions ($b=0.0547$, $SE=0.0178$, $t=3.076$, $p=0.0064$), indicating ~5.6% longer processing times for LVCs. However, no significant difference was observed between LVCs and Same Noun constructions and between Same Noun and Same Verb constructions.

5 Conclusion

This article examined Noun-Verb Constructions (NVCs), which can be subdivided into Function Verb Constructions (FVCs) and Light or Support Verb Constructions (LVCs, SVCs). Our goal was to provide evidence for categorizing NVCs as strong syntactic variants with less natural properties than their base verbs (BVs), and to demonstrate that NVCs are less natural than corresponding free phrases (FPs). Two corpus studies (Petrič 1993, Petrič 1998) showed that NVCs (both FVCs and LVCs) frequently occur in less natural grammatical environments and text types. The experimental study by Basarić and Petrič (2015) found that a subgroup of NVCs – familiar FVCs – were processed faster than free phrases by L2 German learners, supporting the view that FVCs could be stored in the mental lexicon as superlemmas, as demonstrated by Sprenger et al. (2006) for idioms. Additionally, Barking, Mos, and Backus (2025) demonstrated that not all light verbs freely combine with noun phrases (e.g., *make a walk, *do a decision), indicating that LVCs must be mentally stored as lexically specific units to ensure conventional combinations. The second experiment presented in this article provided evidence that constructions containing a nominalization with an inherited verbal argument structure slow down processing compared to free phrases that lack such a noun. We hypothesized that LVCs would show longer response times than constructions in the Same Noun and Same Verb conditions, aligning with our naturalness evaluation, which considers LVCs less natural than FPs. However, our results did not support either our hypothesis or either of the two contrasting psycholinguistic approaches – the Separate Entry Approach or the Underspecification Approach. At both verb offset and 300 ms later, LVCs were significantly slower than Same Verb constructions but not different from Same Noun constructions. This suggests that processing difficulty stems from the deverbal noun rather than the light verb, aligning with lexical semantics research on abstract nouns

and supporting our hypothesis that nominalizations are less natural than their base verbs. If LVCs were stored as whole units, they should have been processed faster, and if difficulty arose from late integration, RT differences should have appeared only 300 ms after verb offset – but neither pattern was observed. Instead, LVC difficulty emerges early, indicating that deverbal nouns drive the increased cognitive load. Our findings, based on Slovenian learners of German, contrast with those of Wittenberg and Piñango (2011), who found evidence for Underspecification in native German speakers. Their results showed no significant difference at verb onset, but 300 ms later, LVC response times increased significantly compared to non-LVCs, suggesting that LVCs require more cognitive resources, supporting the Underspecification Approach. The contrast between Wittenberg and Piñango's (2011) results and our findings suggests that our L2 learners primarily responded to the presence or absence of a resource-intensive nominalization rather than systematically distinguishing between the two similar construction types, LVCs and FPs, as the German native speakers did.

References

- Barking, Marie, Maria Mos, and Ad Backus. 2025. Investigating Language Transfer from a Usage-Based Perspective. *International Journal of Bilingualism* 29 (2). 415–438. <https://doi.org/10.1177/13670069231175629>.
- Basarić, Mirjana, and Teodor Petrič. 2015. Funktionsverbgefüge als Einheiten im mentalen Lexikon. *Jezikoslovlje* 16 (2-3): 169–186.
- Crawley, Michael J. 2013. *The R Book*. Wiley & Sons: Chichester.
- Cvetko Orešnik, Varja, and Orešnik, Janez. 2021. Prispevek naravne skladnje k določanju slogovnih razločkov. *Linguistica* 61(2): 15–29. <https://doi.org/10.4312/linguistica.61.2.15-29>.
- Dressler, Wolfgang U, Wolfgang Ullrich Wurzel, Oswald Panagl, and Willi Mayerthaler. 1987. *Leitmotifs in Natural Morphology*. Benjamins: Amsterdam, New York.
- Institut für Deutsche Sprache, Forschungsstelle Freiburg. 1974. *Heutiges Deutsch: Texte Gesprochener Deutscher Standardsprache II*. München/Düsseldorf: Hueber Verlag.
- Koch, Peter, and Wulf Oesterreicher. 1990. *Gesprochene Sprache in der Romania: Französisch, Italienisch, Spanisch*. Tübingen: Niemeyer.
- Orešnik, Janez. 2008. Natural Syntax: English Relative Clauses1. *Poznań Studies in Contemporary Linguistics* 44 (1): 61–101.
- Orešnik, Janez, Andrej Snedec, Karmen Teržan, and Frančiška Trobevšek-Drobnak. 1990. Introduction to the Subsequent Three Papers in the Present Volume. *Linguistica* 30 (1): 5–12.
- Pavlić, Ivo. 1971. *Statistička teorija i primjena*. Tehnička knjiga: Zagreb.

- Petrič, Teodor. 1993. Funktionsverbgefüge und Nominalisierungsverbgefüge im Deutschen als verstärkte syntaktische Konstruktionen im Vergleich zu stammgleichen Verben und Adjektiven. *Papiere Zur Linguistik* 49 (2): 125–150.
- – . 1994. Zu einigen strukturellen Eigenschaften von Nominalisierungen im Deutschen. *Linguistica* 34 (1): 181–197.
- – – . 1998. Korrelationen zwischen merkmalthaften oder merkmallösen syntaktischen Varianten und verschiedenen Graden der Textsortenkomplexität. *Linguistica* 38 (1): 201–224. <https://doi.org/10.4312/linguistica.38.1.201-224>.
- Polenz, Peter von. 1988. *Deutsche Satzsemantik: Grundbegriffe des Zwischen-den-Zeilen-Lesens*. Berlin, New York: De Gruyter.
- So, Man-Seob. 1991. *Die deutschen Funktionsverbgefüge in ihrer Entwicklung vom 17. Jahrhundert bis zur Gegenwart*. Trier: Wissenschaftlicher Verlag.
- Sprenger, Simone A, Willem JM Levelt, and Gerard Kempen. 2006. Lexical Access During the Production of Idiomatic Phrases. *Journal of Memory and Language* 54 (2): 161–184.
- Storrer, Angelika. 2007. Corpus-Based Investigations on German Support Verb Constructions. *Idioms and Collocations: Corpus-Based Linguistic and Lexicographic Studies*, 164–187.
- Wittenberg, Eva, and Maria Mercedes Piñango. 2011. Processing Light Verb Constructions. *The Mental Lexicon* 6 (3): 393–413.