Entrenched Lexical Patterns: the Russian Construction

в том-то и весь N

Dmitrij Dobrovol’skij*; Ludmila Pöppelb

*a Russian Academy of Sciences, 18/2 ul. Volkhonka, Moscow, 119019, Russia
b Stockholm University, 10 C Universitetsvägen, Stockholm, SE-10691, Sweden

Abstract

The primary theoretical goal of the present study is to verify a hypothesis that fixedness of word combinations is not necessarily connected with non-compositionality. Many constructions formed in accordance with the rules governing the co-occurrence of their elements can nevertheless be retained in memory as separate units. Using large text corpora for the empirical data we are going to analyze the type of construction в том-то и весь N (that’s /just the whole N) which is realized in a wide variety of tokens and demonstrate that some tokens of the construction can be so frequent that they can be considered to be cognitively entrenched units and are preserved in memory as separate units of the language. Such units should be described as separate items of the lexicon. The practical task of the investigation is to refine our notions about the structural peculiarities of в том-то и весь N (that’s /just/ the whole N) and its variant в том-то и весь и N (that’s /just/ the whole N). We are also going to identify some regularities of distribution of fillers of the slot N in both variants of the construction and formulate corresponding rules for such distribution.

Keywords: Cognitive entrenchment; corpus; fixed expressions; constructions.

1. Introduction

A number of related constructions in Russian have similar semantics. Such, for example, are в том-то и N (that’s /just/ the N/ the N is), то-то и N (that’s /just/ the N), в том-то и весь N (that’s /just/ the whole N). They are all very close in meaning, which can be described as ‘what the speaker determines to be the most important element in the interlocutor’s utterance, that which is crucial to an adequate understanding of a given situation.’ These constructions can be used both as a separate utterance and as a matrix clause in sentences of the type в том-то и N, что P (that’s just the N/the N is that P); то-то и N, что P (that’s just the N/the N is that
P): в том-то и весь N, что P (the whole N is that P).

(1) – В том-то и дело, что паники никакой нет! (RNC, 2011)
“The point is that there is no panic!”b
(2) – А кто-нибудь, кроме вас, видел этого консультанта? – То-то и беда, что только я и Берлиоз (RNC).
“And did anyone besides you see this consultant?”
(3) Действительно, проследите иной, даже вовсе и не такой яркий на первый взгляд факт действительной жизни, – и если только вы в силах и имеете глаз, то найдете в нем глубину, какой нет у Шекспира. Но ведь в том-то и весь вопрос: на чей глаз и кто в силах? (RNC)
In truth, if you investigate some fact of real life—even one that at first glance is not so vivid—you’ll find it in, if you have the capacity and the vision, a depth that you won’t find even in Shakespeare. But here, you see, is the whole point: whose vision and whose capacity? (RNC, 1993)

The semantics of these constructions remains basically the same both when used as a separate utterance and when they serve as a main clause introducing an object clause with что (that). Since we have treated в том-то и N and то-то и N (that’s /just/ the N/the N is) in previous publications devoted to the semantics of Russian constructions (Dobrovolskij & Pöppel, 2015a; Dobrovolskij & Pöppel, 2015b), here we will consider в том-то и весь N (that’s /just the whole N).

2. Aims and goals

The theoretical objective of the present article is to test a hypothesis on the nature of lexical co-occurrence. It has traditionally been thought that word groups can be either free word combinations or fixed expressions. The former are compositional; i.e., their overall meaning is derived from the meaning of their constituent parts, and they are generated by speakers according to the rules of semantic combinability. Fixed expressions are non-compositional; i.e., their meaning is not derived, or not fully derived, from the meaning of their constituent parts. With the appearance of large text corpora and the development of new approaches in linguistics such as Construction Grammar (CxG), it has become possible to study the frequency and hence the cognitive entrenchment of word combinations by empirical methods, particularly quantitative analysis.c We hypothesize that fixedness is not necessarily connected with non-compositionality. Many constructions formed in accordance with the rules governing the co-occurrence of their elements can nevertheless be retained in memory as separate units.

Relative to our materials, this means that the type of construction в том-то и весь N (that’s /just the whole N) is realized in a wide variety of tokens such as том-то и весь вопрос (that’s the whole question), в том-то и вся беда (that’s the whole trouble), в том-то и вся штука (that’s the whole point), в том-то и вся проблема (that’s the whole problem), в том-то и весь ужас (that’s the whole horror), etc. Yet certain tokens of в том-то и весь N (that’s /just the whole N) can be so frequent that there is reason to consider them to be cognitively entrenched units and to describe them as separate items of the lexicon and not only as tokens of this construction. Testing this hypothesis is the primary theoretical goal of the present investigation.

The study will also attempt to refine our notions about the structural peculiarities of в том-то и весь N (that’s /just/ the whole N). Obviously, the constituent parts весь (‘whole’ or ‘all’) and и (lit. ‘and’, here: ‘precisely’ as emphasis) can exchange places within the construction. Depending on the filler of the N slot, one or the other word order is preferred. Thus в том-то все и дело (that’s the whole point) sounds much better than в том-то и все дело (that’s the whole point), and в том-то и весь ужас (that’s the whole horror) is better than в том-то весь и ужас (that’s the whole horror). The question arises as to whether this order is random or is governed by

b The English translation is included for the sake of understanding.
c See for this notion, above all, Langacker 1987, Schmid 2007
some sort of rules. If the empirical materials allow us to identify certain regularities of distribution we will formulate the corresponding rules.

3. Data

The empirical data have been collected from the corpus query system Sketch Engine, subcorpus ruTenTen (2011), which contains more than 18 billion tokens. This made it possible to find all instances of the construction в том-то и весь N (that’s /just/ the whole N) and its alternative variant в том-то весь и N (that’s /just/ the whole N) and obtain statistically representative data. The Russian National Corpus (RNC) containing about 500 million tokens was used for comparison. We searched Sketch Engine for “в том-то и весь <X>”, where X is either a pronoun or a noun in all forms of gender and number. On the basis of the concordance thus obtained, frequency graphs for multiword phrases were constructed.

The search system singles out variants with various punctuation marks and arranges them in different lines. Because we are interested only in the lexical filling of the N slot regardless of the structure of the sentence, we processed our findings manually. Cf. Tables 1 and 2.\textsuperscript{d}

Table 1. В том-то весь и N (that’s /just/ the whole N).

<table>
<thead>
<tr>
<th>Word</th>
<th>Total frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>дело (matter/point)</td>
<td>1450</td>
</tr>
<tr>
<td>штука (thing/point)</td>
<td>77</td>
</tr>
<tr>
<td>беда (trouble)</td>
<td>56</td>
</tr>
<tr>
<td>проблема (problem)</td>
<td>46</td>
</tr>
<tr>
<td>фокус (trick)</td>
<td>26</td>
</tr>
<tr>
<td>прелесть (charm)</td>
<td>24</td>
</tr>
<tr>
<td>фишка (thing/point)</td>
<td>22</td>
</tr>
<tr>
<td>соль (&quot;salt&quot;/point)</td>
<td>20</td>
</tr>
<tr>
<td>загвоздка (hitch/trouble)</td>
<td>17</td>
</tr>
<tr>
<td>суть (essence)</td>
<td>16</td>
</tr>
</tbody>
</table>

Table 2. В том-то весь N (that’s /just/ the whole N)

<table>
<thead>
<tr>
<th>Word</th>
<th>Total frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>дело (matter/point)</td>
<td>90</td>
</tr>
<tr>
<td>фокус (trick)</td>
<td>34</td>
</tr>
<tr>
<td>беда (trouble)</td>
<td>30</td>
</tr>
<tr>
<td>прелесть (charm)</td>
<td>24</td>
</tr>
<tr>
<td>вопрос (question)</td>
<td>21</td>
</tr>
<tr>
<td>проблема (problem)</td>
<td>21</td>
</tr>
</tbody>
</table>

\textsuperscript{d} The tables show only 10 most frequent occurrences.
We have found totally 1850 examples with в том-то весь и N (that’s /just/ the whole N) and 438 with в том-то и весь N (that’s /just/ the whole N). The results showed that in the construction в том-то весь и N (that’s /just/ the whole N) the N slot is most frequently filled with the following: дело (matter/point), (1450), штука (thing/point) (77), беда (trouble) (56), проблема (problem) (46), фокус (trick) (26), прелесть (charm) (24), фишка (point/thing) (22), соль (“salt”/point) (78), загвоздка (hitch/trouble) (17), суть (essence) (16). In the construction в том-то и весь N (that’s /just/ the whole N) the most frequent fillers are: дело (matter/point) (90), фокус (trick) (34), беда (trouble) (30), прелесть (charm) (24), вопрос (question) (21), проблема (problem) (21), ужас (horror) (19), суть (essence) (18), штука (thing/point) (18), соль (“salt”/point) (16), смысл (sense/meaning) (15), загвоздка (hitch/trouble) (13), фишка (point/thing). The search of the RNC yielded the following results: в том-то весь и N (that’s /just/ the whole N): 66; в том-то и весь N: 29 (that’s /just/ the whole N) (22.05.2015). In the construction в том-то весь и N (that’s /just/ the whole N) the most frequent filler by a huge margin is дело (matter/point) (52), followed by штука (thing/point) (7). The other hits are sporadic: соль (“salt”/point) (2), фокус (trick) (1), секрет (secret) (1), беда (trouble) (1), загвоздка (hitch/trouble) (1), несчастье (misfortune) (1). The picture for в том-то и весь N (that’s /just/ the whole N) is different – all hits are isolated and the spread statistically irrelevant: ужас (horror) (5), дело (matter/point) (4), фокус (trick) (4), штука (thing/point) (3), суть (essence) (2), (моя) беда ((my) trouble) (2), вопрос (question) (2), несчастье (misfortune) (1), парадокс (paradox) (1), соль (“salt”/point) (1), хитрость (trick/catch) (1), разница (difference) (1), его трагедия (his tragedy) (1), ваша просьба (your request) (1).

4. Discussion

Our working hypothesis was the assumption that in many constructions in which the filling of an open slot is semantically fairly predictable, there are clear preferences in the choice of specific words to fill the slot. On the one hand, such constructions are semantically regular, while on the other, certain of their instances are fixed expressions. Analysis of the empirical data collected from the corpus query system Sketch Engine showed that the structure of constructions of this type is even more complex. The filler of the N slot is always an abstract noun pointing to a serious, problematic factor in the situation. All concrete nouns occurring in this position are used metonymically or metaphorically in an abstract sense. In the в том-то весь и N (that’s /just/ the whole N) group there is a clear nucleus – the construction в том-то все и дело (that’s the whole point) (1450 hits; 0.1 per million). Cf. Fig. 1. We also obtained a long list of isolated hits, e.g., в том-то весь и страх, секрет, парадокс, шик (that’s the whole horror, secret, paradox, chic), в том-то вся и ситуация, интрига, важность, загвоздка (that’s the whole situation, intrigue, importance, riddle), which confirms our hypothesis about the quantitative spread between entrenched fillers and other N. Between the nucleus and the sporadic occurrences there is a gray zone ranging from 77 to 2 hits: e.g. в том-то вся и штука (that’s the whole thing) (77), беда (trouble) (56), проблема (problem) (46); в том-то весь и фокус (that’s the whole trick) (26), в том-то вся и прелесть (that’s the whole charm) (24), фишка (thing/point) (22), соль (“salt”/point) (20). The picture for the variant of this construction в том-то и весь N (that’s /just/ the whole N) is similar with the difference that в том-то и все дело (that’s the whole point) occurs only 90 times. If we disregard instances of в том-то все и дело (that’s the whole point) and в том-то и все дело (that’s the whole point), the occurrence of both variants is approximately the same. Analysis of the RNC materials generally confirmed the findings from ruTenTen. This
is especially obvious with respect to the frequency of в том-то все и дело (that’s the whole point) (52 out of a total of 66 hits).

Analysis of the empirical data allows us to identify the following structural features of the construction в том-то и весь N (that’s /just/ the whole N. One important factor in filler selection is whether it begins with a vowel or a consonant. If a consonant, no preference is shown to either variant, cf. cn. в том-то весь и фокус vs. в том-то и весь фокус (that’s the whole trick); в том-то вся и беда vs. в том-то и вся беда (that’s the whole problem). If, however, the filler begins with a vowel, the variant в том-то и весь N (that’s /just/ the whole N) is preferred. In the RuTenTen corpus, for the variant в том-то и весь N (that’s /just/ the whole N) 11 types of fillers were found in a total of 35 tokens beginning in a vowel, whereas for the variant в том-то весь и N (that’s /just/ the whole N) there were 5 types and 12 tokens.

5. Conclusion

The investigation confirmed our basic theoretical hypothesis on the nature of lexical co-occurrence, namely that fixedness is not necessarily connected with non-compositionality. We have shown that в том-то все и дело (that’s /just/ the whole point) is many times more frequent than all other instances of the construction. This gives reason to assume that it is cognitively entrenched and is preserved in memory as a separate unit of the language.
Acknowledgements

This paper is based on work supported by the RFFI (Russian Foundation for Basic Research) under Grant 13-06-00403 and by the RGNF (Russian Foundation for Humanities) under Grant 15-04-00507 and 15-04-12018.

References


