

Grammatical collocations in English essays written by Czech secondary school students

Věra Sládková
Masaryk University, Czech Republic

This paper presents the results of a study investigating the frequency and accuracy of five types of grammatical collocations (G8E to G8I) according to the BBI Dictionary classification of collocations (Benson et al., 1986) in a Czech learner corpus. CZEMATELC 2017 (44,044 tokens; 2,765 types) is an English language learner corpus consisting of 390 essays from the written part of the national school-leaving examination which leads to a certificate of secondary education in the Czech Republic. The above-mentioned types of collocations were chosen because they correspond to the 'Can Do Statements' of *English Grammar Profile* (EGP), which describes the skills that EFL learners commonly master at CEFR level B1. Language production at this level is also required for successful certification.

Since Czech has only one infinitive form, which is marked by a '-t ending', and since it lacks a gerund construction, most of the verbs in the selected collocations enter different colligation patterns in Czech, that is, they can have a variety of complementation patterns. The exact equivalents of some verbs (i.e. *hope, suggest*) can be complemented only by clauses, and although the exact equivalents of other verbs (i.e. *would like, want, need*) can be complemented by an infinitive, they require the indirect object to become the subject of the dependent clause in case of ditransitive complementation. Creating equivalent sentences in English often requires the use of more complex grammar to avoid errors and confusion. Consequently, Czech students need support in avoiding the use of mother-tongue verb patterns and relying on the regularity of G8E to G8I grammatical collocations in English if they want to express themselves clearly and be understood. However, the majority of English course books do not provide enough examples and exercises to internalize these colligation patterns, and exercises practising grammatical collocations G8H (verb + obj + to-inf) and G8I (verb + obj + inf without 'to') are generally scarce. In this situation, it is important to know exactly which similarities and differences between English and Czech promote or hinder Czech learners' language production, because grammatical collocations (G8E to G8I) are not only a common source of errors, but a key reason for avoidance, which prevents the students from expressing deeper thoughts by forming extended sentences.

Accordingly, the aims of this study were as follows:

1. To determine to what extent Czech students attempted to use the investigated grammatical collocations in their essays.
2. To identify the accurate and inaccurate uses of the selected collocations in light of English varieties (Deshors & Gries, 2016).
3. To specify which particular correctly used grammatical collocations and error patterns are similar to colligation patterns in Czech.
4. To ascertain which of the studied grammatical collocations with particular verbs need to be given a greater emphasis in teaching in order to increase the chances of more learners being able to use verb patterns at B1 level according to the CEFR as expected by EGP.

The analysis was approached from the perspective of Pattern Grammar (Hunston & Francis, 2000) since it focused on particular verbs and investigated all the grammar patterns in which they appeared in the learner corpus. The accuracy of grammatical collocations and the frequency of the grammar patterns were established by manual counting with the help of an

online corpus analytical tool, AntConc 3.4.4w (Anthony, 2014) since the analysis of a raw corpus was deemed more reliable due to interlanguage inconsistencies. Frequency information from the Czech National Corpus (Křen et al. 2015) was used for the comparison of colligation patterns.

The frequency counts of the studied collocations reveal that they can be divided into two groups with similar features. The grammatical collocations G8E (verb + to-inf) (n=247) and G8F (verb + inf without 'to') (n=722) are fairly frequent with a high incidence of accurate uses. Errors that are potentially caused by mother-tongue influence (e.g. omitted lexical verbs after modals, inappropriately used infinitives with or without 'to' or incorrectly inserted objects) appear to be much less frequent than errors which can be loosely attributed to overgeneralization (e.g. verbs followed by gerunds and past tense forms). Grammatical collocations G8G (verb + v-ing) (n=37), G8H (verb + obj + to-inf) (n=9) and G8I (verb + obj + inf without 'to') (n=28) are rather scarce. Moreover, their raw frequencies in the studied corpus rely predominantly on colligation patterns of a few particular verbs. This suggests that these collocations should be taught explicitly with most verbs if the learners are to acquire colligations at CEFR level B1 and use them more frequently.

The results indicate that the learners attempt to produce more complex verb patterns with higher accuracy in well-rehearsed phrases and with some verbs which form colligation patterns which are more easily perceived as similar to those in Czech.

References

- Anthony, L. (2014) *AntConc*. Tokio: Waseda University. Retrieved from: www.laurenceanthony.net/software/anconc
- Benson, M., Benson, E., & Ilson, R. (1986) *The BBI Dictionary of English Word Combinations*. Amsterdam: John Benjamins.
- English Grammar Profile* (2015) Cambridge: Cambridge University Press. Retrieved from <http://englishprofile.org/english-grammar-profile/egp-online>
- Deshors, S. C., & Gries, S. T. (2016) Profiling verb complementation constructions across New Englishes. *International Journal of Corpus Linguistics*, 21(2), 192–218.
- Hunston, S. & Francis, G. (2000) *Pattern Grammar: A corpus-driven approach to the lexical grammar of English*. Amsterdam: John Benjamins.
- Křen, M., Cvrček, V., Čapka, T., Čermáková, A., Hnátková, M., Chlumská, L., Jelínek, T., Kovářiková, D., Petkevič, V., Procházka, P., Skoumalová, H., Škrabal, M., Truneček, P., Vondříčka, P., & Zasina, A. (2015) *SYN2015: reprezentativní korpus psané češtiny*. Praha: Ústav Českého národního korpusu. FFUK. Retrieved from: <http://www.korpus.cz>