

Balkan focus particles in Bulgarian: *bash* and *taman*

Ekaterina Tarpomanova

Sofia University Saint Kliment Ohridski
Sofia, Bulgaria

katya@slav.uni-sofia.bg

Abstract

This study investigates the Bulgarian focus particles *bash* and *taman*, two lexical items of the shared Balkan lexicon, with regard to their role in information structure. Using a corpus-based approach, the paper analyzes their function as particularizers that identify a focused constituent within a set of alternatives. The analysis is based on data from the Bulgarian web corpus CLASSLA-web.bg 2.0. The findings indicate that both particles are usually positioned adjacent to the focused constituent, most often in preposed position, though variation is attested. Overall, the study highlights both the shared and divergent properties of *bash* and *taman*, showing how their etymological origins and semantic development influence their present-day behavior as focus particles in Bulgarian.

Keywords: information structure, focus particles, corpus linguistics

1 Introduction

The present study examines two Bulgarian lexical items that belong to the shared Balkan lexicon, with particular emphasis on their potential function as focus particles. Although morpho-syntactic similarities are considered key features of the Balkan Sprachbund as they prove the deep convergence process that took place between the Balkan languages, the words they share are not with lesser importance in revealing the cultural and spiritual proximity of the Balkan peoples. The lexical items that are object of this study are not just shared loanwords in the Balkan languages as they play an important communicative role as focus particles interacting with the sentence focus and thus contributing for the correct interpretation of the information structure in discourse.

The study adopts a corpus-based approach to the analysis of the two focus particles in Bulgarian, with the aim of establishing their role in the

information structure of the sentence. The theoretical account of focus particles follows the framework proposed by Ekkehard König (1991), while the analysis of Bulgarian focus particles draws on the model developed by Ruselina Nitsolova (2001), as well as two studies by Yovka Tisheva and Marina Dzhonova (2003, 2006) on the Bulgarian focus particle *samo* ‘only’. The corpus analysis has a twofold objective: on the one hand, to provide quantitative data on the co-occurrence of focus particles with different word classes; on the other hand, to analyze – on the basis of authentic corpus examples – the types of sentential constituents that may fall within their scope. To this end, the CLASSLA-web.bg 2.0 is employed – a corpus of Bulgarian web, comprising 5,986,954,232 tokens from a wide range of genres and topics and enriched with linguistic annotation (Pungeršek et al., 2026).

2 Information structure

Information structure (Halliday, 1967) or information packaging (Chafe, 1976) is a central concept in contemporary linguistics, concerned with how speakers organize and present information within an utterance in relation to the communicative context and the common ground of the interlocutors. Rather than focusing on grammatical correctness or semantic content, information structure examines how meaning is shaped by the distribution of given and new information, the management of discourse, and the speaker’s assumptions about the hearer’s knowledge (Halliday, 1967; Lambrecht, 1994; Krifka, 2008). In recent decades, information structure has become increasingly relevant not only in theoretical linguistics and discourse analysis but also in related fields such as computational linguistics. Understanding how information is structured contributes to models of language processing and discourse representation and may

improve natural language generation systems by enhancing the analysis of spoken and written discourse.

Information structure has been analysed within different theoretical frameworks, but at its core, it addresses the ways in which speakers highlight certain elements of a sentence while backgrounding others. Concepts such as *topic* (or *topic – comment*) and *focus* (or *ground – focus*) play a crucial role in structuring discourse. For instance, the difference between *John bought the book* and *John bought the book* lies not in propositional content but in how information is packaged and interpreted.

Cross-linguistically, languages employ diverse strategies to encode information structure. While languages like English rely on intonation and stress, others, such as Japanese, use dedicated particles to mark topic and focus explicitly (Nakamura, 2006). Balkan languages have been classified as discourse-prominent languages as they code the communicative units not only by prosody, but also using syntactic devices such as word order, clitic doubling, subject position, etc. (Krapova, 2008: 181). Many languages use lexical tools to manage information structure, among which focus particles serve as a key device for marking sentence focus.

3 Focus and focus particles

Focus is an important notion in theories of information structure, referring to the part of an utterance that conveys the most informative, non-presupposed, or contrastive content. Broadly speaking, focus identifies what the speaker presents as new, informative, dominant, or contrary-to-expectations in relation to the part of the utterance considered as ground (Vallduví & Engdahl, 1996: 3). It plays a central role in how speakers manage the flow of information and guide the listener's attention to what is communicatively important. Focus has been defined as “that part of the utterance that answers the question under discussion” (Clopper & Tonhauser, 2011:1). Following the strategy of “question-answer congruence test”, questions are used as an important tool to identify the focused constituent (Kroeger, 2017:4). Focus is not a uniform phenomenon but encompasses several subtypes, including information focus and contrastive (identificational) focus. Information focus introduces new information into the discourse,

typically answering implicit or explicit questions (e.g., *What did John buy? – A car*), whereas, contrastive focus highlights alternatives and often implies a correction or opposition (e.g., *John bought a car; not a bike*) (Kiss, 1998; Krifka, 2008).

Focus particles are operators on the semantic structure of a sentence that mark the contrastive focus – an alternative among a set of alternatives that can be implemented in the case. The speaker only has a presupposition for these alternatives before he/she has received the information about the actual event that contrasts with his/her presupposition (Nitsolova, 2008: 479). Focus particles can be classified into several types based on their semantic contribution and pragmatic function. Two main types are to be outlined: inclusive (additive) which include some of the alternatives as possible values of the variable of their scope (*also, too*), and exclusive (restrictive), which imply that none of the alternatives satisfies the respective open sentence (*only*) (König, 1991: 33). König (1991: 122) notices a third type – particles of emphatic assessment of identity (*exactly, precisely*). Following Quirk's classification (Quirk et al., 1972), Nitsolova adopts the term particularizers for the same subclass of focus particles in Bulgarian – *točno* ‘exactly, precisely’, *tákmo* ‘exactly, precisely’, *imenno* ‘namely’, defining them as cases in which the predicate connects with some part of the alternatives (Nitsolova, 2008: 479).

4 *Bash* and *taman* as lexical items of the common Balkan lexicon

The lexical items that are object of analysis, *bash* and *taman*, both meaning ‘exactly, precisely’, belong to the shared Balkan lexicon which contains words that occur in the languages of the Balkan Sprachbund regardless of their origin (Asenova and Detrez, 2021). Similarly to other words of the Balkan lexicon, *bash* and *taman* are also found in Bosnian-Croatian-Montenegrin-Serbian (BCMS).

Among languages of the Balkans, *bash* occurs in Turkish, BCMS, Balkan Slavic (Bulgarian and Macedonian), Slovene and Albanian; it has regional use in Daco-Romanian. In some of the Balkan languages, including Bulgarian, *bash* is described in lexicographic resources as two separate lexical items: as an adjective (and substantive) with the meaning ‘main’ and as an adverb or particle with the meaning ‘exactly,

precisely'. The adjective has been loaned from Turkish *baş* 'head, leader; beginning, basis', but the origin of the adverb/particle is disputable. In some sources, mainly Bulgarian and Serbian, it is stated that the adverb/particle has derived from the adjective, i.e. they are etymologically related and both have Turkish origin (Asenova and Detrez 2021: 304; BER, 1971: 37; ERSJ, 2006: 276), whereas Croatian dictionaries claim Slavic origin of the adverb/particle – *baš* has derived from Proto-Slavic **bъšъ* 'entirely, quite', a petrified form of the noun *bъšъ* 'hit, blow', cognate with the verbs *buhnuti*, *bušiti* 'to hit' (ERHJ 2016: 45-46).

In contemporary Bulgarian *bash* occurs as an uninflected adjective with the meaning 'main, the best, top' (1a) and as focus particle with the meaning 'exactly, precisely, just' (1b).

(1a) *Izbraha **bash** maystorite na vinata v Parvomay i Dalbok izvor.* 'The **top** master winemakers were selected in Parvomay and Dalbok Izvor.'

(1b) – *Krastosa edva li e pravilnata дума, otche. – **Bash** pravilnata si e.* '– Crossed is hardly the right word, Father. – It's **exactly** the right one.'¹

For Serbian, Mišković defines two functions of the focus particle *baš*: specificatory when the particle specifies the focus and has a temporal, locative or deictic expression in its scope (2a), and emphatic when it focuses on certain process, state or quality (2b):

(2a) *Knjiga je **baš** na stolu.* 'The book is **right** on the table.'

(2b) ***Baš** peva.* 'He/she is really singing (beautifully).' (Miskovic, 2001)

As compared to Serbian (and BCMS in general), *bash* in Bulgarian has most often a specifying function, whereas the emphatic function is rather rare. The difference between BCMS and Bulgarian can be illustrated with the sentences below which are translations of *The Ugly Duckling* by Hans Christian Andersen:

(3a) Serb. „**Baš** si ružan!“ , *rekle su mu patke.* „*Ali nama to ne smeta, samo da se ne ženiš nekom od nas!*“ (translation by Stanislav Vinaver)

(3b) Bulg. – *Ti si **mnogo** grozno! – kazaha divite patitsi. – Ala tui ne ni vliza v rabotata, stiga samo*

da ne ti hrumne da se ozhenish za nyakoya oot nasheto semeystvo. (translation by Svetoslav Minkov)

(3c) Eng. “You are **terribly** ugly,” they told him, “but that’s nothing to us so long as you don’t marry into our family.” (translation by Jean Hersholt)

Both Bulgarian and English have chosen an adverb to intensify the quality expressed by the adjective – *mnogo* 'very' and *terribly*, respectively, whereas for Serbian *bash* seems to be a natural choice in this context.

When used independently, *bash* expresses confirmation or agreement with a previous statement (4).

(4) – *Kaza, che tezi neshta sa na Sarina? – **Bash**.* 'You said these are Sarina's things. – **Right**.'

The origin of *taman* is clear – in the Balkan languages, it is loaned from Ottoman Turkish (*tamam* 'whole, complete, accurate; just, exactly, precisely'), where it penetrated from Arabic (*tamām* 'full, entire, complete; completeness, wholeness, perfection') (Asenova and Detrez 2021: 313). It occurs (with different frequency) in almost all contemporary languages and dialects of the Balkans: Turkish, Balkan Slavic (Bulgarian and Macedonian), Balkan Romance (Daco-Romanian, Aromanian and Megleno-Romanian), Albanian, Greek, BCMS. In modern Arabic and Turkish, it has undergone strong pragmaticisation and its initial lexical meaning related to the notion of completeness has developed into a pragmatic one (Ruhi 2013; Al-Hanaktah and Hamdan 2024), with closest translation equivalents in English discourse markers such as *ok*, *right*, *fine*. In the Balkan languages, pragmatic development is not observed. The word *taman* is defined there as an adverb meaning 'exactly'.

In contemporary Bulgarian, *taman* is a colloquial word with several main uses: as an adverb of manner, indicating that something (clothes, shoes) is the right size (5a) or otherwise suitable (5b) in the speaker's view; as an adverb of time, indicating that the action is carried out immediately before or at the moment of speaking (5c), or in sentences in which the situation in the first clause, introduced by *taman*, takes place

¹ All Bulgarian examples are drawn from the CLASSLA-web.bg 2.0 corpus and are transliterated into Latin script. The lexical items under analysis are presented in bold both in the original and in the English translation, where possible, and the constituent falling within their scope is underlined. As noted by Tisheva and Dzhonova (2003: 58), the scope of

focus particles and sentence focus usually coincide, but not always.

immediately before the situation in the second one (5d); as an adverb expressing logical relations (according to the classification in Nitsolova 2008: 450), functioning as a link between two utterances, confirming the preceding statement and adding a new argument in its support or a positive development of the situation (5e); and as a focus particle (5f).

(5a) *Roklyata ti e **taman**.* ‘Your dress **fits perfectly**.’

(5b) *Na men lichno nastroyka 23 na klimatik si mi e **taman**...* ‘Personally, setting 23 on the air conditioner suits me **just fine**...’

(5c) ***Taman** vlizahme v Balgariya.* ‘We were **just** entering Bulgaria.’

(5d) ***Taman** reshih da zatvorya „krazhbite“ i ti gi otvori nanovo.* ‘I had **just** decided to stop the “thefts,” and you opened them up again.’

(5e) *Idi bre, **taman** shte razberesh ot parva raka shto smartnostta e niska.* ‘Come on, go there, **then** you’ll understand firsthand why the mortality rate is low.’

(5f) *... osvobozhdenieto ot sorosoidnata sekta tryabva da zapochne **taman** ot Polsha!* ‘... liberation from the Sorosoid sect must begin **precisely with Poland**!’

Taman is rarely used independently; in such cases it expresses the speaker’s approval that the circumstances in a given situation are good and meet expectations.

(6) – *Arzhentina i Iran poiskaha da se prisaedinyat kam BRIKS. – Da, da, **taman**.* ‘– Argentina and Iran have requested to join BRICS. – Yeah, yeah, **perfect**.’

The review of the etymology and semantic development of *bash* and *taman* shows that, although they have different origins and functions, their usage in contemporary Bulgarian is similar when they function as focus particles.

5 *Bash* and *taman* as particularizers

Both Balkan lexical items in Bulgarian are used as focus particles of the type of particularizers, i.e., they associate with a given alternative within a set of alternatives. Both items are characteristic of informal register and, in contrast to their primary Bulgarian synonyms *tochno* and *takmo* ‘exactly, precisely’, they are expressive and stylistically marked. *Bash* and *taman* occur significantly less frequently compared to their stylistically neutral synonyms: *tochno* has 2,825,380 occurrences, *takmo* – 217,070, *bash* – 66,576, and *taman* (in its

three phonetic variants – *taman*, *tamăn*, and *tamam*) – 21,798. As has been shown, the two particles have different origins and inherit the semantic properties of their source lexemes: *bash* is linked to the concept of centrality or primacy, whereas *taman* is associated with the idea of completeness or perfection.

The present study examines the way in which they participate in the information structure of the utterance by identifying the focus constituent of the sentence, and, additionally, addresses the possible differences in their semantics as focus particles. This perspective directs attention primarily to their syntactic properties – namely, their position within the sentence and their scope, and secondly to their interpretation depending on the context. It is well known that the syntactic behavior of focus particles varies across languages; moreover, within a single language, different focus particles may exhibit distinct syntactic patterns. Focus particles can generally combine with all major constituent types – NP, VP, AP, AdvP, PP, and CP (Herrmann 2014: 185).

In Bulgarian, *bash* and *taman* occur in sentence-initial, sentence-final, and medial positions. They are typically preposed and in contact position with respect to the constituent they associate with (7a); more rarely, they are postposed (7b) or occur at a distance from that constituent (7c).

(7a) *Boyko Borisov **taman tebe** chaka, da mu nosish haber ot Varna...* ‘Boyko Borisov is waiting **exactly for you** to bring him news from Varna.’

(7b) *Na tebe **bash** da ti obyasnyavam...* ‘To you, of all people (lit. exactly), I should explain...’

(7c) *Tiho sya, **taman** Steven Pinker se opitva da podkovava britantsite, che nyamalo takova neshto kato chist kralski algliyski.* ‘Quiet now, Steven Pinker **is just trying** to convince the British that there is no such thing as pure Queen’s English.’

Therefore, most often, though not always, *bash* and *taman* identify the first right constituent as the focus. Table 1 presents the co-occurrence of the two focus particles and the first right constituent, which may belong to different word classes: noun (N), pronoun (Pron), adjective (A), adverb (Adv), verb (V), numeral (Num), conjunction (Conj), or preposition (P). The results can to some extent guide us in determining the scope of the focus particles, but further analysis shows that they can be misleading.

	<i>bash</i>	<i>taman</i>
N	17,108	539
Pron	10,574	2,608
A	4,613	150
(A+N)	(2,140)	(79)
Num	506	1,053
Adv	2,224	423
V	1,231	3,364
P	12,688	1,575
Conj	534	445

Table 1: Co-occurrence of *bash* and *taman* with different word classes.

5.1 *Bash* and *taman* + NP

The first option to be checked is whether the two focus particles can have in their scope a noun phrase with different structure – a common noun, a proper noun, a pronoun, and a noun modified by an adjective. In the combination of the particles with a noun, the overwhelming predominance of occurrences of *bash* is immediately noticeable. In fact, in a large number of cases *bash* is used not as a focus particle but as an adjective (e.g. *bash opozitsionerite* ‘the real oppositionists’). The expression *bash maystor* – the title of a well-known Bulgarian film that has come to be used ironically to refer to a person who is incompetent at their job – alone occurs 1,128 times. However, *bash* can have a scope over an NP with different type of structure – rather rarely if the NP consists of a noun (8a) and much more frequently if it consists of a proper noun (8b), of a pronoun (8c) or an adjective and a noun (8d).

(8a) *Sanchez si e bash krilo. I v Barselona si igraeshe krilo.* ‘Sanchez is **exactly** a winger. He also played as a winger at Barcelona.’

(8b) *Taka che bash Hari e vinoven...* ‘So Harry is the one to blame...’

(8c) *...bash toy ni nabuta da poluchavame samo ruski gaz!* ‘...it was **exactly** him who got us stuck with receiving only Russian gas!’

(8d) *...no faktite sochat, che bash desnite demokrati sa se kultivirali v kachestvoto si na obsluzhvasht personal okolo Todor Zhivkov.* ‘...but the facts suggest that it was **precisely** the right-wing democrats who were cultivated in their role as a kind of service personnel around Todor Zhivkov.’

The review of the results for *taman* shows that, as a focus particle, it much more rarely takes an NP

within its scope, and this is usually a pronoun (9a). In most cases, *taman* is not used as a particularizer, but as an adjective (9b) or adverb (9c). In a number of sentences, although the first right constituent is an NP, as an adverb it modifies a VP or a clause (9d).

(9a) *Makar che ako beshe za byagane prez Dunava, taman toy beshe za tova...* ‘Although, if it had been about escaping across the Danube, he was **just** the right person for it (lit. it was exactly him for it)...’

(9b) *Neshto ne e taman rabotata.* ‘Things aren’t quite **right**.’

(9c) *Ako ne iskate, ne glasuvayte, taman BSP shte izchezne okonchatelno.* ‘If you don’t want to, don’t vote – **that way** BSP will disappear for good.’

(9d) *...da mahnat pone vecheryata, taman izlishni kilogrami shte svalyat.* ‘...they should at least drop dinner – **that way they**’ll lose some extra weight.’

For both *bash* and *taman*, the high frequencies in combination with pronouns are due to the annotation scheme of the corpus: pronominal adverbs (*bash togava* ‘right then’, *bash kogato* ‘exactly when’), as well as reflexive pronouns functioning as derivational or grammatical particles (*taman si trǎgvam* ‘I am just leaving’) or as ethic dative (*bash si e legenda* ‘he is a legend indeed’), are annotated as pronouns. In addition, pronominal clitics cannot serve as the focus of a sentence (*taman go nauchih* ‘I just taught him’).

5.2 *Bash* and *taman* + NumP

When analyzing the possible scopes of *bash* and *taman*, special attention should be paid to numeral phrases (for an analysis of NumP as part of DP, see Cinque 2005), due to the semantic compatibility of the particularizers with numerals. The particularizers take in their scope NumPs with both ordinals (10a, 10b) and cardinals (10c), as well as numeral derivatives (10d). The co-occurrence of *bash* and *taman* with numerals is much more frequent as compared with their use with other quantifiers such as *mnogo* ‘very, many’, *malko* ‘a little’.

(10a) *taka che si smeni iztochnitsite, a Svetlyo si vziat bash 5 bona.* ‘so change your sources, and Svetlyo, they take **exactly** 5 grand.’

(10b) *Mozhem da si predstavim kak shte reagira, ako bash dvete partii-simvoli na statukvoto na prehoda izlyazat na vlast.* ‘We can imagine how he

will react if exactly the two parties that symbolize the status quo of the transition come to power’

(10c) ...*taman* *tretyat den ot tridnevna voyna*. ‘...exactly the third day of the three-day war.’

(10d) ...*ima taman shestkratno uvelichenie na dobivite ot 1981 do 1989 g.* ‘There is exactly a sixfold increase in output from 1981 to 1989.’

In combination with numerals, *taman* occurs more frequently than *bash*, despite the overall frequency of *taman* being approximately three times lower. A plausible explanation lies in the modal nuance associated with *taman*, which is absent in *bash*. When *taman* functions as a particularizer with a NumP in its scope, it additionally conveys the speaker’s subjective stance, typically positive – expressing approval or the fulfillment of prior expectations, and in some cases irony. A distinction between modal and epistemic uses of scalar particles such as *at least* has been discussed in a number of studies (Grosz 2011; Biezma 2013; for the South Slavic languages Stefanov 2026, among others). In the case of the Bulgarian particularizers *bash* and *taman*, the latter exhibits stronger modal properties, likely inherited from its adverbial uses.

5.3 *Bash* and *taman* + AP

In nearly half of the cases, when the first right-hand constituent is an adjective, it forms part of an NP consisting of an adjective and a noun (in Table 1 – A+N). Beyond these instances, the two focus particles interact with adjectival phrases under conditions comparable to those observed with NPs: both particles may take an AP within their scope, although this occurs relatively infrequently and is more common with *bash* than with *taman*.

(11a) *Raykov izglezhda bash idealen za sluchaya*. ‘Raykov seems just perfect for the case.’

(11c) „Pampers“ *peleni ne kupuvam, no „Hagis“ v U. sa si ama taman sashtite kato drugade*. ‘I don’t buy “Pampers” diapers, but “Huggies” in U. are just exactly the same as elsewhere.’

5.4 *Bash* and *taman* + AdvP

Both focus particles may take adverbial phrases within their scope, although *bash* does so more frequently and combines with a wider range of adverb types, including temporal (12a), (12b), locative (12c), and manner adverbs (12d). As observed by Miskovic (2001) for Serbian, when temporal and locative adverbs fall within its

scope, *bash* typically has a specifying function, whereas with manner adverbs it conveys emphasis; the latter use is relatively infrequent in Bulgarian. By contrast, *taman*, as in its interaction with NPs, is more often used as an adverb (12f) than as a particularizer when combined with adverbs. When functioning as a particularizer, it most commonly occurs with temporal adverbs (12e).

(12a) ...*mi bash dnes vashite hora kazaha che ne e agent kato prezidenta...* ‘well, it was exactly today that your people said he is not an agent like the president...’

(12b) ...*taka se omazaha bash predizborno*. ‘so they messed up right before the elections.’

(12c) *Vav videoto mnogo dobre kazvat koy otkade idva i snimkata na kolata na myasto si e bash napreki na idvashtoto semeystvo*. ‘In the video they clearly say who comes from where, and the picture of the car is positioned exactly across the approaching family.’

(12d) *I ti ne mozh da mi kazhesh, che ti e bash ubavo kato kogato beshe po-mlad*. ‘And you can’t tell me it feels exactly as good as when you were younger.’

(12e) *I taman sega shte ti go trusnat v glavata s novi zakoni na ES*. ‘And just now they will dump it on your head with new EU laws.’

(12f) *Prodalzhavay, taman po-skoro shte ni otarvesh ot prisastviето si*. ‘Go on, that way you’ll get rid of your presence sooner.’

5.5 *Bash* and *taman* + VP

With verb phrases, *taman* displays a considerably higher frequency; however, it is almost invariably used as an adverb (13a) rather than as a particularizer. In certain cases, *taman* allows for ambiguity between an adverbial/adjectival and a focus-particle interpretation, with disambiguation relying on prosody in spoken language or broader contextual cues in written discourse (13b). *Bash*, in turn, only rarely takes VPs within its scope. Many apparent instances result from the annotation of participles as verbal forms, although these function as adjectives within NPs (e.g. *bash protestirashtite partii* ‘exactly the protesting parties’). Most commonly, *bash* scopes over the copula, and more rarely over other verbs (13c). A characteristic construction involves the copula combined with a reflexive pronoun in ethical dative, which introduces a subjective stance on the part of the speaker. The predicative complement in such

constructions may be realized by various phrase types, including NP (13d), (13e), AP (13f), NumP (13g), and PP (13h). With both the copula and full verbs, when used as particularizers, *bash* and *taman* select as focused constituent the complement within the VP. In copular sentences, the particularizers move to a sentence-initial position due to prosodic factors in Bulgarian.

(13a) ...vecherniyat prelet **taman** zapochva, chovek tryabva da si otvarya ochite. ‘...the evening migration is **just** beginning; one has to keep their eyes open.’

(13b) **Taman** si e za Kitay. ‘It’s **perfect** for China.’ / **Taman** si e za Kitay. ‘It’s **exactly** for China.’

(13c) A dneshnite arabi **bash** iskat titli, slava, da badat zabelyazani, a ne pari. ‘And what today’s Arabs **really** want is titles, fame, to be noticed, not money.’

(13d) Uff... pāk makedontsite bili s debeli glavi – **bash** sa si balgari! ‘Ugh... and they say Macedonians have thick heads – they are **exactly** Bulgarians!’

(13e) **Bash** si e prerazhdane v dvizhenie. ‘It **really** is reincarnation in motion.’

(13f) **Bash** si e falshiva. ‘It **really** is fake.’

(13g) **Bash** si e 1000MB. ‘It is **exactly** 1000MB.’

(13h) **Bash** si e za tam. ‘It **really** belongs there. (lit. for there)’

5.6 *Bash* and *taman* + PP

Both particles may take prepositional phrases within their scope, with this tendency being particularly pronounced for *bash*. The PPs involved may be headed by a range of prepositions. In the case of *bash*, the highest number of occurrences is found with the highly frequent Bulgarian preposition *na* ‘of, to’, whereas *taman* most commonly combines with *za* ‘for’. This distribution can be attributed to the adjectival semantics of *taman*, which conveys meanings such as ‘suitable (for)’ or ‘perfect (for)’. (14a) Prostitutka nalyetya **bash** na shefa na politsiyata. ‘A prostitute went **straight at the chief of police**.’

(14b) ...ama **bash** v toz sektor nashiyat „aviatsionen“ spetsialist ne otgovarya. ‘...but **in exactly this sector** our “aviation” specialist is not responsible.’

(14c) za da mozhe **bash** za uikenda da ni se stovaryat tezhki gramotevitsi. ‘so that heavy thunderstorms hit us **exactly on the weekend**.’

(14d) ...nisko letyasht voennomorski hidroplan se spuska **taman** kam nego. ‘a low-flying naval seaplane is descending **right toward him**.’

(14e) Nyama osnovaniya Ivo da se pritesnyava ot sada i to **taman** v Sofiya. ‘There are no grounds for Ivo to worry about the court – **especially in Sofia**.’

(14f) ...**taman** kato sherifa ot diviya zapad! ‘...**just** like the sheriff from the Wild West!’

5.7 *Bash* and *taman* + CP

Both *bash* and *taman* may interact with complementizer phrases. However, the results presented in Table 1 should be interpreted with caution, as corpus data involving conjunctions frequently contain cases of misannotated grammatical ambiguity, for instance between conjunctions and prepositions (*bash kato družkata ti* ‘exactly as your friend’) or between conjunctions and particles (*bash pa tam* ‘exactly there’). At the same time, *k*-words (i.e. interrogative and relative pronouns and pronominal adverbs) are included in the results under the category of pronouns.

A comparison between the two particularizers shows that *bash* more frequently takes CPs within its scope than *taman*. Both particles combine with the modal complementizer *da*, which in Bulgarian functions as a substitute for the old Slavic infinitive and has grammaticalized into a preverbal morpheme (Asenova 2002: 151). The sequence *bash da* occurs 237 times, and *taman da* 932 times; however, in these cases *bash* predominantly functions as a particularizer (15a), whereas *taman* is used almost exclusively as an adverb – temporal or expressing logical relation (15b). CPs introduced by the factive complementizer *che* (‘that’) only rarely fall within the scope of *bash* (15c), and there are no such instances with *taman* in the corpus. Both particularizers can take within their scope CPs introduced by *ako* ‘if’ (15d, 15e), but *taman* retains its adverbial meaning in most cases.

(15a) ...**bash** da sa tupali Bayern redovno e malko presileno! Biha gi samo 2 pati. ‘**Exactly that they have beaten Bayern regularly** is a bit exaggerated! They beat them only twice.’

(15b) *I **taman** da si traghnat, Bush kazal: ...*
 ‘And **just** as they were about to leave, Bush said: ...’

(15c) *Ti ne znaesh... hahaha... i samo smenyash nikuviti... hahaha... i taka... **Bash** che ne smenyam nikoga “nikuviti”, a drugi go pravyat ot rusofilska zloba poradi tozi fakt, pisheiki s moya! ‘You don’t know... haha... and you only change nicknames... haha... and so on... **Exactly** that I never change “nicknames”, while others do it out of pro-Russian spite because of this fact, writing with mine!’*

(15d) *Ti ne razbirash – **bash** ako beshe pobedil fashizmat, sega Makedoniya, polovin Gartsiya i tryala Dobrudzha [...] shteshe da e nasha! ‘You don’t understand – exactly if fascism had won, now Macedonia, half of Greece, and all of Dobruja [...] would have been ours!’*

(15e) *Zdravey, kolega, i az ako mozhe da se kandidatiram za stoykata, ponezhe mnogo iskam da si vzema edna, no to tuk mi idvat mnogo skapo i sega si porachah zumeri ot Rusiya, i **taman** ako mi ya dadesh, shte e prefektno. ‘Hello, colleague, if possible, I’d also like to apply for the stand, since I really want to get one, but from here they come out very expensive, and now I’ve ordered buzzers from Russia, and **just** if you give it to me, it will be perfect.’*

6 Conclusion

The analysis of the corpus data show that both *bash* and *taman* can take scope over a wide range of syntactic constituents, including NP, NumP, AP, AdvP, VP, PP, and CP, though with notable differences in frequency and distribution. *Bash* more frequently functions as a focus particle, and does so with all types of phrases except VPs, typically exhibiting a specifying function. Still, when used with NPs disambiguation is always necessary. In contrast, *taman* is more often used as an adverb, which is why it shows a higher frequency with VPs. When functioning as a particularizer, most often with NumPs and PPs, *taman* tends to convey an additional modal nuance, expressing the speaker’s (positive) attitude.

References

Isra'a Al-Hanaktah and Jihad M. Hamdan. 2024. On the Pragmatic Functions of the Idiomatic Expression ‘Tamam’ in Jordanian Arabic.

Journal of Language Teaching and Research, Vol. 15, No. 3.

Petya Asenova. 2002. *Balkansko ezikoznanie*. Veliko Tarnovo: Faber.

Petya Asenova and Raymond Detrez. 2021. *Rechnik na obshtobalkanskata leksika v knizhovnite ezitsi*. Sofia, Universitetsko izdatelstvo “Sv. Kliment Ohridski”.

BER 1971: Vladimir Georgiev, Ivan. Galabov, Yordan Zaimov, Stefan Ilchev. *Balgarski etimologichen rechnik*. Sofia, Izdatelstvo na BAN.

ERHJ 2016: Ranko Matasović, Tijmen Pronk, Dubravka Ivšić, Dunja Brozović Rončević. *Etimološki rječnik hrvatskoga jezika*. 1. svezak. A – Nj. Zagreb, Institut za hrvatski jezik i jezikoslovlje.

ERSJ 2006: Aleksandar Loma (urednik), Marta Bjelitić, Jasna Vlajić-Popović, Marija Vučković, Maja Đokić, Aleksandar Loma, Snežana Petrović. *Etimološki rečnik srpskog jezika, Sveska 2: BA–BD*. Beograd, Institut za srpski jezik SANU.

María Biezma. 2013. Only one ‘at least’: Refining the role of discourse in building alternatives. In K. Shwayder, editor, *Proceedings of Penn Linguistics Colloquium 36*, pages 11–19.

Wallace Chafe. 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In Charles N. Li, editor. *Subject and Topic*, pages 25–55. New York: Academic Press.

Guglielmo Cinque. 2005. Deriving Greenberg’s Universal 20 and Its Exceptions. *Linguistic Inquiry*, 36(3), pages 315–332.

Cynthia G. Clopper and Judith Tonhauser. 2011. On the Prosodic Coding of Focus in Paraguayan Guaraní. In Mary Byram Washburn, Katherine McKinney-Bock, Erika Varis, Ann Sawyer, and Barbara Tomaszewicz, editors. *Proceedings of the 28th West Coast Conference on Formal Linguistics*, pages 249–257.

P. G. Grosz. (2011). A uniform analysis for concessive at least and optative at least. *Proceedings of SALT 21*, pages 572–591.

M. A. K. Halliday. 1967. Notes on Transitivity and Theme in English. *Journal of Linguistics*, Vol. 3, No. 2, pages 199–244.

- Annika Herrmann. 2014. *Modal and Focus Particles in Sign Languages: A Cross-Linguistic Study*. Berlin, Boston, De Gruyter Mouton.
- Katalin É. Kiss. 1998. Identificational Focus and Information Focus. *Language*, 74(2), pages 245-273.
- Ekkehard König. 1991. *The Meaning of Focus Particles: A Comparative Perspective*. London and New York: Routledge.
- Iliyana Krapova. 2008. Obshti cherti na sintaktichniya stroezh na taka narechenata lyava periferiya na izrechenieto v balkanskite ezitsi. *Imeto v pametta na ezika. Sbornik v pamet na profesor Boris Simeonov*, pages 181-198. Sofia, Figura.
- Manfred Krifka. 2008. Basic Notions of Information Structure. *Acta Linguistica Hungarica*. Vol. 55, No. 3-4, pages 243-276.
- Paul Kroeger. 2017. Basic Concepts in Information Structure: Topic, Focus, and Contrast. *GIALens 2017*, Vol. 11, No. 1.
- Knud Lambrecht. 1994. *Information Structure and Sentence Form: Topic, Focus, and the Mental Representations of Discourse Referents*. Cambridge University Press.
- Mirjana Mišković. 2001. The particle *baš* in contemporary Serbian. *Pragmatics*, 11 (1), pages 17-30.
- Hiroaki Nakamura. 2006. Topic-Comment Articulation in Japanese: A Categorical Approach. *Proceedings of the 20th Pacific Asia Conference on Language, Information and Computation*, pages 198-205. Tsinghua University Press.
- Ruselina Nitsolova. 2001. Za kontrastnija fokus s chastici v balgarskiya ezik. *Savremenni sintaktichni teorii*, pages 76-85. Plovdiv.
- Ruselina Nitsolova. 2008. *Balgarska gramatika. Morfologiya*. Sofia, Universitetsko izdatelstvo "Sv. Kliment Ohridski".
- Kuzman Pungershek, Taja and Peter Rupnik, Vít Suchomel, and Nikola Ljubešić. 2026. The Growing Gains and Pains of Iterative Web Corpora Crawling: Insights from South Slavic CLASSLA-web 2.0 Corpora. In submission.
- Randolph Quirk, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. 1972. *A grammar of contemporary English*. Beccles and London, Longman.
- Şükriye Ruhi. 2013. The interactional functions of tamam in spoken Turkish. *Dil ve Edebiyat Dergisi / Journal of Linguistics and Literature*, 10:2, pages 9-32.
- Martin Stefanov. (2026). Restrictive particles as discourse markers in South Slavic languages (contrastive study on the *at least* particles in Bulgarian, Serbian, Croatian, and Slovene). *New Directions in Slavic Studies: Language, Literature, and Political Discourse. Proceedings of the 6th International Conference GranaSlavic2025*, University of Granada, 2 – 4 June 2025, pages 645-654.
- Yovka Tisheva and Marina Dzhonova. 2003. Lexical Markers on the Information Structure Level. In Anissava Miltenova, David Birnbaum, Sarah Slevinski, editors. *Computational Approaches to the Study of Early and Modern Slavic Languages and Texts. Proceedings from the International Conference on Electronic Description and Edition of Slavic Sources*, pages 53-76. Sofia, Boyan Penev Publishing Centre.
- Yovka Tisheva and Marina Dzhonova. 2006. Lexical Tools for Focus Marking – The Case of *samo* in Bulgarian. *Formal Approaches to South Slavic and Balkan Languages – 5*, pages 147-150.
- Enric Vallduví & Elisabet Engdahl. 1996. The Linguistic Realisation of Information Packaging. *Linguistics*, 34(343), pages 459-519.