

Bulgarian Vocative within HPSG framework*

Petya N. Osenova[†] and Kiril Iv. Simov

BulTreeBank Project

<http://www.BulTreeBank.org>

Linguistic Modelling Laboratory, Bulgarian Academy of Sciences

Acad. G. Bonchev St. 25A, 1113 Sofia, Bulgaria

petya@bultreebank.org, kivs@bultreebank.org

June 30, 2002

1 Introduction

Crosslinguistically vocatives are an underexplored linguistic phenomenon and in different languages they can be highly idiosyncratic and complex [Levinson 1987, p.71]. Therefore, the problem, which is discussed in this paper, is not a language-specific one, in spite of the fact that most of the languages have their own repository for marking the role of the addressee in communicative utterances.

In our opinion this linguistic phenomenon needs its adequate treatment because of two main reasons:

1. The vocative is supposed to be present on two levels: syntax and pragmatics. Therefore it needs more elaborate interpretation on the interface side, which, in HPSG, is more developed for morphology/syntax and syntax/semantics than syntax/pragmatics;
2. It will be useful for HPSG-oriented implementations, especially treebanks and dialogue systems.

The paper is structured as follows: in the next section the status of the vocative in Bulgarian is discussed. In section 3 we propose our ideas on a unified treatment of vocatives. In section 4 the HPSG model is given. Section 5 outlines the conclusions and future work.

2 The Status of the Vocative in Bulgarian

Vocatives are assumed to be restricted to the second person usage only. Usually they subsume the following two subtypes: calls (*hey you*) and addresses (*Madam*) [Levinson 1987, p. 71]. Bulgarian vocative role is usually treated within the opposition: vocative form (a remnant of the case paradigm) vs. base nominative form, i.e. with respect to the presence or loss of the special vocative inflections. Hence,

*The work reported here is done within the BulTreeBank project. The project is funded by the Volkswagen Stiftung, Federal Republic of Germany under the Programme "Cooperation with Natural and Engineering Scientists in Central and Eastern Europe" contract I/76 887. The authors wish to thank the Seminar für Sprachwissenschaft of the Eberhard-Karls-Universität, Tübingen, for hosting the writing of this paper, and the Internationales Zentrum of the Eberhard-Karls-Universität, Tübingen, for funding Simov's visit to the Seminar. Also, the authors are grateful to Prof. Erhard Hinrichs, Frank Richter and Manfred Sailer from SfS for helpful advice, discussions and comments on the earlier drafts of this paper.

[†] Also at the Bulgarian Language Division, Faculty of Slavonic languages, St. Kl. Ohridsky University, Sofia, Bulgaria

functionally it includes not only the structural vocative case of the paradigm of some nominals (the singular masculine and feminine nouns and the long forms of some adjectives). As a matter of fact, it can be expressed by: vocative forms, nominal groups in nominative and different particles. We rely on the classification of vocatives, presented in [Ivanova and Nitsolova 1995, pp. 24-29]:

1. vocative particles, which do not nominate the addressee, but only refer to him/her (hej,be)
2. vocatives, which nominate the addressee
 - (a) they function as contact establishers only
 - (b) they qualify the addressee.

In the last case it is assumed [Nitsolova 1984, p. 44] that there is a hidden proposition with the performative verb 'consider somebody to be of some property'. Traditionally vocatives are assigned three interpretations concerning their syntactic position in the sentence:

1. Non-arguments, such as parenthetical elements, which do not participate in syntactic relations [Popov 1983, p. 130] and [Brezinski 2000, p. 94].

(1) Gospodine, dnes ste vali
sir[sg,voc] today it-will-rain[fut,3p,sg]
Sir, it will rain today

2. Subjects, when they are used in imperative, exclamative and optative sentences, agreeing with the verb and obeying the additional requirement not to be duplicated by a pronoun [Popov 1983, p. 129] and [Brezinski 2000, p. 94], [Acad.Gram. 1983, p. 120].

(2) Gospodine, elate nasam
sir[sg,voc] come[imper,pl] here[adv]
Sir, come here

3. Appositions, when they appear together with a pronoun [Nitsolova 1984, p. 43] and [Popov 1983, p. 130].

(3) Vie, gospodine, elate nasam
You[pl,2p] sir[sg,voc] come[imper,pl] here[adv]
You, Sir, come here

We argue that assigning the vocative three distinct syntactic/pragmatic roles is misleading and irrelevant on linguistic grounds. Our contra-arguments are as follows:

1. If vocatives are pure discourse markers and do not contribute to the sentence, how we could explain reasonably the structure-sharing between the vocative form and some parts of the sentence (subject, object, possessives etc).
2. If we assume that in Bulgarian only nominatives can be assigned a subject role, then it is strange to select a context, in which the vocative is assigned such a role. The agreement relations between the subject and the verb is not a pretty strong argument, because there are cases, where:
 - (a) between the explicit subject and the verb there is no agreement relation

- (4) Vsichki izljazohme na razhodka
 All[3p,pl] went[1p,pl] for walk[sg,f]
 All of us went for a walk

(b) the verb can agree with either of the the parts of the subject

- (5) Chast ot studentite vleze/vljazoha
 Part[sg,f] of students-the[pl] came[3p,sg]/[3p,pl]
 Part of the students came

(c) the agreement relation depends on the lexical semantics of the conjuncts in a coordination

- (6) Radost i taga imashe v ochite mu
 Joy[sg,f] and sadness[sg,f] there-was[3p,sg] in eyes[pl] his-clitic[3p,poss]
 There was joy and sadness in his eyes

For HPSG-oriented discussion of the first two types see [Osenova (to appear)].

3. Vocatives can combine in the same way with all types of illocutionary force, including declaratives and interrogatives:

- (7) Gospodine, kakvo tarsite?
 sir[sg,voc] what[interrogative pron] search[pres,2p,pl]
 Sir, what are you searching for?

4. Bulgarian is a pro-drop language, in which the subject is always realized on the verb and has the characteristics of the nominative personal pronouns. In the cases, where *hey you* vocative type is triggered, we can assume that the second person pronoun has two syncretic forms: one for nominative and one for vocative.
5. Vocative is outside the scope of the Left dislocation or Left Periphery phenomena, which in Bulgarian are usually connected with object doubling [Penchev 1993, p. 120] or with complementizers [Krapova and Karastaneva 1999]. Needless to say, its contribution to the Information structure of the utterance needs more elaborate research.

3 Towards a Unified View of the Vocative in Bulgarian

Vocatives play a pragmatic role with respect to the addressee of an utterance. But it is still not explained what the interaction between the syntactic and pragmatic behavior of the vocative is. Here we are not concerned with encoding of the speaker's intentions in BACKGROUND feature [Green 2000] or with metapragmatic phenomena like honorifics. Rather, we concentrate on C-INDICES and their contribution to the adequate formalization of the vocative-sentence relation.

It is interesting to compare how this problem has been dealt with for more practical purposes. We suggest as an example the Verbmobil treebank. In the English HPSG-oriented part the NP, vocatives are treated as adjuncts and therefore attached to the highest sentential level, and the particles are treated as discourse markers [Kordoni 2000, p. 21 and p. 40]. In the German part all of them are treated as discourse markers and therefore they stay unattached [Stegmann, Telljohann and Hinrichs 2000, p. 40]. We propose to combine the two views in one, i.e. to interpret the vocatives as adjuncts and discourse markers at the same time. We need the first, because vocatives very often share syntactic properties with the elements within the sentence and we need the second, because the interaction is done on the super-sentential level.

Another fact that supports our idea is the free word order of the NP vocatives (here, due to space limitations, we do not give all the possibilities of the vocative position in the sentence). One could argue that the ostensive particles, the pronouns or complex vocative groups are more restricted in their vocative distribution, but it does not make a contradiction. All of them perfectly fit into the adjunct interpretation, because adjuncts can be recursive and of different nature as well. As a result of our modular view on this phenomenon, we propose the following types of interrelation between syntactic and pragmatic specificity of the vocatives:

1. the vocative and the expressed/unexpressed subject or object in the sentence refer to the same entity. In the first case the vocative can be considered a subject controller of the matrix verb.

(8) Gospodine, izpusnahte vlaka
 sir[sg,voc] missed[aorist,2p] train-the[sg,masc,def]
 Sir, you missed the train

(9) Gospodine, tarsiat Vi
 sir[sg,voc] look for[present,2p] you-clitic[pl,acc]
 Sir, they are looking for you

2. the vocative can bind some possessive or reflexive pronoun in the sentence

(10) Zena vi, gospodine, se obazhda
 wife[sg,f] your[pron,poss] sir[sg,voc] call[pres,3p,sg]
 Sir, your wife is calling

(11) Gospodine, sprete zena si
 sir[sg,voc] stop[imperativ] wife[sg,f] your[pron,poss,refl]
 Sir, stop your wife

3. the vocative just serves as an ostensive stimulus for the hearer regarding some fact

(12) Gospodine, navan vali
 sir[sg,voc] outside[adverb] rain[present, 3p, sg]
 Sir, it is raining outside

4 Modelling Bulgarian Vocatives in HPSG

In this section we present a formal model of Bulgarian vocative along the lines of [Pollard and Sag 1994] with the necessary modifications and refinements to it. On the syntactic level we treat vocatives as a special kind of adjuncts that contribute to the highest sentence node via their MOD feature. The main reason for choosing such an approach is that the proposition becomes visible for the vocative expression. This is needed, when the vocative shares some properties with inner-sentence elements. The information, specified for the vocative, is added to the value of the ADDRESSEE feature within CONTEXT value of the proposition, ensuring that the vocative coincides with the hearer of the proposition. Hence, in sentences, which explicitly refer to the hearer, the vocative expression is co-indexed with the appropriate syntactic elements within the proposition.

One problem when treating the vocative as an adjunct is the Semantics principle in HPSG. It requires the CONTENT of the mother to be structure-shared with the CONTENT of the semantic head. In head-adjunct phrases the adjuncts are assumed to be semantic heads. In our account of the vocative we change this by stating that vocative cannot be a semantic head.

In order to have our idea working, we need principles. HPSG94 does not introduce any principles, which operate on the c-indices of the daughters of a phrase [Pollard and Sag 1994, p. 337]. At the same time, we are aware of the relevant exploitation of the CONTEXT feature for resolving dialogue fragments within HPSG [Ginzburg, Gregory and Lappin 2001]. For our present purpose (to model the vocative-sentence relation first) we propose the following principle:

Vocative Principle:

In a head-adjunct structure, in which the adjunct is of case vocative, the ADDRESSEE value of the adjunct is token-identical with the ADDRESSEE value of the mother.

Additionally, we change the definition of the semantic head. This change can be regarded as a consequence of the pragmatic nature of the vocative and as shifting its information contribution to CONTEXT:

Vocative Semantic Head:

In a head-adjunct structure, in which the adjunct is of case vocative, the head daughter is the semantic head.

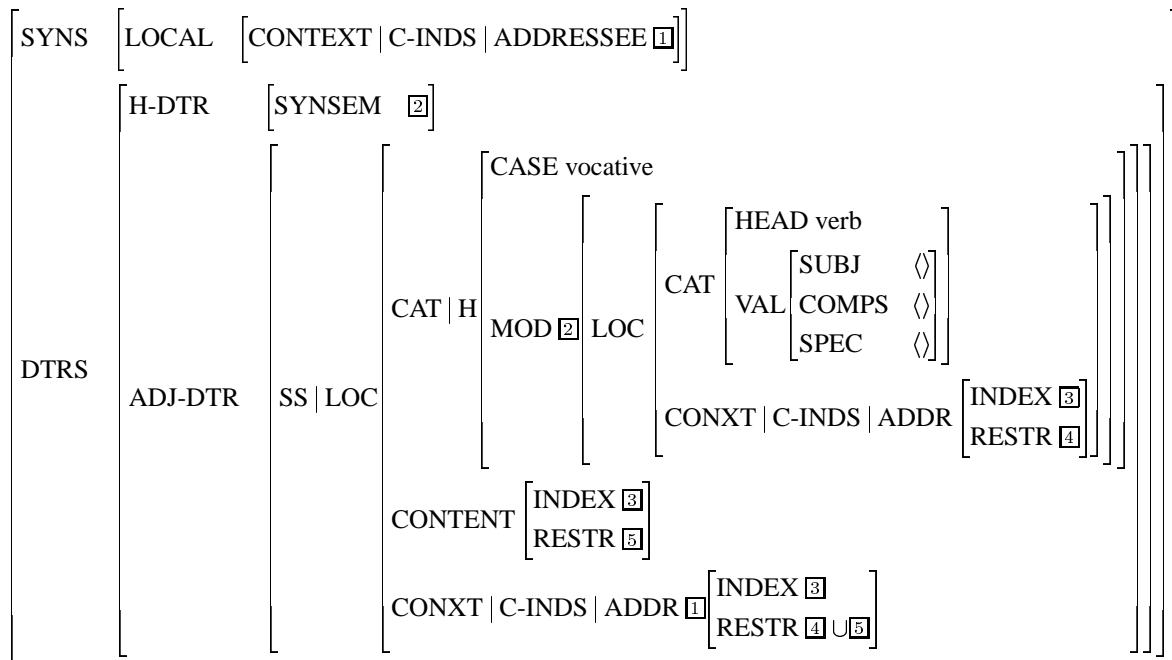
Now it is possible to present the general mechanism:

1. The whole sentence has a DTRS value of sort headed-adjunct-structure, where the head daughter is a saturated verb phrase and the adjunct daughter is the vocative phrase.
2. The appropriate selection mechanism is encoded within the vocative phrase. Thus, the MOD value of the adjunct (the vocative phrase) requires as a head daughter a saturated verb phrase (equal to a sentence).
3. The MOD value of the adjunct is token identical with the SYNSEM value of the head daughter and thus the ADDRESSEE value of the head daughter is available within the structure of the adjunct.
We define that the INDEX value of the ADDRESSEE (ADDR) feature of the adjunct is co-indexed with the INDEX value of the CONTENT feature of the vocative phrase and also with the INDEX within the ADDRESSEE value of the head daughter (see co-reference 3 in the feature structure given on page 6). The restriction (RESTR) of the ADDRESSEE of the vocative phrase is union of the restriction of the CONTENT of the vocative phrase and the restriction of the ADDRESSEE of the head daughter (see co-references 4 and 5 in the feature structure given on page 6).
4. Via the Vocative Principle, defined above, the ADDRESSEE value of the adjunct is token-identical with the ADDRESSEE value of the mother (see co-reference 1 in the picture).

A schematic feature structure is given on page 6.

Below we give a simplified idea about the formal mechanisms, which ensure the interrelation between syntactic and pragmatic specificity within the three vocative types, discussed at the end of the previous section.

1. The vocative and the expressed/unexpressed subject or object in the sentence refer to the same entity in the world. Within the head daughter the subject's CONTENT (or the object's CONTENT) value is structure-shared in appropriate way with the ADDRESSEE in the CONTEXT value.



2. The vocative binds some possessive or reflexive pronoun in the sentence. This binding first takes place within the head sentence where the possessive or reflexive pronoun is bound by the subject of the sentence and the subject's CONTENT is co-indexed in appropriate way to the ADDRESSEE value of the sentence. From there it gets co-indexed with the vocative.
3. The vocative just serves as an ostensive stimulus for the hearer regarding some fact. In this case again we have an appropriate co-indexed value of the ADDRESSEE and vocative CONTENT, but the INDEX of the ADDRESSEE is not structure-shared with any other INDEX in the CONTENT of the sentence.

However, one technical problem remains: how to encode entries of the nominative nominals, which could serve as vocatives as well. In our view this can be stated in two ways: (1) as a disjunct with the MOD feature for all appropriate lexical signs, or (2) via a lexical rule. In our opinion the second mechanism is more appropriate, because practically all nominals can be used in a vocative role, but the first can ensure better treatment of the exceptions.

5 Conclusion

In this paper we present an HPSG-based unified analysis of Bulgarian vocative phrases. At the same time, in our view, the general interaction between the adjunct vocative phrase and the selected sentence tends to be universal. Note the vocative role in other languages like Spanish, Czech, Russian, English, German, Polish etc. Language specific remains the typology of vocatives and the domains of their structure-sharing with different sentential elements.

As a natural future direction of this work we consider extending our idea with respect to the illocutionary force of the selected by the vocative sentence. It is to be done within the more recent two-dimensional reclassification of phrases [Sag 1997] and the proposed structuring of the conversational move types [Ginzburg, Sag and Purver 2001, p. 6].

References

- [Acad.Gram. 1983] Acad.Gram. *Gramatika na bylgarskiia knizhoven ezik. Sintaksis*. Sofia. Bulgaria.
- [Brezinski 2000] Brezinski St. *Balgarski sintaksis*. Sofia.
- [Ginzburg, Gregory and Lappin 2001] Ginzburg J., Gregory H. and Lappin Sh. *SHARDS: Fragment Resolution in Dialogue*. In Bunt H., van der Sliuis, Thijsse eds., Proc. of the 4th International conference on Computational Semantics, Tilburg, pp. 156-172.
- [Ginzburg, Sag and Purver 2001] Ginzburg J., Sag I. and Purver M. *Integrating Conversational Move Types in the Grammar of Conversation*. In Kuhlein P., Reiser H., Zeevat H, eds. Proc. of Bidialog, the 5th Workshop on the Semantics and Pragmatics of Dialogue, Bielefeld.
- [Green 2000] Green G. *The Nature of Pragmatic Information*. In Cann, Ronnie, Claire Grover and Philip Miller, ed. Grammatical Interfaces in HPSG. CSLI Publications, pp. 292, Studies in Constraint-Based Lexicalism 8
- [Ivanova and Nitsolova 1995] Ivanova K, Nitsolova R. *Nie, govoreshtite hora*. Sofia. Bulgaria.
- [Kordoni 2000] Kordoni V. *Stylebook for the English Treebank in VERBMOBIL*. Report 241. Sfs, Universitat Tubingen.
- [Krapova and Karastaneva 1999] Krapova I. Karastaneva T. *On the structure of The CP field in Bulgarian*. Plovdiv, Bulgaria.
- [Levinson 1987] Levinson S. *Pragmatics*. Cambridge Textbooks in Linguistics.
- [Nitsolova 1984] Nitsolova R. *Pragmatichen aspect na izrechenieto v balgarskiia ezik*. Narodna prosveta. Sofia, Bulgaria.
- [Osenova (to appear)] Osenova P. *On Subject-Verb agreement in Bulgarian (An HPSG-based account)*. Proc. of the Fourth Formal Description of Slavic Languages Conference, Potsdam, Germany, November 2001.
- [Pollard and Sag 1994] Carl J. Pollard and Ivan A. Sag. *Head-Driven Phrase Structure Grammar*. University of Chicago Press, Chicago, Illinois, USA.
- [Penchev 1993] Penchev Y. *Balgarski sintaksis*. Plovdiv, Bulgaria.
- [Popov 1983] Popov K. *Sintaksis*. Veliko Tarnovo, Bulgaria.
- [Sag 1997] Sag I. *English relative clause construction* Journal of Linguistics, 33, pp. 431-484.
- [Stegmann, Telljohann and Hinrichs 2000] Stegmann R., Telljohann H. and Hinrichs E. *Stylebook for the German Treebank in VERBMOBIL*. Report 239. Sfs, Universitat Tubingen.