

VERBAL MULTIWORD EXPRESSIONS IN THE BULGARIAN NATIONAL CORPUS

Maria Todorova, Ivelina Stoyanova
Institute for Bulgarian Language, Bulgarian Academy of Sciences

The paper is focused on verbal MWEs and investigates their usage in the Bulgarian National Corpus. The study employs a large dictionary of 27,902 Bulgarian verbal MWEs. The analysis aims at enhancing dictionary description of verbal MWEs, and can also influence the development of methods for automatic identification and classification of verbal MWEs by examining their properties, usage and variability.

Key words: syntax, Bulgarian language, Natural language processing, language technologies

1. Introduction

The investigation of multiword expressions (MWEs) with a view to their computational identification and extraction is a very popular problem in the NLP lately. A lot of work has been done on the classification and tagging of (both verbal and nominal) MWEs. Earlier research on the lexical encoding of idioms and their syntactic variations relied mainly on human intuition (Villavicencio et al. 2004). Similar approaches have been developed for Bulgarian (Todorova 2010, 2015). Though manual description of MWE is reliable, it is laborious and dependable on the linguist's subjective opinion. We invested efforts into a different approach combining large language resources in order to examine objectively the usage and variability of MWEs. We experimented on enriching the dictionary description of Bulgarian verbal multi word expressions automatically, using large corpora such as the Bulgarian National Corpus (BulNC).

Our proposal for the automatic acquisition and encoding of the grammatical and syntactic description of idiomatic expressions is based on data extracted from the corpus. We identify automatically occurrences of MWEs and analyse them in their context. Further, on the basis of the data we deduce the morphological and syntactic properties and variations which we include in their lexicographic description in the Bulgarian Dictionary of MWEs. Semantic analysis is focused on the degree of idiomaticity of

MWEs. This can be used to identify new MWEs (not currently in the dictionary) as well as to automatically classify MWEs into groups – the morphosyntactic and semantic properties of the MWEs vary depending on the degree of idiomaticity, so different groups require different approach in their automatic treatment.

1.1. The Bulgarian National Corpus (BulNC)¹

The Bulgarian National Corpus (BulNC) (Koeva et al. 2011; Koeva et al. 2012) is a large general corpus of Bulgarian currently consisting of approximately 1.2 billion words in more than 240,000 text documents. The corpus reflects the state of the Bulgarian language from the middle of 20th century until the present, including the following types of texts: administrative; popular science; fiction; science; news; informal. This gives us reason to consider that such large resource will be reliable enough to provide examples for different kinds of MWEs. Another advantage of BulNC are the different levels of linguistic annotation. The Bulgarian texts are annotated using the Bulgarian Language Processing Chain: tokenisation and sentence splitting, morphosyntactic tagging and lemmatisation (Koeva, Genov 2011). Certain parts of the corpus have specialised annotation – semantic annotation, clause segmentation, MWE and NE annotation.

1.2. Verbal MWEs

The contemporary research on multiword expressions focuses on verbal MWEs (vMWEs) and especially on the description of their components and structure (Villavicencio et al. 2004; Gregoire 2010). The automatic extraction and tagging of vMWEs poses more challenges when we talk about morphologically rich languages. MWEs in Bulgarian are characterised with a rich inventory of synthetic and analytical verb forms; with a complex and flexible word order. They combine structural peculiarities, such as mandatory and optional components, possibility for insertion of external phrases (clitics, adverbial phrases, noun phrases, etc.), and discontinuous components. The homonymy with free phrases is another problem which occurs when extracting examples from corpora.

2. Method for identification of vMWE in corpora

For the purposes of our work on the identification of vMWEs in BulNC we use a large dictionary of 27,902 MWEs, combined with a set of rules and basic heuristics to manage verb forms, variations in word order and insertions of external elements. We also use frequency analysis and association measures for collocations to identify new MWEs which are not in the dictionary. Similar methodology has been applied by Fazly and Ste-

¹ <http://dcl.bas.bg/bulnc/>

venson (2006) for a particular syntactic type – verb+noun combination which authors characterise as having poor lexical and syntactic flexibility.

Determining a unique set of syntactic patterns appropriate for the recognition of all idiomatic combinations is a difficult task. The forms of an idiomatic combination are not entirely predictable (Sag et al. 2002). That is why we encode within the dictionary framework possible syntactic variations in terms of word order, optional components, and possible insertions both as modifiers of components (e.g., *vzemam vazhno reshenie* ‘make an **important** decision’) or external phrases between the components (e.g., *poemam barzo v druga posoka* ‘change **quickly** direction’).

2.1. Compiling a Bulgarian Dictionary of vMWEs

The Dictionary of Bulgarian vMWEs applied here is part of the Bulgarian dictionary of MWEs (Todorova, Stoyanova 2014) which contains over 86,373 MWEs, both nominal and verbal, extracted from different dictionaries of Bulgarian idioms, as well as entries automatically extracted from the BulNC. The verbal part of the Dictionary is a collection of 27,902 verb idioms.

The vast majority of vMWE entries are manually or semi-automatically collected from various printed and electronic dictionaries of Bulgarian idioms. The sources include the Explanatory Dictionary of Bulgarian (Andreychin et al. 2005) and the Bulgarian WordNet². The Dictionary of Bulgarian vMWEs is compiled automatically after the methodology of Fischer and Keil (1996) adopted for PhraseoLex. The data are being manually verified and normalised and MWE components were automatically POS annotated using the Bulgarian Language Processing Chain (Koeva, Genov 2011).

The dictionary includes a wide range of vMWE types with respect to the number of components (graphic words), the syntactic class and the structure, and the semantics (degree of idiomaticity). Entries are divided into several two main classes: true VP (24,234 MWEs) and sentential MWEs (2,345 MWEs), with multiple structural types and subtypes. The entries are classified into categories with respect to the degree of compositionality and transparency of meaning, called idiomaticity – vMWEs are divided into non-decomposable (e.g., *ritna kambanata* ‘kick the bucket’), semi-decomposable (e.g., *vidya v nova svetlina* ‘see in new light / in a different way’), simple decomposable (e.g., *svalya ot vlast* ‘overthrow from power’), and light verb constructions (e.g., *vzemam reshenie* ‘make a decision’). The idiomaticity was manually assigned to about 1000 entries.

² <http://dcl.bas.bg/bulnet/>

The description of each lexical entry is represented by a frame containing the following information: (a) morphological properties: the lemma of the MWE; POS and grammatical values, often inherited from the head of the phrase; inflection type, describing possible MWE paradigm and morphophonemic variations in MWE forms; (b) structural properties: a linear sequence of MWE components and structural specifics; restrictions on word order; optionality of components; etc.; (c) semantic properties: degree of idiomaticity.

Example 1. Representation of vMWE in the Dictionary

```

1 <MWE lemma="vzemam reshenie" pos="VP" MWE_TYPE="LVC" MWE_WO="free">
2 <word lemma="vzemam" wordform="vzemam" pos="VLIT">
3 <word type="external" lemma="" wordform="" pos="PPz">
4 <word type="external" lemma="ли" wordform="" pos="T">
5 <word type="optional" lemma="важен" wordform="важно" pos="Asno">
6 <word lemma="reshenie" wordform="reshenie" pos="Nsno">
7 </MWE>

```

As shown in Example 1, the dictionary scheme involves description of MWE structure and components and distinguishes between: compulsory components (line 2), optional components (line 5), external elements which are not part of the phrase but can appear between its components (lines 3 and 4), and substitutes which are free but internal for the phrase (e.g., *popadam v racete na **nyakogo*** ‘fall is **someone’s** hands’). Also, the word order of the MWE is described either as fixed or free. This is of particular importance for verbal MWEs which exhibit a variety of structures and word order variations, in order to be able to identify them in text.

As the idiomaticity type of the vMWE to a large degree determines the possibilities for variation, substitutions as well as its paradigm, some automatic checks were performed based on a set of heuristics, to identify possible inconsistencies in the MWE description. Examples of such inconsistencies include non-decomposable MWEs with optional components, or decomposable MWEs with frozen word order.

2.2. Method for identification of vMWE in BulNC

The method for identification of vMWEs from the dictionary involves the following steps:

(1) Preprocessing involving initial annotation – POS tagging and lemmatisation.

(2) For each vMWE entry in the dictionary, we extract from the corpus sentences containing its head verb (identified by its lemma).

(3) In each sentence we check if the other compulsory components of the MWE are present. Components are labelled in the dictionary with the following information: (a) participation in the MWE ‘compulsory’ or ‘optional’; (b) type of paradigm in the MWE – whether they are ‘frozen’ (in which case they are identified by the form they appear with in the MWE) or ‘variable’ (identified in the text by their lemma).

(4) If all components are present, we check if they fall into the same chunk. We consider punctuation and conjunctions to mark chunk boundaries. More precise chunking methods will be considered in the future.

(5) Next step involves checking the distance between any two consecutive components of the MWE against a set threshold to find non-contiguous MWEs. We limited the distance to 3 tokens in order to exclude random occurrences of the same combination of words as the components of the MWE.

The method for identification of new vMWEs which are not in the dictionary currently is limited to MWEs with two components (V+N, V+Adv, V+A) and involves the following steps:

(1) Preprocessing.

(2) In each sentence of the text the verbs are identified. Auxiliary verbs are excluded.

(3) Association measure is calculated between the verb and any of the words in its context within a certain range (currently 3 tokens) but within a single chunk, i.e. stopping at punctuation and conjunctions.

(4) If we identify a pair of words (verb and another component) with high association measure and frequency above certain threshold (currently set to 5 occurrences in BulNC), we list it as a candidate for a MWE, together with all its occurrences which reflect possible variations in terms of word order, optional components, insertions. Although these variations are limited in the case of 2-component MWEs, this analysis will be essential when extending the method to longer MWEs. It is manually confirmed by an expert if the candidate is indeed a MWE.

3. Enhancing dictionary description by analysing vMWEs extracted from the BulNC

The procedures for finding vMWEs in the corpus (described in section 2.2) identify expressions of certain syntactic structure, allowing flexible word order and taking into consideration possible variations in components, optional modifiers, etc. Based on the data extracted from the BulNC, the dictionary description of the MWE can be extended with more information. Corpus data are particularly useful for compiling description of

vMWEs since they exhibit much more variations compared to nominal MWEs, some of which are not systematic and cannot be easily predicted.

The analysis of the occurrences of the vMWE in the BulNC aim at the identification of any irregularities in their paradigm, and information about their structure, lexicogrammatical and morphosemantic properties, grammatical categories, morphophonemic changes of individual components of the MWE in different wordforms. The semi-automatic compilation of MWE description will facilitate the creation of large dictionaries of vMWEs for the purposes of various NLP applications.

By analysing the structure of the MWE we identify its main type – nominal, verbal or sentential, as well as its structural subtype within the main category (e.g., V-NP, V-PP, etc.). A set of heuristics is applied to select the MWE lemma out of all possible occurrences of the MWE in BulNC (Example 2). The purpose is to identify the least marked form: singular rather than plural, present tense rather than other tenses, positive form, 1st person, etc. Priority is given to the head, and then the other components. In Example 2, we identify the least marked form of the head verb (*vdigam*), and then consider the noun (*ramene*), which is in plural and does not occur in singular within the MWE (**vdigna ramo* ‘shrug a shoulder’), so we conclude that the least marked form is *vdigna ramene*. Since the least marked form is rarely the most frequent, in some cases it may not occur in the corpus at all, and this procedure ensures that the lemma assigned is an existing form of the MWE, even if it is not the exact lemma. It is then verified by an expert.

Example 2. Lemma of the MWE

vdigam/V1sr ramene/NNp0 ‘shrug (my) shoulders’

Corpus usage:

vdigam ramene – 1st person, sing., present tense (29 occurrences)

vdigash ramene – 2nd person, sing., present tense (4 occurrences)

vdiga ramene – 3rd person, sing., present tense (116 occurrences)

vdigat ramene – 3rd person, pl., present tense (28 occurrences)

vdigashe ramene – 2/3 person, sing., Imperfect (67 occurrences)

vdigaha ramene – 3rd person, pl., Imperfect (45 occurrences)

Selected lemma: *vdigam ramene*

The lexicogrammatical and morphosemantic properties of MWE components give information about the possible MWE paradigm. It is defined by a set of parameters for each component – nouns have forms for singular and plural, definite and indefinite form, verbs has forms for person, number, tense, mood, etc. The paradigm of the MWE does not usually

allow full realisation of components' wordforms, in fact most components are likely to appear in a frozen form. In Example 2, we see that the verb has forms for person, number, tense, so we conclude that it makes full paradigm, while the noun (*ramene*) occurs only in plural, undetermined, so we conclude that this component has one fixed form.

Variations in word order are also identified by observing corpus data. Although currently, for unification purposes, we adopt a standard verb-NP-PP word order for the lemma, in some cases it may be preferred to select the most frequent word order for the lemma (e.g., *chakam ot umryal pismo*, Example 3).

Example 3. Word order of the MWE

chakam ot umryal pismo 'to wait for a letter from the dead' (to expect pointlessly)

- V PP NP (40 occurrences)

*Da chkash pari ot darjavata e vse edno da **chakash ot umryal pismo**.* (It's pointless to wait for money from the state)

- PP NP V (6 occurrences)

*Pozvanih na 112 – **ot umryal pismo da chakash**...* (I called 112 – it's pointless...)

- V NP PP (2 occurrences)

*Popalvash edno formulyarचे i **chakash pismo ot umryal**.* (You fill in a form and you wait pointlessly.)

The analysis of the occurrences of vMWEs also allows us to identify cases where a component of the phrase can receive a modifier, as well as positions where external insertions are possible. In many cases it is difficult to distinguish internal modifiers from insertions, except when the modifier is in agreement with the respective component (e.g., adjective modifier of a noun). The analysis also allows us to describe the categories of the inserted elements (pronominal or interrogative clitics or whole phrases) and, whenever possible, to label them as either optional modifiers or external elements.

Example 4. Insertion of modifiers and external phrases

hvarlyam pogled 'take a look', V NP (473 occurrences)

- Identified as modifiers as they agree with the noun component *pogled* (bold, underlined)

*Nyamah vreme da **hvarlya posleden pogled** na doma si.* I had no time to take a **final** look at my home.

*Bashta mi mu **hvarli izumen pogled**.* My father took a **stunned** look at him.

- External insertions (underlined)

Hvarli mu pogled i otmina. (He/She) took a look at him/it and passed.

Hvarli li mu pogled? Did you take a look at him/it?

Hvarliam nabarzo pogled nazad kam zhivota si. I take a look quickly back on my life.

- Both modifier and external insertions

Hvarlih mu edin pogled za dovizhdane. I took one look at him for goodbye.

The dictionary information derived automatically needs verification in order to ensure the high quality of the resource. Further, it is problematic to distinguish between the occurrence of the MWE as compared to its free phrase counterpart (if it exists) or mere coincidence of co-occurrence of words, so in many cases expert intervention is needed. Sometimes while the MWE exhibits constraints in the paradigm, the free phrase may show more variability in forms and lead to wrong categorisation of the MWE. Example 5 shows how we can incorrectly deduce that in the MWE *barkam v dzhoba*, the noun *dzhoba* can change in number or assign modifiers, or that external elements can be inserted, while these are not examples of the use of the MWE but of its homonymous free phrase just coincidentally co-occurring words.

Example 5. vMWE vs. free phrase

barkam/V v/P dzhoba/Nsh ‘dig into (someone’s) pocket’ (to cause someone financial losses)

- MWE

Polititsite pak shte barkat v dzhoba na obiknoveniya danakoplatets.
The politicians will again dig into the pocket of the regular tax-payer.

- Free phrase

Ne biva da barkash v chuzhdi dzhobove/Np0. You must not put your hands in other people’s pockets.

Vidyah go da barka v predniya si dzhob. I saw him searching his front pocket.

- Coincidental co-occurrence of *barkam* and *dzhoba* in proximity

Barkashe choveka s raka v dzoba i drugiya sas strannata pricheska.
He confused the person with a hand in his pocket with the other one with the strange haircut.

There are some vMWEs whose components may vary in number and/or lexical realisation. These are rarely identifiable in the corpus, since the search is either token- or lemma-based. Moreover, as the verb is the

head of the vMWE, a different verb means a different vMWE. However, sometimes it is difficult to cover all variants or inefficient to do so since the variations are systematic. The analysis show that these variations are of three types: (a) the pair of verbs from the same stem but different in aspect (perfective and imperfective) can usually both be used in the MWE; (b) components are substituted by a derivationally related word (e.g., the verb is prefixed); and (c) components are replaced by a synonym.

Example 6. Lexical variations of vMWE components

- Perfective and imperfective verbs

Imperfective: *barkam/V v/P dzhoba/Nsh* (to cause someone financial losses)

Perfective: *brakna/V v/P dzhoba/Nsh*

- Derivationally related verbs

chakam ot umryal pismo ‘to wait for a letter from the dead’ (to expect pointlessly)

chakam – *dochakam*, *zachakam*

Nyama da dochakash ot umryal pismo. You won’t be able to wait until you get a letter from the dead.

Podade molbata i zachaka ot umryal pismo. He filed the claim and started waiting for a letter from the dead.

- Synonyms of the verb

izvadya dushata ‘to take (someone’s) soul out’ (to cause someone big distress by pestering or threatening them)

izvadya dushata (67 occurrences) – perfective

izkaram dushata (14 occurrences) – perfective

izvazhdam dushata (4 occurrences) – imperfective

izkarvam dushata (2 occurrences) – imperfective

Izvadi dushata na bashta si da mu kupi fotoapararat. He pestered his father to buy him a camera.

Shte mi izkarash dushata s tezi skandali. You will kill me with these fights.

As the dictionary is compiled from different MWE sources, one MWE can appear several times with several different descriptions, sometimes even more than one lemma: e.g., *pisnalo mi e* (present perfect, as most frequent) – *pisva mi* (present tense), *idva mu otvatre* (it comes naturally to **him**, as most frequently the pronoun is in 3rd person) – *idva mi otvatre* (it comes naturally to me). Procedures for unifying the lemmas and discovering variants of the same vMWE will be implemented in the future.

The analysis of the occurrences in context can also be used to deduce subcategorisation information. Prepositions, such as *na*, *za*, *ot*, within 3 tokens (currently) before the first or after the last component of the MWE are considered together with the NP they introduce, as possible argument candidates, as in Example 7.

Example 7. Prepositional phrases in proximity to the vMWE

vlizam v oganya ‘to go into fire’ (to sacrifice myself)

vlizam v oganya za/P (7 occurrences)

Za priyatel i v oganya vlizam. For a friend I will go into fire.

Te i v oganya vlizat za detsata si. They go into fire for their children.

Example 8 shows that the BulNC, together with some analysis of the grammatical features, can be used to identify possible MWE derivations using derivational patterns (e.g., pairs of verb-noun suffices).

Example 8. Derivations from vMWEs

- Present participle functioning as an adjective

smrazyavam kravta ‘to make someone’s blood freeze’ (to frighten someone)

smrazyavasht kravta (present participle functioning as an adjective) ‘such that makes someone’s blood freeze’

Djeri nadade smrazyavasht kravta voy. (Djery made his blood freeze.)

- Transgressive form, adverbial

hvarlyam pogled ‘to take a look’

hvarlyayki pogled (transgressive form) ‘taking a look’

Toy se usmihna, hvarlyayki y pogled. He smiled while taking a look at her.

- Nominal phrase derived from the vMWE – these can vary in semantics

igraya karti ‘play cards’

igrach na karti ‘card player’ (agent)

igra na karti ‘card game’ (event)

Toy e strasten igrach na karti. He is a keen card player.

We show the variety of information that can be directly observed or deduced from corpus data, regarding the lexicogrammatical, structural and semantic properties of vMWEs. More detailed analysis will allow the more precise extraction of information and dictionary description of vMWEs, which will reduce the need of manual verification of their description.

4. Conclusions and Future work

The paper describes work in progress on the employment of corpus data for the purposes of automatic description of vMWEs. We observe the usage of vMWEs in the Bulgarian National Corpus – a large representative corpus of Bulgarian. The main result consists in enhancing dictionary description of vMWEs by adding new features.

The observations and conclusions can facilitate the development and testing of methods and approaches for identification of vMWEs in text. Future work will be focused on improving the method in terms of precision (solving ambiguity, etc.) and in terms of efficiency, as well as on implementing more advanced methods for MWE recognition. For example, we consider searching for the NP or PP within the vMWE first, as they are often constant or are also MWEs which can occur with different verbs, e.g. *sreshtu techenieto* ‘against the current’ from *pluvam sreshtu techenieto* ‘to swim against the current’ can occur with other verbs *varvya* ‘to walk’, *tragvam* ‘to go’, etc. This can be a successful approach with light verb constructions or those decomposable vMWEs allowing substitutions.

We are also interested in the semantic properties of vMWEs, the degree of idiomaticity and the way it influences the syntactic behaviour of the vMWEs, their argument structure, etc. Further investigations can also focus on the coverage of vMWEs in the BulNC and other corpora of Bulgarian.

REFERENCES

- Andreychin et al. 2005:** Andreychin, L., Georgiev, L., Ilchev, S., Kostov, N., Lekov, I., Stoykov, S., Todorov, T. *Balgarski talkoven rechnik*. Dopolneno i preraboteno izdanie ot D. Popov. Sofia: Nauka i izkustvo, 2005.
- Fazly, Stevenson 2006:** Fazly, Afsaneh, and Stevenson, Suzanne. Automatically constructing a lexicon of verb phrase idiomatic combinations. // *Proceedings of the 11th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, Trento, Italy, April 2006, 337–344.
- Fischer, Keil 1996:** Fischer, I., and M. Keil. Parsing decomposable idioms [Online report]. // *Proceedings of the 16th Conference on Computational linguistics. Vol. 1. Association for Computational Linguistics*, 1996, 388–393. Available from <http://bit.ly/1AUzbBB>.
- Gregoire 2010:** Gregoire, Nicole. DuELME: a Dutch electronic lexicon of multiword expressions. // *Language Resources & Evaluation* (2010) 44, 23–39.
- Koeva et al. 2011:** Koeva, Svetla, Svetlozara Leseva, Borislav Rizov, Ekaterina Tarpomanova, Tsvetana Dimitrova, Hristina Kukova and Maria

- Todorova. Design and Development of the Bulgarian Sense-Annotated Corpus. // *Actas del III congreso de linguistica de corpus. Las tecnologicas de la informacion y las comunicaciones: presente y futuro en el analisis de corpus*. Valencia, Spain, 7–9 April 2011, 143–150.
- Koeva et al. 2012:** Koeva, Svetla, Ivelina Stoyanova, Svetlozara Leseva, Tsvetana Dimitrova, Rositsa Dekova and Ekaterina Tarpomanova. The Bulgarian National Corpus: Theory and Practice in Corpus Design. // *Journal of Language Modelling*, 2012, Vol. 0, No. 1, 65–110.
- Koeva, Genov 2011:** Koeva, Svetla and Angel Genov. Bulgarian Language Processing Chain. // *Proceeding of the Workshop Integration of multilingual resources and tools in Web applications*, Hamburg, 2011.
- Sag et al. 2002:** Sag et al., 2002. Multiword expressions: a pain in the neck for NLP [Online report]. // *Proceedings of the 3rd International Conference on intelligent text processing and computational linguistics (CI-CLing – 2002)*. Berlin, Heidelberg: Springer, 2002, 1–15. Available from <http://bit.ly/1KJc6ou>.
- Todorova 2010:** Todorova, 2010. Todorova Maria. Organizaciya na lingvistichnata informaciya v komputeren morfologichen rechnik na balgarski glagolni frazeologizmi (Organisation of the Linguistic information in Computational Morphological Dictionary of Bulgarian Verb Idioms). // *Proceedings from the Fifth Conference on Lexicology and Lexicography*, 2010, 331–344.
- Todorova 2015:** Todorova, 2015. Todorova, Maria. *Typology and Properties of Multiword Expressions in Bulgarian. Verb idioms*. PHD Thesis, Sofia 2015. http://dcl.bas.bg/Disertacia_M.Todorova/M.Todorova_disertacia.pdf
- Todorova, Stoyanova 2014:** Todorova, M., Stoyanova, I. Razrabotvane na rechnitsi na sastavnite leksikalni ediniti v balgarskiya ezik za tselite na kompyutarnata lingvistika. [Compiling Dictionaries of Bulgarian Multiword Expressions for the Purposes of Computational Linguistics – in Bulgarian]. // *Ezikovi resursi i tehnologii za balgarski ezik*, Academic Press Prof. Marin Drinov, 2014, 185–202.
- Villavicencio et al. 2004:** A. Villavicencio, A. Copestake, B. Waldron, F. Lambeau. The lexical encoding of MWEs. // *Proceedings of the ACL 2004 workshop on multiword expressions: Integrating processing*. Barcelona, Spain, 2004, 80–87.