

Derivational Morphology in an E-Dictionary of Serbian

Duško Vitas¹, Cvetana Krstev²

¹Faculty of Mathematics, University of Belgrade, Studentski trg 16, CS-11000 Belgrade

²Faculty of Philology, University of Belgrade, Studentski trg 3, CS-11000 Belgrade
(vitas|cvetana)@matf.bg.ac.yu

Abstract

In this paper we explore the relation between derivational morphology and synonymy in connection with an electronic dictionary, inspired by the work of Maurice Gross. The characteristics of this relation are illustrated by derivation in Serbian, which produces new lemmas with predictable meaning. We call this regular derivation. We then demonstrate how this kind of derivation is handled in text processing using a morphological e-dictionary of Serbian and a collection of transducers with lexical constraints. Finally, we analyze the cases of synonymy that include regular derivation in one aligned text.

1. Introduction

The use of finite-state automata for the representation of word families is discussed in (Gross, 1988) and illustrated by the set of words derived from the proper name *France*. The derived words are connected in a way that mirrors the syntactic relations maintaining the meaning, so that "...each meaning of the words should be described by means of an elementary sentence and related to other forms by formal transformations" (p. 40). The relation between derivational morphology and synonymy from the syntactic point of view is discussed in (Gross, 1997). Maurice Gross remarks that the synonymy relation is limited to the words that have "approximately the same meaning" and that *belong to the same grammatical category*, and it does not apply to the relations that emerge as a result of some derivational process. This further leads to the discussion of the relation between synonymy and derivational morphology on the level of an elementary sentence and the definition of one semantic equivalence relation between elementary sentences.

These ideas of Maurice Gross can be successfully explored on languages with rich morphological systems, particularly the derivational possibilities of morphological systems of Slavic languages. The significance of these ideas is stressed, for instance, by difficulties encountered in developing semantic networks of the type *WordNet* for Slavic languages (Bulgarian, Czech, Serbian). Since in *WordNet* one synset (set of synonyms) can consist only of literals that belong to the same PoS, it is necessary to enhance the *WordNet* by a set of synsets that encompass literals obtained by derivational processes and that are not lexicalized in English (Pala, 2005).

The analysis of texts aligned on the sentence level shows that derivational relations can produce synonymy relations in one language which do not exist in another. For instance, in Flaubert's novel *Bouvard et Pécuchet* only the string *Bouvard* occurs, while in its Serbian translation 16 different forms of the noun *Buvar* and its possessive adjectives *Buvarov* occur. Some examples are¹:

<tu><tuv>Celui de *Bouvard* était continu, sonore, ...</tuv>
<tuv>*Buvarov* smeh je bio pun, zvučan, ...</tuv></tu>

<tu><tuv>Mais la banlieue, selon *Bouvard*, était ...</tuv>
<tuv>Ali je predgrađe, po *Buvarovom* mišljenju, ...</tuv></tu>

Among derivational processes, particularly in Slavic languages, of special interest are processes for which the meaning of the derived word can be deduced from the meaning of the basic word. This class of derivational processes we will call *regular* (or *structural*) *derivation*. The role of regular derivation will be considered as a systematic means to realising synonymy, in the sense of (Gross, 1997).

Using the principles established in (Gross, 1988), (Gross, 1989), (Courtois et Silberztein, 1990), it is possible to construct electronic dictionaries for languages with rich morphology by developing the system of dictionaries of the type DELA in a similar way to how it has been done for French. However, in these dictionaries regular derivation poses serious methodological problems. If the results of regular derivation are incorporated into dictionaries in a systematic way, it not only multiplies the sizes of both DELAS and DELAF dictionaries and complicates their maintenance, but it also adds considerably to the text ambiguity. If, on the other hand, only those regularly derived lemmas that are represented in traditional (paper) dictionaries enter e-dictionary, it leads to serious inconsistencies, as will be shown in section 4. Finally, if regularly derived lemmas are considered as separate lemmas, the relations between the basic word and its derivatives are lost, and it will be impossible to effectively apply e-dictionaries to the analysis of synonymy relations, as suggested in (Gross, 1997).

In this paper we will illustrate the relation between regular derivation and e-dictionary by the case of Serbian. In section 2 the basic characteristics of Serbian that are related to the e-dictionary construction are given, while the dictionary itself is presented in section 3. The phenomenon of regular derivation in Serbian and some examples are given in section 4, while the way it is processed by transducers with lexical constraints using system Intex (Silberztein, 2004) is given in section 5. Some examples of using this kind of morphological grammars on parallel texts are given in section 6.

¹ The address of parallel corpora developed at Faculty of Mathematics is <http://www.korpus.matf.bg.ac.yu/pkorpus>

2. The Processing of Serbian

The contemporary standard Serbian language is one of the standard languages that have emerged from a common basis, namely the language that was called Serbo-Croatian until 1990 (Popović, 2003). From the computational point of view, certain characteristics of the Serbian language have to be taken into consideration before attempting to process Serbian written texts:

- a. *The use of two alphabets.* A text in Serbian can be written using either the official Cyrillic alphabet or the Latin alphabet, which is widely used. However, the transliteration procedure is not unique in any of the standard coding schemas. Due to this, the lexical resources are encoded in a way that neutralizes both the use of alphabet and encoding.
- b. *Phonologically based orthography.* One consequence of this is that a considerable number of morphophonemic processes are being reproduced in written texts. Moreover, the differences that exist between different variants (Ekavian and Ijekavian) of the standard language are recorded in written texts. For instance, the Serbian equivalents of the English words *child* and *girl* have two standard forms of the nominative singular: *dete*, *devojka* (Ekavian) and *dijete*, *djevojka* (Ijekavian). In this paper we will take only Ekavian into consideration.
- c. *The rich morphological system*, which is reflected both on the inflective and derivational level.
- d. *Free word order* of the subject, predicate, object and other sentence constituents, *special placement of enclitics*, and *complex agreement system*.

These characteristics have a direct impact on the acquisition, preparation, and processing of resources for the Serbian language and make the problem of disambiguation extremely difficult.

The results of the traditional description of the Serbian/Serbo-Croatian grammatical system of can rarely be applied to natural language processing needs. In particular, there are no traditional lexicographic resources that could be directly reused for these purposes. From the linguistic point of view, as the basis of the theoretical framework for the processing of Serbian the integral model of the syntax of Serbian is particularly important (Stanojčić et al., 2002). The concepts of lexicon-grammars and local grammars are also of considerable importance (Gross, 1997). On the technological level, the use of finite-state transducers (FST) for describing the interactions between text and dictionary is crucial, both for morphological and morphosyntactic descriptions.

3. The Morphological e-Dictionary of Serbian

The morphological electronic dictionary of Serbian has been developed by the NLP group, which works at the Faculty of Mathematics at the University of Belgrade. The model adopted for the construction of the morphological electronic dictionary of Serbian has been developed with the direct influence and fruitful help of Maurice Gross and LADL.

For this approach, the starting point is the empirically established and comprehensive classification of the inflective features of lexemes. Each inflective class is uniquely described by the assignment of a numerical code that describes the combination of its inflective endings. For instance, the class **N1** in Serbian designates the set of unmarked endings of the non-animated nouns of the first declension type. Such a classification is based on a factorization of the inflective paradigms, where the right factor describes in a unique way the characteristics of an inflective paradigm (Vitas et al. 2001) and enables a precise and automatic generation of all the forms of the inflective paradigm.

The system of morphological dictionaries consists of dictionaries of simple words (a sequence of alphabetical characters) and simple word forms, a dictionary of compounds (e.g. syntagmas), and a dictionary consisting of FST, used for recognition of unknown words, i.e. words that are not found in other dictionaries of the system but are derived from one or more lemmas that are in them.

As an example, one entry in the Serbian dictionary of simple word forms is:

digitalizaciju, digitalizacija.N600:fs4q

This entry assigns the lemma *digitalizacija* (Engl. digitalization) to the string of characters *digitalizaciju*. This lemma belongs to the inflective class **N600**, which encompasses the nouns of the third declension type that have unmarked endings. The code **fs4q** describes the word form *digitalizaciju* as the accusative case (**4**), singular (**s**) of the feminine gender (**f**), non-animate (**q**), of the lemma *digitalizacija*. The set of syntactic and semantic codes can be added to the lemma after the inflective class code. The following example illustrates the use of syntactic markers:

bojali,bojati.V542+Imperf+It+Ref:Gpm

The word form *bojali* is the plural (**p**) masculine gender (**m**) of the active past participle (**G**) of the verb *bojati se* (Engl. to be afraid), which belongs to the verb inflective class **V542**, and is imperfective (**Imperf**), intransitive (**It**), and reflexive (**Ref**). Similarly, semantic markers can be added, as in the example:

gvozdenoj,gvozden.A6+Mat :aefs3g:aefs7g

where *gvozden* (Engl. ferric, ferrous) is an adjective from the class **A6** with a material mark (**Mat**).

The advantage of such a structure of the e-dictionary is the possibility of consistently applying the theory of finite automata to corpus tagging and lemmatization. An excerpt from the dictionary is given in the Appendix. The present size of the Serbian dictionary of simple words is approximately 77,000 lemmas, while the dictionary of forms contains approximately 1,040,000 word forms. With the addition of special dictionaries of proper names, the number of lemmas reaches nearly 94,000, and the number of word forms 1,170,000. Construction of the dictionary of compounds is still in the initial phase. The main tool for the exploitation of e-dictionaries is the system Intex v. 4.33 (Silberztein, 2004).

4. Regular Derivation in Serbian

The phenomenon of regular derivation in Serbian encompasses various derivational relations that can lead to the change of PoS of the basic word, but do not necessarily do so. Among them, gender motion and amplification of meaning (diminutives and augmentatives) are particularly important and described in (Vitas, 2004). Beside them, for a large number of nouns possessive and relational adjectives exist, verbal nouns exist for all imperfective verbs, etc. The definitions used for these derived lemmas in traditional dictionary clearly show that they represent a special kind of derivation. Namely, in many cases they just mark that there is regular derivation, either by use of grammatical reference, e.g. *diminutive of*, or by use of synonymic reformulation of lemma.

basic lemma	Professor (professor)	lektor (language editor)	rektor (university rector)	pekar (baker)	N2+Hum
Gen	profesorka	lektorka	rektorka	<u>pekarka</u>	N661+Hum
ARel	profesorski	<i>lektorski</i>	<i>rektorski</i>	<u>pekarski</u>	A2+Rel
APos	<i>profesorov</i>	<i>lektorov</i>	rektorov	<u>pekarov</u> <i>/ev</i>	A1+Pos
Gen+APos	<i>profesorkin</i>	<i>lektorkin</i>	<i>rektorkin</i>	<u>pekarkin</u>	A1+Pos
Dim	profesorčić	<i>lektorčić</i>	<i>rektorčić</i>	<u>pekarčić</u>	N28+Hum
Augm	<i>profesorčina</i>	<i>lektorčina</i>	<i>rektorčina</i>	<u>pekarčina</u>	N601+Hum+MG

Table 1. Examples of regular derivations: Gen - gender motion, ARel - relational adjectives, APos - possessive adjectives, Dim - diminutives, Augm - augmentatives

Consider the examples given in Table 1, listed according to the only complete explanatory dictionary of Serbian (RMSMH, 1967). Each column in the Table represents a noun denoting a profession that belongs to the inflective class N2. In each table row examples of certain regular derivation are given. Lemmas in table cells that are not represented in this dictionary are given in italic, while those not represented in the corpus are underlined.

The unsystematic processing of lemmas in traditional lexicographic descriptions is illustrated by the examples from the Table 1. The basic lemmas, *profesor*, *lektor*, *rektor*, *pekar* can be expanded in a similar way by regular derivation, as seen in Table 1, but only some derived lemmas are represented in the explanatory dictionary, not necessarily those occurring in the corpus of contemporary Serbian². For instance, lemma *pekarov* (Engl. belonging to the baker) exists in (RMSMH, 1967), but does not occur in corpus. On the other hand, relational adjectives *lektorski* and *rektorski* are not in (RMSMH, 1967) but they both occur in the corpus. From those derived lemmas represented in the dictionary, definitions are given by regular patterns. For instance, the meaning of

possessive adjective of the noun *X* is defined as "*that which belongs to X*", while for gender motion two possibilities exist: "*woman that is X*" or "*the wife of X*".

The important feature of regular derivation is that if it is produced by suffixation, then for the basic lemma belonging to the certain inflective class it generates the derived lemma belonging to the inflective class precisely defined in the system of dictionaries DELA. So, for all the basic words from the class N2+Hum, words obtained by gender motion using the suffix *-ka* belong to the class N661+Hum³, possessive adjective belongs to the class A1, and relational adjective to the class A2. In general, regular derivations are the feature of inflective class rather than lemma itself, and thus they enhance it in a way that enables their classification in a similar manner as inflectional classes themselves. In other words, if the noun belongs to the class N2+Hum+Act, lemmas given in Table 1 can be derived for it, which independently of the lemma itself always belong to the same, pre-existing class.

The next important feature is that some suffixes that are used in regular derivation can change the meaning of the basic lemma. For instance, the nouns *kašičica*, *karanfilić* are derived from *kašika*, *karanfil* (Engl. spoon, carnation) but they acquire new meanings - tea spoon and clove tree respectively - which cancel the diminutive meaning. In these cases the relation of synonymy between the basic and derived word is broken and derivation cannot be regarded any more as a means to establishing synonymy.

As a consequence of these two properties, the methods to treat regular derivation in the system of e-dictionaries can be established. The first property implies that regular derivation will not introduce new classes in the description of inflective properties, while the second property determines the priorities of derived forms during the lexical recognition.

5. Transducers with lexical constraints

The special transducers with lexical constraints, the so called lexical transducers, incorporated in Intex v.4.33, allow the expression of morphological rules that govern word formation. The input of lexical transducers is used to recognize word forms, while the output is used to compute the corresponding lemma and other grammatical information. The computed output can have the same form as the entries in the DELAF dictionary and can be used in the same way during the lexical recognition. These transducers can be quite complex as they can perform tokenization of word forms into linguistic units. These linguistic units are established on the basis of imposed constraints which are expressed in terms of recognition by e-dictionaries. Furthermore, during the recognition process the values of the recognized linguistic units can be stored in variables, which can later be used for the computation of lemmas and grammatical categories.

² 25 million word corpus developed on the Faculty of Mathematics is used (Vitas, et al. 2003). It can be searched on the web address: <http://www.korpus.matf.bg.ac.yu/korpus/>.

³ Suffix *-ica* can also be used for gender motion. However, obtained lemmas are more characteristic for Croatian.

Numerous lexical transducers have been produced that make use of these derivational patterns to recognize possessive adjectives, diminutives, augmentatives, gender motion, prefixation, etc, analyzed in (Pavlović-Lažetić, et al. 2004).

6. Examples of regular derivations from the aligned text

We will illustrate the usage of regular derivation in Serbian with their occurrences in the aligned text *La Vénus d'Ille* by Prosper Mérimée. These examples show that the translator could choose between at least two solutions, of which one is the basic form and the others are obtained by regular derivation.

First, we will mention the examples of gender motion. The gender motion is in French considered part of the inflective paradigm, while in Serbian it is a separate lemma. In the chosen story the following examples of gender motion occur, and as they are not in the dictionary DELAF they are recognized by the transducers with lexical constraints:

fiancée, fiancé.N+z1:fs =
 verenica, {verenik, verenik.N10+Hum+Ek:ms1v}
 {verenica.N+Hum:fs1v}
 paysannes, paysan.N+z1:fp =
 selxanke, {selxanin, selxanin.N60+Hum:ms1v}
 {selxanka.N+Hum:fp1v}

This type of recognition enables the recognition of lemmas *verenica* (Engl. fiancée), *seljanka* (Engl. country woman) with patterns <verenik> and <selxanin>. As a consequence, all forms that correspond to the French patterns <fiancé> and <paysan> can be retrieved.

Diminutives which exist in Serbian, also recognized by the transducers with lexical constraints, are more complex. The transformation of diminutives can be described by the following rule in one generalization of regular notation:

$$[\langle \text{petit} \rangle \langle \text{N} \rangle]_{\text{fr}} \rightarrow [\langle \text{mali} \rangle \langle \text{N} \rangle]_{\text{sr}} \mid [\langle \text{N} + \text{Dem} \rangle]_{\text{sr}}$$

where subscripts fr and sr denote the French and Serbian constructions respectively. In the following examples the corresponding French and Serbian equivalents are in italics, while the synonymous forms in Serbian are given in brackets. The use of the diminutive is enforced by the diminishment of the noun, which is in French realized by the adjective *petit*:

les maisons de *la petite ville* de Ille = *gradić* (mali grad)
 ç'était un *petit vieillard* vert encore = *stračić* (mali starac)

The same French structure <petit> <N> is in other cases translated by the equivalent construction <mali> <N>, that is the diminutive form is not used, though possible:

Cette *petite bague*-là, ajouta-t-il... = *mali prsten* (prstenčić)
 nous lui ferons un *petit sacrifice* = *mala žrtva* (žrtvica)
 je vois sur le bras un *petit trou* = *rupica* (mala rupa)

However, in the case of

ma *petite femme* = moja *ženica*

the diminutive form has a hypocoristic meaning and, thus, it cannot be replaced by the form <mali> <N>. Similarly, in the example

qui séparait un *petit jardin* de un vaste carré = *mali vrt* or
baštica, but not *vrtić*

the construction <mali> <N> is used instead of the diminutive form *vrtić*, since the latter has lexicalized a different meaning *kindergarten*. If the translator had chosen *bašta* instead of *vrt* (Engl. garden), which are synonyms, then she/he could have used the diminutive form *baštica*, too.

de la première phalange de son *petit doigt* une *grosse bague*
 enrichie de diamants = *mali prst* (≠ *prstić*); *širok prsten*
 (*prstenčina*)

Le petit doigt (Engl. pinkie, little finger) is in French, as in Serbian, a lexicalized compound and it cannot be replaced by the diminutive *prstić*. *Une grosse bague* (Engl. large ring) could have been replaced by the augmentative.

In the above examples, the transformations in Serbian do not change the PoS. As we will show, different types of transformations are used when PoS is changed, as is the case with the derived adjectives and verbal nouns. For instance, it is possible to replace the possessive adjective derived from a certain noun, by the form of that same noun in genitive, according to the following simplified rule:

$$[\langle \text{N} \rangle]_{\text{fr}} \rightarrow [\text{A} + \text{Pos}]_{\text{sr}} \mid [\langle \text{N} : 2\text{s} \rangle]_{\text{sr}}$$

An example which illustrates this point on the chosen text is the use of the noun <marié> in the French text in the form *mariée, marié.N+z1:fs* (Engl. bride). In the Serbian translation, the possessive adjective *mladin* corresponds to it. As it is recognized a transducer with lexical constraint, two lemmas can be associated to it: *mlada.N* (Engl. bride) or *mladin.A+Pos* (Engl. belonging to the bride):

mladinog, {*mlada.N726+Hum:fs1v*} {*mladin.A+Pos:adms2g*}

The examples from the text are:

<tu><tuv>... blanc ... qu'il venait de détacher de *la cheville*
de la mariée.</tuv>

<tuv>... belu traku koju je odvezalo sa *mladinog*
članka.</tuv></tuv>

<tu><tuv>... de s'introduire la nuit dans *la chambre de la*
mariée.</tuv>

<tuv>... da se preko noći uvuku u *mladinu sobu.*</tuv></tuv>

In both cases, the form of the possessive adjective could have been replaced by the form of the corresponding noun in the genitive case *članka mlade* (ankle of the bride), *sobe mlade* (room of the bride). The use of relational adjective follow the similar pattern:

$$[\langle \text{N} \rangle]_{\text{fr}} \rightarrow [\text{A} + \text{PosQ}]_{\text{fr}} \mid \{\text{PREP}\}_{\text{fr}} [\langle \text{N} \rangle]_{\text{sr}}$$

as in the example of the following transformation:
serrurier.N_{fr} → *bravarski.A+PosQ_{fr}* (derived from
bravar.N, Engl. locksmith)

(il paraît que c'était un *apprenti serrurier*):
apprenti.N+z1:ms serrurier.N+z1:ms
 (verovatno je bio *bravarski šegrt*):

bravarski.A+PosQ:adms1g:šegrt.N+Hum:ms1v

A similar transformation is given in the example:

rien ne dispose mieux que *l'air vif des montagnes*
nema bolje stvari za apetit od *planinskog vazduha*

in which the sequence *air*.N:ms *vif des montagne*.N:fp is translated with the sequence *planinskog*.planinski.A+PosQ:adms2g *vazduha*.vazduh.N:ms2q. A different case is :

...les ruisseaux de la montagne y forment des mares infectes.</seg>

...potoci sa planine prave odvratne bare.</seg>

where the literal translation *potoci*.potok.N:mp1q *sa planine*.planina.N:fs2q is used instead of the synonymous *planinski potoci*.

Finally, consider the transformation [$\langle V \rangle$]_{fr} → [$\langle V \rangle$]_{sr} | [N+VN]_{sr}, where VN denotes verbal nouns One example is given by the translation pair:

<tu><tuv>... j'avais entendu force allées et venues ..., les portes s'ouvrir et se fermer, </tuv>
<tuv>... čuo sam mnogobrojne korake tamo-amo ..., otvaranje i zatvaranje vrata,</tuv></tu>

where the translation equivalents are:

ouvrir.V+z1:W → otvaranje.N300+VN:ns1q
fermer.V+z1:W → zatvaranje.N300+VN:ns1q

This could be translated by the synonymous equivalents:

≈ kako se otvaraju i zatvaraju vrata
≈ da se vrata otvaraju i zatvaraju

7. Conclusions

We have shown that lemmas derived by structural derivation in Serbian can be recognized in text without being explicitly recorded in a dictionary. As a consequence it is possible to reorganize the complete inventory of lemmas in Serbian dictionaries, both monolingual and multilingual. Finite state transducers enable a precise classification of processes of structural derivation in a way similar to the classification of inflective phenomena.

8. References

- Courtois, Blandine; Max Silberztein (eds.) 1990. Dictionnaires électroniques du français. *Langue française* 87. Paris: Larousse
- Gross, Maurice. 1988. The Use of Finite Automata in the Lexical Representation of Natural Languages. In: Gross, M., Perrin, D. (Eds.): *Electronic Dictionaries and Automata in Computational Linguistics*, LNCS 337, Berlin: Springer, pp 34-50
- Gross, Maurice. 1989. La construction de dictionnaires électroniques. *Annales des télécommunications*. 44 (1-2), pp. 4-19
- Gross, Maurice. 1997. Synonymie, morphologie dérivationnelle et transformations. *Langages* 128. Paris: Larousse, 72-90
- Pavlović-Lažetić, Gordana., Vitas, D., Krstev, C. 2004. Towards Full Lexical Recognition. In Sojka, P., Kopeček, I., Pala, K. (eds.): *Text, Speech and Dialogue TSD 2004*, LNAI 3206, Berlin: Springer, pp. 179-186

- Popović, Ljubomir. 2003. Od srpskohrvatskog do srpskog i hrvatskog standardnog jezika: srpska i hrvatska verzija. Wien, *Wiener Slavistischer Almanach*, 57, 201-224.
- Pala, Karel.; Sedláček, R. 2005. Enriching WordNet with Derivational Subnets. *CICLing 2005* (Abstract). www.cicling.org/2005/Abstracts/
- RMSMH (1967). *Rečnik srpskohrvatskoga književnog jezika*, vol. 1-6, Beograd-Zagreb: Matica Srpska, Matica Hrvatska
- Silberztein, Max. 1993. Le dictionnaire électronique et analyse automatique de textes: Le système INTEX, Paris: Masson.
- Silberztein, Max. 2004. INTEX Manual, v. 4.33. (<http://intex.univ-fcomte.fr/downloads/Manual.pdf>)
- Stanojčić, Živojin, and Popović, Lj. 2002. *Gramatika srpskoga jezika*. Beograd, Zavod za udžbenike i nastavna sredstva.
- Vitas, Duško, Krstev, Cvetana, and Pavlović-Lažetić, Gordana. 2001. The Flexible Entry. In: Zybatow, G. et al. (eds.): *Current Issues in Formal Slavic Linguistics*. Leipzig: University of Leipzig. 461-468.
- Vitas, Duško 2004. Morphologie dérivationnelle et mots simples. Le cas du serbo-croate. Leclère, Ch., Laporte, E., Piot, M., Silberztein, M. (Eds.): *Lexique, Syntaxe et Lexique-Grammaire / Syntax, Lexis & Lexicon-Grammar. Papers in honour of Maurice Gross. Linguisticae Investigationes Supplementa* 24. Amsterdam/Philadelphia: Benjamins. 629-639.

Appendix

abazxur, abazxur.N1:ms1q:msA4q
abazxura, abazxur.N1:ms2q:mpG2q
.....
abazxure, abazxur.N1:ms5q:mp4q
abazxuri, abazxur.N1:mp1q:mp5q
abazxurima, abazxur.N1:mp3q:mp6q:mp7q
pevati, pevati.V1:W
pevate, pevati.V1:Pyp
pevaju, pevati.V1:Pzp
.....
pevajucxi, pevati.V1:S
pevavsxi, pevati.V1:X
pevao, pevati.V1:Gsm
.....
pevacxu, pevati.V1:Fxs
pevacxesx, pevati.V1:Fys
pevan, pevati.V1:Tms
pevani, pevati.V1:Tmp
iko, iko.PRO12+Indef+ProN+Sr:s1v
ikog, iko.PRO12+Indef+ProN+Sr:s2v:s4v
ikoga, iko.PRO12+Indef+ProN+Sr:s2v:s4v
dve, dva.NUM02+v2+Ek:fp1g:fp4g:fp5g
dveju, dva.NUM02+v2+Ek:fp2g
dvema, dva.NUM02+v2+Ek:fp3g:fp6g:fp7g