

The automation of speech parts inflexion process

E.Boian A.Danilchenco L.Topal

Abstract

The article describes the Roumanian language words automat inflexion system utilized by Roumanian texts spelling checker ROMSP. It contains the nouns and adjectives declination, the verbs conjugation details.

1 Introduction

At present the texts editors, dotted with the function to correct orthographically, as well as an automated translation system etc. have a large expansion. Their work is based on a special dictionary. The dictionary's structure for a highly inflexional language is a topical and difficult problem. Two methods to keep information are known : declarative (to memorize all flexions) and procedural (to obtain the necessary flexions in a special way from one or some base words). Using the declarative method to complete the dictionary it is necessary to obtain all word-forms. There is an analogous problem for the procedural method to derive a new word, for which it is necessary to establish inflexions rules.

It is very difficult to introduce manually all word-forms with the morfologic information (which is necessary for series of applications). In this way it is naturally to try to automatize this process.

Our article describes the Roumanian language words automat inflexion system utilized by Roumanian texts spelling checker ROMSP [1]. Special attention is given to nouns and adjectives declination, and

verbs conjugation because they generate a lot of flexions (for example, 12 for nouns, 20 for adjectives, 35 for verbs).

Being not so numerous pronouns, articles and numerals are not so interesting, moreover they may be introduced in the dictionary preliminary. Evidently that the problem of the complete formalization of the inflection process is very difficult, and we try to solve it partially, in special cases we ask the users supplementary information about the morfological categories, the consonant and vowel alternations.

The investigations of nouns, adjectives and verbs inflexions regularities are our first subject. The consonant and vowel alternations are established in the next section. The subjects of the last two sections are the nouns, adjectives and verbs procedures of automation inflexion.

We use the word "affix", which denotes the notion of both a suffix and an ending, or either a suffix or an ending in accordance with the context.

All grammatical notions used here are described in [2, 3, 4, 5]. The program complex was implemented in TURBO-PASCAL 6.0.

2 Speech parts inflecting regularities

Six parts of speech (of eleven) can be inflected. Nouns, adjectives, articles, numerals and pronouns are declined in accordance with case, number and form. Verb is conjugated in accordance with tense, mood, person etc. The rest – adverb, preposition, conjunction, particle and interjection – is invariable [2, 3, 4]. We investigate in more details about inflecting regularities of nouns, adjectives and verbs.

2.1 The noun

As a part of speech the noun modifies its form in accordance with case and number, but it is invariable in gender [5].

For nouns the following morfologic categories are characteristic: the gender (*GF* – feminine, *GM* – masculine and *NN* – neuter noun), the case (*CNA* – nominative-accusative, *CGD* – genitive-dative, *CV* –

vocative), the number (*NS* – singular and *NP* – plural, singularia-tantum – the nouns have only singular number and pluralia-tantum – those have only plural number), the form (*FI* – indefinite and *FD* – definite), common and proper.

The registered affixes after the examination of the Roumanian nouns are grouped in a lot of series.

The affixes series characterized exclusively for the feminine nouns are the following:

Figure 1

affixes series	Case	NS	NP	Example	
eancă iancă	CNA	eancă	ence	moldoveancă	(Moldavian woman)
	CGD	iancă	ience	italiancă	(Italian woman)
		ence	ence		
		ience	ience		
eală	CNA	eală	eli	căptușeală	(lining)
ială		ială	ieli	sfială	(shyness)
easă		easă	ese	mireasă	(bride)
eară		eară	eri	seară	(evening)
eață		eață	ețe	dimineată	(morning)
oară	CGD	oară	ori	vioară	(violin)
		eli	eli		
		ieli	ieli		
		ese	ese		
		eri	eri		
		ețe	ețe		
ori	ori				
oare	CNA	oare	oare	învățătoare	(teacher)
	CGD	oare	oare		
eană	CNA	eană	ene	moldoveană	(Moldavian woman)
iană	CGD	iană	iene	italiană	(Italian woman)
		ene	ene		
		iene	iene		

The automation of speech parts...

affixes series	Case	NS	NP	Example	
oaie	CNA	oaie	oaie	căsoaie	(shed)
		oaie	oi	foaie	(leaf)
	CGD	oaie	oaie		
		oi	oi		
ană	CNA	ană	ăni	cană	(decanter)
scă		scă	ști	maskă	(mask)
șcă		șcă	ște	gălușcă	(dumpling)
șcă		șcă	ști	cloșcă	(brood hen)
ară		ară	ări	gară	(station)
ară		ară	eri	vară	(summer)
	CGD	ăni	ăni		
		ști	ști		
		ște	ște		
		ști	ști		
		ări	ări		
		eri	eri		
eză	CNA	eză	eze	burgheză	(bourgeois woman)
	CGD	eze	eze		
aie	CNA	aie	ăi	baie	(bath)
	CGD	ăi	ăi		
ică	CNA	ică	ici	bunică	(grandmather)
		ică	ele	rîndunică	(swallow)
	CGD	ici	ici		
		ele	ele		
ață	CNA	ață	ețe	povață	(piece of advice)
		eață	eți	ceață	(mist)
	CGD	ețe	ețe		
		eți	eți		
ate	CNA	ate	ăți	activitate	(activity)
	CGD	ăți	ăți		
ea	CNA	ea	ele	măsea	(tooth)

affixes series	Case	NS	NP	Example	
ia		ia	iele	nuia	(rod)
	CGD	ele	iele		
ie	CNA	ie	i	femeie	(woman)
ie		ie	ii	pălărie	(hat)
ee		ee	ei	idee	(idea)
oe		oe	oe	canoe	(canoe)
	CGD	i	i		
		ii	ii		
		ei	ei		
	oe	oe			
le	CNA	le	i	cale	(road)
	CGD	i	i		
a	CNA	a	ale	para	(money)
	CGD	ale	ale		
consonant +e	CNA	e	i	lume	(world)
	CGD	i	i		

These affixes series characterize numerous noun groups. The series 7 and 9 are an exception: they specify one limited number of words.

The affixes series that belong to the masculine nouns are :

Figure 2

affixes series	Case	NS	NP	Example	
stru	CNA	stru	ștri	ministru	(minister)
	CGD	stru	ștri		
ean	CNA	ean	eni	moldovean	(Moldavian man)
ian		ian	ieni	italian	(Italian man)
eag		eag	egi	moșneag	(oldman)
iag		iag	iegi	sfoiag	(worm)
eac		eac	eci	dovleac	(pumpkin)
iac		iac	ieci	liliac	(lilac)

The automation of speech parts...

affixes series	Case	NS	NP	Example	
	CGD	ean ian eag iag eac iac	eni ieni egi iegi eci ieci		
înt	CNA CGD	înt înt	inți inți	sfint	(saint)
vowel+u	CNA CGD	vowel+u vowel+u	i i	leu	(lion)
ez ăr ăn ăt st	CNA CGD	ez ăr ăn ăt st ez ăr ăn ăt st	ezi eri eni eți ști ezi eri eni eți ști	burghez luceafăr mesteacăn făt arbust	(bourgeois) (star) (birch) (beby) (bush)
l	CNA CGD	l l l l	i li i li	cal nobil	(horse) (nobleman)
u	CNA CGD	u u	i i	codru	(forest)
consonant	CNA CGD	consonant consonant	i i	pom	(tree)
i	CNA CGD	i i	i i	arici	(hedgehog)
e	CNA	e	i	câine	(dog)

affixes series	Case	NS	NP	Example	
	CGD	e	i		
a	CNA	a	a	boa	(boa)
ă		ă	i	popă	(pope)
o		o	i	picolo	(boy)
	CGD	a	a		
		ă	i		
		o	i		

The affixes series 11 appears in paradigm of only some nouns (*popă*, *pasă*, *tată*, *vodă*).

The affixes series which characterize the neuter nouns are:

Figure 3

affixes series	Case	NS	NP	Example	
vowel+u	CNA	u	uri	tablou	(picture)
	CGD	u	uri		
consonant	CNA	consonant	uri	tren	(train)
	CGD	consonant	uri		
u	CNA	u	uri	lucru	(work)
	CGD	u	uri		
o	CNA	o	ale	caro	(caro)
	CGD	o	ale		
e	CNA	e	e	nume	(name)
	CGD	e	e		

The series which specified any limited groups are 4 and 5.

The vocative case characterizes the possibility to be expressed by a lot of affixes. The masculine and neuter nouns *NS* vocative affixes may be : "e" , "ule" , "le" or an affix homonymous with that one of *NS CNA FI*.

For example : "e" – *băiete*, *frate*, *colege*, *doamne*;

"ule" – *băiatule, trenule*;

"le" – *codrule*;

the word "soare" affix coincides with that of *NS CNA FI*.

The affixes which characterize the feminine nouns vocative are :
"o" and affix homonymous with that of *NS CNA FI*.

Example : "o" – *fato, vicleano*;

word's affix "fată" coincides with that of *NS CNA FI*.

The *NP* vocative affix for all genders is "lor" or homonymous with that of *NP CNA FI*.

Example : "lor" – *fraților, doamnelor*;

word affix for "băieți" and "păduri" coincides with that of
NP CNA FI.

Examining the series of affixes discussed above we can establish, in the function of gender in the process of the Roumanian nouns inflexion, a lot of homonymy types A – C, namely:

Homonymies types	a	b	c	d	e	f	for noun (gender)
A	+	+	+	+	+		feminine
B		+	+	+	+	+	masculine
C	+	+				+	neuter noun

where a) – f) are the following homonymies :

a) ($CN = CA$);

b) ($CG = CD = CN = CG = CD = CA$) ;

c) ($CN = CG = CD = CA$);

d) ($CN = CG = CD = CA$);

e) ($CG = CD$);

f) $(CN = CG = CD = CA) = (CN = CG = CD = CA)_{NP}$.

All varieties of nouns examined above belong to the common nouns set.

The declination and articulation of proper nouns coincide with the process of inflexion corresponding to common nouns with analogous morfologic categories [4].

So, nouns classification was performed in the function of the affixes obtained. This classification is pictured in Figures 1–3, which served as a base for nouns inflexion program implementation.

2.2 The adjective

Adjectives decline in accordance with the defined noun in gender, number and case.

M.Manoliu in [2] classifies adjectives in 19 big groups of inflexion. Our classification is more detailed, it contains 26 groups. Adjective flexions forming process is in accordance with the following picture :

Figure 4

affixes series	Gender	Case	NS	NP	Example
icel	GM	CNA	icel	icei	bunicel (grandfather)
		CGD	icel	icei	
	GF	CNA	icică	icele	
		CGD	icele	icele	
iesc	GM	CNA	iesc	iești	crăiesc (emperor)
		CGD	iesc	iești	
	GF	CNA	iească	iești	
		CGD	iești	iești	
ean	GM	CNA	ean	eni	moldovean (Moldavian man)
		CGD	ean	eni	
	GF	CNA	eană	ene	
		CGD	ene	ene	
eancă	GF	CNA	eancă	ence	moldoveancă

The automation of speech parts...

affixes series	Gender	Case	NS	NP	Example
ian	GM	CGD	ence	ence	(Moldavian woman)
		CNA	ian	ieni	italian
iancă	GF	CGD	ian	ieni	(Italian man)
		CNA	iană	iene	
	GF	CGD	iene	iene	
		CNA	iancă	ience	italiancă
esc	GM	CGD	ience	ience	(Italian woman)
		CNA	esc	ești	obștesc
ist	GF	CGD	esc	ești	(public)
		CNA	ească	ești	
	GM	CGD	ești	ești	
		CNA	ist	iști	marxist
GF	CGD	ist	iști	(marxist)	
	CNA	istă	iste		
ixt	GM	CGD	iste	iste	
		CNA	ixt	icști	mixt
	GF	CGD	ixt	icști	(joint)
		CNA	ixtă	ixte	
tor	GM	CGD	ixte	ixte	
		CNA	tor	tori	silitor
GF	CGD	tor	tori	(diligent)	
	CNA	toare	toare		
ept	GM	CGD	toare	toare	
		CNA	ept	epti	destept
	GF	CGD	ept	epti	(clever)
		CNA	eaptă	epte	
înt	GM	CGD	epte	epte	
		CNA	înt	inți	sfint
	GF	CGD	înt	inți	(holy)
		CNA	întă	inte	
CGD	inte	inte			

affixes series	Gender	Case	NS	NP	Example
ru	GM	CNA	ru	ri	acru
		CGD	ru	ri	(sour)
	GF	CNA	ră	re	
		CGD	re	re	
lu	GM	CNA	lu	li	simplu
		CGD	lu	li	(simple)
	GF	CNA	lă	le	
		CGD	le	le	
consonant+e	GM	CNA	e	i	mare
		CGD	e	i	(big)
	GF	CNA	e	i	
		CGD	i	i	
consonant+u	GM	CNA	u	ii	roșu
		CGD	u	ii	(red)
	GF	CNA	ie	ii	
		CGD	ii	ii	
consonant+i	GM	CNA	i	i	mehenghi
		CGD	i	i	(skilful)
	GF	CNA	e	e	
		CGD	e	e	
oi	GM	CNA	oi	oi	vioi
		CGD	oi	oi	(lively)
	GF	CNA	oaie	oaie	
		CGD	oaie	oaie	
ei	GM	CNA	ei	ei	rotofei
		CGD	ei	ei	(plump)
	GF	CNA	ee	ee	
		CGD	ee	ee	
ez	GM	CNA	ez	ezi	portughez
		CGD	ez	ezi	(Portuguese)
	GF	CNA	eză	eze	

The automation of speech parts...

affixes series	Gender	Case	NS	NP	Example
et	GM	CGD	eze	eze	
		CNA	et	eți	încet
	GF	CGD	et	eți	(slow)
		CNA	eată	ete	
ăn	GM	CGD	ete	ete	
		CNA	ăn	eni	zdravăn
	GF	CGD	ăn	eni	(strong)
		CNA	ănă	ene	
eț	GM	CGD	ene	ene	
		CNA	eț	eți	îndrăznet
	GF	CGD	eț	eți	(bold)
		CNA	eață	ețe	
ăț	GM	CGD	ețe	ețe	
		CNA	ăț	eți	șugubăț
	GF	CGD	ăț	eți	(waggish)
		CNA	eață	ețe	
uu	GM	CGD	ețe	ețe	
		CNA	uu	ui	continuu
	GF	CGD	uu	ui	(permanent)
		CNA	uă	ue	
monosilabic	GM	CGD	ue	ue	
		CNA	eu	ei	greu
	GF	CGD	eu	ei	(heavy)
		CNA	ea	ele	
ău	GM	CGD	ele	ele	
		CNA	ău	ăi	rău
	GF	CGD	ău	ăi	(bad)
		CNA	ea	ele	
polisilabic		CGD	ele	ele	
		CNA	ele	ele	

affixes series	Gender	Case	NS	NP	Example
eu	GM	CNA	eu	ei	instantaneu
		CGD	eu	ei	(instantaneous)
	GF	CNA	eie	eie	
		CGD	eie	eie	
ău	GM	CNA	ău	ăi	nătărău
		CGD	ău	ăi	(blockhead)
	GF	CNA	ie	ii	
		CGD	ii	ii	
monosilabic	GM	CNA	ou	oi	nou
		CGD	ou	oi	(new)
	GF	CNA	ouă	oi	
		CGD	oi	oi	
ui	GM	CNA	ui	ui	amărui
		CGD	ui	ui	(bitterish)
	GF	CNA	uie	ui	
		CGD	ui	ui	
ai	GM	CNA	ai	ai	bălai
		CGD	ai	ai	(fair)
	GF	CNA	aie	aie	
		CGD	aie	aie	
ce	GM	CNA	ce	ci	rece
		CGD	ce	ci	(cold)
	GF	CNA	ce	ci	
		CGD	ci	ci	
ci	GM	CNA	ci	ci	dibaci
		CGD	ci	ci	(dexterous)
	GF	CNA	ce	ce	
		CGD	ce	ce	

The automation of speech parts...

affixes series	Gender	Case	NS	NP	Example
el	GM	CNA	el	ei	mititel (small)
		CGD	el	ei	
	GF	CNA	ică	ele	ele
		CGD	ele	ele	
sc	GM	CNA	sc	ști	brusc (sudden)
		CGD	sc	ști	
	GF	CNA	scă	ște	ște
		CGD	ște	ște	
or	GM	CNA	or	ori	ușor (light)
		CGD	or	ori	
	GF	CNA	oară	oare	oare
		CGD	oare	oare	
a	GM	CNA	a	a	gata (ready)
		CGD	a	a	
	GF	CNA	a	a	a
		CGD	a	a	
e	GM	CNA	e	e	cumsecade (decent)
		CGD	e	e	
	GF	CNA	e	e	e
		CGD	e	e	
o	GM	CNA	o	o	caro (of diamonds)
		CGD	o	o	
	GF	CNA	o	o	o
		CGD	o	o	
l	GM	CNA	l	i	gol (naked) nobil (noble)
			l	li	
		CGD	l	i	
			l	li	
	GF	CNA	lă	le	le
		CGD	le	le	
consonant	GM	CNA	consonant	i	bun

affixes series	Gender	Case	NS	NP	Example
		CGD	consonant	i	(good)
	GF	CNA	ă	e	
		CGD	e	e	

In the adjective inflexion process 10 forms appear for *GF*, as well as for *GM*. If an adjective is before a noun, then it obtains a definite article and it declines as the corresponding noun. If an adjective is after a noun, then it changes in accordance with the form (definite or indefinite) and the number. Only feminine adjectives have special forms of case (*CGD*).

The adjective affix for vocative coincides with the corresponding forms of masculine and feminine nouns.

Examining the adjective affixes described above we can establish, in the function of gender, a lot of homonyms types A – B, namely :

Homonymies types	a	b	c	d	e	f	for adjective (gender)
A	+		+	+	+	+	feminine
B	+	+	+	+	+		masculine

where a) – f) is the homonyms as for nouns (see 2.1.)

The results of the analysis performed were utilized for the implementation of the adjective inflexion program.

2.3 The verb

The program complex for execution of generating verb inflexion forms a words-form list in accordance with the infinitive of a verb. This list contains Imperativ, Participiu, Gerunziu flexions and verb flexions of those tenses only, for which changing in accordance with person is characteristic (i.e. Indicativ Prezent, Conjunctiv Prezent, Imperfect, Perfectul Simplu, Mai Mult ca Perfectul).

There are a lot of special literature where different schemes of verb conjugation are analyzed. For example, in [6] the verb set is divided

into 5 types, each of which has its characteristic features. Each type is divided into some classes, characterized by its affix in accordance with the indicated mood or tenses of verbs. Each class (or type, if the division into classes does not exist) in its turn is divided into schemes, in the function of specific vowel alternation, of word roots, of difference of affixes etc. For example, the first type of verbs contains the verbs of the first grammatical group (the verbs with the ending "a" for the infinitive) and some verbs of the second grammatical group (the verbs with the ending "ea" for the infinitive) which conjugate with the suffix "ez" (for Indicativ Present, Singular, and first person) for example, *a vegea* (to watch). This type of verbs has the following characteristics:

- the presence of character "a" at the end of a word in some cases (Infinitiv, Indicativ Present and Conjunctiv Present Plural for some persons, Imperfect, Perfectul Simplu and Mai Mult ca Perfectul have the form : theme + "a" + affix);
- Indicativ Present and Conjunctiv Present Singular have the form : theme + suffix (0/"ez") + affix, where 0 means that they can be without a suffix;
- Gerunziu has the form : theme + "înd", (note, however, that this type contains the verbs *a muia* (to soak), *a înfœia* (to swell), *a tăia* (to cut) etc., which have the corresponding Gerunziu : *muînd*, *înfœind*, *tăînd* etc.).

The two classes, which correspond to the type of the verb above mentioned, differ by the suffixes, which obtain Indicativ Present Singular (the suffixes 0 and "ez" for the corresponding verb from the first and the second classes). The first class is divided into 15 schemes, the second – into 4 schemes.

To clearly imagine these schemes, we will study one representative (Figure 5) in every of 15 schemes from the first class, and its conjugation in Indicativ Present Singular, which can reflect the difference of conjugation between these schemes. The consonant and vowel alternations are indicated in the order of their appearance in a word. The

flexion is obtained by cutting the ending "a" from the infinitive, and executing the set of alternations, and by adding the necessary suffixes.

Figure 5

Verb	pers.I		pers.II		pers.III	
	alter.	affix	alter.	affix	alter.	affix
a cînta (to sing)	—	—	t→Ț	i	—	ă
a căuta (to seek)	ă→a	—	ă→a, t→Ț	i	ă→a	ă
a arăta (to show)	—	—	t→Ț	i	ă→a	ă
a apăra (to defend)	—	—	ă→e	i	—	ă
a apăsa (to push)	—	—	ă→e, s→ș	i	ă→a	ă
a căpăta (to obtain)	ă→a	—	ă→a, ă→e t→Ț	i	ă→a	ă
a asemăna (to compare)	—	—	ă→e	i	e→ea	ă
a boteza (to baptize)	—	—	—	i	—	e
a convoca (to convene)	—	—	—	i	o→oa	ă
a juca (to danse)	u→o	—	u→o	i	u→oa	ă
a afla (to learn)	—	u	—	i	—	ă
a încuia (to lock)	—	—	—	—	—	e
a tăia (to cut)	ă→a	—	ă→a	—	ă→a	e
a înfoia (to swell)	—	—	—	—	o→oa	e

Verb	pers.I		pers.II		pers.III	
	alter.	affix	alter.	affix	alter.	affix
a muia (to soak)	u→o	—	u→o	—	o→oa	e

Note: in cases 9 and 10 for II person there is the alternation [k] → [ʃh], which has the phonetic nature, and graphically has an invariant form.

Figure 5 shows that every one of these schemes has its specific form of conjugation.

Analogous is the situation with other types, classes and schemes of conjugation. The second grammatical group (which contains a verb with the ending "ea") contains one class, which is divided into 3 schemes, the third grammatical group (with the ending "e") contains 2 types, 3 classes and 19 schemes, the fourth grammatical group includes one type, which consists of 4 classes and has 15 schemes. As we see, in [6] the regular verbs set is divided into 56 groups and the verbs list (about 4000) containing the conjugation scheme number is annexed. Although serving as an excellent base for the necessary algorithms elaboration, the work [6] has some inaccuracies. For example, it has no scheme for the verbs such as *a apropia* (to draw), *a învia* (to revive). It proposes that these verbs are conjugated in accordance with the scheme of the verb *a încuia* (to lock), which is not correct, because Indicativ Present, I-st person is *eu încui* (I lock), but – *eu învii* (I revive). With this consideration, Figure 5 is completed with the following string:

N	Verb	pers.I		pers.II		pers.III	
		alter.	affix	alter.	affix	alter.	affix
16	a învia (to revive)	—	i	—	i	—	e

Imperativ Singular for the verbs *a minți* (to lie), *a simți* (to feel), *a crede* (to believe) etc. has a single form *mințe*, *simțe*, *crede* in [6], but these verbs have also other forms *minți*, *simți*, *crezi*. The verb *a insufla*

(to inspire) wrongly is classified in accordance with the conjugation scheme for the verb *a cînta* (to sing) etc.

There are some other works on this aspect. For example, [7], where more verb conjugation schemes are indicated (about 300) in the majority of cases, if there are a lot of forms for one morfologic category (including diverse regionalisms and rarely utilized forms). For example, the verbs *a apropia* (to bring) and *a învia* (to revive) in [7] are classified into different conjugation schemes, because the verb *a apropia* permits the forms *eu apropî* (this is a regionalism) and *eu apropîi*, but the verb *a învia* has only the form *eu învîi*. From [7] we extract both forms of Imperativ Plural for some verbs such as the verb *a naște* (to give birth) (e.g. : *nașteți* and *nășteți*) and other specification.

In spite of the said, the work [6] was used as the base. For us types and classes description is not of a great importance. We have used in full the information about including a verb into one or other grammatical group, which is determined in procedural way, from the verb ending in the infinitive, but for the verbs from grammatical groups I-st and IV-th the information is necessary whether the verb conjugates with or without a suffix. Essentially it is the information whether the verb is personal or impersonal. In the function of an answer it generates 35 or 12 flexions respectively. The answers are obtained from the dialog with the user of the system. The information about including a verb in any group corresponds to including it in any type in accordance with [6], but the division of grammatical group into parts (with or without suffix) corresponds to classes. Thus, we will investigate 6 classes of regular verbs and one of irregular (*a fi* (to be), *a lua* (to take), *a avea* (to have) etc.). For each class, the following procedures were specified:

1. Word-Verb1 , that generates the flexions for the first grammatical group verbs, and that are conjugated without a suffix;
2. Word-Verb1ez, that generates the flexions for the first grammatical group verbs, and that are conjugated with a suffix;
3. Word-Verb2 – for the second grammatical group verbs;
4. Word-Verb3 – for the third grammatical group verbs;

5. Word-Verb4 – for the forth grammatical group verbs, which are conjugated without a suffix ;
6. Word-Verb4esc – for the forth grammatical group verbs, which are conjugated with a suffix;
7. Word-Verb-Nereg – for irregular verbs.

It is clear, that our division of verbs into 7 classes overlaps the one from [6], and it is possible to transfer the division into schemes from [6] into classes in accordance with our division. These schemes are used in our above mentioned procedures.

3 The consonant and vowel alternations in speech parts inflexion process

One of the declination and conjugation particularities characteristic of the Roumanian language is consonant and vowel alternations. They are realized from variations of roots and affixes.

In the process of noun and adjective declination for consonant and vowel alternations, the following regularities were established for indefinite

FI						
NS			NP			
	CNA	CGD	CV	CNA	CGD	CCV
MN	—	—	A0	A3	A3	A3
FN	—	—	—	A1	A1	A1
NN	—	A2	—	A2	A2	A2
MA	—	—	—	A4	A4	A4
FA	A5	A6	—	A7	A7	A7
		A7				

and definite forms.

FD						
	NS			NP		
	CNA	CGD	CV	CNA	CGD	CCV
MN	—	—	A0	A3	A3	A3
FN	—	—	—	A1	A1	A1
NN	—	A2	—	A2	A2	A2
MA	—	A2	—	A4	A2	A2
FA	A5	A6	—	A7	A7	A7
		A7				

where A0 is the set of alternation for *CV* (affix - "le").
The sets of alternations are the following :

- for masculine nouns MN (A1) :
 $t \rightarrow \text{ț}$, $d \rightarrow z$, $s \rightarrow \text{ș}$, $x \rightarrow \text{cș}$, $\text{str} \rightarrow \text{ștr}$, $z \rightarrow j$, $l \rightarrow 0$ (where 0 is empty affix), $\text{sc} \rightarrow \text{șt}$, $\text{ea} \rightarrow e$, $\text{ean} \rightarrow \text{en}$, $\text{ăt} \rightarrow \text{eț}$, $\text{înt} \rightarrow \text{int}$, $\text{eaz} \rightarrow \text{ej}$, $\text{iac} \rightarrow \text{iec}$, $\text{ăn} \rightarrow \text{en}$, $\text{ian} \rightarrow \text{ien}$, $\text{ar} \rightarrow \text{ăr}$, $\text{ia} \rightarrow \text{ie}$;
- for feminine nouns FN (A2) :
 $\text{sc} \rightarrow \text{șt}$, $\text{șc} \rightarrow \text{șt}$, $\text{ea} \rightarrow e$, $a \rightarrow e$, $\text{ean} \rightarrow \text{en}$, $\text{ian} \rightarrow \text{ien}$, $\text{ăn} \rightarrow \text{en}$, $\text{ead} \rightarrow \text{ez}$, $\text{ăt} \rightarrow \text{eț}$, $\text{ăr} \rightarrow \text{er}$, $\text{ăd} \rightarrow \text{ed}$, $\text{ic} \rightarrow \text{el}$, $t \rightarrow \text{ț}$, $d \rightarrow z$, $\text{eat} \rightarrow \text{eț}$, $\text{ăd} \rightarrow \text{ez}$, $\text{east} \rightarrow \text{eșt}$, $\text{asc} \rightarrow \text{ășt}$, $\text{easc} \rightarrow \text{eșt}$, $\text{ia} \rightarrow \text{ie}$, $\text{st} \rightarrow \text{șt}$, $\text{ar} \rightarrow \text{ăr}$, $\text{as} \rightarrow \text{ăs}$, $l \rightarrow 0$, $\text{oas} \rightarrow \text{os}$, $\text{tate} \rightarrow \text{tățî}$, $\text{oa} \rightarrow \text{o}$, $\text{ias} \rightarrow \text{ies}$, $\text{iaț} \rightarrow \text{ieț}$, $\text{eanc} \rightarrow \text{enc}$, $\text{ianc} \rightarrow \text{ienc}$, $\text{șanc} \rightarrow \text{șenc}$, $\text{în} \rightarrow \text{in}$, $\text{eal} \rightarrow \text{el}$, $\text{jeal} \rightarrow \text{jel}$, $\text{șeal} \rightarrow \text{șel}$, $\text{easc} \rightarrow \text{eșt}$;
- for neuter nouns NN (A3) :
 $\text{ea} \rightarrow e$, $\text{înt} \rightarrow \text{int}$, $\text{ăn} \rightarrow \text{en}$, $\text{ăt} \rightarrow \text{eț}$, $\text{o} \rightarrow \text{oa}$, $\text{str} \rightarrow \text{ștr}$, $\text{ia} \rightarrow \text{ie}$, $\text{ăs} \rightarrow \text{eș}$;
- for masculine adjectives MA (A4) :
 $t \rightarrow \text{ț}$, $d \rightarrow z$, $s \rightarrow \text{ș}$, $x \rightarrow \text{cș}$, $\text{st} \rightarrow \text{șt}$, $\text{xt} \rightarrow \text{cșt}$, $z \rightarrow j$, $\text{sc} \rightarrow \text{șt}$, $\text{ea} \rightarrow e$, $\text{ian} \rightarrow \text{ien}$, $\text{ean} \rightarrow \text{en}$, $\text{ăn} \rightarrow \text{en}$, $\text{eat} \rightarrow \text{eț}$, $\text{ead} \rightarrow \text{ez}$, $\text{ăt} \rightarrow \text{eț}$, $\text{ăd} \rightarrow \text{ez}$, $\text{os} \rightarrow \text{oș}$, $\text{est} \rightarrow \text{eșt}$, $\text{oal} \rightarrow \text{ol}$, $l \rightarrow 0$;
- for feminine adjectives FA :

- (A5) $\check{a}t \rightarrow ea\check{t}$, $o \rightarrow oa$, $os \rightarrow oas$, $e\check{t} \rightarrow ea\check{t}$, $or \rightarrow oar$, $el \rightarrow ea$,
 $esc \rightarrow easc$, $iesc \rightarrow iasc$, $ie\check{t} \rightarrow ia\check{t}$, $e \rightarrow ea$;
- (A6) $at \rightarrow et$, $\check{a}d \rightarrow ed$, $\hat{i}nt \rightarrow int$;
- (A7) $ian \rightarrow ien$, $sc \rightarrow \check{s}t$, $ea \rightarrow e$, $ean \rightarrow en$, $es \rightarrow eas$, $el \rightarrow ea$,
 $el \rightarrow ic$, $\check{a}n \rightarrow en$, $esc \rightarrow e\check{s}t$.

The consonant alternations are present in variations of roots and affixes. The word root is ended with a consonant or one group of consonants, which, in inflexing process, are substituted for others.

Examining the sets A1 – A7 we select the following consonant alternations

Figure 6

- t / \check{t} ex. : *burete – bureți*, *boltă – bolți*;
- d / z ex. : *brad – brazi*, *coadă – cozi*;
- s / \check{s} ex. : *urs – urși*;
- $(\check{s})c / (\check{s})t$, $(s)c / (s)t$ ex. : *pușcă – puști*, *muscă – muște*;
- $s(t) / \check{s}(t)$, $s(tr) / \check{s}(tr)$ ex. : *artist – artiști*, *ministru – miniștri*;
- z / j ex. : *obraz – obraji*;
- l / i ex. : *cal – cai*, *gol – goi*;
- $x / c\check{s}$ ex. : *sfinx – sfincși*, *fix – ficși*.

Comparing the set of consonant alternations examined above, which are based on the graphic principle, to those, which are based on the phonetic principle (examined in [4]), we found that they differ in the following consonant alternations :

- c / ci ex. : *sărac – săraci*, *ac – ace*;
- g / gi ex. : *vargă – vergi*, *cîrlig – cîrlige*.

These alternations from the graphical point of view are identical, therefore they are not examined for automat inflexion.

The vowel alternations can be both inside the root and in affixes, obtained in the process of word inflexion. They are produced both in tonic (stressed) and in atonic (unstressed) syllables . There are the some examples of vowel alternations [4]. (Note, that " ' " denotes the stress in word.)

1. The vowel alternations occuring in the tonic syllable :

- a' / e' ex. : *băia't* – *băie'ti*, *șar'pe* – *șe'rpi*;
- ă' / e' ex. : *vă'r* – *ve'ri*, *mă'r* – *meri*;
- î' / i' ex. : *cuvî'nt* – *cuvi'nte*, *vi'nă* – *vi'ne*;
- a' / ă' ex. : *ca'le* – *că'i*, *la'mpă* – *lă'mpi*;
- oa' / o' ex. : *soa're* – *so'ri*, *moa'ră* – *mo'ri*;
- o' / oa' ex. : *ogo'r* – *ogoa're*, *ro'd* – *roa'de*;
- ea' / e' ex. : *crea'stă* – *cre'ste*; *sea'ră* – *se'ri*;
- î' / îi' ex. : *mî'nă* – *mî'ni*;
- e' / i' ex. : *fugăre't* – *fugări'tă*;
- e' / ea' ex. : *des* – *dea'să*.

2. The vowel alternations occuring in atonic syllable:

- ă / e ex. : *sî'mbătă* – *sî'mbete*;
- o / oa ex. : *mijlo'c* – *mijloa'ce*.

3. The vowel alternations presupposing the stress place changing:

- o' / u ex. : *so'ră* – *suro'ri*, *no'ră* – *nuro'ri*.

If we can determine, in procedural way, the atonic or tonic syllables in a word, then the word inflexion process automation problem from vowel alternations point of view can be solved completely. But unfortunately, it is difficult (or may be impossible) to indicate the procedure of stressed or unstressed syllables determination.

We can divide the set of vowel alternations into three groups : absolute and partial regularities, and irregularities.

The absolute regularities are the following:

- common for nouns and adjectives : ea → e, ia → ie ;
- for nouns : ä → e;
- for neuter nouns : o → oa;
- for adjectives : o → oa, e → ea, ie → ia.

The partial regularities are the following:

- common for nouns and adjectives : î → i;
- for feminine nouns: a → ä, a → e, oa → o.

The rest of vowel alternations belongs to the irregularities.

Some roots may be affected by both vowel and consonant alternations. Their possible combinations are pictured in the following figure:

Alt.vowel. alt.conson.	a'/e'	î'/i	a'/ä'	o'/oa'	e'/ea'	ä/e	e/a
t/ṭ	+	+	+	+	+		+
d/z	+		+	+			
s/š				+	+	+	
st(r)/št(r)				+			+
sc/št					+		
z/j					+		
l/0			+		+		

We made some conclusions :

- The masculine nouns do not have vowel alternations, but only consonant ones, mentioned above.
- The feminine nouns have both vowel and consonant alternations.
- The neuter nouns are affected only by vowel alternations.

4 The nouns and adjectives inflexion proceeding

The nouns and adjectives inflexion proceeding is the following.

The word for inflexion is introduced, for it the speech part of this word is indicated. If the word is a noun, then the gender and the number are specified. For this word the base form – *FI, CNA* and *NS* – is indicated. The division of the word into a root and an affix is made in the following way. The affixes specified in Figures 1 – 3 are arranged in the decreased order of their lengths. For each affix a special procedure of noun inflexion is selected. In the word it finds the affix in the decreased order of length. If it coincides with the one specified in Figures 1 – 3, then the corresponding procedure of inflexion is called. In the process of inflexion in automat way it should be produced the consonant alternation (in accordance with Figure 6).

The vowel alternations, which are absolute regularities both for nouns and for adjectives ($ea \rightarrow e$, $ia \rightarrow ie$) are made without a dialog, but for the alternation $\hat{i} \rightarrow i$, which is a partial regularity, the dialog is necessary. (For example: *tînăr – tineri*, *pămînt – pămînturi*).

The vowel alternation which is characteristic only of nouns and which is an absolute regularity (in the majority of cases) $\check{a} \rightarrow e$ it is made without dialog.

For feminine nouns we can distinguish the following vowel alternations, which is a partial regularity: $a \rightarrow \check{a}$, $a \rightarrow e$, $oa \rightarrow o$. The necessary alternation is selected by means of dialog. For example: *cană – căni*, *masă – mese*, *floare – flori*, but for examples: *casă – case*, *masă – mase*, *coasă – coase* a dialog is necessary.

The masculine nouns do not have vowel alternations, but only consonant ones, mentioned above (Figure 6). Therefore the inflexion of masculine nouns is carried out completely automatically. Naturally, the special processing is necessary for a finite set of irregular words, e.g. *tată*, *popă* etc.

For neuter nouns we can specify the following vowel alternations, which are absolute regularity : $o \rightarrow oa$. For example: *cojoc – cojoace*. The problem of this alternation is solved automatically. Here the dialog

is necessary to determine the ending of Plural (for instance, *muzeu* - *muzei*, *scaun* - *scaune*, *dulap* - *dulapuri*, *zmeu* - *zmeie*).

The help of the user is necessary for inflexing a word containing any irregularity. The system proposes some flexions to be examined and corrected. For example, the word "om" will be declined by the system:

FI : NS : CNA - om, CGD - om, CV - om;
NP : CNA - omi, CGD - omi CV - omi;
FD : NS : CNA - omul, CGD - omului, CV - omule;
NP : CNA - omii, CGD - omilor, CV - omilor.

The user will make the corrections for *FI NP CNA, CGD* and *CV*: the *omi* will be changed in the *oameni*, but for *FD NP CNA, CGD* and *CV*: *omii* and *omilor* will be changed respectively in *oamenii* and *oamenilor*.

If a word is an adjective, then the base form - *FI, GM, CNA* and *NS* - is indicated. The word division into a root and an affix is made in analogous way, as in the case of a noun. All affixes of adjective are arranged in the decreased order of their lengths (Figure 4). The procedure of selecting a possible affix is analogous to that one for a noun. If an affix coincides with the one in Figure 4, then we call the corresponding procedure for adjective declination. The consonant alternations were produced (in accordance with Figure 6) without dialog.

For adjectives the absolute vowel alternations are realized automatically, the rest need a dialog. For example, the dialog is necessary for the following vowel alternations, which are a partial regularity: *a* → *ă*, *a*, → *e*, *oa* → *o*. These alternations are frequent for adjective inflexion (for feminine gender forms). For example: *frumos* - *frumoasă*, *nobil* - *nobilă*, *des* - *deasă*, *dens* - *densă*, *măiestru* - *măiastră*.

The problem of irregularities of vowel alternations for adjectives is solved in the same way as for nouns.

5 The procedure of verb inflexion

We show now how the procedure of verb inflexion works examining the activity of one more complicated module Word-Verb1.

The result of every procedure work is the list L of complicated structures, which contains the information for different speech parts. The procedures of this complex complete the fields :

- Prt – speech part, in our case this is constant Verb;
- Wrd – the verb for conjugation in the infinitive (from which we will obtain the generated flexions);
- $Generator$ – the procedure number, which conjugates this indicated word, in our case these are the procedures 1–7 (look 2.3).

The fields, which determine the grammatical categories of flexions for verbs are :

- $Personal$ – in this field we introduce the information if the verb is personal or impersonal (the user answer);
- the $Mood, Tense, Person$ fields indicate corresponding mood, tense and person) and are completed in the function of generated flexion.

The list L (in the majority of cases) contains 35 of elements (Infinitiv – 1, Indicativ Present – 6, Perfectul Simplu – 6, Mai Mult ca Perfectul – 6, Imperfect – 6, Conjunctiv Present – 6, Imperativ – 2, Participiu – 1, Gerunziu – 1). Some verbs of Imperativ Singular or Plural have a lot of forms, more than 2. For example, the verb $a\ veni$ (to came) for Imperativ Singular has the forms $vino, vin, vină$, the verb $a\ naște$ (to be born) for Imperativ Plural: $nașteți, nășteți$. There is an analogous situation for Conjunctiv Present, III-rd person for some verbs which have 2 forms. For example, the verb $a\ da$ (to give): $să dea, să deie$.

For impersonal verbs there are only 12 flexions.

We describe the implementation of procedures using some functions of the TURBO-PASCAL designated for processing the strings :

- Length (S) – it indicates the string S length ;

- Copy (S, n, k) – gives substring of length k from string S from position n ;
- Insert ($S1, S, n$) – the string $S1$ is inserted to the string S from position n ;
- Delete (S, n, k) – in the string S from position n k characters are deleted.

These means are sufficient for the realization of the schemes of conjugation (look 2.3, Figure 5).

Let us introduce the notion of "word root".

Definition 1 *The word root is calculated by the formula :*

$$R = \text{Copy}(S, 1, \text{Length}(S) - k),$$

where S is the string which contains the infinitive of conjugation verb,
 $k = 2$ for the verbs of the second grammatical group (with the ending "ea"),
 $k = 1$ for the rest of the verbs.

If in the verbs conjugation process in the word root there are the consonant or vowel alternations, then we obtain some roots, which we denote with $R, R1, R2, R3, \dots$

The set of affixes consists of 6 strings which are annexed on the corresponding root for I – VI persons (we consider that I – III is for the singular number and IV – VI – for the plural) for obtaining the flexions. For example, to realize of the conjugation scheme of the verb $a \hat{c}înta$ (to sing), $R = "cînt"$ the affixes set for forming Indicativ Prezent (" ", "i", "ă", "ăm", "ați", "ă") is annexed, but for the forming, for example, Indicativ Imperfect it is necessary to have the set ("am", "ai", "a", "am", "ați", "au") etc. Realizing the conjugation scheme of the verb $a \check{c}ăuta$ we obtain the roots $R = "căut"$ and $R1 = "caut"$. With the procedure Proc_Affix ($S1, S2, S3, S4, S5, S6$) we prepare the sets of affixes for both roots. The parameters of the indicated procedure are the necessary set of affixes. We call the procedure

Proc_It (*kon*, *kon+n*, *Pers*, *Mod*, *Timp*: *integer*; *Rd*: *string*)

as often as required. During one call the procedure includes in structures array described above, the elements with numbers *kon*, *kon+1*, *kon+2*, ..., *kon+n*. The fields *Wrd* will contain the flexions obtained from roots, which are transmitted by parameter *Rd*, adding the set of affixes in accordance with the parameter indicated by *Pers*. The fields *Person* will contain *Pers*, *Pers + 1*, *Pers + 2*, ..., *Pers + n*, but the fields *Mood* and *Tense* are completed with constants transmitted by the parameters *Mod* and *Timp* (respectively). We know, that for the generation of any flexion for any tense in accordance with the scheme, corresponding to the verb *a cînta*, it is sufficient to call the procedures *Proc_Affix* with the set of affixes, in accordance with the given flexion (with the specified categories), and *Proc_It* once, because the word root for this given scheme is one for all cases of the necessary morphologic categories. We can remark the following. One question may arise: why do we use for the verb *a cînta* one root sufficient for proceeding flexions, if there exists another: *cînt* (*eu cînt*, *tu cînți*)? The answer is the following: the consonant alternations *t*→*ț*, *d*→*z*, *sc*→*șt*, *st* →*șt*, *șc*→*șt*, which is necessary to be executed, to obtain the flexions for II-nd person, Indicativ Prezent Singular and Conjunctiv Prezent (sometimes and the flexions for III-rd person, for example, *eu urăsc*, *tu urăști*, *el urăște*) are automated by two auxiliary procedures. This is possibly because these alternations do not depend on the scheme of the conjugated verb, but only on the presence of one from the left parts for I-st person Indicativ Prezent Singular. These are the unique alternations which do not require the introduction of a new root.

Let us recall the verb *a căuta*. The set of affixes for Indicativ Prezent is same as for the verb *a cînta*. The call for the procedure *Proc_It* will be the following:

Proc_It (*kon*, *kon+2*, *Pers*, *Mood_Indic*, *Tense_Prezent*, *R1*);

Proc_It (*kon*, *kon+1*, *Pers*, *Mood_Indic*, *Tense_Prezent*, *R*);

Proc_It (*kon*, *kon*, *Pers*, *Mood_Indic*, *Tense_Prezent*, *R1*).

In this way we obtain all the flexions of the conjugated verb, but with the exception of Imperativ, Participiu and Gerunziu. For the latter cases we use a special procedure

Proc_Alt_Timp (*Pers:integer*;

RPart, SufPart, RGer, SufGer : string; Var L : Itemlist),

where *Itemlist* is a type of structures list *L* mentioned above.

By parameter *Pers* we transmit the information about Imperativ:

if *Pers* = 2 then Imperativ Singular has only 2 forms, which coincides with Indicativ Prezent Singular II-nd person and Imperativ Plural which coincides with Indicativ Prezent Plural II-nd person;

if *Pers* = 3 then it has only 2 forms, but Imperativ Singular coincides with Indicativ Prezent Singular III-rd person;

if *Pers* = 5 then Imperativ Singular has only 2 forms, one of which coincides with Indicativ Prezent Singular II-nd person, but the second – with Indicativ Prezent Singular III-rd person, and the Imperativ Plural – usually – coincides with Indicativ Prezent Plural II-nd person.

The parameter *RPart* obtains the value of the root, which in concatenation with the parameter *SufPart*, forms Participiu, but the parameter *RGer* in concatenation with the parameter *SufGer* forms Gerunziul. At the end of this procedure an array of structures *It* for all verbs flexions is linked into list *L* (this list was mentioned above), which will be transmitted to other procedures.

In this manner the conjugation schemes are realized. But it is not easy as it seems at the first sight. There are many exceptions even within one scheme limits. For example: the verb *a scădea* (to subtract) is conjugated as the word *a cădea* (to fall), however Imperativ Singular of the verb *a cădea* coincides with Indicativ Prezent Singular II-nd person – *cazi*, but the verb *a scădea* has Imperativ Singular – *scade*, which coincides with Indicativ Prezent Singular III-rd person. These exceptions were considered diuring the programming of the given conjugation scheme. Unfortunately these exceptions are not always indicated in [6]. Some collisions were deleted consulting [7, 8]. Another difficulty is the formalization of the verbs set description which refers to one or another conjugation scheme.

Let us introduce the notion of "root predaffix" and consider the example of this formalization.

Definition 2 *Predaffix with the length n ($n \leq Length(R)$) is the substring Copy ($R, Length(R) - (n - 1), n$), i.e. it is the root R "tail" of length n .*

The verbs conjugation schemes *a tăia* (to cut), *a înfoia* (to swell), *a muia* (to soak), *a apropia* (to approach) have a lot of common proprieties and they differ greatly from other conjugation schemes. Therefore these schemes are joined into one case .

1. If predaffix of the length 1 is constant "i", then 2 affixes subsets variants appears for forming Indicativ Prezent (we examine namely it, because other flexion tenses forms are realized very easily) :
if $R[Length(R) - 1]$ does not coincide with one of the characters "a", "u", "o", "e", then we use the following set of affixes : "i", "i", "e", "em", "aṭi", "e";
else (if it coincides) – it uses the following set : " ", " ", "e", "em", "aṭi", "e". To clarify the set of roots we have to remember the definition of root R and we can examine the following subcases:

- if predaffix of the length 3 coincides with string "täi" (for example: the verbs *a tăia*, *a străătăia* etc.), then we obtain again one new root ($R1$), substituting the last vowel "ă" for "a";
- if $R[Length(R) - 1]$ coincides with character "o", then we obtain the root $R1 = Insert("a", R, Length(R))$;
- if $R[Length(R) - 1] = "u"$, but $R[Length(R) - 2]$ does not coincide with some of the characters "c", "e", "ṭ" and $R \neq "împui"$, the root $R1$ is obtained on the substitution of the character "u" with "o";
- the rest of verbs from this case have one root.

The procedure *Proc_It* remains to be used correctly (for each case in a special way).

2. Let us consider how to formalize the verbs set which conjugates in accordance with the verb scheme *a afla* (to hear) and to obtain Indicativ Present. The formalization of this set is made in the following way. If the preaffix of length 2 coincides with one of the strings "FL", "CR", "PL", "NU", "TR", "BL", "RL", "TU" (for example: *a insufla* (to inspire), *a consacra* (to consecrate), *a contempla* (to contemplate), *a intra* (to enter), *a umbla* (to walk), *a urla* (to howl), *a continua* (to continue) and so on), then the affixes set will be : "u", "i", "ă", "ăm", "ați", "ă". The roots set (with the exception of the verb *a lătra* (to bark), which has 2 roots : $R = \text{"lătr"}$ and $R1 = \text{"latr"}$), consists of one element and consequently forms Indicativ Present by procedure *Proc_It* (*kon*, *kon+5*, *Pers*, *Mood_Indic*, *Tense_Prezent*, *R*). For the verb *a lătra* we obtain two roots and the procedure *Proc_It* is called in a analogous way as in the case of the verb *a căuta* (to look for).

3. The verbs which conjugate in accordance with conjugation scheme of the verb *a apăsa* (to press) have to satisfy the condition: the preaffix of length 3 has to coincide with the one of the strings "PĂS", "FĂS", "FĂT", "BĂR", "MĂT", "BĂT", "VĂȚ", "FĂȚ", "ĂRS", "PĂL", or "SĂL". For example, the verbs *a învăța* (to study), *a înfășa* (to swaddle), *a vărsa* (to spill) and so on.

4. The verbs set, for which verbs conjugation schemes coincide with those for the verb *a căpăta* (to receive), satisfies the following condition :
 $(Length(R) > 3)$ and $(R[Length(R) - 1] = \text{"ă"})$ and $(R[Length(R) - 3] = \text{"ă"})$ or $(R[Length(R) - 4] = \text{"ă"})$.
 For example, the verbs are *a cățara* (to climb), *a vătăma* (to harm), *a scărmana* (to snatch), *a scăpăra* (to sparkle) and so on.

5. If the preaffix of length 3 coincide with the one of the following strings "PĂD", "MĂN", "TĂN", "PĂN", "MĂT", "ZEM", "PĂR", "RĂȚ", "MĂR" (and this condition is verified after case

3), then the verb is conjugated in accordance with the scheme for the verb *a asemăna* (to compare).

For example: the verbs are : *a depăna* (to spin), *a fremăta* (to rustle), *a pieptăna* (to comb), *a rezema* (to lean), *a lepăda* (to drop), *a legăna* (to rock) and so on.

Besides those cases, other 5 sets are distinguished in analogous ways, but the corresponding conditions are much longer, therefore we do not indicate them here.

The rest of the verbs from the first grammatical group (which does not satisfy the conditions of these cases) conjugates in accordance with the conjugate scheme of the verb *a cânta* (to sing).

It is necessary to note again that within each conjugation scheme limits (i.e. within the mentioned 10 cases limits) the subcases, which reflect the conjugation specific features of some verbs subgroups, are selected. For example, the verbs *a cățara* (to climb), *a văicăra* (to lament), *a vătăma* (to harm), *a tărăgăna* (to dally) (it may be other) which belong to the II-nd case, for Indicativ Prezent II-nd person Singular, do not substitute the last but one vowel "ă" for "e" (*cațări, vatămi* and so on), but for the this case the rest of the verbs have the vowel alternation $\check{a} \rightarrow e$ (*capeți, scarmeni*).

Consequently we concluded that the formalization of cases and subcases is a rather difficult problem, moreover if we consider the development and extension of a natural language, the usage of different forms with the same grammatical categories (for example: *comenzi* and *comanzi*), then it is clear that the results of the program's work must be examined and corrected if it is necessary by the user and only after that the word-forms are included into a dictionary.

References

- [1] Cojocaru S., Evstunin M., Ufnarovski V. Detecting and Correcting Spelling Errors for the Roumanian language. The Computer Science Journal of Moldova, 1993, n. 1, p. 3-21.

- [2] Toşa A. Elemente de morfologie. Bucureşti: ed. stiinţ. şi encicl., 1983. 299 p.
- [3] Zdrenghea M. Limba română contemporană. Morfologia. Cluj, 1977. 256 p.
- [4] Guţu_Romalo V. Morfologie structurală a limbii române. Bucureşti: ed. Acad. Rep. Soc. România, 1968. 364 p.
- [5] Iordan I., Guţu_Romalo V., Niculecu A. Structura morfologică a limbii române contemporane. Bucureşti: ed. stiinţ., 1967. 356 p.
- [6] Irimia D. Verbul. Bucureşti: Junimea, 1977. 356 p.
- [7] Lombard A., Gadei C. Dictionnaire morphologique de la langue roumaine. Bucureşti: ed. Acad. Rep. Soc. România, 1981. 415 p.
- [8] ORTO. Dicţionar ortografic cu elemente de ortoepie şi morfologie. Chişinău : Red. Principală a Enciclopediei Sovietice Moldoveneşti, 1990. 608 p.

E.Boian, A.Danilchenko, L.Topal,
Institute of Mathematics,
Academy of Sciences of Moldova,
5 Academiei str.,
Chişinău, 277028, Moldova
e-mail: 22lena@math.moldova.su

Received February 15, 1993