

"Categorization and analogical change: The case of athematic 1sg *-m* in the Slavic languages"<sup>1</sup>

by Laura A. Janda, UNC Chapel Hill

Forthcoming in a  
refereed volume

0.0 Introduction

Every inflected language carries some seemingly useless morphological baggage bequeathed to it by previous generations in the form of irregular inflectional affixes. These relics of what were once productive and systematic linguistic distinctions often fade into oblivion, although they can persist for hundreds of years. However, it is not the case that all defunct morphology must wither and die. Languages are capable of recycling nearly extinct morphemes and using them either to restore distinctions that have eroded or even to build entirely new systems of distinctions. There are many examples of this type of analogical extension, yet it has received so little attention that it lacks a standard term. This type of analogical extension may be called "upstream" (i.e., against the "expected" flow of morphology which tends to eliminate irregular forms in favor of regular forms) or "exaptative," a term borrowed by Lass (1990) from biologist Stephen J. Gould. I will illustrate exaptative analogical extension with a chapter from the history of the Slavic languages: the spread of 1sg *-m* from a handful of athematic verbs to much or all of the verbal lexicon in the West and South Slavic languages.

"Exaptation" appears to have both a variety of sources and a variety of explanations. There are several different kinds of marginalized morphemes that can serve as source material for this process because a grammatical morpheme can become marginal in one of several ways: a) when the paradigm that uses the morpheme becomes limited to only a few lexical items (as in the case of the athematic 1sg *-m*), b) when the paradigm is lost (e.g., the former *ũ*-stems of Slavic, and c) when the grammatical category that the morpheme represents is lost (e.g., the loss of the dual as a grammatical category in most of Slavic). The categorial status of marginal morphemes is thus compromised and they are available for recategorization. Furthermore, the pattern of recategorization is quite logical. Morphemes in situation a), where the paradigm is limited to a few lexical items, still represent active linguistic categories in live paradigms, and they will remain stable or decline unless phonological (and morphological) events "conspire" to change their status within their category. This is the case of the spread of athematic *-m*.

In addition to the framework of cognitive linguistics, involving linguistic categorization based on prototypes and radial networks, I invoke Andersen's (1973) model of language change via abduction and subsequent deduction. The synchronic state of a language presents learners of the younger generation with data from which they abduce their grammar by making decisions about which linguistic categories are relevant and which characteristics are prototypical of those categories. There is inevitably some mismatch between the grammars abduced by successive generations, and this is the ultimate source of language change. Thus, in the analysis that follows, I will present the synchronic state of each relevant language just prior to the spread of 1sg *-m* and discuss the abductive innovations that promoted *-m* from the status of a marginal morpheme to that of prototypicality, leading to the deduction that *-m* should be more widespread in the verbal lexicon.

#### 1.0 Preconditions common to West and South Slavic

IndoEuropean had two sets of verbal endings, one for thematic and another for athematic verbs. Athematic verbs further distinguished between primary endings for the present active indicative and secondary endings for all other categories. For the 1sg, the primary ending was *-m-i* and the secondary was the same as the primary, except that it lacked the "hic et nunc" particle *-i* and was merely *-m*. ProtoSlavic inherited the primary ending *-mi* for its five athematic verbs and developed an ending consisting of *-ōm* for the 1sg of the thematic verbs. The presence of the *-i* particle in the athematic forms vs. its absence in the thematic forms had serious consequences for the syllabification and subsequent fate of 1sg forms. The law of rising sonority (in effect until the disintegration of Slavic unity) occasioned the simplification of consonant clusters and the monophthongization of diphthongs as described below:

athematic verbs:

1sg forms:  $CV.C$  (root) + *m-i* > simplification of consonant clusters >  
 $CV + .m-i$  > fall of jers >  $CV + m$

thematic verbs in *\*ole* :

1sg forms:  $CV.C$  (root) + *ōm* > monophthongization of diphthongs >  
 $CVC + q$

The period symbolizes the position of a syllable boundary. Where *m* was in syllable initial position it was retained, but in syllable final position *m* was monophthongized with the

theme vowel *o* to yield the nasal vowel *q*. The close of the common Slavic period is associated with the elimination of short *i, u* (called "the fall of the jers") which removed the final *-i* particle from the athematic verbs, thus opposing two very dissimilar expressions of 1sg. Note that monophthongization not only removed the 1sg marker *-m* but also altered the role of the theme vowel since it was no longer a separate entity in this form. A similar monophthongization took place in the thematic third plural forms with the same consequences for the theme vowel:

thematic verbs in *\*o/e* :

3pl forms:  $CV.C + o + n.tũ >$  monophthongization of diphthongs  $>$   
 $CVC + qtũ >$  fall of jers  $>$   $CVC + qt$

The development in thematic verbs in *\*i*, though not identical, also effectively stripped the theme vowel of its independent role in non-past morphology:

thematic verbs in *\*i* :

1sg forms:  $CV.C + i + o-m >$  *i* becomes non-syllabic when not syllable  
 peak  $>$   $CV.Cj + om >$  deiotation and monophthongization of  
 diphthongs  $>$   $CVČ + q^2$

3pl forms:  $CV.C + i + ntũ >$  monophthongization of diphthongs  $>$   $CVC + etũ$   
 $>$  fall of jers  $>$   $CVC + et$

Thus by the time of late Common Slavic the role of the theme vowel had been compromised and we can speak of an independent theme vowel only in the 2sg, 3sg, 1pl, and 2pl forms.

The deiotation of the stem final consonant in 1sg forms of thematic verbs in *\*i* created a morphophonemic alternation that opposed the 1sg to all other forms (e.g. 1sg in *proš-* vs. all other forms in *pros-* for the verb 'beg'), and the palatalization of stem final velars before the theme vowel *e* in the 2sg-2pl forms created another morphophonemic alternation opposing the 1sg and 3pl to all other forms (e.g. 1sg and 3pl in *pek-* vs all other forms in *peč-* for the verb 'bake'). As we shall see, in most languages of West and South Slavic (i.e., everywhere but Polish) it was abduced that desinences, not morphophonemic alternations, were the prototypical means for distinguishing person and number in the non-past. The spread of *-m* in these languages is associated with the elimination of morphophonemic alternations.

The fate of the 1sg markers cannot be discussed outside of the context of the verbal system to which they belonged, so it is necessary to outline the paradigms of this system

here. At about the time of the fall of the jers, sequences of *-VjV-* contracted to yield *-V̄-*, and this included sequences in which the *j* was the stem final consonant and the second vowel was the theme vowel, thus removing the original theme vowel and producing new conjugational paradigms. Contraction was not carried out uniformly in this territory, but certain generalizations can be made:

-- contraction was avoided wherever it would disfigure the shape of the root (e.g., forms of a verb like 'bark' in *laj-ole-* were never contracted)

-- contraction did not take place in the 1sg and 3pl

-- Bulgarian and Macedonian contracted only sequences of *-aje-* in verbal forms, never *-ěje-*

-- Slovene and Serbo-Croatian merged the verbs of 'becoming' in original *-ěje-* with the class II verbs (in *-i-*) prior to the time of the spread of *-m* and therefore do not show reflexes of contraction for these verbs

-- Polish avoided contracting any sequence of *-aje-* or *-ěje-* that was identified with a suffix denoting imperfectivization or the deadjectival derivation of a verb meaning 'become X' (e.g. *poznajesz* 'you meet,' *leniwiejesz* 'you become lazy').

Contraction created new long vowels. Length will be marked in the paradigms below only where it occurs in all languages except Macedonian and Bulgarian (which lost this distinction prior to the spread of *-m*).

After contraction had taken place, the following system of non-past conjugations was present in West and South Slavic:

**Class I: -o-/-e-**

1sg	-q	1pl	-e-mV
2sg	-e-š	2pl	-e-te
3sg	-e	3pl	-q̄(t)

This class continues the IndoEuropean type of conjugation with the ablauting theme vowel *-o-/-e-*. The final vowel in the 1pl is a recent innovation, and varies according to language: in Polish it is *y*; in Czech, Slovak, Macedonian, and Bulgarian it is *e*; and in Slovene and Serbo-Croatian it is *o*. The final *t* is present in the 3pl forms only in those languages that do not have phonemic length to distinguish the 3pl from the 1sg.

**Class II : i**

1sg	-q	1pl	-i-mV
2sg	-i-š	2pl	-i-te

3sg -i                      3pl -ē(t)

In this class the theme vowel is *i* and this vowel is long in all of West and South Slavic but Polish, Macedonian, and Bulgarian.

**Class III** : -ěj + -o-/e-

1sg	-ějq	1pl	-ěj-mV
2sg	-ěj-š	2pl	-ěj-te
3sg	-ěj	3pl	-ějq̄

This is a new conjugation class created by the contraction of original sequences of -ėje- in the 2sg-2pl forms to produce a new theme vowel of -ěj-. This class does not exist in Macedonian and Bulgarian, and is limited to only three non-derived verbs in Polish. Verbs in these three languages with uncontracted -ėje- are kept in class I.

**Class IV**: -aj + -o-/e-

1sg	-ajq	1pl	-ā-mV
2sg	-ā-š	2pl	-ā-te
3sg	-ā	3pl	-ajq̄(t)

Class IV is likewise a new conjugation class arising from the contraction of -aje- to produce the new theme vowel -ā-. The Polish verbs in original -aje- that did not contract remained in class I.

**Class V**: athematic<sup>3</sup>

<i>věm, věš, vě, věmV, vě(s)te, vědě(t)</i>	'know'
<i>jěm, jěš, jě, jěmV, jě(s)te, jědě</i>	'eat'
<i>(i)mām, (i)māš, (i)mā, (i)māmV, (i)māte, (i)majq(t)</i>	'have'
<i>dām, dāš, dā, dāmV, dā(s)te, dadē(t)</i>	'give'

The parallels between the athematic verbs 'know' and 'eat' and class III (both types have the vowel *ě* preceding the person-number desinences) on the one hand and between the athematic verbs 'have' and 'give' and class IV (both types have the vowel *ā* preceding the person-number desinences) on the other hand were striking and were instrumental in the spread of 1sg -m. Up until this point, distinctive athematic desinences had been eliminated in order to bring the athematic verbs into conformity with the thematic paradigms. Specific events in this trend include:

-- changing the 2sg athematic desinence from -si to -š

-- loss of final -t in the 3sg and (except in Macedonian and Bulgarian) 3pl, paralleling its loss in the thematic verbs

-- loss of *-s-* in the 2pl desinence, e.g. *věste* > *věte* 'you know' (except in Slovene)

The spread of the one remaining distinctive athematic desinence to thematic verbs was to mark a dramatic departure from this trend.

One more piece of pertinent background information: the morphophoneme *m* was present in the following pronominal roots and verbal desinences, all of which signaled first person:

- all of the oblique forms of the 1sg pronoun
- the nominative forms of the 1pl pronoun: *my* (everywhere but most Bulgarian and Macedonian dialects)
- the 1pl non-past marker, also used in 1pl imperative formation: *-mV*
- the 1pl marker for both the imperfect and the aorist in *-m*

Thus the choice of *m* as a marker of first person had plenty of support in the grammar inherited by early West and South Slavic.

To summarize, the following are preconditions shared by all languages that ultimately spread 1sg *-m*:

- 1sg non-past forms end in *-q* or *-m*
- *m* is used elsewhere in pronominal and verbal morphology to mark first person
- independent theme vowels are present in 2sg-2pl only
- morphophonemic stem alternations distinguish 1sg and 3pl from other forms in class I, and 1sg from other forms in class II
- contraction creates new theme vowels *-ě-* (most languages) and *-ā-* (all languages) which parallel the predesinential vowels in the athematic verbs
- up until this point the trend has been to level out distinctive athematic desinences in favor of thematic desinences.

Although all of West and South Slavic extended the use of *-m*, this process occurred several centuries after the disintegration of Slavic. Sound changes that took place in the intervening time played a decisive role in determining what new categorial associations could be abduced in the grammars of individual languages. Both the ultimate extent of 1sg *-m* and the route of cognitive leaps that brought about its extension varied greatly from language to language. There are three Slavic languages in which *-m* was universalized: Slovak, Slovene, and Macedonian. This does not mean, however, that the

course or even the outcome of this process was the same in these three languages, as we shall see. Although the discussion will focus primarily on these languages, parallels will be drawn to the remaining Slavic languages.

## 2.0 The spread of *-m* in Slovak (and Czech, Polish, and Serbo-Croatian)

Here are the relevant phonological changes peculiar to Old Slovak (cf. Krajčovič 1988):

$g > h$	(circa fall of jers)
$\bar{e}, \bar{e} > \bar{e}, e$	(circa fall of jers)
$\bar{q} > \bar{u}; q > u; \bar{e} > \bar{a}; \bar{e} > \bar{a} / P\_ , \text{ elsewhere } > a$	(10th c)
$\bar{e} > ie$	(13th c)
$\bar{a} > ia / C'_-$	(15th c)
depalatalization of palatalized consonants	(15th c)

Here is the inventory of Old Slovak verb classes at the time *-m* spread:

Class I is divided into two major subclasses according to the length of the reflex of the original theme vowel in the 2sg-2pl forms, which will be referred to as V<sub>1</sub> (since the theme vowel is no longer an independent morpheme throughout verbal paradigms).

Class Ia: (sg) *-u, -eš, -e*, (pl) *-eme, -ete, -ū*<sup>4</sup>

This class includes verbs

-- in *\*-no/-ne* following a C

*vydýchnu, vydýchneš, (vydýchnu >) vydýchnu* 'exhale'

-- in *\*-jo/-je-* following a C

*ukážu, ukážeš, (ukážu >) ukážu* 'show'

The shortenings observed in the 3pl forms of these verbs are caused by the rhythmic law of Slovak and will be discussed below.

Class Ib: (sg) *-u, -ieš, -ie*, (pl) *-ieme, -iete, -ū*

This class includes

- verbs in *\*-no/-ne* following a *V* or syllabic *r* or *l*
  - minu, minieš, minū* 'pass'
- non-suffixed stems
  - nesu, nesieš, nesū* 'carry'
  - peku* (> *peču*), *pečieš, pekū* (> *pečū*) 'bake'
  - tru, trieš, trū* 'rub'
  - beru, berieš, berū* 'take'
- verbs with stem in *-Vj-*
  - kryju, kryjieš* (> *kryješ*), *kryjū*<sup>5</sup> 'cover'
- verbs in *-ova-* / *-uj-*
  - kupuju, kupujieš* (> *kupuješ*), *kupujū* 'buy'

The shortenings observed in the last two subtypes will also be discussed below.

Class II: (sg) *-u, -īš, -ī*, (pl) *-īme, -īte, -ā*

This class includes

- verbs in *-ī-*
  - prošu, prosīš, prosā* (> *prosia*) 'beg'

Class III: (sg) *-eju, -ieš, -ie*, (pl) *-ieme, -iete, -ejū*

This class includes

- non-derived verbs
  - smeju, smieš, smejū* 'dare'
- derived verbs, especially with the meaning 'becoming'
  - beleju, belieš, belejū* 'become white'

Class IV: (sg) *-aju, -āš, -ā*, (pl) *-āme, -āte, -ajū*

This class includes

- verbs in original *\*-aje-*



	<i>hl'adaju, hl'adāš, hl'adajū</i>	'look'
-- derived imperfectives in *-aje-		
	<i>klaňaju sa, klaňāš sa (&gt; klaniaš sa), klaňajū sa</i>	'bow'
Class V:	<i>viem, vieš, vie, vieme, viete, veď ā (&gt; vedia)</i>	'know'
	<i>jēm, jěš, jē, jēme, jēte, jeď ā (&gt; jedia)</i>	'eat'
	<i>(i)mām, (i)māš, (i)mā, (i)māme, (i)māte, (i)majū</i>	'have'
	<i>dām, dāš, dā, dāme, dāte, daď ā (&gt; dadia, later &gt; dajū)</i>	'give'

V<sub>1</sub> of 'eat' is long *-ē-* (in all but the 3pl) but is diphthongized to *-ie-* and then the first part of the diphthong coalesces with the contiguous glide *-j-*, thus yielding *jem, ješ, je, jeme, jete, jedia*.

I will suggest that fourteenth century learners of Slovak faced with this verb system made the following abduction about their verbal morphology. The non-past forms of all verbs could be segmented as:

stem - V<sub>1</sub> - (C) - (V<sub>2</sub>)

where the stem ends in a consonant and has the minimal shape C(C), V<sub>1</sub> is any vowel or diphthong, and the last two segments are optional. This abduction carries with it the assumption that the predesinential vowel of the athematic verbs is V<sub>1</sub>, not a root vowel.

All forms of the non-past paradigm have the same schematic shape, regardless of conjugation class, except the 1sg and 3pl which are not uniform, giving:

1sg	stem - V <sub>1</sub>	classes I, II
	stem - V <sub>1</sub> - C - V <sub>2</sub>	classes III, IV
	stem - $\bar{V}_1$ - C	class V
3pl	stem - $\bar{V}_1$	classes I, II
	stem - V <sub>1</sub> - C - $\bar{V}_2$	classes III, IV, V

This fact singles out the 1sg as the form with the most variants.

If we compare the paradigms of class Ib and the contracted classes (classes III and IV) with the athematic class (class V) in terms of both the schematic shapes of the forms and the quality and quantity of  $V_1$ , we find that the following generalization can be made:

--  $V_1$  is long *-ie-* or *-ā-* in all forms but the 3pl, with the exception of:

-- the 1sg of classes III, IV, and Ib

The logical consequence of observing this generalization is that the 1sg forms of classes III, IV, and Ib are exceptional, whereas the 1sg forms in the athematic class are prototypical. Further, this generalization imposes a hierarchy of prototypicality judgements on the entire verbal system:

generalization holds

class V

\*

\*

generalization holds, but

-- the 1sg is exceptional

classes Ib, III, IV

\*

\*

generalization holds, but

-- the 1sg is exceptional

-- quality of  $V_1$  varies

class II

\*

\*

generalization holds, but

-- the 1sg is exceptional

-- quality of  $V_1$  varies

-- quantity of  $V_1$  varies

class Ia

The formerly marginalized *-m* has received a great boost in status, and is now licensed to spread to all classes, beginning with those with a  $V_1$  of *-ie-* or *-ā-*, as we see in:

Phase 1

- Abductions: -- *m* is a marker of first person  
 -- there is a category of verbs with (underlying)  
 long  $\bar{V}_1 = \bar{a}$  or *ie*  
 -- class V is prototypical of this category

Deduction: -- classes Ib, III, IV use the prototypical athematic  
1sg in  $-\bar{V}_1 + m$

Phase 1 (plus subsequent sound changes) deduces the modern Slovak forms:

<i>miniem</i>	'I pass'	<i>nesiem</i>	'I carry'
<i>pečiem</i>	'I bake'	<i>triem</i>	'I rub'
<i>beriem</i>	'I take'	<i>(kryjiem &gt; ) kryjem</i>	'I cover'
<i>(kupujiem &gt; ) kupujem</i>	'I buy'	<i>smiem</i>	'I dare'
<i>beliem</i>	'I become white'	<i>hľadám</i>	'I look'
<i>(klaňám sa &gt; ) klaniam sa</i> 'I bow'			

based on the prototypical pattern of

<i>viem</i>	'I know'	<i>(jēm &gt; ) jem</i>	'I eat'
<i>((i)mām &gt; ) mám</i>	'I have'	<i>dám</i>	'I give'

The next step in the process of spreading 1sg  $-m$  is the removal of the constraint on the quality of  $V_1$ . Now any verb with a long  $V_1$  is subject to the influence of the prototypical athematic class. This brings class II with its  $V_1 -\bar{i}-$  on board for:

#### Phase 2

- Abductions: --  $m$  is a marker of first person  
-- there is a category of verbs with (underlying) long  $\bar{V}_1$   
-- class V is prototypical of this category
- Deduction: -- classes Ib, II, III, IV use the prototypical  
athematic 1sg in  $\bar{V}_1 + m$

and Phase 2 deduces *prosím* 'I beg.'

The lifting of the quantity restriction on  $V_1$  permits the last group of verbs (class Ia) to come under the influence of class V. Three factors gradually eroded the distinctive value of length in  $V_1$ :

- the rhythmic law of Slovak
- the presence of morphologized exceptions to the rhythmic law
- the shortening of diphthongs after  $j$

According to the rhythmic law of Slovak, wherever there are two consecutive long syllables (i.e., syllables containing a long vocalic segment or a diphthong), the second is shortened, as in class Ia: (*vydýchnu* >) *vydýchnu* 'they exhale,' (*ukážu* >) *ukážu* 'they show.' A parallel example from class IV is (*pomáháš* >) *pomáháš* 'you help.' Whenever the final syllable in the root of a verb was long, V<sub>1</sub> was shortened, yielding forms schematically parallel to those in class Ia:

Class Ia (with long stem syllable)	C $\bar{V}$ C - $\check{V}$ <sub>1</sub> -
other classes (with long stem syllable)	C $\bar{V}$ C - $\check{V}$ <sub>1</sub> -.

The recognition of certain short vocalic segments as underlyingly long was compromised by morphologized exceptions to the rhythmic law<sup>6</sup> and shortenings of long syllables attributable not to the rhythmic law, but to the absorption of the first part of diphthongs beginning in *i* by a preceding *j*. Particularly important here was the shortening of V<sub>1</sub> *ie* > *e* in position after *-j-*, as in *kupujěš* > *kupujieš* > *kupuješ* 'you buy' and *jěš* > *jieš* > *ješ* 'you eat.' The underlying quantity of V<sub>1</sub> was no longer reliably recoverable, and it ceased to play a determining role in the spread of *-m*, as we see in:

### Phase 3

- Abductions: -- *m* is a marker of first person  
 -- length of V<sub>1</sub> is not distinctive in classifying verbs  
 -- there is one category of verbs  
 (rhythmic law is not automatic; there are other shortenings; underlying length cannot be determined)
- Deduction: -- all verbs use the prototypical athematic 1sg in *-V<sub>1</sub> + m*, where V<sub>1</sub> can be long or short

which deduces: *vydýchnem* 'I exhale,' *ukážem* 'I show.'

The spread of *-m* in Czech follows a similar path, except that length did not develop in class Ib and Czech lacked the rhythmic law and other shortenings of V<sub>1</sub>, so there was no change in the role of quantity in this process. Athematic 1sg *-m* is thus spread only to classes II, III, and IV in Czech, by a series of abductive steps roughly equivalent to phases 1 and 2 (modified by the exception of class Ib). The spread of *-m* is

even more circumscribed in Polish, where class II did not have a long theme vowel and the process is limited to the equivalent of phase 1, producing 1sg forms in *-m* only in class III and class IV verbs (and recall that these classes were smaller because Polish did not carry out contraction as fully). Polish did not spread *-m* specifically to verbal classes (I and II) that had morphophonemic stem alternations; instead Polish enhanced these alternations to distinguish the 1sg and 3pl from all other forms in both classes of verbs:

class I:           *biorę* 'I take,' *bierzesz* 'you take,' *biorą* 'they take,'

          where we see *r* alternating with *rz* and *o* with *e*

class II:           *proszę* 'I beg,' *prosisz* 'you beg,' (*\*prosią >*) *proszą* 'they beg,'

          where we see the alternation of *sz* with *ś*.

Early prosodic changes produced in thirteenth century Serbo-Croatian a system with long  $V_1$  in all conjugation classes, so quantity was never an issue. The spread of *-m* followed a two-phase development, parallel to the first two phases in Slovak, with the exception that original *ě* did not merge with *e* in time for this change (or in some dialects ever), which means that *-m* spread to classes III and IV (constrained only by quality of  $V_1$ ) in phase 1 and to classes I and II in phase 2. 1sg *-m* now marks all verbs in Serbo-Croatian, with only two exceptions, the modal *moğu* 'I can' and the modal/auxiliary *hoću* 'I want/will.'

### 3.0 The spread of *-m* in Slovene

The strategy in Slovene was to regeneralize the theme vowel, and the 1sg and 3pl forms were perceived as: a) failing to conform to the otherwise universal non-past shape of stem + theme vowel + desinence, b) lacking unitary expression, c) lacking sufficient distinctiveness (since for many verbs the 1sg and 3pl had the same form). The spread of *-m* helped to solve all of these problems, and was accompanied by innovations in the 3pl forms. Phonemic vowel length did not play any role in the spread of *-m* in Slovene. Note that Slovene has preserved the dual forms in its conjugations, as indicated in the inventory of classes below.

The following relevant changes are particular to Old Slovene, and all of them took place in the 9th-10th cc. (cf. Lencek 1982):

$\bar{q}, q > o; \bar{e}, e > e$

$\bar{e}, \check{e} > e$

Here is an inventory of verbal classes in Slovene just prior to the spread of 1sg *-m* :

Class I: (sg) *-o, -eš, -e*, (du) *-eva, -eta, -eta*, (pl) *-emo, -ete, -o*

This class includes:

-- verbs in *\*-no-/-ne-*

*izdihno, izdihneš, izdihno* 'exhale'

*mino, mineš, mino* 'pass'

-- non-suffixed stems

*neso, neseš, neso* 'carry'

*peko/pečo, pečeš, peko/pečo* 'bake'

*tro, treš, tro* 'rub'

*bero, bereš, bero* 'take'

-- verbs in *\*-jo-/-je-* following a C

*pokažo, pokažeš, pokažo* 'show'

-- verbs with stem in *-Vj*

*krijo, kriješ, krijo* 'cover'

-- verbs in *-ova- / -uj-*

*kupujo, kupuješ, kupujo* 'buy'

Class II: (sg) *-o, -iš, -i*, (du) *-iva, -ita, -ita*, (pl) *-imo, -ite, -e*

This class includes

-- verbs in *-i-*

*prošo, prosiš, prose* 'beg'

-- the verbs of 'becoming'

*(bělěju, bělěš, bělějū) > belo, beliš, bele* 'become white'

Class III: (sg) *-ejo, -eš, -e*, (du) *-eva, -eta, -eta*, (pl) *-emo, -ete, -ejo*

This class includes non-derived verbs

*smejo, smeš, smejo* 'dare'

Class IV: (sg) *-ajo, -aš, -a*, (du) *-ava, -ata, -ata*, (pl) *-amo, -ate, -ajo*

This class includes

-- verbs in original *\*-ajV-*

*gledajo, gledaš, gledajo* 'look'

-- all derived imperfectives (except those in *-ova-* / *-uj-*)

*klanjajo se, klanjaš se, klanjajo se* 'bow'

Class V:

*vem, veš, ve, veva, vesta, vesta, vemo, veste, vede* 'know'

*jem, ješ, je, jeva, jesta, jesta, jemo, jeste, jede* 'eat'

*imam, imaš, ima, imava, imata, imata, imamo, imate, imajo* 'have'

*dam, daš, da, dava, dasta, dasta, damo, daste, dade* 'give'

If we make the generalization that the fourteenth century verb system presents a unified non-past conjugation, based on a pattern of stem + theme vowel + desinence, where there is one desinence per form in the paradigm, and each desinence consists of either a zero or a consonant that may or may not be followed by a vowel, we observe the following exceptions:

-- four verbs (from original class V) have *-s-* in their 2du, 3du, and 2pl desinences

-- the theme vowel is missing in the 1sg and 3pl forms of classes I and II

-- the following are not uniform

1sg, which has the exponents

*-o* classes I & II

*-theme vowel-j-o* classes III & IV

*-theme vowel-m* class V

3pl, which has the exponents

*-o* class I

*-e* class II

*-theme vowel-j-o* classes III & IV & *imajo*

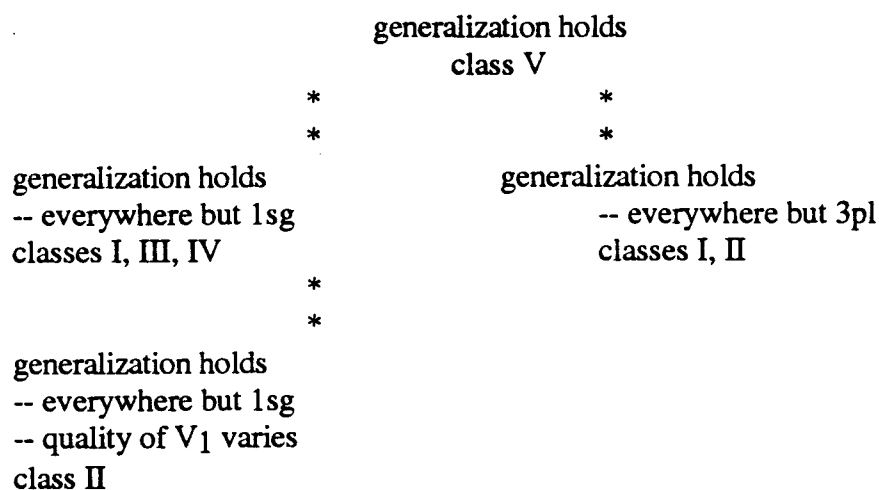
*-theme vowel-d-e* *dade, vede, jede*

In classes I, III, and IV, the 1sg and 3pl forms are identical. The spread of *-m* combined with the generalization of the theme vowel restored distinctiveness and produced verbal paradigms unique in the Slavic world.

Here is a formulation of the generalization and the hierarchical organization it imposes on the verbal system:

Generalization: the prototypical shape of a non-past verbal form is

stem + *-e/-a-* (theme vowel) + desinence



#### Phase 1

- Abductions: -- the vowels *-a-* and *-e-* in class V are theme vowels equivalent to the theme vowels in classes I, III, and IV  
 -- the 1sg forms in class V are prototypical because they conform to the pattern of stem + theme vowel + desinence
- Deduction: -- 1sg forms of verbs in classes I, III, and IV have the shape stem + theme vowel + *m*

Phase 1 (plus subsequent sound changes) deduces

<i>izdihnem</i>	‘I exhale’	<i>minem</i>	‘I pass’
<i>nesem</i>	‘I carry’	<i>pečem</i>	‘I bake’



<i>tarem</i>	'I rub'	<i>berem</i>	'I take'
<i>pokažem</i>	'I show'	<i>krijem</i>	'I cover'
<i>kupujem</i>	'I buy'	<i>smem</i>	'I dare'
<i>gledam</i>	'I look'	<i>klanjam sa</i>	'I bow'
based on the pattern of			
<i>vem</i>	'I know'	<i>jem</i>	'I eat'
<i>imam</i>	'I have'	<i>dam</i>	'I give.'

## Phase 2

- Abductions: -- all theme vowels have equal status,  
regardless of quality  
-- the 1sg forms in class V are prototypical  
because they conform to the pattern of  
stem + theme vowel + desinence
- Deduction: -- 1sg forms of all verbs have the shape  
stem + theme vowel + *m*

Phase 2 deduces: *prosim* 'I beg,' *belim* 'I become white.'

In Slovene the underlying assumption was that all non-past verb forms have a theme vowel (rather than a stem enlargement, as is likely the case with V<sub>1</sub> in West Slavic and Serbo-Croatian). This assertion is made on the basis of the changes that took place in phase 3 in Slovene, a phase that probably overlapped with phases 1 and 2.

## Phase 3

- Abduction: -- the 3pl form has the prototypical shape  
stem + theme vowel + *jo*
- Deduction: -- verbs in all classes have 3pl forms  
conforming to this shape

Phase 3 deduces forms such as: *izdihnejo* 'they exhale,' *nesejo* 'they carry,' *kupujejo* 'they buy,' *prosijo* 'they beg,' *vejo* 'they know,' *jejo* 'they eat,' *dajo* 'they give.'

The existence of forms like *izdihnejo* 'they exhale' and *kupujejo* 'they buy' confirm that the extension of theme vowels was the driving force behind these changes. If Slovene were concerned only with preserving distinctiveness, there would have been no need to

modify the original forms in *\*izdihno* and *\*kupujo*. Thus although the spread of *-m* in Slovene might appear at first glance to be part of an areal phenomenon shared with Slovak, Serbo-Croatian, and Macedonian, this result was obtained via a different set of abductions than in neighboring languages.

#### 4.0 The spread of *-m* in Macedonian (and Bulgarian)

In Macedonian, in contradistinction to Slovak and particularly to Slovene, the role of the theme vowel was not altered in the course of the spread of *-m*. Instead, sound changes created a situation in which all 1sg non-past verb forms ended either in *-a* or in *-am*. It was abducted that *-a* and *-am* were allomorphs of a single morpheme, and the more distinctive of the two, namely *-am*, was selected as the prototypical allomorph and extended accordingly.

Here are the relevant changes that yielded the sound system of Macedonian at the time of the spread of 1sg *-m* (cf. Koneski 1983, Janda & Friedman 1994):

loss of distinctive length	(prior to 13th c.)
$y > i$	(12-13th cc.)
$e > \check{e}; q > a$	(12-13th cc.)
$(j)\check{e} > ja$ in initial position	
$\check{e} > e$ elsewhere	(14-15th cc. or later, dep. on dialect)
$x > v / \_ C$	(16th c.)

Here is an inventory of Macedonian verb classes at the time of the spread of *-m* :

Class I: (sg) *-a, -eš, -e*, (pl) *-eme, -ete, -at*

This class includes

-- verbs in *\*-no/-ne*

*izdivna, izdivneš, izdivnat* 'exhale'

*mina, mineš, minat* 'pass'

-- non-suffixed stems

*nesa, neseš, nesat* 'carry'

*(peka >) peča, pečeš, (pekat >) pečat* 'bake'

*tra, treš, trat (> trija, triješ, trijat)* 'rub'

<i>bera, bereš, berat</i>	'take'
-- verbs in *-jo-/je- following a C	
<i>ukaža, ukažeš, ukažat</i>	'show'
-- verbs with stem in -Vj	
<i>krija, kriješ, krijat</i>	'cover'
-- verbs in *-ěje-	
<i>smeja, smeješ, smejat</i>	'dare'
<i>beleja, beleješ, belejat</i>	'become white'

Class II: (sg) -a, -iš, -i, (pl) -im(e), -ite, -at

This class includes:

-- verbs in -i-	
<i>proša, prosiš, prosat</i>	'beg'

Class III: This class is not distinct in Macedonian, for contraction of original -ěje- clusters did not take place, so all of its members remain in class I.

Class IV: (sg) -aja, -aš, -a, (pl) -ame, -ate, -ajat

This class includes

-- verbs in original *-aje- and derived imperfectives	
<i>gledaja, gledaš, gledajat</i>	'look'
<i>klanjaja se, klanjaš se, klanjajat se</i>	'bow'
-- suffixed verbs in original *-ova-	
<i>kupuvaja, kupuvaš, kupuvajat</i>	'buy'

Class V: [\**věděti* 'know' was lost early in the history of Macedonian]

<i>jam, jaš, ja, jame, jate, jadet/jadat</i>	'eat'
<i>imam, imaš, ima, imame, imate, imajat</i>	'have'
<i>dam, daš, da, dame, date, dadet/dadat</i>	'give'

Note that the loss of *\*věděti* 'know,' coupled with the change *jě-* > *ja-* in the verb for 'eat' meant that the predesinential vowel in all the athematic verbs was *-a-*.

To summarize, the pattern of non-past verb forms in Macedonian at the time was the following:

1sg	stem - <i>a</i>	classes I & II
	stem - <i>aj - a</i>	class IV
	stem - <i>am</i>	<i>imam, dam, jam</i>
2sg-2pl	stem - theme V - desinence	all classes
3pl	stem - <i>at</i>	classes I & II, <i>dadat, jadat</i>
	stem - <i>aj - at</i>	class IV, <i>imajat</i>

The verb *\*imati* 'have' was identified with class IV, and *-am* was generalized to class IV verbs. At this point, all 1sg forms were in either *-a* or *-am*, and this occasioned the spread of *-am* to all verbs. Subsequent to the generalization of *-am*, constraints on the shape of non-past verbal stems led to adjustments in the conjugation of the remaining two athematic verbs and non-syllabic verb types in class I. The generalization that could be drawn by language learners at this time and the hierarchies it entailed were:

Generalization: 1sg forms have the shape		
syllabic stem	+	<i>am</i>
*		*
*		*
valid for most verbs		valid for class V
*		*
*		*
not valid for <i>jam, dam,</i> and non-syllabic stems in class I		final segment <i>-m</i> is missing in classes I, II, III, IV

The changes motivated by this generalization can be summarized as follows:

#### Phase 1

Abductions: -- there are three classes of verbs, based on the theme vowel present in the 2sg, 3sg, 1pl, and 2pl forms

- *ima* 'have' is a prototypical class IV verb
- Deductions: -- all verbs have a theme vowel of *-e-*, *-i-*, or *-a-* in the 2sg, 3sg, 1pl, and 2pl forms
- class IV verbs have 1sg in *-am* (i.e., their paradigm follows that of *ima* 'have')

Phase 1 deduces *gledam* 'I look,' *kupuvam* 'I buy,' *klanjam se* 'I bow.'

Phase 2

- Abduction: -- the 1sg desinence is *-a(m)*, and *-am* is prototypical
- Deduction: -- all verbs have the 1sg desinence *-am*

Phase 2 (plus subsequent loss of *j* in sequences of *-eja-*) deduces:

<i>izdivnam</i>	'I exhale'	<i>minam</i>	'I pass'
<i>nesam</i>	'I carry'	<i>pečam</i>	'I bake'
<i>trijam</i>	'I rub'	<i>beram</i>	'I take'
<i>ukažam</i>	'I show'	<i>krijam</i>	'I cover'
<i>smeam</i>	'I dare'	<i>beleam</i>	'I become white'
<i>prosam</i>	'I beg'		

Phase 3

- Abductions: -- non-past verb forms are minimally disyllabic, following the pattern:
- |         |                                                            |
|---------|------------------------------------------------------------|
| 1sg     | (C <sub>n</sub> )VC <sub>n</sub> - <i>am</i>               |
| 2sg-2pl | (C <sub>n</sub> )VC <sub>n</sub> - theme V - desinence     |
| 3pl     | (C <sub>n</sub> )VC <sub>n</sub> - ( <i>aj</i> ) <i>at</i> |
- i.e., the verbal stem is syllabic
- the 3pl forms *dadat*, *jadat* belong to class I (parallel to *idat*, *vedat*, *kradat*, *kladat*, etc.) and have the stems *dad-*, *jad-*

Deductions: -- unprefixated resonant stems are all given syllabic shapes, e.g. *tr-* > *trij-* in *trijam* 'I rub'  
 -- the verbs 'give' and 'eat' have the following class

I conjugation:

<i>dadam</i>	<i>jadam</i>
<i>dadeš</i>	<i>jadeš</i>
<i>dade</i>	<i>jade</i>
<i>dademe</i>	<i>jademe</i>
<i>dadete</i>	<i>jadete</i>
<i>dadat</i>	<i>jadat</i>

Bulgarian underwent a phase 1 nearly identical to that of Macedonian, but because *q* > *ə* rather than *a*, the spread of *-m* stopped with the class IV verbs, since there was no motivation for the type of abduction that led to phase 2 in Macedonian.

5.0 Why *-m* did not spread in East Slavic Throughout West and South Slavic, the contractions that produced classes III and especially IV were crucial in facilitating the abduction that these classes could be categorized with the athematic verbs. East Slavic never experienced contraction and this precluded the spread of *-m*. The trend was instead to regularize athematic verbs as members of class I (cf. Russian *imet* 'have' and *vedat* 'know' which are regular, but *dat* 'give' and *est* 'eat' which have retained the athematic pattern). Morphophonemic alternations involving stress and stem final consonants play an important role in verbal conjugations in East Slavic, contrary to what we observe in West and South Slavic.

## 6.0 Conclusion

Analogy is a reinterpretation of categorial status. Items that are marginal in the morphology of one generation can be abduced to be prototypical by successive generations when sound changes have altered the perceptible relations among items in a category. This paper has traced the various routes by which an originally non-prototypical marker of 1sg was reinterpreted as prototypical for some or all of the verbal lexicon of the West and South Slavic languages.

<sup>1</sup>This research was aided by a grant from the Joint Committee on Eastern Europe of the American Council of Learned Societies and the Social Science Research Council. I would like to thank Victor Friedman and Craig Melchert for comments and corrections. All errors in the present text can be attributed only to me.

<sup>2</sup>The term "deiotation" is used here to refer to both the loss of *j* and the palatalization of the preceding consonant. The symbol Č is used to indicate the result of this process, which is either a palatal or a labial followed by palatal *l*.

<sup>3</sup>The fifth athematic verb, *byti* 'be' has been eliminated from discussion in this paper because this verb is highly irregular and has an idiosyncratic development in each Slavic language.

<sup>4</sup>The conjugations of verbs list segments following the stem of forms in the traditional order of 1sg, 2sg, 3sg, 1pl, 2pl, 3pl, although in the case of Slovene 1du, 2du, 3du have been inserted between the singular and the plural. The examples show only the 1sg, 2sg, 3pl, since the shape (excluding desinence) of the 3sg, 1pl, and 2pl (and in the case of Slovene, 1du, 2du, 3du) is the same as that of the 2sg.

<sup>5</sup>Stanislav (1967) and Krajčovič (1988) do not posit inherited length in the theme vowel for the subtypes in *-j-*, although they do posit length for the remaining subtypes. There are compelling reasons to include these two types here. One is structural: all other verbs with an original theme vowel *-e-* not preceded by an original consonant cluster belong to this class, (thus we see *pečieš* (< \*pek-e-ši) 'you bake,' but not *plačeš* (< \*plak-je-ši) 'you cry' in this group). The other reason is phonological: the initial portion of the diphthong *-ie-* was absorbed by the preceding *-j-* before any attestations are available, so it is impossible to prove that these two subtypes did not have a long V<sub>1</sub>.

<sup>6</sup>Examples of morphological environments that do not show the expected shortenings of the rhythmic law (the first two cited by Kenstowicz & Rubach 1987):

denominal (possessive) adjectives like	<i>kohútí</i>	'rooster's'
neuter collective nouns like	<i>prútie</i>	'twigs'
deverbal nouns like	<i>prianie</i>	'wish'
class II 3pl verb forms like	<i>páčia</i>	'they please'
verbs of 'becoming' like	<i>zmúdriet'</i>	'become wise'
idiosyncratic lexemes like	<i>úpiet'</i>	'lament.'

## WORKS CITED

- Andersen, Henning. 1973. "Abductive and deductive change." *Language* 49: 765-793.
- Janda, Laura A. & Victor Friedman. 1994. "About the *ja-* in *makedonskiot jazyk*: the fate of initial \**ě-* and \**ę-* in Macedonian." *Journal of Slavic Linguistics* 2.
- Kenstowicz, Michael & Jerzy Rubach. 1987. "The phonology of syllabic nuclei in Slovak." *Language* 63: 463-497.
- Koneski, Blaže. 1983. *A historical phonology of the Macedonian language*. Heidelberg: Winter.
- Krajčovič, Rudolf. 1988. *Vývin slovenského jazyka a dialektológia*. Bratislava: Slovenské pedagogické nakladateľstvo.
- Lass, Roger. 1990. "How to do things with junk: exaptation in language evolution." *Linguistics* 26: 79-102.
- Lencek, Rado. 1982. *The Structure and History of the Slovene Language*. Columbus: Slavica.
- Stanislav, Jan. 1967. *Dejiny slovenského jazyka. II. Tvaroslovie*. Bratislava: Slovenská Akadémia Vied.