0. Introduction

In June-July 1994, at the behest of the International Conference on Former Yugoslavia (ICFY) and under the sponsorship of the Council of Europe, an extraordinary census was carried out in the Republic of Macedonia. This was the first census taken after Macedonia became an independent country. The political conditions that led up to this event are beyond the scope of this article, but have been discussed elsewhere (Friedman 1996). Of significance for this article is the fact that in accordance with article 35 of the much-debated special law that governed the operation of that census, Romani, together with Macedonian, Albanian, Turkish, Vlah, and Serbian (the former Serbo-Croatian), was one of the six official languages in which the census was conducted. Romani was recognized as an official minority language in the 1991 Macedonian constitution insofar as the Romani people are explicitly named as a nationality of Macedonia in the constitution's preamble and Article Seven of that document guarantees nationality language rights. Nonetheless, the published materials connected with the 1994 census represent the first official use of Romani in Macedonian government documents, and they are thus intimately connected with the standardization of Literary Romani in that country. My paper will discuss the language of the census documentation in the context of Romani language standardization in the Republic of Macedonia. These documents show the increasing emergence of the Arli dialect of Skopje as the base of the standard language with certain compromises made in the direction of other dialects.

Beginning with the publication of Jusuf and Kepeski's Romani gramatika (1980), there have been sporadic efforts concerning the standardization of Romani for use in Macedonia, and since the Republic of Macedonia declared independence in 1991 there has been a significant increase in such activities. In November 1992 a conference was
sponsored by the Ministry of Education of the Republic of Macedonia for introducing Romani as a subject of study in elementary schools (Friedman 1995), and since November 1993 an independent Romani language newspaper Romano Sumnal 'Romani World' edited by Oskar Mahmut has appeared thrice (17.XI.93, 10.XII.93, 1.IV.94; see Friedman Forthcoming). There have also been other publications in Romani in the Republic of Macedonia, such as translations of Macedonian literature, e.g. Petroski (1989, 1992). Romani is used in radio and television broadcasting -- both private and state-sponsored -- and it has been used in both film and drama, but this has not resulted in a published written record. A textbook by Šaip Jusuf for teaching Romani as a language of study at the third grade level was officially published and announced to the public on 8 April 1996, but its release was delayed due to financial complications. The census materials were thus one of the few official uses of Romani to which large numbers of native speakers could be exposed at the time.

I should note here that complete and final figures on the number of people choosing to be censused in Romani, as well as those declaring Romani as their mother tongue are not yet available (as of June 1996). The preliminary 1994 figure for those declaring Romani as their nationality (Romani nacionalikano priperipe, Macedonian nacionalna pripadnost) was 43,732 or 2.3% of the total population of 1,936,877 (Nova Makedonija 15.IX.94, p.1). However, mother tongue and nationality do not represent a one-to-one correspondence, as demonstrated by such fact that Muslim is a nationality category but not a linguistic one.2 In some locations people with Romani as their native language chose to be censused in another language, usually Macedonian. Nonetheless, according to informal reports large numbers of Romani speakers at least in Western Macedonia -- which is where ethnic and nationality issues are particularly sensitive -- chose to be censused in Romani.3 This accords with my own observations during the census.4

The census materials that constitute the basis of this study consist of a 100-page instruction manual for census takers and three census forms -- one for each individual, one
for each household and one for agricultural holdings -- consisting of 6, 13 and 4 pages, respectively, for a total of 123 pages. In the case of the census forms, all questions were bilingual, with Macedonian first and Romani second. The instruction manual was entirely in Romani but translated from the Macedonian original by Šaip Jusuf and Mehmed Nedžat. Owing to the nature of the material, the linguistic data here is of a very particular type, since it represents a specific level of bureaucratic and professional terminology seeking to project an image of complete neutrality. Thus, for example, there are no first or second person verb forms in the entire corpus. All instructions refer to the census taker and the censused persons, and thus all verb forms are third person active, passive or middle. Similarly, the vocabulary itself has a high proportion of words not occurring in everyday conversation.

In the exposition which follows, I shall examine some of the most salient linguistic features of the census documents in the context of earlier relevant literary documents, viz. Jusuf and Kepeski (1980) and Romano Sumnal as treated in Friedman (1985 and Forthcoming, respectively). Wherever possible, reference will be made to the relevant sections of those articles. A complete analysis of the language of the census documents is beyond the scope of this paper, but those features focused on are symptomatic of the broader concerns of Romani language standardization in the Republic of Macedonia. These features demonstrate that the normativization of a Romani standard in the Republic of Macedonia is making consistent and perceptible progress.

1. ORTHOGRAPHY AND PHONOLOGY

1.01 Orthography and Typography

The orthography of the census forms is the Latin-based one generally in use in Eastern Europe and elsewhere (cf. Kenrick 1981, Hancock 1995). In general the quality of typesetting and proofreading was good, although there were occasional careless slips, e.g. majmut for majbut 'more', some inconsistencies in word divisions, e.g. nadžanela, omanuš for na džanela 'does not know', o manuš 'the person', some pronunciation-influenced
spellings, e.g. anaf, čhip for anav 'name', čhib 'language', where final devoicing occurs in some dialects under Macedonian influence. But the Romani norm is already sufficiently established that these spellings can be identified as mistakes rather than variants. Similarly, the occasional spelling phuv for phuv 'earth' represents an Arli dialectal pronunciation (aspirated voiceless bilabial stop > voiceless bilabial affricate). On the whole, the orthography of the census forms showed a significant consistency (cf. Friedman 1985:§1.0, Forthcoming:§1.01).

1.02 Schwa

The position of schwa in standard Romani is still a subject of dispute; this is especially so in Macedonia, where schwa is a marginal phoneme in all the Romani dialects and generally occurs in loanwords. Jusuf and Kepeski (1980) proposed the symbol <ä>, but this has not been widely adopted. The census forms vary among three solutions, all of which also occur in Macedonian. The prescribed Macedonian representation for schwa is an apostrophe, and this is used in the Romani forms such as s'ra 'column'. Despite this prescription, however, in the Macedonian press there are occasions where schwa will simply be omitted, generally when followed by a sonorant, and the same phenomenon occurs in the Romani census form in items such as jardmi 'help' (< Turkish yardım), gndinela 'think' (< Romanian gînd). In the case of the triplet haz'rdipe/hzrdipe/hazardipe 'preparation' (< Turkish hazı 'ready') we have in the first instance the standard Macedonian-type solution, in the second a leveling of both vowels to schwa with zero orthographic indication, and in the third the standard Macedonian treatment based on the west central dialects (schwa > /a/) (cf. Friedman 1985:§1.1, Forthcoming:§1.02).

1.03 Dental and palatal jotation

The census forms show considerably more consistency in the treatment of palatalized dentals and fronted velars than pervious efforts. Thus, for example, velar plus front vowel is consistently rendered as such: čhavenge 'children' (dat.), khere 'at home'. There is some variation in the representation of jotated dentals, although certain principles are
discernible. Thus, for example, the substantive **buti** 'work' has oblique forms **bukjake** (dat.), **bukengoro** (gen.) while the participle **dendo** 'given' has 3 sg. aor. **dengja**, 3 pl aor. **dende**. The middle verb **mothovgol** 'declare' (<thov-d>) has 3 sg aor. **mothovgja** and 3 pl aorist **mothovge**. In the case of the substantive, jotation is indicated throughout the oblique stem by spelling with a velar and the use of <j> before a back vowel, in the aorist forms, however, the underlying dental sometimes appears (cf. Friedman 1985:§1.4, Forthcoming:§1.03).

1.04 Jotation in feminine substantives

Certain oblique feminine stems in Romani are jotated in some dialects but not in others. The census materials are consistent in their use of the non-jotated forms of these nouns, e.g. **čhibakiri**, 'language' (gen.), **čhiba** 'languages' (cf. Friedman 1985:§2.1,Forthcoming:§1.04).

1.05 The treatment of /j/

The census materials show much greater consistency than previous efforts in the differentiation /j/ and /i/, i.e. writing <j> for the glide and <i> for the vowel, as in the following examples: **šaj** (only rarely **šai**) 'it is possible', **duj** 'two', **dujto** 'second', **sajbije** 'owner', **haibe** 'food', **leibe** 'permission'. For the instrumental singular, which consistently shows the Arli form, <j> is written only when the preceding vowel is not front, e.g. **lilea** 'sheet, form' (instr.), **čhibaja** 'language' (instr.). In the case of **tejsi** (< te isi 'if it is'), we seem to have a specific orthographic contraction (cf. Friedman 1985:§1.3,Forthcoming:§105).

1.06 The oppositions h/x

As was recommended at the 1992 codification conference (Friedman 1995), there is no graphic representation of the opposition between uvular /x/ and glottal /h/, which is occurs in some Romani dialects but is absent in the Arli dialect and does not go back to an original distinction in the parent language. In the census materials only <h> is used: **haibe**
'eating', *haljovela pes* 'it is understood, *hramonela pes* 'is written', etc., (cf. Friedman 1985:§1.2, Forthcoming:§1.06).

1.07 Clear vs palatal /l/

A considerable advance has been made in the representation of clear /l/, which, as in Macedonian, is automatic before front vowels. In previous documents, the digraph <lj> was used for both palatal /l/ before back vowels and inconsistently for clear /l/ before front vowels. In the census materials, the digraph /lj/ is used only before back vowels: *liil* 'sheet, form', *lela* 'take', *džangla* 'known' (cf. Friedman 1985:§1.5, Forthcoming:§1.08).

1.08 Aspiration

Romani aspirated consonants do not occur word finally and before another consonant, so a question of Romani orthography is whether or not to represent underlying aspiration in environments where it is neutralized. The census materials are consistent in not representing underlying aspiration in environments of neutralization, but occasionally fail to indicate aspiration when it should be spelled: *dikkeribe* 'examination", *dikhibe* 'viewing', but *mothovgja* 'declared' vs *motovgjo* 'declares'; *mukhibaja* vs *mukibaja* 'with permission' (cf. Friedman Forthcoming:§1.09).

1.09 The treatment of intervocalic and final -s- in inflections

In general intervocalic and final inflectional /s/ is lost in the Arli dialects of Macedonia, e.g. in the instr. sg., acc. anim. sg., 3 sg. aor., etc. The census materials are quite consistent in using the Arli forms, e.g. *čhibaja* 'language' (instr.), *lilea* 'sheet, form' (instr.), *kergja* 'did' (3 sg. aor.), *manuše* 'person' (acc.), but the 3 sg. reflexive pronoun, which also functions as an intransitive marker as in Macedonian, does show variation between Arli and non-Arli forms, e.g. *hramonela pe/pes* 'it is written' (cf. Friedman 1985:§1.3, Forthcoming:§1.10).

1.10 Instrumental plural (n+s at morpheme boundaries)

The change of /s/ to /c/ after /n/ at the morpheme boundary in the instr. pl. is consistently spelled: *manušencar* 'person' (instr. pl.) (cf. Friedman Forthcoming:§1.11).
1.11 Monosyllabic preposition + definite article

Romani has a number of monosyllabic prepositions that normally occur with the definite article. In some orthographies the article is separated from the preposition by an apostrophe or a hyphen, but in the census materials the two morphemes are consistently spelled together: *ki Republika Makeonija* 'in the Republic of Macedonia', *avazijal e taro dženo 6* 'in accordance with article 6', *bičhaldo pi buti* 'sent to work'. The preposition *baš-* 'for', which sometimes behaves like a monosyllabic preposition, is lexicalized in the census documents as *baši: karana baši i prezencija* 'reason for presence', *mukibajin baši o ačhovibe* 'with permission for residence' (Friedman 1985:§2.3, Forthcoming:§1.12).

2. MORPHOLOGY

2.01 The shape of nominative third person and possessive pronouns and the nominative plural definite article

The third person pronouns and definite articles are all consistently Arli: *ov, oj, ola, ple, o manuša* 'he, she, they, their (masc.), their (obl.), the people'. (Friedman 1985:§§2.3, 2.4, Forthcoming:§§2.01, 202).

2.04 Comparatives and superlatives

The census materials show an attempt to integrate Arli and non-Arli forms in their use of comparatives. While the superlative is indicated by the Arli prefix *em-* (of Turkish origin), the comparative uses both *po-* (of Macedonian origin) and *maj-* (of Romanian origin), this latter being able to mark both the comparative and the superlative in Vlah dialects: *pobut* 'more', but also *majbut*, e.g. *l berš thaj majbut* '1 year and over', *majhari* 'less', *majsereki* 'more rare', *emtikno* 'smallest', *emuči* 'highest' (Friedman 1985:§2.2, Forthcoming:§2.08).

2.05 Case usage after prepositions

The census materials show a mix of oblique case forms and nominative case with and without an additional preposition after certain prepositions of adverbial origin, e.g. *avrijal e phuvjatar* and *avrijal tari phuv* and *avrijal i phuv* all meaning 'outside the country'. From a dialectal point of view, Džambaz favors oblique case usage and Arli favors the more
analytical nominative usage, but clearly this is an area in which the compilers of the
census materials felt free to use variation.

2.06 Agreement interference

In some instances, the materials show interference from Macedonian that has implications
for the inflectional system. One such instance was in the use of adjectives borrowed from
Macedonian, where Macedonian plural -i and feminine -a inflectional endings were used
instead of the native Romani -e and -i, respectively: e.g. taro privatnikani karane 'for
private reasons' (Macedonian od privatni pričini), kvalitetna evidentiya 'qualified
documentation' (Macedonian kvaliteta evidenica).

2.07 Genitives

The language of the census materials shows a preference for preposed long genitives:
pretprijatengere bujakerere ranika 'business unit of the enterprise', avrijal phuvjakoro
manuš so dela buti 'foreign employer', Republika Makedonijakere raštrole oficijalnikane
dživdipaskere thane ki akaja adresa 'citizen of the Republic of Macedonia with official
place of residence at this address'. On the rare occasion of a postposed genitive, the long
form is still used, e.g. butikeribe e manušesekoro so parabela ple familija ki Republika
Makedonija 'employment of the person who supports his family in the Republic of
Macedonia' (cf. Friedman 1985:§2.3, Forthcoming:§2.03)

2.08 Abstract nominal derivation

The census forms seemed to follow a consistent policy of using -be to derive abstract
nouns from verbal stems and -pe to derive abstract nouns from non-verbal stems thus
assigning morphological functions to these two variants. Some examples are given here:
derverbal abstract nouns: leibe 'receipt', bijanibe 'birth', deibe 'giving', butikeribe
'employment', pučhibe 'question', peribe 'completion', dikhibe 'vision', kamibe 'desire',
haibe 'eating', mothovibe 'declaration', avibe 'arrival', džaibe 'going', bešibe 'stay', ačhovibe
'residence, sojourn', polagibe 'passing [an exam]' (from Macedonian polaganje),
specijaliziribe 'advanced study', registriribe 'census'
non-deverbal abstract nouns: *raštralipe* 'citizenship', *hazardipe* 'preparation', *avazjalipe* 'agreement', *čhavoripe* 'childhood'. The one exception in this category appears to be *dživdipe* 'residence, living', which, however, unlike the other deverbal nouns, has a participial base. (Friedman 1985:§2.1, Forthcoming:§2.09)

### 3.0 Syntax and Semantics

#### 3.01 Conditionals

The census materials are fairly consistent in using the Arli conditional formed by the subjunctive marker *te* plus the aorist, although usually reinforced with the older Turkism *eger* 'if': *Eger nesavo manuš te mulo* 'If someone has died', *Eger o manuš te nakhļja taro than ko than* 'if the person has emigrated', *eger o manuš te na džangļja* 'if the person does not know'. On occasion, the Macedonian construction *ako* 'if' plus aorist or present is also used *Ako o manuš meningja i adresa* 'if a person has changed address', *Ako e manušeske isi = Eger tejsi e manušeske* 'if the person has'. On occasion *eger* is used without *te*. These constructions all translate the same types of Macedonian condition (*ako* plus active or passive perfect or present), so the usages of the different Romani constructions appear to be conditioned by the desire for stylistic variation. (Friedman Forthcoming:§3.01)

#### 3.02 Vocabulary

The vocabulary of the census materials is remarkably creative, considering the type of language that was required. In the case of those shibboleths that serve to identify a given dialect with native speakers, Arli forms are consistently used, e.g. *thaj* 'and', *agjaar* 'thus'. The methods of vocabulary enrichment are of five types: neologisms and other constructs using native materials, Indicisms (neologisms based on Hindi or Sanskrit), Turkisms (items utilizing elements borrowings from Turkish during five centuries of Ottoman rule in Macedonia that are still in the colloquial registers of most Balkan languages), Macedonianisms (borrowings and calques based on Macedonian some of recent origin others probably older), and internationalisms (Greco-Latinate vocabulary that has entered
many languages of the world via the languages of the Great Powers, e.g. English, French and German). The following examples are illustrative of these five types.

Neologisms: *sastakeribe* 'recovery', *than ko pani başi mačhe* 'fish trap', *kombi vordona* "mini-van", *majčače* "or, that is to say" (Macedonian *odnosno*, German *bzw.*)

Indicism: *raštralipe* 'citizenship'

Turkisms: *mahalkerdo* 'population', *kabil bukjače* 'employable', *kašta emišea* 'orchard', *(these first three items combine Turkisms with native material)*, *sajbi je* 'owner', *barabarluko* 'unit/union'; *misafırlukoskoro* 'of the host', *askerluko* 'period of military servuce', *vakti* (periodi) 'time, period', *hamami* 'bathroom', *kenefi* 'toilet', *manzili ja kati* 'floor (storey)', *hali* 'situation'

Macedonianisms: *priperipe* 'affiliation' (Mac. *pripadnost*); *pretpijatije* 'enterprise', *stepen* (digra) 'degree'; *višne* 'sour cherry', *plugoja* 'plows'; *lela pe ko dikhibe* 'it is taken into consideration' (Macedonian *se zema predvid*)

Internationalisms: *pesticidija* 'pesticides', *edukacijakoro* 'educational', *registriribe* 'census' (Macedonian *registriranje* 'registration'), (cf. Friedman Forthcoming:§3.04).

### 4.0 Conclusion

The creation of standard languages is intimately connected with the maintenance of identities in the contexts of nation-states. Whether the language in question is that of the nation that lays claim to constituting the state or that of a national minority, the standard language itself is a vehicle for access to power and resources. While some argue that this can result in elitist practices, the counter-argument is that in the context of assimilatory pressures, standard languages are a necessary vehicle for access to education that does not result in loss of identity. In the case of Romani in the Republic of Macedonia, efforts by Romani speakers themselves, while conscious of the international movement, are taking place within a national context as a necessary first step. In the fifteen years since the first significant publication in this direction, significant progress has been made. In terms of the purposes for which it was initiated, the 1994 Macedonian census was a
statistical success and a political failure (see Friedman 1996). However, the Romani materials that resulted from it demonstrate success for the process of the standardization of Romani. These materials show an emerging Arli dialectal base with certain elements of compromise with other dialects, increasing orthographic and grammatical consistency, possibilities for stylistic variation, and a broad range of vocabulary building techniques making significant use of native material without becoming lost in purism.
1 Efforts at standardizing Romani elsewhere as well as on an international basis are considerably older, but they are beyond the concerns of this paper. See Hancock (1975, 1993, 1995), Kenrick (1981), Cortiade (1991) as well as Friedman (1995) and Matras (1996).

2 In the 1981 census, of 43,125 who declared Romani nationality, 36,399 declared Romani as their mother tongue. However, 1,697 declaring Albanian nationality declared Romani mother tongue, as did 316 claiming Macedonian nationality, 94 claiming Turkish nationality, 308 claiming Muslim nationality, 530 claiming Yugoslav nationality, 14 claiming Serb nationality, 2 claiming Vlah nationality, and 1,280 'others' (people claiming some other nationality, no nationality, a regional identity, or giving a facetious answer, e.g. lightbulb) for a total of 37,780 declaring Romani mother tongue in 1981. (Savezni zavod za statistiku 1988).

3 The following raw figures for use of Romani census forms were made available to me by Dr. Svetlana Antonovska, head of the Bureau of Statistics of the Republic of Macedonia. They are not to be taken as definitive, but they do give an idea of the use of the census forms.

<table>
<thead>
<tr>
<th>District</th>
<th>Total (P-1)</th>
<th>Romani (P-1)</th>
<th>Prelim. Total</th>
<th>Prelim. Romani</th>
</tr>
</thead>
<tbody>
<tr>
<td>Čair (Skopje)</td>
<td>88616</td>
<td>22500</td>
<td>85489</td>
<td>12706</td>
</tr>
<tr>
<td>Karpoš (Skopje)</td>
<td>126896</td>
<td>3300</td>
<td>125756</td>
<td>1969</td>
</tr>
<tr>
<td>Tetovo</td>
<td>179851</td>
<td>2500</td>
<td>174748</td>
<td>2407</td>
</tr>
<tr>
<td>Gostivar</td>
<td>112576</td>
<td>1100</td>
<td>108189</td>
<td>2092</td>
</tr>
<tr>
<td>Kičevo</td>
<td>54767</td>
<td>110</td>
<td>53044</td>
<td>1393</td>
</tr>
<tr>
<td>Kumanovo</td>
<td>127639</td>
<td>6200</td>
<td>126543</td>
<td>3121</td>
</tr>
<tr>
<td>Total</td>
<td>690345</td>
<td>35710</td>
<td>673769</td>
<td>23688</td>
</tr>
</tbody>
</table>
Explanation: The first column represents the total number of basic (P-1) census forms turned in for six of the 34 municipalities of Macedonia as reported between 25 and 31 July 1994 (Zapisnik za primo-predavanje na isečcite od Obrazec P-1). The second column represents the number of Romani forms turned in. The third column gives the total number counted for the municipality as reported in the preliminary results (Republika Makedonija, Zavod za statiststika, Popis '94, Prvi rezultati Soopštenie 2, 28 December 1994). The fourth column gives the preliminary total of those declaring Romani nationality (romska nacionalnost). Discrepencies between the P-1 and preliminary totals are connected with incomplete or otherwise disqualified census forms or other mechanical corrections. The discrepancies between Romani census forms used and declared nationality, however, also reflect the fact that in some municipalities more people declared Romani nationality without requesting to be censused in Romani, while in others more people were censused in Romani but declared some other nationality. The municipality of Skopje is divided into five districts. The district of Čair includes the predominantly Romani suburb of Šuto Orizari (Šutka). The preliminary total number declaring Romani nationality in the five districts of Skopje was 20,966.

4I was registered as an official observer of the 1994 Macedonian census in connection with my duties as a policy and political analyst for the analysis and assessment unit of the office of the Special Representative of the Secretary General to UNPROFOR (changed to UNPREDEP in Macedonia in 1995) from June-August 1994.

References cited


