

## The Thousand Accents' Island: Languages and Dialects in the Township of Semnan

Alexander Kolbitsch

M.A., Uppsala University, Sweden

Mirjam Kolbitsch

B.Sc., Uppsala University, Sweden

In cooperation with University of Uppsala, Sweden & University of Sistan and Baluchestan, Iran.

### Abstract

*This research aims to describe the distribution of speech forms in the township of Semnan. With this – to our knowledge – first detailed dialect survey of the region, we propose suggesting linguistic groupings of the local speech varieties, their boundaries, their relation to each other and to Mazandarani as well as an approximation of the number of speakers for each variety. In each of eleven reference points we interviewed knowledgeable informants, administered 30-40 questionnaires, collected material in the local speech variety and videotaped short stories or sentences, as well as two to four word lists with 220 words each. We transcribed these phonetically and with the help of WordSurv computer software compared them to each other as well as to Persian (Farsi) and Mazandarani. To determine the distribution of speech forms, we take into account the following criteria: previous research, the opinion of knowledgeable informants, the speakers' perception and lexicon- as well as phonostatic similarity. Our estimate of numbers of speakers is based on previous research, the estimate of knowledgeable informants as well as locals and on the National Census from 2006. We came to the conclusion that there are three languages spoken in the township of Semnan: Shahmirzadi, which is part of the Mazandarani group (ca. 6 000 speakers in the township of Semnan), Semnani (a continuum with approximately 30 000 speakers) and Sangesari (ca. 16 500 speakers).*

**Key words:** language, accent, dialect, Semnan, linguistic grouping.

## Preface

"The Thousand Accents' Island", that is what the people of Semnan call their own region. And indeed, it is a rather small island, surrounded by desert and mountains, on which a remarkable number of languages and dialects can be found also because it is situated on the Silk Road. Whether or not there really are 1 000 accents or actually Thousand-and-One is explored in this paper...

This paper represents the dialect-survey-part of a larger sociolinguistic survey, which itself is part of a joint research project between the University of Uppsala (U.U.), Sweden, and the University of Sistan and Baluchestan (U.S.B.), Iran, with the theme, "Language, identity and society - a documentation of minority languages in Iran, their sociolinguistic milieu and the role of the language in individual and group identity."<sup>1</sup> The project has been carried out with the help of eight students from U.S.B.<sup>2</sup>.

---

1-We want to express our gratitude to Prof. Dr. Carina Jahani from the department of Linguistics and Philology at U.U., to the chancellor of U.S.B. Prof. Dr. Ahmad Akbari, to Dr. Giti Mousapour Negari and Dr. Pakzad Yousefian from the Department of English Language and Literature, to Prof. Dr. Adam Torkamanzehi, Mr. Yahya Khaikhaei and Mr. Ali Esmaeli from the International Office of U.S.B. and Dr. Mohammad Reza Ashory as well as Dr. 'Esmat Esmā'ili from Semnan University. We want to thank our excellent teacher in sociolinguistics, Dr. Ted Bergman. We are very thankful for advice and feedback to Dr. Mark Karan, Mr. Behrooz Barjasteh Delforooz and Mr. Esmail Abbasi. We hope that our research can be a positive contribution towards understanding, cultural interaction as well as unity for the Iranian nation.

2-Mrs. Farideh Okati, Ms. Rayhanne Nooraeei-Nia, Ms. Marjan Amirabadi-Zadeh, Ms. Manije Khazaei, Mr. Aliakbar Choopani, Mr. Hasanali Kadkhoda, Mr. Bakhtiyar Seddighy-Nejad, Mr. Vahid Mohamadi

## 1. Introduction

### 1.1. Abbreviations of Place Names and Map Index

AF	Aftar, Aftari = Aftare = Afdari, افتر
AL	Ala = Kala, Alai = Kelayi, علاء
AS	Asad Abad, آسَد آباد
ATD	Artad, آرْتَد
AV	Arvane = Arvone, اروانه
BI	Biabanak, بیابانک
CS	Chashm, چشم
CT	Chehel Tan = Cheheltan, چهلتن
DA	Darjazin, درجزین
DL	Delazian = Delaziun, دلازیان
DS	Deh Sufian, ده صوفیان
EA	Eslam Abad, اسلام آباد
EZ	Emamzade Abdullah, امامزاده عبدالله
FA	Farsi = Persian, فارسی
HA	Hasan Abad, حَسَن آباد
HDA 1&2 <sup>1</sup>	Haji Abad, حاجی آباد
HK	Hiku, هیکو
JV	Joveyn, جوین

---

1-One HDA is near Ala and another one is near Lasjerd.

KA	Kheyr Abad, خیر آباد
KK	Khatir Kuh, ختیر کوه
LA	Lasjerd = Lasgerd, لاسجرد
MA	Mazandaran = Mazanderan, مازندران
MM	Mumen Abad, مومن آباد
NE	Nezami, نظامی
PO	Poshte, پُشته
RA	Rokn Abad, رُکن آباد
SA	Sufi Abad, صوفی آباد
SE	Semnan <sup>1</sup> , سمنان
SH	Shahmirzad, Shahmirzadi = Shamzayi, شهمیرزاد
SO	Sorkhe, سُرخه
SS	Mehdi Shahr <sup>2</sup> = Sangesar = Sengiser = Sangsar = Sangisar = Sengsar, مهدی شهر
TA	Taleb Abad, طالب آباد
YA	Yousef Abad = Kandu = Kendu = Kond, کندو

- ... Reference point.  
 ⊙ □ ... Place, mentioned by informants, but not visited and investigated.  
 — ... Country border.  
 — — ... Main road.

1-Note that “Semnan” is at the same time name of the province (ostan), of the township (shahrestan) and of the city.

2-In 1979 Sangesar was renamed Mehdi Shahr. To avoid confusion we use the short form SS in our maps, as the language spoken there is still called “Sangesari”.

 ... Provincial border.

## 1.2. Aim

This research aims to contribute to the wider goal of researching all the speech varieties in Iran. Our approach is a linguistic and sociolinguistic one, focusing on structural (in our case lexical) comparison and language attitudes. The aim of this paper is to describe the distribution of speech forms in the township<sup>1</sup> of Semnan. With this – to our knowledge – first detailed dialect survey of the region we propose suggesting linguistic groupings of the local speech varieties, their boundaries, their relation to each other and to Mazandarani, as well as an approximation of the number of speakers for each variety.

## 1.3. Fieldwork

In March 2008 we conducted fieldwork in the township of Semnan with the help of eight linguistics students from U.S.B.. Semnan is situated 220 km East of Tehran.

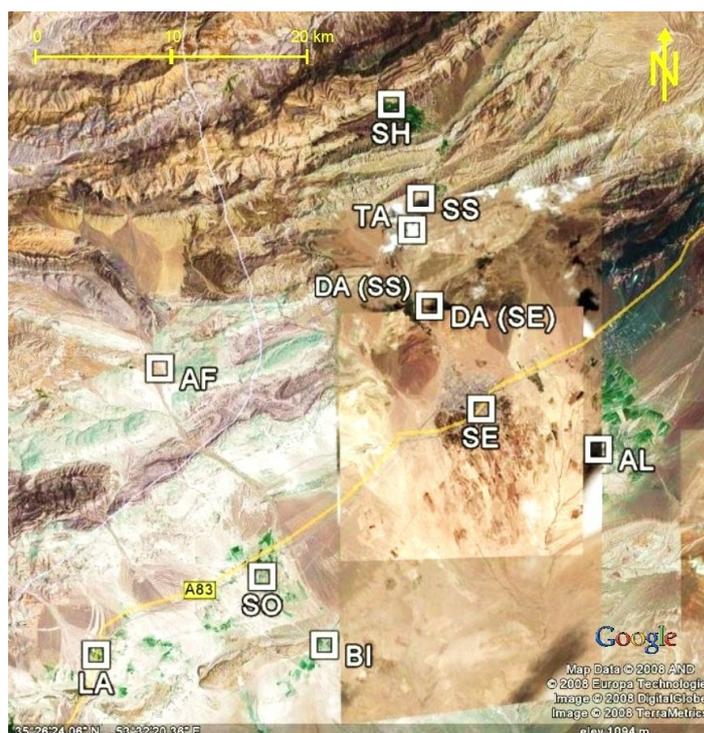
We selected the following reference points: Shahmirzad, Mehdi Shahr (= Sangesar), Darjazin, Taleb Abad, Atari, Aftar, Semnan City (Northern part and Bazaar), Ala, Biabanak, Sorkhe and Lasjerd. Our decision was based on the following criteria: size<sup>2</sup>, all central villages according to Gharib (1981)<sup>3</sup> and all places, which give their names to speech forms mentioned in important existing research. The choice includes both small and large, as well as rural and urban places.

---

1-In Iran each ostan (province) is divided into several shahrestans (townships). Ostan Semnan (570 835 inhabitants) consists of the three shahrestans: Semnan (186 159), Damqan (81 993) and Shahrud (225 007). (Statistical Centre of Iran, <http://www.sci.org.ir/>)

2-All places, which according to the National Census from 2006 (Statistical Centre of Iran, <http://www.sci.org.ir/>) have more than 1 000 inhabitants were chosen, except Kheyr Abad, Rokn Abad, Mowmen Abad and Shahrak Afghane which are satellite cities of Semnan (city).

3-The intention of Gharib's survey was to locate strategically important villages for the development of the region with respect to infrastructure, size and trade.



**Map 1: Chosen Reference Points**

In Darjazin, where most people speak Sangesari, there also live a number of Semnani speakers. Therefore the square in the map has two labels: DA (SS) and DA (SE). When visiting Atari we found the place to be deserted and subsequently removed it from the list.

In each of the reference points we did the following:

- Talked to knowledgeable informants<sup>1</sup>.
- Collected material<sup>2</sup> in the local language.

---

1-We asked them for literature, other knowledgeable informants we should meet, number of speakers, their perception of how the speech forms should be grouped and then video-recorded them reading a standard word list in their variety.

2-Magazines, newspapers, books, CDs and tapes.

- Video-recorded two to four standard word lists<sup>1</sup>.
- Video-recorded short stories or sentences in the local speech variety.
- Administered 30-40 questionnaires<sup>2</sup> to a certain sample of inhabitants as follows:

We did quota sampling according to three criteria: gender, age and educational level. We asked 15 men, 15 women, ten informants aged 9-25, ten informants aged 26-50, ten informants aged 51-100, ten illiterate informants, ten informants educated in schools and ten informants educated in universities. Literate informants were asked to fill in the questionnaires anonymously themselves and to illiterate informants the questionnaire was administered orally.

#### 1.4. Data Analysis

We took only those answers from questionnaires into consideration, that were given by informants who grew up in the reference point, had not lived in another location for more than three years, and whose father and mother came from there. We rejected those questionnaires which contained contradicting or illogical answers. If the questions Q8 and Q9 (Appendix B) were answered with “a little” or “most words”, we regarded that as indicating that the informants perceive the speech form as a different language. We regarded the answers “a lot” and “some words”, as indicating that the informant perceives the speech form as a different dialect, and the answers “completely” and “no difference” as the same dialect. Only when 50% or more of the informants agreed on one answer, we considered it relevant for the result. We applied the same procedure for question Q11 (Appendix B). When defining names and borders of speech forms we regarded those villages

---

1-See appendix C). Some of the informants wrote down the words in their own speech variety before they read them out and we recorded them.

2-See appendix B). We took the questions from Callister (1998), Chapter “3.4.6.1. Linguistic Similarity”, revised and translated them into Persian with the help of the assistants.

belonging to a language or dialect, that were mentioned in answers to questions Q12 - Q18 (Appendix B) in three or more questionnaires.

We also took into consideration small villages that were mentioned to be similar or identical to a reference point in answers to the questions in the beginning of this chapter.

For each reference point we produced a slightly simplified phonetic transcription<sup>1</sup> from the filmed standard word lists. When merging differing word lists from one reference point we eliminated Persian loanwords in the first place and secondly chose those words that occurred most frequently among the lists from that same place. With the help of the computer software WordSurv we counted the percentage of apparent cognates and phonostatic similarity<sup>2</sup> of apparent cognates from each of the speech forms investigated, including Persian<sup>3</sup> and Mazandarani<sup>4</sup>. To draw the line between language and dialect we applied the standard of 70% in lexicostatic similarity as upper confidence limit, which means that values below 70% certainly indicate a different language. We interpreted values between 70% and 83% of lexicostatic similarity as disputable with regard to their relevance for deciding between language and dialect. Numbers above that value were regarded indicating that the speech varieties surely constitute one language.

To complement the lexicostatic comparison, which is often regarded not to be sufficiently reliable, we also determined the phonostatic similarity. WordSurv counts the degree of difference between sounds that occur in variation in apparent cognates by adding up the number of minimal steps in

---

1-The phonetic inventory of our transcription can be found in appendix D).

2-This is done by determining the degree of difference (number of minimal steps in difference of articulation) between any of the sounds that occur in variation in apparent cognates.

3-For that purpose we transcribed video-recordings of our assistants, which were taken when they read out the Persian words to informants.

4-We transcribed a recording that was done by other linguists from a Mazandarani speaker, living in Europe. The high similarity to Persian suggests that his Mazandarani is heavily persified.

difference of articulation<sup>1</sup>. To facilitate the comparison of results and drawing linguistic maps, we used the computer software Google Earth<sup>2</sup>.

Our final decision on language- and dialect groupings is, besides lexico- and phonostatic similarity, based on previous research, the opinion of knowledgeable informants and the perception of locals.

### 1.5. Constraints

The discussion about languages and dialects in the township of Semnan would benefit from a linguistic analysis of all the speech forms in the region, from a thorough comparison of those speech forms and from intelligibility testing<sup>3</sup>, all of them being beyond the scope of this paper. We continue our research, focusing on other sociolinguistic features, among them language attitudes, language use, bilingualism and language vitality. The relation of languages and dialects in the township of Semnan to other Iranian languages is another task that remains to be researched. Both existing research and informants in the region mention similarities to Mazandarani/Taberi<sup>4</sup> (Qasemi Shahrizadi 1379: 24), Gilaki (Anjoman-e Resālat-e Qalam-e Sorkhe 1386 AH: 56 and Esmā'ili 1384 AH: 18), Lori (Anjoman-e Resālat-e Qalam-e Sorkhe 1386 AH: 56), Azari (Esmā'ili 1384 AH: 18), Gorgani (Esmā'ili 1384 AH: 18), Dari, Ancient Persian, Middle Persian (Esmā'ili 1384 AH: 13), Pahlavi (Anjoman-e Resālat-e Qalam-e Sorkhe 1386 AH: 56 and Qasemi Shahrizadi 1379: 24), Tati (Esmā'ili 1384 AH: 18 and Windfuhr 1965: 4) and Raji (Windfuhr 1965: 4). According

---

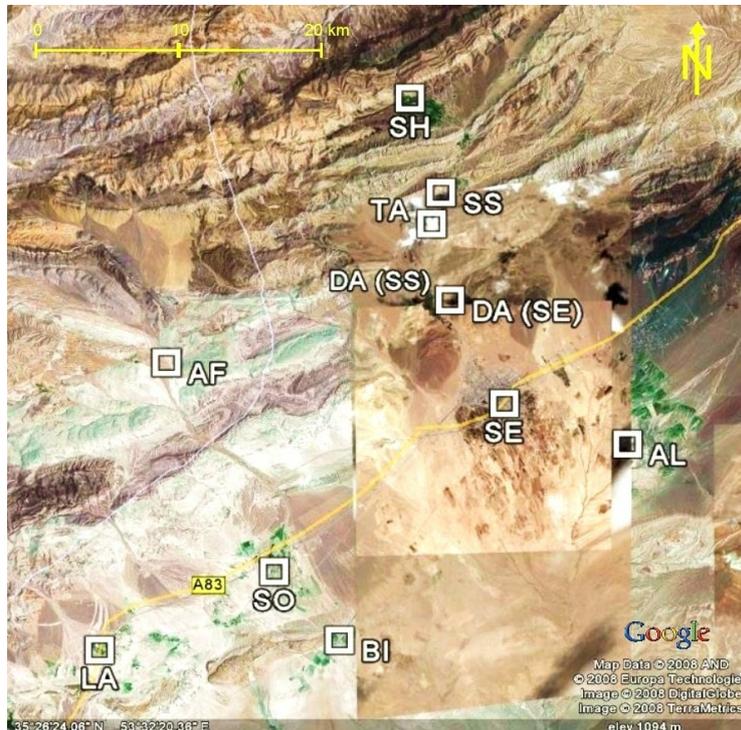
1-Therefore a lower degree of difference represents a closer relationship.

2-Permission for publication was given by Google Earth (Reference Number: #362158266).

3-Since it was very hard to find people who cannot speak Persian in these regions, we suspect that bi- or multilingualism is wide spread. When doing intelligibility testing, special attention has to be given to that.

4-or Tabrestani.

to two independent informants in Mehdi Shahr even a Sangesari – Sistani dictionary exists. The status of Shahmirzadi in the context of Mazandarani languages has to be investigated in detail.



## 2. Languages and Dialects in the Township of Semnan

### 2.1. Publications and Research

The early explorers of the region are Khanikov (1858), Houtum Schindler (1876)<sup>1</sup>, Bernhard Dorn (1859), Zukovskij (1888), M. James Bassett, M. A. Querry (1896), J. de Morgan and Geiger<sup>2</sup>, Oskar Mann (1903) and A. Christensen (1915-1935), who is mentioned frequently in connection with Semnani (Majidi 1980: 3f; Windfuhr 1965: 5f).

During the 20<sup>th</sup> century Georg Morgenstierne published two linguistic articles on Semnani in the Norwegian Journal of Linguistics in 1958 and 1960, followed by Manoochehr Sotude's dictionary in 1342 AH, Gernot Ludwig Windfuhr's study of verbal morphemes in Sangesari from 1965 and a dictionary of Sangesari with a grammatical outline from the same author published together with Cheragh Ali Azami in 1972. Majidi published his description of Semnani in 1980 and Homādokht Homayun published his of Aftari in 1371 AH. A significant event was the "First National Conference on Semnani Speech Variety" at the University of Semnan in January 2004. The editors of the compendium of this conference are 'Esmā'ili and Mostafa Jabbari. The department of Speech Therapy in the faculty of Therapy at Mazandaran University of Medical Science carried out a study on bilingualism among Semnani children. Z. Eftekhari, A. Sadollahi and F. Kasbi were involved in this project. Brand new are Esmail Hemmati's description of Semnani "Sarzamin Gujeshha" from 1386 AH and the "Dictionary of Semnan Ancient Dialect" in two volumes, by Mohammad Hasan Javaheri and Parviz Pezhom Shariati from 1386 AH. The Semnan area is one main focus of Donald Stilo's Northwest Iranian Project at the Max Planck Institute for Evolutionary Anthropology in Leipzig<sup>3</sup>. Windfuhr's

---

1-"Bericht über den semnanischen Dialekt".

2-"Grundriss der Iranischen Philologie".

3-<http://www.eva.mpg.de/lingua/research/northwest-iranian.php>

new publication “The Iranian Languages”, planned for March 2009, will certainly make a significant contribution.

Other, not specifically linguistic but nevertheless important, publications for sociolinguistic research are Connell's (1969) and Panāhi Semnāni's (1383 AH) description of the people of Semnan. In 1981 Gharib studied the infrastructure of the region from the point of view of town planning. Very useful is the publication of the National Census on the Internet by the Statistical Centre of Iran (<http://www.sci.org.ir/>) from 2006.

## 2.2. Classification of Languages and Dialects in Previous Research

The Ethnologue<sup>1</sup> classifies Semnani (smy, 21 099 speakers), Lasgerdi (lsa), Sangisari (sgr) and Sorkhei (sqo) as Semnani languages and designates the following lineage to this group: Semnani – Northwestern – Western – Iranian – Indo-Iranian – Indo-European. Shahmirzadi is classified together with Gilaki and Mazandarani in the Caspian group<sup>2</sup>.

- **Shahmirzadi:** Christensen has three groups: 1) Semnani, 2) Sangesari and Sorkhei-Lazgerdi (which he regards to be similar to dialects between Kashan and Esfahan) and 3) Shamirzadi, similar to Mazandarani.

The similarity between Shahmirzadi and Mazandarani has also been recognized by Hemmati<sup>3</sup> (1386 AH: 90) and Qasemi Shahmirzadi (1379 AH: 24), who points out that Shahmirzad has very close ties to Mazandaran because many inhabitants were born there and because it used to belong to the province of Mazandaran. He writes that people from Shahmirzad understand 90% of Mazandarani whereas they only understand 40% of the languages spoken in the region of Semnan.

---

1-[http://www.ethnologue.com/show\\_family.asp?subid=90021](http://www.ethnologue.com/show_family.asp?subid=90021)

2- Caspian – Northwestern – Western – Iranian – Indo-Iranian – Indo-European.

3-Hemmati (1386 AH: 90) also mentions Deh Sufian in connection with Shahmirzadi and Mazandarani.

Hoseini in Hemmati (1386 AH: 89) investigates four small villages<sup>1</sup> on the border between Semnan and Mazandaran, thus shedding light on the transition.

- **Sangesari:** The independence of Sangesari to Shahmirzadi has been pointed out by Zukovskij (Windfuhr 1965: 4).

The article “Sangiseri language” ascertains the independence of Sangesari from Sorkhei, Lasjerdi, Biabanaki and Aftari<sup>2</sup>.

- **Aftari – Sorkhei – Lasjerdi:** In his article “AFTARĪ” in Encyclopaedia Iranica<sup>3</sup> Windfuhr groups Aftari together with Sorkhei-Lasjerdi and relates all of them to Sangesari, Shahmirzadi and furthermore to Tabari/Mazandarani:

Aftarī is closely related to Soṛajī-Lāsgerdī; together with Sangesarī and Šahmīrzādī these dialects form a distinct semicircle around Semnān and are related to labarī/Māzandarānī and, more remotely, to Harzanī in north Azerbaijan and Zāzā in eastern Anatolia (...). Aftarī with Soṛajī and Lāsgerdī developed certain *Sprachbund* features with Semnānī,...

Windfuhr here suggests that Aftari, Sorkhei and Lasjerdi developed similarities to Semnani due to proximity.

Hemmati (1386 AH: 90) and the article “Dialects of the Semnan Region” (1977: 101) also confirm the relation between Aftari and Lasjerdi-Sorkhei.

- **Semnani:** Majidi (1980: 19) and Sotude (1342 AH: 2)

1-Chashm, Prur, Aftar, Deh Sufian.

2-It does so by pointing out certain grammatical features like the distinction between male and female verb forms.

3<http://www.iranica.com/newsite/index.isc?Article=http://www.iranica.com/newsite/articles/v1f6/v1f6a028.html>

write that Biabanaki, Rokn Abadi, Sufi Abadi, Kheyf Abadi and Alai are dialects of Semnani. Maqsudi (1381 AH: 35) on the other hand emphasizes that their speech is quite different to Semnani.

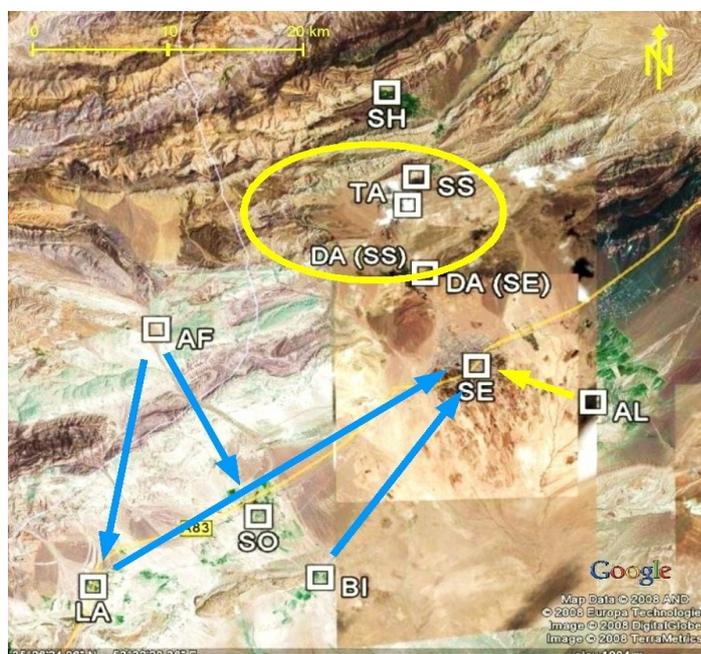
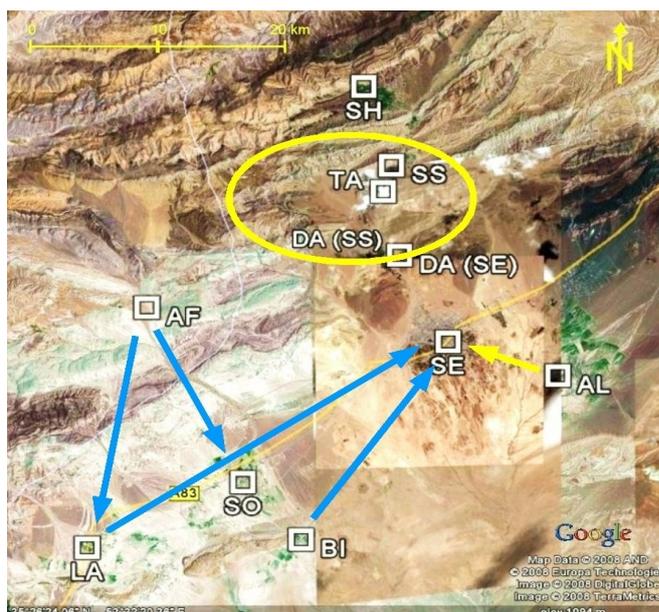
### 2.3. Survey Results

#### 2.3.1. Questionnaires: Perception of Informants

More than 50% of the relevant informants, who completed a questionnaire in each reference point, agreed that the following places have similar or identical speech:

**Table 1: Perception of Informants Regarding Similar and Identical Speech Varieties**

<u>AF</u> : similar: LA, SO, AV	<u>SS</u> : identical: DA, TA
<u>AL</u> : similar or identical: XA; identical: SE	<u>SE</u> : -
<u>BI</u> : similar or identical: SE	<u>SH</u> : -
<u>DA</u> : identical: SS, TA	<u>SO</u> : -
<u>LA</u> : similar: SE	<u>TA</u> : identical: DA, SS



**Map 2: Questionnaires: Perception of Informants Regarding Similar and Identical Speech Varieties**

... Identical

... Similar

Results suggest that people regard Sangesari, Taleb Abadi and Darjazini as belonging together and that people from Ala regard their speech to be identical to Semnani. Aftari is perceived to be similar to Lasjerdi and Sorkhei. Lasjerdi and Biabanaki are perceived to be similar to Semnani.

More than 50% of the questionnaires from Aftar and from Shahmirzad indicate that people understand Mazandarani completely. For all the other reference points more than 50% of the questionnaires indicate that Mazandarani is either not understood at all or just a little. In no reference point did more than 50% of the informants reply that they understand Mazandarani "a lot".

The following villages were mentioned in three or more questionnaires in answers to the questions Q12 - Q18 (Appendix B).

We also included small villages that were perceived to have similar or identical speech forms to a reference point (above). Then we included them in the table even if they were not mentioned three times or more in answers to the questions above.

**Table 2: Places Associated with Reference Points**

Reference Point	Associated Places
AF	AV, MM, Emamzade Abdullah
AL	YA
BI	AS
DA	EA, CT, Enqelab Shahrak, Shahrake Amir Almomenin, Poshte Bala va Pain Many informants regard the language spoken in DA to be Sangesari and even think that it is better spoken there. SS is regarded as the centre of the language.

LA	BI, SO, AF, AS, DV, Shomali <sup>1</sup> , HDA 2
SS	SE, Mazandarani, Gilaki, SH, Ziarat, ATD DA and TA are perceived to have the same language. Most of the times SS is regarded to be the centre of Sangesari language but several informants mentioned DA as the place, where Sangesar is spoken best <sup>2</sup> .
SE	DL, HA, RA, XA and “Southern villages”, SA, SH, SS, DA, AL Whereas SE is often regarded to be the centre of SE language, the villages XA, RA, HA, DA were mentioned to have the best SE dialect, most probably because many immigrants live in the city.
SH	Mazandarani, Gilaki, CS, DS, KK, Fulad Mahale <sup>1</sup> , SS, HK One informant mentioned that SH has lots of trade with Mazandaran.
SO	SE, BI, NE, AF, SS, SH, AL
TA	Ziarat, SE, MA Though mostly SS is regarded to be the centre of Sangesari some informants thought that TA is the place where the language is spoken best.

### 2.3.2. Lexicostatistic and Phonostatistic Analysis

**Table 3: Lexicostatistic Similarity**

Variety	FA	MA	SH	SS	TA	DA - SS	DA - SE	SE	AL	BI	LA	AF	SO
FA	100%	85%	77%	66%	72%	72%	74%	72%	73%	78%	73%	74%	70%
MA	85%	100%	81%	70%	76%	76%	72%	73%	74%	77%	77%	78%	72%
SH	77%	81%	100%	76%	78%	78%	72%	76%	77%	78%	76%	80%	75%
SS	66%	70%	76%	100%	85%	87%	73%	72%	78%	74%	73%	79%	76%
TA	72%	76%	78%	85%	100%	92%	76%	76%	79%	81%	80%	81%	77%
DA - SS	72%	76%	78%	87%	92%	100%	78%	77%	81%	80%	79%	83%	78%
DA - SE	74%	72%	72%	73%	76%	78%	100%	80%	82%	83%	82%	79%	75%
SE	72%	73%	76%	72%	76%	77%	80%	100%	86%	85%	84%	81%	81%
AL	73%	74%	77%	78%	79%	81%	82%	86%	100%	88%	85%	83%	83%
BI	78%	77%	78%	74%	81%	80%	83%	85%	88%	100%	90%	83%	82%
LA	73%	77%	76%	73%	80%	79%	82%	84%	85%	90%	100%	80%	83%
AF	74%	78%	80%	79%	81%	83%	79%	81%	83%	83%	80%	100%	80%
SO	70%	72%	75%	76%	77%	78%	75%	81%	83%	82%	83%	80%	100%

1-Mazandarani or Gilaki.

2-The reason for that could be that SS as a city has more immigrants.

**Table 4: Phonostatistic Similarity<sup>2</sup> in Degrees of Difference**

Variety	FA	MA	SH	SS	TA	DA - SS	DA - SE	SE	AL	BI	LA	AF	SO
FA	100%	85%	77%	66%	72%	72%	74%	72%	73%	78%	73%	74%	70%
MA	85%	100%	81%	70%	76%	76%	72%	73%	74%	77%	77%	78%	72%
SH	77%	81%	100%	76%	78%	78%	72%	76%	77%	78%	76%	80%	75%
SS	66%	70%	76%	100%	85%	87%	73%	72%	78%	74%	73%	79%	76%
TA	72%	76%	78%	85%	100%	92%	76%	76%	79%	81%	80%	81%	77%
DA - SS	72%	76%	78%	87%	92%	100%	78%	77%	81%	80%	79%	83%	78%
DA - SE	74%	72%	72%	73%	76%	76%	100%	80%	82%	83%	82%	79%	75%
SE	72%	73%	76%	72%	76%	77%	80%	100%	86%	85%	84%	81%	81%
AL	73%	74%	77%	78%	79%	81%	82%	86%	100%	88%	85%	83%	83%
BI	78%	77%	78%	74%	81%	80%	83%	85%	88%	100%	90%	83%	82%
LA	73%	77%	76%	73%	80%	79%	82%	84%	85%	90%	100%	80%	83%
AF	74%	78%	80%	79%	81%	83%	79%	81%	83%	83%	80%	100%	80%
SO	70%	72%	75%	76%	77%	78%	75%	81%	83%	82%	83%	80%	100%

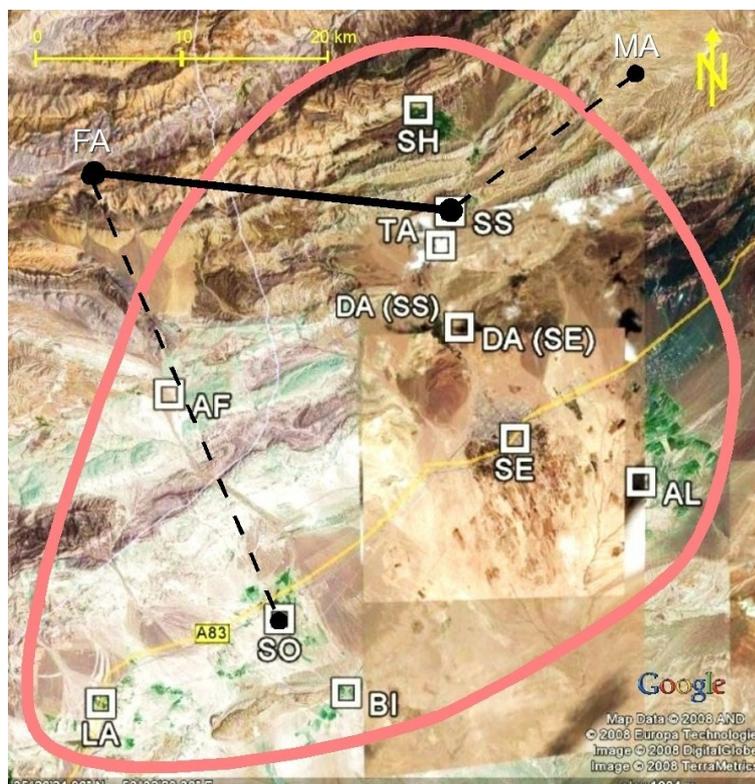
In the following we depict the results of the lexico- and phonostatistic analysis of the word lists in form of maps. The different maps show groupings at a certain value of similarity. Map three, showing groupings at 84% in lexicostatistic similarity and 80 degrees of difference in the phonostatistic comparison<sup>3</sup>, depicts the peaks of similarity. Map four represents the upper confidence limit, below which speech varieties most certainly constitute different languages.



**Map 3: Peaks of Similarity: Lexicostatistic Similarity of 84% and above and Phonostatistic Similarity of 80 Degrees and Below**

- 1-We did not consider this place in drawing the map because it is rather far to the North East of the surveyed region.
- 2-The lower the number, the greater the similarity (compare footnote 21).
- 3-These numbers emerged from the data to be a suitable threshold to identify the peaks of similarity.

Map three shows that there is a dialect continuum ordered in a circle around Sorkhe: Semnani (of Semnan (city)) – Alai (which could as well be classified as an accent of Semnani of Semnan (city)) – Biabanaki – Lasjerdi – Aftari. The value of lexicostatistic and phonostatistic similarity<sup>1</sup> of each speech variety in this continuum is higher to the neighbouring speech forms than to any other. Besides Aftari the continuum spreads out in the plain along the road and the railway.



#### 4: Lexicostatistic Similarity of 70% and Above

Map four shows that from the point of view of lexicostatistics (which is only one of our criteria), only Sangesari can be declared a language different from Persian (indicated here with a

1-For Aftari this is only true for phonostatistic similarity.

solid line), for sure. Just on the limit with 70% (indicated here with dotted lines) are Sorkhei - Persian and Sangesari - Mazandarani. These can also almost certainly be regarded as distinct languages.

### 2.3.3. Estimate of Numbers of Speakers

The estimate of numbers of resident speakers<sup>1</sup> is based on the estimate of various informants and on the National Census from 2006, assuming that in Semnan (city) approximately 10% of the inhabitants speak Semnani (many people from other regions moved to Semnan to find work<sup>2</sup>). Other cities above 5 000 inhabitants were assumed to have 70% of speakers and villages 80%. A number of very small locations (all of them had less than 254 and most of them less than 100 inhabitants) have not been taken into account in this estimate. It is more the percentage of speakers in Semnan (city) that could make a considerable difference. The assumption that approximately 10% of the inhabitants of Semnan City speak Semnani is based on the estimate of two independent knowledgeable informants.

**Table 5: Estimate of Numbers of Speakers**

Name	Estimate of Informants	Including	Estimate, based on the National Census 2006	FINAL
Semnani of Semnan	10 000 – 15 000 <sup>3</sup>	City, SA, DA	10% of city (124 999) + 80% of	15 000

1-In the estimate we include nomads but exclude speakers who live in other parts of the country. Connell (1967: 83) writes that people from the mountainous regions strive to move not only to Semnan but to bigger and more prestigious cities like Tehran. Therefore the actual number of speakers would be higher, if those who emigrated would be taken into consideration.

2-According to Gharib (1981: 97) rural migration in the township of Semnan is especially high because the terrain in the desert to the South and the mountains to the North is nearly unusable for agriculture and in addition to that the climate is extremely dry and harsh (Gharib 1981: 101).

3-Knowledgeable Informants; 13<sup>th</sup> and 18<sup>th</sup> of March 2008.

(city)			remaining Bachsh <sup>1</sup> Markazi (5 060 <sup>2</sup> ) + 400 in DA = 12 500 + 4 048 = 16 948	
Alai (part of Semnani of Semnan?)	1 800 <sup>3</sup>	AL, YA, HDA 1, XA, RA, HA, DL	80% of 5 813 = 4 650	4 500
Biabanaki	100 <sup>4</sup> - 2 000	BI	80% of 352 = 282	300
Lasjerdi	2 000	AS, HDA 2 DV, Ich	80% of 1 251 = 1 001	1 000
Aftari	1 800 – 2 000 <sup>5</sup>	AF, MM, AV, EZ	80% of 2 897 = 2 318	2 200
Sorkhei	7 000 - 10 000 <sup>6</sup>	SO, NE	70% of 9 089 = 6 362	7 000
<b>Semnani, Total</b>				<b>30 000</b>
Sangesari of	7 000 –	SS	70% of 15 581 <sup>8</sup> =	10 000

1-A bakhsh is an administrative political sub-division of a township.

2-Bakhsh Markazi (135 872) – Semnan (city) (124 999) – Alai (5 813) = 5 060.

3-Member of Village Council of Ala; 27<sup>th</sup> of March 2008.

4-The informant said that about 3 000 people moved from Biabanak to Tehran, 500 to Esfahan and Shiraz and a few went abroad for educational purposes. 16<sup>th</sup> of March 2008.

5-Member of the village council of Aftar; 22<sup>nd</sup> of March 2008.

6-The knowledgeable informant told us that approximately another 10 000 people moved to cities like Tehran and Semnan. 17<sup>th</sup> of March 2008.

7-The informant probably exaggerated. He also said that about 70 000 people from Mehdi Shahr live in other places in Iran. 27<sup>th</sup> of March 2008.

8-SS (20 581) – TA (est. 5 000) = 15 581.

Mehdi Shahr	30 000 <sup>7</sup>		10 907	
Taleb Abadi (Sangesari)	1 000 – 6 000 <sup>1</sup>	TA	TA not found in the census, probably with SS; 70% of est. 5 000 = 3 500	3 500
Darjazini (Sangesari)	1 500 <sup>2</sup> - 2 600	DA, CT, PO	80% of 3 970 <sup>3</sup> = 3 176	3 000
<b>Sangesari, Total</b>				<b>16 500</b>
<b>Shahmirzadi</b>	4 000 - 15 000; summer: 40 000 - 45 000 <sup>4</sup>	SH, CS, DS, KK, HK, ATD	70% of 8 207 = 5 745	<b>6 000 in the Township of Semnan</b>

A feasible conclusion could be that Semnani has approximately 30 000 speakers, Sangesari approximately 15 000 and Shahmirzadi (in the township of Semnan) approximately 6 000. The latter belongs according to the attitude of speakers, previous research and lexico- as well as phonostatic similarity to Mazandarani. Its position within the context of Mazandarani speech varieties remains to be determined. It might be part of a larger group there.

### 3.conclusion

1-Informants, 24<sup>th</sup> of March 2008.

2-Informants estimated that there live about 1 500 Sangesari speakers and about 400 Semnani speakers in Darjazin.

3-DA (4 370) – 400 Semnani speakers = 3 970.

4-A representative of the Ministry for Education told us on the 26<sup>th</sup> of March 2008 that there are about 10 000 – 15 000 people living in Shahmirzad and that during the summer approximately another 30 000 Shahmirzadi speakers come.

### 3.1. Result

Existing literature (Chapter 2.2.), knowledgeable informants (Appendix A) and the perception of local speakers (Ch. 2.3.1.) concur that Sangesari is independent from Semnani.

Taking into account the results of lexico- and phonostatic comparison (Ch. 2.3.2.) the most genuine conclusion would be that all speech forms in the township of Semnan originate from Aftari<sup>1</sup> (or a proto-language similar to Aftari) and are closely related to Mazandarani. However, Aftar is a remote village in the mountains, with less than 1 000 inhabitants and might not even be known to people in all regions of the township of Semnan.

Taking into account people's perception (Ch. 2.3.1.), the number of speakers (Ch. 2.3.3.), and the existing infrastructure<sup>2</sup>, it becomes obvious that Semnan (city) and Mehdi Shahr have to play major roles.

Considering the fact that the lexico- and phonostatic similarity of Semnani to Sangesari is quite low (Ch. 2.3.2.) we decided that it is most natural to propose that there are two languages (Semnani and Sangesari) in the region besides Shahmirzadi, which according to literature (Ch.2.2.), knowledgeable informants (Appendix A), people's perception (Ch. 2.3.1.) and lexico- as well as phonostatic comparison, (Ch. 2.3.2.) is part of the Mazandarani group.

Our final decision therefore is that besides **Shahmirzadi** (ca. 6 000 speakers in the township of Semnan<sup>3</sup>) there exists the **Semnani continuum** (ca. 30 000) and the language of **Sangesari** (ca. 16 500) in the township of Semnan. The

---

1-The lexicostatic similarity of Aftari to all other reference points except Sangesari and Semnani of Darjazin is 80% or above. Even this latter distinction disappears in the phonostatic analysis.

2-Based on his investigation of infrastructure, existing markets, central location and number of inhabitants in the township of Semnan Gharib (1981: 169) chose the city of Semnan as the main centre of the region and Sorkhe, Mehdi Shahr and Fazlabad (to be developed in future) as semi-centres.

3-These number has to be redefined in the context of Mazandarani.

Semnani continuum is ordered in a circle around Sorkhe<sup>1</sup> (ca. 7 000) - which itself does not show any particular affiliation with any of the speech forms in the continuum, but is related to all of them – as follows: Semnani of Semnan (city)<sup>2</sup> (ca. 15 000), Alai<sup>3</sup> (ca. 4 500) (which could as well be regarded to be an accent<sup>4</sup> of Semnani of Semnan), Biabanaki (ca. 300), Lasjerdi<sup>5</sup> (ca. 1 000) and Aftari<sup>6</sup> (ca. 2 200). Sangesari has the following dialects or accents: Sangesari of Mehdi Shahr (ca. 15 000), Taleb Abadi (ca. 3 500) and Darjazini<sup>7</sup> (ca. 3 000). Although an important road is going from Semnan to Mazandaran through Mehdi Shahr and although Semnani and Sangesari meet in Darjazin they show surprisingly low similarity. Aftar and Darjazin truly are two remarkable villages, in a way opposites to each other. The speech variety of Aftar is either a mixture or (close to) a common ancestor of all the speech forms in this township; while Darjazin, not far away, also situated in the mountains, is a place where two distinct languages meet and still maintain their own characteristics.

Another way to view the linguistic situation is to divide the region into two axes. One from the North to the South, coming from Mazandaran and one from the East to the West, following the railway to Tehran. The first axis shows great linguistic diversity, which is not surprising, as it goes through the mountains. The second one, much more homogeneous, is spread out in the plain between the mountain chain and the desert.

---

1-Sorkhei is probably also spoken in Nezami.

2-Possible accents: Semnani of Darjazin, Sufi Abadi, Mahallati, Chup Masjedi, Bazaar, Balae Bazaar, Esfanjane, Sallase, Si Sar and Shashu.

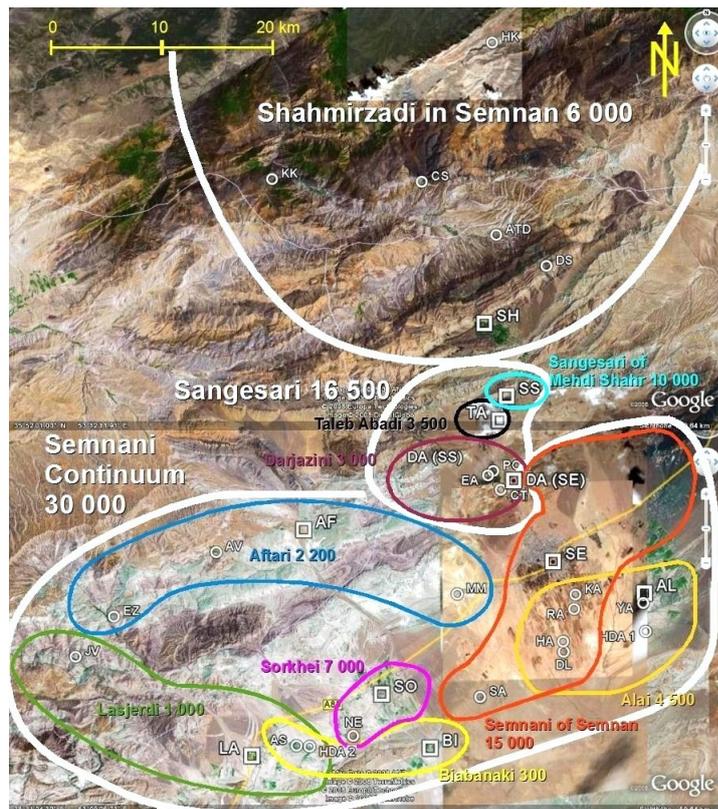
3-Probably also spoken in: Yousef Abad (Kendu), Haji Abad (near Ala), Kheyr Abad, Rokn Abad, Hasan Abad and Delazian.

4-We do not have adequate data to distinguish between dialects and accents.

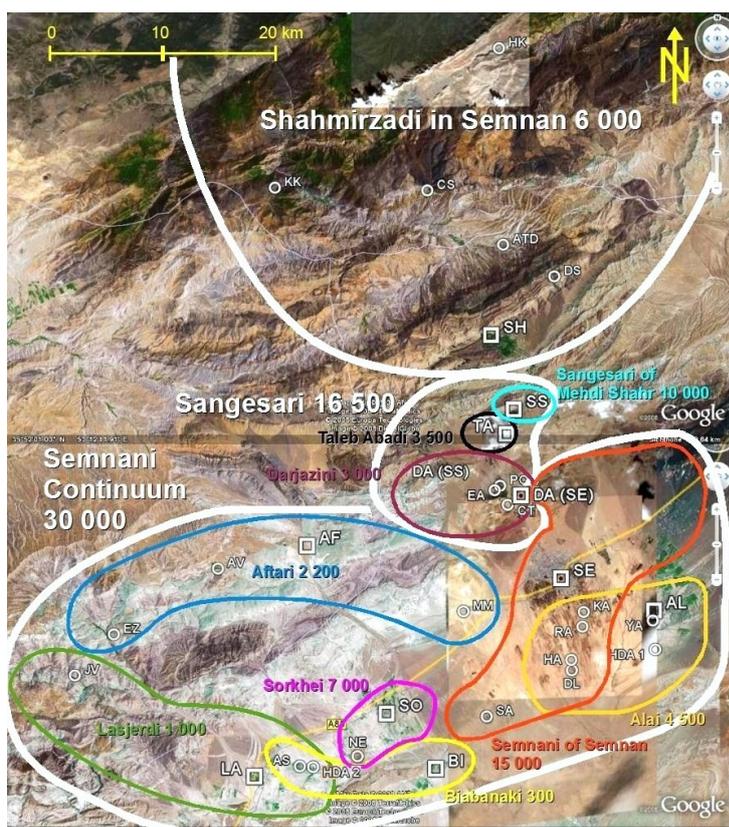
5-Probably also spoken in: Asad Abad, Haji Abad (near Lasjerd), Joveyn and Ich.

6-Probably also spoken in: Mumen Abad, Arvane and Emamzade Abdullah.

7-Probably also spoken in: Chehel Tan.



**Map 5: Languages and Dialects/Accents in the Township of Semnan**



### 3.2. Reasons for Linguistic Diversity

It is most fascinating to see how many different speech forms exist in the small area of the township of Semnan. The mountains might be one main reason, but clearly the question remains, how all the “thousand accents” came there. In the following we present a few hints that we came across in our research, which could be a good starting point for further research aiming at answering this question.

### 3.2.1. The Silk Road

During the second millennium Semnan<sup>1</sup> was a well established centre on the Silk Road, and therefore saw masses of people coming through. They had two possible intentions: trade or war. For the region the traders were more beneficial than the armies: Semnan was frequently destroyed<sup>2</sup> and rebuilt (Connell 1969: 21ff). Curzon already in 1892 noted that Semnan was famous for the “unintelligibility of its speech” (Connell 1969: 23). Since 1957 Semnan has a rail road that connects it with Tehran and Mashhad, two of the most important cities in Iran. Each year masses of people come through Semnan when they go on pilgrimage to Mashhad.

### 3.2.2. Nomads

The township of Semnan is a place where a large number of nomads can still be found. During our stay there we met and interviewed nomads in Aftar and Mehdi Shahr. Connell (1969: 93) describes the migration of nomads living in Taleb Abad. The number of migrating nomads is getting less but still thousands of them move in summer and winter.

### 3.2.3. Economy

Semnan has lots of mineral resources. Many immigrants work in mines and quarries as well as in industry. Places like Aftar have lively trade with Mazandaran<sup>3</sup> and many habitations that we investigated are situated on one of the two roads connecting Semnan to the North.

---

1-Semnan was probably founded during the Sassanid period (224-641 AD).

2-Invasions:

<u>Tenth Century:</u>	Ziyanids of Mazandaran and Dailamites
<u>1036:</u>	Ghozz tribes from Turkey
<u>1221:</u>	Mongol Invasion
<u>End of 14<sup>th</sup> Century:</u>	Timurid troops

3-Informant in Aftar; 22<sup>nd</sup> of March 2008.

## Bibliography

- Anjoman-e Resālat-e Qalam-e Sorkhe (ed.) (1386 AH). *Sarzamin-e āftāb. Negāhi be tārikh, joghrāfiā va mashāhir-e sorxe*, Qom: Zamzam Hedāyat. Azami, Cheragh Ali, and Windfuhr, Gernot Ludwig (1972). *A Dictionary of Sangesari; With a Grammatical Outline*, Tehran: Sherkat-e Sahami-ye Katabhaye Jibi. Callister, Sandra (1998). "Sociolinguistic Approaches To Dialect Surveying In Papua New Guinea", in Loving, R. (ed.) *Language Variation and Survey Techniques*, Ukarumpa: Summer Institute of Linguistics. Christensen, Arthur (1935). *Contributions ? la Dialectologie Iranienne II*, Kobenhavn: Levin & Munksgaard.
- Connell, John (1969). *Semnan: Persian City and Region*, London: University College London Expedition to Iran. Esmā'ili, 'Eṣ mat (1384 AH). *Majmu 'e-ye maqālāt-e nokhostin hamāyesh-e melli-ye pazhuheshi dar guyesh-e Semnāni*, Semnān: Dāneshgāh-e Semnān. Gharib, Freidou (1981). *Zentralorte im Rahmen einer künftigen Regionalstruktur in der Provinz Semnan/Iran*, Stuttgart: Fakult?t für Architektur und Stadtplanung der Universit?t Stuttgart. Google (2008). *Google Earth 4.3.7284.3916* (beta).
- Gordon, Raymond G. (ed.) (2005). *Ethnologue: Languages of the World. Fifteenth edition*, Dallas: SIL International. Hemmati, Esmā'il (1386 AH). *Semnān: Sarzamin-e Guyeshhā*, Semnān: Bustān-e andishe. Homayun, Homādokht (1371 AH). *Guyesh-e Aftari*, Mo'assese-ye momāle'āt o tāqiqāt-e farhangi. Javaheri, Mohammad Hasan, and Shariati, Parviz Pezhom (1386 AH). *Vāzhenāme guyesh-e bāstāni Semnāni*, Semnān: Ābrokh.
- Majidi, Mohammed-Reza (1980). *Strukturelle Beschreibung des iranischen Dialekts der Stadt Semnan. Phonetik, Morphologie, Syntax, Texte*, Hamburg: Helmut Buske Verlag. Maqsudi, Mōammad Yusef (1381 AH). *Negāhi be Shāhmīrzād*, Tehrān: Andishehā-ye Gouharbār. Morgenstierne, G. (1985). "Notes on Semnani", *Norsk Tidsskrift for Sprogvidenskap*, 18: 91-117.
- Morgenstierne, Georg (1960). "Stray notes on Persian dialects", in *Norsk Tidsskrift for Sprogvidenskap*, 19. Oslo: Universitetsforlaget, pp. 73-141, 107f. Panāhi Semnāni, A□mad (1383 AH). *Ādāb va rosum-e mardom-e Semnān*, Tehrān: Pazhuheshgāh-e 'olum-e ensāni va mo□āle'āt-e farhangi.
- Qāsemi Shāhmīrzādi, 'Abdullāh (1379 AH). *Shāhmīrzād-shenāsi*, Tehran: Morshed. Sotude, Manuchehr (1342 AH). *Farhang-e Semnāni-Sorkhei-Lāsgardi-Sangesari-Shahmirzādi*, Tehrān: Chāpkhāne-ye dāneshgāh-e Tehrān. Statistical Centre of Iran (2006). *Population and Housing Census 28. Oct - 17. Nov 2006*. White, Art, and Skoglund, Michael, and McCauley, Nathan, and Roller, Ben (2006). *WordSurv 6*, Taylor University and SIL.
- Windfuhr, Gernot Ludwig (1965). *Verbalmorpheme im Sangesari. Ein Beitrag zur neuiranischen Dialektkunde*, Hamburg: Philosophische Fakult?t der Universit?t Hamburg. Windfuhr, Gernot Ludwig (2009). *The Iranian*

*Languages*, Routledge. (1977). "Dialects of the Semnan Region", in *Norsk Tidsskrift for Sprogvidenskap*, 31. Oslo: Universitetsforlaget.

## Appendix

### A) Information Given by Knowledgeable Informants

- On the 13<sup>th</sup> of March 2008 we interviewed Mr. Mohammad Javaheri, a well known Semnani poet and artist. He suggested the following groupings:
  - 1) Semnani – Biabanaki – Alai – Atari
  - 2) Lasjerdi – Sorkhei
  - 3) Aftari is somewhere in between Lasjerdi, Sorkhei and Semnani
  - 4) Darjazini – Taleb Abadi – Sangesari
  - 5) Shahmirzadi belongs to Mazandarani
- On the 17<sup>th</sup> of March 2008 we talked to the researcher and publisher of the quarterly journal "Surjan Salam", Mr. Mohammad Reza Eslami. He grouped Sorkhei together with Lasjerdi and Biabanaki with Semnani.
- On the 23<sup>rd</sup> of March 2008 informants in Darjazin told us that there are two groups of people living in their village: Sangesari speakers (about 80%) and Semnani speakers (about 20%) and that these two languages are usually not mixed in conversation.
- On the 26<sup>th</sup> of March 2008 a representative of the Ministry of Education in Shahmirzad kindly showed us around in the city and told us that Shahmirzad used to belong to Mazandaran not too long ago, that there is a main road leading from Semnan to Mazandaran and that many people from Shahmirzad moved to Mazandaran to find work. They keep their identity by having their own mosques and tekies<sup>1</sup>. He reckoned that Shahmirzadi was more similar to Mazandarani than to Sangesari. In summer about 30 000 Shahmirzadi people return from

---

1-These are religious places associated with the Shiite saint Emam Husseyn.

Tehran, Mazandaran and Mashhad because of the pleasant climate. Another group of people (presumably from surrounding villages) moved to Shahmirzad for education and work. He thought that the main reasons for the diverse linguistic situation in the township of Semnan were immigrants moving in and the people's close attachment to their language.

- The following dialects of Semnani, spoken in the city itself were mentioned by two informants in Semnan (city): Mahallati, Chup Masjedi, Bazaar, Balae Bazaar, Esfanjane, Sallase, Si Sar, Shashu.
- On the 27<sup>th</sup> of March 2007 the head of the village of Ala said that about 20 000 people emigrated from Ala to Tehran and that there are many Afghani and Baluchi immigrants living in Ala. He regarded Alai to be very similar to Semnani. Sorkhei and Lasjerdi (mentioned together) were very different and Sangesari was possible to understand but not to speak for them.

### **B) Questionnaire**

The following questions of the questionnaire are relevant to this paper:

Q1) Where were you born?

Q2) Where did you grow up?

Q3) Where do you live now?

Q4) Except in this place, where else did you live before, for more than one year? How many years?

Q5) Father comes from:

Q6) Mother comes from:

Q7) In which cities or villages do people speak like you?

Q8) How much do people here understand of the dialect there? (1 = a little, 2 = a lot, 3 = completely)

Q9) How different is your dialect to the dialect there? (1 = most words, 2 = some words, 3 = no difference)

Q10) The language or dialect of which city or village do you understand the most?

Q11) When you go to Mazandaran, how much of the people's language do you understand? (0 = nothing, 1 = a little, 2 = a lot, 3 = completely)

Q12) Is there any village, far away from here, where people speak in your language or dialect?

Q13) Do other villages, where your language or dialect is spoken, have different names for it?

Q14) What do they call the language?

Q15) What do they call themselves?

Q16) Where is your dialect spoken the best?

Q17) Why there?

Q18) Where is it best to live for learning your language or dialect?

### C) Standard Word List

head, hair, eye, nose, ear, mouth, tooth, tongue, neck, throat, arm, hand, finger, nail, stomach (belly), navel, back, leg, knee, foot, skin, bone, blood, heart, man, woman, child, father, mother, brother, sister, uncle (maternal), name, chief (tribal), dog, goat, chicken, bovine (cow-bull), horn (cow), tail, camel, lion, snake, fish, bird, ant, spider, scorpion, louse, tree, branch, leaf, bark, root, flower, seed, grass, sky, cloud, sun, moon, night, star, wind, earth, hill, rock, sand, dust, pebble, water, dew, river, fire, smoke, ember, year, summer, winter, village, field, path, house, bed, hole, rubbish, clothing, saddle, pot, meat, salt, oil, egg, milk, hunger, thirst, rope, iron, knife, war, one, two, three, four, five, six, seven, eight, nine, ten, eleven, twenty, one hundred, much, little, all, good, bad, old, new, hot, cold, tall, short, long, short, heavy, light, full, empty, clean, dirty, dry, red, black, white, leave<sup>1</sup>, come, arrive, get up, sit, lie down, fall, walk, run, swim, fly, see, hear, smell, give birth, die, sleep, blow, whistle, swell, suck, spit, cough, vomit, bark, bite, eat, drink, want, fear, know, think, count, suffer (have pain), laugh, cry, say, ask, sing, dance, play, give, show, send, buy, marry, fight, kill, steal, take, bring, look for, find, push, pull, tie, hit,

---

1-All verb forms were elicited in third person singular simple past.

cut, scrape, press, wash, burn, throw, pour, bury, hide, work, sweep, weave, cultivate, cook, here, near, there, far, to the right, to the left, now, yesterday, tomorrow, where, when, how, who, what, I / me, you / you (thou / thee), he-she / him-her.

#### D) Phonetic Inventory

General (Persian): i, e<sup>1</sup>, e<sup>2</sup>, u, o, v, <sup>u</sup>, <sup>v</sup>; p, b, d, t, s, z, <sup>h</sup>, <sup>ll</sup>, a, k, f, v, x, v<sup>3</sup>, h<sup>x</sup>, l, m, n, ŋ, j.

Specific:

<u>SE</u> : ?, ʁ, ə	<u>DA (Semnani)</u> : ?, ?, ə
<u>SS</u> : ə, ə	<u>BI</u> : y
<u>SH</u> : y, ?, ?	<u>SO</u> : y, ?, ?, ə
<u>AF</u> : ? <sup>''</sup>	<u>AL</u> : y, ?, ?, ə
<u>LA</u> : ?, ə	<u>TA</u> : ?, ʁ, ə, ə
<u>DA (Sangesari)</u> : ə, ə	<u>Mazandarani</u> : ?, ə

---

1-Including [ɛ].

2-Including [a] or [ʔ].

3-Including [q] or [G].