Exploring Heritage Language Variation & Change in Toronto

Naomi Nagy

St. Petersburg, Russia 1916
Toronto, Canada
Faeto, Italy
Lviv, Ukraine 1913
Seoul, Korea 1967
Calabria, Italy 1908
Hong Kong, China 1972
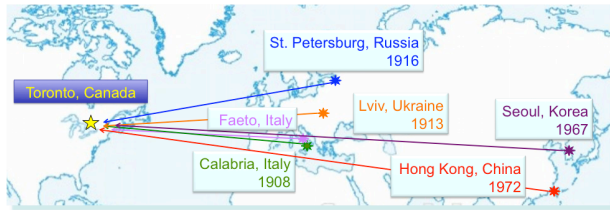
This project is made possible by:

Collaborators
Yoonjung Kang, Univ. of Toronto
Alexei Kochetov, Univ. of Toronto
James Walker, York University

Funding sources
Social Science & Humanities
Research Council of Canada
Work-study Program

Research assistants & Students
Tiina Rebane
Olivia Yu
Ka-man Wong
Karen Chan
Joyce Fok
Rachel Yip
Melania Hrycyna
Taisa Hewka
Ulyana Bila
Vanessa Bertone
Awet Tekeste
Rosanna Calla
Marco Covi
Iryna Kulyk
Yiqing Sim
Anna Shalaginova
Natalia Lapinskaya
Meghan Hollett
Dongkeun Han
Minji Cha
Sheila Chung
Janyce Kim
Shinjung Park

---

2 types of sociolinguistics

Macrolinguistic

study of the languages used and how they are distributed

Ex: 1.5% of Torontonians report two Mother Tongues (= are bilingual as toddlers) (2006 census)

Microlinguistic

study of variable patterns within a language or dialect

Ex: Torontonian youth use "be like" for 58% of their Verbs of Quotation (Tagliamonte & D'Arcy 2004:501)

Microlinguistic   **(TH):**   ← A linguistic variable

% use of stop variant

Reflection of Social Processes in Linguistic Structure

% use of stop variant

Lowest class

Highest class
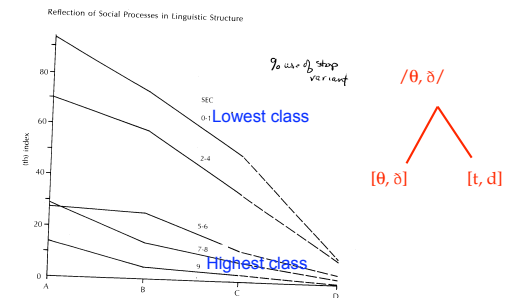
/θ, ð/
[θ, ð]   [t, d]

Fig. 4.1. Class stratification of a linguistic variable with stable social significance: (th) in thing, through, etc. Socioeconomic class scale: 0-1, lower class; 2-4, working class; 5-6, 7-8, lower middle class; 9, upper middle class. A, casual speech; B, careful speech; C, reading style; D, word lists.

Labov (1972: 113)

Languages in Canada

English(es)

"To contemplate a map of Canada while thinking of its languages is to induce an optical illusion."

(Mackey 1998:13)

Languages
French(es)

Mother Tongue, 2006 census

Mother Tongue	Canada	Ontario	Quebec
English	18 million	8 million	600,000
French	7 million	500,000	6 million
Chinese (all types)	1 million	500,000	63,000
Italian	500,000	300,000	100,000
Cree	79,000	4,000	13,000

Mother Tongue – 2006 Census, part 1

	Canada	N.B.	Que.	Ont.
			number	
Total population	31,241,030	719,650	7,435,905	12,028,895
Single responses[1]	30,848,270	714,490	7,339,495	11,853,565
English	17,882,775	463,190	575,555	8,230,705
French	6,817,655	232,975	5,877,660	488,815
Non-official languages	6,147,840	18,320	886,280	3,134,045
Chinese	1,012,065	2,160	63,415	482,570
Cantonese	361,450	295	9,850	181,820
Mandarin	170,950	505	7,770	75,335
Hakka	4,415	10	85	2,805
Chinese, n.o.s.	456,705	1,270	44,740	215,345
Italian	455,040	590	124,820	282,750
German	450,570	1,935	17,855	158,000
Polish	211,175	220	17,305	140,890
Spanish	345,345	1,040	108,790	160,275

11/24/09   http://www40.statcan.ca/l01/cst01/demo11b-eng.htm   7

Mother Tongue – 2006 Census, part 2

	Canada	N.B.	Que.	Ont.
			number	
Total population	31,241,030	719,650	7,435,905	12,028,895
Punjabi	367,505	55	11,905	152,645
Ukrainian	134,500	140	5,395	48,310
Arabic	261,640	970	108,105	114,730
Dutch	128,900	1,290	3,620	68,180
Tagalog (Pilipino)	235,615	330	11,785	117,365
Greek	117,285	275	41,845	61,330
Vietnamese	141,630	205	25,370	67,150
Cree	78,855	0	13,340	3,495
Inuktitut (Eskimo)	32,380	0	9,615	390
Other non-official languages	1,956,060	8,900	288,405	1,120,655
Multiple responses[2]	392,760	5,160	96,405	175,330
English and French	98,625	4,450	43,335	32,690
English and non-official language	240,005	560	16,200	131,290
French and non-official language	43,335	120	31,350	7,790
English, French and non-official language	10,790	30	5,520	3,565

11/24/09   http://www40.statcan.ca/l01/cst01/demo11b-eng.htm   8

% Multilingual studies vs. % multilingual people in the world (North American sample)

Nagy & Meyerhoff 2009

11/24/09   HLVC / SLUGS 2009   9

Comparison of factors in some language-contact studies

Table 1. Community level factors cited in accounting for variation

	IC?	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	N
Geographic																	
Geographic domain	y	x	x		x		x		x	x	x	x			x		9
Geographic overlap	y		x				x			x	x				x		5
Contributing geog.	y									x							1
Political																	
Pol. relation of grps.	y		x		x	x					x	x			x		6
Political domain	y			x			x										3
Other grps. present	y						x								x		2
Demographic																	
Size of subord. grp.	y		x		x				x	x					x		5
Both sexes present?	y								x	x					x		3
Size of dom. group	y			x					x								2
Chronology																	
Period of contact	y			x	x	x	x	x	x	x	x	x	x				11
Period of influence	y												x				1

Nagy 1996

11/24/09   HLVC / SLUGS 2009   10

Comparison of factors in some language-contact studies, cont.

Language use												
Biling. in sub. group	y	x	x		x				x	x		x	6
Other lgs. spoken	y			x		x	x	x			x		5

Do... 4
Ur... 4
Bil... 3
Or... 3
Cu...
Int... 7
Co... 5
Co... 2
Co... 2
Mi...
Ca... 6
Lg... 5
Sta... 3
Lg... 2
Personality | | | | | | x | | | | | | 1
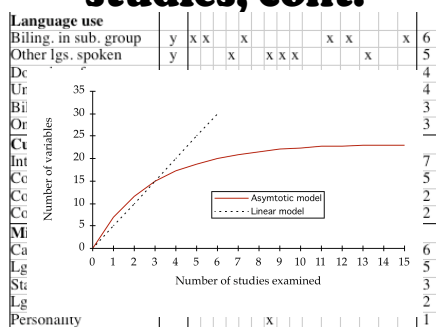
11/24/09   HLVC / SLUGS 2009   11

Communities examined

1. Marathi/Hindi contact in Nagpur, India (Pandharipande 1982:97)
2. Brahui-Balochi contact situation (Indic) (Thomason & Kaufman 1988:70)
3. Uzbek-Tadzhik contact in Soviet Union (*ibid* 70-1)
4. French and Norse influence on English (*ibid* 263-303)
5. Asia-Minor Greek (*ibid* 215-222)
6. Uralic interference in Slavic and Baltic (*ibid* 238-250)
7. Ma'a (*ibid* 223-227)
8. Chinook Jargon (*ibid* 256-262)
9. Michif (*ibid* 228-232)
10. Mednyj Aleut (*ibid* 233-237)
11. Afrikaans (*ibid* 251-255)
12. Norman French & medieval English (van Coetsem 1988:129-135)
13. Afrikaans (*ibid* 129-135)
14. Korlai Portuguese creole in India (Clements 1992:41-52)
15. Basque, Gascon, and French interaction (Haase 1992:343-4)

11/24/09   HLVC / SLUGS 2009   12

Problem

Studies of contact-induced language variation vary widely in methods & contexts.
This prohibits generalizable findings.

Solution

Consistent methods and context, while varying pairs of languages in contact

The language quilt — TORONTO STAR

Italian
Chinese
Cantonese
Punjabi
Portuguese
Spanish
Tagalog
Urdu
Tamil
Polish

Toronto Star 30 Dec. 2007

Long-range questions

LINGUISTIC:

• **Are cross-linguistic generalizations possible about the types of features, structures, rules or constraints that are borrowed earlier and more often ?**

• **If so, what are they?**

Long-range questions

SOCIOLINGUISTIC:

• **How are social factors relevant ?**

• **Do the same (types of) speakers lead changes in both/all their languages ?**

• **Or do speakers choose to use one language or the other for this social "work" ?**

General framework

Contrasting languages
Consistent methodology

Collect
Collaborate
Compare
Quantify
Conclude

Plan

1. **Establish communities of interest** — 2009 √
2. **Interview & record speakers (~ 1 hour each)** — Almost done
3. **Transcribe broadly** — This year
4. **Analyze variables in each language** — 2010-2012
5. **Compare trends across languages** — 2011-2012
6. **Develop generalized framework for linguistic change in a multilingual metropolis** — 2013

Data & organization methods

Corpus design

6 languages
- 3 generations / language
- 3 age groups / generation
- 4 speakers / age group

= **240 speakers**
- Balanced for sex
- Varying in fluency, usage, and ethnic orientation
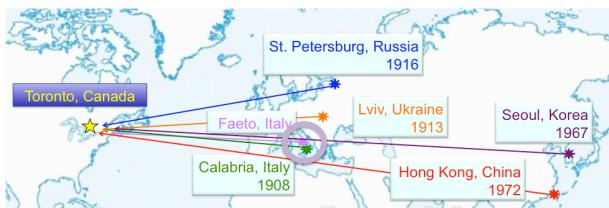
* 3 tasks
- Sociolinguistic interview / Conversation
- Ethnic Orientation Questionnaire
- Picture description task (or Reading Passage)

Languages

St. Petersburg, Russia
1916

Toronto, Canada

Faeto, Italy

Lviv, Ukraine
1913

Seoul, Korea
1967

Calabria, Italy
1908

Hong Kong, China
1972

Years
= founding of first Toronto church for each group, indicative of the establishment of a "community."

Insider interviewers

- Native speakers
- Local
- Outgoing, friendly, careful
- Working in pairs

A big thank you to the interviewers:
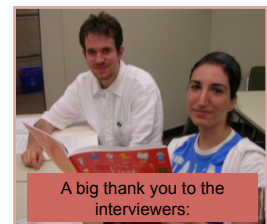
Karen Chan
Joyce Fok
Melania Hrycyna
Taisa Hewka
Awet Tekeste
Rosanna Calla
Marco Covi

Iryna Kulyk
Anna Shalaginova
Sheila Chung
Janyce Kim

Community networks

R1F82A
R1M80B
R1M79A
R2M56A
Holy Trinity Russian Orthodox Church, Henry Street
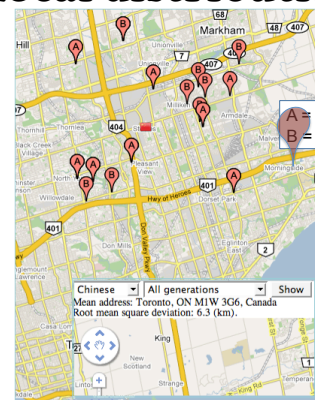R2M50A
R1M34A
R3M56A
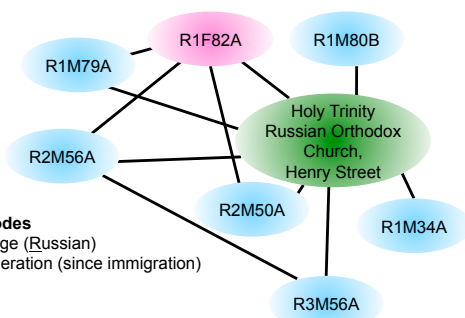
Speaker codes
R = Language (Russian)
1,2,3 = Generation (since immigration)
M, F = Sex
= Age
A,B – just to distinguish multiple speakers in same category

Local distribution

A = 1st generation
B = 2nd generation

Chinese | All generations | Show
Mean address: Toronto, ON M1W 3G6, Canada
Root mean square deviation: 6.3 (km).

Map created by Anna Shalaginova

Data collection methods

Sociolinguistic Interview

- "Guided conversation"
- Designed to elicit relaxed, conversational speech
- Variety of topics to find speaker's interests
- Minimize the effects of a person (stranger) with a tape recorder and microphone asking questions

Why did your family move here?
  Because of work?
  Because of community roots?
  To be close to other Italians? Close to relatives?
Do you know where your family came from?
  When did they come here? Why did they come?
Do you remember hearing stories about how your family came to Toronto? …
Was it hard for them to get set up here?

Adapted from Labov 1984

Ethnic Orientation Questionnaire

- Identità etnica
  – Ti identifichi come Italiano? Canadese? Italo-Canadese?
  – La maggioranza dei tuoi amici sono italiani?
  – La gente nel tuo quartiere e italiana?
  – Quando eri piccolo/a i tuoi compagni di scuola erano italiani? I tuoi amici?...
- Lingua
  – Parli italiano? Parli bene? A Che livello diresti? Parli italiano spesso? Quante volte per giorno/settimana/mese?
  – Dove hai imparato l'italiano? A casa? A scuola?
  Preferisci parlare italiano o in Inglese? …
- Scelta delle lingue
  – Che lingua parla la tua famiglia quando siete tutti insieme ?
  – Che lingua parli con i tuoi amici?
  – Che lingua usi quando parli di cose personali? Quando sei arrabbiato/a?…
- Cultura
- Genitori
- Moglie/Marito/fidanzato/a
  - La cultura italiana
  - Discriminazione

Adapted from Keefe & Padilla 1987, Walker & Hoffman 2008

Ethnic Orientation Questionnaire

A. Ethnic identification
  1. Do you think of yourself as Italian, Canadian or Italian-Canadian?
  2. Are most of your friends Italian?
  3. Are people in your neighbourhood Italian?...
B. Language
  1. Do you speak Italian? How well? How often?
  2. Where did you learn Italian? At home? In school?
  3. Do you prefer to speak Italian or English?
  4. Do you prefer to read and write in Italian or English? …
C. Language choice
  1. What language does your family speak when you get together?
  2. What language do you speak with your friends?
D. Cultural heritage
E. Parents
F. Partner
G. Italian culture
H. Discrimination
  I. Italian culture
  J. Discrimination

Adapted from Keefe & Padilla 1987, Walker & Hoffman 2008

EOQ Results (Korean)

SPEAKER	ETHNIC IDENTIFICATION QUESTIONS				
	A1	A2	A3	A4	TOTAL
K1F39A	2		1		3
K1F48A	2	1	0		3
K1F73A	1				1
K1M45A	1	2	0	2 (Inf.)	4
K1M63A	2	2	0	2	6
K2F21B	1	2	1	1	3
K2F22A	2	1			3
K2F22B	2	1	1	0	4
K2M22A	0	1	1	0	2

0=Canadian
1=Korean-Canadian
2=Korean

Picture description task

Amery & Cartwright 1987
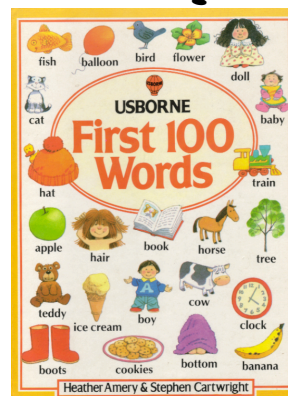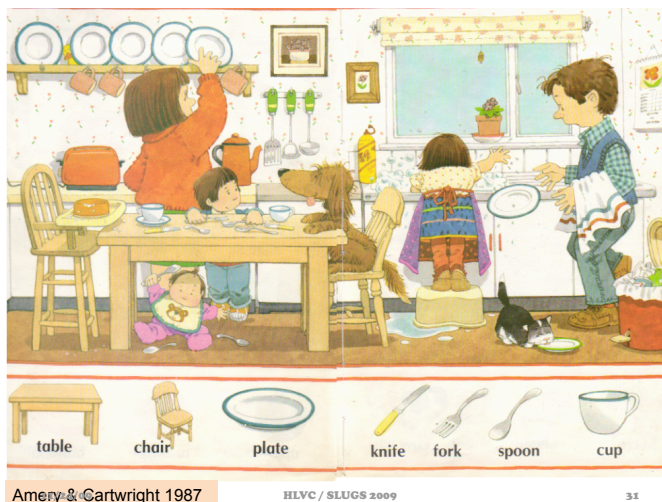
Pronunciation variation

76 year old male	14 year old female	
lu tawuǐin	lu tawuǐinə	
la sɛdʒ	la sédʒə	
lu kartəlléh	lu kartəlléh	
lu kutte'ɣ	lu kuttéjə	
la brókk(əl)	la brókkələ	
la kəʎí'ɣ	la kəʎíjə	
la tátz	la tátzə	

Data analysis methods

Time-aligned orthographic transcription

www.lat-mpi.eu/tools/elan
Initial transcription is broad
Additional tiers added for details.

Candidate cross-linguistic variables

* **Phonetic**
  * Voice Onset Time   Korean – Shinjung Park
* **Phonological**
  * Word-final C deletion & devoicing
* **Morphological**
  * Case and gender marking
  * Classifiers
  * Variable subject presence (Pro-drop)   Russian-Meghan Hollett
* **Syntactic**
  * Word order
* **Lexical**
  * Borrowing   Korean-Sheila Chung, Faetar-me
  * Use of (home country) archaisms
* Other ideas ?

Sample Variable

Null Subject Presence / Absence

I1M75A
Ø Avevo 14 anni e mia moglie ce ne aveva 13.
Ø (I) ...

Io ho ...
I said ...

R3F25A
Я очень медленно читаю, …
I read very slowly

Ø Начала
Ø (I) gradua...
this.

K2F22B
할머니가 (내 이름을) 지으신것 같에요.
My grandmother named (me).

Ø 동생이랑 오빠 이름이랑 다 지으셨어요.
Ø (she) also named my younger and old brothers.

Null-subject variation in Faetar

Faetar Ø-pronouns according to age and sex
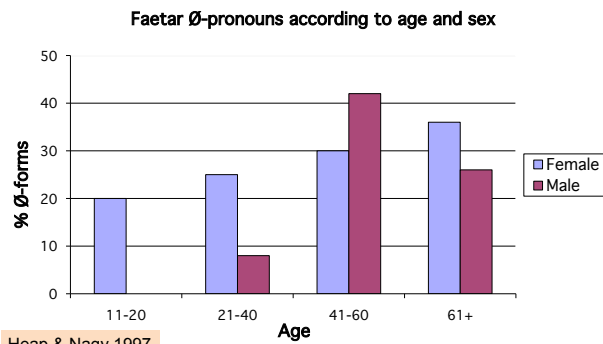
Heap & Nagy 1997

Sample findings:

Lexical change in

Faeto   vs.   Toronto

http://assets.espn.go.com/photo/
2008/1115/travel_g_toronto_580.jpg

Map of Italy and France

France

Francoprovençal region

Italy   Faeto

Roma

Napoli   Bari

Non-Italian dialects
- Albanian
- Provençal
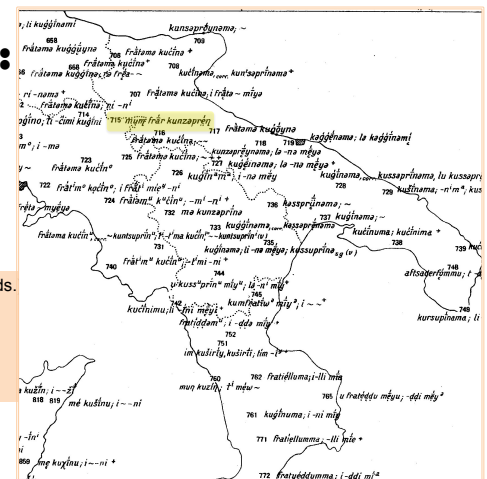- Catalan
- Greek
- German
- Serbo-Croation
- Ladin
- Sardinian
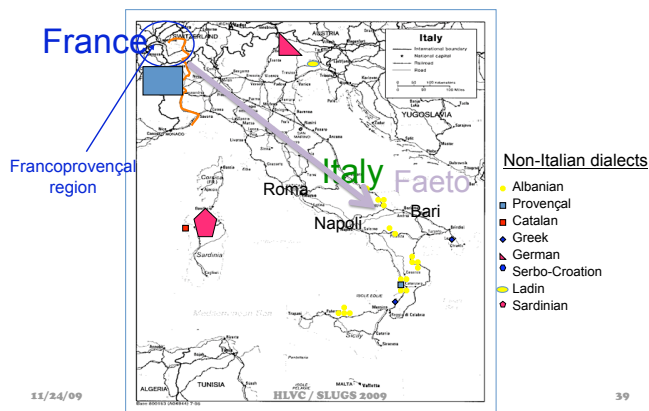
Faeto's location: #715

Jaberg, K. & J. Jud. (eds.)
1928-40.
Sprach- und Sachatlas Italiens und der südschweiz. Zofingen:
Ringier.

Saluti a Maria

Cecile
[dʒi me səsíl tu tə kwannáj ʎo ja sun túttə parɛnn]
I am Cecile. You know. Yes, yeah. We are all family.

Peppine
[tok und a diʃ a mari undə diʃ i at maná ənə bɛllə fiʎa a faít]
What to say to Maria? We should say she sent a nice girl to Faeto.

Concettine
[kəmár kuntʃətín kumə t a dir meɣ]
[t a dir sta bunn]   [və sta bunn]

"Godmother Concettine. What more should I say?"
Say, "Be well."   "You are well."

---

table   chair   plate   knife   fork   spoon   cup

Amery & Cartwright 1987

Invariably Italian words
(N = 13)

English	Faetar	Italian
cup	tatzə	tazza
sheep	pɛgɔrə	pecora
chicken	dʒəlinnə	gallina
train	trɛn	treno
dice	daddə	dado
slipper	pantuoflə,	pantofola
	ʃkarpə	scarpe
dress	vɛst	veste
jacket	dʒakkɛttə	giacchetta
house	kas, kaz	casa
(car	magənə	macchina)
(bicycle	bitʃiklɛttə	bicicletta)
(truck	kamiə	camione)

Invariably Francoprov. words
(N = 4)

English	Faetar	Francoprov.	Italian
chair	sɛdʒ	sɛʒ	sedia
spoon	kuʎijə	kuʎi	cucchiaio
cow	vatʃə	vats, vaʃ	vacca
pig	kaijunnə	?	maiale

Duraffour 1969

Italian and FP forms used
(N = 9)

English	Faetar-F	Francoprovençal	Faetar-I	Italian
table	bufɛttə	bufe(t)	tawolə	tavola
	tab	tabla	tawolinə	tavolina
plate	kartolɛh		piat	piatto
knife	kuttej	kute	kutɛl	cutello
fork	brokələ	bigorə ?	fɔrket	forchetta
Grandma	mowə		nonnə	nonna
Grandpa	biaran		nonnə	nonno
hat	tʃappej	ʃape	kapəlinə	cappello
airplane	parekiə		arioplan	aereoplano
donkey /	tʃuwa	ʃuva	tʃutʃ	ciucio
horse	dʒument	ʒumà	dʒument	giumenta

Duraffour 1969

Language Indices

- **Francoprovençal index**
  - the ratio of Francoprovençal-based words to the number of responses given
  - Tall blue bars = Faetar maintenance
- **Italian index**
  - the ratio of Italian-based words to the number of responses given
  - Taller red bars = language loss
- Indices are the combined values for the 9 words for which the lexical choices varied.

Nagy 1994

Overall results (Faeto)

60% {
- 48% Francoprovençal
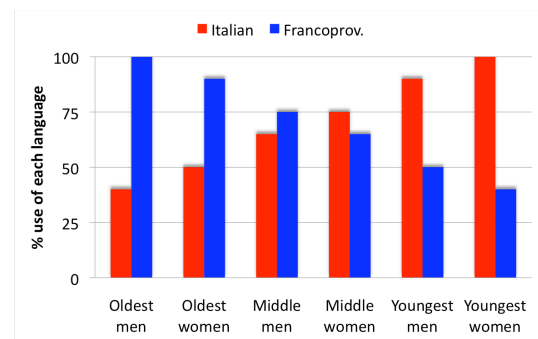- 12% Speaker produced both forms, or an ambiguous form

Nagy 1994

- 40% Italian

Not bad for an Endangered Language.
But how are these forms distributed in the population?

Predicted pattern

Nagy 1994

Generational effect on a phonological variable (final deletion)

(more like French/FP)

(more like Italian)

% without final vowel

100, 90, 80, 70, 60, 50, 40, 30, 20, 10

- Male
- Female

Age: 61-80, 41-60, 21-40, 11-20

Nagy & Reynolds 1997

Actual (lack of) pattern

- Italian
- Francoprov.

% use of each language

100%, 75%, 50%, 25%, 0%

81+ M, 81+ F, 61-80 M, 61-80 F, 41-60 M, 41-60 F, 21-40 M, 21-40 F, 11-20 M, 11-20 F

Nagy 1994

Faeto vs. Toronto

Retention of Francoprovençal words

% use of FP words

100%, 75%, 50%, 25%, 0%

N=31   N=8

Faeto-older, Faeto - younger, Toronto-older *1st generation*, Toronto-younger *2nd generation*

No significant difference by sex, age, or location (Fischer's Exact Test, 2-tailed)

Faeto vs. Toronto

Retention of Francoprovençal words

% use of FP words

100%, 75%, 50%, 25%, 0%

N=31   N=8

Faeto-older, Faeto - younger, Toronto-older, Toronto-younger

- Francoprovençal
- English

Differentiation by social class

N=31

Effect of social class on word choice

- Italian
- Francoprovençal

0.6, 0.4, 0.2, 0

Upper, Middle, Lower

Francoprovençal, Italian
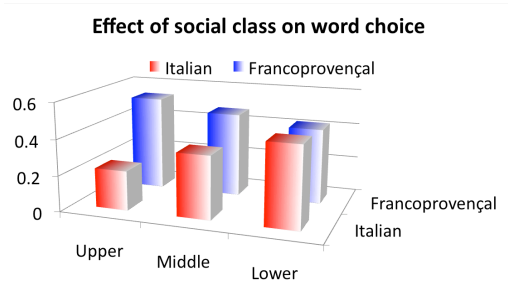
It's a big project …

… and you're invited to collaborate

❑ What could you do?
  - ❑ Recruit speakers
  - ❑ Interview
  - ❑ Transcribe
  - ❑ Analyze data

❑ How could you do it?
  - ❑ Course work (LIN 201, HUM 199, LIN 451…)
  - ❑ Work-study
  - ❑ Research-assistantship
  - ❑ Independent Study
  - ❑ Informal research (e.g., summer project)

Naomi.Nagy@utoronto.ca

What speakers do we still need?

- **Have a look at**
  recordkeeping data/
  speaker_dist_all_lgs.doc

References

Duraffour, A. 1969. *Glossaire des patois francoprovençaux*. Paris: CNRS.

Heap, D. & N. Nagy. 1997. Subject Pronoun Variation in Faetar and in Franco-Provençal. NWAVE 26.

Jaberg, K. & J. Jud. (eds.) 1928-40. *Sprach- und Sachatlas Italiens und der südschweiz*, Zofingen: Ringier.

Keefe, S. & A. Padilla. 1987. *Chicano Ethnicity*. Albuquerque, NM: University of New Mexico Press.

Labov, W. 1984. Field methods of the Project on Linguistic Change and Variation. In J. Baugh & J. Sherzer (eds). *Language in Use*. Englewood Cliffs: Prentice Hall. 28-54.

Nagy, N. 2009. Heritage Language Variation and Change in Toronto.
  http://individual.utoronto.ca/ngn/heritage_lgs.htm.

Nagy, N. 2000. *Faetar*. Munich. Lincom Europa.

Nagy, N. 1994. Lexical change and language contact. *Penn Review of Linguistics* 18: 117-132. Philadelphia: University of Pennsylvania Department of Linguistics.

Nagy, N. & M. Meyerhoff. 2008. The social life of sociolinguistics. In M. Meyerhoff & N. Nagy (eds). *Social Lives in Language: Sociolinguistics and Multilingual Speech Communities*. Amsterdam: John Benjamins.

Nagy, N. & B. Reynolds. 1997. Optimality theory and variable word-final deletion in Faetar. *Language Variation and Change* 9.1:37-56.

Nagy, N., J. Walker, A. Kochetov and Y. Kang. 2009. Heritage language change and variation in Toronto. NWAV, Ottawa.

Toronto Star. The language quilt. Dec. 30, 2007.

Walker, J. & M. Hoffman. 2008. Language Contact, Linguistic Variation and Ethnic Identity in Toronto English 2008. http://www.yorku.ca/jamesw/projects.htm.

Abstract

Given that over half of the world's population is multilingual from childhood (Tucker 1999), it's strange that in quantitative variationist studies the trend is decidedly to examine one language at a time, essentially treating speakers as monolingual (Nagy & Meyerhoff 2008)—although we recognize that significant exceptions exist (e.g. Poplack 1980, Poplack & Meechan 1998). Even in Toronto, touted as the "most multilingual city in the world," two major projects examining ethnic effects on language focus exclusively on English (Tagliamonte 2007, Walker & Hoffman 2008). Complete understanding of how linguistic variation is used to construct identity requires examining multilingual speakers' full repertoires, building on what we have learned from studying speakers' monolingual facets and patterns of code-mixing. To remedy this, we have initiated the **Heritage Language Variation and Change Project** which acomplements English-focused corpus-development projects in Toronto by examining variation and inter-generational change in 7 (of the ~100) heritage languages spoken in the city: Cantonese, Faetar (an endangered Apulian Francoprovençal variety), Italian, Korean, Russian, Ukrainian, and Urdu.

This project addresses questions such as: Which features, structures, rules or constraints are cross-linguistically relevant to borrowing? Which are borrowed earlier and more often? Which social/demographic factors are cross-linguistically relevant to borrowing? Do the same (types of) speakers lead changes in HLs and in English? Is leadership in language change inherent, or do leaders choose to use one language for this social "work"?

The purpose of this paper is to describe our goals and the methods involved in constructing a large multilingual corpus for the purpose of understanding contact-induced language change. Specifically, we illustrate how our methodology ensures comparability and continuity across communities, languages, and fieldworkers in the data-collection stage and uses ELAN (www.lat-mpi.neu/tools/elan), a tool for creating and manipulating time-aligned tiered annotations that eliminates the need for narrow transcription. Our data collection strategy includes hiring fieldworkers from within each community, recording a Labovian-style (Labov 1984) sociolinguistic interview of 60-90 minutes during which participants describe their background, language experiences, demographic characteristics, and linguistic attitudes, and recount personal narratives; complete an Ethnic Orientation Questionnaire (Keffe & Padilla 1987); and complete a picture-identification task to ensure inter-speaker comparability for a set of lexical items.

We will illustrate these methods via description of some cross-generational patterns of change that have been noted in the Italian and Faetar components of our corpus. In our initial fieldwork season (summer 2009) we are conducting 40 interviews in each HL, including first-, second- and third-generation speakers, balanced for a variety of demographic factors. These speakers' variable patterns are compared to published descriptions of the relevant homeland varieties (e.g. Loporcarlo 1997; Nagy 2000; Trumper 1997) to illustrate geographic variation.

An additional goal of this presentation is to encourage further collaboration with scholars interested in using these data.