Exploring Heritage Language Variation & Change in Toronto

Naomi Nagy

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)

Languages in Canada

English(es)
French(es)

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

(Mackey 1998:13)

Languages in Canada

Aboriginal Languages
French(es)
English(es)
Heritage Languages

Mother Tongue, 2006 census

<table>
<thead>
<tr>
<th>Mother Tongue</th>
<th>Canada</th>
<th>Ontario</th>
<th>Quebec</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>18 million</td>
<td>8 million</td>
<td>600,000</td>
</tr>
<tr>
<td>French</td>
<td>7 million</td>
<td>500,000</td>
<td>6 million</td>
</tr>
<tr>
<td>Chinese (all types)</td>
<td>1 million</td>
<td>500,000</td>
<td>63,000</td>
</tr>
<tr>
<td>Italian</td>
<td>500,000</td>
<td>300,000</td>
<td>100,000</td>
</tr>
<tr>
<td>Cree</td>
<td>79,000</td>
<td>4,000</td>
<td>13,000</td>
</tr>
</tbody>
</table>
% Multilingual studies vs.
% multilingual people in the world
(North American sample)

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

% Multilingual studies
% Multilingual people

Comparison of factors in some
language-contact studies

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 Balic (ibid 238-250)
7. Ma'a (ibid 223-227)
8. Chinook Jargon (ibid 256-262)
9. Michif (ibid 228-232)
10. Mednuy Alen (ibid 233-237)
11. Afrikaans (ibid 251-255)
13. Afrikaans (ibid 129-135)
15. Basque, Gascon, and French interaction (Haase 1992:343-4)
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

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?

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
**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**

- Corpus design
  - 6 languages
  - 3 generations / language
  - 3 age groups / generation
  - 4 speakers / age group
  - Balanced for sex
  - Varying in fluency, usage, and ethnic orientation
  - 3 tasks
    - Sociolinguistic interview / Conversation
    - Ethnic Orientation Questionnaire
    - Picture description task (or Reading Passage)

---

**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**

- Holy Trinity Russian Orthodox Church, Henry Street

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**

Map created by Anna Shalaginova

**Languages**

- Corpus design
  - 6 languages
  - 3 generations / language
  - 3 age groups / generation
  - 4 speakers / age group
  - Balanced for sex
  - Varying in fluency, usage, and ethnic orientation
  - 3 tasks
    - Sociolinguistic interview / Conversation
    - Ethnic Orientation Questionnaire
    - Picture description task (or Reading Passage)

---

**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**

- Holy Trinity Russian Orthodox Church, Henry Street

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**

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

**Ethnic Orientation Questionnaire**

- **Identità etnica**
  - Ti identifichi come Italiano? Canadese? Italo-Canadese?
  - La maggioranza dei tuoi amici sono italiani?
  - La gente nel tuo quartiere è italiana?
  - Quando eri piccolo/i i tuoi compagni di scuola erano italiani? I tuoi amici? ...
- **Lingua**
  - Parli italiano? Parli bene? A che livello diresti? 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?

**EOQ Results (Korean)**

<table>
<thead>
<tr>
<th>SPEAKER</th>
<th>ETHNIC IDENTIFICATION QUESTIONS</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>K1F39A</td>
<td>A1 2 A2 1 A3 1 A4 TOTAL 3</td>
<td></td>
</tr>
<tr>
<td>K1F48A</td>
<td>2 1 0 3</td>
<td></td>
</tr>
<tr>
<td>K1F73A</td>
<td>1 0 1</td>
<td></td>
</tr>
<tr>
<td>K1M45A</td>
<td>1 2 0 2 (Inf.) 4</td>
<td></td>
</tr>
<tr>
<td>K1M63A</td>
<td>2 2 0 2 6</td>
<td></td>
</tr>
<tr>
<td>K2F21B</td>
<td>1 2 1 1 3</td>
<td></td>
</tr>
<tr>
<td>K2F22A</td>
<td>2 1 1 3</td>
<td></td>
</tr>
<tr>
<td>K2F22B</td>
<td>2 1 1 0 4</td>
<td></td>
</tr>
<tr>
<td>K2M22A</td>
<td>0 1 1 0 2</td>
<td></td>
</tr>
</tbody>
</table>

**Picture description task**

Adapted from Keefe & Padilla 1987, Walker & Hoffman 2008
**Pronunciation variation**

<table>
<thead>
<tr>
<th>76 year old male</th>
<th>14 year old female</th>
</tr>
</thead>
<tbody>
<tr>
<td>lu tawuli'n</td>
<td>lu tawuli'n</td>
</tr>
<tr>
<td>la sed'y</td>
<td>la sed'y</td>
</tr>
<tr>
<td>lu kartallih</td>
<td>lu kartallih</td>
</tr>
<tr>
<td>lu kutte'y</td>
<td>lu kutte'y</td>
</tr>
<tr>
<td>la brokkala</td>
<td>la brokkala</td>
</tr>
<tr>
<td>la ka'di'</td>
<td>la ka'di'</td>
</tr>
<tr>
<td>la tätz</td>
<td>la tätz</td>
</tr>
</tbody>
</table>

**Data analysis methods**

- Phonetic
- Voice Onset Time
- Phonological
- Word-final C deletion & devoicing
- Morphological
- Case and gender marking
- Classifiers
- Variable subject presence (Pro-drop)
- Syntactic
- Word order
- Lexical
- Borrowing
- Use of (home country) archaisms
- Other ideas?

**Candidate cross-linguistic variables**

<table>
<thead>
<tr>
<th>Phonetic</th>
<th>Voice Onset Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word-final C deletion &amp; devoicing</td>
<td>Morphological</td>
</tr>
<tr>
<td>Case and gender marking</td>
<td>Classifiers</td>
</tr>
<tr>
<td>Variable subject presence (Pro-drop)</td>
<td>Syntactic</td>
</tr>
<tr>
<td>Word order</td>
<td>Lexical</td>
</tr>
<tr>
<td>Borrowing</td>
<td>Use of (home country) archaisms</td>
</tr>
<tr>
<td>Other ideas?</td>
<td>Korean – Shinjung Park</td>
</tr>
</tbody>
</table>

**Sample Variable**

**Null Subject Presence / Absence**

Korean – Shinjung Park

<table>
<thead>
<tr>
<th>Korean – Shinjung Park</th>
<th>Russian – Meghan Hollett</th>
</tr>
</thead>
<tbody>
<tr>
<td>I1M75A</td>
<td>Ø Avevo 14 anni e mia moglie ce ne aveva 13.</td>
</tr>
<tr>
<td>Ø (I)</td>
<td>Ø (I) was 14 and my wife was only 13…</td>
</tr>
<tr>
<td>I said</td>
<td>¥ I said, “OK, I’ll do it.”</td>
</tr>
<tr>
<td>Korean-Sheila Chung, Faetar-me</td>
<td>I read very slowly,…</td>
</tr>
<tr>
<td>My grandmother named (me).</td>
<td>Ø (she) also named my younger and old brothers.</td>
</tr>
<tr>
<td>Ø (she) also named my younger and old brothers.</td>
<td>Ø (she) gradually started to do some research on this.</td>
</tr>
</tbody>
</table>
Null-subject variation in Faetar

Faetar Ø-pronouns according to age and sex

Sample findings:

Lexical change in

Faeto vs. Toronto

Map of Italy and France

Faeto’s location:

#715

Saluti a Maria

[di' me sasli tu ta kwana' ajo ja sun tatta parinn]
I am Cecile. You know. Yes, yeah. We are all family.

[lok und a dij a mari undo dij i at man a na bëü bëü a falt]
What to say to Maria? We should say she sent a nice girl to Faeto.

[ka'mar kun'taen kuma t a dir mey]
Say, “Be well.”

“Godmother Concettine. What more should I say?”

Say, “You are well.”
Invariably Italian words
(N = 13)

<table>
<thead>
<tr>
<th>English</th>
<th>Faetar</th>
<th>Italian</th>
</tr>
</thead>
<tbody>
<tr>
<td>cup</td>
<td>tatza</td>
<td>tazza</td>
</tr>
<tr>
<td>sheep</td>
<td>prgrara</td>
<td>pecora</td>
</tr>
<tr>
<td>chicken</td>
<td>dlulinna</td>
<td>gallina</td>
</tr>
<tr>
<td>train</td>
<td>trren</td>
<td>treno</td>
</tr>
<tr>
<td>dice</td>
<td>dadda</td>
<td>dado</td>
</tr>
<tr>
<td>slipper</td>
<td>pantufo</td>
<td>pantofola</td>
</tr>
<tr>
<td>dress</td>
<td>vest</td>
<td>veste</td>
</tr>
<tr>
<td>jacket</td>
<td>dsakketta</td>
<td>giacchetta</td>
</tr>
<tr>
<td>house</td>
<td>kas, kaz</td>
<td>casa</td>
</tr>
<tr>
<td>(car)</td>
<td>magona</td>
<td>macchina</td>
</tr>
<tr>
<td>(bicycle)</td>
<td>bifi</td>
<td>bicicletta</td>
</tr>
<tr>
<td>(truck)</td>
<td>kamia</td>
<td>camione</td>
</tr>
</tbody>
</table>

Invariably Francoprov. words
(N = 4)

<table>
<thead>
<tr>
<th>English</th>
<th>Faetar</th>
<th>Francoprov.</th>
<th>Italian</th>
</tr>
</thead>
<tbody>
<tr>
<td>chair</td>
<td>sedz</td>
<td>sez</td>
<td>sedia</td>
</tr>
<tr>
<td>spoon</td>
<td>kujiø</td>
<td>kuji</td>
<td>cucchio</td>
</tr>
<tr>
<td>cow</td>
<td>vatšø</td>
<td>vats, vaf</td>
<td>vacca</td>
</tr>
<tr>
<td>pig</td>
<td>kaijunnø</td>
<td>?</td>
<td>maiale</td>
</tr>
</tbody>
</table>

Italian and FP forms used
(N = 9)

<table>
<thead>
<tr>
<th>English</th>
<th>Faetar-F</th>
<th>Francoprovençal</th>
<th>Faetar-I</th>
<th>Italian</th>
</tr>
</thead>
<tbody>
<tr>
<td>table</td>
<td>bufetta</td>
<td>bufe(t)</td>
<td>tawola</td>
<td>tawola</td>
</tr>
<tr>
<td>plate</td>
<td>kartolch</td>
<td>tabla</td>
<td>tavolina</td>
<td>tavolina</td>
</tr>
<tr>
<td>knife</td>
<td>kuttej</td>
<td>kute</td>
<td>piat</td>
<td>piatto</td>
</tr>
<tr>
<td>fork</td>
<td>brokala</td>
<td>bigoro ?</td>
<td>fišket</td>
<td>forchetta</td>
</tr>
<tr>
<td>Grandma</td>
<td>mowa</td>
<td>nonna</td>
<td>nonna</td>
<td>nonna</td>
</tr>
<tr>
<td>Grandpa</td>
<td>biaran</td>
<td>jape</td>
<td>kapolina</td>
<td>cappello</td>
</tr>
<tr>
<td>airplane</td>
<td>parekia</td>
<td>jüva</td>
<td>arioplan</td>
<td>aereoplan</td>
</tr>
<tr>
<td>donkey/</td>
<td>jjuwa</td>
<td>juva</td>
<td>tʃutʃ</td>
<td>ciucio</td>
</tr>
<tr>
<td>horse</td>
<td>dʒument</td>
<td>gumà</td>
<td>dʒument</td>
<td>giumenta</td>
</tr>
</tbody>
</table>

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.

Overall results (Faeto)

60% Francoprovençal
48% Speaker produced both forms, or an ambiguous form
12% Italian
40% Italian

Not bad for an Endangered Language.
But how are these forms distributed in the population?
Generational effect on a phonological variable (final deletion) (more like French/FP) (more like Italian)

Nagy & Reynolds 1997

Actual (lack of) pattern

Nagy 1994

Faeto vs. Toronto

Retention of Francoprovençal words

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

Faeto vs. Toronto

Retention of Francoprovençal words

Differentiation by social class

Effect of social class on word choice

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

Abstract

Given that over half of the world’s population is multilingual from childhood (Tucker 1999), it’s strange that in quantitative variational 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, hailed 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 speaker’s 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 accompanies English-focused corpus-development projects in Toronto by examining variation and inter-generational change in 7 (~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 context-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.org/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 elicit personal narratives; complete an Ethnic Orientation Questionnaire (Keefe & 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 completing a picture-identification task to elicit 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. An additional goal of this presentation is to encourage further collaboration with scholars interested in Heritage Language Variation and Change Project.

References

http://individual.utoronto.ca/ngn/heritage_lgs.htm.