Morphology generation for Swiss German dialects

Yves Scherrer

LATL
University of Geneva
Switzerland

SFCM 2011, Zurich

Outline

Outline

Standard German and Swiss German

Standard German:

Warst du auf dem Markt einkaufen?
Were you at the market shopping?
'Have you been shopping at the market?'

Swiss German Dialects:

ZH Bisch uf em Mèrt gsy go poschte?
SZ Bisch ufem Märcht gsi go poschte?
SO Besch ofem Märed gsy go boschte?
OW Bischt ufm Märt gsii go ichoifä?
SG Bisch uf dä Märt go poschtä?
BE Bisch uf e Märit ga kömerle?
FR Büschu z'Määret gsi ga iichufe?
WS Bisch dü uf um Markt gsii ga ichöifu?

http://als.wikipedia.org/wiki/Alemannischer_Beispielsatz

Differences between Standard German and Swiss German:

- Preterite tense replaced by present perfect
- *go/ga* + infinitive construction

Differences within Swiss German dialects:

- No standardized spelling
- Phonetics
- Lexicon
- Morphology
- Syntax

Standard German and Swiss German

Standard German:

Warst du auf dem Markt einkaufen?
Were you at the market shopping?
'Have you been shopping at the market?'

Swiss German Dialects:

ZH Bisch uf em Mèrt gsy go poschte?

SZ Bisch ufem Märcht gsi go poschte?

SO Besch ofem Märed gsy go boschte?

OW Bischt ufm Märt gsii go ichoifä?

SG Bisch uf dä Märt go poschtä?

BE Bisch uf e Märit ga kömerle?

FR Büschu z'Määret gsi ga iichufe?

WS Bisch dü uf um Markt gsii ga ichöifu?

http://als.wikipedia.org/wiki/Alemannischer_Beispielsatz

Differences between Standard German and Swiss German:

- Preterite tense replaced by present perfect
- *go/ga* + infinitive construction

Differences within Swiss German dialects:

- No standardized spelling
- Phonetics
- Lexicon
- Morphology
- Syntax

Standard German and Swiss German

Standard German:

Warst du auf dem Markt einkaufen?
Were you at the market shopping?
'Have you been shopping at the market?'

Swiss German Dialects:

ZH Bisch uf em Mèrt gsy go poschte?

SZ Bisch ufem Märcht gsi go poschte?

SO Besch ofem Märed gsy go boschte?

OW Bischt ufm Märt gsii go ichoifä?

SG Bisch uf dä Märt go poschtä?

BE Bisch uf e Märit ga kömerle?

FR Büschu z'Määret gsi ga iichufe?

WS Bisch dü uf um Markt gsii ga ichöifu?

http://als.wikipedia.org/wiki/Alemannischer_Beispielsatz

Differences between Standard German and Swiss German:

- Preterite tense replaced by present perfect
- *go/ga* + infinitive construction

Differences within Swiss German dialects:

- No standardized spelling
- Phonetics
- Lexicon
- Morphology
- Syntax

Standard German and Swiss German

Standard German:

Warst du auf dem Markt einkaufen?
Were you at the market shopping?
'Have you been shopping at the market?'

Swiss German Dialects:

ZH Bisch uf em Mèrt gsy go poschte?

SZ Bisch ufem Märcht gsi go poschte?

SO Besch ofem Märed gsy go boschte?

OW Bischt ufm Märt gsii go ichoifä?

SG Bisch uf dä Märt go poschtä?

BE Bisch uf e Märit ga kömerle?

FR Büschu z'Määret gsi ga iichufe?

WS Bisch dü uf um Markt gsii ga ichöifu?

http://als.wikipedia.org/wiki/Alemannischer_Beispielsatz

Differences between Standard German and Swiss German:

- Preterite tense replaced by present perfect
- *go/ga* + infinitive construction

Differences within Swiss German dialects:

- No standardized spelling
- Phonetics
- Lexicon
- Morphology
- Syntax

Standard German and Swiss German

Standard German:

Warst du auf dem Markt einkaufen?
Were you at the market shopping?
'Have you been shopping at the market?'

Swiss German Dialects:

ZH	Bisch uf em Mèrt gsy go poschte?
SZ	Bisch ufem Märcht gsi go poschte?
SO	Besch ofem Märed gsy go boschte?
OW	Bischt ufm Märt gsii go ichoifä?
SG	Bisch uf dä Märt go poschtä?
BE	Bisch uf e Märit ga kömerle?
FR	Büschu z'Määret gsi ga iichufe?
WS	Bisch dü uf um Markt gsii ga ichöifu?

http://als.wikipedia.org/wiki/Alemannischer_Beispielsatz

Differences between Standard German and Swiss German:

- Preterite tense replaced by present perfect
- *go/ga* + infinitive construction

Differences within Swiss German dialects:

- No standardized spelling
- Phonetics
- Lexicon
- Morphology
- Syntax

Standard German and Swiss German

Standard German:

Warst du auf dem Markt einkaufen?
Were you at the market shopping?
'Have you been shopping at the market?'

Swiss German Dialects:

ZH Bisch uf em Mèrt gsy go poschte?

SZ Bisch ufem Märcht gsi go poschte?

SO Besch ofem Märed gsy go boschte?

OW Bischt ufm Märt gsii go ichoifä?

SG Bisch uf dä Märt go poschtä?

BE Bisch uf e Märit ga kömerle?

FR Büschu z'Määret gsi ga iichufe?

WS Bisch dü uf um Markt gsii ga ichöifu?

http://als.wikipedia.org/wiki/Alemannischer_Beispielsatz

Differences between Standard German and Swiss German:

- Preterite tense replaced by present perfect
- *go/ga* + infinitive construction

Differences within Swiss German dialects:

- No standardized spelling
- Phonetics
- Lexicon
- Morphology
- Syntax

Standard German and Swiss German

Standard German:

Warst du auf dem Markt einkaufen?
Were you at the market shopping?
'Have you been shopping at the market?'

Swiss German Dialects:

ZH Bisch uf em Mèrt gsy go poschte?

SZ Bisch ufem Märcht gsi go poschte?

SO Besch ofem Märed gsy go boschte?

OW Bischt ufm Märt gsii go ichoifä?

SG Bisch uf dä Märt go poschtä?

BE Bisch uf e Märit ga kömerle?

FR Büschu z'Määret gsi ga iichufe?

WS Bisch dü uf um Markt gsii ga ichöifu?

http://als.wikipedia.org/wiki/Alemannischer_Beispielsatz

Differences between Standard German and Swiss German:

- Preterite tense replaced by present perfect
- *go/ga* + infinitive construction

Differences within Swiss German dialects:

- No standardized spelling
- Phonetics
- Lexicon
- Morphology
- Syntax

Standard German and Swiss German

Standard German:

Warst du auf dem Markt einkaufen?
Were you at the market shopping?
'Have you been shopping at the market?'

Swiss German Dialects:

ZH	Bisch uf em Mèrt gsy go poschte?
SZ	Bisch ufem Märcht gsi go poschte?
SO	Besch ofem Märed gsy go boschte?
OW	Bischt ufm Märt gsii go ichoifä?
SG	Bisch uf dä Märt go poschtä?
BE	Bisch uf e Märit ga kömerle?
FR	Büschu z'Määret gsi ga iichufe?
WS	Bisch dü uf um Markt gsii ga ichöifu?

http://als.wikipedia.org/wiki/Alemannischer_Beispielsatz

Differences between Standard German and Swiss German:

- Preterite tense replaced by present perfect
- *go/ga* + infinitive construction

Differences within Swiss German dialects:

- No standardized spelling
- Phonetics
- Lexicon
- Morphology
- Syntax

Standard German and Swiss German

Standard German:

Warst du auf dem Markt einkaufen?
Were you at the market shopping?
'Have you been shopping at the market?'

Swiss German Dialects:

ZH Bisch uf em Mèrt gsy go poschte?

SZ Bisch ufem Märcht gsi go poschte?

SO Besch ofem Märed gsy go boschte?

OW Bischt ufm Märt gsii go ichoifä?

SG Bisch uf dä Märt go poschtä?

BE Bisch uf e Märit ga kömerle?

FR Büschu z'Määret gsi ga iichufe?

WS Bisch dü uf um Markt gsii ga ichöifu?

http://als.wikipedia.org/wiki/Alemannischer_Beispielsatz

Differences between Standard German and Swiss German:

- Preterite tense replaced by present perfect
- *go/ga* + infinitive construction

Differences within Swiss German dialects:

- No standardized spelling
- Phonetics
- Lexicon
- Morphology
- Syntax

Standard German and Swiss German

Standard German:

Warst du auf dem Markt einkaufen?
Were you at the market shopping?
'Have you been shopping at the market?'

Swiss German Dialects:

ZH	Bisch uf em Mèrt gsy go poschte?
SZ	Bisch ufem Märcht gsi go poschte?
SO	Besch ofem Märed gsy go boschte?
OW	Bischt ufm Märt gsii go ichoifä?
SG	Bisch uf dä Märt go poschtä?
BE	Bisch uf e Märit ga kömerle?
FR	Büschu z'Määret gsi ga iichufe?
WS	Bisch dü uf um Markt gsii ga ichöifu?

http://als.wikipedia.org/wiki/Alemannischer_Beispielsatz

Differences between Standard German and Swiss German:

- Preterite tense replaced by present perfect
- *go/ga* + infinitive construction

Differences within Swiss German dialects:

- No standardized spelling
- Phonetics
- Lexicon
- Morphology
- Syntax

Standard German and Swiss German

Standard German:

Warst du auf dem Markt einkaufen?
Were you at the market shopping?
'Have you been shopping at the market?'

Swiss German Dialects:

ZH Bisch uf em Mèrt gsy go poschte?

SZ Bisch ufem Märcht gsi go poschte?

SO Besch ofem Märed gsy go boschte?

OW Bischt ufm Märt gsii go ichoifä?

SG Bisch uf dä Märt go poschtä?

BE Bisch uf e Märit ga kömerle?

FR Büschu z'Määret gsi ga iichufe?

WS Bisch dü uf um Markt gsii ga ichöifu?

http://als.wikipedia.org/wiki/Alemannischer_Beispielsatz

Differences between Standard German and Swiss German:

- Preterite tense replaced by present perfect
- *go/ga* + infinitive construction

Differences within Swiss German dialects:

- No standardized spelling
- Phonetics
- Lexicon
- Morphology
- Syntax

Standard German and Swiss German

Standard German:

Warst du auf dem Markt einkaufen?
Were you at the market shopping?
'Have you been shopping at the market?'

Swiss German Dialects:

ZH Bisch uf em Mèrt gsy go poschte?

SZ Bisch ufem Märcht gsi go poschte?

SO Besch ofem Märed gsy go boschte?

OW Bischt ufm Märt gsii go ichoifä?

SG Bisch uf dä Märt go poschtä?

BE Bisch uf e Märit ga kömerle?

FR Büschu z'Määret gsi ga iichufe?

WS Bisch dü uf um Markt gsii ga ichöifu?

http://als.wikipedia.org/wiki/Alemannischer_Beispielsatz

Differences between Standard German and Swiss German:

- Preterite tense replaced by present perfect
- *go/ga* + infinitive construction

Differences within Swiss German dialects:

- No standardized spelling
- Phonetics
- Lexicon
- Morphology
- Syntax

Standard German and Swiss German

Standard German:

Warst du auf dem Markt einkaufen?
Were you at the market shopping?
'Have you been shopping at the market?'

Swiss German Dialects:

ZH Bisch uf em Mèrt gsy go poschte?

SZ Bisch ufem Märcht gsi go poschte?

SO Besch ofem Märed gsy go boschte?

OW Bischt ufm Märt gsii go ichoifä?

SG Bisch uf dä Märt go poschtä?

BE Bisch uf e Märit ga kömerle?

FR Büschu z'Määret gsi ga iichufe?

WS Bisch dü uf um Markt gsii ga ichöifu?

http://als.wikipedia.org/wiki/Alemannischer_Beispielsatz

Differences between Standard German and Swiss German:

- Preterite tense replaced by present perfect
- *go/ga* + infinitive construction

Differences within Swiss German dialects:

- No standardized spelling
- Phonetics
- Lexicon
- Morphology
- Syntax

Standard German and Swiss German

Standard German:

Warst du auf dem Markt einkaufen?
Were you at the market shopping?
'Have you been shopping at the market?'

Swiss German Dialects:

ZH Bisch uf em Mèrt gsy go poschte?

SZ Bisch ufem Märcht gsi go poschte?

SO Besch ofem Märed gsy go boschte?

OW Bischt ufm Märt gsii go ichoifä?

SG Bisch uf dä Märt go poschtä?

BE Bisch uf e Märit ga kömerle?

FR Büschu z'Määret gsi ga iichufe?

WS Bisch dü uf um Markt gsii ga ichöifu?

http://als.wikipedia.org/wiki/Alemannischer_Beispielsatz

Differences between Standard German and Swiss German:

- Preterite tense replaced by present perfect
- *go/ga* + infinitive construction

Differences within Swiss German dialects:

- No standardized spelling
- Phonetics
- Lexicon
- Morphology
- Syntax

Goal

Generate Swiss German dialect word forms… → Morphology

…from Standard German data → Phonetics, lexicon

Outline

Swiss German word generation

Standard German root
+ Morphosyntactic features

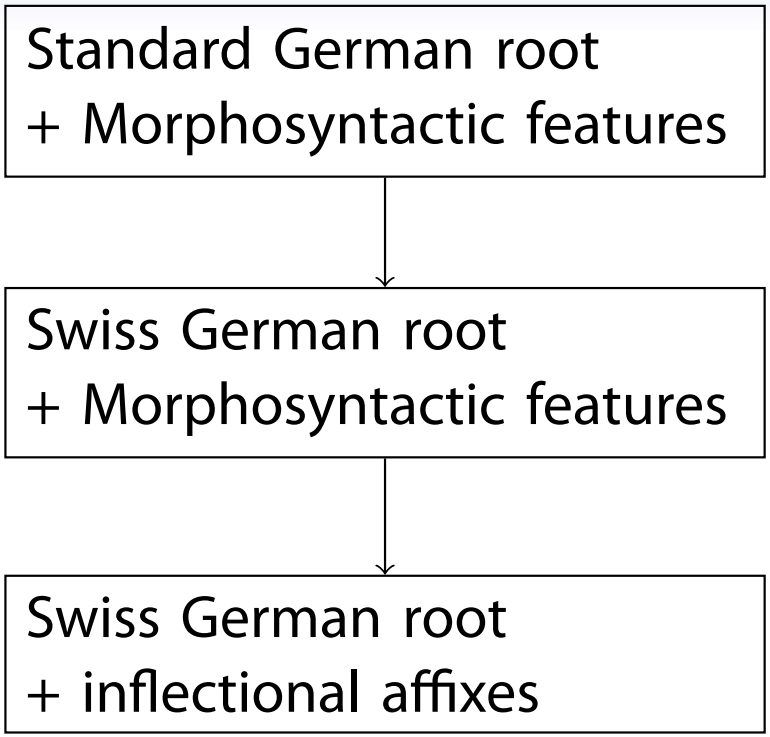
↓

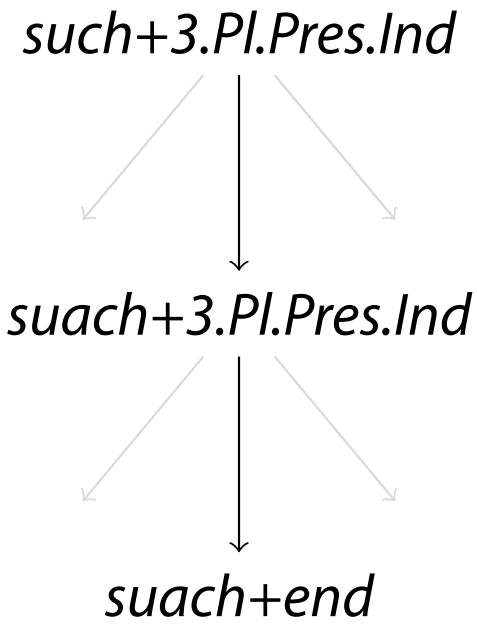
Swiss German root
+ Morphosyntactic features

↓

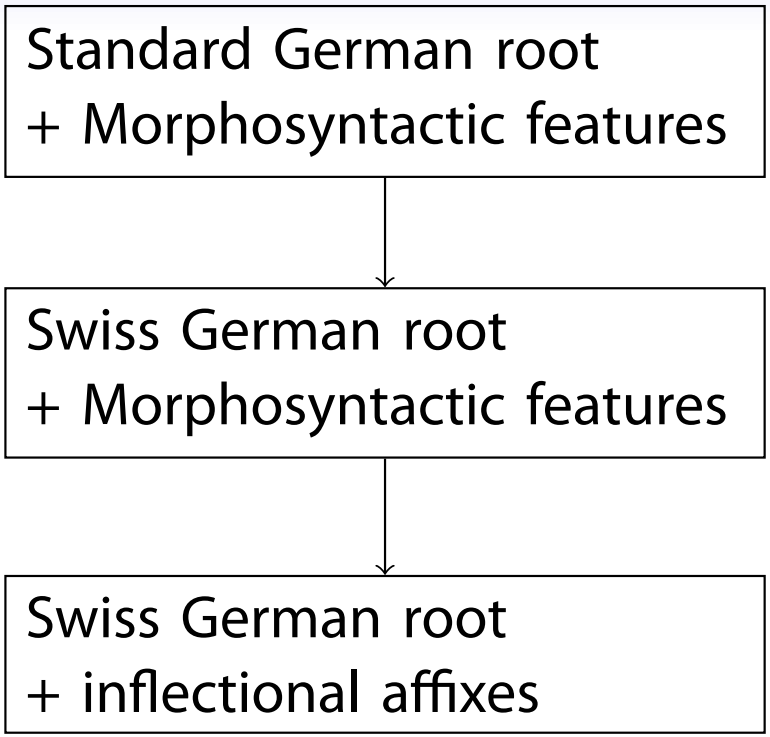
Swiss German root
+ inflectional affixes

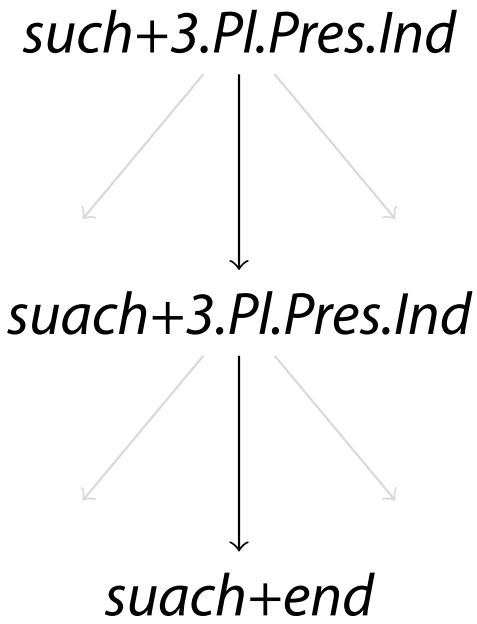
Swiss German word generation





Swiss German word generation

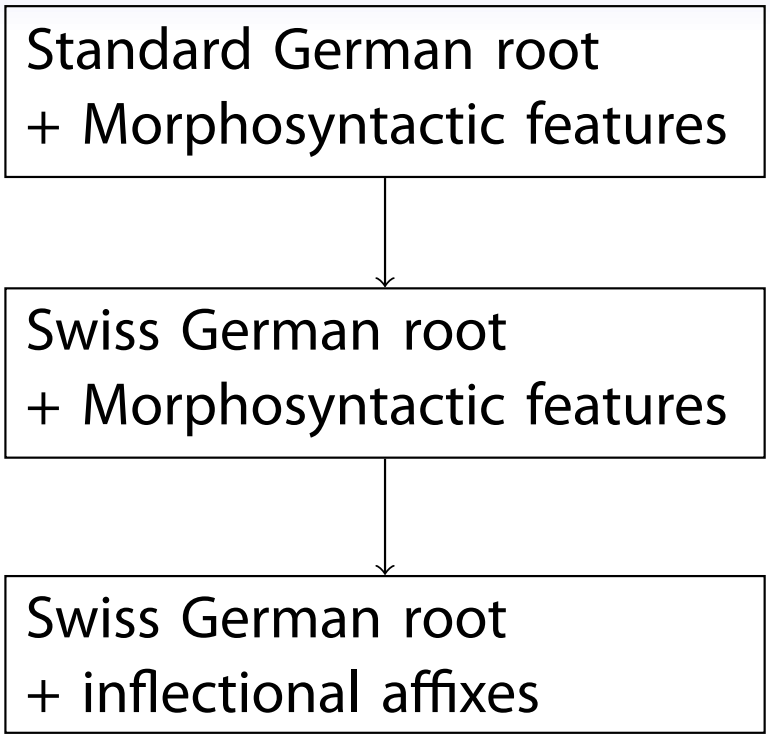


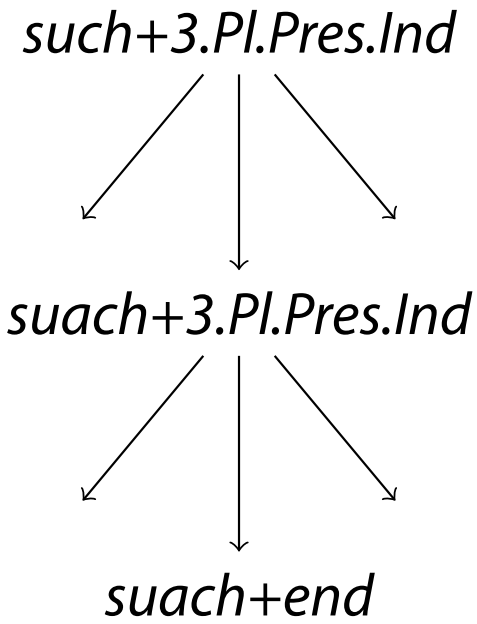


A cross-lingual approach:

Apply **phonetic and lexical transformations** to the Standard German root to obtain a dialect root, then add inflection.

Swiss German word generation





A multi-dialect approach:

Transformation rules yield different outputs for different dialects.
The rules are **georeferenced**, i.e. linked to dialectological maps.

Variables and variants

A **variable** is any linguistic phenomenon whose realisation varies along the geographical axis. The different realisations are called **variants**.

Examples:

- A phonetic transformation
- A lexical transformation
- An affixation rule

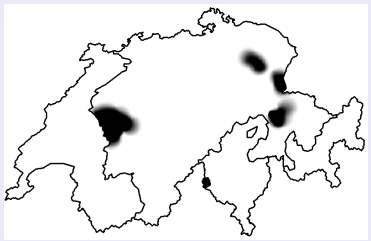
Variables and variants

A **variable** is any linguistic phenomenon whose realisation varies along the geographical axis. The different realisations are called **variants**.

Example of a phonetic transformation:

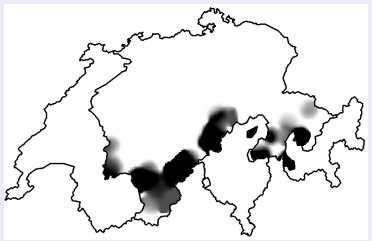
Variable: Standard German *-nd* in word-final position
e.g. *Hund 'dog', Kind 'child'*

Variants: *-nn* [n], *-ng* [ŋ], *-nt* [nt], *-nd* [nd̥]

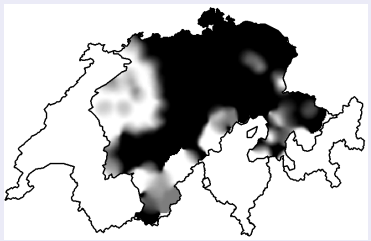


nd → *nn*

nd → *ng*



nd → *nt*



keep *nd*

Black: p = 1

White: p = 0

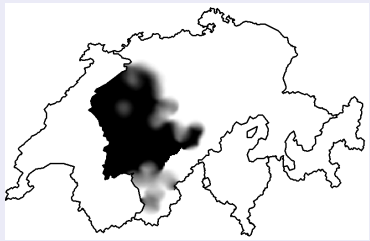
Variables and variants

A **variable** is any linguistic phenomenon whose realisation varies along the geographical axis. The different realisations are called **variants**.

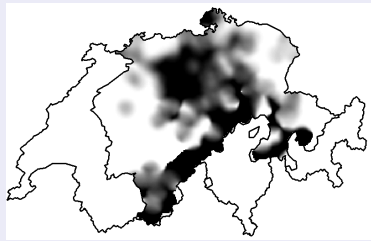
Example of a lexical transformation:

Variable: Standard German *immer 'always'*

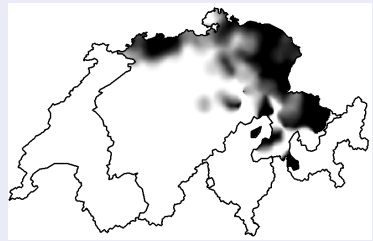
Variants: *geng, immer, all, …*



immer → *geng*



immer → *immer*



immer → *all*

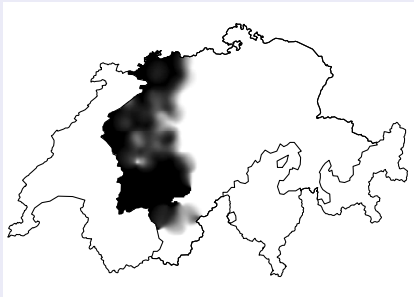
Variables and variants

A **variable** is any linguistic phenomenon whose realisation varies along the geographical axis. The different realisations are called **variants**.

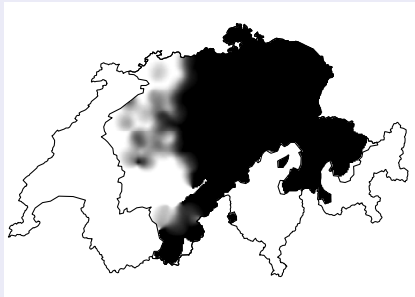
Example of an affixation rule:

Variable: Weak nominative singular adjective suffix (e.g. *die schwarz**e** Katze* 'the black cat')

Variants: *-i*, *-0*



ADJA.Nom.Sg.Weak → *-i*



ADJA.Nom.Sg.Weak → *-0*

Map data

Sprachatlas der deutschen Schweiz (SDS):

Data collection: 1939-1958
600 inquiry points in German-speaking Switzerland

Publication: 1962-1997
8 volumes, 1500 hand-drawn maps

Phonetics: 2 volumes, 400 maps
→ 60 maps scanned, digitized and converted

Morphology: 1 volume, 250 maps
→ 110 maps scanned, digitized and converted

Lexicon: 5 volumes
→ 30 maps scanned, digitized and converted

Map conversion: Interpolate point maps to continuous surface maps with kernel estimator method

Map data

Sprachatlas der deutschen Schweiz (SDS):

Data collection: 1939-1958
600 inquiry points in German-speaking Switzerland

Publication: 1962-1997
8 volumes, 1500 hand-drawn maps

Phonetics: 2 volumes, 400 maps
$\rightarrow$ 60 maps scanned, digitized and converted

Morphology: 1 volume, 250 maps
$\rightarrow$ 110 maps scanned, digitized and converted

Lexicon: 5 volumes
$\rightarrow$ 30 maps scanned, digitized and converted

Map conversion: Interpolate point maps to continuous surface maps with kernel estimator method

Map data

Sprachatlas der deutschen Schweiz (SDS):

Data collection:	1939-1958 600 inquiry points in German-speaking Switzerland
Publication:	1962-1997 8 volumes, 1500 hand-drawn maps
Phonetics:	2 volumes, 400 maps $\rightarrow$ 60 maps scanned, digitized and converted
Morphology:	1 volume, 250 maps $\rightarrow$ 110 maps scanned, digitized and converted
Lexicon:	5 volumes $\rightarrow$ 30 maps scanned, digitized and converted
Map conversion:	Interpolate point maps to continuous surface maps with kernel estimator method

Implementation

Rules are stored in a database with the following tables:

- phoneticVariables, lexicalVariables, affixVariables:
  - List the conditions when a specific rule has to be applied
    - Regular expressions
    - Morphosyntactic features
- phoneticVariants, lexicalVariants, affixVariants:
  - Specify the changes to be made
    - Simple suffixation
    - Regular expression replacement
    - Addition of morphosyntactic features
  - Specify the area in which the variants are valid
    - Pointer to a map file

Implementation is done with Python scripts, no finite state toolkit

- More flexible (include map filtering with linear algebra)
- But slower and not bidirectional (only generation mode)

Implementation

Rules are stored in a database with the following tables:

- phoneticVariables, lexicalVariables, affixVariables:
- List the conditions when a specific rule has to be applied
  - Regular expressions
  - Morphosyntactic features
- phoneticVariants, lexicalVariants, affixVariants:
- Specify the changes to be made
  - Simple suffixation
  - Regular expression replacement
  - Addition of morphosyntactic features
- Specify the area in which the variants are valid
  - Pointer to a map file

Implementation is done with Python scripts, no finite state toolkit

- More flexible (include map filtering with linear algebra)
- But slower and not bidirectional (only generation mode)

Implementation

Rules are stored in a database with the following tables:

- phoneticVariables, lexicalVariables, affixVariables:
- List the conditions when a specific rule has to be applied
  - Regular expressions
  - Morphosyntactic features
- phoneticVariants, lexicalVariants, affixVariants:
- Specify the changes to be made
  - Simple suffixation
  - Regular expression replacement
  - Addition of morphosyntactic features
- Specify the area in which the variants are valid
  - Pointer to a map file

Implementation is done with Python scripts, no finite state toolkit

- More flexible (include map filtering with linear algebra)
- But slower and not bidirectional (only generation mode)

Problematic cases

Lexical restrictions

- Phonetic context and morphosyntactic features are not always sufficient:

  Haus → *Huus* 'house' but *Baum* → *Baum* 'tree'

  - Two distinct vowels in Middle High German, have fallen together in Modern Standard German, but remained distinct in Swiss German.
- Solution: use word lists.

Short verbs

- Verbs with monosyllabic infinitive and plural forms
  - Two in Standard German (*sein, tun*)
  - Up to a dozen in Swiss German (but not all of them are short in all dialects…)
- Solution:
  - Generate "long" or "short" root and add a specific feature
  - Generate "long" or "short" affixes depending on feature

Problematic cases

Lexical restrictions

- Phonetic context and morphosyntactic features are not always sufficient:

  Haus → *Huus* 'house' but *Baum* → *Baum* 'tree'

  - Two distinct vowels in Middle High German, have fallen together in Modern Standard German, but remained distinct in Swiss German.
- Solution: use word lists.

Short verbs

- Verbs with monosyllabic infinitive and plural forms
  - Two in Standard German (*sein, tun*)
  - Up to a dozen in Swiss German (but not all of them are short in all dialects…)
- Solution:
  - Generate "long" or "short" root and add a specific feature
  - Generate "long" or "short" affixes depending on feature

Outline

Experimental setup

1. Generate a bilingual dictionary Standard German ↔ Swiss German
   - Extract leaf nodes from Standard German TIGER corpus (morphosyntactically annotated)
   - Apply transformation rules
   - Store resulting dialect words with respective areas of validity
2. Take 100 sentences in five dialects, and translate each word back to Standard German (dictionary lookup).
   - The Swiss German Wikipedia provides texts in several dialects:

     BA Basel
     BE Bern
     OS Eastern Switzerland
     WS Wallis
     ZH Zurich
3. Check if the result is correct.

Experimental setup

1. Generate a bilingual dictionary Standard German ↔ Swiss German
   - Extract leaf nodes from Standard German TIGER corpus (morphosyntactically annotated)
   - Apply transformation rules
   - Store resulting dialect words with respective areas of validity
2. Take 100 sentences in five dialects, and translate each word back to Standard German (dictionary lookup).
   - The Swiss German Wikipedia provides texts in several dialects:
     - BA Basel
     - BE Bern
     - OS Eastern Switzerland
     - WS Wallis
     - ZH Zurich
3. Check if the result is correct.

Experimental setup

1. Generate a bilingual dictionary Standard German ↔ Swiss German
   - Extract leaf nodes from Standard German TIGER corpus (morphosyntactically annotated)
   - Apply transformation rules
   - Store resulting dialect words with respective areas of validity
2. Take 100 sentences in five dialects, and translate each word back to Standard German (dictionary lookup).
   - The Swiss German Wikipedia provides texts in several dialects:

     BA Basel
     BE Bern
     OS Eastern Switzerland
     WS Wallis
     ZH Zurich
3. Check if the result is correct.

Results

Without geographical filtering

- In this scenario, when a Basel dialect word is analyzed, the system may also return derivations that are only valid in the region of Bern.

Types					Tokens				
BA	BE	OS	WS	ZH	BA	BE	OS	WS	ZH
42%	40%	41%	25%	45%	62%	57%	60%	44%	65%

With geographical filtering

- Retain only analyses that obtain a minimal probability of 10% in the most representative city of the respective dialect area.

Types					Tokens				
BA	BE	OS	WS	ZH	BA	BE	OS	WS	ZH
37%	29%	27%	17%	40%	57%	47%	44%	30%	58%

Results

Without geographical filtering

- In this scenario, when a Basel dialect word is analyzed, the system may also return derivations that are only valid in the region of Bern.

Types					Tokens				
BA	BE	OS	WS	ZH	BA	BE	OS	WS	ZH
42%	40%	41%	25%	45%	62%	57%	60%	44%	65%

With geographical filtering

- Retain only analyses that obtain a minimal probability of 10% in the most representative city of the respective dialect area.

Types					Tokens				
BA	BE	OS	WS	ZH	BA	BE	OS	WS	ZH
37%	29%	27%	17%	40%	57%	47%	44%	30%	58%

Discussion

- Many compound nouns and proper nouns are not recognized because they are not in the TIGER corpus, and therefore not in the generated dictionary.
  - Upper bound: about 70% of types / 80% of tokens
- Figures correlate with perceived distance from Standard German
  - Northern lowland dialects (BA, ZH) achieve higher performances than the alpine dialect (WS).
- The reference city dialect is not necessarily the one used by the Wikipedia author.
  - The geographical filtering method may be too aggressive.

Error analysis

Spelling differences

- Generated: *bestaat* Corpus: *bestaht* 'consists' (ZH)
- The generated word follows the Dieth spelling guidelines, the word encountered in the corpus is closer to Standard German spelling.

Missing lexical rules

- Generated: *Chirche* Corpus: *Chile* 'church' (ZH)
- The lexical rule *Kirche* → *Chile* is missing. The generated form only uses generic phonetic rules.

Diachronic change

- Generated: *zäägt, zaagt, zeigt* Corpus: *zaigt* 'shows' (OS)
- The generated variants were frequently used at the time of the SDS inquiries, but are perceived as obsolete and marginal today.

Error analysis

Spelling differences

- Generated: *bestaat* Corpus: *bestaht* 'consists' (ZH)
- The generated word follows the Dieth spelling guidelines, the word encountered in the corpus is closer to Standard German spelling.

Missing lexical rules

- Generated: *Chirche* Corpus: *Chile* 'church' (ZH)
- The lexical rule *Kirche* $\rightarrow$ *Chile* is missing. The generated form only uses generic phonetic rules.

Diachronic change

- Generated: *zäägt, zaagt, zeigt* Corpus: *zaigt* 'shows' (OS)
- The generated variants were frequently used at the time of the SDS inquiries, but are perceived as obsolete and marginal today.

Error analysis

Spelling differences

- Generated: *bestaat* Corpus: *bestaht* 'consists' (ZH)
- The generated word follows the Dieth spelling guidelines, the word encountered in the corpus is closer to Standard German spelling.

Missing lexical rules

- Generated: *Chirche* Corpus: *Chile* 'church' (ZH)
- The lexical rule *Kirche* $\rightarrow$ *Chile* is missing. The generated form only uses generic phonetic rules.

Diachronic change

- Generated: *zäägt, zaagt, zeigt* Corpus: *zaigt* 'shows' (OS)
- The generated variants were frequently used at the time of the SDS inquiries, but are perceived as obsolete and marginal today.

Outline

Conclusion

Key contributions

- A cross-lingual model to benefit from a large number of cognates and a similar morphological inventory.
- Multi-dialectal coverage by using existing dialectological resources.

Potential applications

- Machine translation
- Morphosyntactic analysis to enhance information retrieval
- Speech recognition and synthesis

Future work

- Reimplementation with a finite state toolkit: bidirectionality, speed
- Better handling of spelling variants

Conclusion

Key contributions

- A cross-lingual model to benefit from a large number of cognates and a similar morphological inventory.
- Multi-dialectal coverage by using existing dialectological resources.

Potential applications

- Machine translation
- Morphosyntactic analysis to enhance information retrieval
- Speech recognition and synthesis

Future work

- Reimplementation with a finite state toolkit: bidirectionality, speed
- Better handling of spelling variants

Conclusion

Key contributions

- A cross-lingual model to benefit from a large number of cognates and a similar morphological inventory.
- Multi-dialectal coverage by using existing dialectological resources.

Potential applications

- Machine translation
- Morphosyntactic analysis to enhance information retrieval
- Speech recognition and synthesis

Future work

- Reimplementation with a finite state toolkit: bidirectionality, speed
- Better handling of spelling variants

Outline

Demo

http://latlntic.unige.ch/~scherrey/

- Interactive point maps
- Interactive surface maps
- Morphology generation for five predefined dialects
- Word-by-word translation to any dialect
- Dialect identification