Morphological Analysis and Generation in Computer-Assisted Teaching of Indigenous Languages

Michael Gasser
School of Informatics
Indiana University
Goals
Goals

• Computational linguistics and (computer-assisted) language instruction
Goals

- Computational linguistics and (computer-assisted) language instruction
- Towards collaboration among fields in the development of teaching materials for indigenous languages
Intelligent computer-assisted language instruction
Intelligent computer-assisted language instruction

- The program models the learner.
Intelligent computer-assisted language instruction

• The program models the learner.
  - Based on computational theories (from AI, cognitive science) of language learning
Intelligent computer-assisted language instruction

• The program models the learner.
  - Based on computational theories (from AI, cognitive science) of language learning
• The program includes a model of the language.
Intelligent computer-assisted language instruction

- The program models the learner.
  - Based on computational theories (from AI, cognitive science) of language learning
- The program includes a model of the language.
  - Based on work in computational linguistics
Intelligent computer-assisted language instruction

• The program models the learner.
  - Based on computational theories (from AI, cognitive science) of language learning

• The program includes a model of the language.
  - Based on work in computational linguistics
    - Syntax and semantics
Intelligent computer-assisted language instruction

- The program models the learner.
  - Based on computational theories (from AI, cognitive science) of language learning

- The program includes a model of the language.
  - Based on work in computational linguistics
    - Syntax and semantics
    - Lexicon
Intelligent computer-assisted language instruction

• The program models the learner.
  - Based on computational theories (from AI, cognitive science) of language learning

• The program includes a model of the language.
  - Based on work in computational linguistics
    - Syntax and semantics
    - Lexicon
    - Discourse pragmatics
Intelligent computer-assisted language instruction

• The program models the learner.
  - Based on computational theories (from AI, cognitive science) of language learning

• The program includes a model of the language.
  - Based on work in computational linguistics
    • Syntax and semantics
    • Lexicon
    • Discourse pragmatics
    • Morphology
Morphologically complex languages
Morphologically complex languages

- Inflection
Morphologically complex languages

- Inflection
  - Tense, aspect, mood
Morphologically complex languages

• Inflection
  - Tense, aspect, mood
  - Subject and object agreement
Morphologically complex languages

• Inflection
  - Tense, aspect, mood
  - Subject and object agreement
    - *you see me*
      - *kinawilo* (K’iche’)
      - *qhawawanki* (Quechua)
Morphologically complex languages

- Inflection
  - Tense, aspect, mood
  - Subject and object agreement
    - you see me
      kinawilo (K’iche’)
      qhawawanki (Quechua)

- Derivation
Morphologically complex languages

- **Inflection**
  - Tense, aspect, mood
  - Subject and object agreement
    - *you see me*
    - *kinawilo* (K’iche’)
    - *qhawawanki* (Quechua)

- **Derivation**
  - Part-of-speech to part-of-speech
Morphologically complex languages

• Inflection
  - Tense, aspect, mood
  - Subject and object agreement
    - you see me
    - kinawilo (K’iche’)
    - qhawawanki (Quechua)

• Derivation
  - Part-of-speech to part-of-speech
  - Transitivity and argument structure
Morphologically complex languages

- **Inflection**
  - Tense, aspect, mood
  - Subject and object agreement
    - you see me
      - *kinawilo* (K’iche’)
      - *qhawawanki* (Quechua)

- **Derivation**
  - Part-of-speech to part-of-speech
  - Transitivity and argument structure
  - Time, movement, manner
Morphologically complex languages

- **Inflection**
  - Tense, aspect, mood
  - Subject and object agreement
    - *you see me*
    - *kinawilo* (K’iche’)
    - *qhawawanki* (Quechua)

- **Derivation**
  - Part-of-speech to part-of-speech
  - Transitivity and argument structure
  - Time, movement, manner
    - *rimay ‘speak’, rimaykuy ‘explain’* (Quechua)
Morphologically complex languages

- **Inflection**
  - Tense, aspect, mood
  - Subject and object agreement
    - \textit{you see me}
    - \textit{kinawilo} (K’iche’)
    - \textit{qhawawanki} (Quechua)

- **Derivation**
  - Part-of-speech to part-of-speech
  - Transitivity and argument structure
  - Time, movement, manner
    - \textit{rimay} ‘speak’, \textit{rimaykuy} ‘explain’ (Quechua)
    - \textit{llank’a} ‘work’, \textit{llank’arpary} ‘work completely’ (Quechua)
Morphologically complex languages
Morphologically complex languages

• Noun incorporation
Morphologically complex languages

• Noun incorporation
  - *mepitetatsihetixakítetuiritositixiaximekaitikaku*
    ‘however, they already wanted to start ordering us to bring them the *dishes*’ (Huichol, Iturrioz & Gómez López)
Morphology as a fundamentally different way of organizing concepts
Morphology as a fundamentally different way of organizing concepts

- Small number of units combined through a large number of productive and semi-productive processes
Morphology as a fundamentally different way of organizing concepts

- Small number of units combined through a large number of productive and semi-productive processes
- Words in Nuxalk, words in English
Morphology as a fundamentally different way of organizing concepts

- Small number of units combined through a large number of productive and semi-productive processes
- Words in Nuxalk, words in English
  - *smnmnmuuuc* ‘mute’
Morphology as a fundamentally different way of organizing concepts

- Small number of units combined through a large number of productive and semi-productive processes
- Words in Nuxalk, words in English
  - smnmnuuc ‘mute’
  - nusʔaˈqulikʷsta ‘prison’
Morphology as a fundamentally different way of organizing concepts

- Small number of units combined through a large number of productive and semi-productive processes
- Words in Nuxalk, words in English
  - smnmmnuuc ‘mute’
  - nusʔaʔulikʷsta ‘prison’
  - snuɬnusɬtɬ ‘twins’
Morphology as a fundamentally different way of organizing concepts

- Small number of units combined through a large number of productive and semi-productive processes
- Words in Nuxalk, words in English
  - smnmnuuc ‘mute’
  - nusʔaʔulikʷsta ‘prison’
  - snuʔnusltɬ ‘twins’
  - kiʔnuus ‘eight’
Morphology as a fundamentally different way of organizing concepts

• Small number of units combined through a large number of productive and semi-productive processes
• Words in Nuxalk, words in English
  - smnmnmnuuc ‘mute’
  - nusʔaʔulikʷsta ‘prison’
  - snuɬnuslɬsta ‘twins’
  - kiłnuus ‘eight’
  - ?itɬuɬčuuɬ ‘naked’
Morphology as a fundamentally different way of organizing concepts

- Small number of units combined through a large number of productive and semi-productive processes
- Words in Nuxalk, words in English
  - smmnmnmuuc ‘mute’
  - nusʔaʔulikʷsta ‘prison’
  - sнуʔnusltɬ ‘twins’
  - ɬiʔnuus ‘eight’
  - ʔitɬuɬčuuɬ ‘naked’
  - nutixcnalsiixʷ ‘uvula’
Morphology as a fundamentally different way of organizing concepts

• Small number of units combined through a large number of productive and semi-productive processes

• Words in Nuxalk, words in English
  - smnmmmcuc ‘mute’
  - nusʔaʔulikʷsta ‘prison’
  - snuʔnusltɬ ‘twins’
  - kiʔnuus ‘eight’
  - ?itluʔčuuɬ ‘naked’
  - nutixcnalsiiχʷ ‘uvula’
  - squpuc ‘beard’
Teaching morphology
Teaching morphology

- Students are made “aware” of the structure of words
Teaching morphology

- Students are made “aware” of the structure of words
- Computer-assisted morphology instruction
Teaching morphology

• Students are made “aware” of the structure of words
• Computer-assisted morphology instruction
  - The software makes word structure explicit
Teaching morphology

- Students are made “aware” of the structure of words
- Computer-assisted morphology instruction
  - The software makes word structure explicit
  - But how?
Teaching morphology

• Students are made “aware” of the structure of words
• Computer-assisted morphology instruction
  - The software makes word structure explicit
  - But how?
    • Through canned examples
Teaching morphology

• Students are made “aware” of the structure of words
• Computer-assisted morphology instruction
  - The software makes word structure explicit
  - But how?
    • Through canned examples
    • Through knowledge of the morphology
Computation morphology
Computation morphology

• Morphological analysis
Computation morphology

• Morphological analysis
  - Converts a surface form to a lexical/grammatical form
Computation morphology

• Morphological analysis
  - Converts a surface form to a lexical/grammatical form
  - A surface form is analyzed into its constituent morphemes
Computation morphology

- **Morphological analysis**
  - Converts a *surface* form to a *lexical/grammatical* form
  - A surface form is analyzed into its constituent morphemes
    - *kinawilo* → *k-in-aw-il-o*
Computation morphology

• Morphological analysis
  - Converts a surface form to a lexical/grammatical form
  - A surface form is analyzed into its constituent morphemes
    - $\text{kinawilo} \rightarrow \text{k-in-aw-il-o}$
  - A surface form is analyzed into a representation of its grammatical features
Computation morphology

- Morphological analysis
  - Converts a surface form to a lexical/grammatical form
  - A surface form is analyzed into its constituent morphemes
    - \textit{kinawilo} $\rightarrow$ \textit{k-in-aw-il-o}
  - A surface form is analyzed into a representation of its grammatical features
    - \textit{kinawilo} $\rightarrow$
      
      \begin{verbatim}
      [root='il',
       abs=[prs=1,num=sing],
       erg=[prs=2,num=sing,-form],
       tam=incmpl]
      \end{verbatim}
Computation morphology
Computation morphology

• Morphological generation
Computation morphology

• Morphological generation
  - An abstract representation of a word’s grammatical and lexical properties is converted into a surface form
Computation morphology

• Morphological generation
  - An abstract representation of a word’s grammatical and lexical properties is converted into a surface form
    - [root='il',
      abs=[prs=1,num=sing],
      erg=[prs=2,num=sing,-form],
      tam=incmpl]
    → kinawilo
Computational morphology theory
Computational morphology theory

• Finite-state transducers
Computational morphology theory

• Finite-state transducers
  - Morphological knowledge (grammatical and lexical) represented in the form of a network of states and transitions between them
Computational morphology theory

- Finite-state transducers
  - Morphological knowledge (grammatical and lexical) represented in the form of a network of states and transitions between them
  - An analysis of a word is a path through the network
Computational morphology theory

- **Finite-state transducers**
  - Morphological knowledge (grammatical and lexical) represented in the form of a network of states and transitions between them
  - An analysis of a word is a path through the network
  - Analysis and generation handled by the same network
Computational morphology theory

• Finite-state transducers
  - Morphological knowledge (grammatical and lexical) represented in the form of a network of states and transitions between them
  - An analysis of a word is a path through the network
  - Analysis and generation handled by the same network
  - Ambiguous inputs yield multiple paths through the network
Computational morphology theory

• Finite-state transducers
  - Morphological knowledge (grammatical and lexical) represented in the form of a network of states and transitions between them
  - An analysis of a word is a path through the network
  - Analysis and generation handled by the same network
  - Ambiguous inputs yield multiple paths through the network
• Feature-structure representations
Computational morphology theory

• Finite-state transducers
  - Morphological knowledge (grammatical and lexical) represented in the form of a network of states and transitions between them
  - An analysis of a word is a path through the network
  - Analysis and generation handled by the same network
  - Ambiguous inputs yield multiple paths through the network

• Feature-structure representations
  - Grammatical knowledge takes the form of structured representations, like those we’d use for syntax
Teaching morphology using morphological analyzers/generators
Teaching morphology using morphological analyzers/generators

• Robust
Teaching morphology using morphological analyzers/generators

- Robust
  - Can respond to forms that are not included in its database
Teaching morphology using morphological analyzers/generators

• Robust
  - Can respond to forms that are not included in its database
• Extensible
Teaching morphology using morphological analyzers/generators

- Robust
  - Can respond to forms that are not included in its database
- Extensible
  - Can easily accommodate new lexical items and new rules
An analyzer/generator for K’iche’ and Quechua verbs
An analyzer/generator for K’iche’ and Quechua verbs

• Analysis
An analyzer/generator for K’iche’ and Quechua verbs

- Analysis
  - Morpheme segmentation
An analyzer/generator for K’iche’ and Quechua verbs

- Analysis
  - Morpheme segmentation
  - Lexical/grammatical representation
An analyzer/generator for K’iche’ and Quechua verbs

• Analysis
  - Morpheme segmentation
  - Lexical/grammatical representation
• Accommodates multiple analyses, surface outputs
An analyzer/generator for K’iche’ and Quechua verbs

• Analysis
  - Morpheme segmentation
  - Lexical/grammatical representation

• Accommodates multiple analyses, surface outputs
  • (to program)
Transitivity, perspective taking in K’iche’
Transitivity, perspective taking in K’iche’

- Presenting L2 from an L2 perspective
Transitivity, perspective taking in K’iche’

- Presenting L2 from an L2 perspective
- A K’iche’ verb root can take multiple suffixes, yielding different stems that convey different perspectives on the same event.
Transitivity, perspective taking in K’iche’

• Presenting L2 from an L2 perspective
• A K’iche’ verb root can take multiple suffixes, yielding different stems that convey different perspectives on the same event.
• Ergative syntax
Transitivity, perspective taking in K’iche’

• Presenting L2 from an L2 perspective
• A K’iche’ verb root can take multiple suffixes, yielding different stems that convey different perspectives on the same event.
• Ergative syntax
  - ABSOLUTIVE: intransitive subjects, transitive objects
    ERGATIVE: transitive subjects
Transitivity, perspective taking in K’iche’

• Presenting L2 from an L2 perspective
• A K’iche’ verb root can take multiple suffixes, yielding different stems that convey different perspectives on the same event.
• Ergative syntax
  - ABSOLUTIVE: intransitive subjects, transitive objects
  - ERGATIVE: transitive subjects
    - Ke’akunaj.
      You cure them.
Transitivity, perspective taking in K’iche’

- Presenting L2 from an L2 perspective
- A K’iche’ verb root can take multiple suffixes, yielding different stems that convey different perspectives on the same event.
- Ergative syntax
  - ABSOLUTIVE: intransitive subjects, transitive objects
  - ERGATIVE: transitive subjects
    - *Ke’akunaj.*
      - You cure them.
    - *Ke’warik.*
      - They sleep.
Transitivizing and intransitivizing suffixes in K’iche’
Transitivizing and intransitivizing suffixes in K’iche’

- Causative
Transitivizing and intransitivizing suffixes in K’iche’

- **Causative**
  - Converts an intransitive to a transitive verb, adding a participant
Transitivizing and intransitivizing suffixes in K’iche’

- **Causative**
  - Converts an intransitive to a transitive verb, adding a participant
  - Intransitive ABSOLUTIVE remains ABSOLUTIVE; ERGATIVE added
Transitivizing and intransitivizing suffixes in K’iche’

- **Causative**
  - Converts an intransitive to a transitive verb, adding a participant
  - Intransitive **ABSOLUTIVE** remains **ABSOLUTIVE**; **ERGATIVE** added
    - *Ke’atínik.*
      - *They bathe.*
Transitivizing and intransitivizing suffixes in K’iche’

- **Causative**
  - Converts an intransitive to a transitive verb, adding a participant
  - Intransitive **ABSOLUTIVE remains ABSOLUTIVE; ERGATIVE** added
    - *Ke’atinik.*
      *They bathe.*
    - *Ke’awatinisaj.*
      *You bathe them.*
Transitivizing and intransitivizing suffixes in K’iche’
Transitivizing and intransitivizing suffixes in K’iche’

• Passive
Transitivizing and intransitivizing suffixes in K’iche’

• Passive
  - Converts a transitive to an intransitive verb, foregrounding the ABSOLUTIVE (object) participant
Transitivizing and intransitivizing suffixes in K’iche’

• Passive
  - Converts a transitive to an intransitive verb, foregrounding the ABSOLUTIVE (object) participant
  - ABSOLUTIVE remains ABSOLUTIVE; ERGATIVE is deleted (or demoted)
Transitivizing and intransitivizing suffixes in K’iche’

- Passive
  - Converts a transitive to an intransitive verb, foregrounding the ABSOLUTIVE (object) participant
  - ABSOLUTIVE remains ABSOLUTIVE; ERGATIVE is deleted (or demoted)
    - *Ke’akunaj.*
    - *You cure them.*
Transitivizing and intransitivizing suffixes in K’iche’

• Passive
  - Converts a transitive to an intransitive verb, foregrounding the ABSOLUTIVE (object) participant
  - ABSOLUTIVE remains ABSOLUTIVE; ERGATIVE is deleted (or demoted)
    • Ke’akunaj.
      You cure them.
    • Ke’kunaxik.
      They are cured.
Transitivizing and intransitivizing suffixes in K’iche’

- **Passive**
  - Converts a transitive to an intransitive verb, foregrounding the ABSOLUTIVE (object) participant
  - **ABSOLUTIVE** remains **ABSOLUTIVE**; **ERGATIVE** is deleted (or demoted)
    - *Ke’akunaj.*
      You cure them.
    - *Ke’kunaxik.*
      They are cured.
    - *Ke’awatinisaj.*
      You bathe them.
Transitivizing and intransitivizing suffixes in K’iche’

- Passive
  - Converts a transitive to an intransitive verb, foregrounding the ABSOLUTIVE (object) participant
  - ABSOLUTIVE remains ABSOLUTIVE; ERGATIVE is deleted (or demoted)
    - Ke’akunaj.
      You cure them.
    - Ke’kunaxik.
      They are cured.
    - Ke’awatinisaj.
      You bathe them.
    - Ke’atinisaxik.
      They are bathed.
Transitivizing and intransitivizing suffixes in K’iche’
Transitivizing and intransitivizing suffixes in K’iche’

• Antipassive
Transitivizing and intransitivizing suffixes in K’iche’

• Antipassive
  - Converts a transitive to an intransitive verb, foregrounding the **ERGATIVE** (subject) participant
Transitivizing and intransitivizing suffixes in K’iche’

• Antipassive
  - Converts a transitive to an intransitive verb, foregrounding the ERGATIVE (subject) participant
  - ERGATIVE becomes ABSOLUTIVE; original ABSOLUTIVE is deleted (or demoted)
Transitivizing and intransitivizing suffixes in K’iche’

• Antipassive
  - Converts a transitive to an intransitive verb, foregrounding the ERGATIVE (subject) participant
  - ERGATIVE becomesABSOLUTIVE; original ABSOLUTIVE is deleted (or demoted)
    - Ke’akunaj.
      You cure them.
Transitivizing and intransitivizing suffixes in K’iche’

• Antipassive
  - Converts a transitive to an intransitive verb, foregrounding the ERGATIVE (subject) participant
  - ERGATIVE becomes ABSOLUTIVE; original ABSOLUTIVE is deleted (or demoted)
    - *Ke’akunaj.*
      You cure *them*.
    - *Katkunanik.*
      You cure (somebody).
Transitivizing and intransitivizing suffixes in K’iche’

- **Antipassive**
  - Converts a transitive to an intransitive verb, foregrounding the **ERGATIVE** (subject) participant
  - **ERGATIVE** becomes **ABSOLUTIVE**; original **ABSOLUTIVE** is deleted (or demoted)
    - *Ke’akunaj.*
      *You cure them.*
    - *Katkunanik.*
      *You cure (somebody).*
    - *Ke’awatinisaj.*
      *You bathe them.*
Transitivizing and intransitivizing suffixes in K’iche’

• Antipassive
  - Converts a transitive to an intransitive verb, foregrounding the ERGATIVE (subject) participant
  - ERGATIVE becomes ABSOLUTE; original ABSOLUTE is deleted (or demoted)
    • Ke’akunaj.
      You cure them.
    • Katkunanik.
      You cure (somebody).
    • Ke’awatinisaj.
      You bathe them.
    • Katatinisanik.
      You bathe (somebody).
Transitivizing and intransitivizing suffixes in K’iche’

• Antipassive
  - Converts a transitive to an intransitive verb, foregrounding the ERGATIVE (subject) participant
  - ERGATIVE becomes ABSOLUTIVE; original ABSOLUTIVE is deleted (or demoted)
    - Ke’akunaj.  
      You cure them.
    - Katkunanik.  
      You cure (somebody).
    - Ke’awatinisaj.  
      You bathe them.
    - Katatinisanik.  
      You bathe (somebody).
  - (to program)
Conclusions
Conclusions

- Many indigenous languages of the Americas feature complex morphology.
Conclusions

- Many indigenous languages of the Americas feature complex morphology.
- Teaching morphology means making word structure explicit.
Conclusions

- Many indigenous languages of the Americas feature complex morphology.
- Teaching morphology means making word structure explicit.
- Intelligent computer-assisted teaching of morphology incorporates rule-based analysis and generation of words, bringing robustness and extensibility.
Conclusions

• Many indigenous languages of the Americas feature complex morphology.
• Teaching morphology means making word structure explicit.
• Intelligent computer-assisted teaching of morphology incorporates rule-based analysis and generation of words, bringing robustness and extensibility.
• Computational techniques applied to K’iche’ could be integrated into teaching materials for K’iche’ or other languages.
Thank you!
¡Maltyox!
¡Añaychaykichis!