Computational Linguistics and Mayan Languages

Michael Gasser
School of Informatics
Indiana University

Languages

Languages

- A language is

Languages

- A language is
  - A set of conventions shared by a linguistic community concerning pronunciation, vocabulary, grammar, and usage

Languages

- A language is
  - A set of conventions shared by a linguistic community concerning pronunciation, vocabulary, grammar, and usage
    - Potentially different conventions for different functions/genres: informal conversation, formal oratory

Languages

- A language is
  - A set of conventions shared by a linguistic community concerning pronunciation, vocabulary, grammar, and usage
    - Potentially different conventions for different functions/genres: informal conversation, formal oratory
  - A set of lexical and grammatical categories that constitute a way of slicing up the “world”, an interpretation of reality

Languages

- A language is
  - A set of conventions shared by a linguistic community concerning pronunciation, vocabulary, grammar, and usage
    - Potentially different conventions for different functions/genres: informal conversation, formal oratory
  - A set of lexical and grammatical categories that constitute a way of slicing up the “world”, an interpretation of reality
    - *Eat* in English, etc., multiple words in Mayan languages depending on the nature of the eaten thing (*eat fruit*, *eat meat*, *eat soft thing*)

Languages

- A language is
  - A set of conventions shared by a linguistic community concerning pronunciation, vocabulary, grammar, and usage
    - Potentially different conventions for different functions/genres: informal conversation, formal oratory
  - A set of lexical and grammatical categories that constitute a way of slicing up the “world”, an interpretation of reality
    - *Eat* in English, etc., multiple words in Mayan languages depending on the nature of the eaten thing (*eat fruit*, *eat meat*, *eat soft thing*)
    - Shape vs. material biases in English, etc. and Yucatec Mayan

Languages

- A language is
  - A set of conventions shared by a linguistic community concerning pronunciation, vocabulary, grammar, and usage
    - Potentially different conventions for different functions/genres: informal conversation, formal oratory
  - A set of lexical and grammatical categories that constitute a way of slicing up the "world", an interpretation of reality
    - *Eat* in English, etc., multiple words in Mayan languages depending on the nature of the eaten thing (*eat fruit*, *eat meat*, *eat soft thing*)
    - Shape vs. material biases in English, etc. and Yucatec Mayan
    - Ergative vs. nominative languages

Languages

- A language is
  - A set of conventions shared by a linguistic community concerning pronunciation, vocabulary, grammar, and usage
    - Potentially different conventions for different functions/genres: informal conversation, formal oratory
  - A set of lexical and grammatical categories that constitute a way of slicing up the "world", an interpretation of reality
    - *Eat* in English, etc., multiple words in Mayan languages depending on the nature of the eaten thing (*eat fruit*, *eat meat*, *eat soft thing*)
    - Shape vs. material biases in English, etc. and Yucatec Mayan
    - Ergative vs. nominative languages
  - All of the available data in the language (uses by native speakers/writers in context)

Linguistic revolutions

Linguistic revolutions

- Linguistic revolutions — Writing, Printing, Computing

Linguistic revolutions

- Linguistic revolutions — Writing, Printing, Computing
  - Changed the nature of language

Linguistic revolutions

- Linguistic revolutions — Writing, Printing, Computing
  - Changed the nature of language
  - Changed individual languages

The First and Second Linguistic Revolutions: Writing, Printing

The First and Second Linguistic Revolutions: Writing, Printing

- Changes in languages that were written

The First and Second Linguistic Revolutions: Writing, Printing

- Changes in languages that were written
  - New functions, genres (bookkeeping, the letter, the receipt, the novel, the parking ticket, the crib sheet)

The First and Second Linguistic Revolutions: Writing, Printing

- Changes in languages that were written
  - New functions, genres (bookkeeping, the letter, the receipt, the novel, the parking ticket, the crib sheet)
  - New processing constraints and possibilities: new grammatical constructions

The First and Second Linguistic Revolutions: Writing, Printing

- Changes in languages that were written
  - New functions, genres (bookkeeping, the letter, the receipt, the novel, the parking ticket, the crib sheet)
  - New processing constraints and possibilities: new grammatical constructions
  - Increase in the database

The Second Linguistic Revolution: Printing

The Second Linguistic Revolution: Printing

- By greatly facilitating copying, greatly increased speed and range of access to knowledge for literate native readers of written languages, but access to printed material still restricted even among this population

The Second Linguistic Revolution: Printing

- By greatly facilitating copying, greatly increased speed and range of access to knowledge for literate native readers of written languages, but access to printed material still restricted even among this population
- Increased population capable of creating accessible knowledge, but, because of costs, left this still quite limited

Privileged and disadvantaged languages

Privileged and disadvantaged languages

- At the dawn of the Digital Revolution

Privileged and disadvantaged languages

- At the dawn of the Digital Revolution
  - Most languages remained essentially unwritten

Privileged and disadvantaged languages

- At the dawn of the Digital Revolution
  - Most languages remained essentially unwritten
  - Most speakers of many written languages were illiterate

Privileged and disadvantaged languages

- At the dawn of the Digital Revolution
  - Most languages remained essentially unwritten
  - Most speakers of many written languages were illiterate
  - Among written languages, the availability of material varied greatly (not necessarily as a function of the number of speakers: Danish, Finnish, Hebrew vs. Hindi, Bengali, Indonesian, Swahili)

Privileged and disadvantaged languages

- At the dawn of the Digital Revolution
  - Most languages remained essentially unwritten
  - Most speakers of many written languages were illiterate
  - Among written languages, the availability of material varied greatly (not necessarily as a function of the number of speakers: Danish, Finnish, Hebrew vs. Hindi, Bengali, Indonesian, Swahili)
  - Many speakers of disadvantaged languages relied on other languages for important functions such as accessing archived information

Privileged and disadvantaged languages

- At the dawn of the Digital Revolution
  - Most languages remained essentially unwritten
  - Most speakers of many written languages were illiterate
  - Among written languages, the availability of material varied greatly (not necessarily as a function of the number of speakers: Danish, Finnish, Hebrew vs. Hindi, Bengali, Indonesian, Swahili)
  - Many speakers of disadvantaged languages relied on other languages for important functions such as accessing archived information
  - Speakers of disadvantaged languages remained handicapped in participation in national and global debates about the future

The Third Linguistic Revolution: the Digital Revolution

The Third Linguistic Revolution: the Digital Revolution

- Changes in the affected languages

The Third Linguistic Revolution: the Digital Revolution

- Changes in the affected languages
  - New functions and genres (email, chat, blogs, text messaging)

The Third Linguistic Revolution: the Digital Revolution

- Changes in the affected languages
  - New functions and genres (email, chat, blogs, text messaging)
  - Massive increase in the database

The Third Linguistic Revolution: the Digital Revolution

- Changes in the affected languages
  - New functions and genres (email, chat, blogs, text messaging)
  - Massive increase in the database
- The promise for disadvantaged languages

The Third Linguistic Revolution: the Digital Revolution

- Changes in the affected languages
  - New functions and genres (email, chat, blogs, text messaging)
  - Massive increase in the database
- The promise for disadvantaged languages
  - Democratic nature of the Internet

The Third Linguistic Revolution: the Digital Revolution

- Changes in the affected languages
  - New functions and genres (email, chat, blogs, text messaging)
  - Massive increase in the database
- The promise for disadvantaged languages
  - Democratic nature of the Internet
  - Possibility of bypassing printing

The Third Linguistic Revolution: the Digital Revolution

- Changes in the affected languages
  - New functions and genres (email, chat, blogs, text messaging)
  - Massive increase in the database
- The promise for disadvantaged languages
  - Democratic nature of the Internet
  - Possibility of bypassing printing
  - Possibility of partially automating translation

The Third Linguistic Revolution: the Digital Revolution

- Changes in the affected languages
  - New functions and genres (email, chat, blogs, text messaging)
  - Massive increase in the database
- The promise for disadvantaged languages
  - Democratic nature of the Internet
  - Possibility of bypassing printing
  - Possibility of partially automating translation
  - Computer-assisted literacy training

The “Information Society”

The “Information Society”

- World Summit on Information Society principles (Geneva, 2003; Tunis, 2005)

The “Information Society”

- World Summit on Information Society principles (Geneva, 2003; Tunis, 2005)
  - 3: “The ability for all to access and contribute information, ideas and knowledge is essential in an inclusive Information Society.”

The “Information Society”

- World Summit on Information Society principles (Geneva, 2003; Tunis, 2005)
  - 3: “The ability for all to access and contribute information, ideas and knowledge is essential in an inclusive Information Society.”
  - 8: “The Information Society should be founded on and stimulate respect for cultural identity, cultural and linguistic diversity, traditions and religions, and foster dialogue among cultures and civilizations. The promotion, affirmation and preservation of diverse cultural identities and languages ... will further enrich the Information Society. ”

The reality

The reality

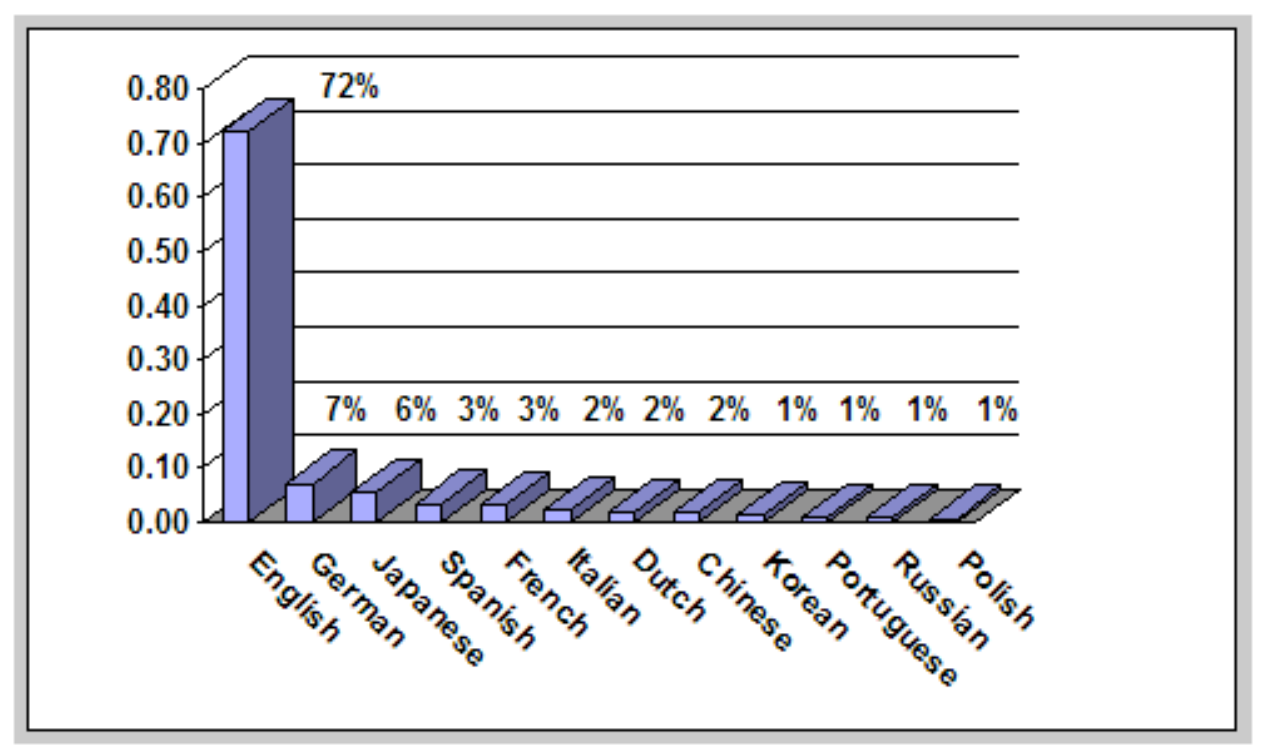
- One language dominates the Internet (~70% of web pages). 12 languages account for ~97% of all web pages. (O'Neill, Lavoie, Bennett)

The reality

- One language dominates the Internet (~70% of web pages). 12 languages account for ~97% of all web pages. (O'Neill, Lavoie, Bennett)
- Distribution corresponds closely to that in the world's libraries.

The reality

- One language dominates the Internet (~70% of web pages). 12 languages account for ~97% of all web pages. (O'Neill, Lavoie, Bennett)
- Distribution corresponds closely to that in the world's libraries.



The reality: Wikipedia

The reality: Wikipedia

- 251 languages

The reality: Wikipedia

- 251 languages
- English: 1,743,312 articles

The reality: Wikipedia

- 251 languages
- English: 1,743,312 articles
- Swedish: 222,821 articles

The reality: Wikipedia

- 251 languages
- English: 1,743,312 articles
- Swedish: 222,821 articles
- Quechua: 2,102 articles

The reality: Wikipedia

- 251 languages
- English: 1,743,312 articles
- Swedish: 222,821 articles
- Quechua: 2,102 articles
- 11 other indigenous American languages (only 4 with more than 100 articles)

Language and the Internet

Language and the Internet

- Even some communities that share a language other than English (e.g., Panjabi speakers) use English for email and chat. (Paolillo)

Language and the Internet

- Even some communities that share a language other than English (e.g., Panjabi speakers) use English for email and chat. (Paolillo)
- Even for languages such as Spanish, resources may focus on music, dancing, food, shopping (in fact may be catering to “cultural tourists”). (Clark & Gorski)

Language and the Internet

- Even some communities that share a language other than English (e.g., Panjabi speakers) use English for email and chat. (Paolillo)
- Even for languages such as Spanish, resources may focus on music, dancing, food, shopping (in fact may be catering to “cultural tourists”). (Clark & Gorski)
- The adoption of encoding standards and keyboards for some scripts lags behind.

Language and the Internet

- Even some communities that share a language other than English (e.g., Panjabi speakers) use English for email and chat. (Paolillo)
- Even for languages such as Spanish, resources may focus on music, dancing, food, shopping (in fact may be catering to “cultural tourists”). (Clark & Gorski)
- The adoption of encoding standards and keyboards for some scripts lags behind.
- Programming and markup languages are based on English. (Paolillo)

The Linguistic Digital Divide

The Linguistic Digital Divide

- Relative lack of documents and computational resources in disadvantaged languages

The Linguistic Digital Divide

- Relative lack of documents and computational resources in disadvantaged languages
- Denies revolutionary advances in access, creation, collaboration to the majority

The Linguistic Digital Divide

- Relative lack of documents and computational resources in disadvantaged languages
- Denies revolutionary advances in access, creation, collaboration to the majority
- Inhibits the participation of the majority in solving urgent national and international problems

The Linguistic Digital Divide

- Relative lack of documents and computational resources in disadvantaged languages
- Denies revolutionary advances in access, creation, collaboration to the majority
- Inhibits the participation of the majority in solving urgent national and international problems
- Exaggerates class divisions within linguistic communities

The Linguistic Digital Divide

- Relative lack of documents and computational resources in disadvantaged languages
- Denies revolutionary advances in access, creation, collaboration to the majority
- Inhibits the participation of the majority in solving urgent national and international problems
- Exaggerates class divisions within linguistic communities
- Diminishes the role of many languages

The Linguistic Digital Divide

- Relative lack of documents and computational resources in disadvantaged languages
- Denies revolutionary advances in access, creation, collaboration to the majority
- Inhibits the participation of the majority in solving urgent national and international problems
- Exaggerates class divisions within linguistic communities
- Diminishes the role of many languages
  - "Strong" international, national, and regional languages (Thai, Tamil, Amharic, Swahili, etc.)

The Linguistic Digital Divide

- Relative lack of documents and computational resources in disadvantaged languages
- Denies revolutionary advances in access, creation, collaboration to the majority
- Inhibits the participation of the majority in solving urgent national and international problems
- Exaggerates class divisions within linguistic communities
- Diminishes the role of many languages
  - "Strong" international, national, and regional languages (Thai, Tamil, Amharic, Swahili, etc.)
  - Languages already marginalized within their own countries

Causes of the LDD

Causes of the LDD

- Lack of users

Causes of the LDD

- Lack of users
- Lack of power and financial resources

Causes of the LDD

- Lack of users
- Lack of power and financial resources
- Linguistic imperialism, chauvinism

Bridging the LDD

Bridging the LDD

- Have everybody learn English.

Bridging the LDD

- Have everybody learn English.
  - It doesn't work. (Brock-Utne : *The Recolonization of the African Mind*)

Bridging the LDD

- Have everybody learn English.
  - It doesn't work. (Brock-Utne : *The Recolonization of the African Mind*)
  - It relegates all other languages to a secondary role: violates WSIS Principle 8.

Bridging the LDD

- Have everybody learn English.
  - It doesn't work. (Brock-Utne : *The Recolonization of the African Mind*)
  - It relegates all other languages to a secondary role: violates WSIS Principle 8.
- Create documents in and tools for under-represented languages, including mediating interpreters

A role for translation?

A role for translation?

- Translation and the spread of knowledge in the Middle Ages

A role for translation?

- Translation and the spread of knowledge in the Middle Ages
  - Greek to Arabic in 8th-10th century Baghdad

A role for translation?

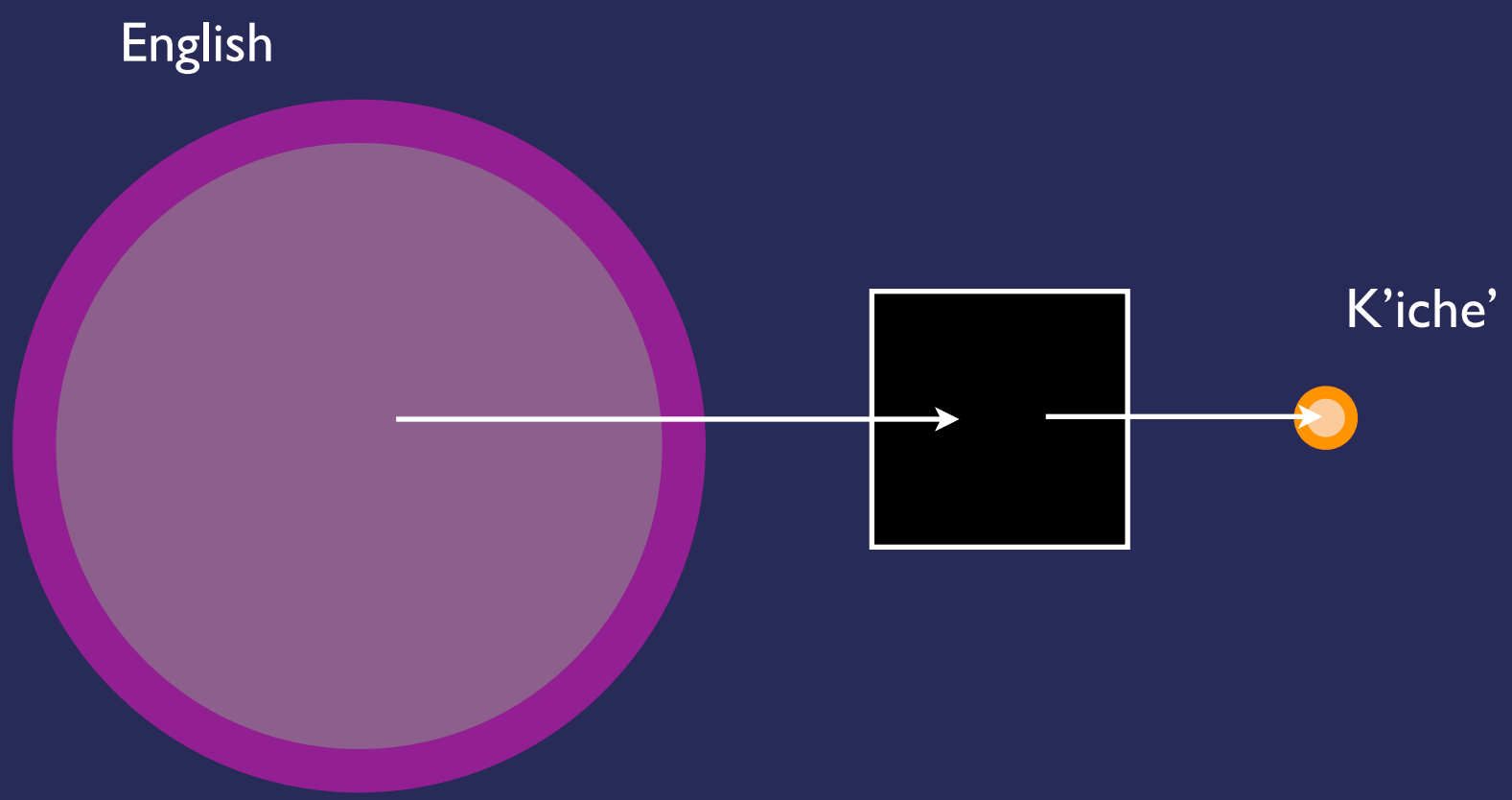
- Translation and the spread of knowledge in the Middle Ages
  - Greek to Arabic in 8th-10th century Baghdad
  - Arabic and Hebrew to Latin and Spanish in 12th-13th century Toledo

A role for translation?

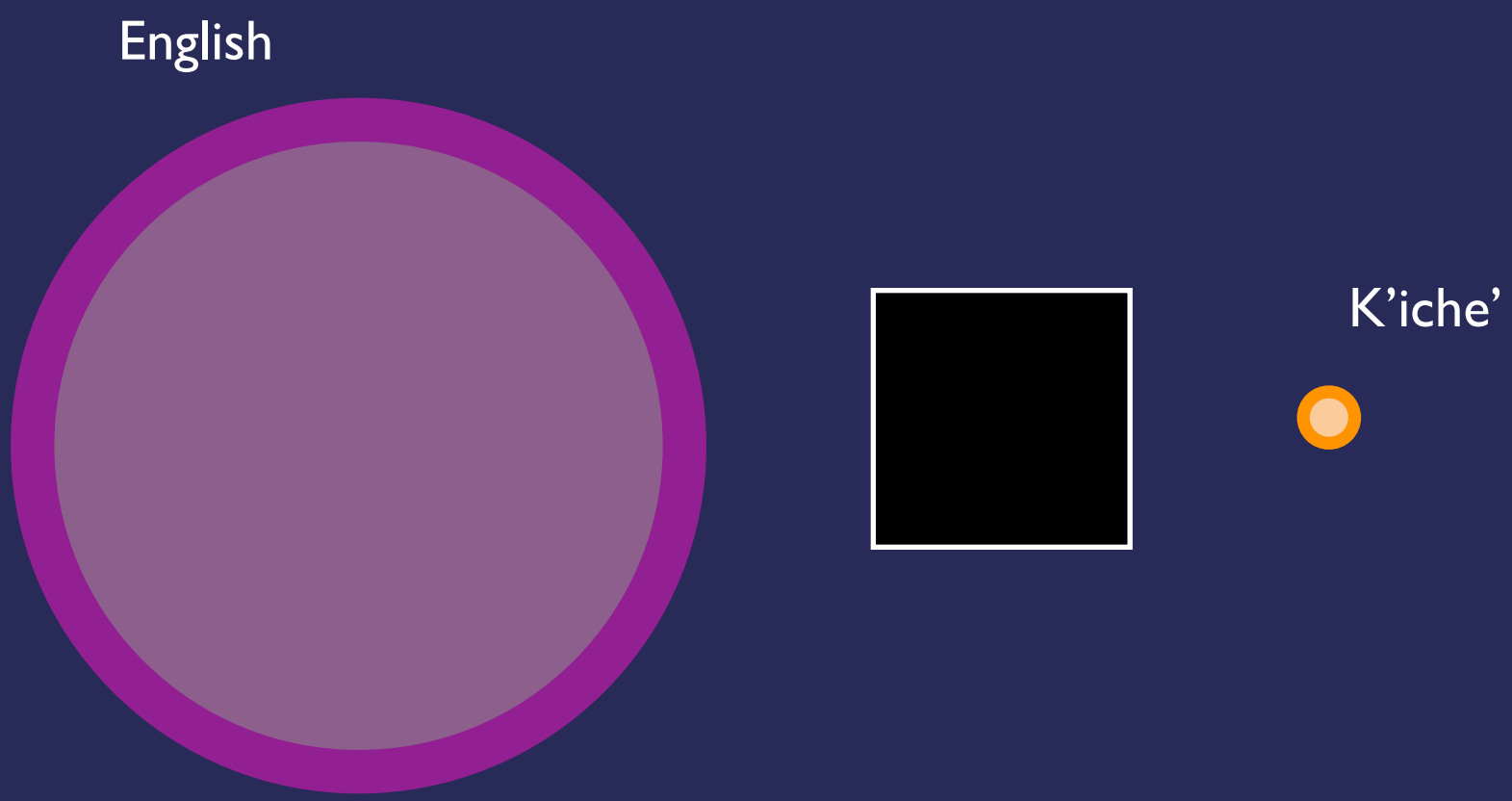
- Translation and the spread of knowledge in the Middle Ages
  - Greek to Arabic in 8th-10th century Baghdad
  - Arabic and Hebrew to Latin and Spanish in 12th-13th century Toledo
- Translation could bridge the divide by making more documents available in disadvantaged languages and giving speakers/writers of these languages a voice.

Translation and the LDD

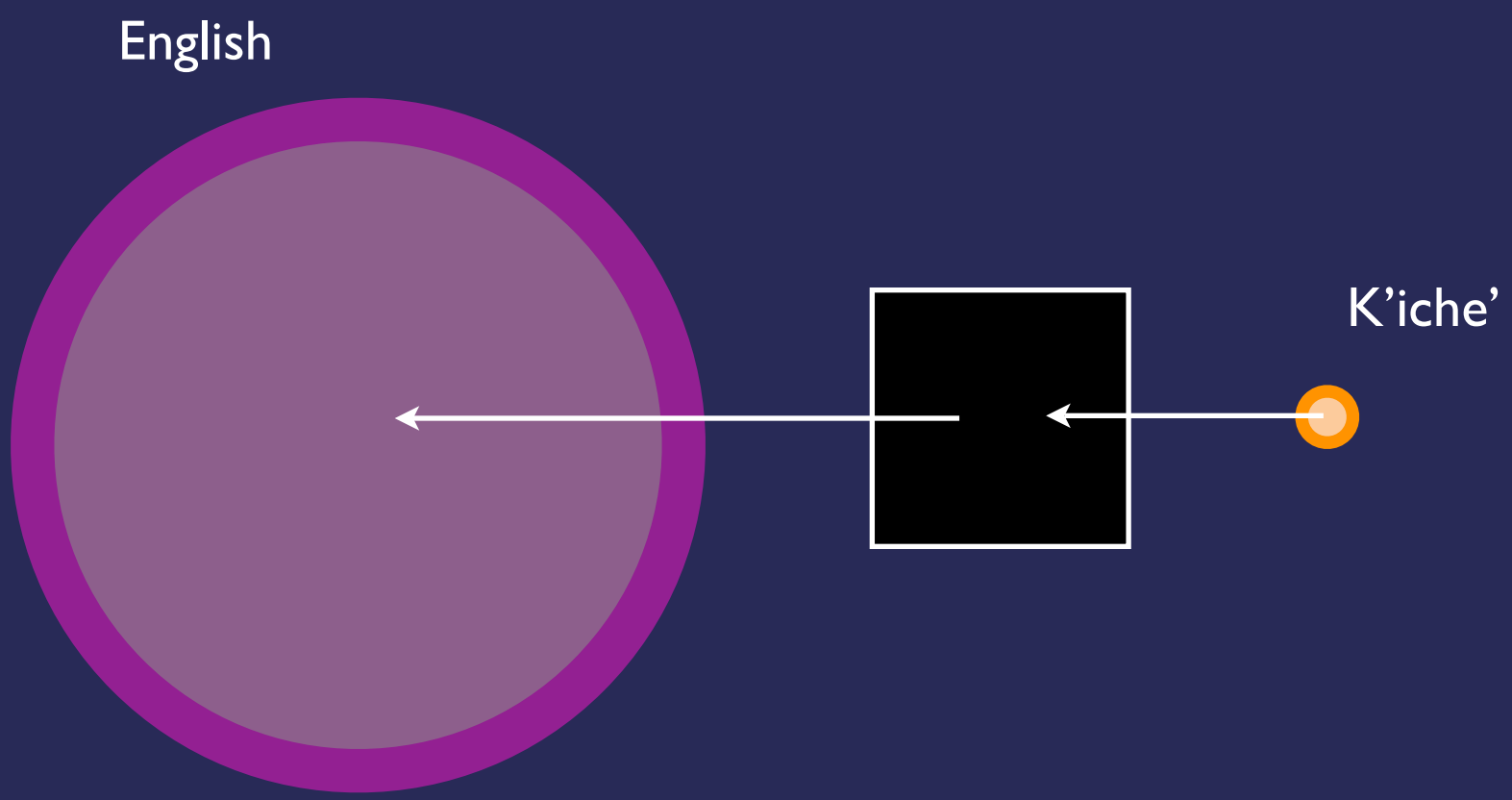
Translation and the LDD



Translation and the LDD



Translation and the LDD



Translation

Translation

- The enormity of the problem
  - Hundreds of languages
  - Millions of documents

Translation

- The enormity of the problem
  - Hundreds of languages
  - Millions of documents
- Machine translation

Overview of machine translation

Overview of machine translation

- As in other fields within computational linguistics, two classes of methods

Overview of machine translation

- As in other fields within computational linguistics, two classes of methods
  - Symbolic/linguistic

Overview of machine translation

- As in other fields within computational linguistics, two classes of methods
  - Symbolic/linguistic
    - Grammatical rules and explicit lexicons for each languages; explicit correspondence rules between the languages

Overview of machine translation

- As in other fields within computational linguistics, two classes of methods
  - Symbolic/linguistic
    - Grammatical rules and explicit lexicons for each languages; explicit correspondence rules between the languages
    - Based on grammars, dictionaries, and linguistic theories

Overview of machine translation

- As in other fields within computational linguistics, two classes of methods
  - Symbolic/linguistic
    - Grammatical rules and explicit lexicons for each languages; explicit correspondence rules between the languages
    - Based on grammars, dictionaries, and linguistic theories
  - Statistical

Overview of machine translation

- As in other fields within computational linguistics, two classes of methods
  - Symbolic/linguistic
    - Grammatical rules and explicit lexicons for each languages; explicit correspondence rules between the languages
    - Based on grammars, dictionaries, and linguistic theories
  - Statistical
    - Co-occurrence regularities between words (within and between languages) are learned.

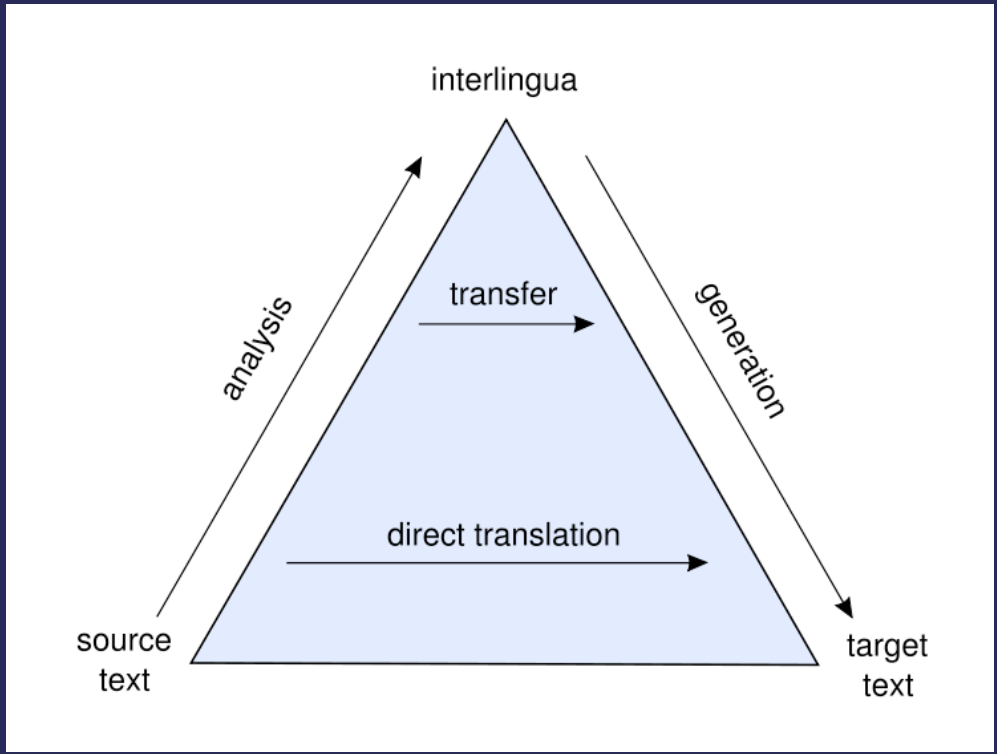
Overview of machine translation

- As in other fields within computational linguistics, two classes of methods
  - Symbolic/linguistic
    - Grammatical rules and explicit lexicons for each languages; explicit correspondence rules between the languages
    - Based on grammars, dictionaries, and linguistic theories
  - Statistical
    - Co-occurrence regularities between words (within and between languages) are learned.
    - Based on large corpora of texts (monolingual and bilingual) and on theories from machine learning (AI) and bayesian statistics

Overview of machine translation

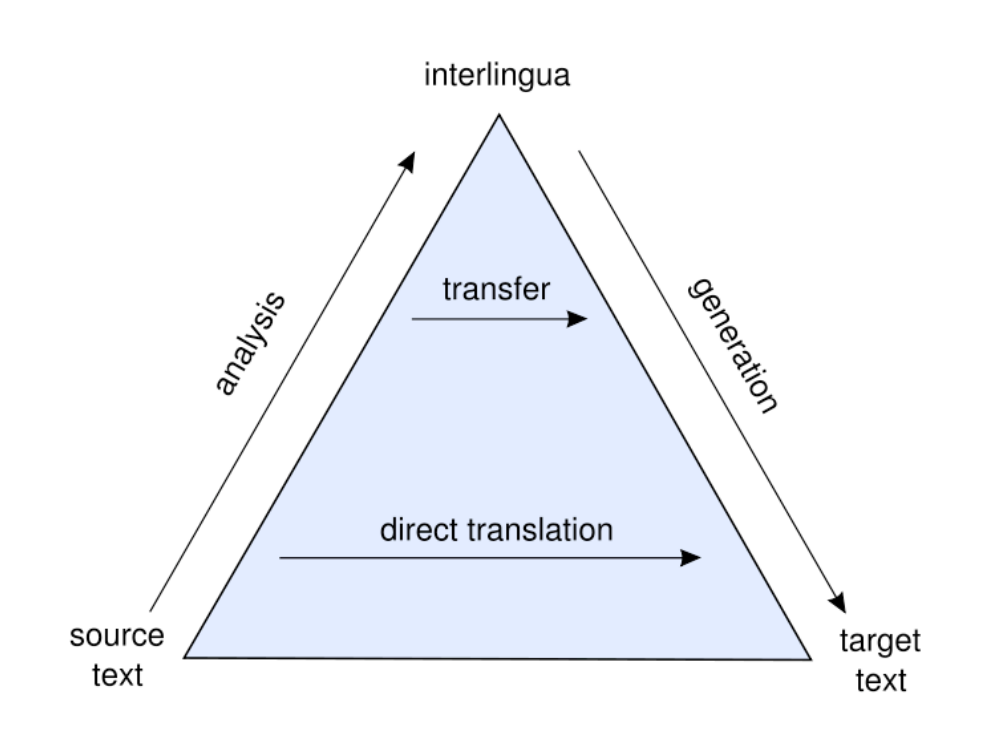
- As in other fields within computational linguistics, two classes of methods
  - Symbolic/linguistic
    - Grammatical rules and explicit lexicons for each languages; explicit correspondence rules between the languages
    - Based on grammars, dictionaries, and linguistic theories
  - Statistical
    - Co-occurrence regularities between words (within and between languages) are learned.
    - Based on large corpora of texts (monolingual and bilingual) and on theories from machine learning (AI) and bayesian statistics
- The problem of integrating the two classes of methods

Overview of machine translation



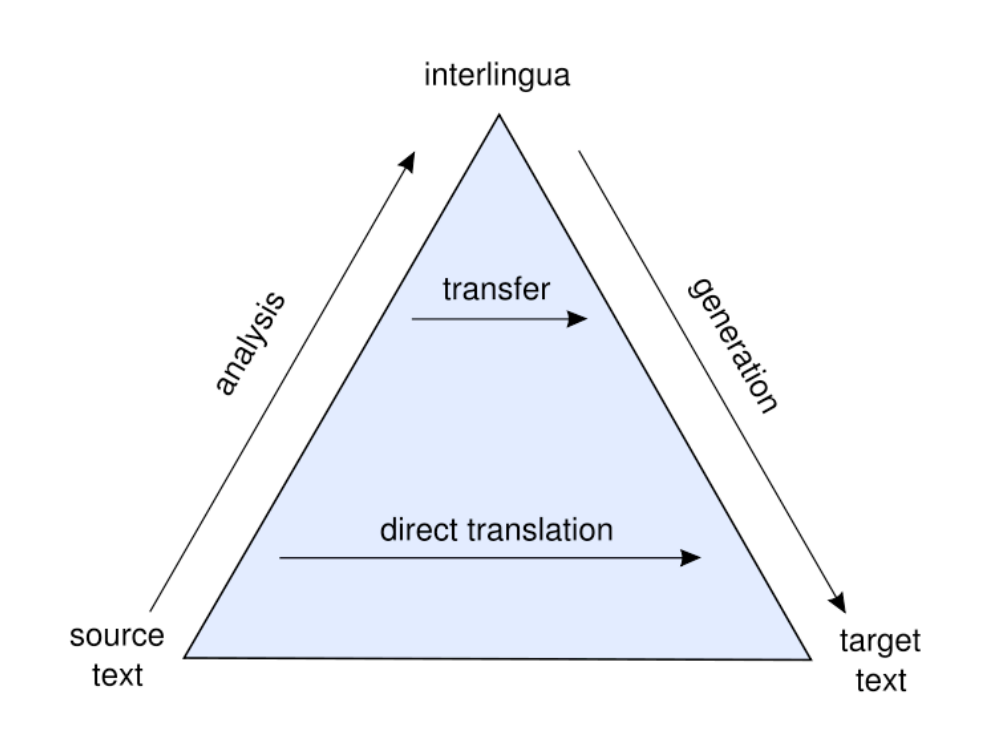
Overview of machine translation

- Transfer architectures



Overview of machine translation

- Transfer architectures
- Interlingua architectures



Machine translation

Machine translation

- Sophisticated machine translation relies on

Machine translation

- Sophisticated machine translation relies on
  - Explicit grammars and lexicons and

Machine translation

- Sophisticated machine translation relies on
  - Explicit grammars and lexicons and
  - Training the system on monolingual and bilingual texts

Machine translation

- Sophisticated machine translation relies on
  - Explicit grammars and lexicons and
  - Training the system on monolingual and bilingual texts
- Quality of translation depends on

Machine translation

- Sophisticated machine translation relies on
  - Explicit grammars and lexicons and
  - Training the system on monolingual and bilingual texts
- Quality of translation depends on
  - Distance between languages

Machine translation

- Sophisticated machine translation relies on
  - Explicit grammars and lexicons and
  - Training the system on monolingual and bilingual texts
- Quality of translation depends on
  - Distance between languages
  - Breadth of content domain

Machine translation

- Sophisticated machine translation relies on
  - Explicit grammars and lexicons and
  - Training the system on monolingual and bilingual texts
- Quality of translation depends on
  - Distance between languages
  - Breadth of content domain
- In general, human intervention is still required

Machine translation

- Sophisticated machine translation relies on
  - Explicit grammars and lexicons and
  - Training the system on monolingual and bilingual texts
- Quality of translation depends on
  - Distance between languages
  - Breadth of content domain
- In general, human intervention is still required
- Because of the background knowledge that seems essential for translation, the original goals of MT will probably never be realized.

Machine translation

- Sophisticated machine translation relies on
  - Explicit grammars and lexicons and
  - Training the system on monolingual and bilingual texts
- Quality of translation depends on
  - Distance between languages
  - Breadth of content domain
- In general, human intervention is still required
- Because of the background knowledge that seems essential for translation, the original goals of MT will probably never be realized.
- Toward appropriate and efficient forms of collaboration between people and machines (Kay)

L^3: long-term goals

L^3: long-term goals

- Translation

L^3: long-term goals

- Translation
  - Between “disadvantaged” languages (DLs) of the Global South and “privileged” languages of the Global North

L^3: long-term goals

- Translation
  - Between “disadvantaged” languages (DLs) of the Global South and “privileged” languages of the Global North
  - Among the the DLs

L^3: long-term goals

- Translation
  - Between “disadvantaged” languages (DLs) of the Global South and “privileged” languages of the Global North
  - Among the the DLs
- Computational tools to aid in teaching the DLs

L^3: long-term goals

- Translation
  - Between “disadvantaged” languages (DLs) of the Global South and “privileged” languages of the Global North
  - Among the the DLs
- Computational tools to aid in teaching the DLs
- Software to facilitate the creation of virtual communities of “experts” (teachers, writers, students, linguists) on DLs

L^3: long-term goals

- Translation
  - Between “disadvantaged” languages (DLs) of the Global South and “privileged” languages of the Global North
  - Among the the DLs
- Computational tools to aid in teaching the DLs
- Software to facilitate the creation of virtual communities of “experts” (teachers, writers, students, linguists) on DLs
  - Strengthening the language

L^3: long-term goals

- Translation
  - Between “disadvantaged” languages (DLs) of the Global South and “privileged” languages of the Global North
  - Among the the DLs
- Computational tools to aid in teaching the DLs
- Software to facilitate the creation of virtual communities of “experts” (teachers, writers, students, linguists) on DLs
  - Strengthening the language
  - Providing texts for training MT system

L^3: long-term goals

- Translation
  - Between “disadvantaged” languages (DLs) of the Global South and “privileged” languages of the Global North
  - Among the the DLs
- Computational tools to aid in teaching the DLs
- Software to facilitate the creation of virtual communities of “experts” (teachers, writers, students, linguists) on DLs
  - Strengthening the language
  - Providing texts for training MT system
  - Providing feedback for MT system

L^3

L^3

- Collaboration between

L^3

- Collaboration between
  - Computational linguists and

L^3

- Collaboration between
  - Computational linguists and
  - Members of the linguistic communities themselves who

L^3

- Collaboration between
  - Computational linguists and
  - Members of the linguistic communities themselves who
    - Define the content areas for translation and

L^3

- Collaboration between
  - Computational linguists and
  - Members of the linguistic communities themselves who
    - Define the content areas for translation and
    - Are responsible for the quality of the final translations

Machine translation and disadvantaged languages

Machine translation and disadvantaged languages

- Much research on privileged languages
  - English, French, German, Spanish, Russian, Dutch, Portuguese, Italian, Chinese, Japanese, Korean

Machine translation and disadvantaged languages

- Much research on privileged languages
  - English, French, German, Spanish, Russian, Dutch, Portuguese, Italian, Chinese, Japanese, Korean
- Some research on major languages of the Global South, in countries with significant research facilities, and on languages deemed “critical” by the US government
  - Arabic, Farsi, Hindi

Machine translation and disadvantaged languages

- Much research on privileged languages
  - English, French, German, Spanish, Russian, Dutch, Portuguese, Italian, Chinese, Japanese, Korean
- Some research on major languages of the Global South, in countries with significant research facilities, and on languages deemed “critical” by the US government
  - Arabic, Farsi, Hindi
- For the majority of languages, we only have at best dictionaries and a few other resources

The situation in Guatemala

The situation in Guatemala

- Roughly half of the population of 12,800,000 is indigenous (Mayan), speaking ~20 languages, with from 1,000,000+ to ~1,000 speakers.

The situation in Guatemala

- Roughly half of the population of 12,800,000 is indigenous (Mayan), speaking ~20 languages, with from 1,000,000+ to ~1,000 speakers.
- A significant number of Mayans are monolingual.

The situation in Guatemala

- Roughly half of the population of 12,800,000 is indigenous (Mayan), speaking ~20 languages, with from 1,000,000+ to ~1,000 speakers.
- A significant number of Mayans are monolingual.
- Most Mayans are not literate in their mother tongue; literacy in Spanish is more common.

The situation in Guatemala

- Roughly half of the population of 12,800,000 is indigenous (Mayan), speaking ~20 languages, with from 1,000,000+ to ~1,000 speakers.
- A significant number of Mayans are monolingual.
- Most Mayans are not literate in their mother tongue; literacy in Spanish is more common.
- Officially the languages are recognized and promoted; in practice there is not much actual support for this.

The situation in Guatemala

- Roughly half of the population of 12,800,000 is indigenous (Mayan), speaking ~20 languages, with from 1,000,000+ to ~1,000 speakers.
- A significant number of Mayans are monolingual.
- Most Mayans are not literate in their mother tongue; literacy in Spanish is more common.
- Officially the languages are recognized and promoted; in practice there is not much actual support for this.
- There are now some bilingual schools.

The situation in Guatemala

The situation in Guatemala

- Semi-governmental organization, the Academia de las Lenguas Mayas de Guatemala, oversees language-related issues, is involved in translation, education, production of materials, but there is little funding for its work.

The situation in Guatemala

- Semi-governmental organization, the Academia de las Lenguas Mayas de Guatemala, oversees language-related issues, is involved in translation, education, production of materials, but there is little funding for its work.
- Other independent organizations working on Mayan language issues, including Asociación Ajb'atz' Enlace Quiché, an NGO dedicated to using technology for teaching and strengthening the languages.

The situation in Guatemala

- Semi-governmental organization, the Academia de las Lenguas Mayas de Guatemala, oversees language-related issues, is involved in translation, education, production of materials, but there is little funding for its work.
- Other independent organizations working on Mayan language issues, including Asociación Ajb'atz' Enlace Quiché, an NGO dedicated to using technology for teaching and strengthening the languages.
- Small number of online resources, a few monolingual and bilingual texts, bilingual dictionaries, teaching materials.

From *Loq'aläj täq Mayab' kunab'äl* (*Mayan Medicine*)

ANIX

UPETIK UB'ANTAJIK

We q'ayes kunab'äl ri', man kk'iy ta ulöq utukel. Je wa' kpetik: ri rulewal are täq le xoral, le uxäq laj täq xerxob' uwach.

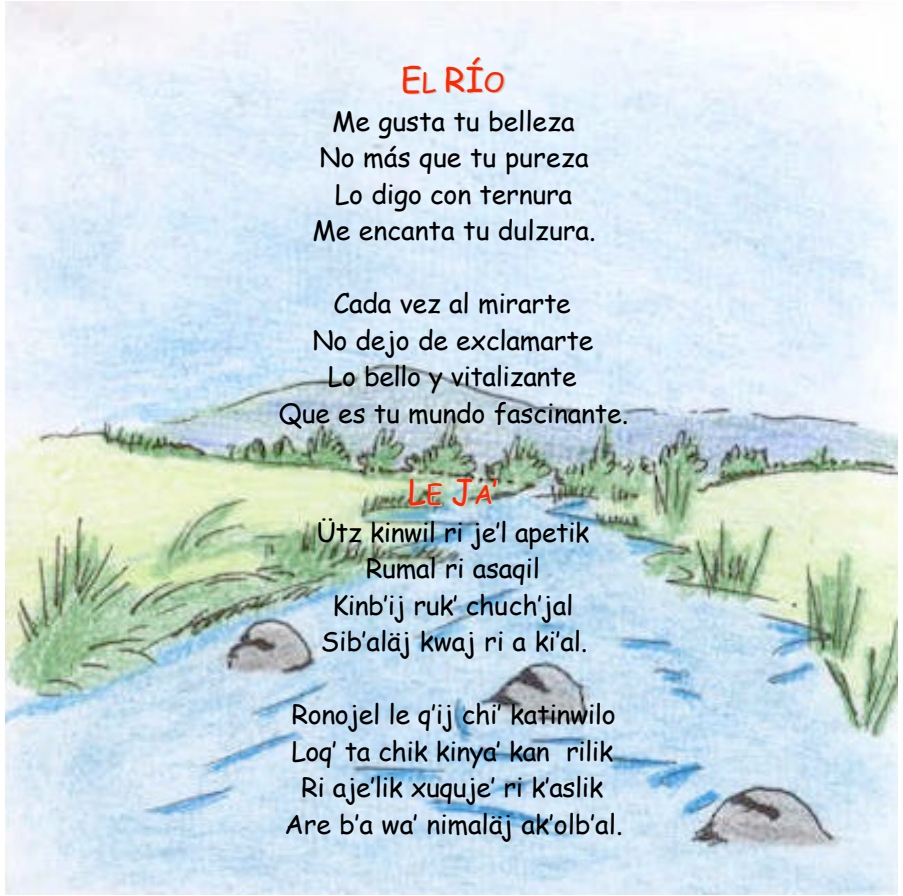
RI KUKUNAJ

- Pamaj.
- Ib'och'.
- Kuk'iysaj upam le tu',

RI UKOJIK

- **Chi rech le pamaj:** Kpoq'owisäx jun laj jub'utzaj pa jun xa'r k'a te k'u ri' ktzaq jub'iq' tzam ruk', ktijow jun qumb'äl ronojel nïm aq'ab' b'elejeb' q'ij ktijowik.
- **Chi rech le ib'och':** Kpoq'owisäx jun laj jub'utzaj pa jun xa'r, ktijow xäq pa chi kech kajb'äl, jun qumb'äl ktijowik ronojel q'ij, lajuj q'ij ktaqexïk.
- **Chi rech upam le tu':** Ktijowik are chi' ke'l ulöq ri ixöq pa tuj, kutïj jun qumb'äl chi rech pa waq'ib' q'ij.

From *Poemas infantiles*



The situation in Guatemala

The situation in Guatemala

- Internet cafes fairly common, home computers not

The situation in Guatemala

- Internet cafes fairly common, home computers not
- Cellphones everywhere

The situation in Guatemala

- Internet cafes fairly common, home computers not
- Cellphones everywhere
- In order to benefit from the Digital Revolution, Mayan language communities need

The situation in Guatemala

- Internet cafes fairly common, home computers not
- Cellphones everywhere
- In order to benefit from the Digital Revolution, Mayan language communities need
  - Access to technology

The situation in Guatemala

- Internet cafes fairly common, home computers not
- Cellphones everywhere
- In order to benefit from the Digital Revolution, Mayan language communities need
  - Access to technology
  - Literacy in their mother tongues

The situation in Guatemala

- Internet cafes fairly common, home computers not
- Cellphones everywhere
- In order to benefit from the Digital Revolution, Mayan language communities need
  - Access to technology
  - Literacy in their mother tongues
  - Many more documents online

The situation in Guatemala

- Internet cafes fairly common, home computers not
- Cellphones everywhere
- In order to benefit from the Digital Revolution, Mayan language communities need
  - Access to technology
  - Literacy in their mother tongues
  - Many more documents online
  - Ways to interact with speakers of other languages

L^3 and Mayan languages: short-term goals

L^3 and Mayan languages: short-term goals

- Morphological parsers and generators for four largest languages (K'iche', Kaqchikel, Mam, Q'eqchi')

L^3 and Mayan languages: short-term goals

- Morphological parsers and generators for four largest languages (K'iche', Kaqchikel, Mam, Q'eqchi')
- Translation of simple sentences in a narrow content domain among these languages.

Thank you!
¡Maltyox!
¡Matyox!
¡Chjonta!