

# Morphology <sup>1</sup>

---

- Introduction
- Morphology
- Morphological Analysis (MA)
- Using FS techniques in MA
- Automatic learning of the morphology of a language

# Morphology <sup>2</sup>

---

- Morphology
  - Structure of a word as a composition of morphemes
  - Related to word formation rules
  - Functions
    - Inflection
    - Derivation
    - Composition
- Result of morphologic analysis
  - Morphosyntactic categorization (POS)
    - e.g. Parole tagset (VMIP1S0), more than 150 categories for Spanish
    - e.g. Penn Treebank tagset (VBD), about 30 categories for English
  - Morphological features
    - Number, case, gender, lexical functions

# Morphology <sup>3</sup>

---

- Morphologic analysis
  - Decompose a word into a concatenation of morphemes
  - Usually some of the morphemes contain the meaning
    - One (root or stem) in flexion and derivation
    - More than one in composition
  - The other (affixes) provide morphological features
- Problems
  - Phonological alterations in morpheme concatenation
  - Morphotactics
    - Which morphemes can be concatenated with which others

# Morphology <sup>4</sup>

---

- Problems
- Affixes
  - Suffixes, prefixes, infixes, interfixes
- Inflectional affixes ≠ derivational affixes
- Derivation implies sometimes a semantic change not always predictable
  - Meaning extensions
  - Lexical rules
- A derivational suffix can be followed by an inflectional one
  - love ⇒ lover ⇒ lovers
- Inflection does not change POS, sometimes derivation does
- Inflection affects other words in the sentence
  - agreement

# Morphology <sup>5</sup>

---

- **Morphotactics**
  - Word formation rules
  - Valid combinations between morphemes
    - Simple concatenation
    - Complex models root/pattern
    - Language dependency regularity
- **Phonological alterations (Morphophonology)**
  - Changes when concatenating morphemes
  - Source: Phonology, morphology, orthography
  - variable in number and complexity
  - e.g. vocalic harmony

# Morphology <sub>6</sub>

---

## Morphemes

- 1 morpheme:  
Evitar ( verb to avoid)
- 2 morphemes:
  - evitable = evitar + able (adj: can be avoided)
- 3 morphemes:
  - inevitable = in + evitar + able  
(adj: cannot be avoided)
- 4 morphemes:
  - inevitabilidad = in + evitar + able + idad  
(noun: cannot be avoided)

# Morphology 7

---

## Inflectional Morphology

- number
- house houses
- cheval chevaux
- casa casas
- verbal form
  - walk walks walked walking
  - amo amas aman ...
- gender
  - niño niña

# Morphology <sup>8</sup>

---

## Derivational Morphology

- **Form**
- Without change      barcelonés
- Prefix      inevitable
- Suffix      importantísimo
- **Source**
- verb => adjective      tardar      => tardío
- verb => noun      sufrir      => sufrimiento
- noun => noun      actor      => actorazo
- noun => adjective      atleta      => atlético
- adjective => adjective      rojo      => rojizo
- adjective => adverb      alegre      => alegremente



# Morphological Analysis 1

---

## Types of morphological analyzers

### Formaries

- Dictionaries of word forms
- + efficiency
- + Languages with few variants (e.g. English)
- + extensibility
- + Possibility of building and maintenance from a morphological generator
- Languages with high flexive variation
- derivation, composition
- FS techniques
- FSA
  - 1 level analyzers
- FST
  - > 1 level analyzers

# Morphological Analysis <sup>2</sup>

---

## Morphological analyzers of two levels

- General model for languages with morpheme concatenation
- Independence between lingware and analyzer
- Valid for analysis and generation
- Distinction between lexical and superficial levels
- Parallel rules for morphophonology
- Simple implementation

# Morphological Analysis <sup>3</sup>

---

- Morphological rules
  - Define the relations between characters (surface) and morphemes and map strings of characters and the morphemic structure of the word.
- Spelling rules
  - Perform at the level of the letters forming the word. Can be used to define the valid phonological alterations.
- Ritchie, Pulman, Black, Russell, 1987

# Morphological Analysis <sup>4</sup>

---

- input:
- form
- output
- lemma + morphological features

Input	Output
cat	cat + N + sg
cats	cat + N + pl
cities	city + N + pl
merging	merge + V + pres_part
caught	(catch + V + past) or (catch + V + past_part)