

- Université de Bouira

Traitement automatique des langues

M1 ISIL, semestre 1

Taha Zerrouki

Programme

- **1. Introduction**
- **2. Les niveaux de traitement**
 - **Codage des caractères**
- **3. Le niveau lexical**
- **4. Le niveau syntaxique**
- **5. Le niveau sémantique**
- **6. Le niveau pragmatique**
- **7. Les applications du TALN**
-

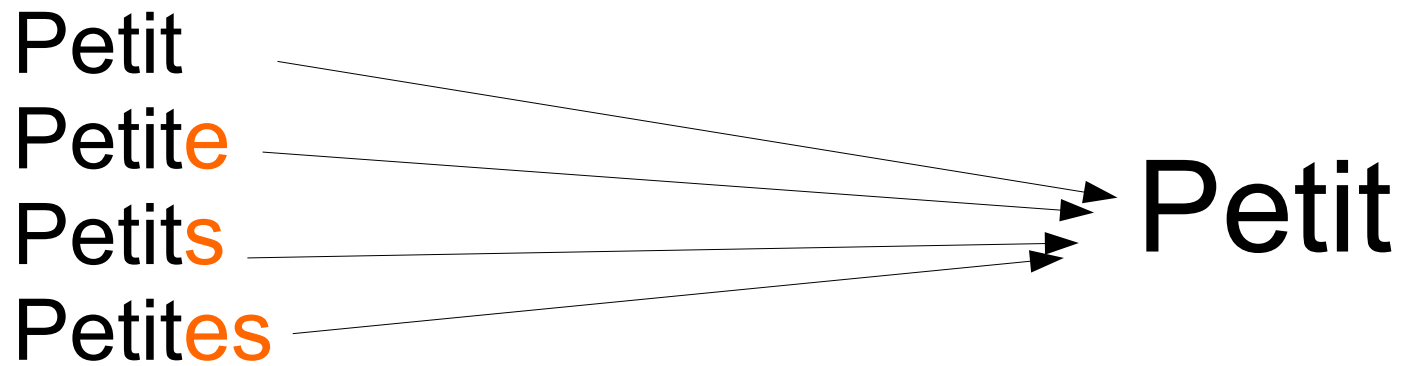
Définition

Lemmatisation

- La **lemmatisation** désigne l'**analyse lexicale** du contenu d'un texte regroupant les **mots** d'une **même famille**.

Exemple

Forme canonique



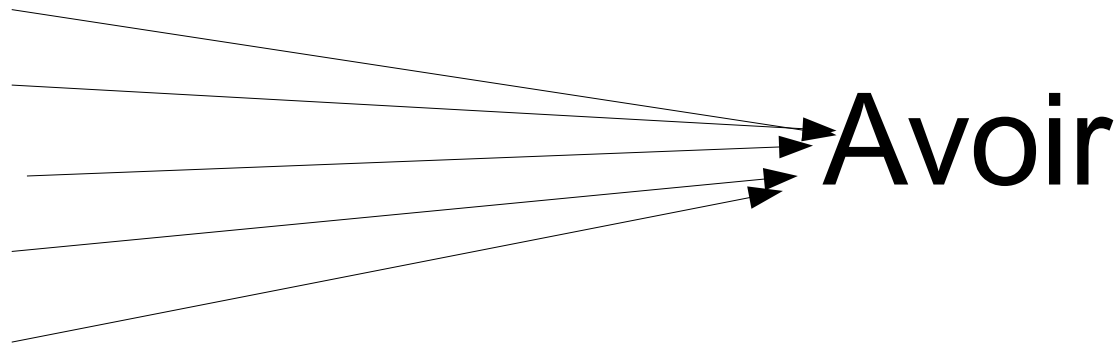
Lemmatisation

- Chacun des mots d'un contenu se trouve ainsi **réduit** en une entité appelée **lemme** (forme canonique).
- La lemmatisation **regroupe les différentes formes** que peut revêtir un mot, soit : le nom, le pluriel, le verbe à l'infinitif, etc

Exemple

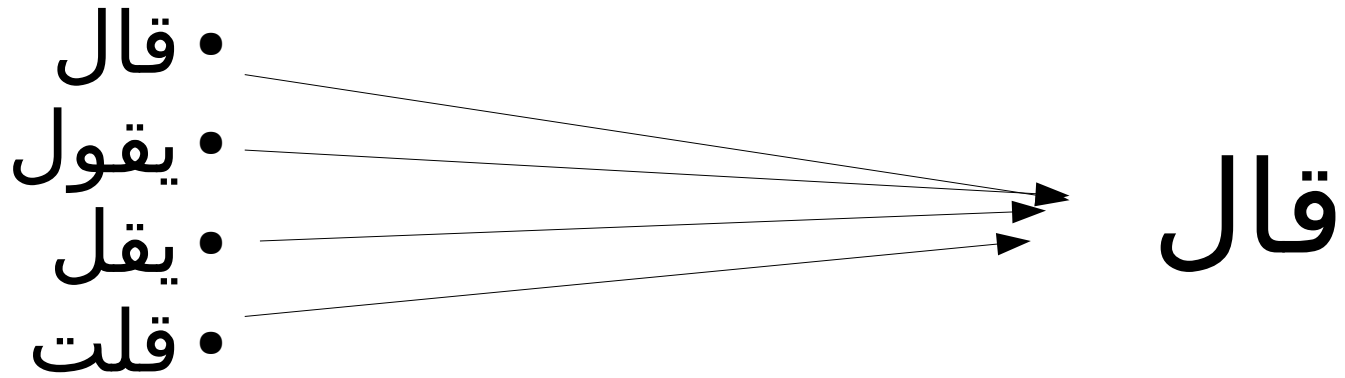
Forme canonique

- ai
- as
- a
- avons
- avions



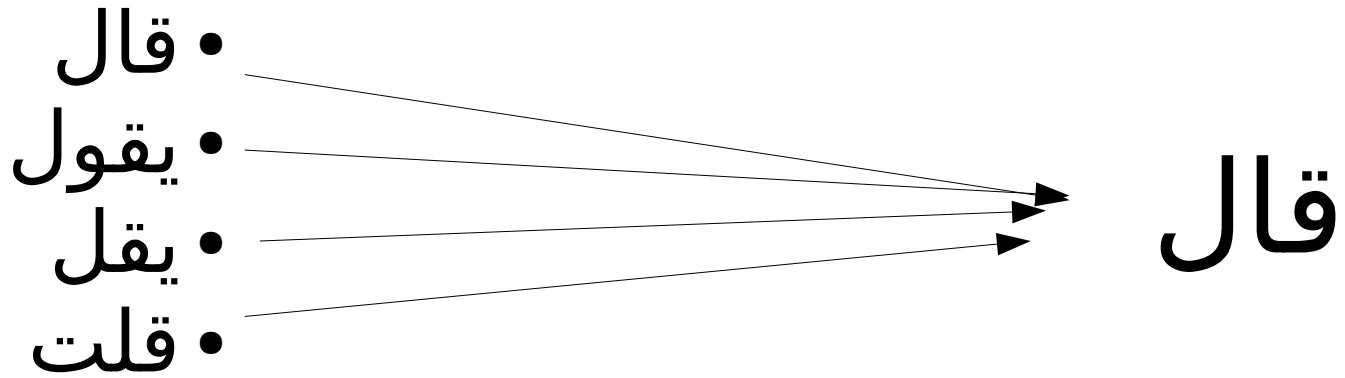
Exemple

Forme canonique



Exemple

Forme canonique



Lemmes

- Les mots (**lemmes**) d'une langue utilisent plusieurs formes en fonction de :
 - leur genre (masculin ou féminin),
 - leur nombre (un ou plusieurs),
 - leur personne (moi, toi, eux...),
 - leur mode (indicatif, impératif...)
 - donnant ainsi naissance à plusieurs formes pour un même lemme.

Exercice

- Donner les lemmes des mots suivants
- مَدَارِسُ
- عَنَادِل
- إِيمَارِيغِن
- fut (français)
- better (english)

La morphologie الصرف

- La morphologie est la branche de la linguistique qui étudie les **types et la forme** des mots.

الصرف La morphologie

- Il existe deux types de la morphologie
- La morphologie flexionnelle (en interne)
- La morphologie dérivationnelle (en externe)

• علم الصرف

- التصريف

- الاشتقاق

La morphologie الصرف

- Il existe deux types de la morphologie
- La morphologie flexionnelle (en interne)
 - enfant => enfants
 - Petit => petite
- La morphologie dérivationnelle (en externe) :
 - Jouer => joueur

الصرف La morphologie

• التصريف

- أكل = < يأكل، أكلت أكلنا، يؤكل

- طالب = < طالبة، طلبة

• الاشتقاق

- طلب = < طالب، مطلوب

La morphologie flexionnelle

التصريف

- En morphologie, on nomme flexion l'ensemble des modifications subies aux mots d'une langue flexionnelle pour dénoter les traits grammaticaux voulus.
- À la différence de la dérivation, la flexion **ne crée pas de nouveaux mots.**

La morphologie flexionnelle

التصريف

- Il existe deux grandes catégories de flexions:
- La déclinaison pour le système nominal
 - Les noms y changent généralement de forme selon le **genre**, le **cas** ou le **nombre**, etc;
- la conjugaison pour les verbes,
 - qui varient généralement en personne, nombre, genre, temps, voix et mode, etc.

La morphologie flexionnelle

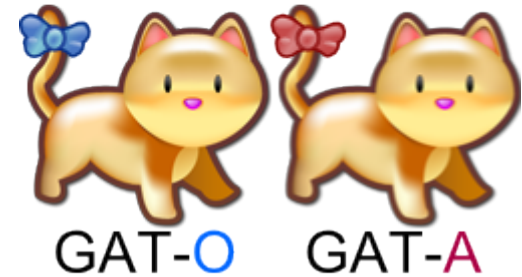
التصريف

La déclinaison pour le système nominal

- Les noms y changent généralement de forme selon le **genre**, le **cas** ou le **nombre**, etc;

Exemple de déclinaison

La déclinaison de
mot chat en
espagnole



Exercice

Donner les déclinaisons possible des mots

- Étudiant

- طالب

- Anelmad انلماذ

La morphologie flexionnelle

التصريف

- la conjugaison pour les verbes,
 - qui varient généralement en personne, nombre, genre, temps, voix et mode, etc.

Exemple

الماضي

أنا كتبتُ

أنت كتبتِ

أنت كتبتِ

هو كتبَ

هي كتبتُ

المضارع

أنا أكتبُ

أنت تكتبُ

أنت تكتبين

هو يكتب

هي تكتب

Morphologie flexionnelle

- La génération des nouvelles forme se fait par les méthodes suivantes
 - L'affixation: الزيادة
 - l'ajout des préfixes et des suffixes
 - La duplication التضعيف
 - Dupliquer le mot ou bien une partie de mot
 - Altération الإبدال
 - Changer une lettre ou bien une partie du mot
 - Variation super segmental النطق
 - Changer l'intonation du mot

Morphologie flexionnelle

- L'affixation:
 - l'ajout des préfixes et des suffixes

ويأكلونها

Enclitique + suffixe + stem + préfixe + proclitique

لاحقة ثانية + لاحقة + stem + سابقة + سابقة أولية

Morphologie flexionnelle

- Le terme "affixe" regroupe un ensemble de cinq éléments :
- Préfixes, placés avant leur base
- Infixes, placés à l'intérieur d'une base, ou entre une base et un affixe d'un terme déjà lexicalisé
- Suffixes, placés après la base

•
وياً كلونها

Morphologie flexionnelle

- La duplication التضعيف
 - Dupliquer le mot ou bien une partie de mot
 - زلزل
 - فعّل
 - كيف كيف (Dardja)
 - Ping-pong
 - Bye-bye
 - zigzag

Morphologie flexionnelle

- Altération
 - Changer une lettre ou bien une partie du mot

يَقُولُ يُقَالُ قَالٌ

قُلْتُ قَالَتْ قَالٌ

Morphologie flexionnelle

- Altération الإبدال
 - Changer une lettre ou bien une partie du mot

Franc **c** => Fran**che**

Adjectif **f** => Adjecti**ve**

Morphologie flexionnelle

- Variation super segmental
 - Changer l'intonation du mot

تعديل النطق

Exercice

- Donner des exemples pour chaque méthode
 - L'affixation:
 - La duplication
 - Altération
 - Variation super segmental

Exercice

- Décrire les méthodes de flexion pour les mots suivants :

- رجل argaz => irgazen

- جبل adhrar => idhourar

- عناكب

- لم يَعُدْنَ

-

La morphologie الصرف

- Il existe deux types de la morphologie
- La morphologie flexionnelle (en interne)
 - enfant => enfant**s**
 - Petit => petite**e**
- La morphologie dérivationnelle (en externe) :
 - Jouer => joue**ur**

La morphologie dérivationnelle

- La morphologie dérivationnelle (الاشتقاق en externe) :

- Jouer => joueur

- La dérivation lexicale, ou encore dérivation, est un des procédés de formation des mots, au même titre que le néologisme ou l'emprunt.

- الاشتقاق هو أحد طرق تكوين الكلمات

La morphologie dérivationnelle

الاشتقاق

- Les méthodes de dérivation
 - L'affixation
 - Les templates
 -

La morphologie dérivationnelle

الزيادة

- L'affixation
 - Refaire
 - Final

طبيعية •

مدرسة •

استعمل •

آقشيش •

تقشيشت •

La morphologie dérivationnelle

الميزان

- Les templates : est une méthode non contaténative
- Basé sur les patrons

علم

عالم

معلومة

استعلام

إعلام

Méthode de formation des mots

- Composition
- Troncation
- Mot-valise
- Acronyme

Formation des mots

- Composition المركب المَرْجِي
- La composition permet de jumeler deux items lexicaux pour en former un nouveau.
- Le sens de ce nouvel item peut porter sur l'un des items de sa fabrication
 - Plate-forme
 - TimeOut
 - حَضْرَ مَوْت
 - بوجمعة

Formation des mots

- Troncation الترخيم
- La troncation permet de raccourcir un item lexical.
- Les éléments tronqués peuvent être laissés de la sorte, être suffixés ou dupliqués (cas des hypocoristiques).

Biblio

Restau

يا صاح

شمسو

موح

Formation des mots

- Mot-valise النحت
- Procédé qui consiste à coller le début d'un mot à la fin d'un autre pour créer un nouveau mot, souvent à but humoristique ou satyrique

– franglais : français + anglais ;

– transistor : transfer + resistor ;

بِسْمِ اللّٰهِ : بِسْمَلَة

لا حول ولا قوة إلا بالله : حوقلة

Formation de mot

- Acronyme :

الاختصار

- Un acronyme est un sigle passé dans le langage courant : ovni, radar, sida, ...
- **S**yndrome d'**I**mmunod**d**éficiency **a**cquise
Le radar : **R**Adio **D**etection **A**nd **R**anging)

