

Rule-based Approaches to the Automatic Generation of Puns Based on Given Names in French

Mathieu DEHOUCK¹ Marine DELABORDE^{1,2}

¹Lattice, CNRS, ENS-PSL, USN

²LT2D, CY Cergy Paris University

CHum, COLING - January 19, 2025

Plan

- ▶ Why Rules in the Age of AI?
- ▶ Current Method (v1)
- ▶ Current Limitations
- ▶ Work in progress (v2)
- ▶ Challenges
- ▶ Perspectives

Introductory Example

- (1) Monsieur et Madame Gator ont un fils. Comment s'appelle-il ?
Mr. and Mrs. Gator have a son. What is his name?

Ali, car, Ali Gator (alligator).

[a.li.ga.tɔʁ]

Ali, because, Ali Gator (alligator).

Why Rules in the Age of AI?

- ▶ Humans: Orality is the main language modality, written text is secondary;
- ▶ Language Models: Writing is the primary modality, and tokens rarely correspond to actual phonological or prosodic units.

Chat GPT 4 (French version) - December 19, 2024

Raconte-moi une blague Monsieur Madame



Avec plaisir ! 😊

Monsieur et Madame Propre ont un fils.
Comment s'appelle-t-il ?

Jean ! (Parce que... Jean Propre 😊)

< 2/2 > 🔊 📄 👍 🗨️ ↻

Pourquoi est-ce drôle

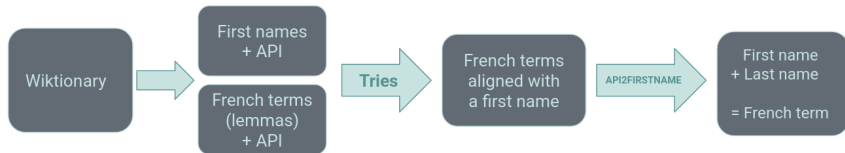


C'est drôle parce que c'est un jeu de mots basé sur l'expression "Jean Propre", qui ressemble à "gens propres" en français, mais aussi parce que **Monsieur Propre** est une figure célèbre des produits de nettoyage. Le fils "Jean" devient un clin d'œil humoristique à ce double sens : un prénom banal qui fait écho au produit ménager.

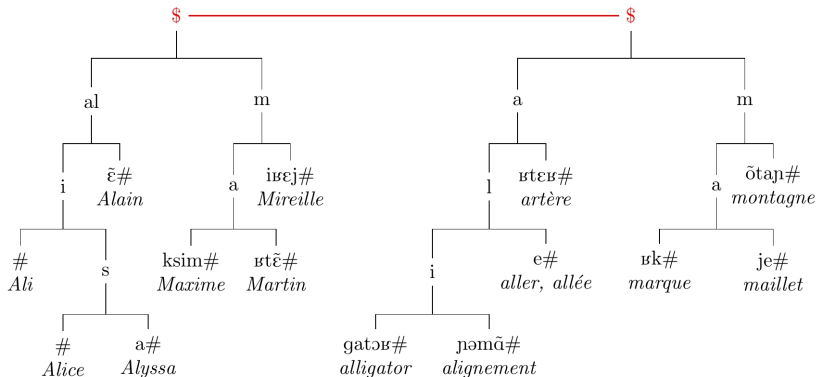
L'humour vient du mélange d'une situation absurde et d'une référence connue ! 😊

🔊 📄 👍 🗨️ ↻

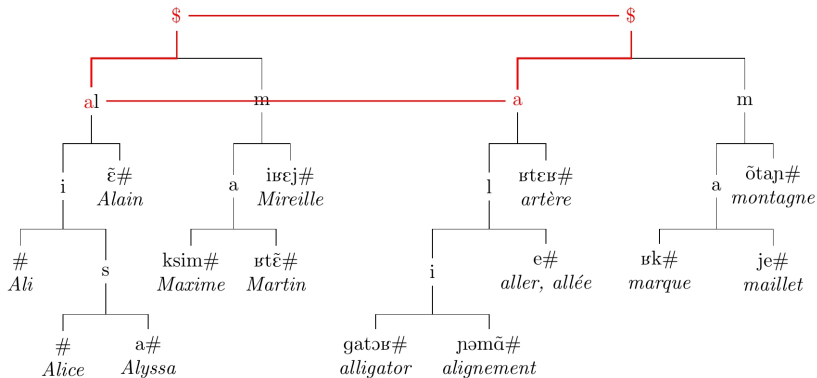
Current Method (v1)



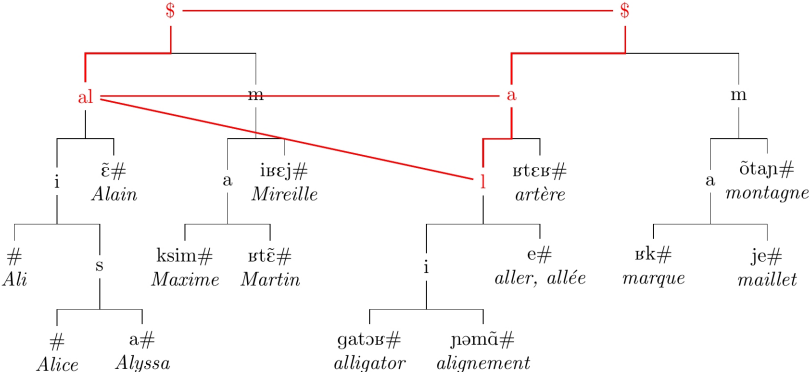
Trie Alignment



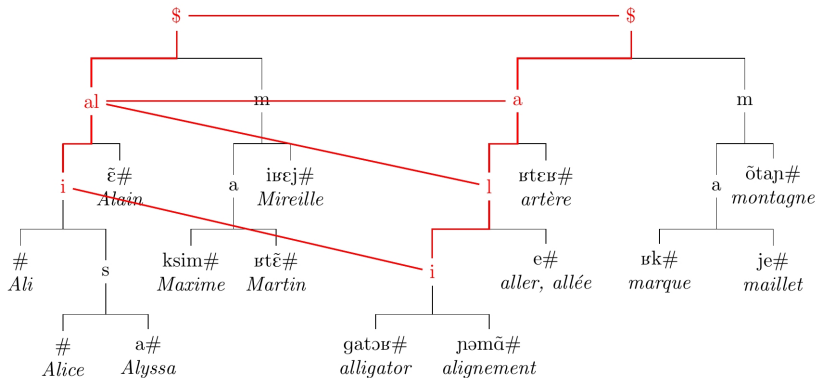
Trie Alignment



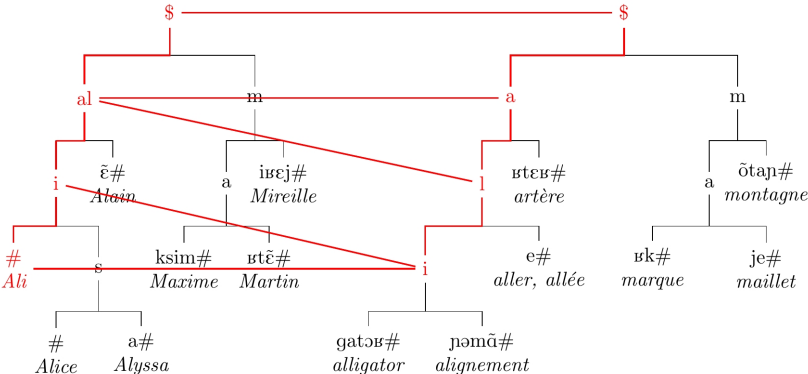
Trie Alignment



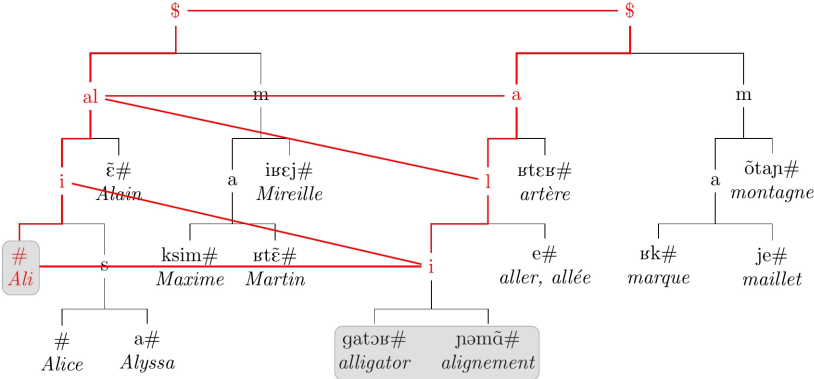
Trie Alignment



Trie Alignment



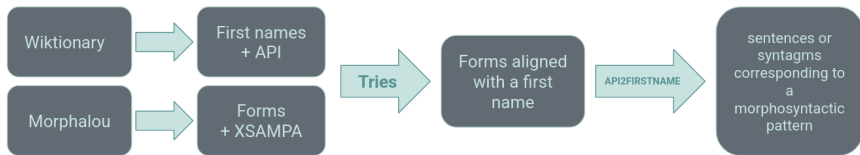
Trie Alignment



Current Limitations

- ▶ Finite vocabulary: Wiktionary French terms and French first names;
- ▶ No control over meaning: Wiktionary French terms;
- ▶ No control over structure: a first name may or may not lead to a pun;
- ▶ No compositionality: only one first name.

Work in progress (v2)



Work in progress (v2)

- ▶ Syntactic structure: Morphalou's entries are morphologically analyzed → chains of morphologically informed tries; Extract syntactic phrases from Universal Dependencies treebanks;
 - 1 DET Gender=G0|Number=N0
 - 2 NOUN Gender=G0|Number=N0
 - 3 VERB Mood=Ind|Number=N0|Person=3|Tense=T0
- ▶ Relaxation of phonological constraints: phonologically constrained mid vowel pairs [e/ɛ], [o/ɔ], [ø/œ].

Semantic consistency

- (2) Monsieur et Madame Golaux ont deux enfants. Comment s'appellent-ils ?
Mr. and Mrs. Golaux have two children. What are their names?
- Laura et Valérie, car, Laura Valérie Golaux (l'or avalé rigolo).
[lo.va va.le.ʁi go.lo] [l ɔʁ a.va.le ʁi.go.lo].
- Laura and Valérie, because, Laura Valérie Golaux (the swallowed funny gold).

Syntactically sound, but not funny...

More on Pronunciation Challenges

- ▶ French has a lot of sandhis (liaison) that are not trivial;
- ▶ Some jokes work better (or at all) with some accent, either actual or deliberately prototypical. Not only could it be offensive, it is really hard to get well.

Perspectives

- ▶ Using LLM to ensure some form of semantic coherence;
- ▶ Beside user annotation, do an actual annotation campaign to assess the degree of humor in puns;
- ▶ Application outside of purely humorous settings: didactic, mnemonics...