

Mind your p's and q's, or the peregrinations of apostrophes in 17th Century English

Odile Piton NLP Laboratoire Marin Mersenne

University of Paris I

Hélène Pignot Senior lecturer in English,
Laboratoire Marin Mersenne University of Paris I

Outline

- **Introduction**
- **Brief historical review: Italy, France, Great Britain**
- **Apostrophes in 17th Century English**
 - Analysis and transcription
 - Use of morphological and syntactical graphs
 - Discussion about the methodology
- **Conclusion**

Introduction

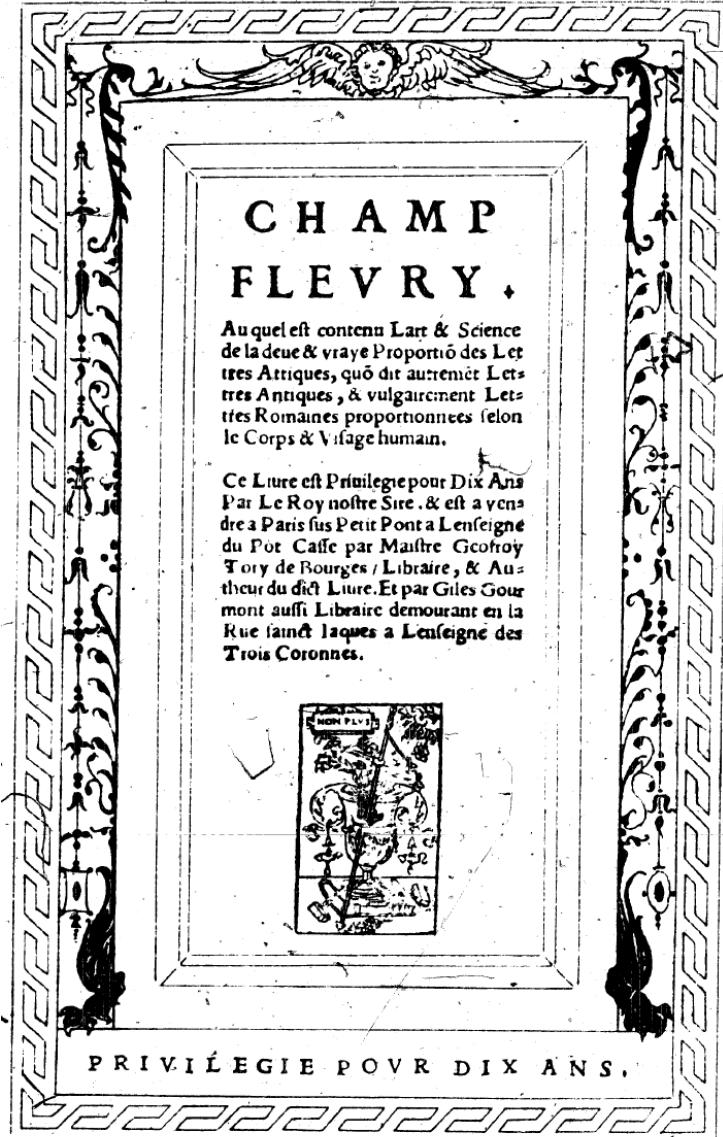
- Appearance of apostrophes in printed texts
 - A brief historical review by Jacques André « Funeste destinée : l'apostrophe détournée » published in *Graphê* in March 2008
 - An internal punctuation mark for verbs or nouns
- a guiding principle to introduce the main transformations of nouns and verbs

Italy 1501



- A punctuation mark that used to be **handwritten**
- **Aldo Manuzio** (Venise) introduces it in the first published Italian grammar *Regole della lingua fiorentina*.
- In Italian it marks the elision of a letter

**Tu nedra' Italia et l'honorata rind
Canzon; ch'a gliocchi miei celd et contend
Non mar, non poggio, o fiume;**



France 1529

- Geofroy Tory, in *Champ Fleury* (1529), a treatise about typography, was the first to use « **ung point crochu quon appelle Apostrophus** »

cōme iay dict, vng point crochu quon appelle Apostrophus. Cōme en disant Viden' Satin' Vui' pour & en Lieu de dire, Videsne, Satisfue, & Vifue.

France 1531 - 1533

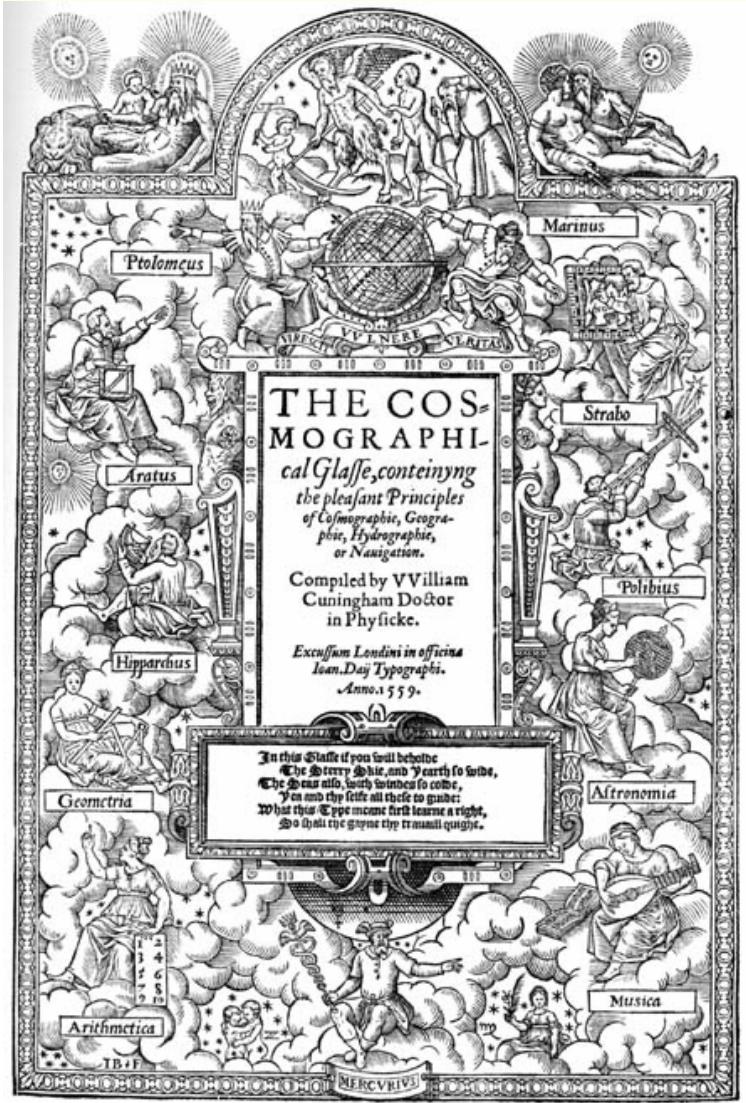
- Jacques Dubois, in his grammar *Isagōge* recommends the use of the Greek apostrophe to mark the omission of a letter.
m'amic' vt, pro ma amic'.

tu n'es qu'un badin: pro tu n'e cs qu'e vn badin.

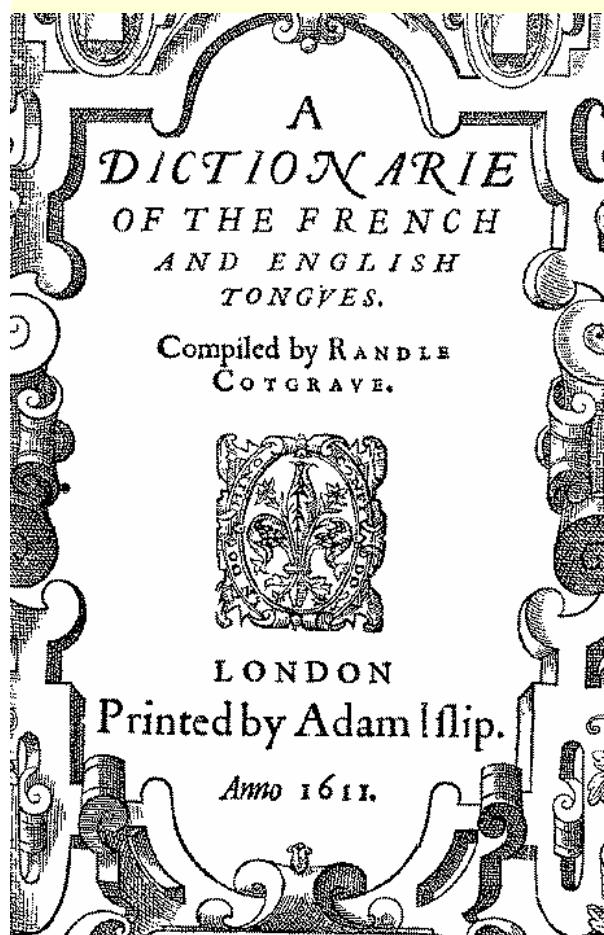
- 1533 first font made by Augereau
- In the meantime it appears in other books
- Its use becomes widespread, and in French its insertion becomes obligatory.

le ami becomes *l'ami*

Great Britain 1559



- It appears in *The Cosmographical Glasse* by Williams Cunningham.
- It is used mainly in contractions
 - don't for do not
- and as a genitive
 - John's dog, is the contraction of the archaic genitive Johnes dog



Definition 1611

- *A Dictionarie of the French and English Tongves*

Compiled by **Randle Cotgrave**

To apostrophise =

To cut off the last vowel of a word

Apostropher. To auert, wry, turne aside; also, to revoke, or call backe; also, to apostrophise; to cut off (by an Apostrophe) the last vowel of a word.

Comparison

Language	French	Italian	English	German	Spanish
Number of apostrophes	10 000	6 000	500	some	some
Number of sentences	9 000	9 000	9 000	9 000	9 000

Apostrophes and computers

- Several codes : 39 (not respected by Anglo-Saxons), then 169, Unicode U+2018 or U+2019
 - NooJ 2018 and 2019 but impossible to type both with the keyboard
- Be careful, you might lose some results

Apostrophes in our corpus

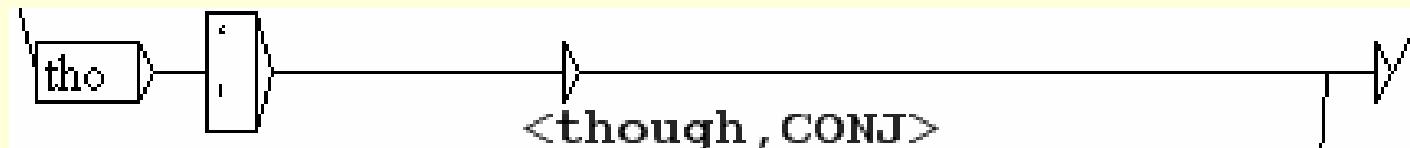
- **Corpus**
 - Accounts of English travellers in Greece and the Ottoman Empire ([NooJ 2008](#))
- **Apostrophes in our corpus**
 - Poetry: not to be modified
 - Verbs: forms in 't, 'd
 - Nouns: recognise 's, introduce it --if possible
 - Other cases : contraction, compounds, elision
 - Adjectives like hair'd: not to be modified

contraction, compounds, elisions

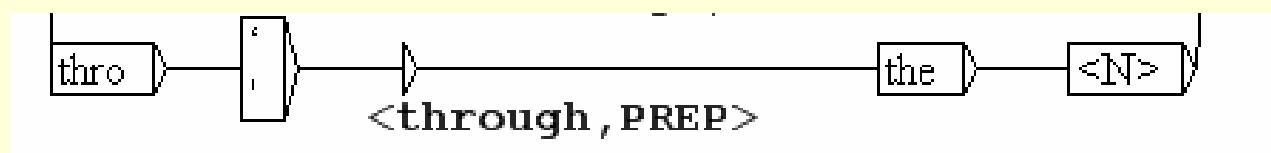
- Contracted forms
 - ‘tis for “it is” or it’s (13)
 - Annotate the word form as a sequence of annotations:
 - tis → <it,PRO+3+s> <is,be,V+PR+3+s>
- Archaic forms as in monyworth for "money's worth"
 - include them in the dictionary
- Elisions
 - though is spelt tho'
 - through is spelt thro'

Graphs or entries in the dictionary

Conjunction: Though is spelt tho'
tho',though,CONJ



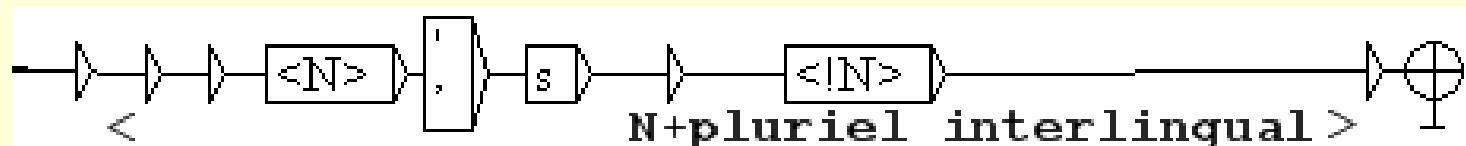
Prepositions: thro' for through
thro', through,PREP



→ BUT problem with different apostrophes

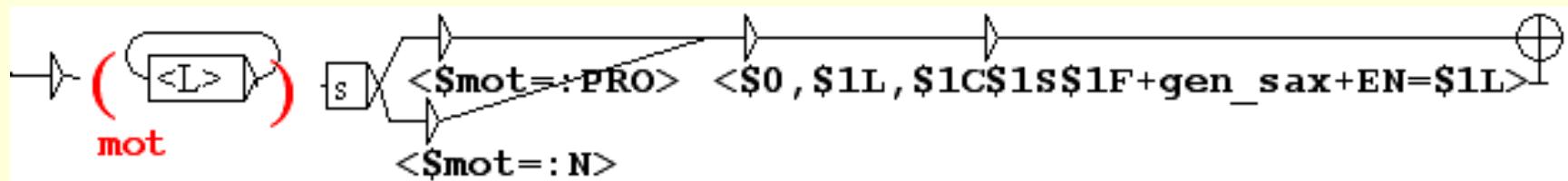
Nouns: 's = genitive or interlingual plural?

- Genitive: formerly without an apostrophe as in
 - “from the *womens* apartment”;
 - “out of their wives and childrens mouths”
- But interlingual plural of foreign nouns too:
 - *Egoumeno's*,
 - *Bassa's* (Pashas),
 - *piazza's*
- Spotted thanks to a NooJ grammar



Nouns : genitive how to add 's ?

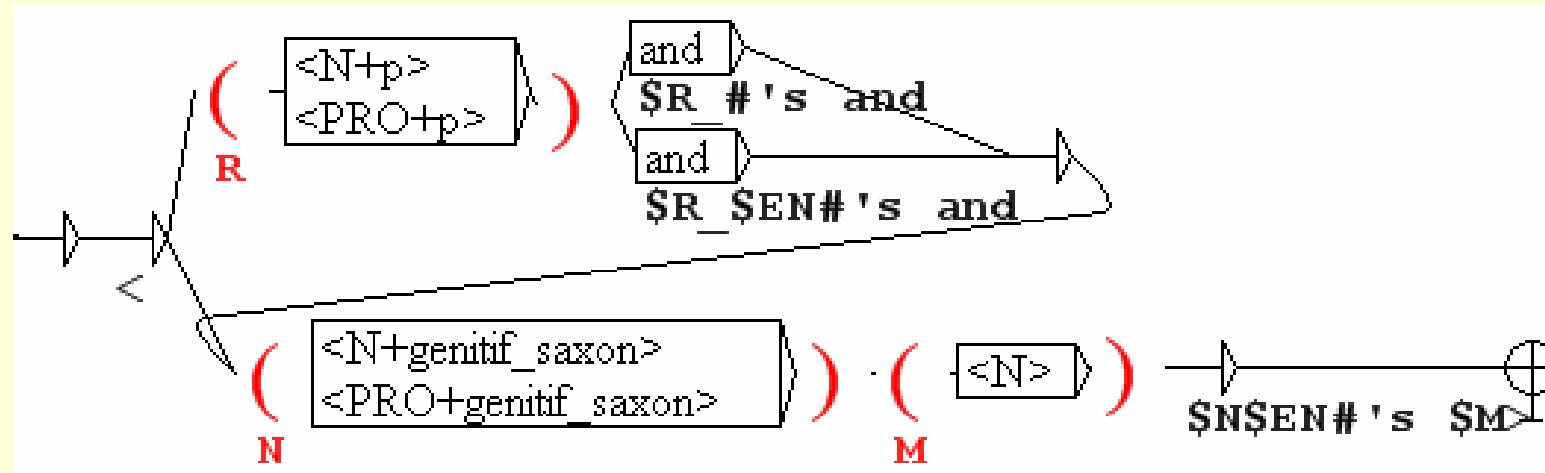
- Unknown words ending in s,
 - another : “one **another** company”
 - childrens : “their wives and **childrens** mouths”
- Morphological graph



- But **wives** is also a genitive !!!
 - the conjunction will help to recognise it with a syntactical graph
 - but we cannot automatise other cases
- N.B. wives' (GB) or (wives's) US

Nouns: morphological + syntactical graphs

- This graph can recognise
 - either noun + noun : <N+gen_sax> <N>
 - or pronoun + noun: <PRO+gen_sax> <N>



- **wives and childrens mouths** → <**wives's and children's mouths**>
- **another's company** → <**another's company**>

Nouns : Problems

- We may recognise unknown words, but what about known words????
- mans in « *no other mans errors could draw* »
« mans » is not identified as a genitive, because it is wrongly «recognised » as the **verb to man**.
transitive verb = tenir (switchboard, desk); armer [qch] en hommes (ship).

Verbs : preterite and PP

- **verbs ending in p, k, x, ss, sh**

→had two forms: normal form in ed

shortened form in t or 't

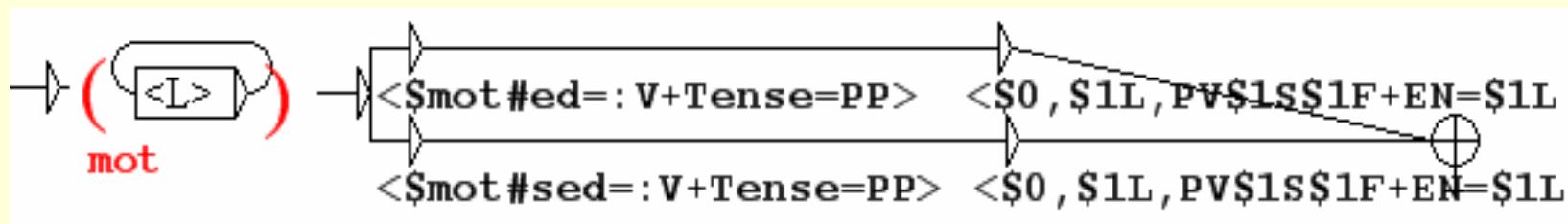
For regular forms and for shortened forms in t, we have a morphological graph

→we use a morphological graph and then a syntactical one

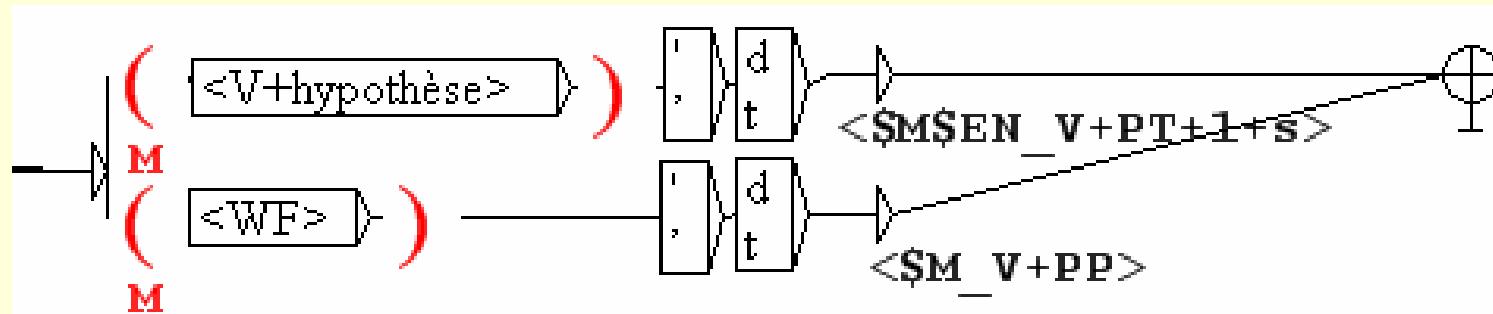
→different cases: KNOWN words and UNKNOWN words need different transformations

Verbs: two steps

- morphological graph
 - Recognises an unknown word by adding ed, ted or sed



- syntactical graph



Verbs: a KNOWN word like establish

establish can have 3 forms :

- “and *establish'd* the same number”,
- “the *establish't* doctrine”,
- “and for ever *established* the adoration”.
- FIRST STEP **the dictionary** recognises **establish** as the bare infinitive or the simple present of the verb to establish
- So the syntactical graph must modify the result of the dictionaries
 - SECOND STEP **the syntactical graph** **builds** **established**

Verbs: an UNKNOWN word like judg

- judg'd: “this being *judg'd*”
 - the form « judg » is NOT recognised
 - FIRST STEP: the morphological graph tests some hypotheses
 - “**judg**” + ed → **judged** = the PP of judge
 - **judg,judge,PV+Tense=PP+ EN=judge.**
 - At that point we don't know if judg is followed by 'd in the text
 - SECOND STEP the syntactical graph will confirm it and we get:
 - **judg,judge,V+Tense=PP+ EN=judge**

Verbs : UNKNOWN word as profes

- *profes'd , profest, professed*
- « profes » is UNKNOWN
- **FIRST STEP** a graph deletes a mute e and suggests
profes,prof,N+Nb=p+Distribution=Hum+EN=prof
another graph adds /sed/,
profes,profess,PV+Tense=PP+EN=profess
(the full graph « recognises » a preterite and an adjective).
- **SECOND STEP** the syntactical graph invalidates the noun , and validates the verb
- **profes,profess,V+Tense=PP+EN=profess**

Verbs: case of an UNKNOWN word, e.g. imbrac

- To be able to recognise *imbrac'd* we must combine two transformations
 - **imbrac'd → embrac'd → embraced.**
several morphological graphs with different levels of priority
- Our corpus has 90 occurrences of 'd or 't.

Conclusion

- **Verbs**
- **It is sometimes difficult to process verbs**
- **The order of application of the graphs is important**
- **Nouns**
- **We failed to process some genitives**
- **We must improve the grammar of the syntagm**