

# Annotations



---

**His MMSE was 23/30 on 15 January 2008.**



# Annotations



His MMSE was 23/30 on 15 January 2008.

0...5...10...15...|...|...|...|...



# Annotations



His MMSE was 23/30 on 15 January 2008.

0...5...10...15...|...|...|...|...





# Annotations



His MMSE was 23/30 on 15 January 2008.

0...5...10...15...|...|...|...|...



Id	Type
1	sentence

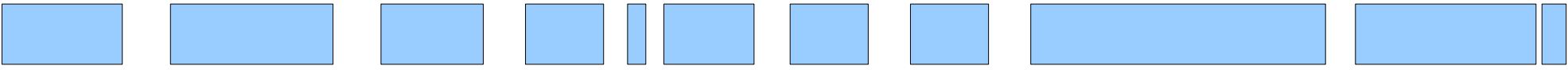


# Annotations



His MMSE was 23/30 on 15 January 2008.

0...5...10...15... | ... | ... | ... | ...



Id	Type
1	sentence

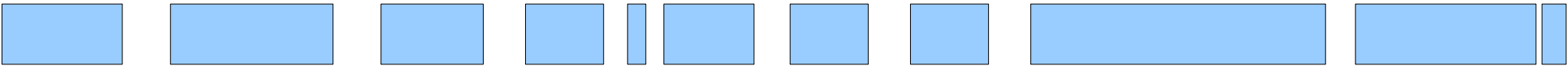


# Annotations



His MMSE was 23/30 on 15 January 2008.

0...5...10...15... | ... | ... | ... | ...



Id	Type
1	sentence
2	token
3	token
4	token
5	token
6	token
7	token



# Annotations

His MMSE was 23/30 on 15 January 2008.  
 0...5...10...15...|...|...|...|...

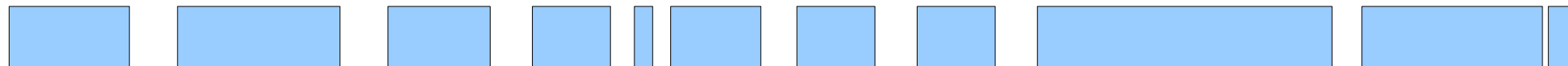


Id	Type	Start	End
1	sentence	0	39
2	token	0	3
3	token	4	8
4	token	9	12
5	token	13	15
6	token	15	16
7	token	16	18



# Annotations

His MMSE was 23/30 on 15 January 2008.  
 0...5...10...15...|...|...|...|...



Id	Type	Start	End	Features
1	sentence	0	39	
2	token	0	3	pos=PP
3	token	4	8	pos=NN
4	token	9	12	pos=VB
5	token	13	15	pos=CD
6	token	15	16	pos=SM
7	token	16	18	pos=CD





# Annotations

His MMSE was 23/30 on 15 January 2008.

0...5...10...15... | ... | ... | ... | ...



Id	Type	Start	End	Features
1	sentence	0	39	
2	token	0	3	pos=PP
3	token	4	8	pos=NN
4	token	9	12	pos=VB      root=be
5	token	13	15	pos=CD      type=num
6	token	15	16	pos=SM      type=slash
7	token	16	18	pos=CD      type=num

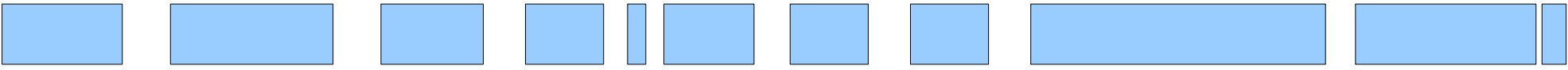


# Dictionary lookup



His MMSE was 23/30 on 15 January 2008.

0...5...10...15...|...|...|...|...



Month

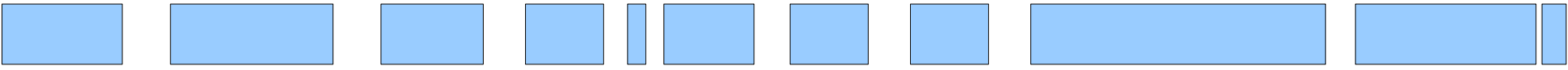


# Dictionary lookup



His MMSE was 23/30 on 15 January 2008.

0...5...10...15...|...|...|...|...



MMSE

Month



# Limitations of dictionary lookup

---

- Dictionary lookup is designed for finding simple, regular terms and features
- False positives
  - *“He may get better”*
  - *“Mother is a smoker”*
  - *“He often burns the toast, setting off the smoke alarm”*
- Cannot deal with complex patterns
  - For example, recognising e-mail addresses using just a dictionary would be impossible
- Cannot deal with ambiguity
  - I for Iodine, or I for me?



# Pattern matching

---

- The early components in a GATE pipeline produce simple annotations
  - Token, Sentence, Dictionary lookups
- These annotations have features
  - Token kind, part of speech, major type...
- Patterns in these annotations and features can suggest more complex information
- A pattern matching language, JAPE, is used to find these patterns

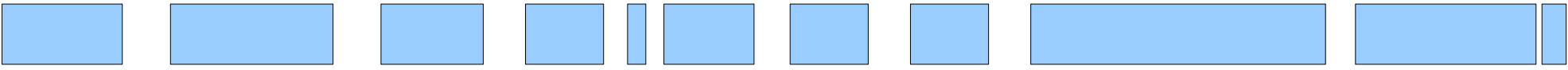


# Patterns



His MMSE was 23/30 on 15 January 2008.

0...5...10...15...|...|...|...|...



MMSE

Month

{number} {Month} {number}

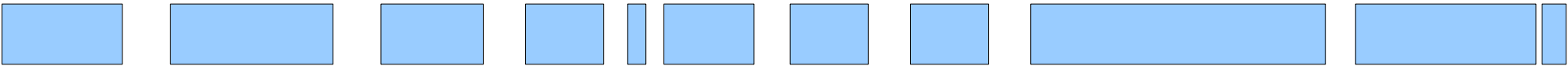


# Patterns



His MMSE was 23/30 on 15 January 2008.

0...5...10...15...|...|...|...|...



MMSE

Month

Date

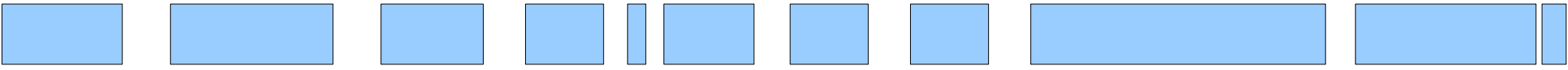


# Patterns



His MMSE was 23/30 on 15 January 2008.

0...5...10...15...|...|...|...|...



MMSE

Month

{number} {slash} {number}

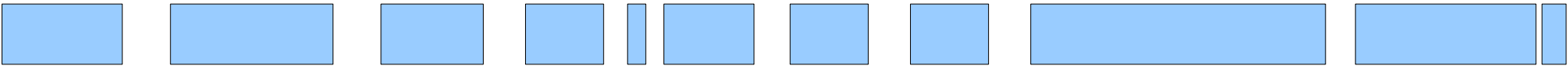




# Patterns

His MMSE was 23/30 on 15 January 2008.

0...5...10...15...|...|...|...|...



MMSE

Month

Score

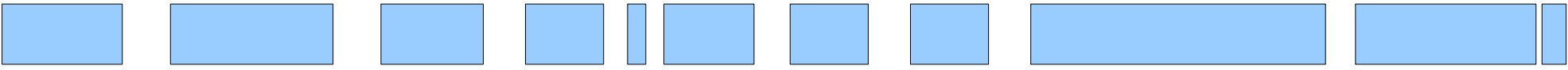


# Patterns



His MMSE was 23/30 on 15 January 2008.

0...5...10...15...|...|...|...|...



MMSE

Month

Score

Date

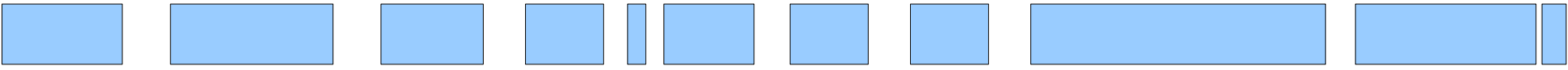


# Patterns



His MMSE was 23/30 on 15 January 2008.

0...5...10...15...|...|...|...|...



MMSE

Month

Score

Date

{MMSE} {BE} {Score} {?} {Date}

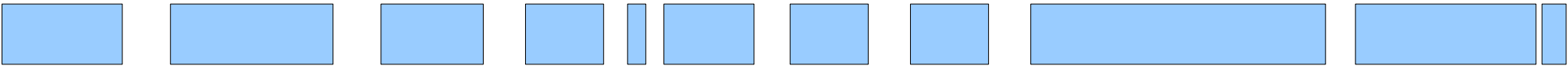


# Patterns



His MMSE was 23/30 on 15 January 2008.

0...5...10...15...|...|...|...|...



MMSE

Month

Score

Date

MMSE with score and date



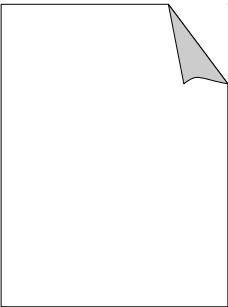
# Patterns are general

---

- MMSE was 23/30 on 15 January 2009
- Mini mental was 25/30 on 12/08/07
- MMS was 25/30 last week
- MMSE is 25/30 today
- With adaptation
  - MMSE 25 out of 30
  - Long range dependencies on dates

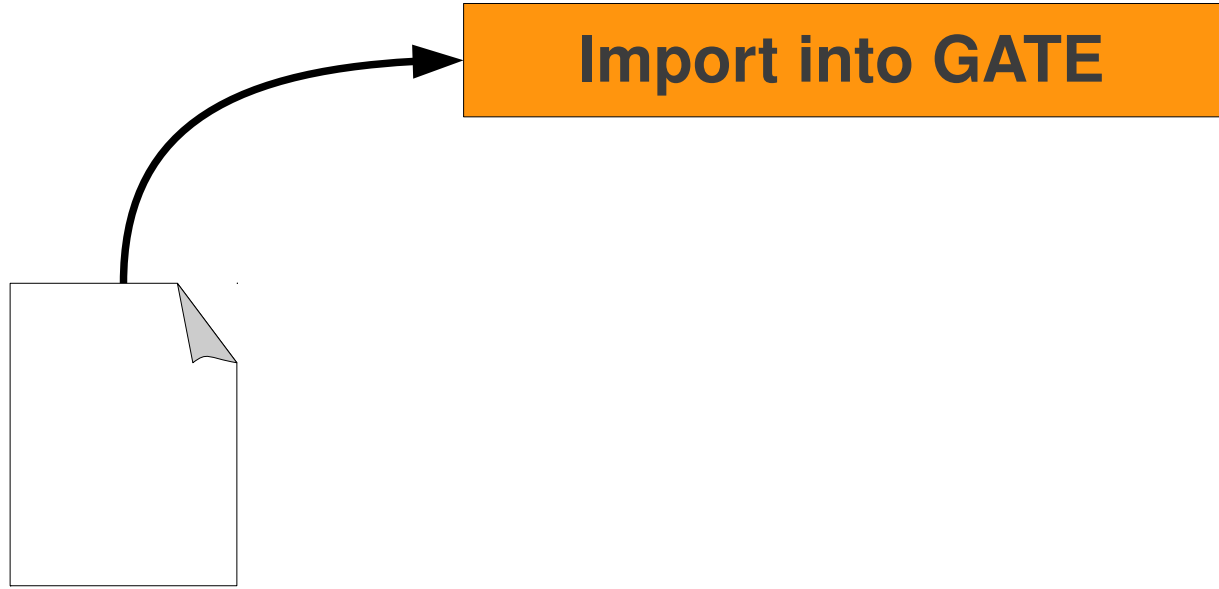


# MMSE pipeline



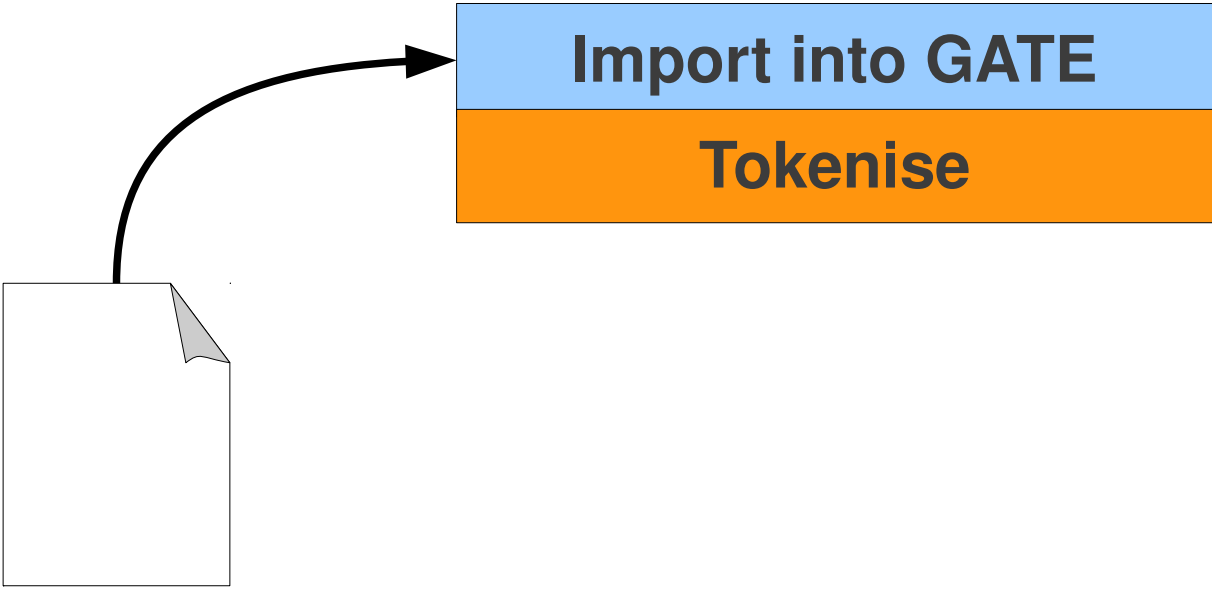


# MMSE pipeline





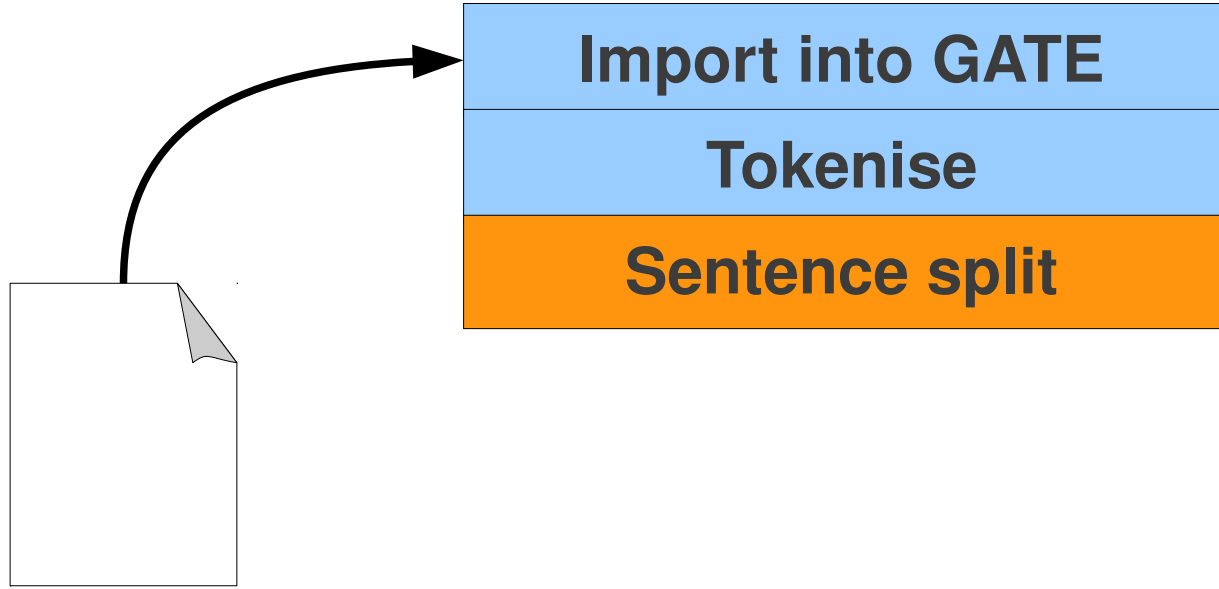
# MMSE pipeline





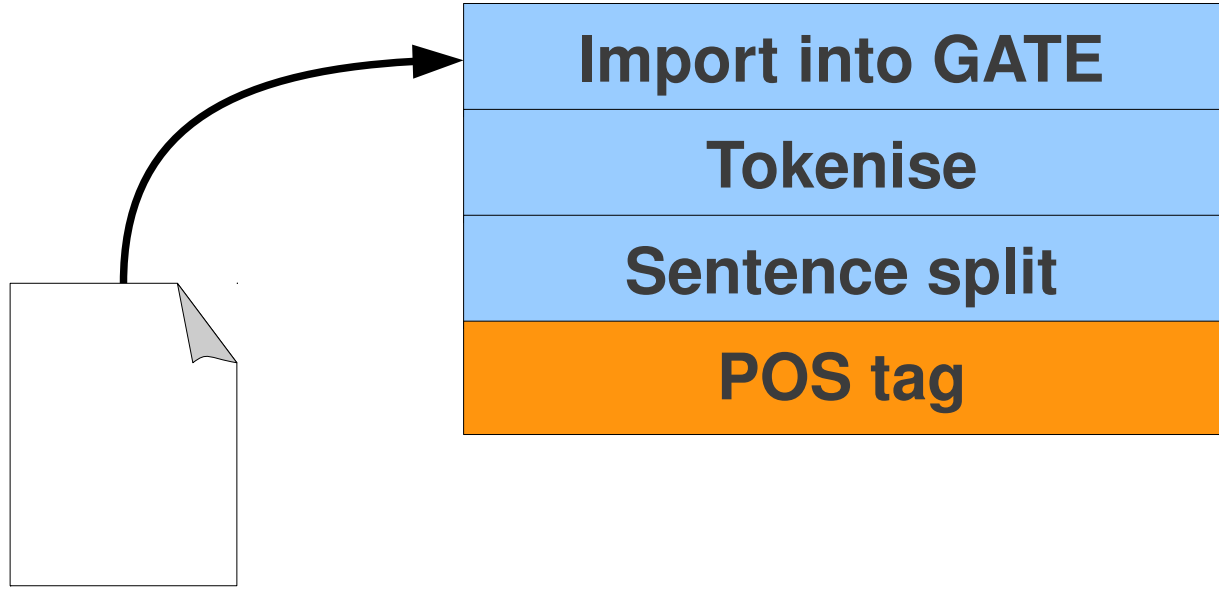


# MMSE pipeline



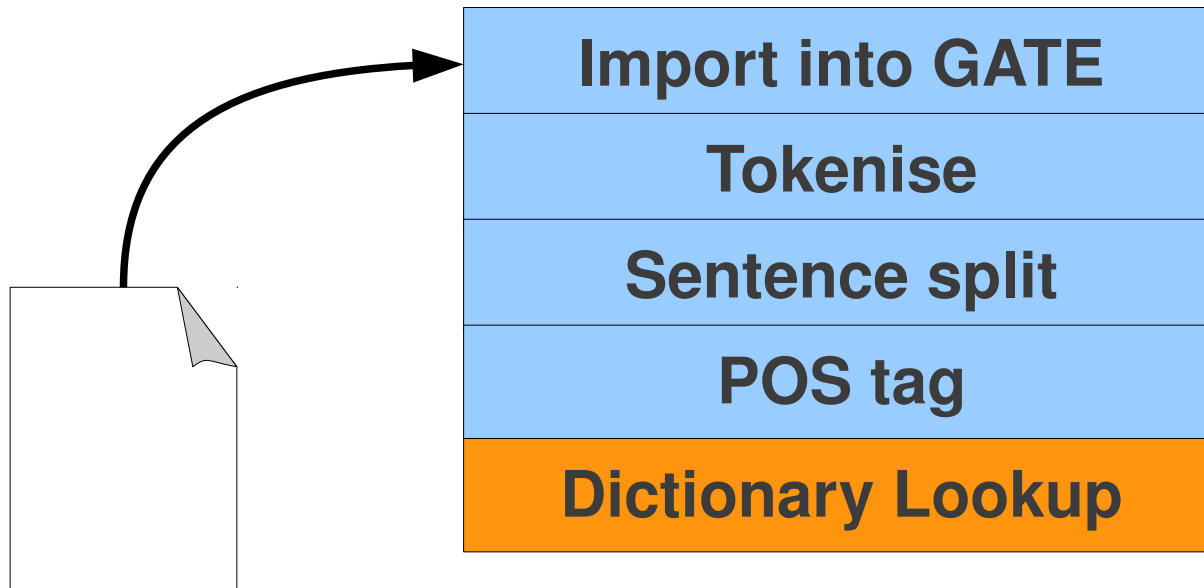


# MMSE pipeline



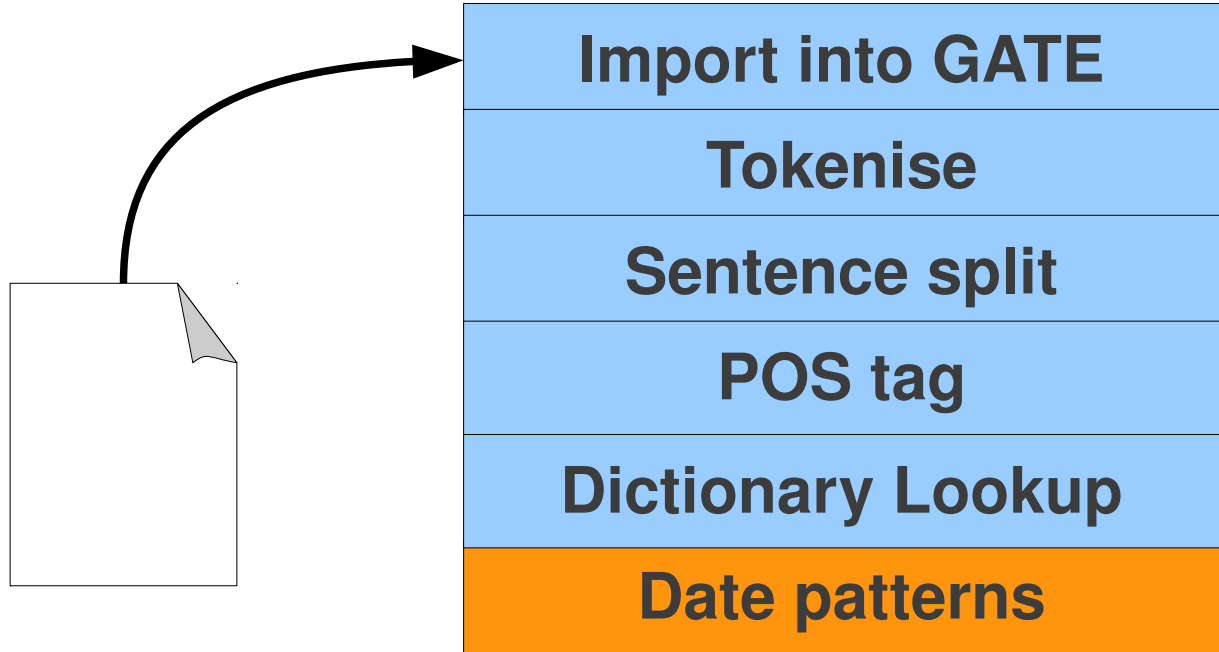


# MMSE pipeline



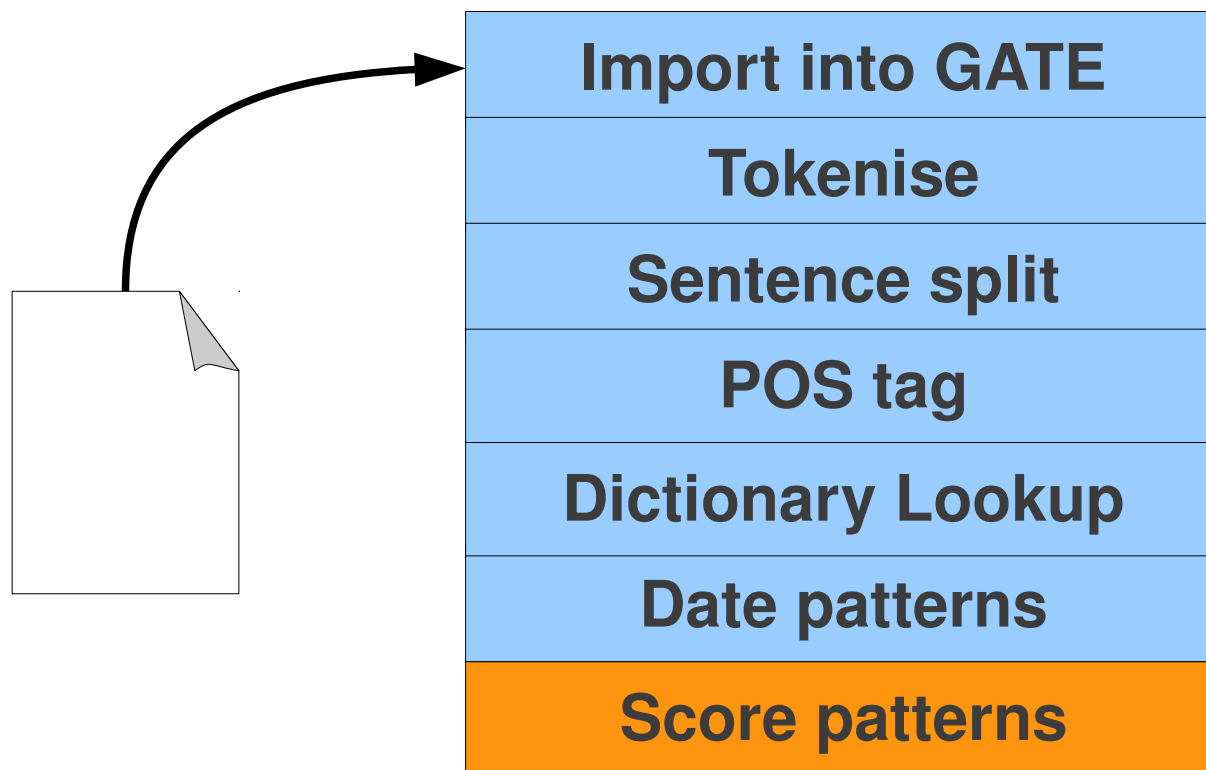


# MMSE pipeline



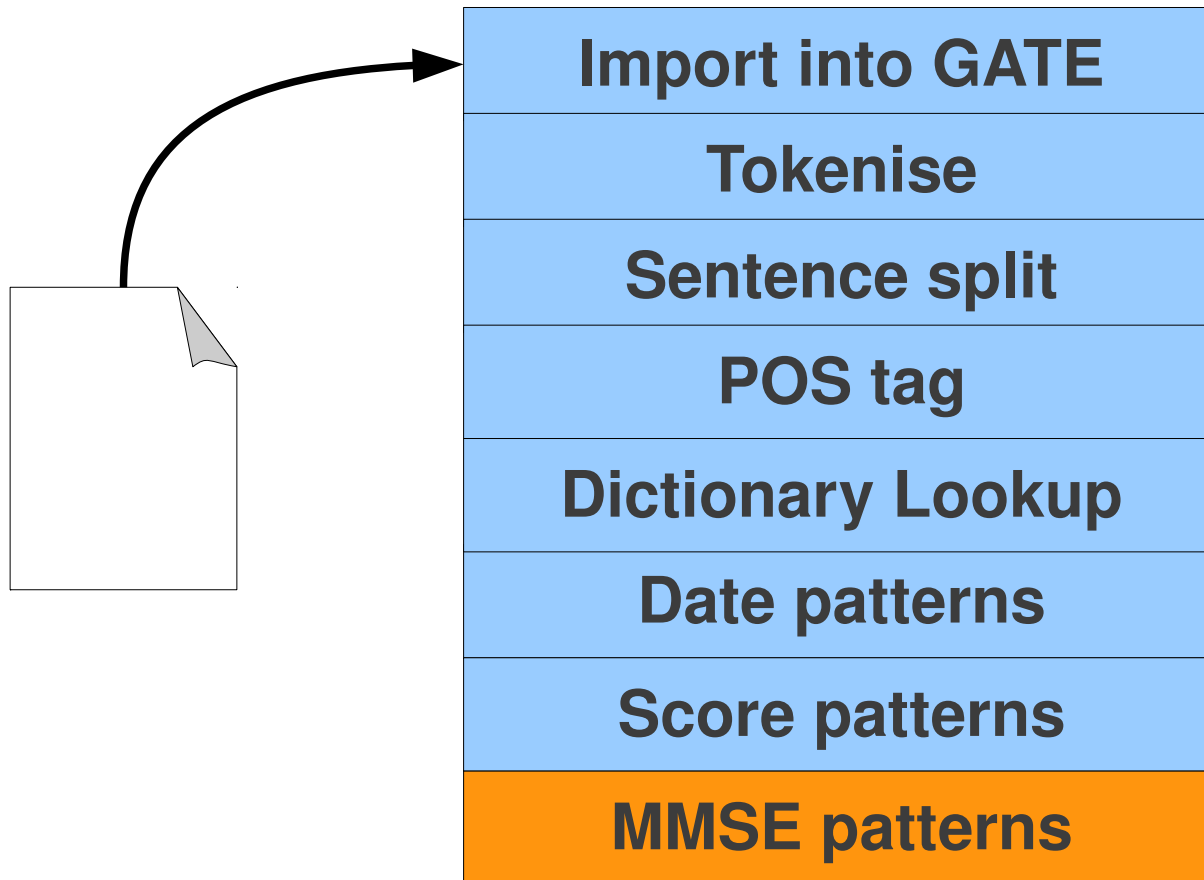


# MMSE pipeline





# MMSE pipeline





# MMSE pipeline

