

## Types & tokens

- (1) De nombreux visiteurs la voient dans la glace de la chambre.

## Types & tokens

(1) De nombreux visiteurs la voient dans la glace de la chambre.

```
no token
1 De
2 nombreux
3 visiteurs
4 la
5 voient
6 dans
7 la
8 glace
9 de
10 la
11 chambre
12 .
```

## Types & tokens

(1) De nombreux visiteurs la voient dans la glace de la chambre.

```
no token
1 De
2 nombreux
3 visiteurs
4 la
5 voient
6 dans
7 la
8 glace
9 de
10 la
11 chambre
12 .
```

12 tokens = taille du corpus

## Types & tokens

(2) De nombreux visiteurs la voient dans la glace de la chambre.

no token	type
1 De	De
2 nombreux	nombreux
3 visiteurs	visiteurs
4 la	la
5 voient	voient
6 dans	dans
7 la	la
8 glace	glace
9 de	de
10 la	la
11 chambre	chambre
12 .	.

12 tokens = taille du corpus

10 types avant normalisation

## Types & tokens

(3) De nombreux visiteurs la voient dans la glace de la chambre.

no token	type
1 De	de
2 nombreux	nombreux
3 visiteurs	visiteurs
4 la	la
5 voient	voient
6 dans	dans
7 la	la
8 glace	glace
9 de	de
10 la	la
11 chambre	chambre
12 .	.

12 tokens = taille du corpus

9 types après normalisation

## Types & tokens

(4) De nombreux visiteurs la voient dans la glace de la chambre.

no token	type	lemme
1 De	de	de
2 nombreux	nombreux	nombreux
3 visiteurs	visiteurs	visiteur
4 la	la	le
5 voient	voient	voir
6 dans	dans	dans
7 la	la	le
8 glace	glace	glace
9 de	de	de
10 la	la	la
11 chambre	chambre	chambre
12 .	.	PONCT

12 tokens = taille du corpus

9 types après normalisation

9 lemmes

## Types & tokens

(5) De nombreux visiteurs la voient dans la glace de la chambre.

no token	type	lemme	catégorie
1 De	de	de	DET
2 nombreux	nombreux	nombreux	ADJ
3 visiteurs	visiteurs	visiteur	NOM
4 la	la	le	CLI
5 voient	voient	voir	VER
6 dans	dans	dans	PRP
7 la	la	le	DET
8 glace	glace	glace	NOM
9 de	de	de	PRP
10 la	la	la	DET
11 chambre	chambre	chambre	NOM
12 .	.	PONCT	PONCT

12 tokens = taille du corpus

9 types après normalisation

9 lemmes

11 items lexicaux – paires (lemme, cat)