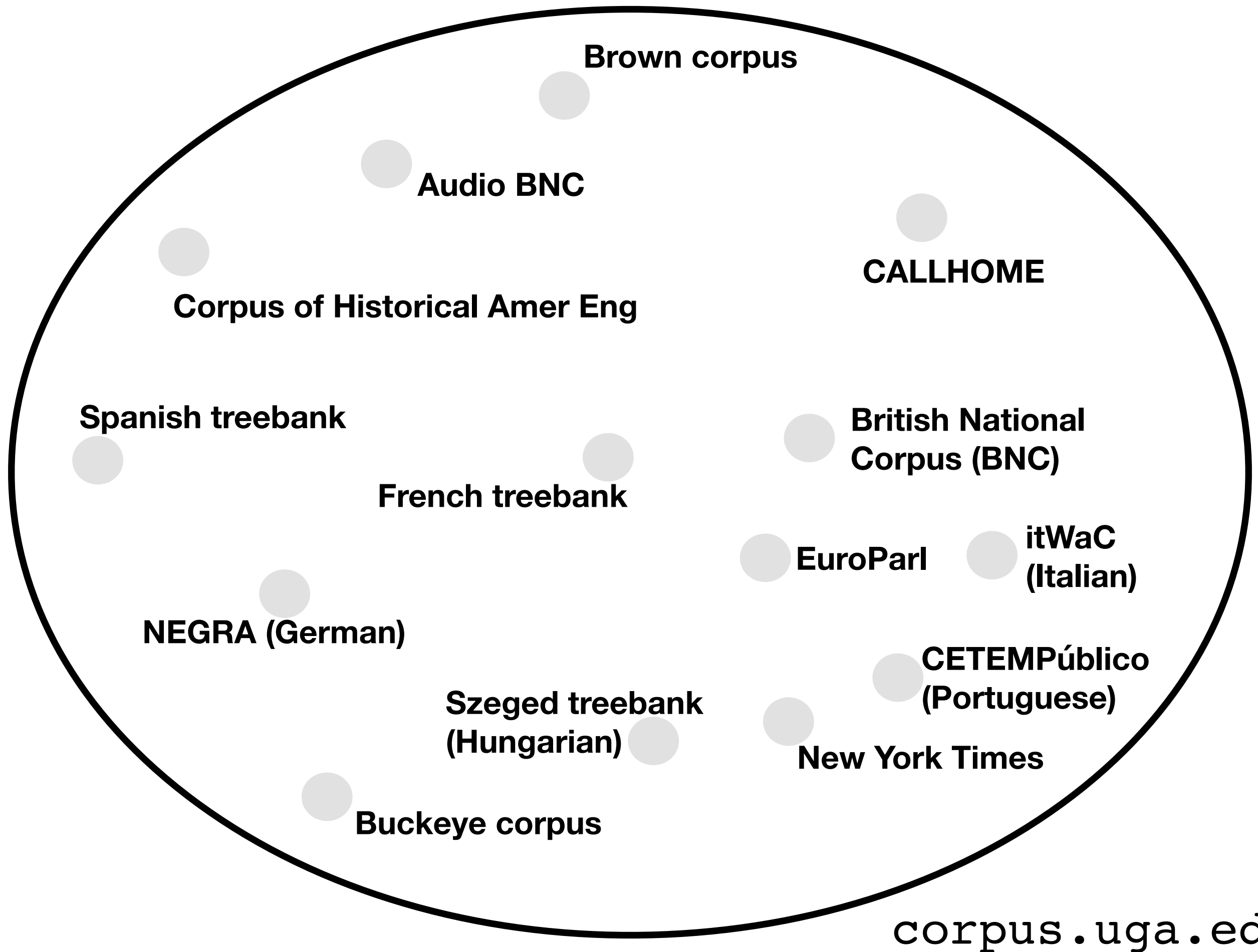


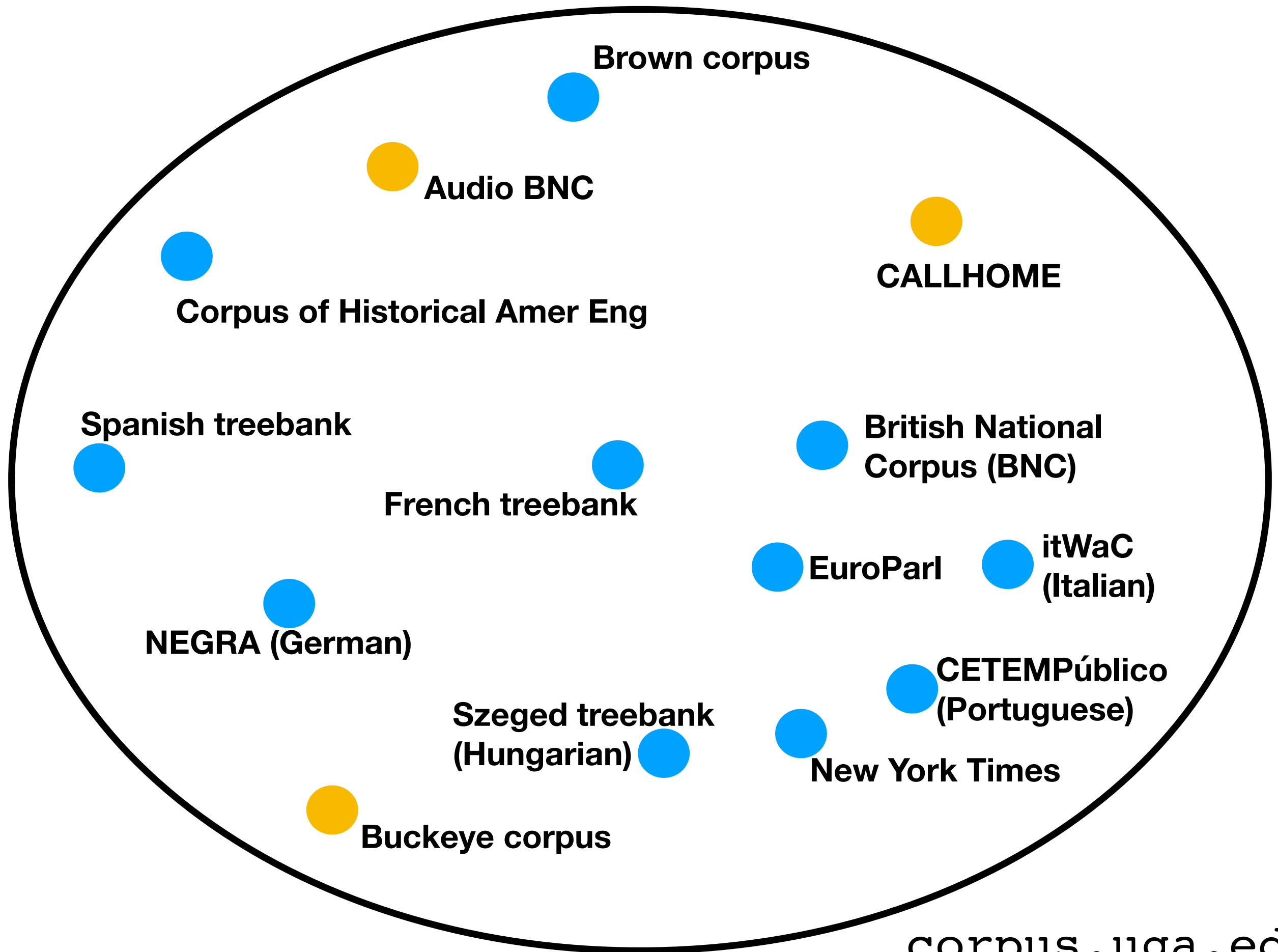
Text Corpora You Can Use

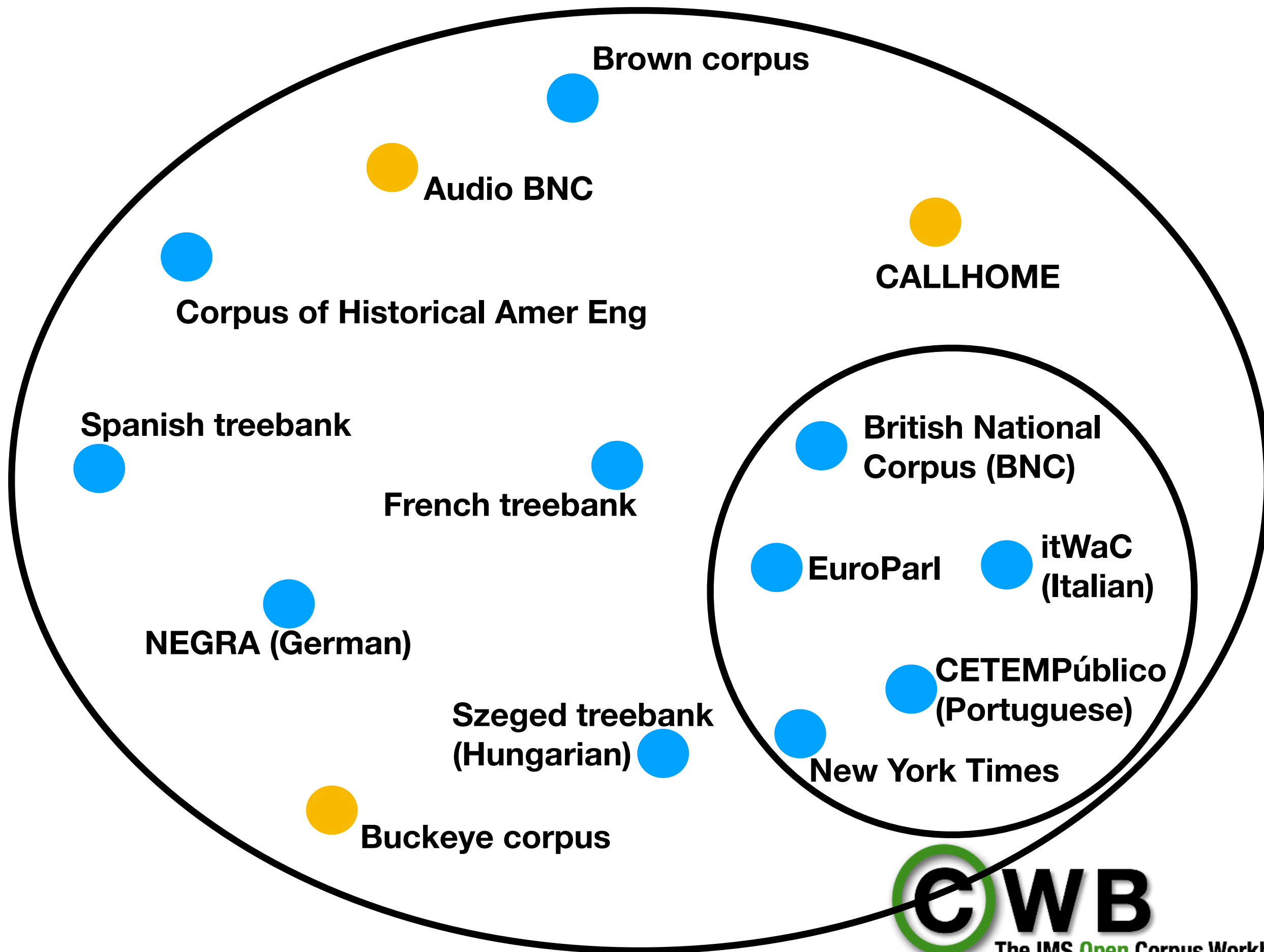
John Hale



UNIVERSITY OF
GEORGIA
Department of Linguistics







text domain	publication date	medium
imaginative (19%)	1960-74 (2%)	book (57%)
natural & pure science (4%)	1975-1984 (5%)	periodical (32%)
applied science (8%)	1985-1993 (91%)	misc. published (4%)
social science (16%)	unknown (2%)	misc. unpublished (5%)
world affairs (20%)		to-be-spoken (1%)
commerce & finance (8%)		
arts (7%)		
belief and thought (3%)		
leisure (14%)		

text domain	publication date	medium
imaginative (19%)	1960-74 (2%)	book (57%)
natural & pure science (4%)	1975-1984 (5%)	periodical (32%)
applied science (8%)	1985-1993 (91%)	misc. published (4%)
social science (16%)	unknown (2%)	misc. unpublished (5%)
world affairs (20%)		to-be-spoken (1%)
commerce & finance (8%)		
arts (7%)		
belief and thought (3%)		
leisure (14%)		

This concordance was generated by:

User: unknown (unknown)
Date: Fri Jan 11 15:19:14 2019
Corpus: bnc
British National Corpus (XML edition)
Subcorpus: BNC:Last
Number of Matches: 5494
Left display context: 25 characters
Right display context: 25 characters

Query text:

BNC; "lovely" ;

117507: nally fulfilled , as many **lovely** ones . Amis also likes t
301274: ; Venica) which is quite **lovely** . So it should be at £10
338909: . `` Conroy , angel ! How **lovely** to see you ! " She twin
341748: calmly . `` It will be a **lovely** surprise for them both w
345721: if you like . I 've got a **lovely** mirror . It 's much bett
352106: ` No . It would have been **lovely** , but it could n't reall
352869: nd saying I think he 's a **lovely** gentleman ! " `` Oh he
355826: id say Mr Merrivale was a **lovely** young man ! " They were
387382: aking it for us . It 's a **lovely** photograph --- I still c
398912: " `` Ooh , that would be **lovely** . Let me get the little
404026: coffee ? " `` No , it 's **lovely** , thank you . " `` Do n
422343: either . She 's a really **lovely** woman but she 's not ver

	female	male
<i>lovely</i> attestations	1265	751
corpus size	11983120	11983120

contingency table for *lovely*

type

The second character of a verb tag marks the type of verb as follows:

B	Forms of <i>be</i> (VBB VBD VBG VBI VBN VBZ)
D	Forms of <i>do</i> (VDB VDD VDG VDI VDN VDZ)
H	Forms of <i>have</i> (VHB VHD VHG VHI VHN VHZ)
M	Other modal verbs (VM0)
V	Lexical verb (VVB VVD VVG VVI VVN VVZ)

Inflection

The third character of a verb tag marks the verb inflection as follows:

B	base form finite
D	past tense
Z	3rd person sing present
N	past participle
I	infinitive
G	present participle

find agentless passives



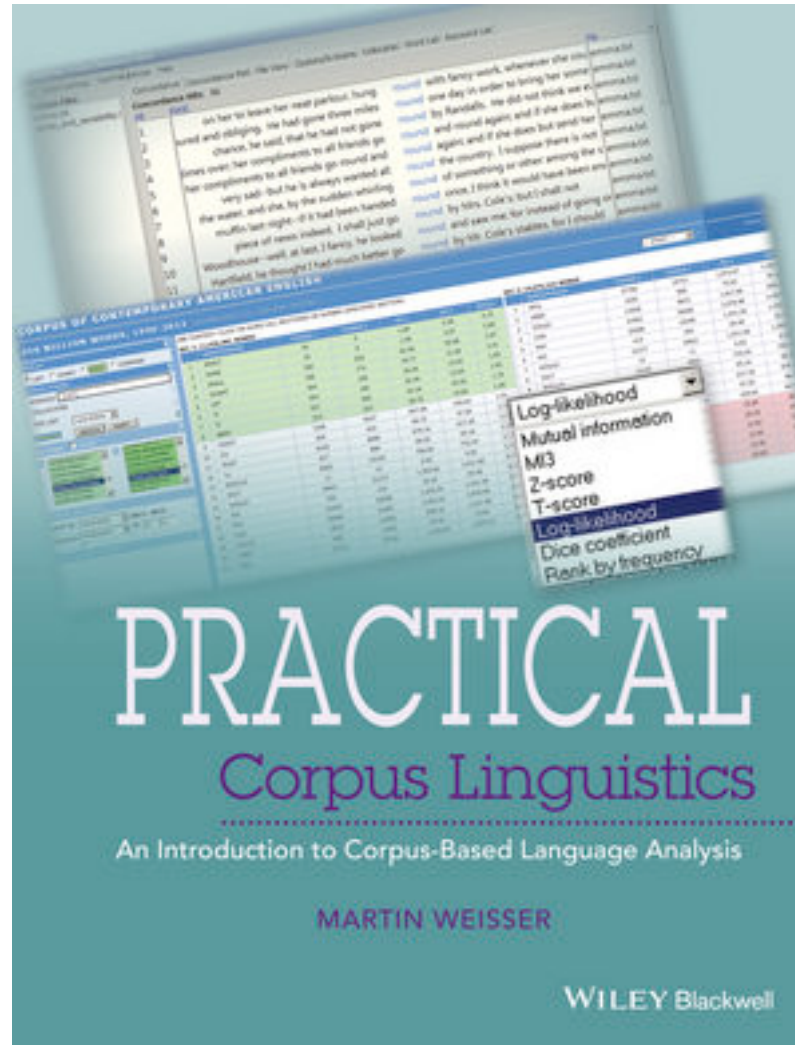
Variation across speech and writing

Douglas Biber

Department of English, Northern Arizona University



find coordination



```
a:[ pos = "N.*|J.*|V.*|AV.?" & pos != "NP0"]  
  [ word="and|an"%c & pos="CJC"]  
  [pos=a.pos]
```

distinguish word meanings



Hanks, a lexicographer ... hypothesized that *strong* is an intrinsic quality whereas *powerful* is an extrinsic one.

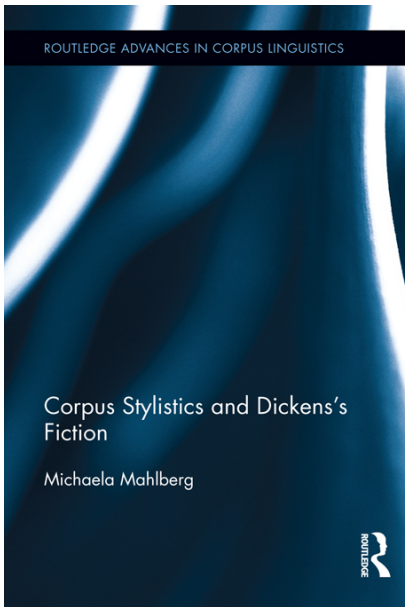
Thus, for example, any worthwhile politician or cause can expect *strong* supporters, who are enthusiastic, convinced, vociferous, etc.,

but far more valuable are *powerful* supporters, who will bring others with them.

Dickens' *as if*

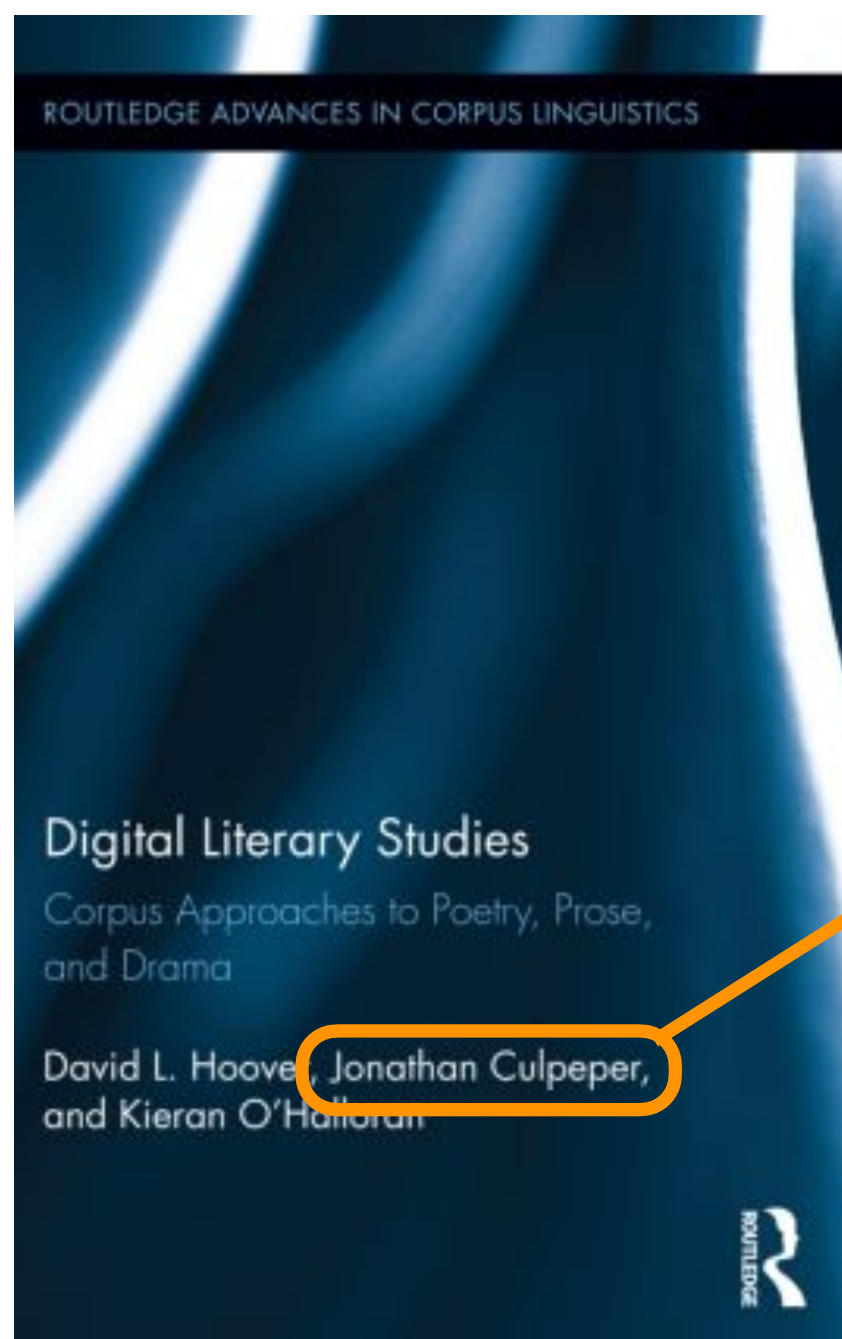
Table 7.3 Meaning groups of left-collocates of *as if* (frequent collocates)

Meaning group	Collocates
Action verbs	<i>made, turned, went, stood, shook, stopped, came</i>
Body part nouns	<i>head, hand, hands, eyes, face, back, mouth, arms, arm, lips</i>
Settings	<i>door, fire, room, side, chair, wall</i>
Manner	<i>manner, way, air, seemed</i>
LOOK	<i>looked, looking, look, looks</i>
SPEAK	<i>said, speaking, spoke, voice</i>
FEEL	<i>felt, feel</i>
Other	<i>time, moment, man, round, down, again, little, almost, very, great, now, well, quite, still</i>



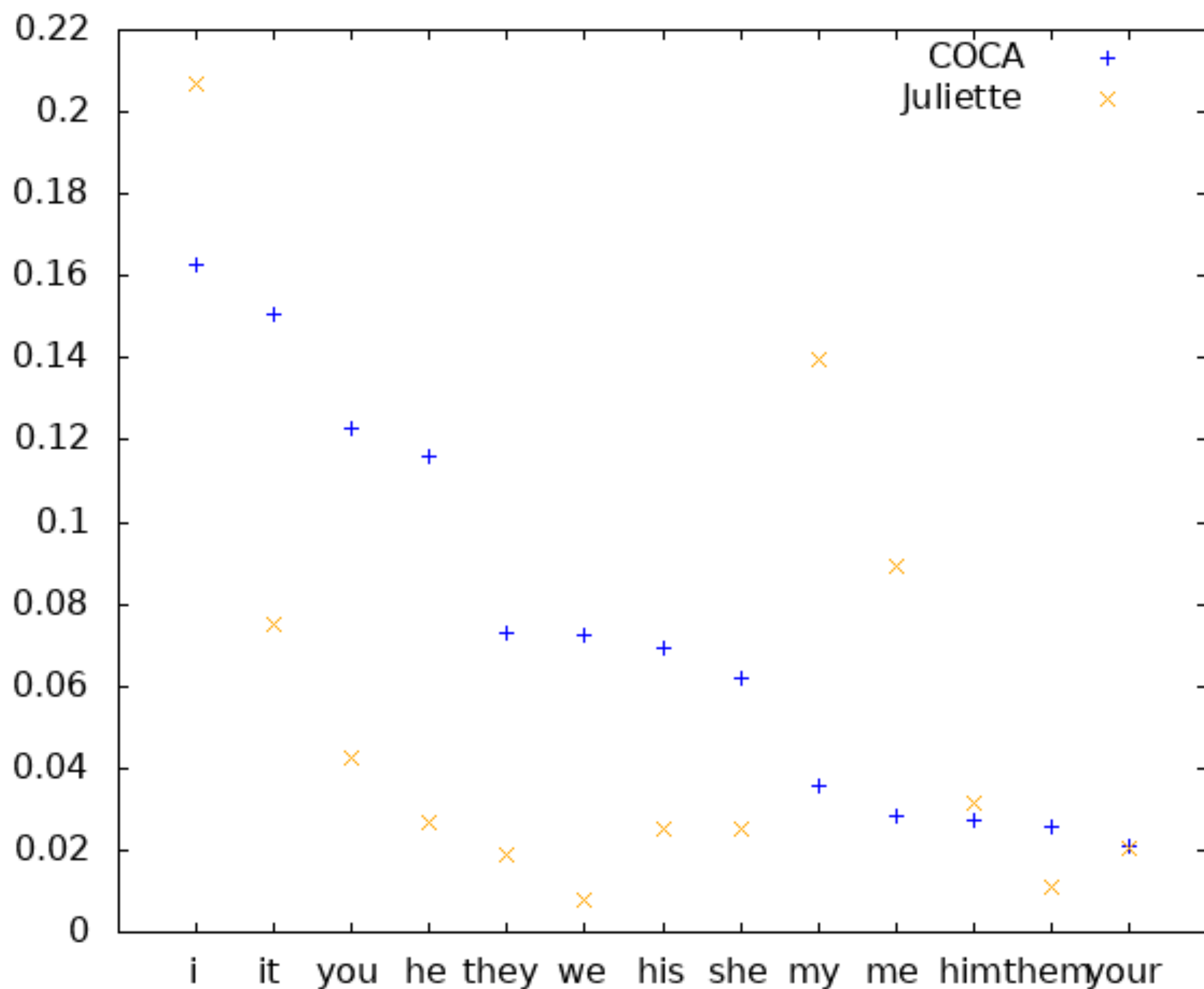
yes, the **body** is a big deal!

all lexical verbs and common nouns	3580
<hr/>	
action verbs	205
body part nouns	335
Settings	107
Manner	125
LOOK	292
SPEAK	120
FEEL	102
other	298
<hr/>	



relative frequency

COCA
Juliette



Thank you!

<http://research.franklin.uga.edu/linglab/corpora>