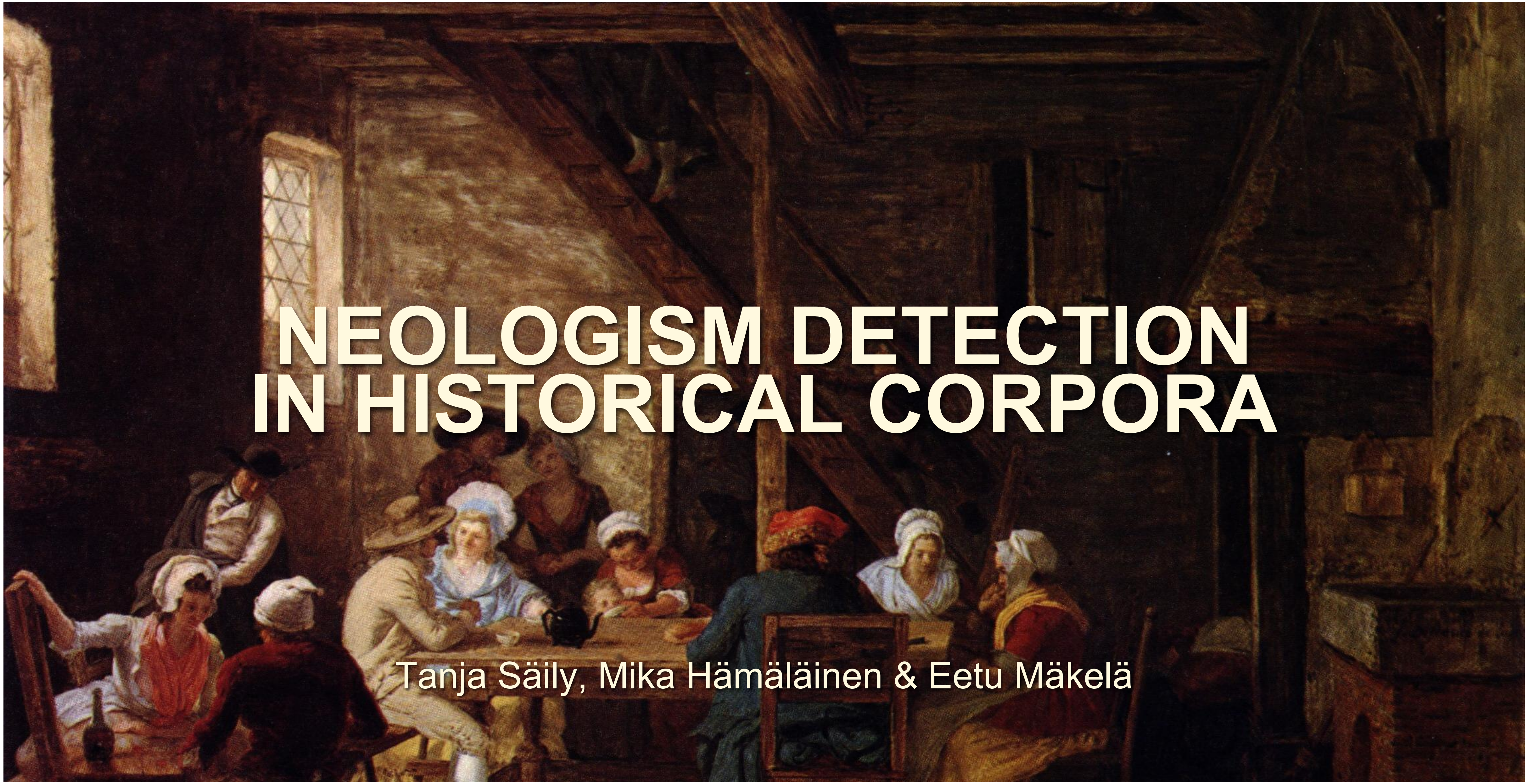


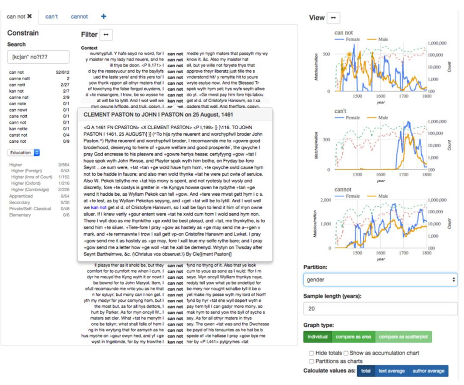
NEOLOGISM DETECTION IN HISTORICAL CORPORA

Tanja Säily, Mika Hämäläinen & Eetu Mäkelä

INTRODUCTION

STRATAS PROJECT

- *Interfacing structured and unstructured data in sociolinguistic research on language change* (Academy of Finland, DIGIHUM, 2016–2019)
  - blogs.helsinki.fi/stratas-project/
- NATAS subproject:
  Social embedding of neologisms in early English correspondence
  - Previous research: mostly lexicographical data, bias towards well-known authors
  - *Corpora of Early English Correspondence* (CEEC): wide social spectrum, speech-like genre



[kc]an* no?t??

can not	52/612
canne nott	2
can nott	2/27
kan not	2/7
canne not	2/9
can note	0/1
can nowt	0/1
cane nott	0/1
cann not	0/1
kan notte	0/1
canst not	0/4
cane not	0/9

Higher	3/564
Higher (Foreign)	0/43
Higher (Inns of Court)	1/152
Higher (Oxford)	1/216
Higher (Cambridge)	2/226
Apprenticed	0/64
Secondary	0/30
Private/Self: Classical	0/48
Elementary	0/6

Context

wurshypfull. Y hafe seyd no word, for I	can not	medle yn hygh maters that passyth my wy
y maister ne my lady had neuere, and he	can not	know it, &c. Also my maister hat
ill thys be doon. <P II,171> I	can not	ell, but ye wille not foryete thys that
d by the resseyuour and by the bayllyfs	can not	approve theyr liberatz just tille the s
ued the laste yere/ and this yere to/ I	can not	vnderstond hit/ y remytte hit to youre
yow thynk vppon all othyr maters that I	can not	wryte esylye now. And the Blessed Tr
of towchyng the false forgyd euydens, I	can not	spek wyth hym yet; hys wyfe seyth allwe
d +te masangere, I trow, be so wysse he	can not	do yt. +Ge mwst pay him fore hijs labou
at will be to lytill. And I wot well we	kan not	get xl d. of Cristofyre Hanswm, so I xa
myn owune fyflode, and truli, cosyn, I	can not	gadere that well. And therffore, cosyn,

CLEMENT PASTON to JOHN I PASTON on 25 August, 1461

<Q A 1461 FN CPASTON> <X CLEMENT PASTON> <P I,199> [} [\116. TO JOHN PASTON I 1461, 25 AUGUST\] }] (^To hijs rythe reuerent and worchypfwll broder John Paston.^) Rythe reuerent and worchypfwll broder, I recomawnde me to +gowre good broderhood, desieryng to herre of +goure welfare and good prosperite', the qwyche I pray God encresse to his pleswre and +gowre hertys hesse; certyfyyng +gow +tat I haue spok wyth John Rwsse, and Playter spak wyth him bothe, on Fryday be-fore Seynt ...ce sum were, +tat +tan +ge wold haue hym hom, +te qwyche xwld cause hym not to be hadde in fauore; and also men wold thynke +tat he were put owte of seruice. Also W. Pekok tellythe me +tat hijs mony is spent, and not ryotesly but wysly and discretly, fore +te costys is gretter in +te Kyngys howse qwen he rydythe +tan +ge wend it hadde be, as Wyllam Pekok can tell +gow. And +tere wee mwst gett hym i c s. at +te lest, as by Wyllam Pekokys seyyng, and +get +tat will be to lytill. And I wot well we kan not get xl d. of Cristofyre Hanswm, so I xall be fayn to lend it him of myn owne siluer. If I knew verily +gour entent were +tat he xwld cum hom I wold send hym non. There I wyll doo as me thynkithe +ge xwld be best plesyd, and +tat, me thynkythe, is to send him +te siluer. +Tere-fore I pray +gow as hastely as +ge may send me a-+gen v mark, and +te remnawnte I trow I xall gett vp-on Cristofire Hanswm and Lwket. I pray +gow send me it as hastely as +ge may, fore I xall leue my-selfe rythe bare; and I pray +gow send me a letter how +ge woll +tat he xall be demenyd. Wrytyn on Twsday after Seynt Barthelmwe, &c. (\Christus vos obseruet.\) By Cle[{ment Paston{]

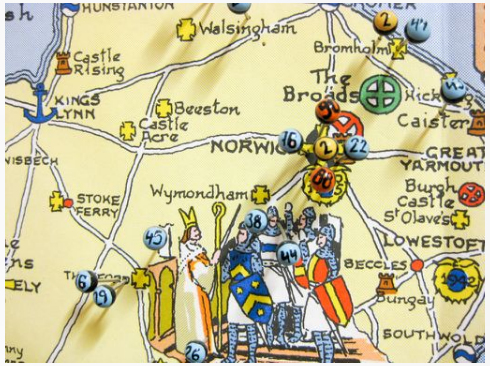
ll plasys ther as it shold be, but they	can not	fynd no thyng of it. Also that ye look
comfort for to cumfort me when I cum. I	can not	cum to youe as sone as I wuld: ffor I m
dyr he meuyd the Kyng wyth it or nowt I	can not	seye. Myn oncyll Wylliam thynkys naye,
be bownd for to John Maryot. Item, I	can not	redyly tell yow what ye be endettyd for
efull recomaunde me vnto you as he that	can not	be mery nor nought schalbe tyll it be o
n for syluyr; but mony can I non get. I	can not	yet make my pesse wyth my lord of Norff
yth my modyr for your comyng hom, but I	can not	fynd by hyr +tat she wyll depert wyth e
the most but, as for all hys dettors, I	can not	pay hem tyll I can gadyr more mony, so
hurt by Parker. As for myn oncyll W., I	can not	mak hym to send yow the byll of syche s
maters set cler. What +tat he menyth I	can not	sey. As for all othyr maters in thys
one be takyn; what shall falle of hem I	can not	sey. The qwen +tat was and the Dwchesse
ng in his wrytyng that for asmych as he	can not	be payd of his tenauntes as he hat be b
hus myche on +gour owyn hed, and yf +ge	can not	spede of +te hatlase I pray +gow bye me
wyst in Ingelonde, for by my trowthe I	kan not	her by <P I,441> pylgrymes +tat

gender

20

☐ Hide totals ☐ Show as accumulation chart
☐ Partitions as charts

CEEC



- Personal letters, c. 1400–1800
  - 1,180 writers, 11,713 letters, 5.2 million words
  - Compiled for historical sociolinguistics: **metadata** on letters, writers, recipients (e.g. gender, social rank)
- Compiled by T. Nevalainen, H. Raumolin-Brunberg et al. at the University of Helsinki
  - Based on published editions of letters
- SCEEC = Standardized-spelling version using VARD2 (excluding 15th century)

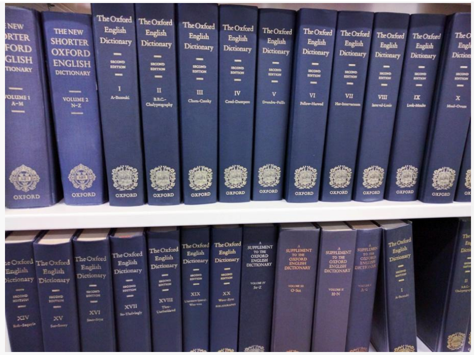
www.helsinki.fi/varieng/CoRD/corpora/CEEC/

RESEARCH QUESTIONS

1. **Who** are the innovators? Which social groups do they represent?
2. **How do the new words spread** socially, geographically and diachronically?
3. **Which semantic domains** do the neologisms represent?
4. **Why** are the neologisms created and established? Can they be linked to:
   - Specific historical events?
   - Changes in culture & society?
   - Social meanings?

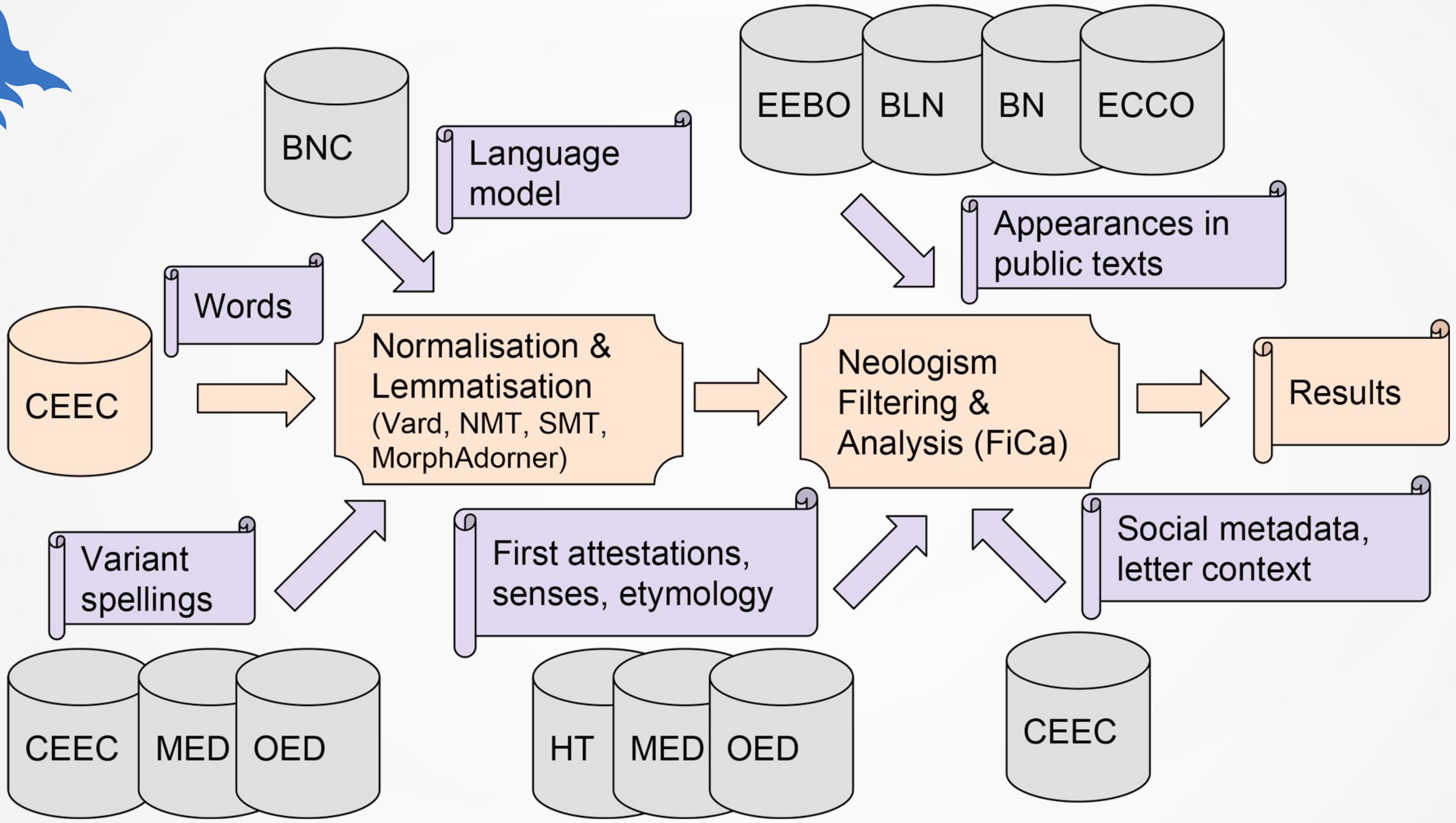
A general View of the City of LONDON and the River Thames | Vue Generale de la Ville de LONDRES et de la Tamise

BIG-DATA APPROACH TO ANALYSING NEOLOGISMS



- Automatically map each word in the corpus to lexicographical data and contemporary published texts, compare first attestation dates
  - Spelling variation: SCEEC not enough, additional **normalization** required
- Automatic retrieval of related **lexicographical data**
  - *Oxford English Dictionary* (OED), *Historical Thesaurus* (HT), *Middle English Dictionary* (MED)
- Automatic retrieval of data from **databases of contemporary published texts**
  - *Early English Books Online* (EEBO), *Eighteenth Century Collections Online* (ECCO), *British Library Newspapers* (BLN), *Burney & Nichols Collections* (BN)
- **Interface** for pruning the possible neologisms found, exploring social factors

CURRENT PIPELINE


EEBO
BLN
BN
ECCO
BNC
Language model
Appearances in public texts
Words
CEEC
Normalisation & Lemmatisation (Vard, NMT, SMT, MorphAdorner)
Neologism Filtering & Analysis (FiCa)
Results
Variant spellings
First attestations, senses, etymology
Social metadata, letter context
CEEC
MED
OED
HT
MED
OED
CEEC

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

INITIAL MAPPING TO OED

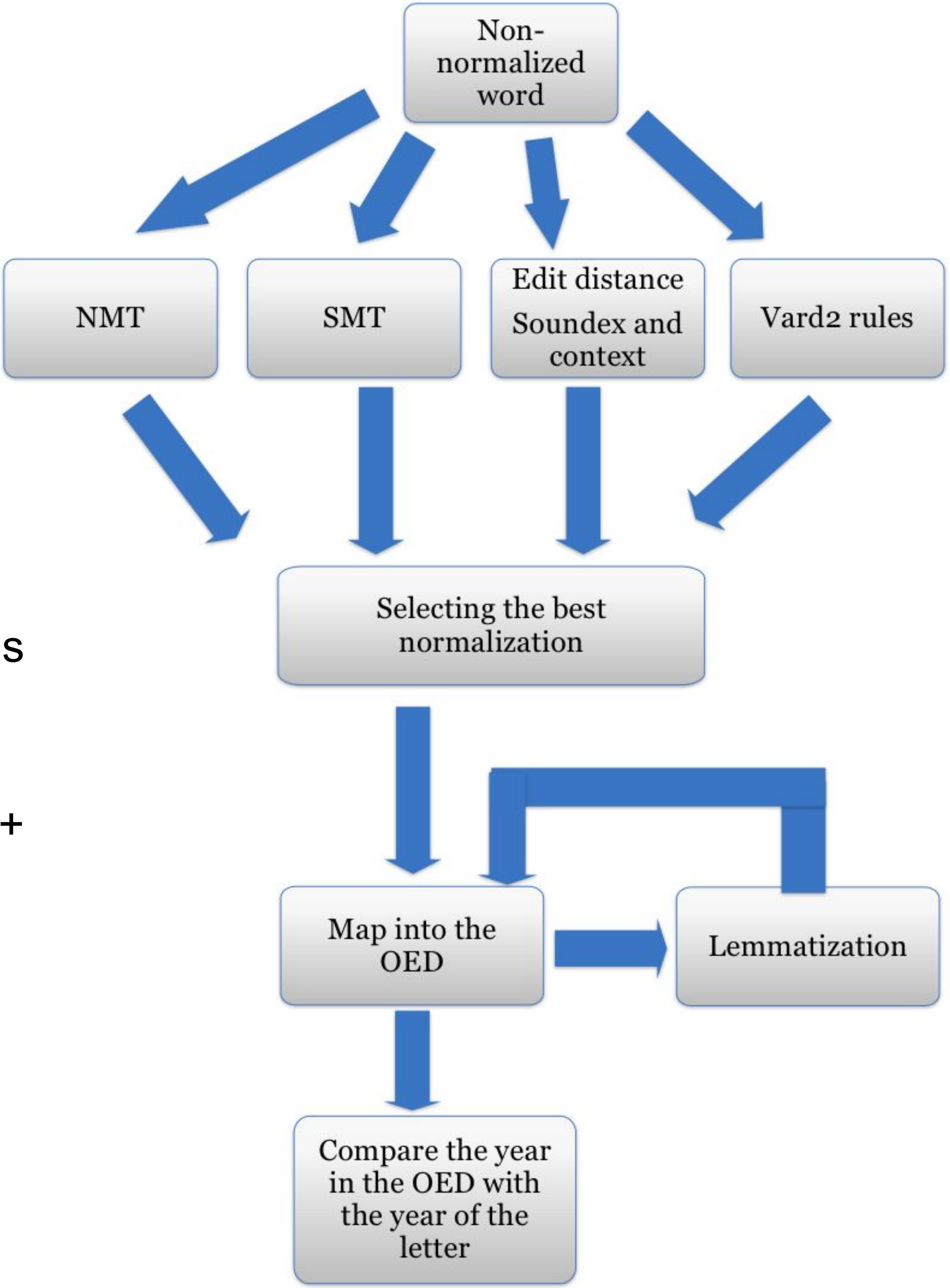
- **Prepare corpus**: convert to Unicode, remove most punctuation, tokenize
- Attempt to **lemmatize** with NLTK (based on Princeton WordNet)
- **Map** lemmas to OED (local JSON version)

- Extend already performed VARD2 normalizations to 15th century
- Use MorphAdorner to automatically **normalize further**
- **Map** again
- Successfully mapped: c. 50,000 word forms, unmapped: c. **100,000**

ADDITIONAL NORMALIZATION

(HÄMÄLÄINEN ET AL. 2018)

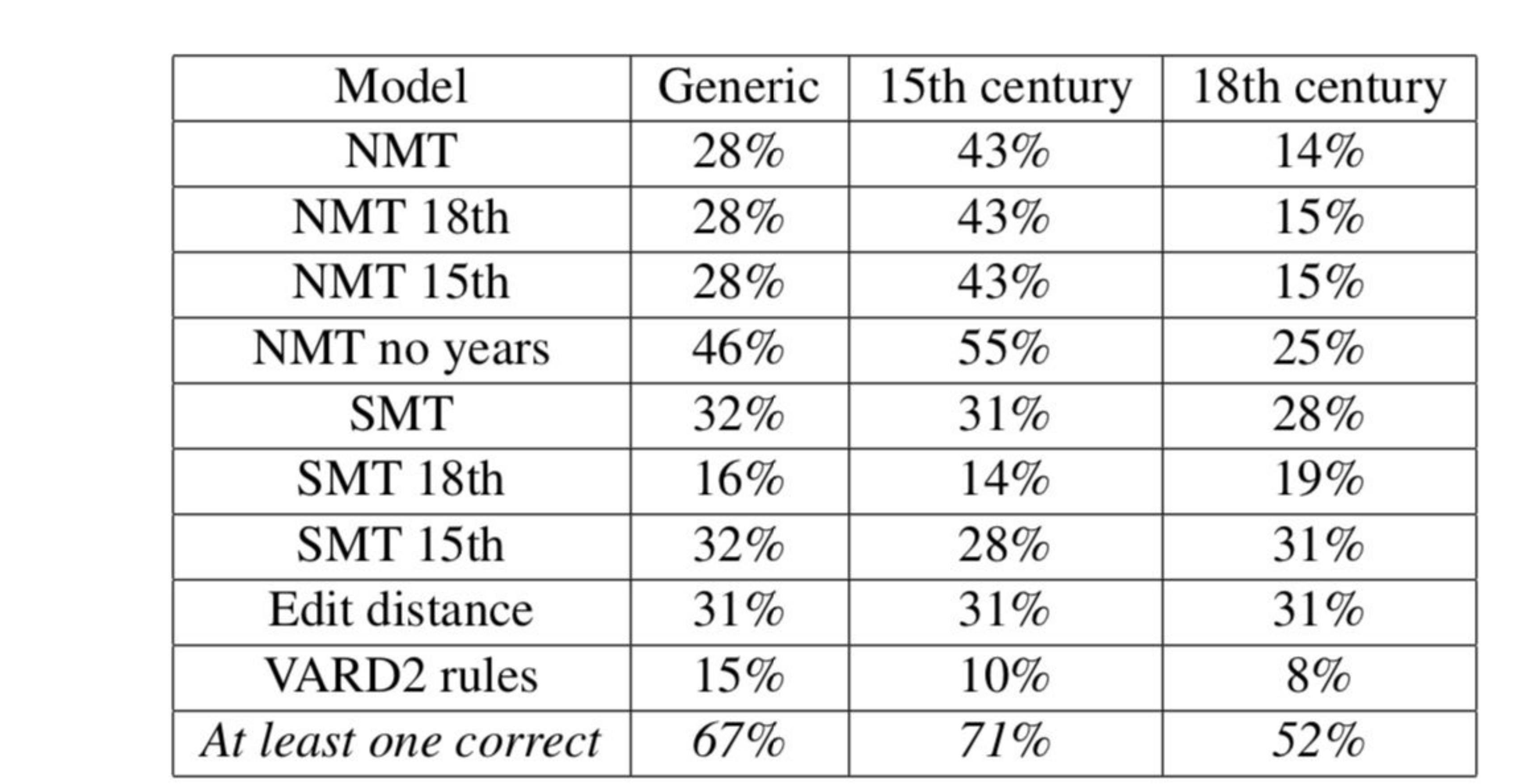
- Idea: use machine translation!
  - **NMT** = neural machine translation (OpenNMT)
  - **SMT** = statistical machine translation (Moses)
  - Use known VARD2 normalizations, MED and OED as input
  - Character-based; language model = BNC
- Levenshtein **edit distance** + filter by semantic similarity + Soundex pronunciation by edit distance
- Extend **VARD2** normalization rules to all words
  - 58 rules, e.g. “u → v anywhere”



EXAMPLES FROM THE 18TH CENTURY

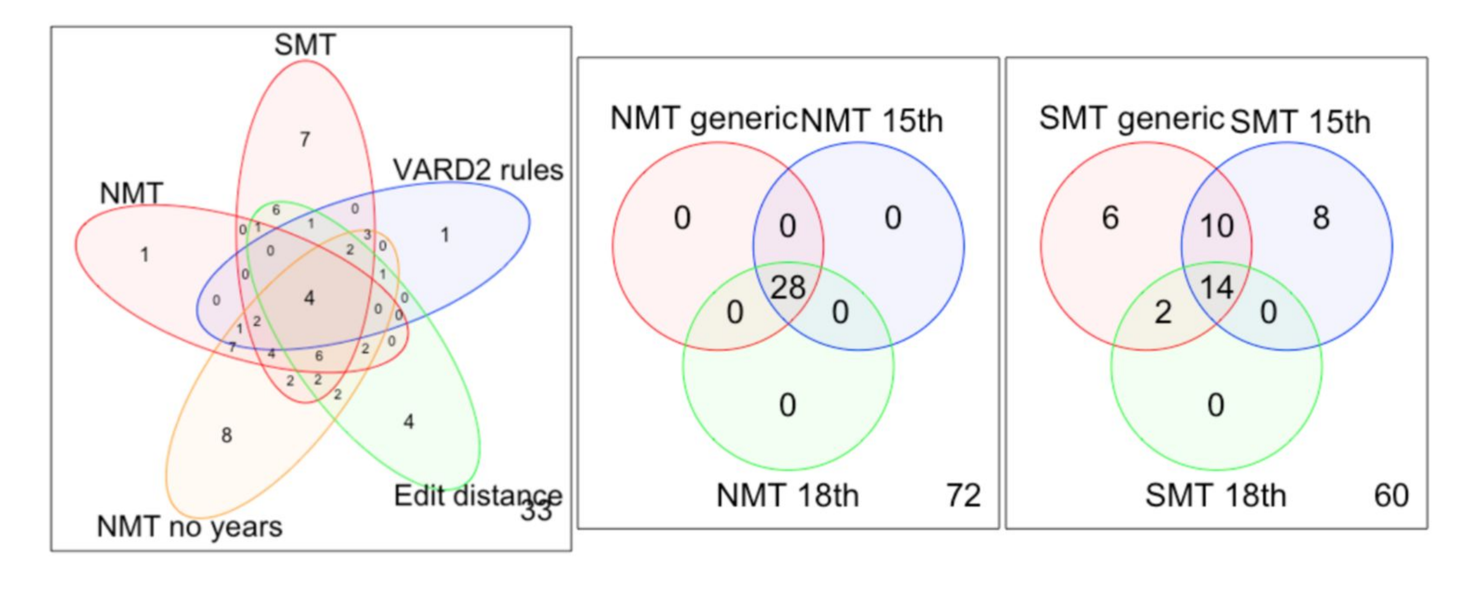
Word	NMT	SMT	Edit distance	VARD2	Correct
elyzian	elizian	elysian	elysian		Elysian
arch-type	archedip	archetype	archetype		archetype
kindelyer	kindler	kinder	kindlier		kindlier
supp'd	supped	supply	supply	cupped	supped
prospity	prosperity	prosperity	prosperity		prosperity
affectionett	affectiont	affection	affectionate		affectionate
driction	driction	diction	direction		direction
octvo	ctuoctoo	octal	oct, oct.		octavo
mic'lemas	micelemas	miles	micklemas		Michaelmas
midetrenian	midetrenian	mid-on			Mediterranean

Results



Model	Generic	15th century	18th century
NMT	28%	43%	14%
NMT 18th	28%	43%	15%
NMT 15th	28%	43%	15%
NMT no years	46%	55%	25%
SMT	32%	31%	28%
SMT 18th	16%	14%	19%
SMT 15th	32%	28%	31%
Edit distance	31%	31%	31%
VARD2 rules	15%	10%	8%
At least one correct	*67%*	*71%*	*52%*

Overlap


SMT
7
NMT
VARD2 rules
1
1
6
0
0 1
1
3
0
0
2
0
0
1
4
0
0
0
0
2
1
7
4
6
2
0
2
2
2
8
4
Edit distance
33
NMT no years
NMT generic
NMT 15th
0
0
0
28
0
0
0
NMT 18th
72
SMT generic
SMT 15th
6
10
8
14
2
0
0
SMT 18th
60

SELECTING THE BEST NORMALIZATION

- **Voting**: the normalization with the most votes from the different methods wins
- **Weighted voting**: methods are weighted based on their accuracy in a test set
- **Markov chain** trained on the BNC

For the 18th-century test set:

- Cases in which at least 1 normalization method is correct: 52%
- Accuracy of methods of picking the best normalization: 22–24% ☹
  - Not good enough to be used yet
  - We are dealing with the most difficult cases; easy cases dealt with in the initial mapping

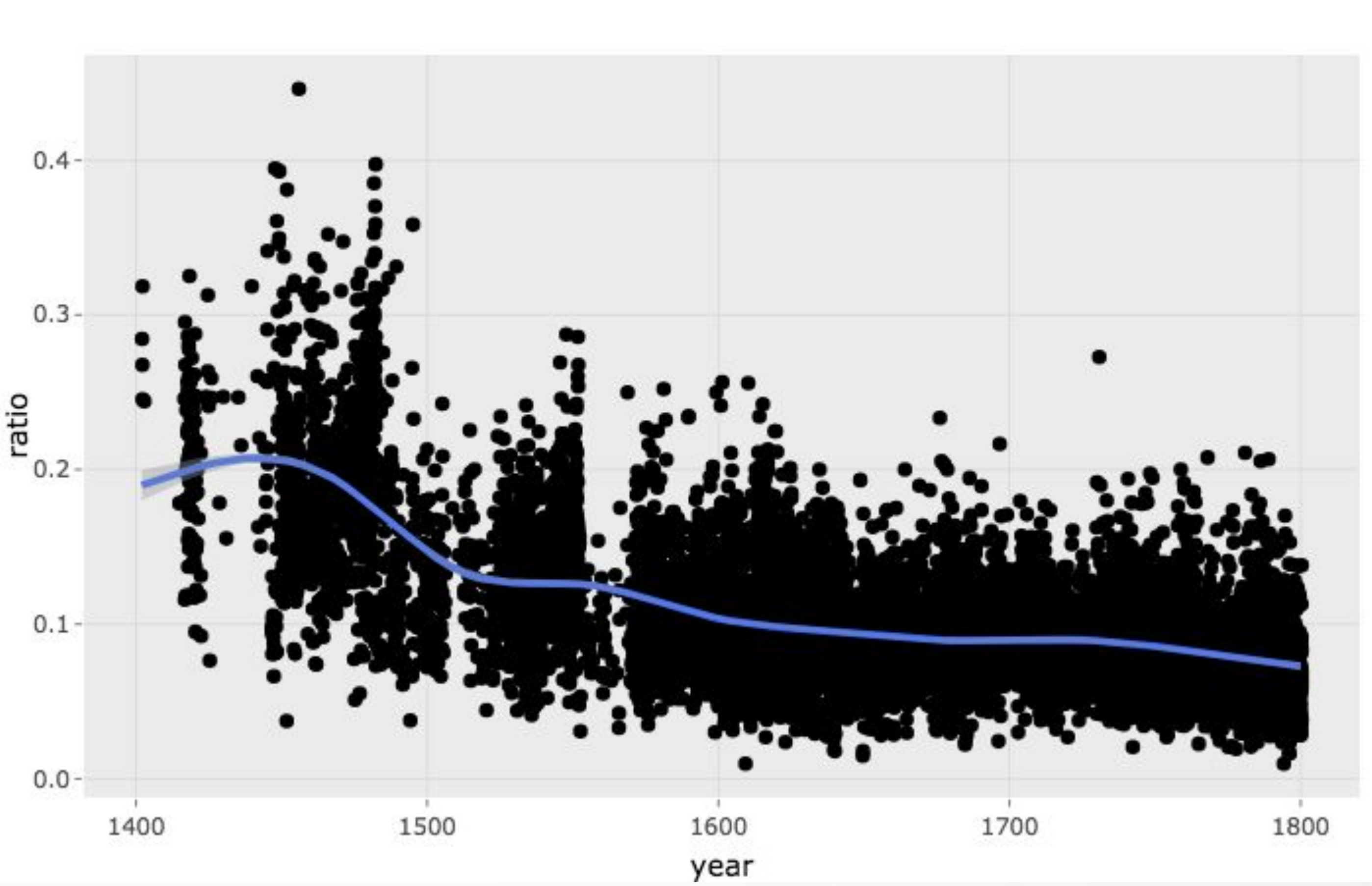
RETRIEVING RELATED DATA

- **OED**: all fields except pronunciation
- **MED**: years, link to OED
- **EEBO**, **ECCO**, **BLN**, **BN**: term frequency & document frequency before & after CEEC 1st attestation
- **CEEC**: metadata on letter, sender, recipient

- Scripts written by Mika Hämäläinen
- Data currently collected in a massive Excel file…

FILTERING STATISTICS

- 151 210 distinct word forms originally
- Proportion of words that aren't in dictionaries decreases through time (= spelling standardizes)
- 34 352 distinct word forms first seen in the long 18th century
- 31 659 of these are not tagged foreign, proper or superscript



FILTERING STATISTICS

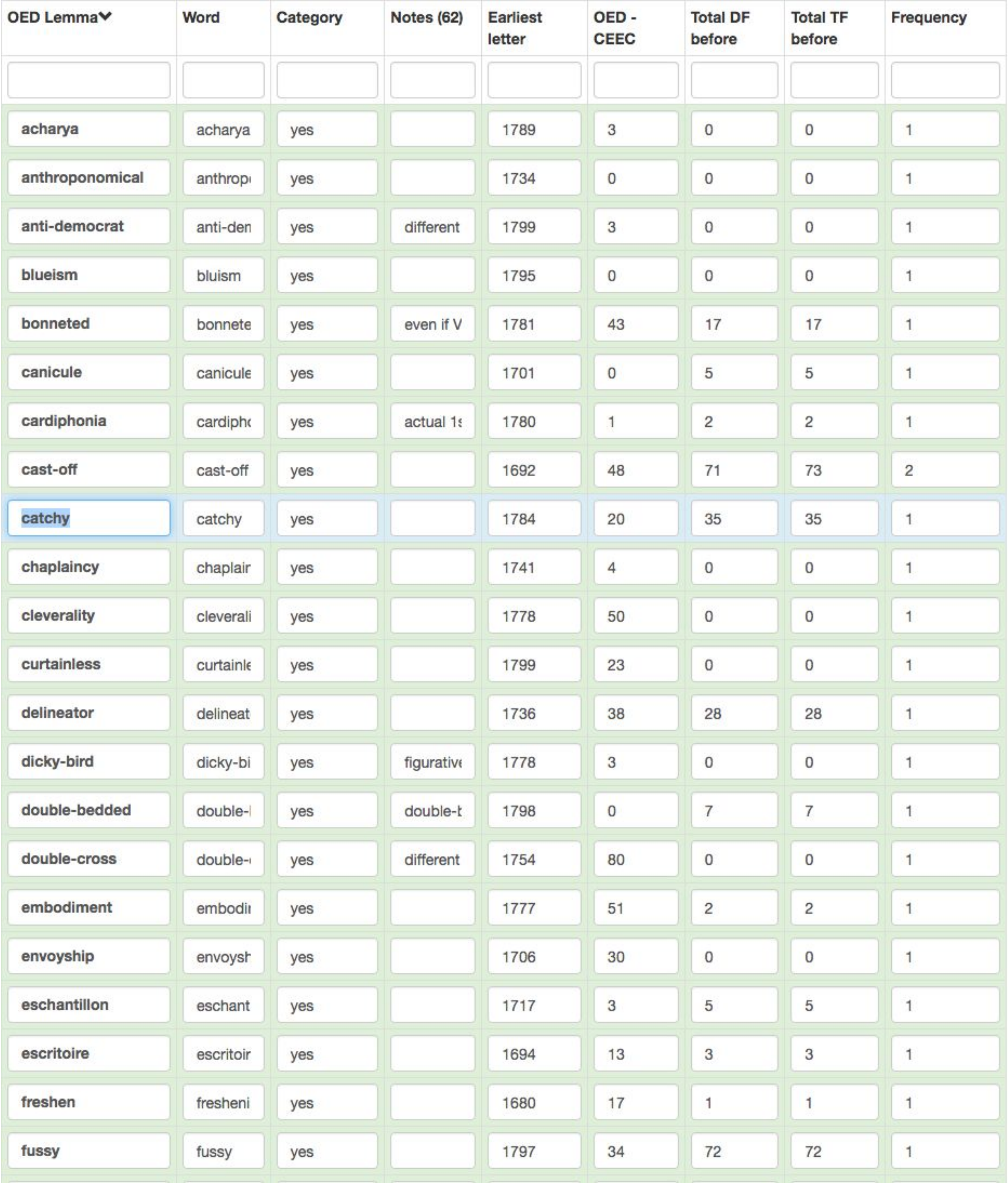
- Of the 31 659, 8 813 directly match a dictionary word
- 1 297 more can additionally be normalized to a dictionary word
- Out of these 10 110, 487 appear in the letter corpus earlier than or in the same year as their earliest dictionary attestation

On the other hand:

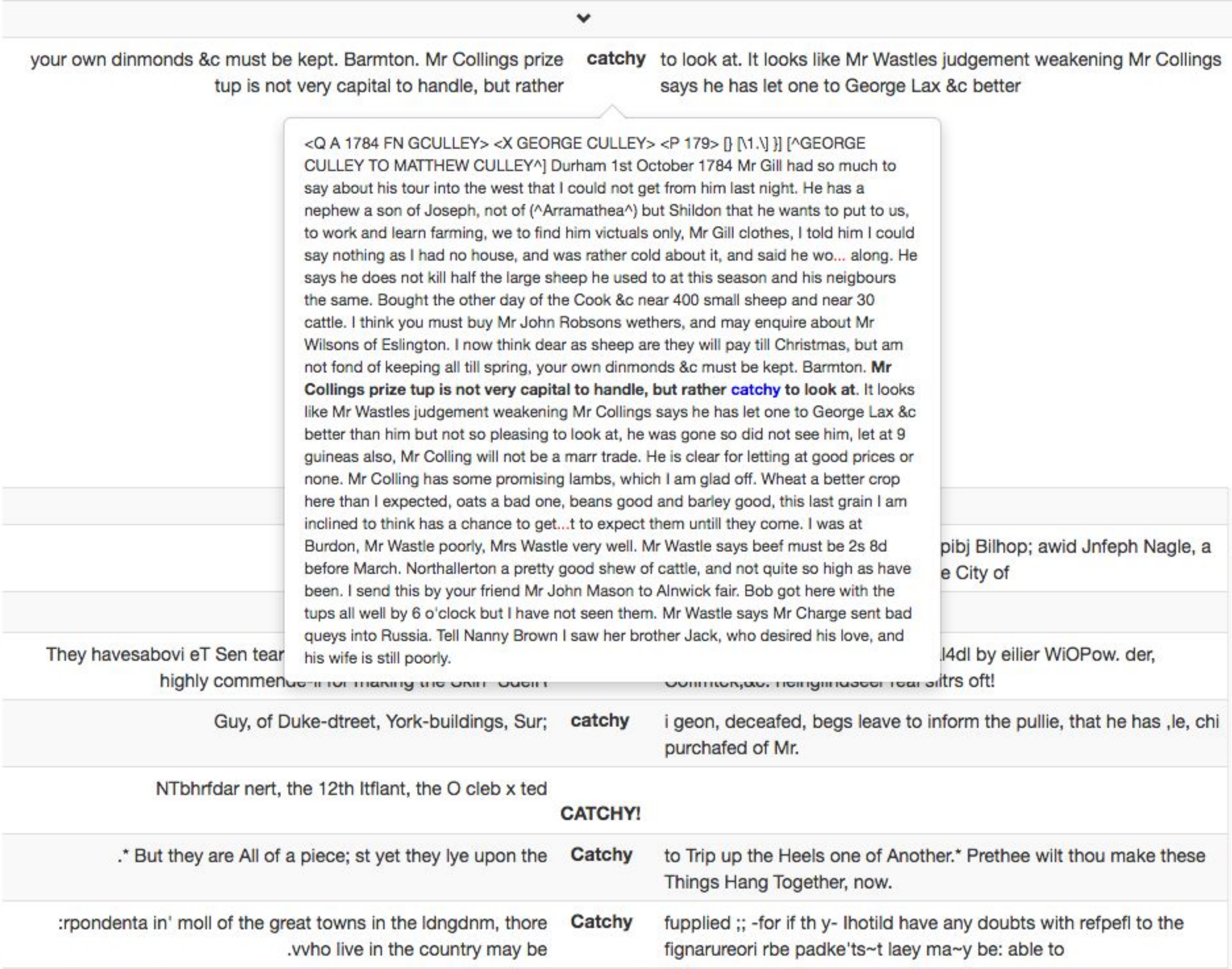
- Out of the 21 549 that couldn’t be mapped to a dictionary word, 12 245 appear in 100 or less documents in the comparison corpora before their first letter appearance
- Out of those, 2 540 appear in 100 or more documents later
- Overall, out of the 21 549, 9 115 appear in 100 or less documents in the comparison corpora at any time

INTERFACE FOR FILTERING THE NEOLOGISM CANDIDATES

- FiCa (Filtering and Categorization)
- Developed by Eetu Mäkelä; see Säily et al. (in press)



OED Lemma	Word	Category	Notes (62)	Earliest letter	OED - CEEC	Total DF before	Total TF before	Frequency
acharya	acharya	yes		1789	3	0	0	1
anthroponomical	anthrop	yes		1734	0	0	0	1
anti-democrat	anti-den	yes	different	1799	3	0	0	1
blueism	bluism	yes		1795	0	0	0	1
bonneted	bonnete	yes	even if V	1781	43	17	17	1
canicule	canicule	yes		1701	0	5	5	1
cardiphonia	cardiphc	yes	actual 1s	1780	1	2	2	1
cast-off	cast-off	yes		1692	48	71	73	2
catchy	catchy	yes		1784	20	35	35	1
chaplaincy	chaplair	yes		1741	4	0	0	1
cleverality	cleverali	yes		1778	50	0	0	1
curtainless	curtainle	yes		1799	23	0	0	1
delineator	delineat	yes		1736	38	28	28	1
dicky-bird	dicky-bi	yes	figurative	1778	3	0	0	1
double-bedded	double-l	yes	double-t	1798	0	7	7	1
double-cross	double-i	yes	different	1754	80	0	0	1
embodiment	embodii	yes		1777	51	2	2	1
envoyship	envoysh	yes		1706	30	0	0	1
eschantillon	eschant	yes		1717	3	5	5	1
escritoire	escritoir	yes		1694	13	3	3	1
freshen	fresheni	yes		1680	17	1	1	1
fussy	fussy	yes		1797	34	72	72	1



your own dinmonds &c must be kept. Barmton. Mr Collings prize tup is not very capital to handle, but rather **catchy** to look at. It looks like Mr Wastles judgement weakening Mr Collings says he has let one to George Lax &c better

<Q A 1784 FN GCULLEY> <X GEORGE CULLEY> <P 179> [} [\1.\] }] [^GEORGE CULLEY TO MATTHEW CULLEY^] Durham 1st October 1784 Mr Gill had so much to say about his tour into the west that I could not get from him last night. He has a nephew a son of Joseph, not of (^Arramathea^) but Shildon that he wants to put to us, to work and learn farming, we to find him victuals only, Mr Gill clothes, I told him I could say nothing as I had no house, and was rather cold about it, and said he wo... along. He says he does not kill half the large sheep he used to at this season and his neigbours the same. Bought the other day of the Cook &c near 400 small sheep and near 30 cattle. I think you must buy Mr John Robsons wethers, and may enquire about Mr Wilsons of Eslington. I now think dear as sheep are they will pay till Christmas, but am not fond of keeping all till spring, your own dinmonds &c must be kept. Barmton. **Mr Collings prize tup is not very capital to handle, but rather catchy to look at**. It looks like Mr Wastles judgement weakening Mr Collings says he has let one to George Lax &c better than him but not so pleasing to look at, he was gone so did not see him, let at 9 guineas also, Mr Colling will not be a marr trade. He is clear for letting at good prices or none. Mr Colling has some promising lambs, which I am glad off. Wheat a better crop here than I expected, oats a bad one, beans good and barley good, this last grain I am inclined to think has a chance to get...t to expect them untill they come. I was at Burdon, Mr Wastle poorly, Mrs Wastle very well. Mr Wastle says beef must be 2s 8d before March. Northallerton a pretty good shew of cattle, and not quite so high as have been. I send this by your friend Mr John Mason to Alnwick fair. Bob got here with the tups all well by 6 o'clock but I have not seen them. Mr Wastle says Mr Charge sent bad queys into Russia. Tell Nanny Brown I saw her brother Jack, who desired his love, and his wife is still poorly.

pibj Bilhop; awid Jnfeph Nagle, a e City of

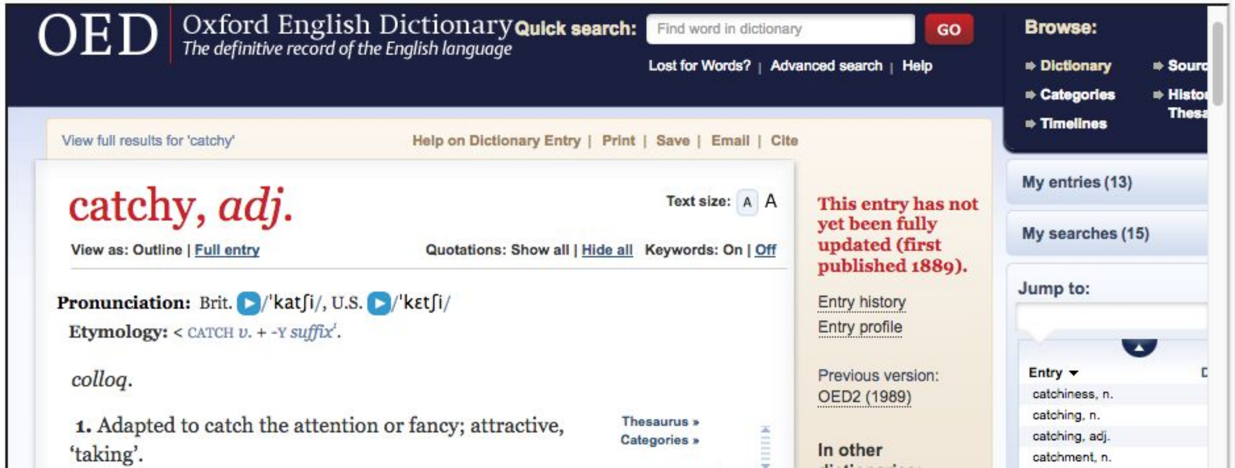
They havesabovi eT Sen tear highly commende ... l4dl by eilier WiOPow. der, ... ltrs oft!

Guy, of Duke-dtreet, York-buildings, Sur; **catchy** i geon, deceafed, begs leave to inform the pullie, that he has ,le, chi purchafed of Mr.

NTbhrfdar nert, the 12th Itflant, the O cleb x ted **CATCHY!**

.* But they are All of a piece; st yet they lye upon the **Catchy** to Trip up the Heels one of Another.* Prethee wilt thou make these Things Hang Together, now.

:rpondenta in' moll of the great towns in the ldngdnm, thore .vvho live in the country may be **Catchy** fupplied ;; -for if th y- lhotild have any doubts with refpefl to the fignarureori rbe padke'ts~t laey ma~y be: able to



catchy, ***adj.***

View as: Outline | Full entry   Quotations: Show all | Hide all   Keywords: On | Off

This entry has not yet been fully updated (first published 1889).

Previous version: OED2 (1989)

Pronunciation: Brit. /ˈkatʃi/, U.S. /ˈkɛtʃi/

Etymology: < CATCH *v.* + -Y *suffix*¹.

colloq.

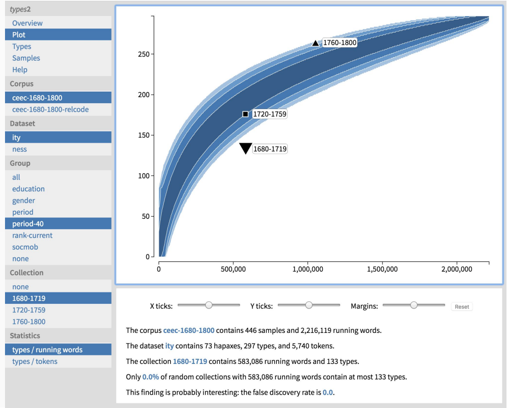
1. Adapted to catch the attention or fancy; attractive, 'taking'.

In other

INTERFACE FOR ANALYSING SOCIAL ASPECTS OF NEOLOGISMS

- Currently: combination of FiCa and Excel

- Need to develop an all-in-one interface that also provides visualizations and statistical analyses



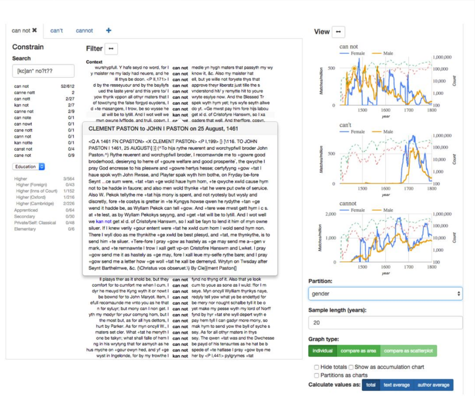
The corpus ceec-1680-1800 contains 446 samples and 2,216,119 running words.

The dataset ity contains 73 hapaxes, 297 types, and 5,740 tokens.

The collection 1680-1719 contains 583,086 running words and 133 types.

Only 0.0% of random collections with 583,086 running words contain at most 133 types.

This finding is probably interesting: the false discovery rate is 0.0.



Partition: gender

Sample length (years): 20

Graph type:

□ Hide totals □ Show as accumulation chart

□ Partitions as charts

Calculate values as:

PILOT STUDY

Social aspects of 18th-century neologisms

18TH-CENTURY NEOLOGISMS IN THE CEEC

- CEEC, long 18th century (1680–1800)
  - 315 writers, 4,945 letters, 2.2 million words
- Criteria:
  - CEEC 1st attestation ≤ OED 1st attestation
  - Can occur in max 100 contemporary published texts before CEEC 1st attestation
- Automated procedures → only 220 candidates for human to filter in interface
  - 81 neologisms found

WHO ARE THE INNOVATORS?

- Surprisingly many neologisms compared to number of running words:
  - Twining (13), clergyman
  - Austen, Bentham, Gray (4), authors
  - Jackson, St. Michel (2)
- Surprisingly few:
  - Lady Mary Wortley Montagu (1)
- Social networks:
  - Twining & Burneys
  - Jackson & St. Michel (related to Pepys)

Name	# of neologisms
Twining, Thomas	13
Austen, Jane	4
Bentham, Jeremy	4
Burney, Frances	4
Gray, Thomas	4
Cowper, William	3
Lennox, Sarah	3
Burney, Charles	2
Jackson, John	2
St. Michel, Balthasar	2
…	2

EXAMPLES

I must not omit acquainting you, Sir, That upon Opening his Body (which the **uncommonness** of his Case required of us, for our own satisfaction as well as Publick Good) there was found in his Left Kidney a nest of no less than 7 stones, of the most irregular Figures your imagination can frame, and weighing together 4 1/2 ounces …

(John Jackson to John Evelyn, 1703; OED 1705)

Thomas Twining and Charles Burney

(© NPG; NPG D39475, NPG 3884)

But I don’t recollect any single word in our language but *Tune* that expresses the *intune-ness* of an interval. *Intonation* is rather more scientific & **jargonic** than I like.

(Thomas Twining to Charles Burney, 1781; OED 1819)

(*intune-ness* not in OED!)

SOCIAL RANKS

- Surprisingly many neologisms compared to number of running words:
  - Lower clergy (due to Twining)
  - **Other non-gentry** (lowest category)
    - John Jackson (2), farmer's son, upwardly mobile
    - Ignatius Sancho (2), son of a slave, upwardly mobile
    - George Culley (1), farmer
- Surprisingly few:
  - Royalty

Social rank	# of neologisms
Professional	26
Clergy, lower	22
Gentry, lower	13
Nobility	6
Gentry, upper	6
Other non-gentry	5
Clergy, upper	1
Merchant	1
Royalty	1

EXAMPLES

Ignatius Sancho & Princess Elizabeth

(Wikimedia Commons)

Mr Collings prize tup is not very capital to handle, but rather **catchy** to look at.

(George Culley to Matthew Culley, 1784; OED 1831)

We hope he is well, and enjoys this fine weather unplagued by flies, and **unbitten** by fleas.

(Ignatius Sancho to Mrs. Cocksedge, 1779; OED 1794)

Our dear mother is well but hurried, my sister very **fussy** & agitated, the rest of the family in full trim though *heart full* from the thoughts of so soon being seperated, with laughing faces to keep up one another's spirits.

(Elizabeth Hanover to the Prince of Wales, 1797; OED 1831)

GENDER, AGE, EDUCATION; REGISTER

- More data from men, more neologisms by men
  - More advanced statistics needed
- **Fewer** neologisms by **younger** (10–29) and **less well educated** people (secondary / apprentice)
  - Previous research on Dutch: highest lexical productivity among highly educated older men (Keune 2012)

- Surprisingly **many** neologisms in **letters to close friends**, fewer to nuclear family members
  - Consistent with “bulge theory” (Wolfson 1990); less stable relationship triggers more creative language use (cf. Säily 2018, -*ity*)

SEMANTICS

- **People:** emotion, mental capacity, attention & judgement, behaviour, manner of action
  - *ill-natured*, *cleverality*, *nidgetty*, *missish*, *fussy*
- **Society:** communication, trade, work, faith, authority
  - *escritoire*, *knick-knackatory*, *wagon-way*, *chaplaincy*, *envoyship*
- **World:** action, space
  - *godsend*, *unstow*

# of neologisms per HT category

6	the world » action or operation
6	the mind » emotion
5	society » communication
4	the world » space
4	the mind » mental capacity
4	the mind » attention and judgement
3	society » trade and finance
3	society » occupation and work
3	society » faith
3	society » authority

…

CONCLUSION

- **Big-data approach**: massive databases & automated pipeline → possible to quickly discover dozens of neologisms in millions of words of running text
- What are we **missing**? Homonyms, zero derivation, MWUs, …
  - Spelling variation still an issue, could disproportionately affect lower ranks
  - Actual 1st attestations? Are these innovators or just early adopters?
- **Future** work:
  - Expand analysis to all research questions, entire time period;
    also analyse words not in the OED, check instances in contemporary published texts
  - Improve normalization & pipeline;
    develop methods & interface for analysing social aspects & spread of neologisms

REFERENCES

- Hämäläinen, Mika, Tanja Säily, Jack Rueter, Jörg Tiedemann & Eetu Mäkelä (2018). “Normalizing early English letters to Present-day English spelling”. Beatrice Alex, Stefania Degaetano-Ortlieb, Anna Feldman, Anna Kazantseva, Nils Reiter & Stan Szpakowicz (eds.), *Proceedings of the 2nd Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature* (ACL Anthology W18-45), 87–96. Stroudsburg, PA: Association for Computational Linguistics.
- Keune, Karen (2012). *Explaining Register and Sociolinguistic Variation in the Lexicon: Corpus Studies on Dutch*. PhD dissertation, Radboud University Nijmegen.
- Säily, Tanja (2018). “Change or variation? Productivity of the suffixes -*ness* and -*ity*”. Terttu Nevalainen, Minna Palander-Collin & Tanja Säily (eds.), *Patterns of Change in 18th-century English: A Sociolinguistic Approach* (Advances in Historical Sociolinguistics 8), 197–218. Amsterdam: John Benjamins.
- Säily, Tanja, Eetu Mäkelä & Mika Hämäläinen (in press). “Explorations into the social contexts of neologism use in early English correspondence”. *Pragmatics & Cognition*, special issue on the dynamics of lexical innovation.
- Wolfson, Nessa (1990). ”The bulge: a theory of speech behavior and social distance”. *Penn Working Papers in Educational Linguistics* 2(1). 55–83.

THANK YOU!

acharya, anthroponomical, anti-democrat, blueism, bonneted, canicule, cardiphonia, cast-off, catchy, chaplaincy, cleverality, curtainless, delineator, dicky-bird, double-bedded, double-cross, embodiment, envoyship, eschantillon, escritoire, freshen, fussy, godsend, grumpy, guimpe, hummingly, hydrogenate, idlish, impracticability, incomed, incontestably, inexact, inside-outness, internment, intrepid, jargonic, jumpable, keyless, kibitka, knick-knackatory, letteret, malformation, mariturient, mevrouw, missish, monotonous, moon, moonery, nidgetty, non-papist, pacifist, paperless, pheasantry, pushery, rishi, schoolmasterishness, scratch-back, shockingly, silentious, slushy, spidery, sprawly, squeezy, stiffish, sweet-hearted, tawdrily, trickster, truantism, unailing, unbitten, unclassed, uncommonness, undefeat, unenjoyable, uneventful, ungown, uninsured, uninteresting, unstowed, wagon-way, yester-evening

Special thanks to Oxford University Press and the Middle English Compendium for sharing their lexicographical data (OED, HT; MED)