NEW METHODS FOR ANALYSING SUFFIX COMPETITION ACROSS REGISTERS

HOW -ITY GAINED GROUND ON -NESS IN EARLY MODERN ENGLISH

Tanja Säily, Paula Rodríguez-Puente and Jukka Suomela

HELSINGIN YLIOPISTO HELSINGFORS UNIVERSITET UNIVERSITY OF HELSINKI



- Large-scale expansion of vocabulary
 - Borrowing + highly productive use of word-formation processes
 - Vernacularization, standardization (Nevalainen 1999: 358; Nevalainen & Tieken-Boon van Ostade 2006: 301–303)
- Middle English: intensive borrowing, decline of native affixal system
 - Beginning of EModE period: still a lot of freedom of choice in affix usage, parallel derivatives



RIVAL SUFFIXES: -ITY AND -NESS

- Nominal suffixes, usually derive abstract nouns from adjectives
 - E.g. productive \rightarrow productivity or productiveness
- -ness native, -ity borrowed from French (+ Latin) in Middle English
 - More sociolinguistic variation in the productivity of -ity (Säily 2014); prestige, learnedness
- Early Modern English: register variation in the use of the suffixes?



- Analysis of -ness and -ity in 17 EModE registers
 - More comprehensive than previous studies (e.g. Cowie 1998)
- Material from 3 corpora
 - Corpus of English Dialogues (CED, 1560–1760)
 - Penn-Helsinki Parsed Corpus of Early Modern English (PPCEME, 1500–1710)
 - EModE section of the Corpus of Historical English Law Reports, (CHELAR, 1535–1999)
- Hypothesis: -*ity* more learned \rightarrow more common in specialized written registers

HELSINGIN YLIOPISTO HELSINGFORS UNIVERSITET UNIVERSITY OF HELSINKI



REGISTERS STUDIED

(RODRÍGUEZ-PUENTE 2020; CF. CULPEPER & KYTÖ 2010)

		Informal		Formal
Speech-related	Speech-like	Diaries		
		Letters, private		
	Speech-based	Trial proceedings		
		Witness depositions		
	Speech-purposed	Drama		Sermons
Writing-based and writing-purposed				Bible
				Educational treatise
			(Auto)biography	History
				Law (statutes and law reports)
			Travelogue	Letters, non-private
				Medicine
				Philosophy
				Science



- -ity gains ground on -ness in all registers!
 - Change seems to start from formal written registers, spread towards speech-related ones
- Related to a general trend towards a more literate style in the 18th century?
 - Biber & Finegan (1997), McIntosh (1998: 23–24)
- More statistical evidence needed \rightarrow collaboration
 - Rodríguez-Puente, Säily & Suomela (forthcoming)



HELSINGIN YLIOPISTO

HELSINGFORS UNIVERSITET

MORPHOLOGICAL PRODUCTIVITY

- The readiness with which an element enters into new combinations (Bolinger 1948)
- Quantitative measures (e.g. Baayen 1993; Cowie & Dalton-Puffer 2002):
 - Number of different words containing the morpheme in a corpus (types)
 - Number of types occurring only once in the corpus (hapax legomena)
 - Number of types not occurring in previous periods (**new types**)
- Problem: Difficult to compare across (sub)corpora
 - Different amounts of data from different periods & registers
 - Type-based measures grow nonlinearly with corpus size → normalization not justifiable



UNIVERSITY OF HELSINKI Humanistinen tiedekunta











- Compare each subcorpus with subcorpora of equal size, randomly sampled from the corpus as a whole
- Automatically provides a measure of statistical significance

• Problems:

- Comparisons over time still difficult;
 x-axis = corpus size, not time period
- Only measures variation within a morpheme, not between morphemes



- Combining data from corpora covering partly different periods of time (+ registers)
- Corpus of English Dialogues (CED), 1560–1747
- Penn-Helsinki Parsed Corpus of Early Modern English (PPCEME), 1500–1719
- Corpus of Historical English Law Reports (CHELAR), 1544–1748



- Force time on the x-axis and see what it requires from the method
- Use a sliding window for time periods to combine the corpora
- Compare competing morphemes as if they formed a **linguistic variable**
 - Calculate proportion of -*ity* types out of all -*ity* and -*ness* types





- Corpora: CED + PPCEME
- Blue = writing-based & purposed, orange = speech-related registers
- 100-year sliding window, 25-year intervals

• Problems:

- Different amounts of data from registers → comparability?
- Statistical significance?

HELSINGFORS UNIVERSITET UNIVERSITY OF HELSINKI





• 1600–1699

HELSINGIN YLIOPISTO HELSINGFORS UNIVERSITET UNIVERSITY OF HELSINKI



FOCUS ON A SPECIFIC PERIOD...



- 1600-1699
- Return corpus size to x-axis
- Compare each register with random subcorpora of the same size sampled from the entire corpus
 - Proportion of -*ity* statistically significantly high in writing-based & purposed registers but still low in speech-related registers

HELSINGIN YLIOPISTO HELSINGFORS UNIVERSITET UNIVERSITY OF HELSINKI



DO THE SAME FOR ALL PERIODS IN BOTH REGISTERS...



HELSINGIN YLIOPISTO HELSINGFORS UNIVERSITET UNIVERSITY OF HELSINKI



HELSINGIN YLIOPISTO

HELSINGFORS UNIVERSITET UNIVERSITY OF HELSINKI

- Proportion of -*ity* increases over time
- Writing-based and purposed registers: rapid increase from late 16th century onwards
- Speech-related registers: lagging behind until late 17th century
- 1650–1749: proportion of *-ity* close to 50%
- Potentially related to a general trend towards a more literate style in the 18th century (Biber & Finegan 1997; McIntosh 1998)





EXTREME EXAMPLE: LAW



- Corpora: CED + PPCEME + CHELAR
- Proportion of -*ity* statistically significantly high in "Law" register (statutes from PPCEME, law reports from CHELAR) throughout the Early Modern English period
- Did the change start here?
- Might reflect an overall preference for Romance vocabulary in the register since Middle English



- Register a key factor in the competition of -ness and -ity in the Early Modern period
 - Rise of -ity starts in formal written registers, spreads towards speech-related ones
- New methods → more reliable & accurate results
 - Enable investigation of register, social & intralinguistic factors in competing processes
 - Method of periodization usable in any diachronic study
- Robust statistical methods tend to require a lot of data \rightarrow larger corpus needed to:
 - Analyse each individual register
 - Consider productivity based on hapax legomena or new types



Baayen, R. H. 1993. On frequency, transparency and productivity. In Geert Booij & Jaap van Marle (eds.), Yearbook of morphology 1992, 181–208. Dordrecht: Kluwer.

Biber, Douglas & Edward **Finegan**. 1997. Diachronic relations among speech-based and written registers in English. In Terttu Nevalainen & Leena Kahlas-Tarkka (eds.), *To explain the present: Studies in the changing English language in honour of Matti Rissanen*, 253–275. Helsinki: Société Néophilologique.

Bolinger, Dwight L. 1948. On defining the morpheme. Word 4. 18-23.

Cowie, Claire. 1998. *Diachronic word-formation: A corpus-based study of derived nominalizations in the history of English*. Cambridge: University of Cambridge dissertation.

Cowie, Claire & Christiane **Dalton-Puffer**. 2002. Diachronic word-formation and studying changes in productivity over time: Theoretical and methodological considerations. In Javier E. Díaz Vera (ed.), *A changing world of words: Studies in English historical lexicography, lexicology and semantics*, 410–437. Amsterdam: Rodopi.

Culpeper, Jonathan & Merja Kytö. 2010. Early Modern English dialogues: Spoken interaction as writing (Studies in English Language). Cambridge: CUP.

McIntosh, Carey. 1998. The evolution of English prose 1700–1900: Style, politeness, and print culture. Cambridge: CUP.

Nevalainen, Terttu. 1999. Early Modern English lexis and semantics. In Roger Lass (ed.), *The Cambridge history of the English language, III:* 1476–1776, 332–458. Cambridge: CUP.

Nevalainen, Terttu & Ingrid **Tieken-Boon van Ostade**. 2006. Standardisation. In Richard Hogg & David Denison (eds.), *A history of the English language*, 271–311. Cambridge: CUP.

Rodríguez-Puente, Paula. 2020. Register variation in word-formation processes: The development of *-ity* and *-ness* in Early Modern English. *International Journal of English Studies* 20(2). 145–167.

Rodríguez-Puente, Paula, Tanja Säily & Jukka Suomela. Forthcoming. New methods for analysing diachronic suffix competition across registers: How *-ity* gained ground on *-ness* in Early Modern English. Submitted to *International Journal of Corpus Linguistics*.

Säily, Tanja. 2014. Sociolinguistic variation in English derivational productivity: Studies and methods in diachronic corpus linguistics. Helsinki: Société Néophilologique.

Säily, Tanja & Jukka Suomela. 2009. Comparing type counts: The case of women, men and *-ity* in early English letters. In Antoinette Renouf & Andrew Kehoe (eds.), *Corpus linguistics: Refinements and reassessments*, 87–109. Amsterdam: Rodopi.

Säily, Tanja & Jukka Suomela. 2017. types2: Exploring word-frequency differences in corpora. In Turo Hiltunen, Joe McVeigh & Tanja Säily (eds.), Big and rich data in English corpus linguistics: Methods and explorations. Helsinki: VARIENG. <u>https://varieng.helsinki.fi/series/volumes/19/saily_suomela/</u>

HELSINGIN YLIOPISTO HELSINGFORS UNIVERSITET UNIVERSITY OF HELSINKI