

# Comparing type counts

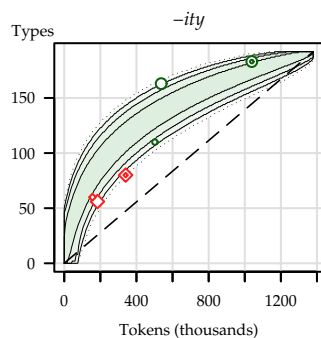
## The case of women, men and *-ity* in early English letters

Tanja Säily and Jukka Suomela

### Supplementary material

#### A Normalisation vs. permutations

Normalisation (dashed line) would not work here: it would indicate that women use *-ity* more variously than people in general



#### B Comparing hapax counts

The same method generalises to other measures of productivity

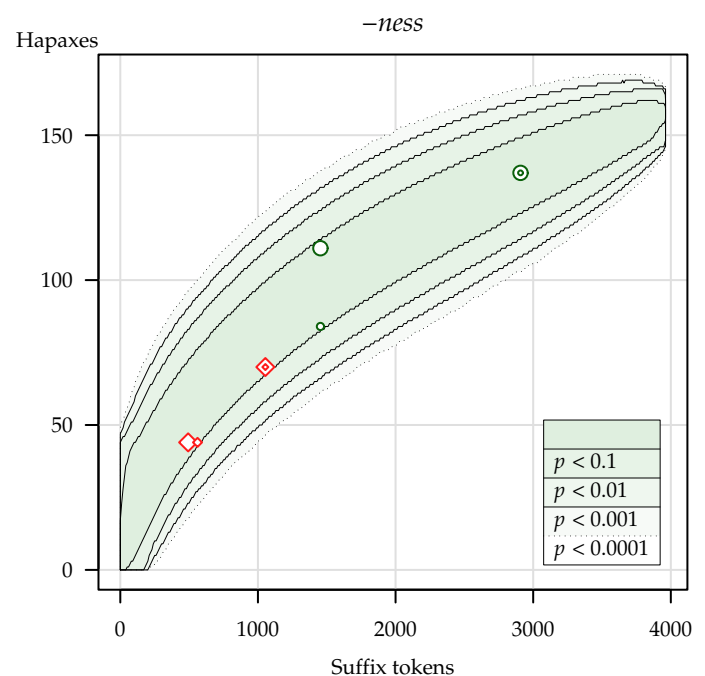
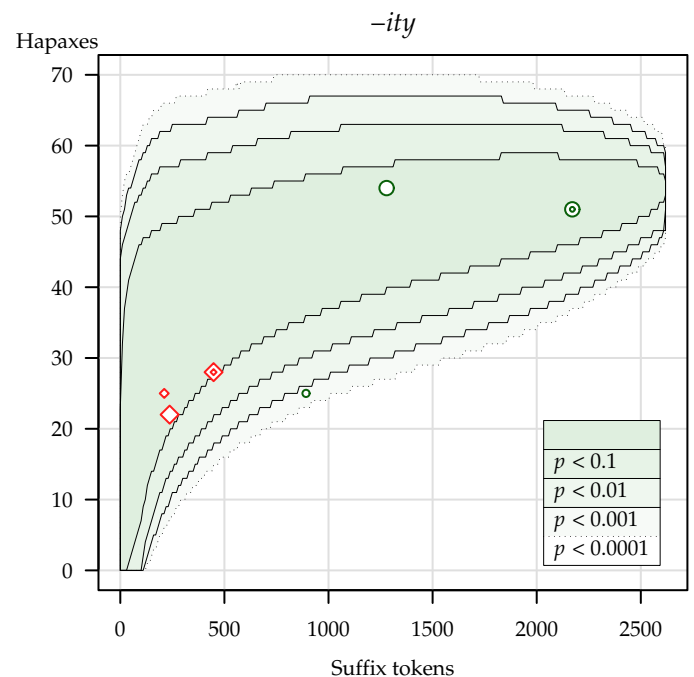
For the figures on the right, we have only counted tokens containing the suffix under study, and plotted the accumulation of hapax legomena instead of all types; cf. Baayen and Lieber's (1991) category-conditioned degree of productivity

- Examples of *-ity* hapaxes in the corpus: *instability, singularity, capability, imbecility, mutuality, absurdity*
- Examples of *-ness* hapaxes: *generousness, procliveness, shamefastness, devoutness, soberness, feverishness*

#### C Software running times

Examples of running times on a desktop PC with a 2.4-GHz Pentium 4 processor; accumulation of *-ity* types/hapaxes; 412 subcorpora and 192 *-ity* types; output resolution 277 slots:

- 1.7/2.3 seconds for 20 000 permutations
- 82/113 seconds for 1 000 000 permutations



#### Bibliography

**Baayen, R.H. and R. Lieber (1991)** Productivity and English derivation: a corpus-based study. *Linguistics* 29, pp. 801–843.

**Good, P. (2005)** *Permutation, Parametric, and Bootstrap Tests of Hypotheses*. 3rd edition. Springer Series in Statistics. Berlin: Springer-Verlag.