Analogy in suffix rivalry: the case of English *-ity* and *-ness*

24.01.2020

Introduction

The paper argues for Analogical Modelling of rivalry in the case of *-ity* vs. *-ness*, from a synchronic and diachronic perspective.

-ity and *-ness*

The distribtion of both forms is neither fully random, nor fully determined from other morphological properties:

- Both can attach to simple bases and Latinate suffixes (e.g. *-able*)
- Only *-ness* attaches to words with Germanic suffixes (e.g. *-ing*)
- *-ity* seems to be preferred in many cases

connectabil*ity*	connectable
nordic*ity*	nordic
metaphoric*ity*	metaphoric
prescriptiv*ity*	prescriptive
picayun*ity*	picayune
exciting*ness*	exciting
generic*ness*	generic
blokish*ness*	blokish
commutative*ness*	commutative
norse*ness*	norse

-ity and *-ness*

-ness is a germanic suffix while *-ity* entered into the English Language later. The question then emerges:

- How has the productivity of *-ity* and *-ness* evolved?
- How are nouns assigned to either *-ity* or *-ness*?

Data and coding

- The paper uses a hand-compiled dataset from the OED with date of first attestation.
- The dataset contained with a total of 2771 items
- Manual coding of syntactic category of the base
- Manual coding of base suffixes (or lack thereof)
- Manual coding for transparency

OED neologisms: general overview

	N -ity	N -ness	N overall
20th century:	344 (61%)	220 (39%)	564 (100%)
19th century:	733 (49%)	759 (51%)	1,492 (100%)
18th century:	306 (43%)	408 (57%)	714 (100%)

Syntactic category of the base Example words (century of first attestation)

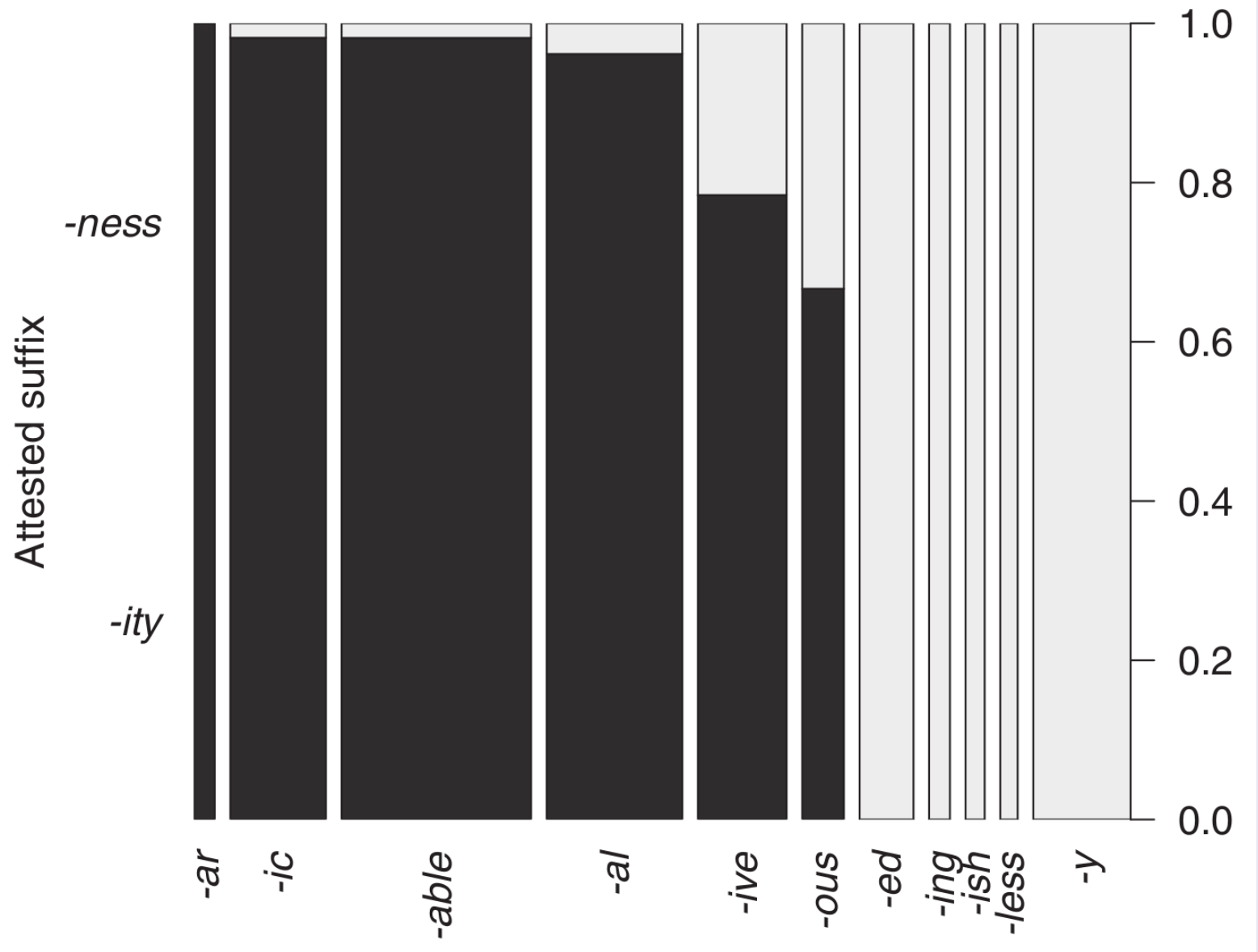
category	example
adjective	phrase lovability (19th century), nerdishness (20th century
noun	perspectivity (20th century), moneyness (20th century)
verb	relaxity (18th century), oughtness (19th century)
adverb	onceness (19th century)
preposition	betweenity (18th century), betweenness (19th century)
pronoun	I-ness (19th century)
wh-pronoun	whenness (20th century)
particle	notness (20th century)
bound form	iracundity (19th century), arity (20th century, based on the
phrase	know-nothingness (19th century)

Table 3. -ity/-ness *derivatives by syntactic category of the base, twentieth century*

	-ity derivatives		*-ness* derivatives	
Syntactic category of the base	N	%	N	%
adjective	326	94.8%	186	84.5%
noun	7	2.0%	14	6.4%
bound form	11	3.2%	0	
phrase	0		10	4.5%
minor categories (adverb, preposition, pronoun)	0		10	4.5%
Total	344	100.0%	220	100.0%

Productivity

By suffix


Attested suffix
-ness
-ity
1.0
0.8
0.6
0.4
0.2
0.0
-ar
-ic
-able
-al
-ive
-ous
-ed
-ing
-ish
-less
-y

By transparency

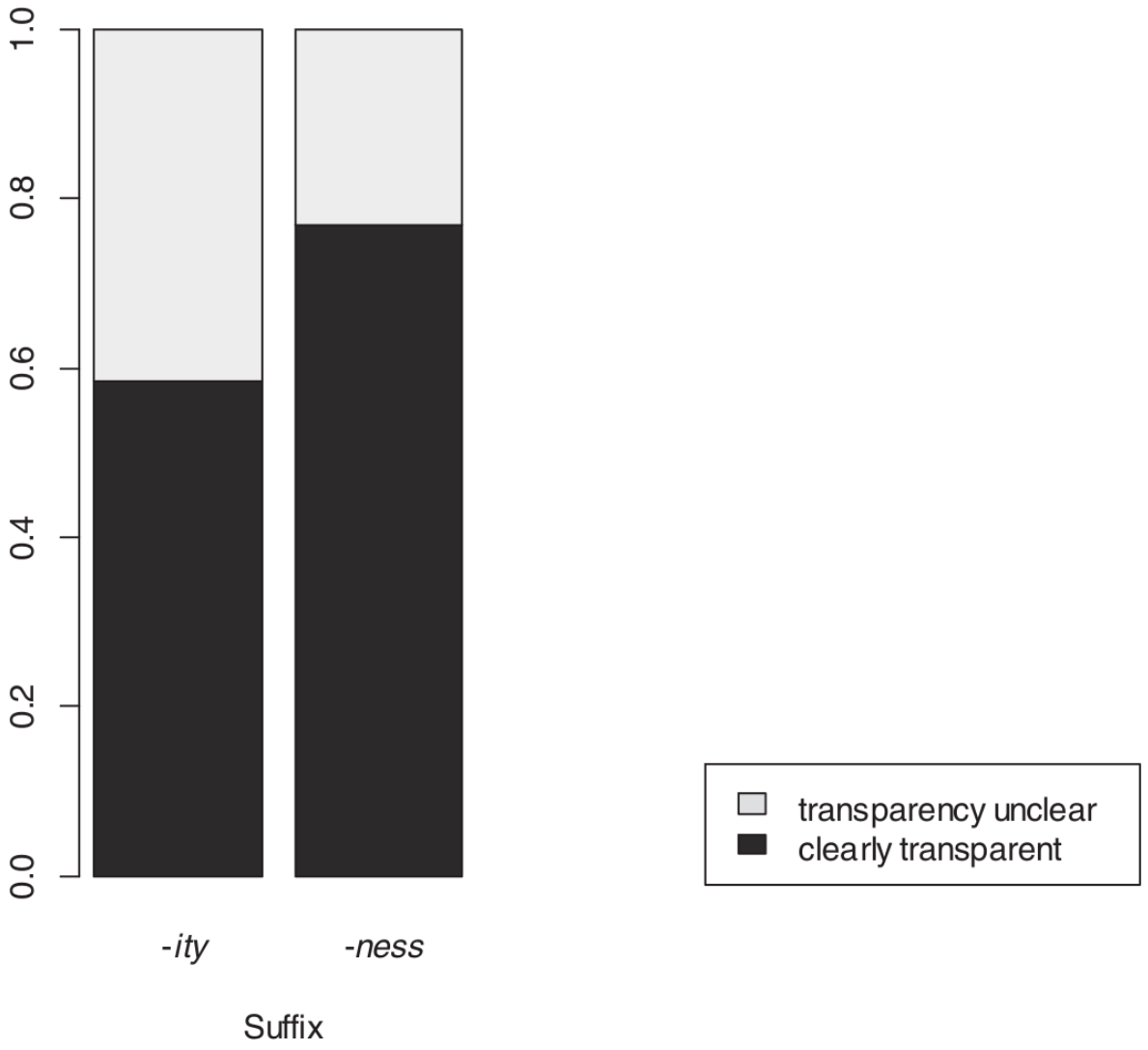


Figure 2. Transparency of *-ity* and *-ness* derivatives, twentieth century (N = 564)

Diachrony

Development

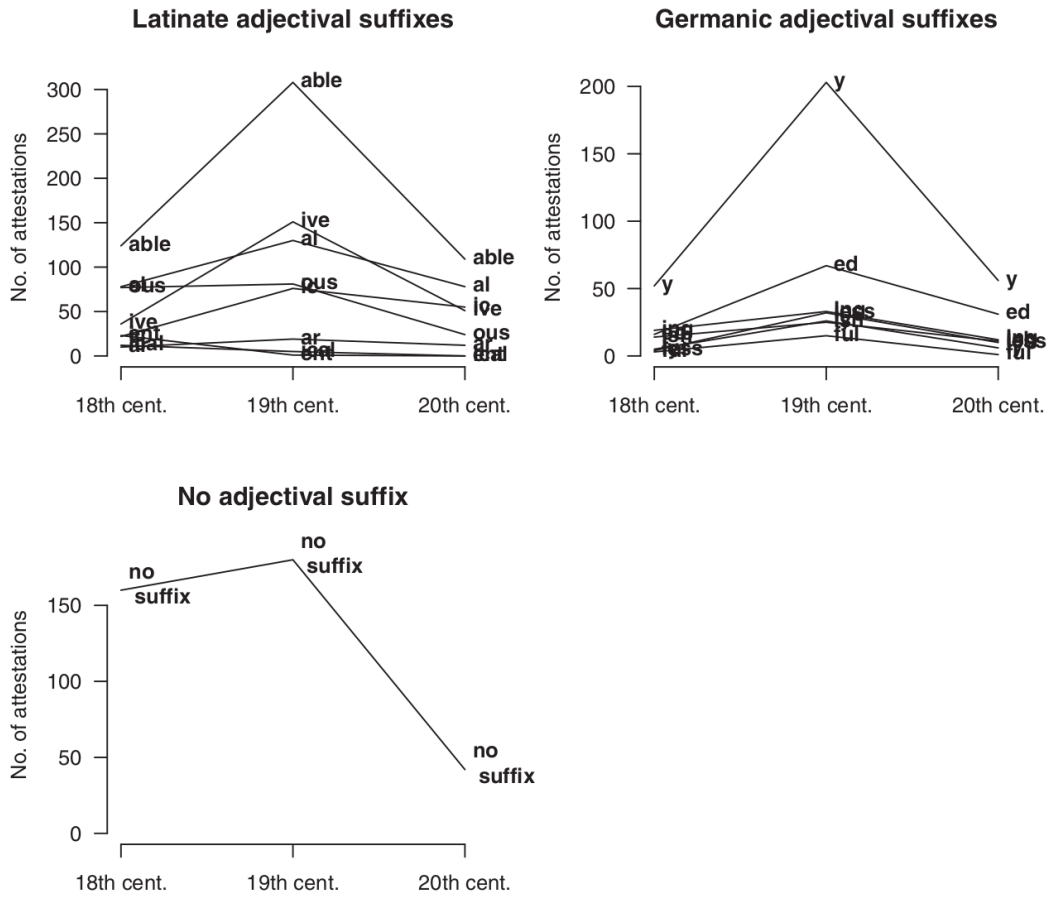


Figure 3. Morphological categories of adjectival bases, diachronically ($N_{18th\ cent.} = 673$; $N_{19th\ cent.} = 1{,}379$; $N_{20th\ cent.} = 512$)

Development

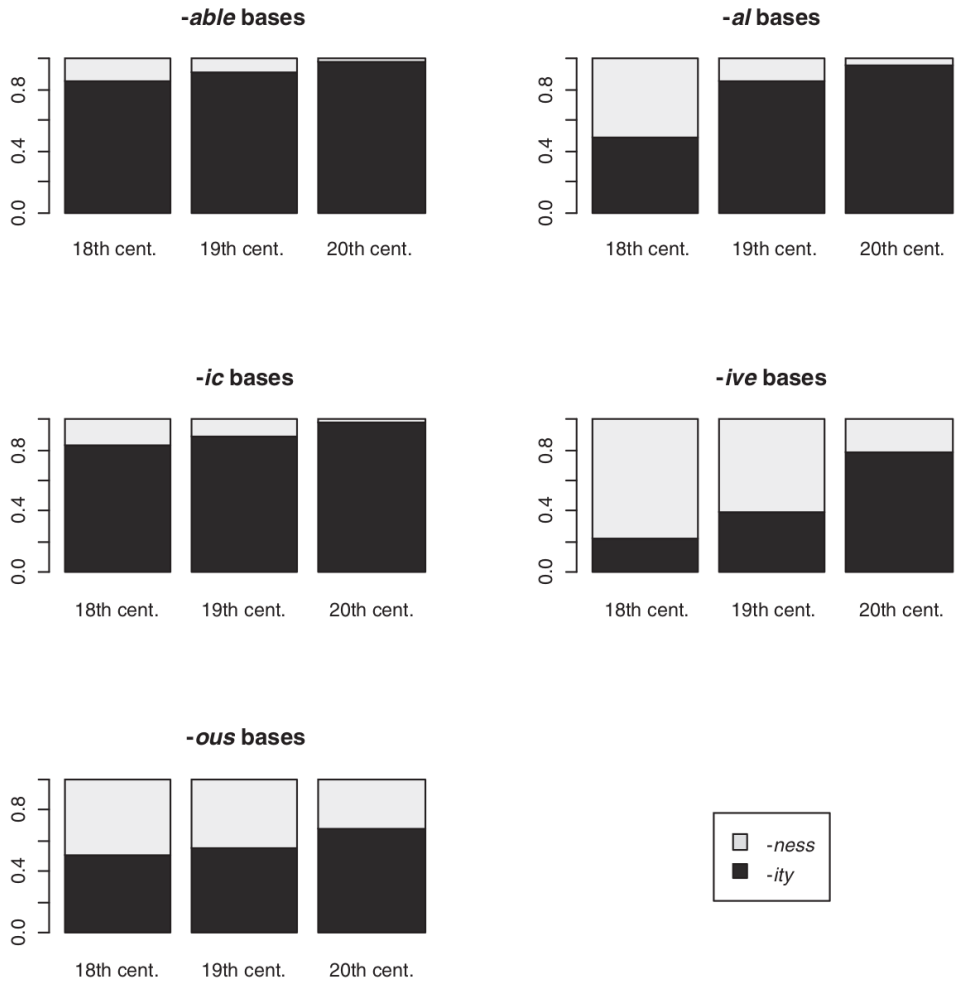


Figure 4. The distribution of *-ity* and *-ness* neologisms among frequent Latinate categories, diachronically (N = 1,425)

Analogical modelling

- new complex words are formed from their bases on the basis of existing base-derivative pairs in the mental lexicon
- analogy happens online (*)
- a specific type of exemplar-based approach (*)

The AM model

- it is a classification task
- class assignment happens on the basis of similarity
- features are hand-coded (phonology of last two syllables + syntactic information)

Issues...

- Morphological information is confused with phonological information

The AM model

Exemplars in the lexicon

item	onset-σ_{pen}	nuc-σ_{pen}	coda-σ_{pen}	onset-σ_{fin}	nuc-σ_{fin}	coda-σ_{fin}	synt	suffix
directivity	r	e	k	t	ɪ	v	word	ity
selectivity	l	e	k	t	ɪ	v	word	ity
overprotectiveness	t	e	k	t	ɪ	v	word	ness
norseness	=	=	=	n	ɔː	s	word	ness
informativeness	m	ə	=	t	ɪ	v	word	ness
normativeness	m	ə	=	t	ɪ	v	word	ness

Analogical set

item	onset-σ_{pen}	nuc-σ_{pen}	coda-σ_{pen}	onset-σ_{fin}	nuc-σ_{fin}	coda-σ_{fin}	synt	suffix
directivity	r	e	k	t	ɪ	v	word	ity
selectivity	l	e	k	t	ɪ	v	word	ity
overprotectiveness	t	e	k	t	ɪ	v	word	ness

New word: *perspective* + -x?

onset-σ_{pen}	nuc-σ_{pen}	coda-σ_{pen}	onset-σ_{fin}	nuc-σ_{fin}	coda-σ_{fin}	synt
sp	e	k	t	ɪ	v	word

Figure 8. The general architecture of an analogical model

The AM model

The crucial feature that distinguishes AM from many other pertinent models is that the degree of similarity that is relevant for exemplars to be included in the analogical set is decided for each new word individually. The rationale that underlies the procedure is that while the model will always incorporate maximally similar items, items with lower degrees of similarity will be incorporated only if that incorporation does not lead to greater uncertainty with respect to the classification task.

Set up

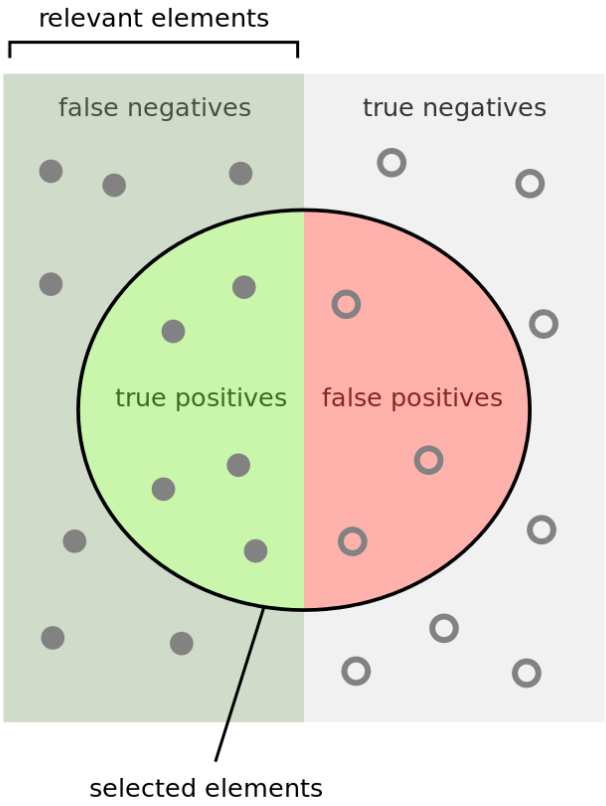
Two kinds of experiment:

- Synchronic: LOO-CV
- Diachronic: train on one century, test on the next century

Some issues:

- are the items true neologisms?
- are the lexicons really representative of speakers' lexicons?

F score



How many selected items are relevant?

$$\text{Precision} = \frac{\text{true positives}}{\text{selected elements}}$$

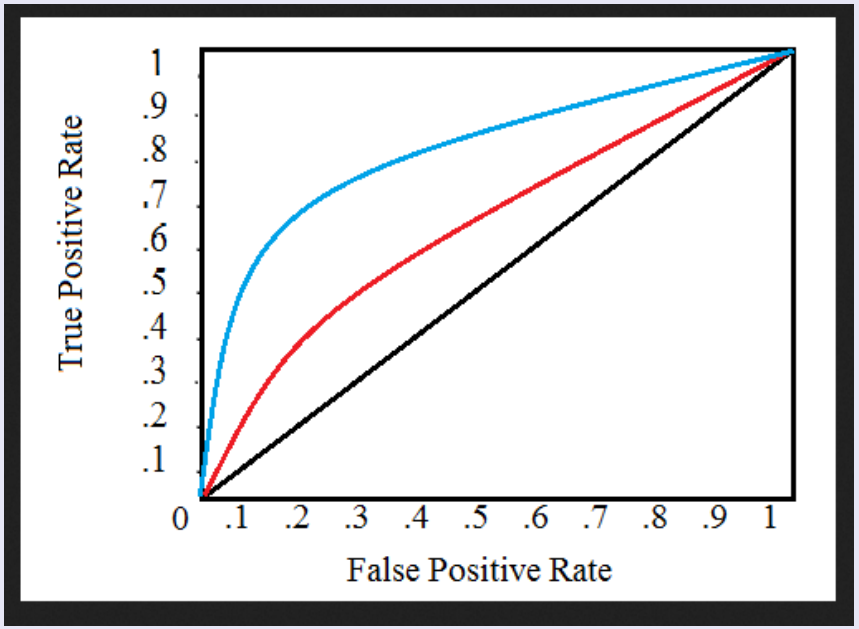
How many relevant items are selected?

$$\text{Recall} = \frac{\text{true positives}}{\text{relevant elements}}$$

F score

$$F_1 = \left(\frac{2}{\mathrm{recall}^{-1} + \mathrm{precision}^{-1}} \right) = 2 \cdot \frac{\mathrm{precision} \cdot \mathrm{recall}}{\mathrm{precision} + \mathrm{recall}}.$$

C statistic


True Positive Rate
1
.9
.8
.7
.6
.5
.4
.3
.2
.1
0
.1
.2
.3
.4
.5
.6
.7
.8
.9
1
False Positive Rate

Synchronic performance

Table 5. *The predictive power of the synchronic simulation (lexicon: twentieth century, test set: twentieth century, N = 564). For the probabilistic AM predictions: C = 0.92*

F-score, macro-averaged:	0.88
% correct predictions (overall):	88.65%
F-score for -*ity*:	0.91
% correct -*ity*:	93.31%
F-score for -*ness*:	0.85
% correct -*ness*:	81.36%

Diachronic performance

Table 6. *The predictive power of the diachronic simulation (lexicon: nineteenth century, test set: twentieth century, $N_{test\ set} = 564$). For the probabilistic AM predictions: $C = 0.89$*

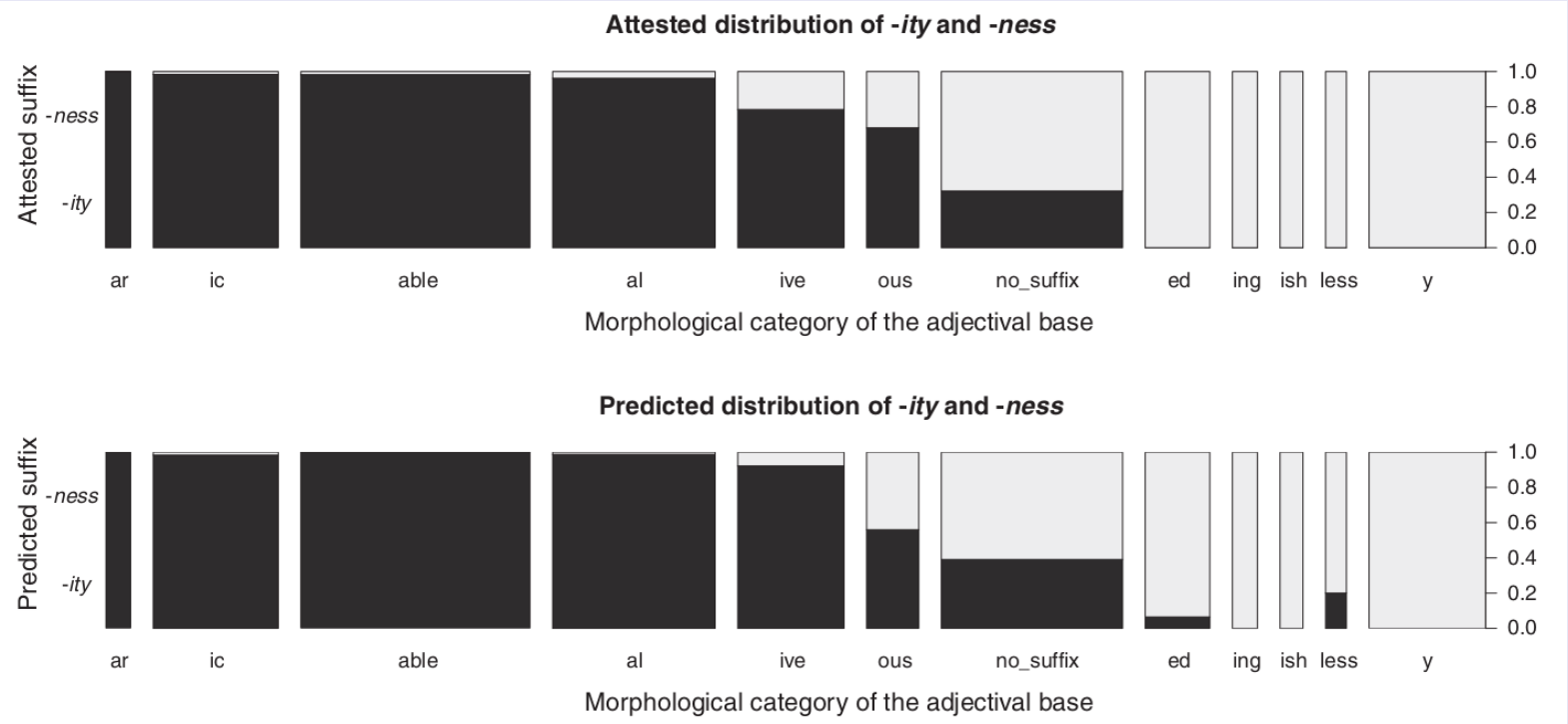
F-score, macro-averaged:	0.85
% correct predictions (overall):	85.82%
F-score for *-ity*:	0.88
% correct *-ity*:	82.56%
F-score for *-ness*:	0.83
% correct *-ness*:	90.91%

Diachronic performance

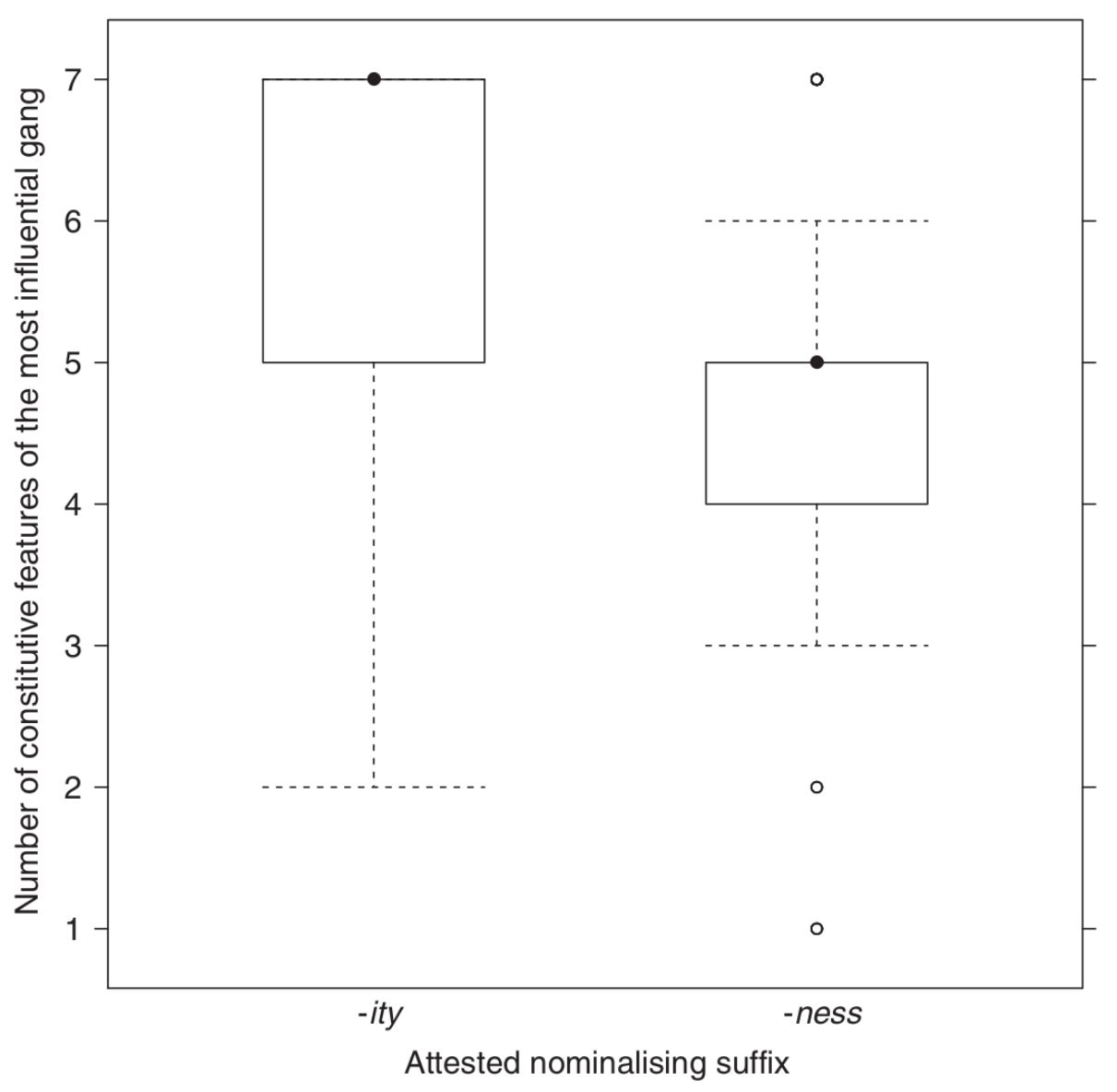
Table 7. *The predictive power of the diachronic simulation (lexicon: eighteenth century, test set: nineteenth century, $N_{test\ set} = 1{,}492$). For the probabilistic AM predictions: $C = 0.82$*

F-score, macro-averaged:	0.78
% correct predictions (overall):	79.69%
F-score for *-ity*:	0.76
% correct *-ity*:	69.85%
F-score for *-ness*:	0.80
% correct *-ness*:	86.56%

Domain-specific productivity

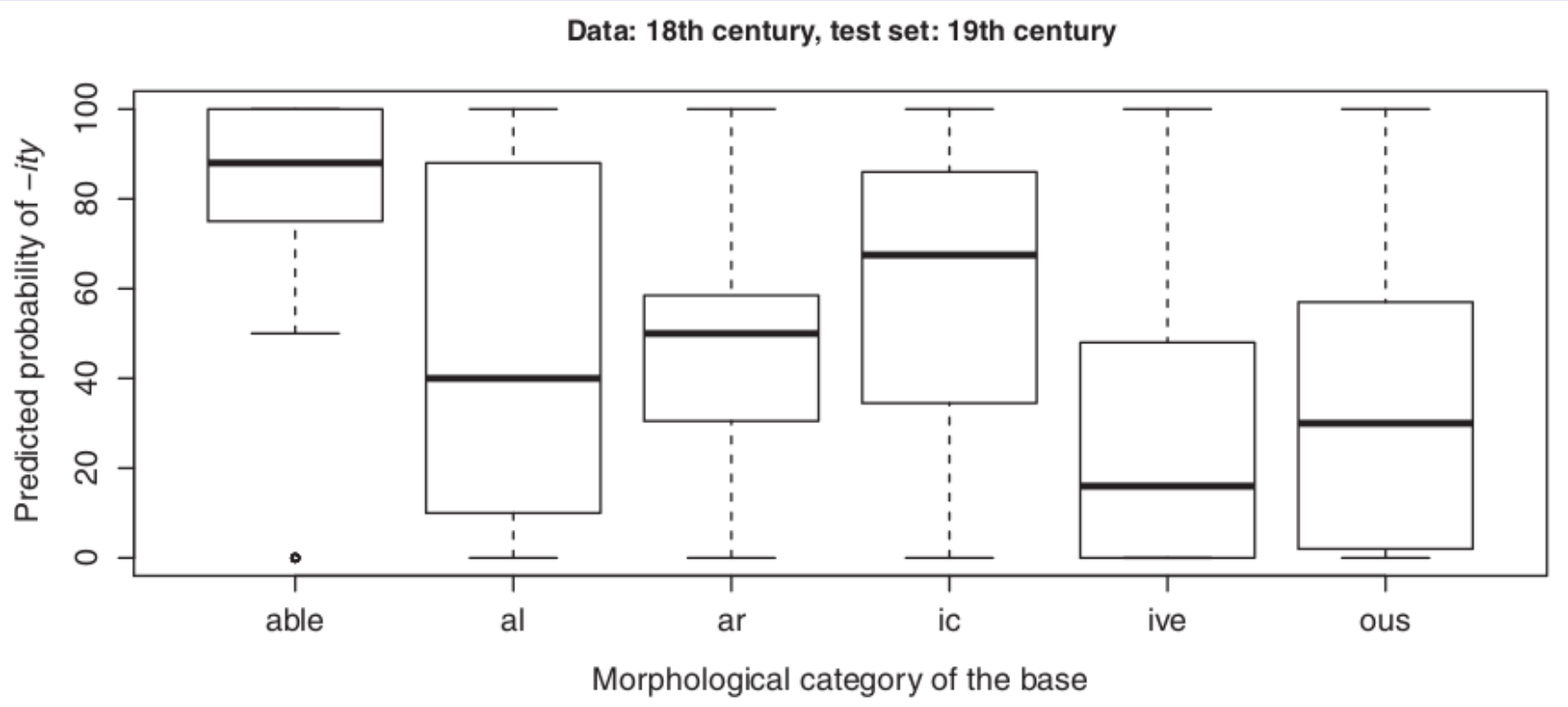


Productivity of *-ity* vs. *-ness*

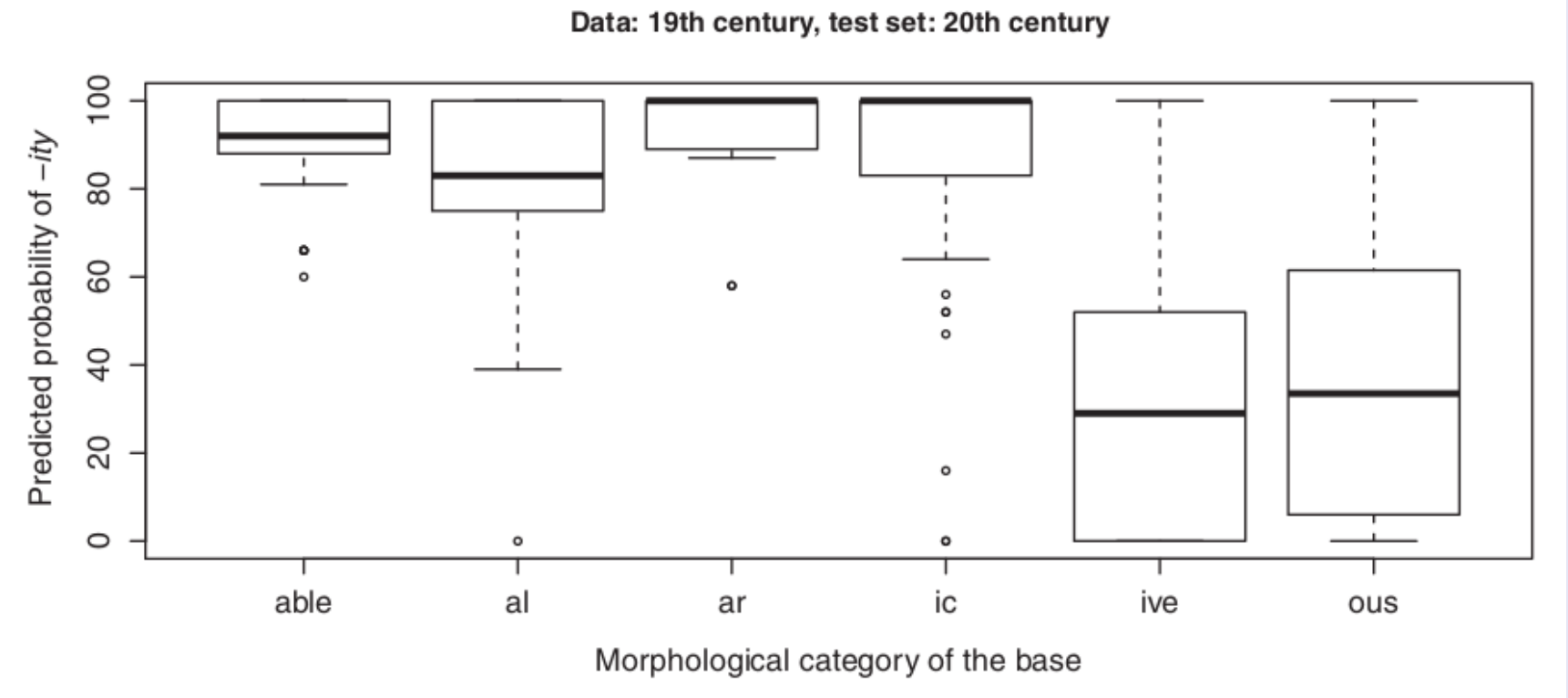


The findings of the analysis thus show that differences in productivity profiles between -ity and -ness emerge from differences in the similarity structure that is relevant for the classification in the lexicon. For -ity derivatives, classification is very local, i.e. dominated by highly similar exemplars.

Diachrony


Data: 18th century, test set: 19th century
Predicted probability of –ity
0
20
40
60
80
100
able
al
ar
ic
ive
ous
Morphological category of the base

Diachrony


Data: 19th century, test set: 20th century
Predicted probability of –ity
0
20
40
60
80
100
able
al
ar
ic
ive
ous
Morphological category of the base

Diachrony


Data: 20th century, test set: 20th century
Predicted probability of –ity
0
20
40
60
80
100
able
al
ar
ic
ive
ous
Morphological category of the base

Gang changes

The number of bases ending in the sequence [nəbl̩] rose from four in the eighteenth-century data (alienability, inalienability, ponibility, interponibility) to 29 in the nineteenth century (e.g. retainability, unamenability, assignability, fashionability).

Since all four eighteenth-century [nəbl̩] bases take -ity as a nominaliser, these four exemplars, acting as a gang for 29 new words, exerted a disproportionately strong pressure towards -ity among -able bases.

Conclusion

Productivity profiles

- *-ity* and *-ness* are semi-complementary
- among the bases that allow for both, there are clear preferences (*-ity* tends to be favored)
- *-ity* is strongly preferred in latinate bases
- analogy plays a clear role

Diachrony

- productivity of *-ity* has been increasing
- productivity of *-ness* has been dencreasing
- this is guided by analogy