

Lexical convergence between Belgian & Netherlandic Dutch

-same virtues, different sins-

Jocelyne Daems, Kris Heylen & Dirk Geeraerts
QLVL, KU Leuven

1. Standardisation of Dutch, stratification and lexical convergence

Standardisation history of Dutch in Belgium and in the Netherlands

NeDu: continuous independent development of standard variant of Dutch
 ⇔ BeDu: after a long delay, promotion of convergence with NeDu norm

Stratification of Dutch in Belgium and in the Netherlands

NeDu: fuzzy border between colloquial and standard Netherlandic Dutch
 ⇔ BeDu: Colloquial Belgian Dutch (CBD) as a supra-regional, informal variant, widely apart and easily (morpho-syntactically & phonetically) distinguishable

Convergence and divergence in the Dutch lexicon (Geeraerts et al. 1999)

⇒ 2x15 concepts (clothing & football) ⇒ 2 varieties (BeDu & NeDu)
 e.g. **JEANS** jeans - spijkerbroek ⇒ 3 points in time (1950-70-90)
 ⇒ 2 stylistic-stratificational dimensions (shop windows & national magazines)

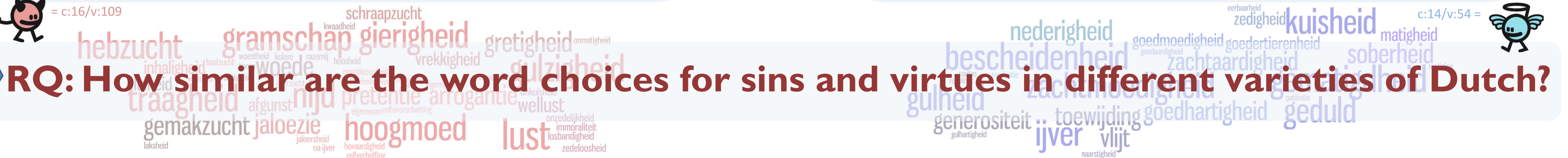
Conclusion: synchronic stratification: larger distance between CBD & SBD
 diachronic convergence: BeDu's normative orientation towards NeDu

2. Data retrieval, manual disambiguation and final data collection

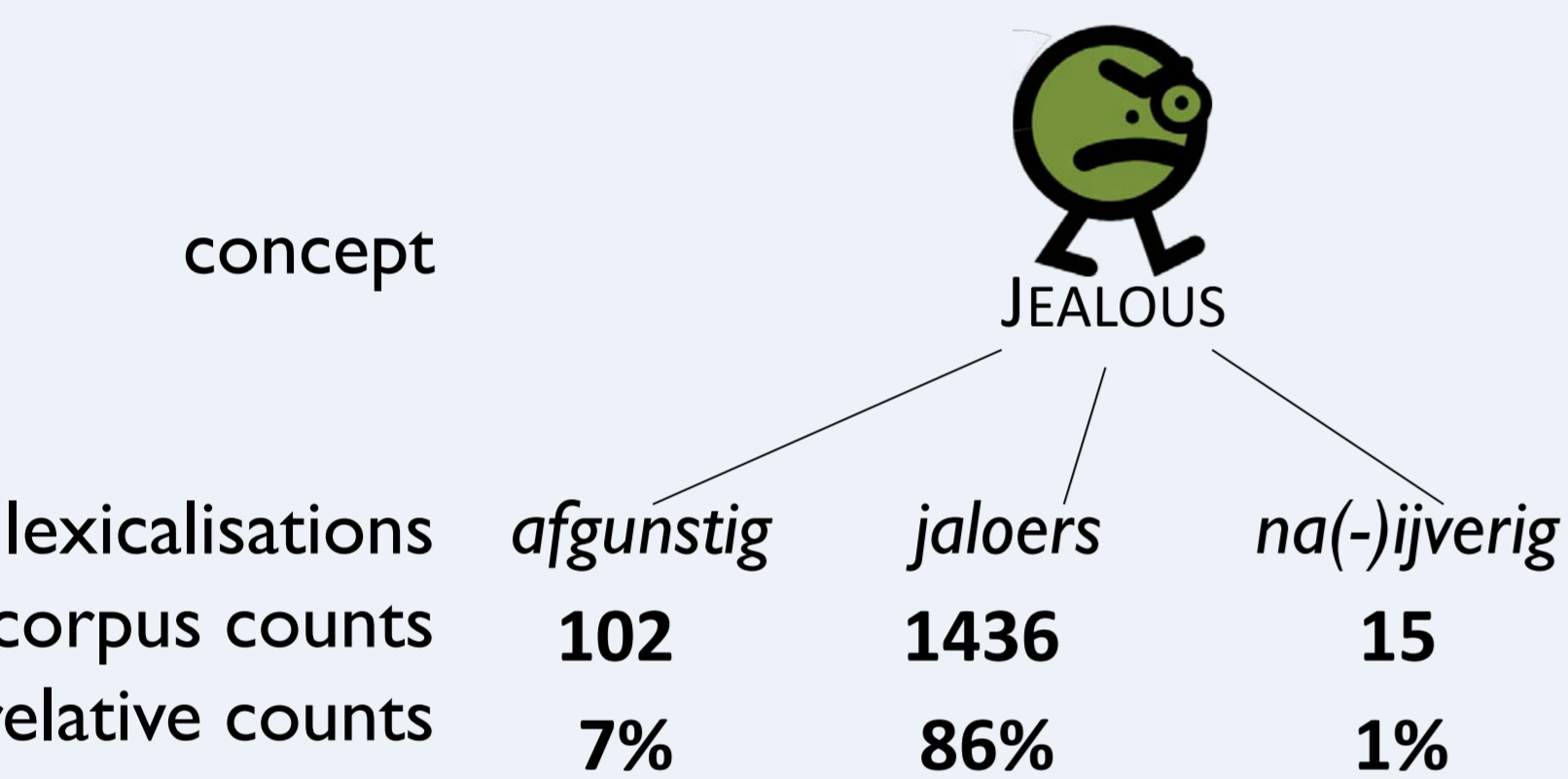
Data retrieval: python script to browse corpora ('99-'04) tagged by Alpino-parser

	newspaper	Usenet	Total	sin	virtue	Total
BeDu	373 M	18 M	391 M	noun 34.867	12.460	47.327
NeDu	161 M	24 M	185 M	adj 65.725	21.630	87.355
Total	534 M	42 M	576 M	Total 100.592	34.090	134.682

Example retrieval – manual disambiguation	Clean hit?
Onze overeenkomst geldt in goede en in kwade dagen. (Our agreement applies in good and in bad days.)	wrong meaning
Mijn vrouw, allez, mijn vriendin was dan kwaad. (My wife, I mean, my girlfriend would then be angry.)	clean hit
Het mysterie van het kwaad in de wereld. (The mystery of evil in the world.)	wrong PoS



3. Methodology: profile-based uniformity measure

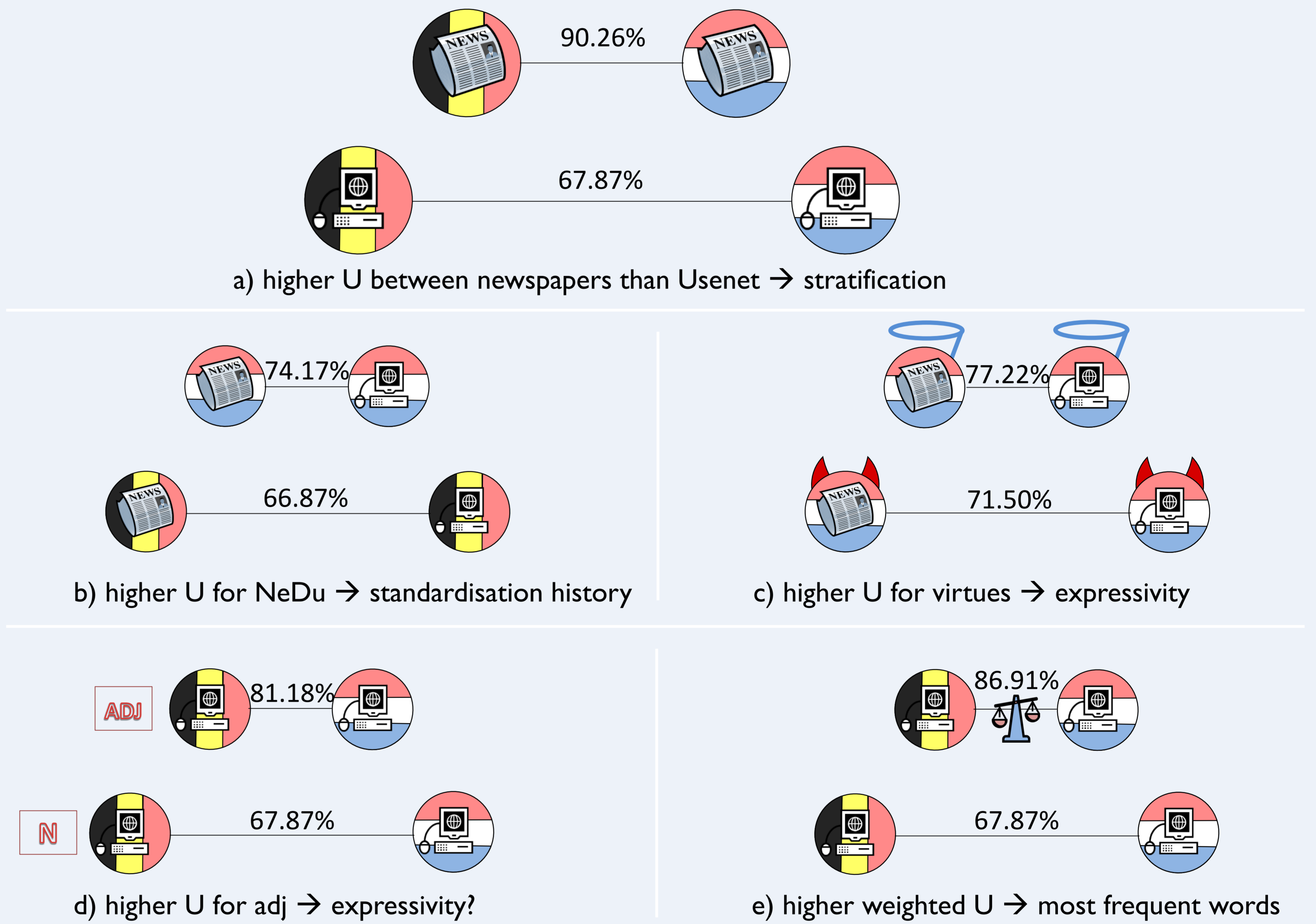


$$U_Z(Y_1, Y_2) = \sum_{i=1}^n \min(F_{Z,Y_1}(x_i), F_{Z,Y_2}(x_i))$$

	Y ₁ , Y ₂		F _{Z,Y₂} (x _i)		
Z	JEALOUS	newsp.	rel.	Usenet	rel.
x ₁	afgunstig	102	7%	56	14%
	ijverzuchtig	0	0%	1	0%
	jaloers	1436	92%	351	86%
	na(-)ijverig	15	1%	0	0%

⇒ U_{JEALOUS}(newsp., Usenet) = 7% + 0% + 86% + 1% = 93%
 (+ optional concept frequency weighting)

4. Results and interpretation



5. Conclusions, attenuations, related studies and further research

Confirmation: convergence newspapers & synchronic stratification

High U between BeDu & NeDu newspapers (~ standardised language)
 Lower U between newspapers & Usenet in BeDu (< delayed standardisation)
 Higher uniformity values for adjectives, virtues (< expressivity) & weighted values

Attenuations

Weighted U affected by high frequent concepts GRAMSCHAP, HOOGMOED, GEDULD
 No evidence found for influence of normative literature, endogeneity, age on U

Related studies with more concrete lexical fields have lower U_(newspapers)

Traffic 69,86% – IT 79,08% – clothing 82,02% – person reference nouns 83,97%

Future research: data retrieval and data collection

Less manual disambiguation (time & labour intensive) < Semantic Vector Spaces
 ⇒ A more stratified and complete image: other registers and more profiles

6. References

Geeraerts, Dirk. 2003. "Cultural models of linguistic standardization." In: Dirven, Roslyn and Pütz (eds.), *Cognitive Models in Language and Thought. Ideology, Metaphors and Meanings*, 25–68. Berlin: Mouton de Gruyter.

Geeraerts, Dirk. 2011. "Colloquial Belgian Dutch." In: Soares da Silva, Torres and Gonçalves (eds.), *Linguas Pluricentricas. Variação Linguística e Dimensões Sociocognitivas*, 61–74. Braga: Publicações da Faculdade de Filosofia, Universidade Católica Portuguesa.

Geeraerts, Dirk, Stefan Grondelaers and Dirk Speelman. 1999. *Convergentie en divergentie in de Nederlandse woordenschat: een onderzoek naar kleding- en voetbaltermen*. Amsterdam: P.J. Meertens-Instituut.

Grondelaers, Stefan, Dirk Geeraerts, Dirk Speelman & José Tummers. 2001. "Lexical standardisation in internet conversations. Comparing Belgium and The Netherlands." In Fontana, McNally, Turell and Vallduví (eds.), *Proceedings of the First International Conference on Language Variation in Europe*, 90–100. Barcelona.

Zenner, Eline, Jocelyne Daems & Dirk Geeraerts. 2013. "Pluricentric Languages and Exocentric Norms: Lexical Convergence between Belgian Dutch and Netherlandic Dutch." Workshop presentation at SLE 2013, Split, Croatia.