

How does productivity relate to semantic coherence in the Evaluative Binominal Construction?

A cross-linguistic corpus study in Dutch, English, French and Italian

Isa Hendriks, Valentina Piunno, Kristel Van Goethem

CogLing Days 2024
University of Antwerp
12-13 December 2024



UNIVERSITÀ
DEGLI STUDI
DI BERGAMO



Outline

1. The Evaluative Binominal Construction: state of the art
2. Research goal and questions
3. Data and method of analysis
4. Corpus results
5. Discussion and conclusions

The Evaluative Binominal Construction (EBC)

- **Examples:**

- DU *een schat van een kind* ‘a lovely child’ (lit. a treasure of a child)
- EN *a hell of a day*
- FR *cette saleté de médicament* ‘this dirty medication’ (lit. this dirt of medication)
- IT *quel tesoro di bambino* ‘that lovely child’ (lit. that treasure of child)

- **Structure:** [DET1 N1 PREP (DET2) N2]

- Semi-schematic construction
 - Only the preposition is fixed
 - All other slots are subject to variation

- **Non-compositional meaning:**

- “N2 showing to an extraordinary degree characteristic features of N1 (subjective-evaluative)” (translated from Verhagen 2005: 202)
 - N1 → extraordinary (physical or abstract) quality (metaphorical value)
 - N2 → entity to which this quality is attributed (referential head)

State of the art

- **Studies on single languages:**

- Dutch (a.o. Paardekooper 1956, Verhagen 2005, **Van Goethem 2023**)
- English (a.o. Aarts 1994, 1998, **ten Wolde 2023**)
- French (a.o. Regula 1965, Larrivée 1994)

- **Cross-linguistic studies:**

- Germanic and Romance languages (a.o. Foolen 2004)
- Romance languages (a.o. Tutescu 1969, **Piunno in press**)

→ red color = corpus-based studies

→ Corpus-based cross-linguistic studies are still scarce

Research goal & questions

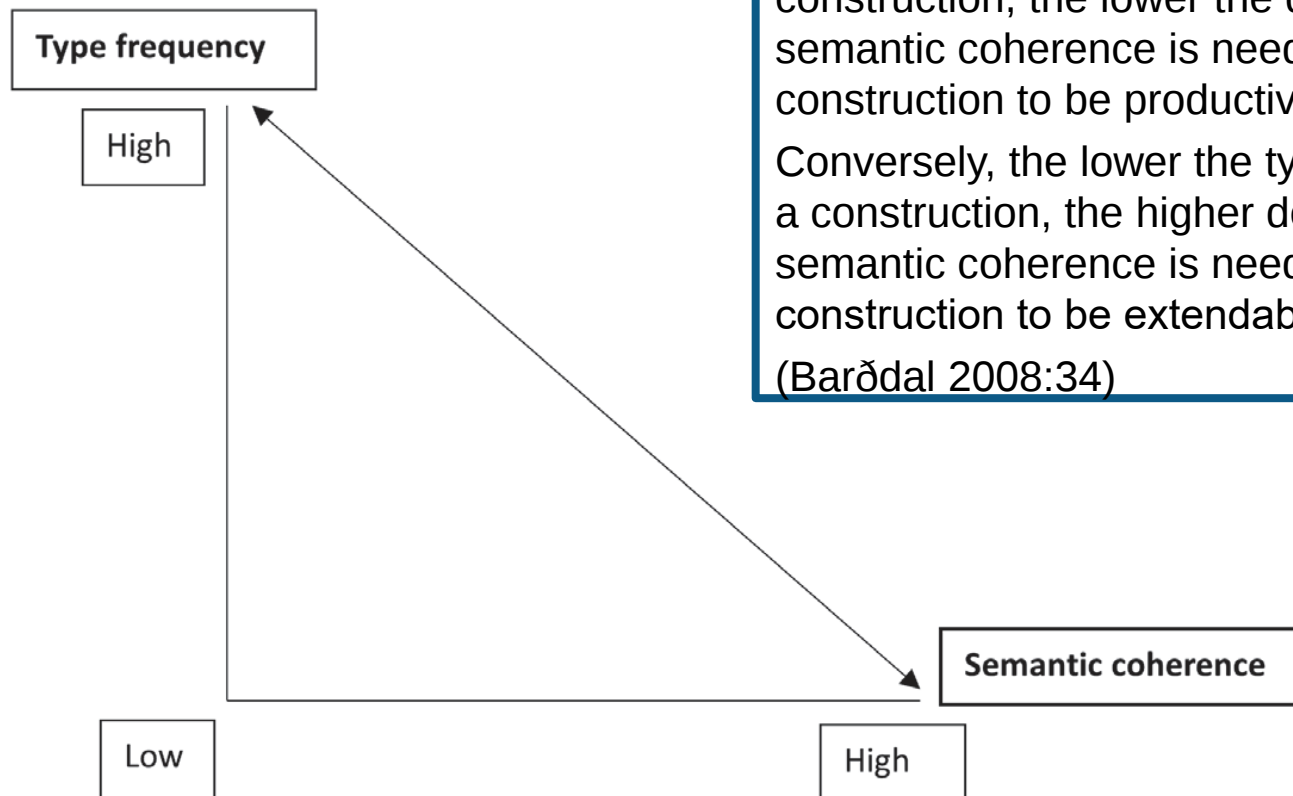
Research goal

- A corpus-based cross-linguistic constructionist analysis of the EBC (Dutch, English, French, Italian)

Research questions

- **RQ1:** What are the main **cross-linguistic similarities and differences** between the EBC in the four languages under study?
- **RQ2:** What is the relation between **productivity** and **semantic coherence** in the EBC in the four languages under study?
 - *semantic coherence*: internal consistency between the members of the schema (Barðdal 2008: 27)

Hypothesis: Inverse correlation between productivity and semantic coherence



“[...] the higher the type frequency of a construction, the lower the degree of semantic coherence is needed for a construction to be productive.

Conversely, the lower the type frequency of a construction, the higher degree of semantic coherence is needed for a construction to be extendable.”

(Barðdal 2008:34)

Barðdal (2008:35)

Corpus data

Sketch Engine: <https://www.sketchengine.eu>

- TenTen corpus family
- corpus data collected from the Internet

two Germanic languages (English and Dutch)

- | | | |
|--------------|------------------------|-----------------------|
| • nlTenTen20 | (6,836,979,371 tokens) | June and July 2020 |
| • enTenTen21 | (52 billion words) | October–December 2021 |

two Romance languages (French and Italian)

- | | | |
|--------------|-------------------------|---------------|
| • frTenTen23 | (23.8 billion words) | 2022 and 2023 |
| • itTenTen20 | (14,514,566,714 tokens) | 2019 and 2020 |

Concordance function – CQL queries

Combination of quantitative and qualitative analysis

Data extraction

General analysis of the **top 20** EBC patterns per language

- General CQL Query:
 - e.g., English [word="a"] [tag = "noun.*"] [word="of"] [word="a"]
- Extract top 20 EBC patterns per language
- General analysis of frequency and semantic properties of top 20

Detailed analysis of the **top 5** EBC patterns per language

- More specific CQL query including N1 lemma
 - e.g., English [tag="det.*"]? [lemma="hell"] [lemma="of"]
 - Italian: [DET 1="quello|quel|questo|quella|quelle|quegli|quei|questi|queste"]
- Random samples of 1000 occurrences per subpattern
- Manual check until 100 relevant occurrences per pattern
- Final datasets of 5x 100 relevant occurrences per language
- = **2000 relevant occurrences in total**

Only for Dutch:
detailed analysis of
the top 20 EBC
patterns

Data annotation and analysis

- **Annotation in Excel in terms of:**
 - **Formal properties** (Det 1, Det 2, number, complexity N2)
 - **Evaluative semantics construction:** ameliorative, pejorative, augmentative, diminutive or combination
 - **Semantics N2:** person, animal, object, place, concept
 - **Lemma N2** ~ productivity scores: TTR, PP

RESULTS:

1. Cross-linguistic similarities and differences

Frequency: top 20

Dutch	English	French	Italian
<i>een dijk van een</i>	<i>a hell of a</i>	<i>ce/cette putain de</i>	<i>quel capolavoro di</i>
<i>een pracht van een</i>	<i>a heck of a</i>	<i>ce/cette drôle de</i>	<i>quel genio di</i>
<i>een schat van een</i>	<i>a gem of a</i>	<i>cette saloperie de</i>	<i>quel maiale di</i>
<i>een joekel van een</i>	<i>a whale of a</i>	<i>ce con de</i>	<i>quello schifo di</i>
<i>een draak van een</i>	<i>a beast of a</i>	<i>cette saleté de</i>	<i>quel pazzo di</i>
<i>een beest van een</i>	<i>a ghost of a</i>	<i>ce monstre de</i>	<i>quella merda di</i>
<i>een juweel van een</i>	<i>a monster of a</i>	<i>ce diable de</i>	<i>quell' amore di</i>
<i>een kanjer van een</i>	<i>a devil of a</i>	<i>ce salaud de</i>	<i>quell' idiota di</i>
<i>een droom van een</i>	<i>a mess of a</i>	<i>ce connard de</i>	<i>quel porco di</i>
<i>een parel van een</i>	<i>a cracker of a</i>	<i>ce bijou de</i>	<i>quel coglione di</i>
<i>een knaller van een</i>	<i>a steal of a</i>	<i>cet imbécile de</i>	<i>quel bastardo di</i>
<i>een dot van een</i>	<i>a jewel of a</i>	<i>ce fou de</i>	<i>quel diavolo di</i>
<i>een boom van een</i>	<i>a nightmare of a</i>	<i>cet idiot de</i>	<i>quel simpaticone di</i>
<i>een beer van een</i>	<i>a whirlwind of a</i>	<i>cet abruti de</i>	<i>quella schifezza di</i>
<i>een kast van een</i>	<i>a rollercoaster of a</i>	<i>ce joyau de</i>	<i>quell' inferno di</i>
<i>een monster van een</i>	<i>a mockery of a</i>	<i>cette connerie de</i>	<i>quella gioia di</i>
<i>een kei van een</i>	<i>a bitch of a</i>	<i>cette merde de</i>	<i>quel matto/mattacchione di</i>
<i>een wolk van een</i>	<i>a blast of a</i>	<i>cet enfoiré de</i>	<i>quel poveraccio di</i>
<i>een topper van een</i>	<i>a prince of a</i>	<i>cette salope de</i>	<i>quell' imbecille di</i>
<i>een reus van een</i>	<i>a rocket of a</i>	<i>ce cochon de</i>	<i>quel tesoro di</i>

Top 20: Semantics cxn

Dutch	English	French	Italian
een dijk van een	a hell of a	ce/cette putain de	quel capolavoro di
een pracht van een	a heck of a	ce/cette drôle de	quel genio di
een schat van een	a gem of a	cette saloperie de	quel maiale di
een joekel van een	a whale of a	ce con de	quello schifo di
een draak van een	a beast of a	cette saleté de	quel pazzo di
een beest van een	a ghost of a	ce monstre de	quella merda di
een juweel van een	a monster of a	ce diable de	quell'amore di
een kanjer van een	a devil of a	ce salaud de	quell' idiota di
een droom van een	a mess of a	ce connard de	quel porco di
een parel van een	a cracker of a	ce bijou de	quel coglione di
een knaller van een	a steal of a	cet imbécile de	quel bastardo di
een dot van een	a jewel of a	ce fou de	quel diavolo di
een boom van een	a nightmare of a	cet idiot de	quel simpaticone di
een beer van een	a whirlwind of a	cet abruti de	quella schifezza di
een kast van een	a rollercoaster of a	ce joyau de	quell' inferno di
een monster van een	a mockery of a	cette connerie de	quella gioia di
een kei van een	a bitch of a	cette merde de	quel matto/mattacchione di
een wolk van een	a blast of a	cet enfoiré de	quel poveraccio di
een topper van een	a prince of a	cette salope de	quell' imbecille di
een reus van een	a rocket of a	ce cochon de	quel tesoro di

Ameliorative

Pejorative

Augmentative

Diminutive

Evaluative semantics: summary

Dutch

- Mostly **ameliorative** and **augmentative** connotations
 - *een **pracht** van een kerstkaart* 'a beauty of a Christmas card'
 - *een **joekel** van een auto* 'a huge car' (lit. a whopper of a car)

English

- Mix of **pejorative** and **ameliorative** values:
 - *a **hell** of a job*
 - *a **gem** of a book*

French / Italian

- Clearly more **negative** connotations + **informal** language
 - FR *la **putain** de patrie* 'that fucking fatherland'
 - IT *quello **schifo** di città* 'that disgusting city'
- **Mirative** value of *drôle* (expressing surprise)
 - FR *Quelle **drôle** d'idée* 'what a crazy idea'

Formal properties: general tendencies

DET 1

- Wide variation in DET1 position in Germanic:
 - DU **vier** *schatten van kleinkinderen* 'four darlings of grandchildren'
 - ENG **my** *hell of a house*
- Preference for a demonstrative/exclamative in Romance
 - FR **ces** *saletés de laveurs de vitres* 'those filthy window cleaners'
 - IT **quell'** *amore di bambino* 'that lovely child'

DET 2

- Indefinite article or absent in Germanic (e.g., *an absolute whale of time*)
- Less frequently present in Romance

N1

- often nominalized adjective in the Romance languages
 - FR *quel(le)* **drôle** *de N*
 - IT *quell'***idiot***a di N*

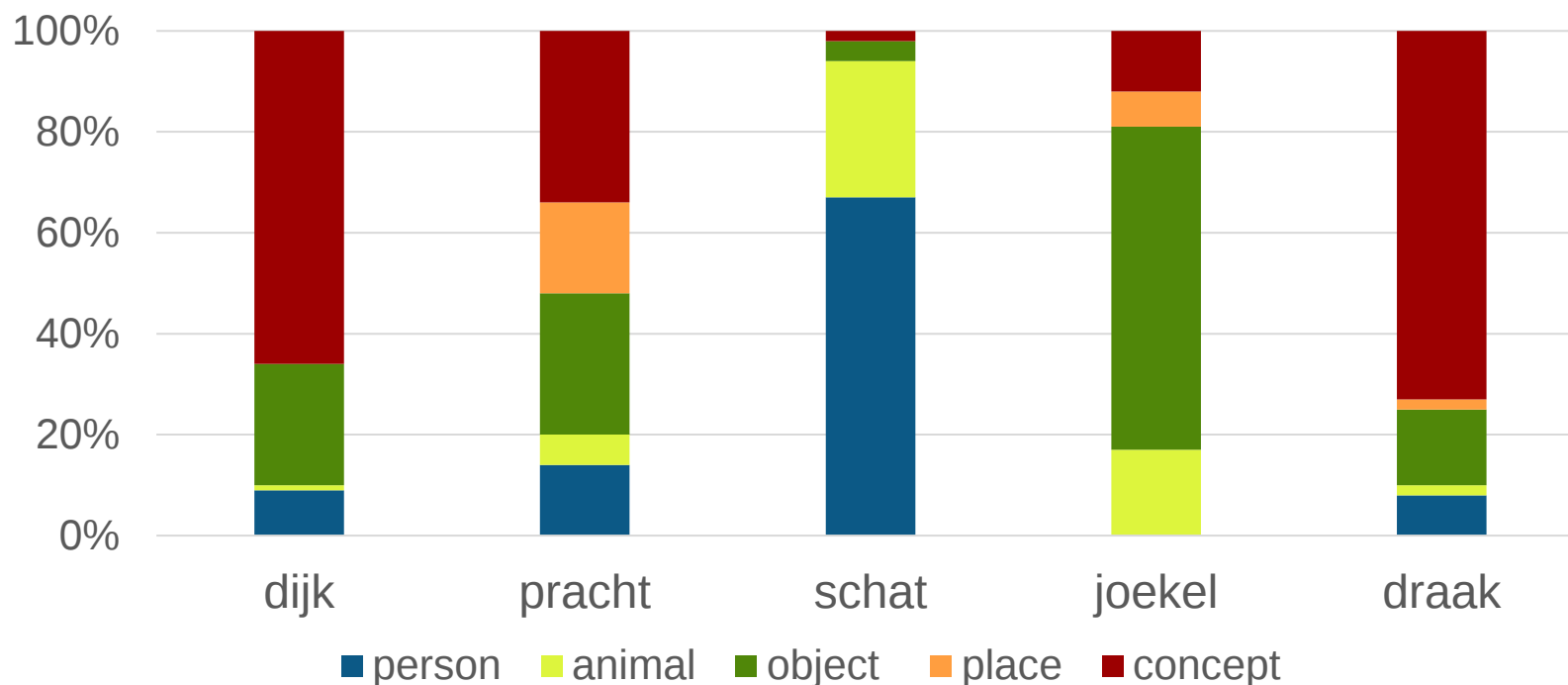
N2

- not only simplex nouns, also compounds and NP
 - FR *ces saloperies de* **toilettes turques** 'these dirty Turkish toilets'
 - DU *een schat van een* **Engelse nanny** 'a sweetheart of an English nanny'

RESULTS:

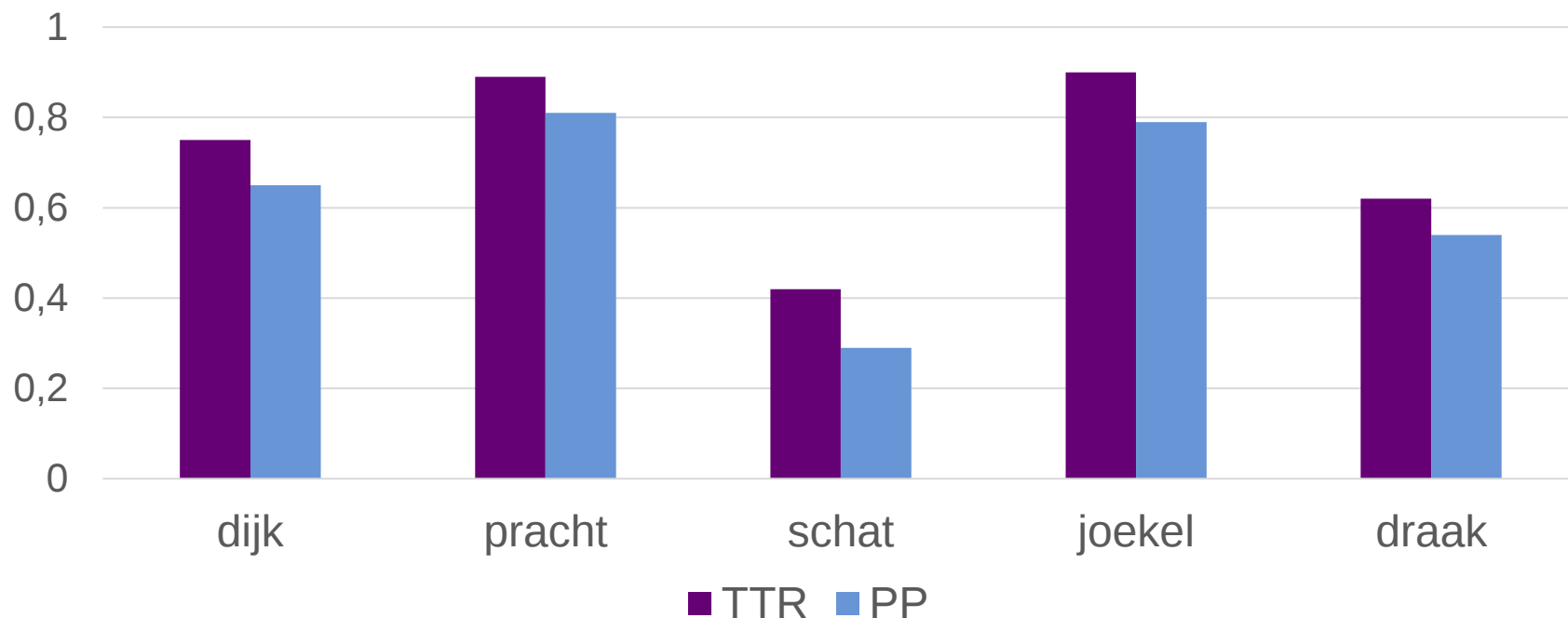
2. Relation between productivity and semantic coherence

Dutch: Semantics N2



- *een dijk van een **stem*** (n=11) 'a powerful voice' (lit. a dike of a voice)
- *een pracht van een **verhaal*** (n=4) 'a beautiful story' (lit. a beauty of a story)
- *een schat van een **hond*** (n=15) 'a lovely dog' (lit. a treasure of a dog)
- *een joekel van een **auto*** (n=3) 'a huge car' (lit. a whopper of a car)
- *een draak van een **wedstrijd*** (n=20) 'a terrible match' (lit. a dragon of a match)

Dutch: Productivity scores

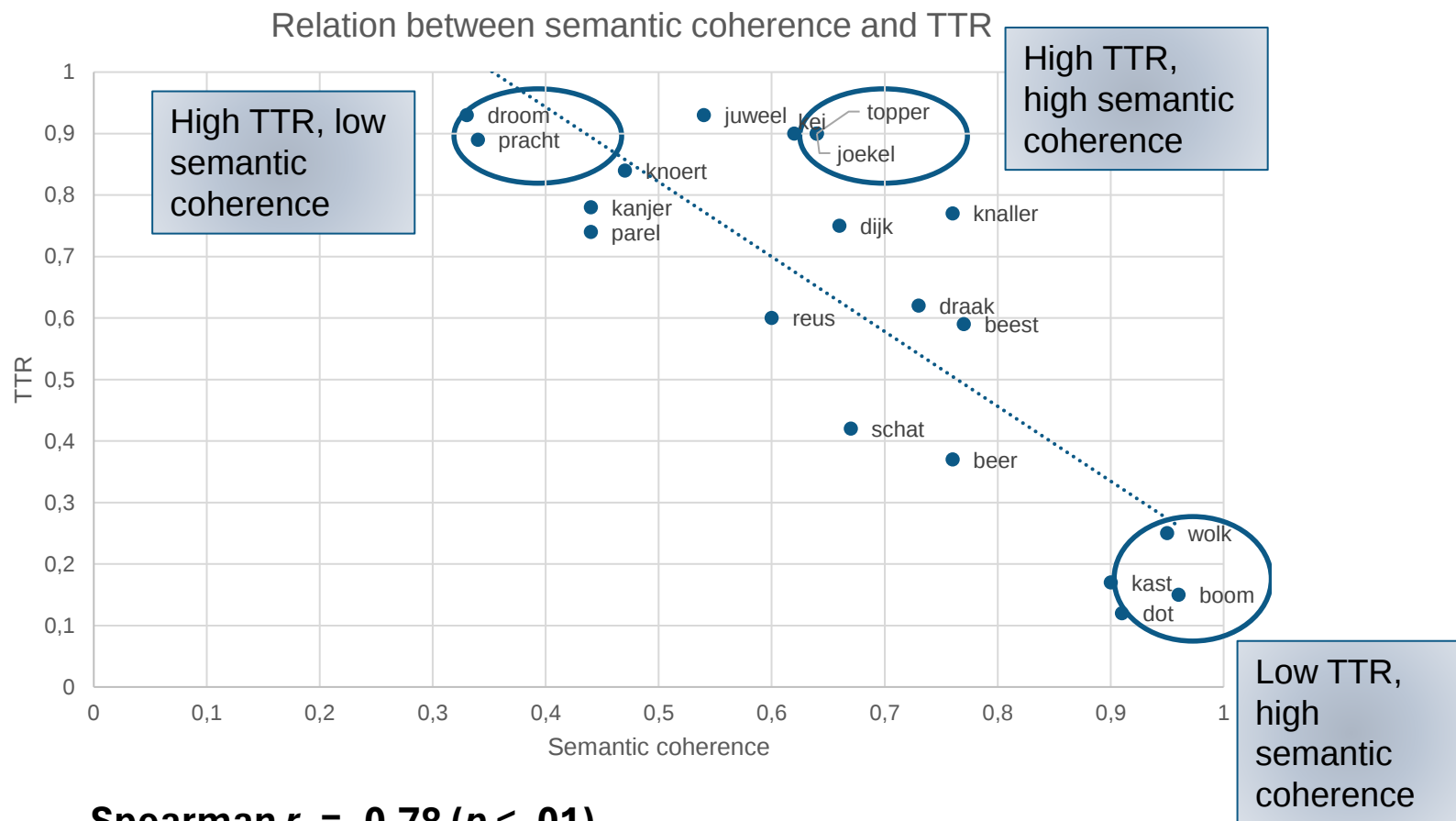


Examples of hapaxes:

- een dijk van een **smartphone**
- een pracht van een **bruidstaart**
- een schat van een **meisjespapegaai**
- een joekel van een **courgette**
- een draak van een **coach**

‘a powerful smartphone’
 ‘a beauty of a wedding cake’
 ‘a sweetheart of a girl parrot’
 ‘a giant zucchini’
 ‘a dragon of a coach’

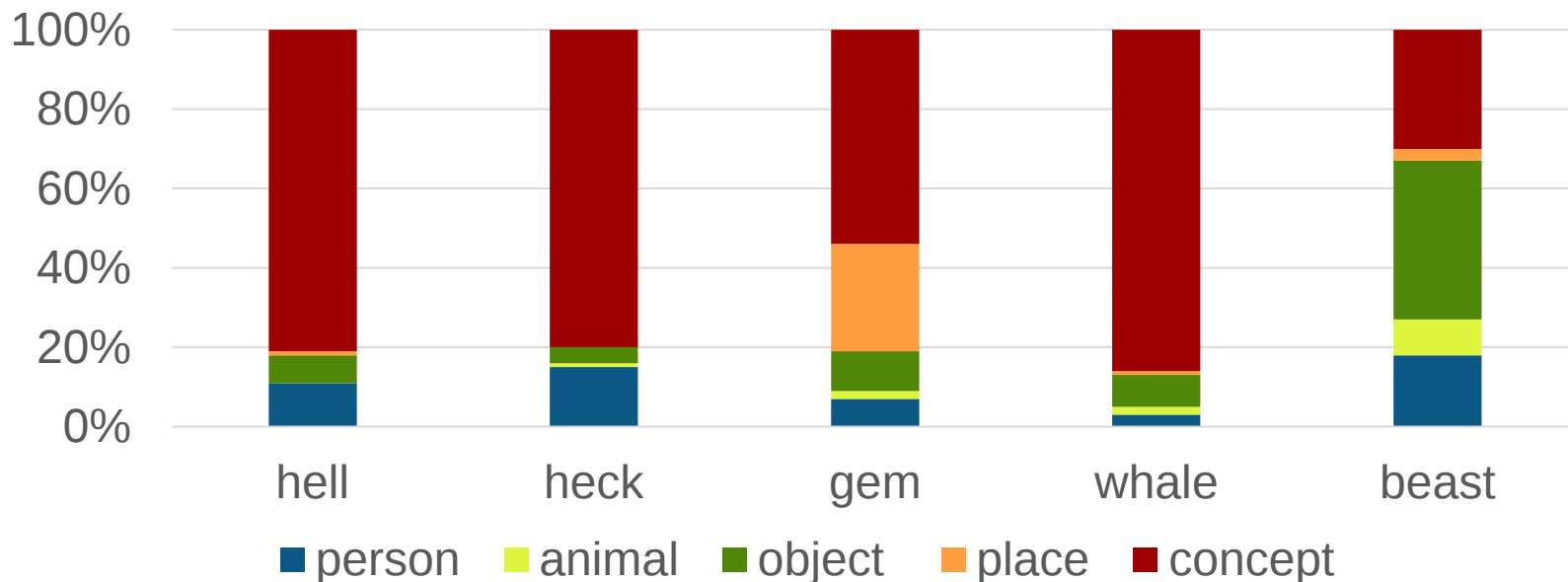
Dutch: Relation between productivity and semantic coherence (20 items)



Spearman $r_s = -0.78$ ($p < .01$)

→ Strong negative relationship between TTR and semantic coherence

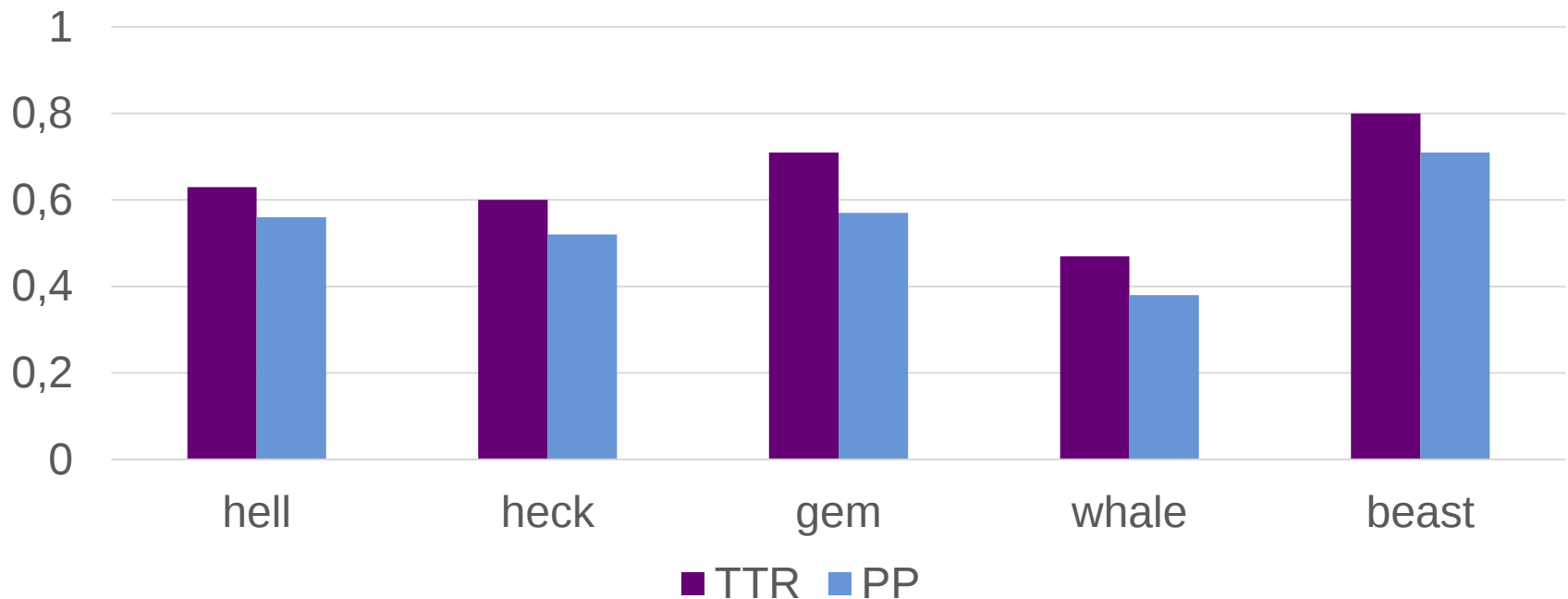
English: Semantics N2



Examples:

- a hell of a *time* (n=5)
- a heck of a *time* (n=10)
- a gem of a *book* (n=8)
- a whale of a *time* (n=39)
- a beast of a *machine* (n=8)

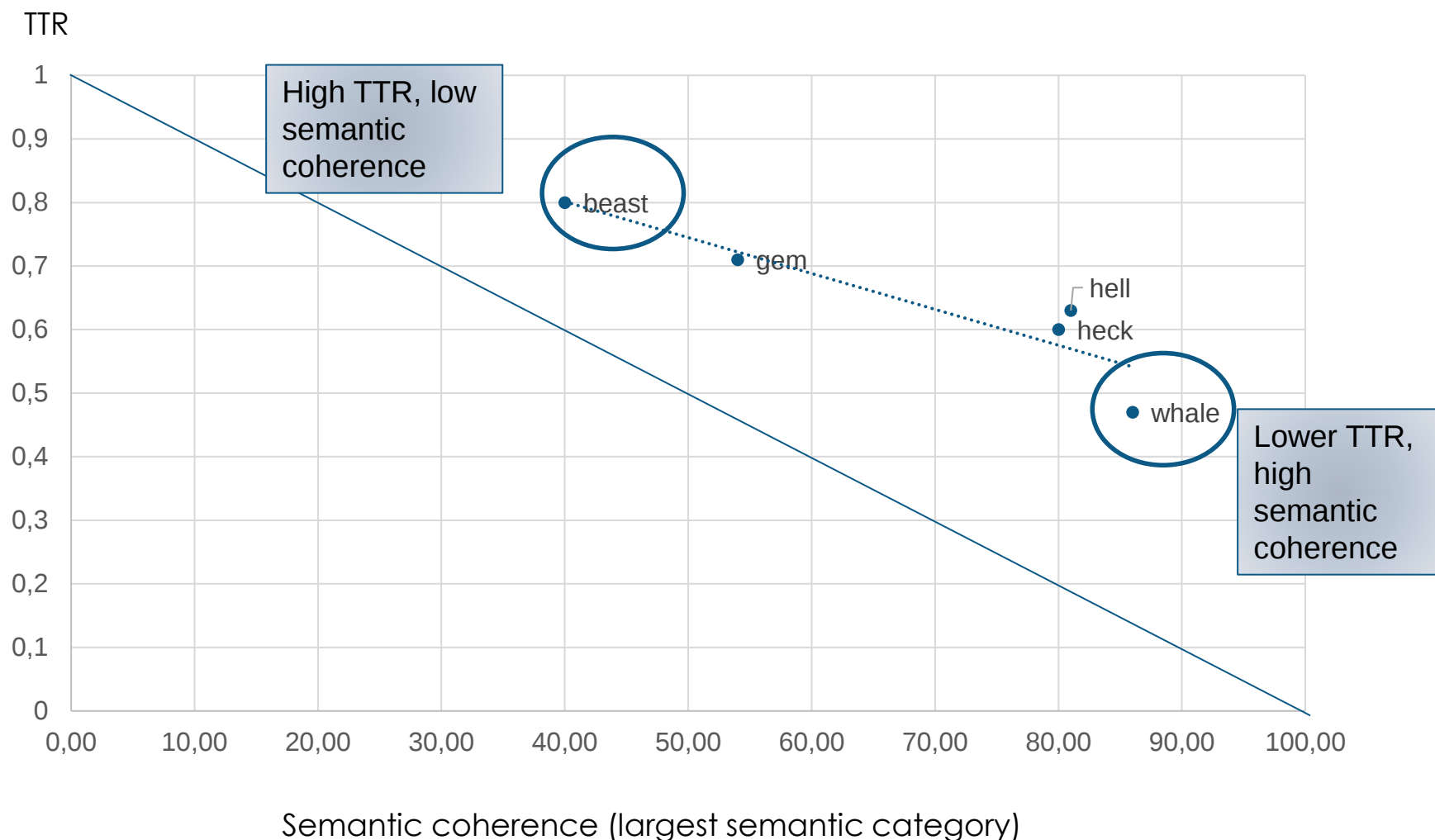
English: Productivity



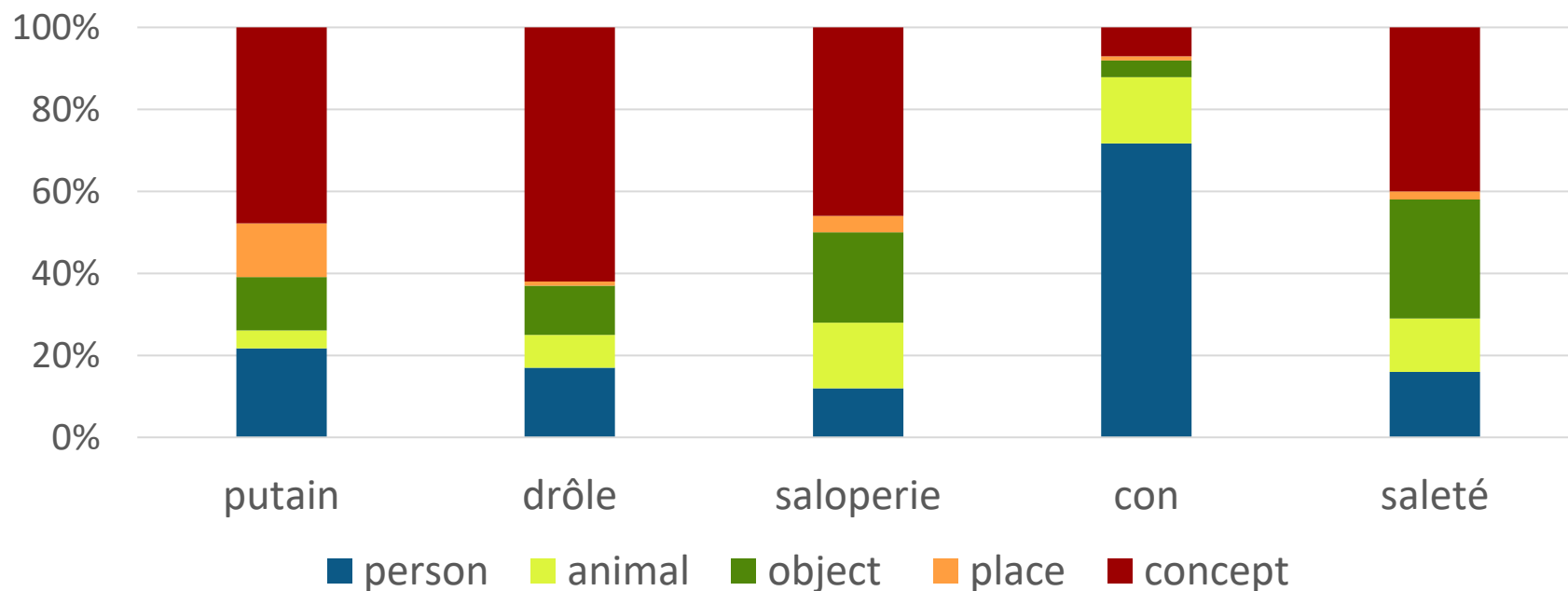
- **Examples of hapaxes:**

- a hell of a *beating*
- one heck of a *cliffhanger*
- a gem of a *town*
- a whale of a *job*
- a beast of a *kayak*

English: Relation between productivity and semantic coherence?



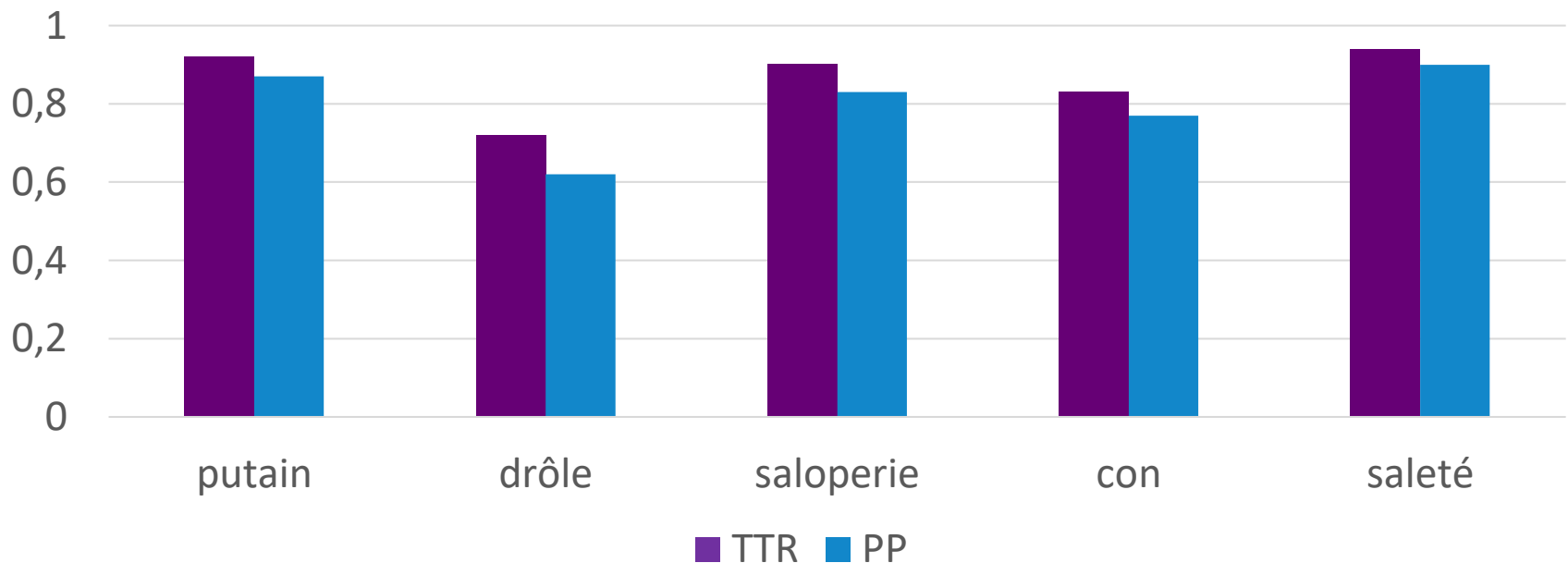
French: Semantics N2



Examples

- *Putain de **merde** !* (n=5) 'Fucking hell!'
- *Quelle drôle d'**idée** !* (n=8) 'What a crazy idea!'
- *Cette salope de **maladie*** (n=4) 'this filthy disease'
- *Un con de **Français*** (n=8) 'an idiot of a Frenchman'
- *Cette saleté de **maladie*** (n=4) 'this filthy disease'

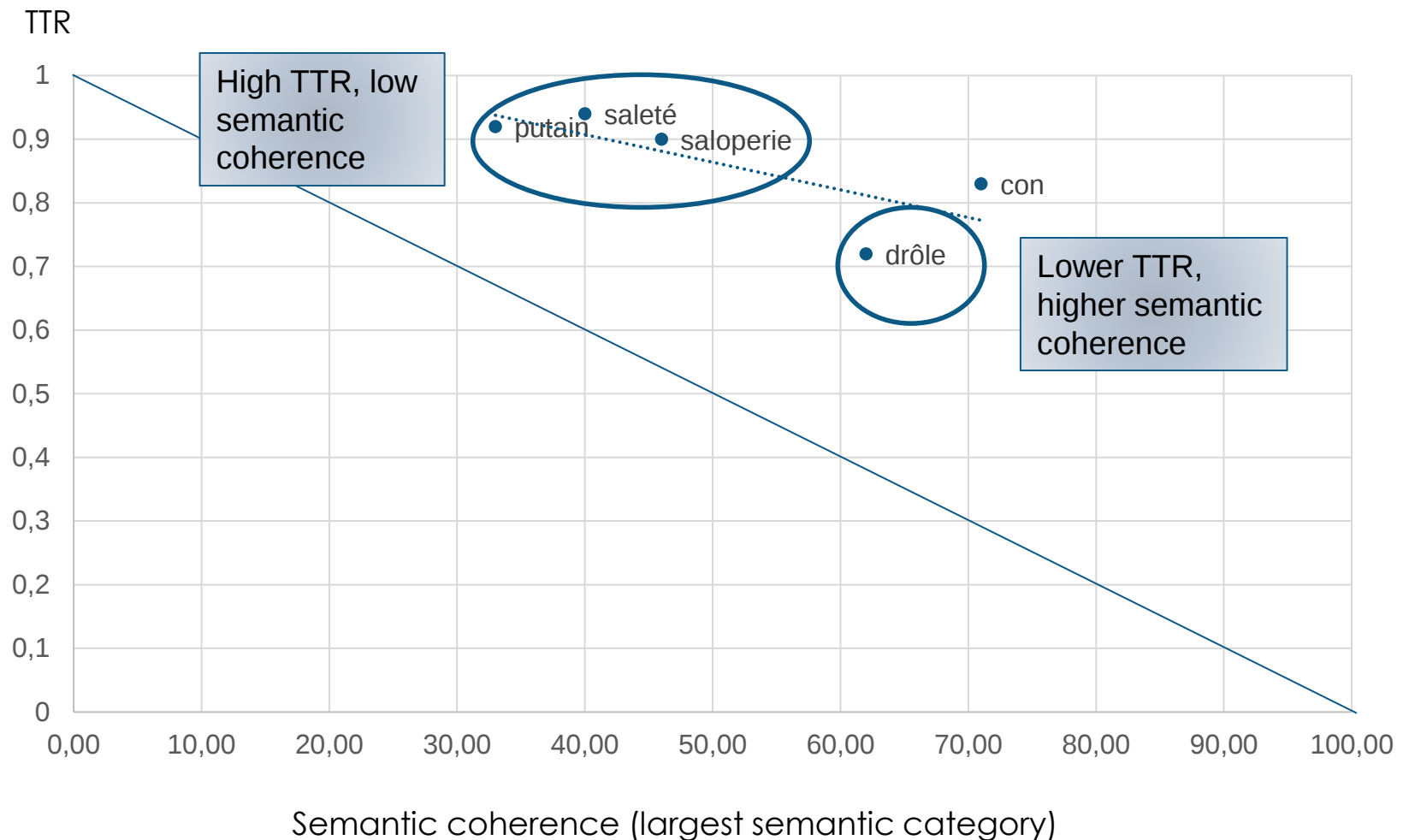
French: Productivity



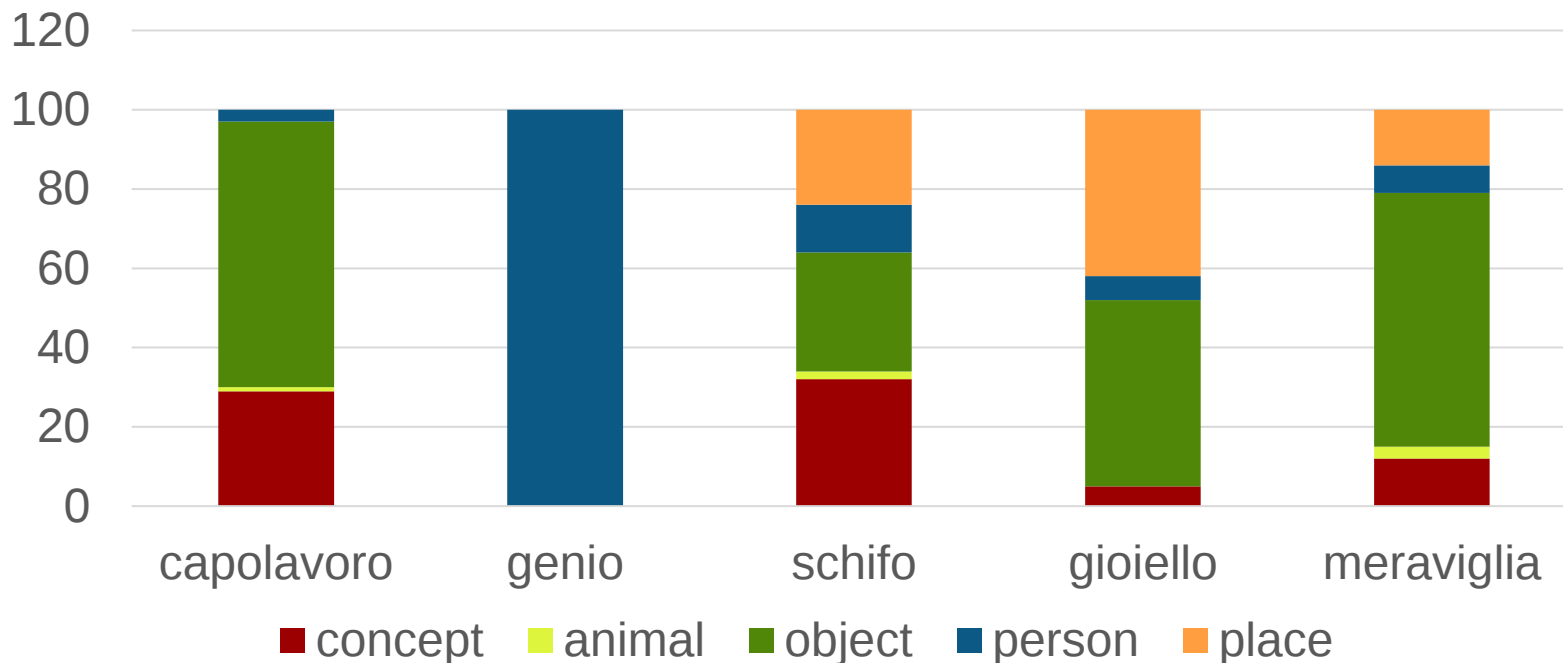
• Examples of hapaxes:

- *Ce putain de **romantisme*** 'this fucking romanticism'
- *Un drôle de **rêve*** 'a crazy dream'
- *Cette saloperie d'**Excel** !* 'this shitty Excel'
- *ces cons de **cyclistes*** 'these idiots of cyclists'
- *une saleté de **bactérie*** 'a nasty bacterium'

French: Relation between productivity and semantic coherence?



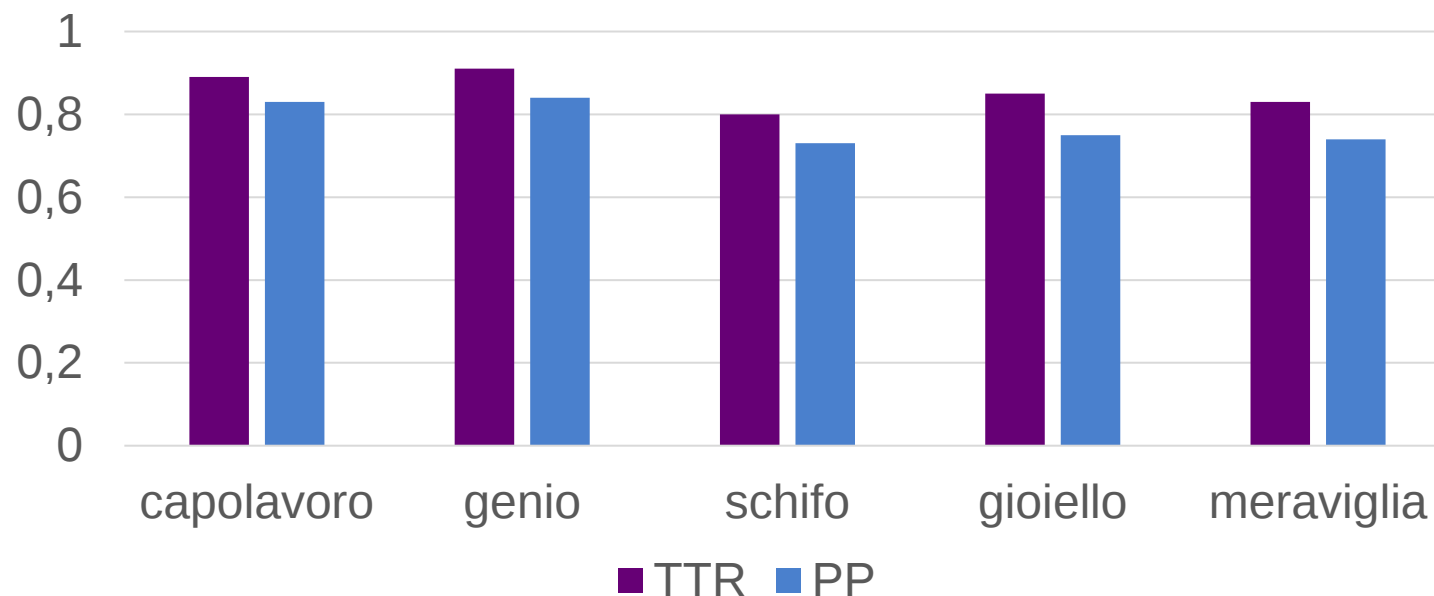
Italian: Semantics N2



- *quel capolavoro di film* (n=5)
- *quel genio di mia madre* (n=3)
- *questo schifo di mondo* (n=9)
- *quel gioiello di città* (n=6)
- *questa meraviglia di torta* (n=5)

‘that masterpiece of a film’
 ‘that genius of my mother’
 ‘this disgust of world’
 ‘that jewel of a city’
 ‘this wonder of a cake’

Italian: Productivity



- **Examples of hapaxes:**

- *quel capolavoro di **disco***
- *quel genio di **ministro***
- *questo schifo di **strada***
- *questo gioiello di **video***
- *quella meraviglia di **auto***

‘that masterpiece of disco’

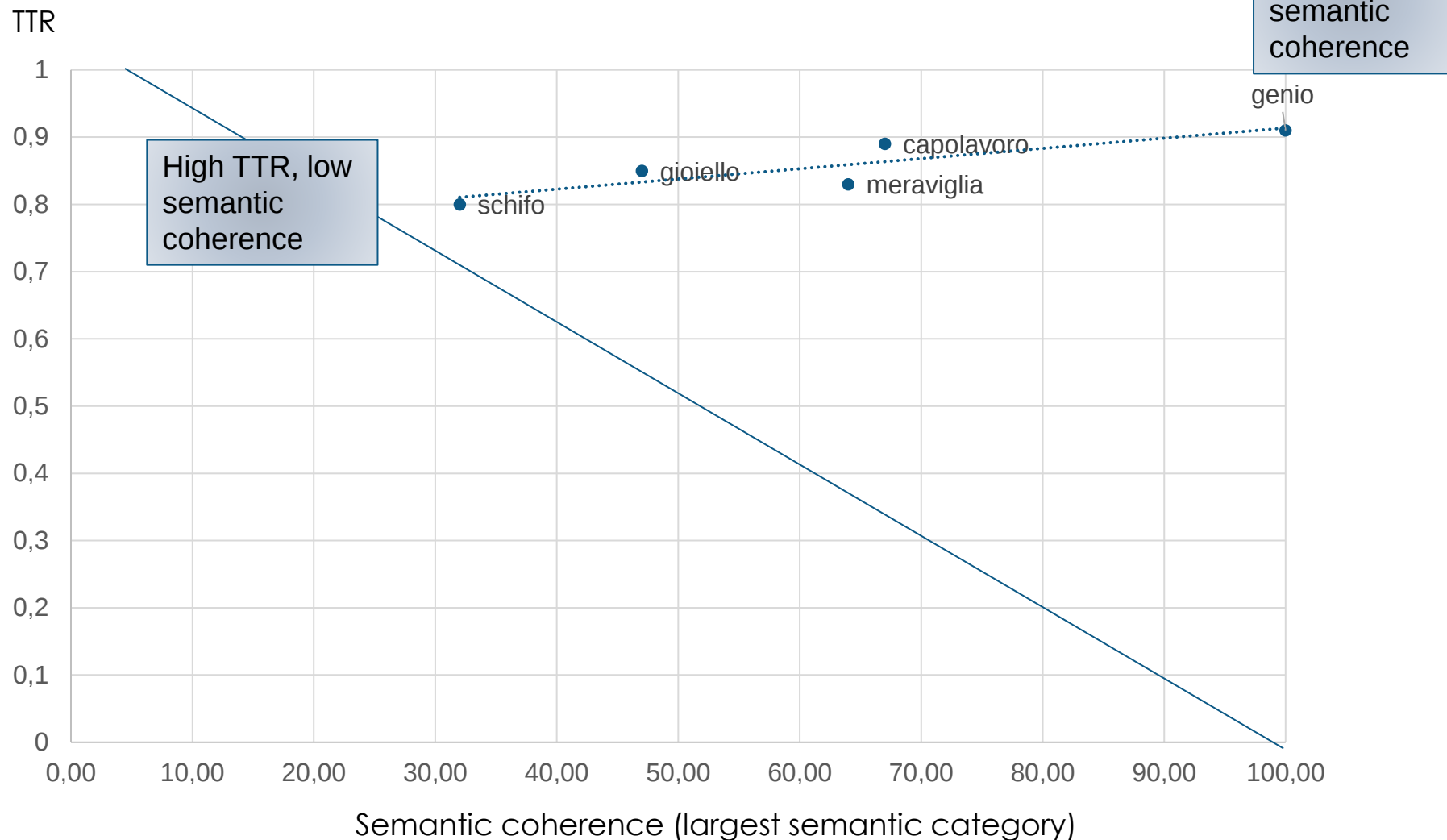
‘that genius of ministry’

‘this disgust of a road’

‘this jewel of a video’

‘that wonder of a car’

Italian: Relation between productivity and semantic coherence?



Discussion and conclusions

- **RQ1: Cross-linguistic similarities and differences?**

- Comparable schematic pattern [DET1 N1 PREP (DET2) N2]
- Comparable non-compositional evaluative semantics
- But also: subtle cross-linguistic differences:
 - **Formal differences:** more variation in DET1 position in Germanic than in Romance; DET2 mostly absent in Romance
 - **Semantic specialization:** ameliorative/augmentative connotation in Dutch, mixed profile in English, pejorative in French and Italian

!!! In addition, **intra-linguistic variation** of the semi-filled subpatterns with respect to semantic specialization and productivity

Discussion and conclusions

- **RQ2: Relation between productivity and semantic coherence?**
- **Inverse correlation between productivity and semantic coherence often confirmed (statistically confirmed for Dutch data set with 20 items)**
 - High productivity and low semantic coherence (= flexible schemas)
 - DU [*een pracht van een N*]
 - FR [*ce(tte) putain/saloperie/saleté de N*]
 - Low productivity and high semantic coherence (= fixed schemas)
 - DU [*een boom van een PERSON*]
 - ENG [*a whale of a CONCEPT*]

Discussion and conclusions

- **RQ2: Relation between productivity and semantic coherence?**

- **But also, exceptions to this tendency:**

- Specific subpatterns combine high productivity and high semantic coherence:
 - DU [*een joekel van een* OBJECT]
 - FR [*ce con de* PERSON]
 - IT [*quel genio di* PERSON]

→ **A semantically constrained construction can still be productive!**

- Larger datasets and further statistical analysis needed to confirm these results

References

- Aarts, B. (1994). The syntax of binominal noun phrases in English. *Dutch Working Papers in English Language and Linguistics* 30, 1-28.
- Aarts, B. (1998). Binominal noun phrases in English. *Transactions of the Philological Society* 9(1), 117-158.
- Barðdal, J. (2008). *Productivity: Evidence from case and argument structure in Icelandic*. Amsterdam: John Benjamins.
- Foolen, A. (2004). Expressive binominal NPs in Germanic and Romance languages. In: G. Radden & K.-U. Panther (red.), *Studies in linguistic motivation*. Berlin: De Gruyter, 75-100.
- Larrivée, P. (1994). Quelques hypothèses sur les structures syntaxique et sémantique de *Ce fripon de valet*. *Revue québécoise de linguistique* 23(2), 101-113.
- Paardekooper, P.C. 1956. Een schat van een kind. *De nieuwe taalgids* 49. 93-99.
- Piunno (in press) *Quel genio di lingua*. Analisi quantitativa e qualitativa delle costruzioni fraseologiche valutative dell'italiano e confronto con lo spagnolo e il francese. In E. Lavric, C. Konecny, C. Konzett-Firth & M. Messner (eds.). *Actes du IXe Colloque International 'Linguistique contrastive germano-romane et intraromane'*, Berlin: Frank & Timme.
- Regula, M. (1965). Zur Analyse der Fügung *ce fripon de valet*. *Linguistica* 7/1, 57-59
- ten Wolde, E. (2023). *The English Binominal Noun Phrase: A Cognitive-Functional Approach*. Cambridge: Cambridge University Press.
- Tutescu M. (1969). Le type nominal *ce fripon de valet*. *Revue de linguistique romane* 33, 299-316.
- Van Goethem, K. (2023). *Een pracht van een constructie vs. een prachconstructie*. Een vergelijkende corpusstudie van de syntactische en morfologische Expressieve Binominale Constructie. *Nederlandse Taalkunde* 28(2), 325-348.
- Verhagen, A. (2005). Constructiegrammatica en 'usage-based' taalkunde. *Nederlandse Taalkunde* 10, 197-222.

