



**GHENT  
UNIVERSITY**

# INTER- AND INTRALINGUAL SUBTITLING IN DUTCH-SPEAKING FLANDERS

## INSIGHTS FROM A MULTIMETHODOLOGICAL LINGUISTIC RESEARCH PROJECT

---

Gert De Sutter & Lynn Prieels  
17/01/2019

# INSTITUTIONAL CONTEXT



## ACADEMIC BIBLIOGRAPHY

Add pu

Search 200 years of publications by Ghent University researchers.

- **FWO project conducted by**

### **Lynn Prieels at Ghent University**

- Measuring norm conformity in Belgian-Dutch subtitling: a multivariate corpus study
- 2014-2018
- PhD defended on 22 June 2018
- Open-access publication:  
<https://biblio.ugent.be/publication/8565224>

[Search results](#) > Balancing between language policy and...



### Balancing between language policy and language reality : a corpus-based multivariate study on linguistic norm adherence in Belgian-Dutch subtitling

Lynn Prieels (UGent)

(2018)

Author

Lynn Prieels (UGent)

Promoter

Gert De Sutter (UGent)

Organization

Department of Translation, Interpreting and Communication

#### Downloads

Proefschrift LynnPrieels ok.pdf

full text | [open access](#) | PDF | 2.60 MB



Download

#### Citation

# LINGUISTIC CONTEXT

- **The language situation in Dutch-speaking Flanders is complex; it is characterised by late standardisation, pluricentricity and demotisation.**

# LINGUISTIC CONTEXT

- **The language situation in Dutch-speaking Flanders is complex; it is characterised by late standardisation, pluricentricity and demotisation.**
  - Whereas The Netherlands started developing a standard language from the 16<sup>th</sup> century onwards, the **northern part of Belgium was isolated** from this development for centuries due to foreign occupation and a ‘francocentric’ language policy.
  - After Dutch was officially accepted as one of Belgium’s languages in the early 20<sup>th</sup> century, it was decided (after some time) to **adopt the standard language developed in The Netherlands**, and this (exogeneous) standard language was massively spread through educational programmes at schools and through radio and tv shows.
  - Side effect of this ‘hyperstandardisation’: **strong standard language ideology** (the only best language is the standard language), stigmatisation of substandard varieties, strong metalinguistic awareness.

# LINGUISTIC CONTEXT

- **The language situation in Dutch-speaking Flanders is complex; it is characterised by late standardisation, pluricentricity and demotisation.**
  - Until the 1990's the linguistic-normative centre in the entire Dutch language area (The Netherlands and Belgium) was located in the so-called *Randstad* (Amsterdam, Rotterdam, Utrecht, Leiden): **monocentrism**; typical Flemish words and constructions were considered non-standard.
  - From the 1990's onwards there is a gradual shift towards **bicentrism**, with a separate linguistic centre for the Netherlandic part of the language area and one for the Flemish part (Brussels-Leuven-Antwerp): typical Flemish words and constructions were considered standard language in Belgium.



General Standard  
Dutch



Netherlandic  
Standard Dutch



Belgian Standard  
Dutch



Non-standard  
Dutch

# LINGUISTIC CONTEXT

- **The language situation in Dutch-speaking Flanders is complex; it is characterised by late standardisation, pluricentricity and demotisation.**
  - Whereas the written variety of standard language is uncontested (highly prestigious, highly valued and widely used by language users), the **spoken variety of the standard language is under pressure:** highly valued ('pure', 'professional', 'beautiful'), highly prestigious, but rarely used ('unnatural').
  - Since the 1990's a new, colloquial variety has developed, with both features of the standard language and the Flemish dialects: this colloquial variety, which is called '**tussentaal**' is widely used, for informal but also increasingly for formal interaction; it is used by teachers, politicians, VIPs, in private and professional life as well as on television and radio. Despite its popularity (it is perceived as the most natural spoken language variety, associated with 'comfortable, closeby, authenticity, coolness'), it is **still a stigmatised variety** (the language of the sloppy, of those who are easily satisfied).

# OBJECTIVES

- **General objective of research project: to what extent does the Flemish public broadcaster VRT use tussentaal colloquialisms in its subtitles? And which contextual parameters influence the occurrence of colloquialisms?**
  - Hypothesis 1: on the one hand, the public broadcaster is legally obliged to use (and spread) the standard language, both in its written and spoken form (language policy), hence subtitlers consistently use standard language only.
  - Hypothesis 2: on the other hand, tussentaal is omnipresent on television (fiction, comedy, interviews with VIPs...), subtitlers will have a tendency to use tussentaal features in their subtitles (either as a consequence of an unconscious transfer or consciously in order to retain authenticity).

# OBJECTIVES

- **General objective of research project: to what extent does the Flemish public broadcaster VRT use tussentaal colloquialisms in its subtitles? And which contextual parameters influence the occurrence of colloquialisms?**
  - Subproject 1: To what extent does VRT use (subtle) lexical colloquialisms, how does this compare to other types of translations and non-translations, and which contextual parameters influence the choice between lexical colloquialisms and ‘standardisms’?
  - Subproject 2: To what extent does VRT use (salient) lexical and grammatical colloquialisms and which contextual parameters influence the choice between lexical colloquialisms and ‘standardisms’?
  - Subproject 3: To what extent are colloquialisms in VRT subtitles influenced by the spoken input?
  - (Subproject 4: What do subtitlers themselves think of their choices?)

# DATA AND METHOD

- **Corpus-based approach**

- We used two existing corpora and one self-compiled corpus to answer the research questions: **Dutch Parallel Corpus** (10M tokens, translations and their source texts in several genres from and into Belgian Dutch), **SONAR Corpus** (which contains a VRT subtitles component: ca. 2M words, 2000-2005, closed subtitles, several genres), **our own self-made parallel subtitles corpus** (10-minute-transcriptions and parallel subtitles of 40 Flemish television programmes across 5 genres, 2014-2016).

- **Profile-based approach**

- We extract colloquialisms in combination with its standard alternatives in order to control for topical differences (e.g. standard *handtas* vs. colloquial *sacoche*) [Speelman et al. 2003]

- **Multifactorial approach**

- We analyse different linguistic profiles simultaneously with profile-based **correspondence analysis** (R-package *corregp*; Plevoets 2015).



# DATA AND METHOD

## 1. Lexical-paradigmatic profiles

- e.g. *appelsien* – *sinaasappel* ‘orange’
- Source: *Van Dale dictionary (2015)* – label ‘BE, spreektaal’

## 2. Constructional-paradigmatic profiles

- e.g. *durven + inf*      *durven + te + inf*    ‘to dare (to)’
- Source: *website Dutch Language Union*

## 3. Morphological-paradigmatic profiles

- e.g. *liefke*      *liefje*    ‘darling’
- Source: *website Dutch Language Union*

## 4. Syntagmatic profiles

- e.g. *aux + part + inf* / *aux + inf + part* / *part + aux + inf* ‘verbal end group’
- Source: *website Dutch Language Union*

# DATA AND METHOD

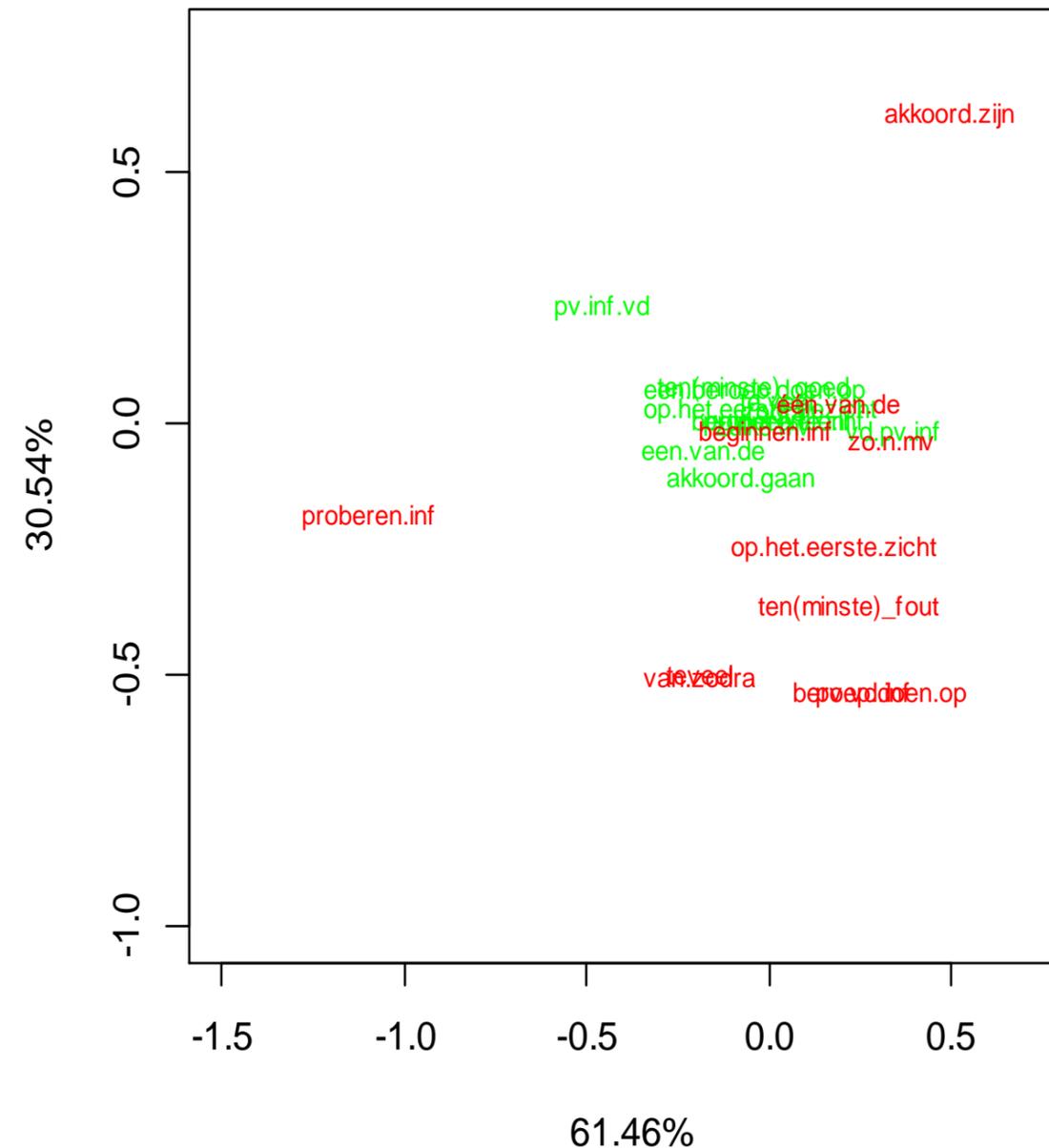
- All extracted data are **manually validated**: mutual interchangeability of variants in 1 profile
- And manually annotated for 3 contextual parameters:
  - program genre (Creeber 2008): fiction, children's television, documentaries, light entertainment, comedy (*coding instructions available, interannotator agreement = 100%*).
  - Source language: English, Netherlandic Dutch, Belgian Dutch
  - Speaker status: **actor**/interviewee, voice-over/presenter/interviewer (**non-actor**) (see next slide for principal dimensions)

<b>Genre</b>	<b>Main purpose</b>	<b>Target audience</b>	<b>Cast</b>
<b>Fiction</b>	entertain	all ages	Actors
<b>Comedy</b>	make laugh	all ages	Actors
<b>Children's television</b>	Inform + entertain	children	Non-actors
<b>Light entertainment</b>	Inform + entertain	all ages	Non-actors
<b>Documentaries</b>	inform	all ages	Non-actors

Profiles	Variety	Non-translated Dutch	Translated Dutch (<EN)	Translated Dutch (<FR)	Subtitles
akkoord gaan met	SN	45	11	30	32
akkoord zijn met	TT	6	0	1	13
part + aux + inf	SN	29	16	19	91
aux + inf + part		33	19	59	41
aux + part + inf	TT	27	3	3	20
zulke + plural noun	SN	57	39	19	173
zo'n + plural noun	TT	3	0	0	11
een beroep doen op	SN	87	29	108	31
beroep doen op	TT	18	3	7	6
proberen te + inf	SN	19	9	15	117
proberen + inf	TT	2	1	2	1
op het eerste gezicht	SN	13	10	4	32
op het eerste zicht	TT	3	0	0	5
beginnen te + inf	SN	14	6	10	54
beginnen + inf	TT	9	2	4	26
zodra	SN	93	37	103	225
van zodra	TT	13	1	5	9

# RESULTS SUBPROJECT 1

- Subproject 1: To what extent does VRT use (subtle) lexical colloquialisms, how does this compare to other types of translations and non-translations, and which contextual parameters influence the choice between lexical colloquialisms and 'standardisms'?

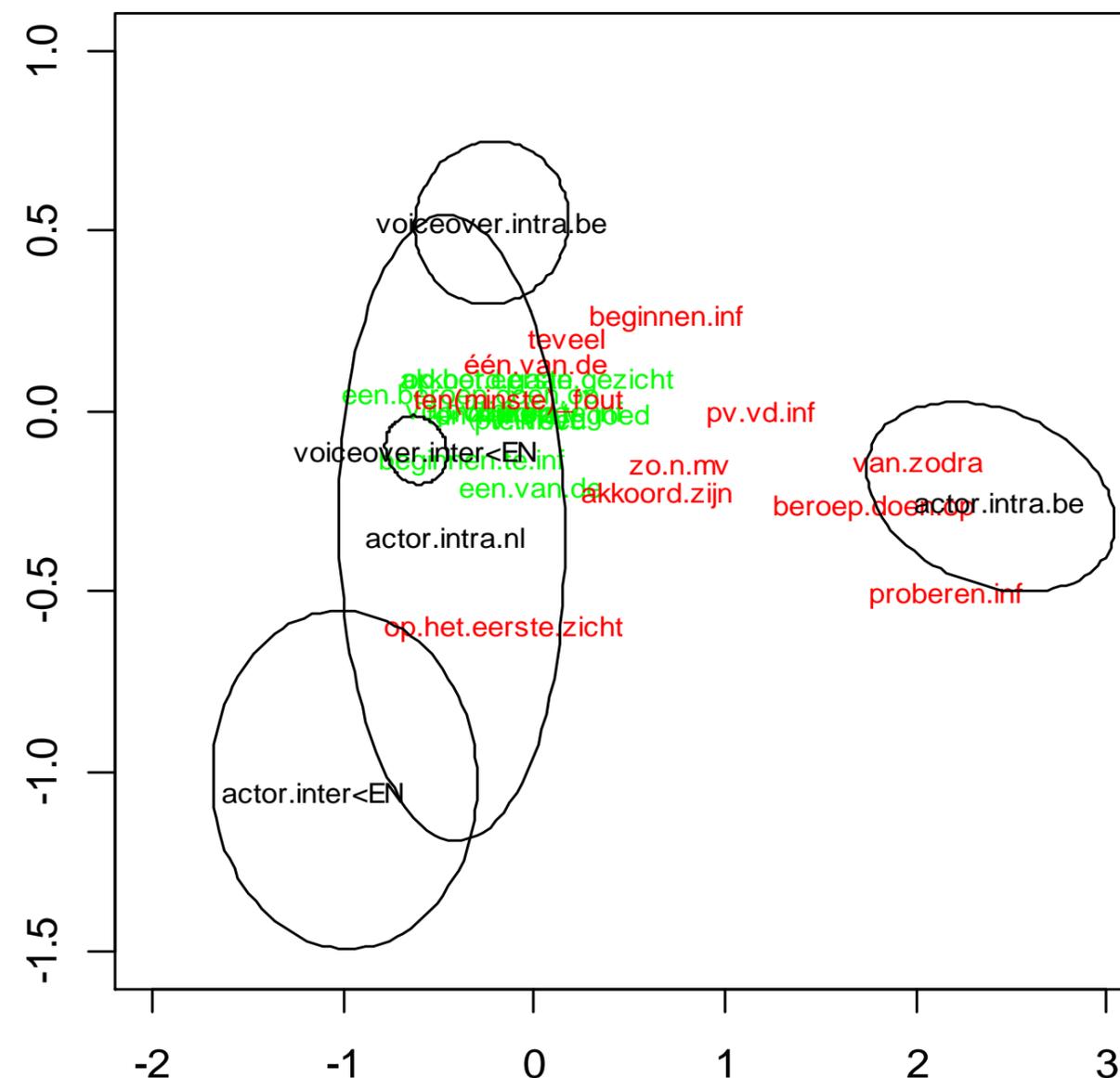






# RESULTS SUBPROJECT 1

- Subproject 1: To what extent does VRT use (subtle) lexical colloquialisms, how does this compare to other types of translations and non-translations, and which contextual parameters influence the choice between lexical colloquialisms and ‘standardisms’?



# RESULTS SUBPROJECT 1

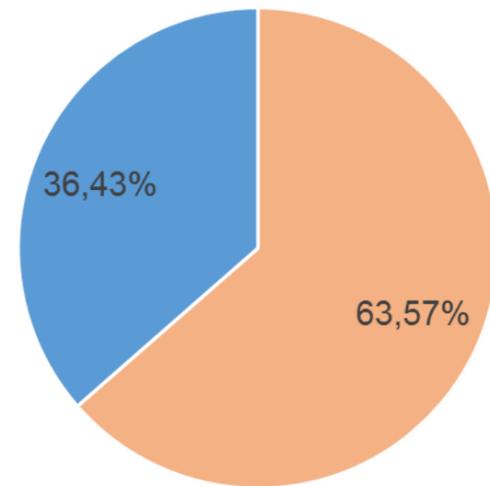
- *Source language of the speaker* is an important factor in the linguistic choices of subtitlers.
  - Interlingual subtitles of English speakers
  - Intralingual subtitles of Netherlandic speakers
  - Intralingual subtitles of Belgian speakers
- *Program genre* is an important factor in the linguistic choices of subtitlers.
  - Documentaries and children's television
  - Light entertainment and fiction
  - Comedy



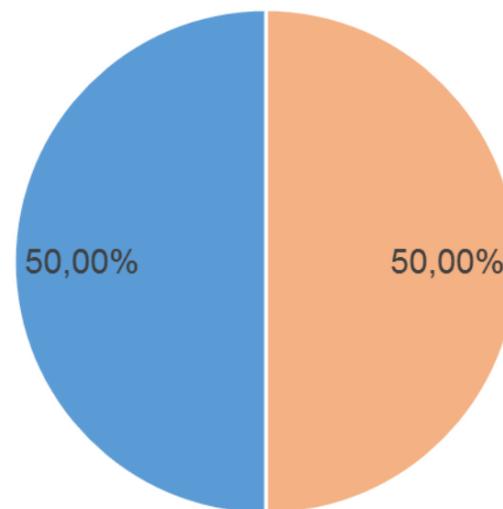
# RESULTS SUBPROJECT 3

- Subproject 3: To what extent are colloquialisms in VRT subtitles influenced by the spoken input?

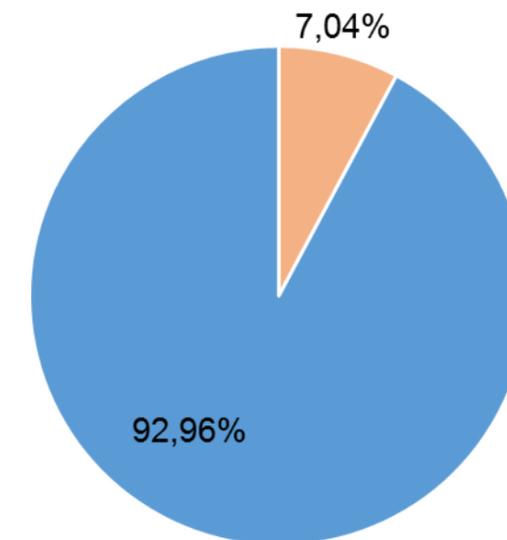
Lexical



Syntagmatic

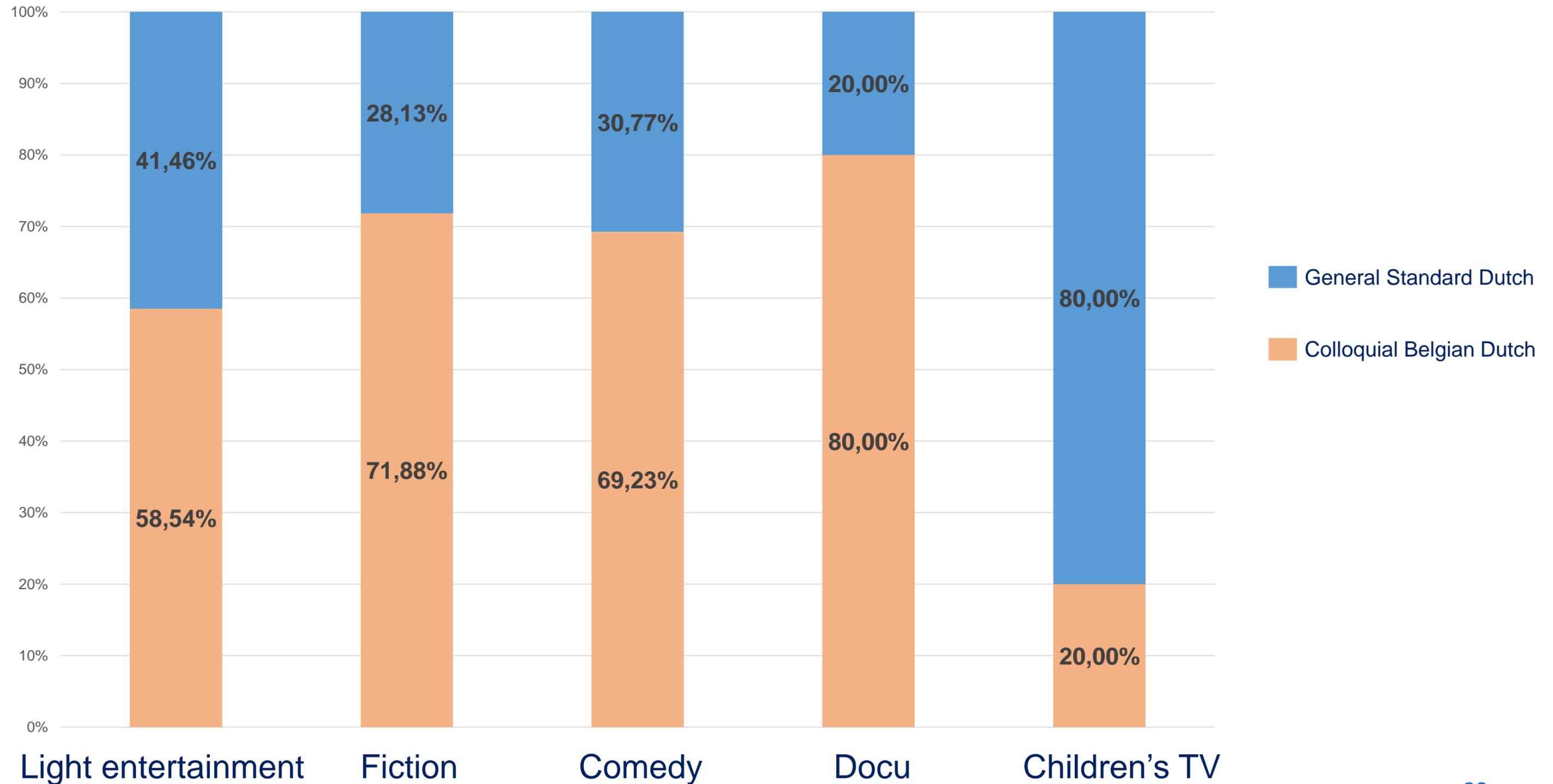


Morphological



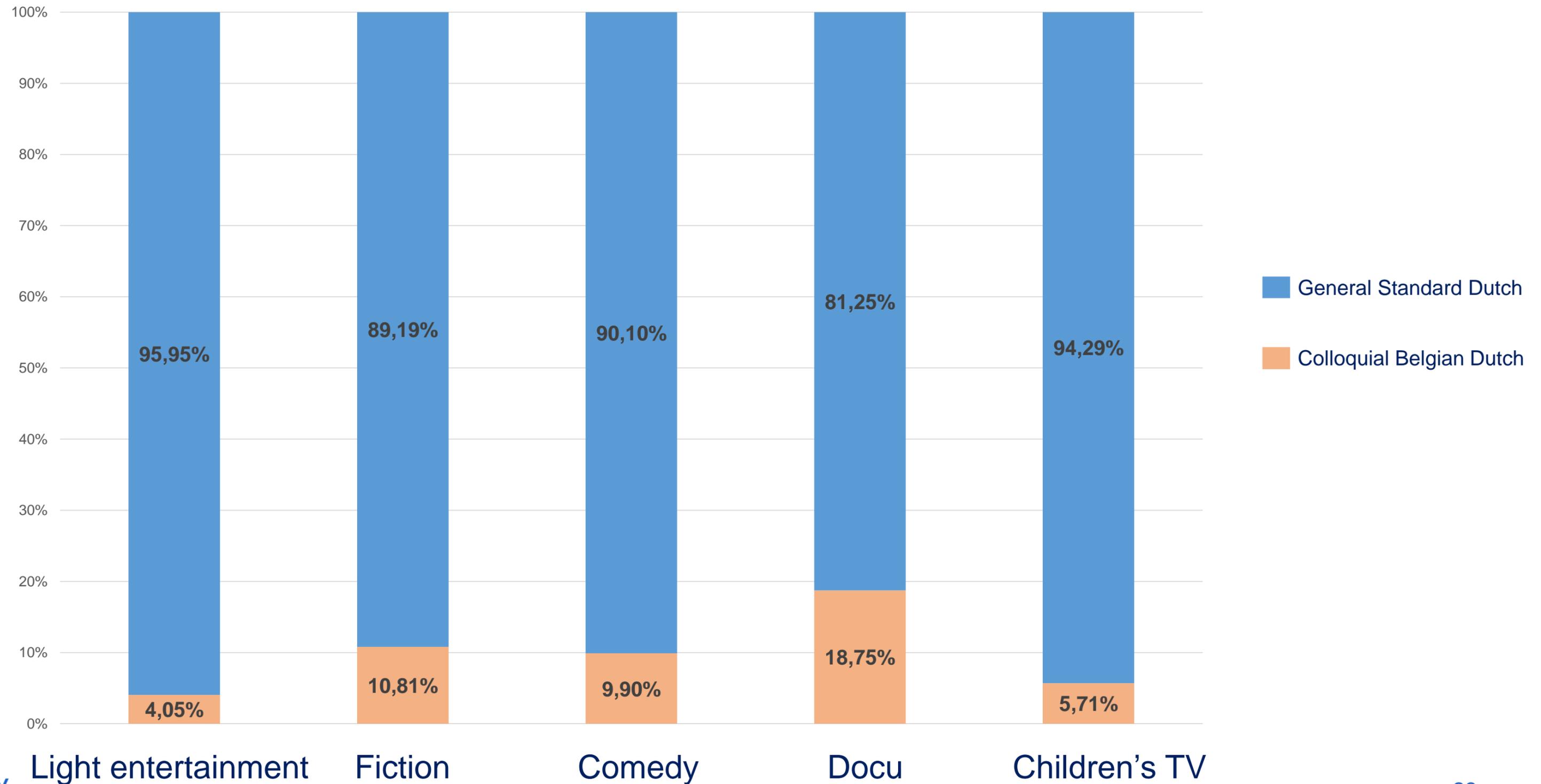
# THE RESULTS

## Lexical-paradigmatic profiles



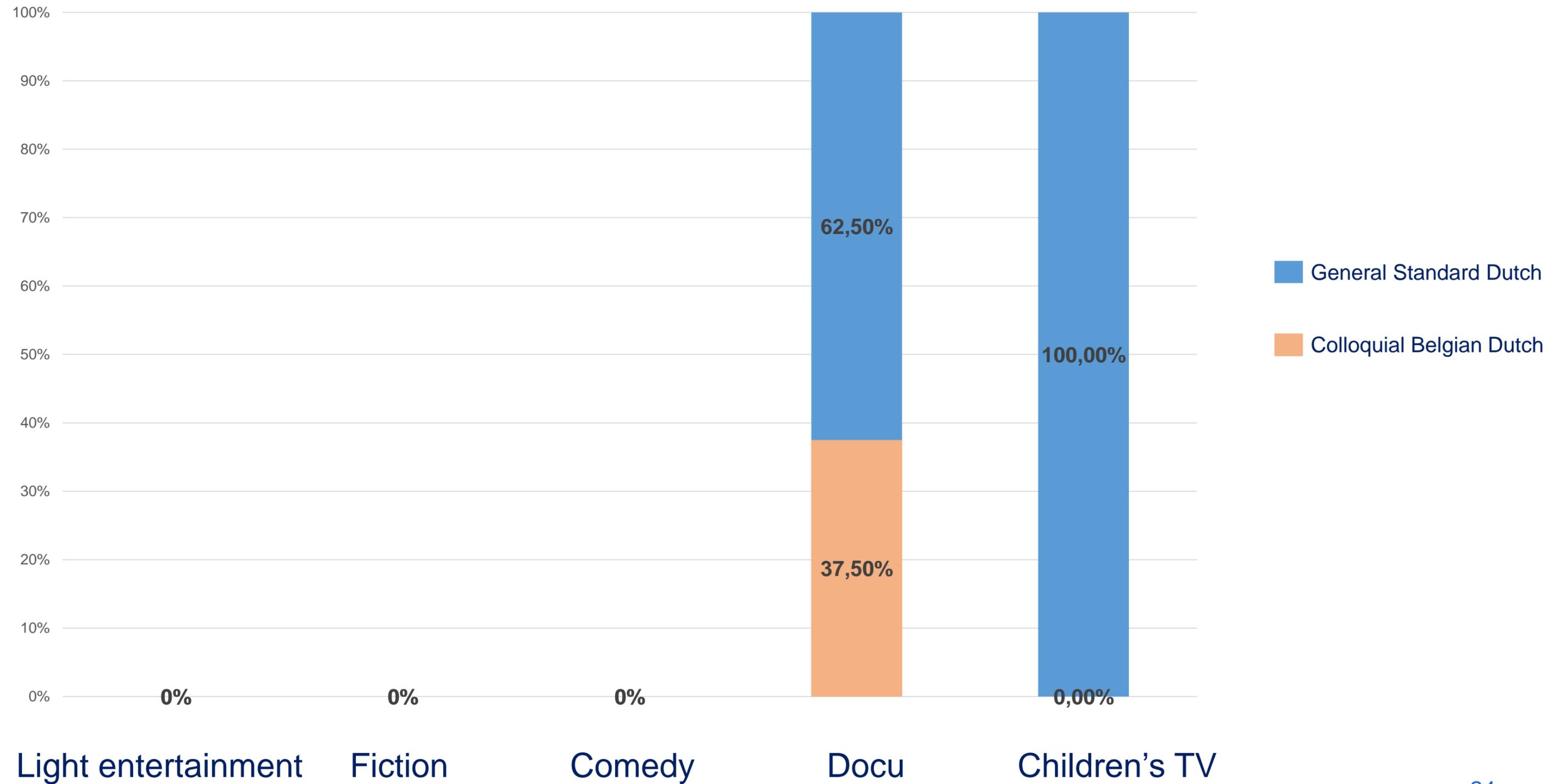
# THE RESULTS

## Morphological-paradigmatic profiles



# THE RESULTS

- Syntagmatic profiles



# RESULTS SUBPROJECT 4

Subproject 4: What do subtitlers themselves think of their choices? (Interviews and live observation)

- Onze papa deed dat ook met *den ajuin*. (speech)
- Onze papa deed dat ook met **ajuin**. (subtitle)
  
- Goh *wete*, ik vond *da* toen een heel **schoon** kind, maar als ik nu foto's zie, dan denk ik van: oh, zo **schoon** was *dieje* precies toch *nie*. (speech)
- Ik vond dat toen een **mooi** kind, maar als ik nu foto's zie, denk ik: zo **schoon** was die toch niet. (subtitle)

Tussentalige woorden moeten normaal gezien altijd overgenomen worden, maar *schoon* **klinkt veel "Vlaamser"** dan *ajuin*, waardoor je als ondertitelaar sneller de neiging hebt om die variant op te schonen.

# RESULTS SUBPROJECT 4

- Is er al cava? Ah, ja. *Wilde* mij *is* een **glaske** geven? (speech)  
Is er al cava? Ah, ja. Wil je mij eens een **glaasje** geven? (subtitle)
- Euh ja, dan hebben wij niks te doen hè **make**, of wel? (speech)
- Ja, euh... Dan hebben wij niks te doen, hè **make**? Of wel?(subtitle)

Het verkleinwoord op *-ke* wordt **sporadisch toegelaten** in fictieprogramma's, waar de personages vaak zulke **koosnaampjes** gebruiken om hun familie of geliefden aan te spreken.

# RESULTS SUBPROJECT 4

- Om **de Luc** e plezier te doen, *da's* alles. (speech)
- Om **Luc** een plezier te doen. Dat is alles. (subtitle)
  
- En die hogere risicopremie, die zal **moeten betaald worden**. (speech)
- Die risicopremie zal **moeten betaald worden**. (subtitle)

Bij last-minute ondertiteling kan het gebeuren dat **onopvallende** grammaticale fouten door **tijdsdruk** toch in de ondertitels worden overgenomen, omdat ze door de ondertitelaar **niet opgemerkt** worden.

# RESULTS SUBPROJECT 4

- Although we were allowed to interview some of the VRT subtitlers and to look what they do while subtitling, it was our impression that they felt rather uncomfortable (cf. strong standard language ideology).
- The VRT subtitlers do not seem to be supported very much by the policy-makers at VRT:
  - **Subtitle style guide** (VRT Stijlboek Teletekst ondertiteling (2009))
    - Lexicon: *tussentaal* is reproduced in the subtitles as much as possible (authenticity!)
    - Syntax: ungrammatical constructions are corrected (e.g. ge/gij/u; meiske; de Jan)
    - Note: the presenter is considered to use General Standard Dutch. If not, *tussentaal* lexicon is reproduced in the subtitles.

# GENERAL CONCLUSIONS

- The use of Colloquial Belgian Dutch or *tussentaal* is no longer limited to the spoken registers, since it also appears in the speech-to-text language of an institution that is considered important in policymaking.
- Especially in informal/entertaining contexts, lexical *tussentaal* features are tolerated to a large extent in the subtitles (grammatical features to a much lesser extent).

# Gert De Sutter

Associate Professor

DEPARTMENT OF TRANSLATION,  
INTERPRETING AND COMMUNICATION

E gert.desutter@ugent.be

T +32 9 331 19 59

[www.ugent.be](http://www.ugent.be)

 Ghent University

 @ugent

 Ghent University