

Syntactic variation along a continuum of subtitles: preliminary results on pronominal placement and third-person anaphoric direct objects

Doctoral student: Lívia Oliveira Azevedo
(Federal University of São Carlos - UFSCar)
Supervisor: Caroline Carnielli Biazolli, PhD
(Federal University of São Carlos - UFSCar)



Background

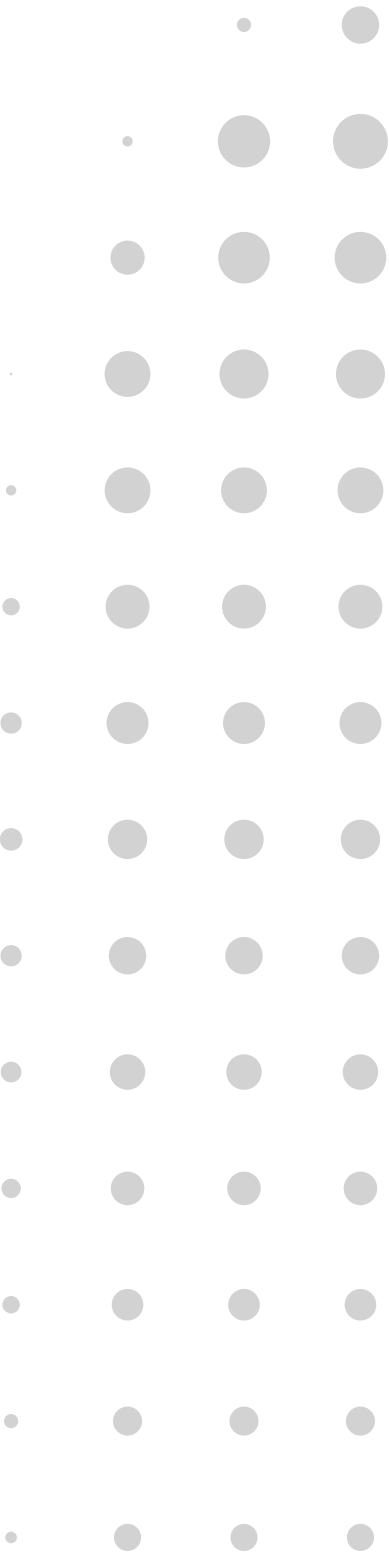
- Language Variation and Change (Labov, 1972/2008, 1994, 2001; Weinreich *et al.*, 1968/2006);





Background

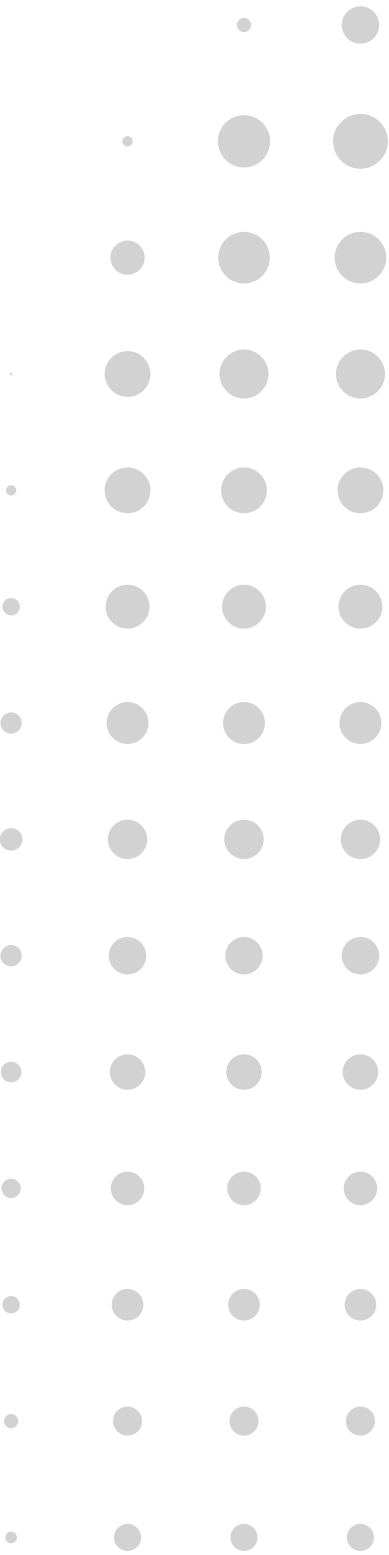
- Language Variation and Change (Labov, 1972/2008, 1994, 2001; Weinreich *et al.*, 1968/2006);
- Textual-discursive genres (Bakhtin, 1979/2016; Marcuschi, 2010);





Background

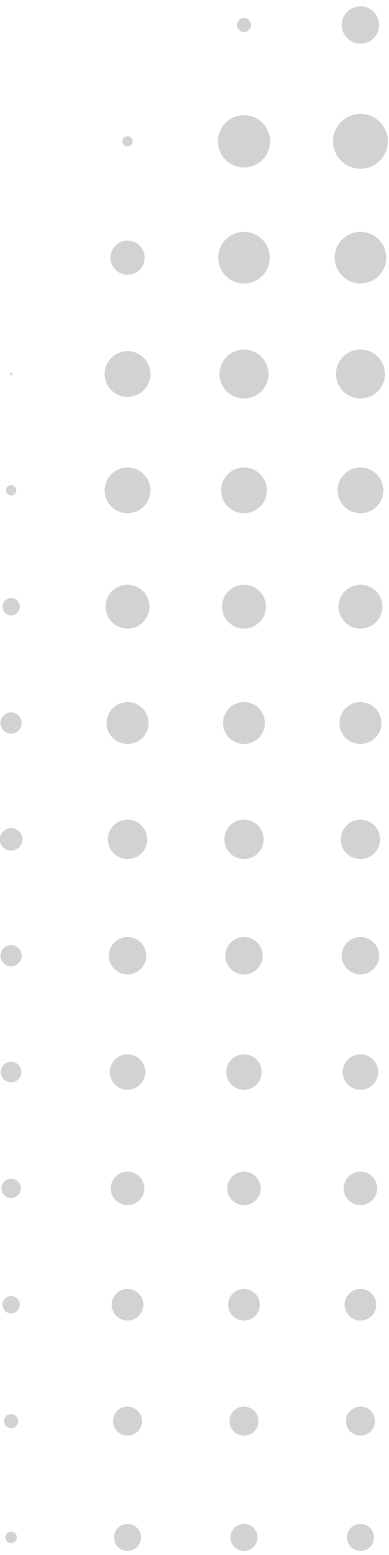
- Language Variation and Change (Labov, 1972/2008, 1994, 2001; Weinreich *et al.*, 1968/2006);
- Textual-discursive genres (Bakhtin, 1979/2016; Marcuschi, 2010);
- Variation and genres (Azevedo, 2024; Biazolli & Berlinck, 2021; Lima, 2022; Vieira & Lima, 2019);





Background

- Language Variation and Change (Labov, 1972/2008, 1994, 2001; Weinreich *et al.*, 1968/2006);
- Textual-discursive genres (Bakhtin, 1979/2016; Marcuschi, 2010);
- Variation and genres (Azevedo, 2024; Biazolli & Berlinck, 2021; Lima, 2022; Vieira & Lima, 2019);
- *Continua* (Bortoni-Ricardo, 2004, 2005, 2012; Marcuschi, 2010; Romaine, 1982/2009);





Background

- Language Variation and Change (Labov, 1972/2008, 1994, 2001; Weinreich *et al.*, 1968/2006);
- Textual-discursive genres (Bakhtin, 1979/2016; Marcuschi, 2010);
- Variation and genres (Azevedo, 2024; Biazolli & Berlinck, 2021; Lima, 2022; Vieira & Lima, 2019);
- *Continua* (Bortoni-Ricardo, 2004, 2005, 2012; Marcuschi, 2010; Romaine, 1982/2009);
- Linguistic norms and ideologies (Faraco, 2008; Lucchesi, 2002/2012; O'Neill, 2024).



Objectives

- Investigate pronominal placement and the realization of the third-person anaphoric accusative in audiovisual subtitles of different series from Netflix, taking into account the proposed *continuum*;





Objectives

- Investigate pronominal placement and the realization of the third-person anaphoric accusative in audiovisual subtitles of different series from Netflix, taking into account the proposed *continuum*;
- Report the stages of *corpus* composition and the extraction of the collected data;





Objectives

- Investigate pronominal placement and the realization of the third-person anaphoric accusative in audiovisual subtitles of different series from Netflix, taking into account the proposed *continuum*;
- Report the stages of *corpus* composition and the extraction of the collected data;
- Present the preliminary results obtained through initial statistical analyses.





The phenomena



Promoninal placement (cf. Biazolli, 2016; Lima, 2022; Pagotto, 1992; Pereira, 1981; Vieira, 2002, among others)



The phenomena



Promoninal placement (cf. Biazolli, 2016; Lima, 2022; Pagotto, 1992; Pereira, 1981; Vieira, 2002, among others)

- Proclisis

Se eu me apaixonar e ficar noiva,
será com alguém que nunca vi. (LBS01E01)



The phenomena

Promoninal placement (cf. Biazolli, 2016; Lima, 2022; Pagotto, 1992; Pereira, 1981; Vieira, 2002, among others)

- Proclisis

Se eu me apaixonar e ficar noiva,
será com alguém que nunca vi. (LBS01E01)

- Enclisis

Jonathan, pare! Pare! Você vai machucá-lo! (STS01E06)



The phenomena



Third-person anaphoric direct object (cf. Azevedo, 2024; Cyrino, 1994; Duarte, 1986; Freire, 2005; Lima, 2022; Marques de Sousa, 2021)



The phenomena



Third-person anaphoric direct object (cf. Azevedo, 2024; Cyrino, 1994; Duarte, 1986; Freire, 2005; Lima, 2022; Marques de Sousa, 2021)

- Accusative clitic

NA1: Uma lontra-marinha.

Sua pelagem grossa **a** mantém aquecida
na água fria. (OPS01E04)



The phenomena

Third-person anaphoric direct object (cf. Azevedo, 2024; Cyrino, 1994; Duarte, 1986; Freire, 2005; Lima, 2022; Marques de Sousa, 2021)

- Accusative clitic

NA1: Uma lontra-marinha.

Sua pelagem grossa **a** mantém aquecida na água fria. (OPS01E04)

- Null object

C1: Esqueci minha mochila.

NB1: Claro que esqueceu **[ø]**. (GFS01E08)



The phenomena



- Lexical pronoun

M1: Deixe-o ir!

T1: Para trás ou eu corto **ele**. (STS01E06)



The phenomena

- Lexical pronoun

M1: Deixe-o ir!

T1: Para trás ou eu corto **ele**. (STS01E06)

- Anaphoric noun phrase

D1: Pode ir, você tem problemas.

C2: Tenho **problemas**. (LBS01E04)

The TV shows



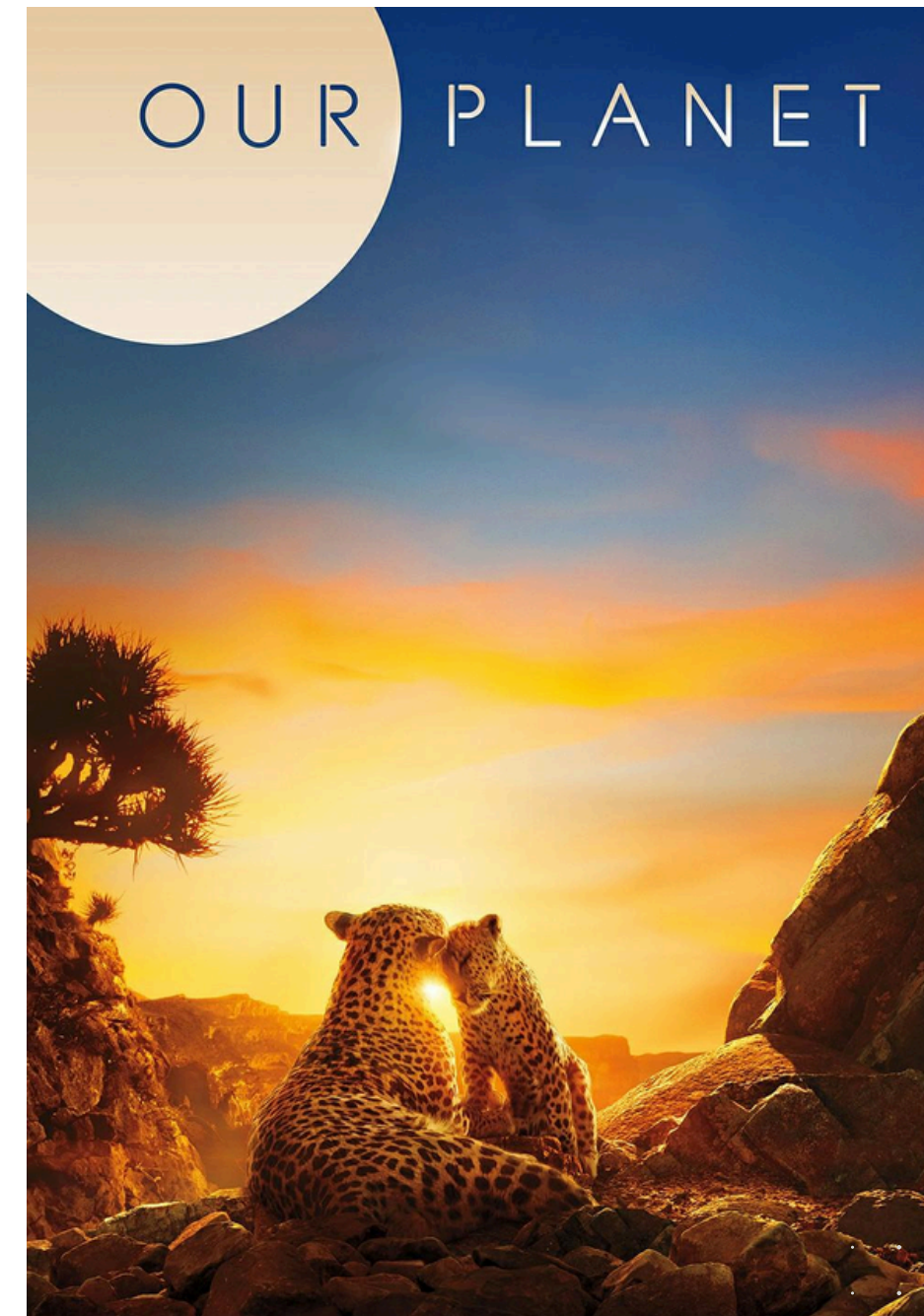
The TV shows



The TV shows



The TV shows





Methodology

Corpus **compilation**

Subtitle extraction
(*Netflix - subtitle
downloader*)

30,000 words per TV
show, collected in
the order of the
episodes

Methodology

Corpus compilation

Subtitle extraction
(*Netflix - subtitle
downloader*)

30,000 words per TV
show, collected in
the order of the
episodes



Grace and Frankie

32,962 words

Love Is Blind

33,944 words

Stranger Things

32,162 words

Our Planet

30,598 words

Total

129,666 words



Methodology

Corpus **compilation**

Subtitle extraction
(*Netflix - subtitle
downloader*)

30,000 words per TV
show, collected in
the order of the
episodes

Methodology

Corpus compilation

Subtitle extraction
(*Netflix - subtitle
downloader*)

30,000 words per TV
show, collected in
the order of the
episodes

Development of the continuum

Designing the
continuum based on
the parameters for
genre comparison
proposed by Biber
and Conrad (2009)

Methodology

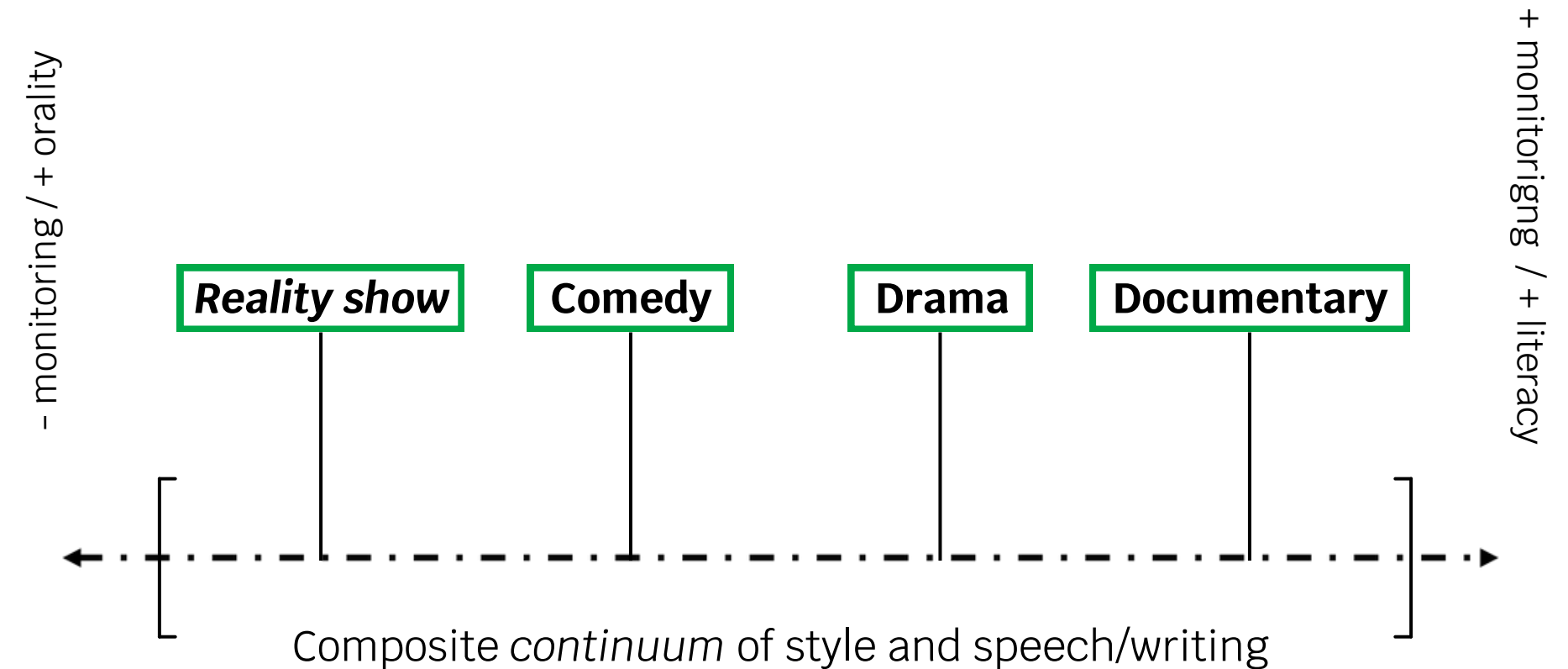
Corpus compilation

Subtitle extraction
(*Netflix - subtitle
downloader*)

30,000 words per TV
show, collected in
the order of the
episodes

Development of the *continuum*

Designing the
continuum based on
the parameters for
genre comparison
proposed by Biber
and Conrad (2009)



Methodology

Corpus compilation

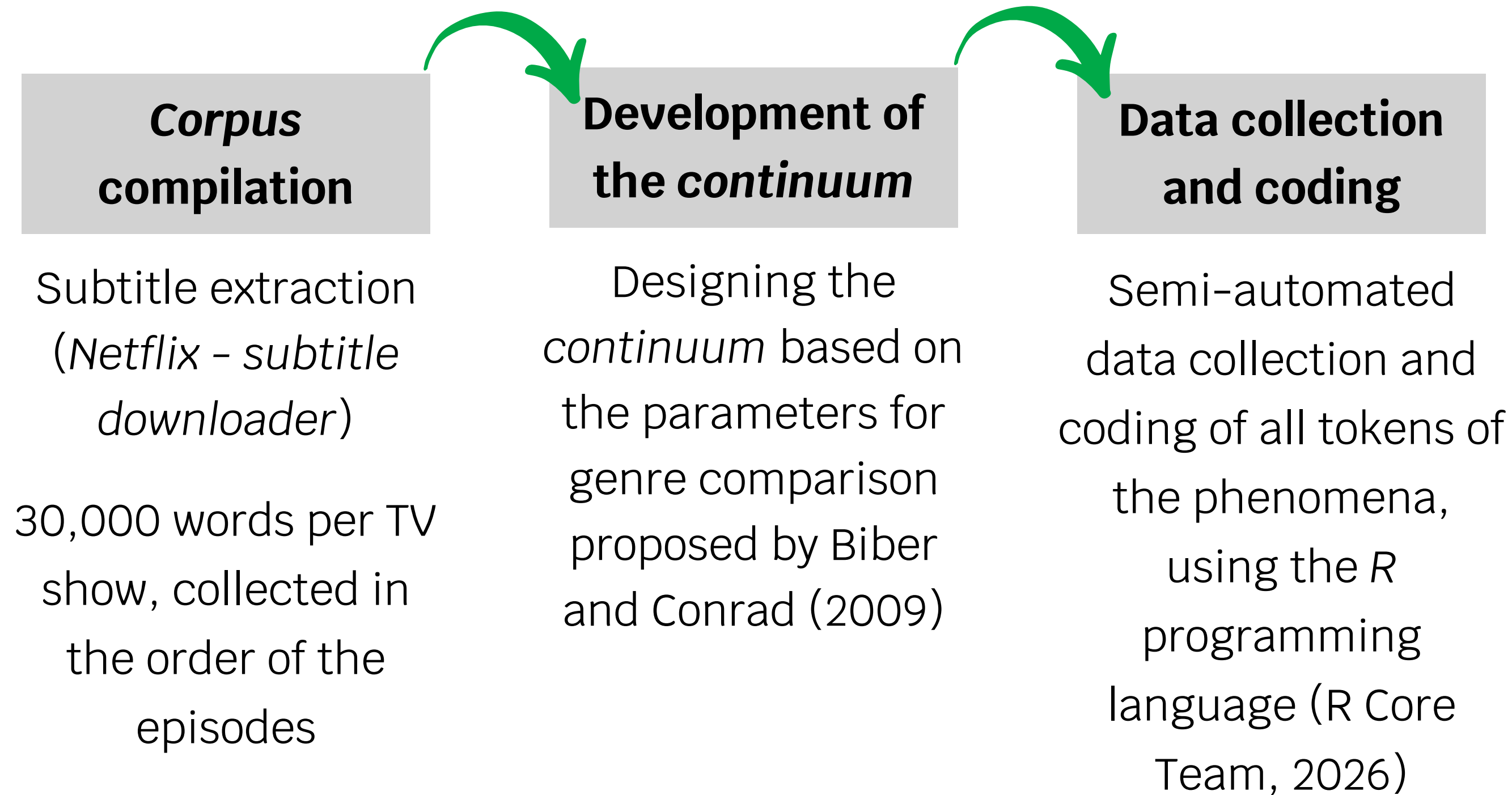
Subtitle extraction
(*Netflix - subtitle
downloader*)

30,000 words per TV
show, collected in
the order of the
episodes

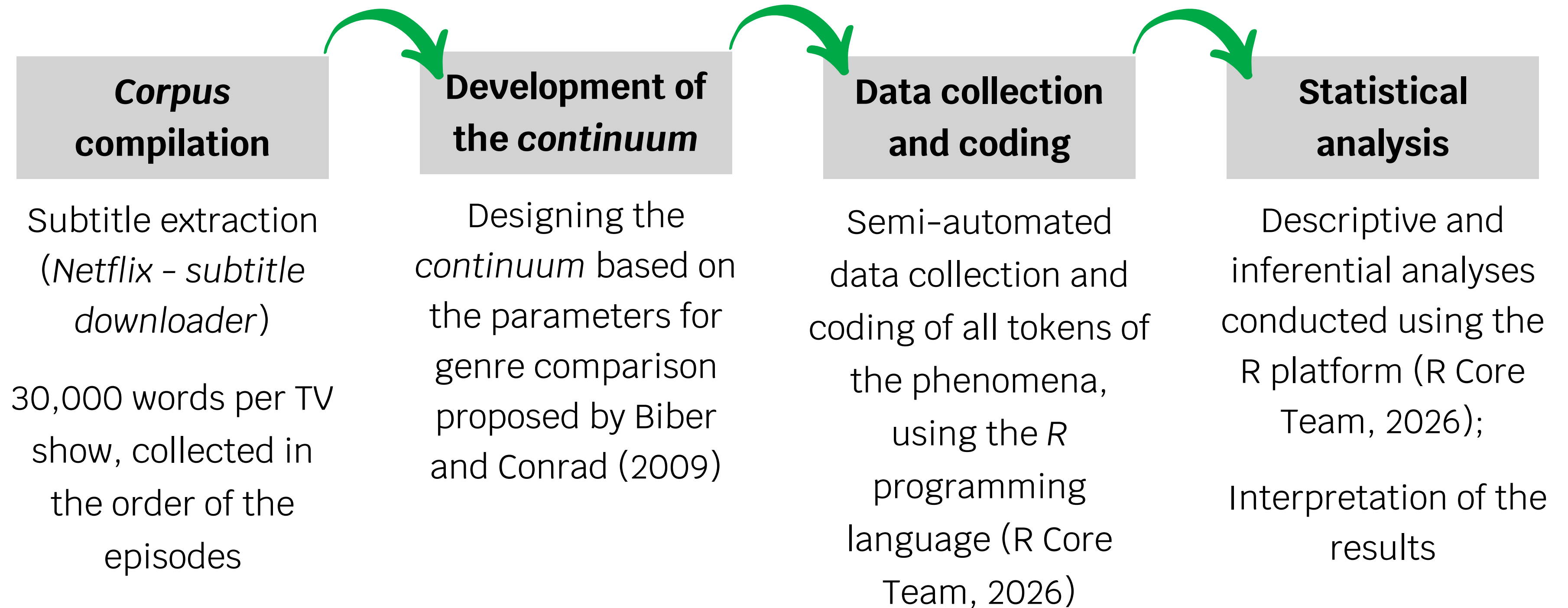
Development of the continuum

Designing the
continuum based on
the parameters for
genre comparison
proposed by Biber
and Conrad (2009)

Methodology



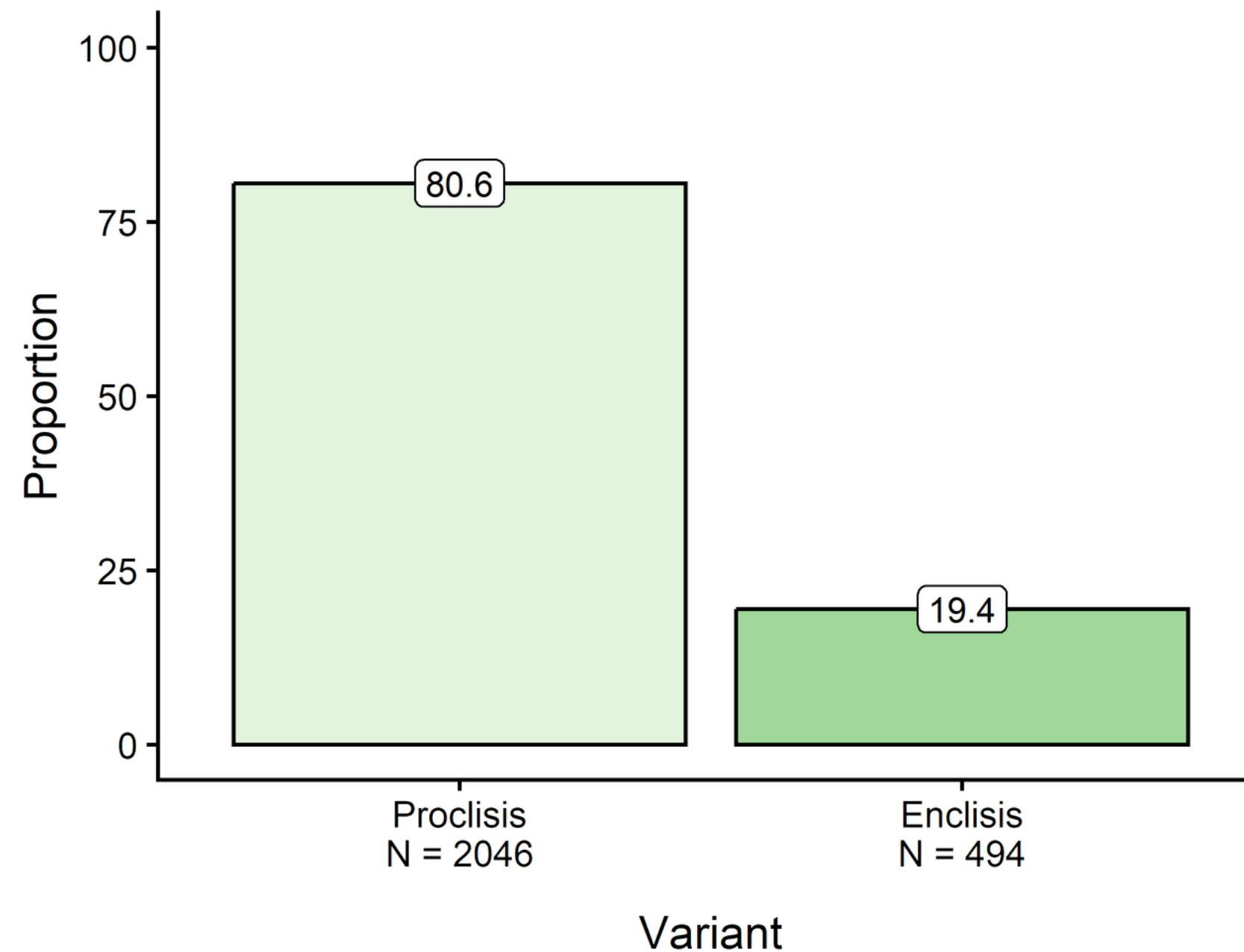
Methodology



Preliminary results

Figure 1

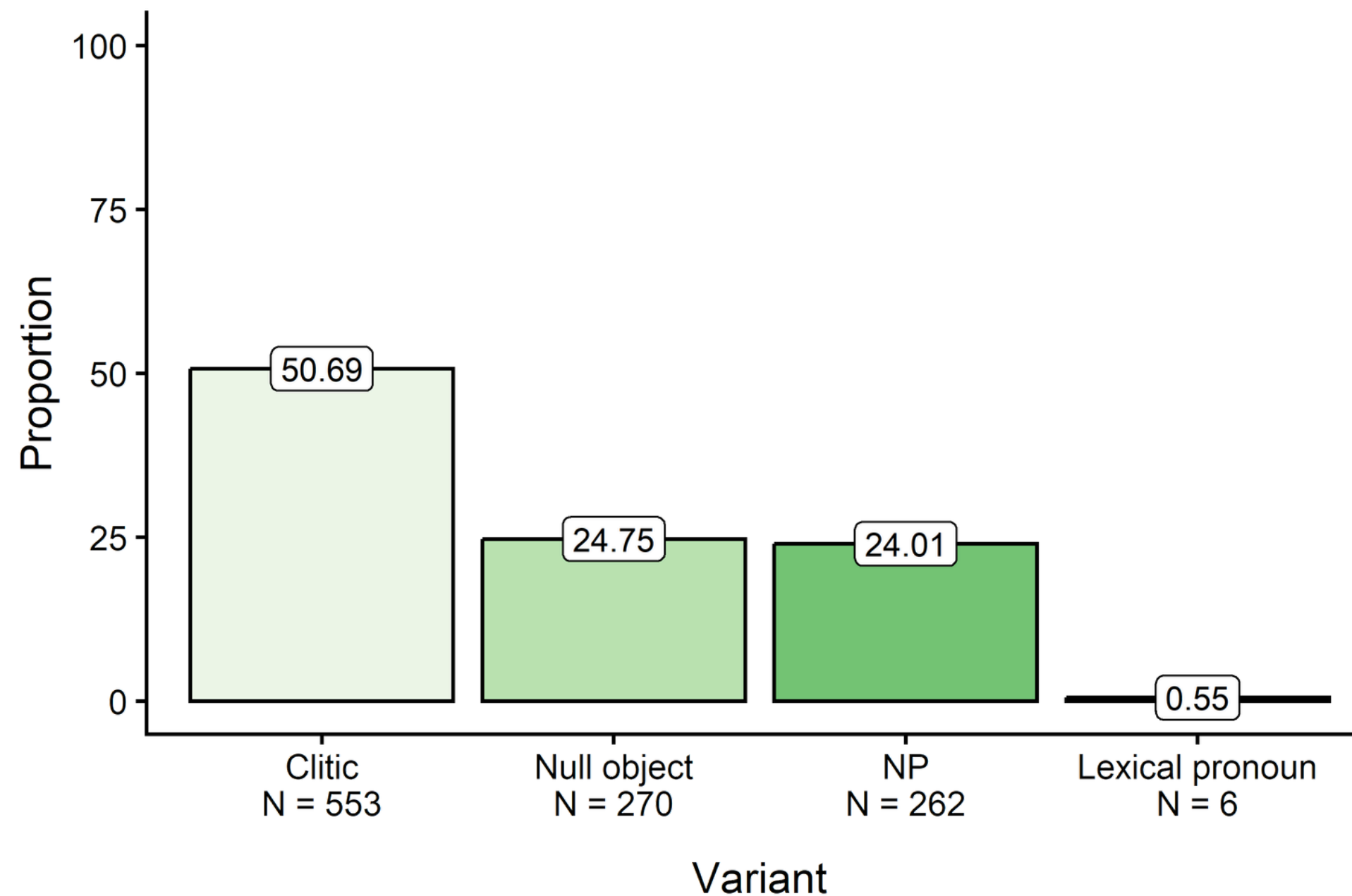
General distribution of pronominal placement variants in *Netflix* subtitles (N = 2540)



Preliminary results

Figure 2

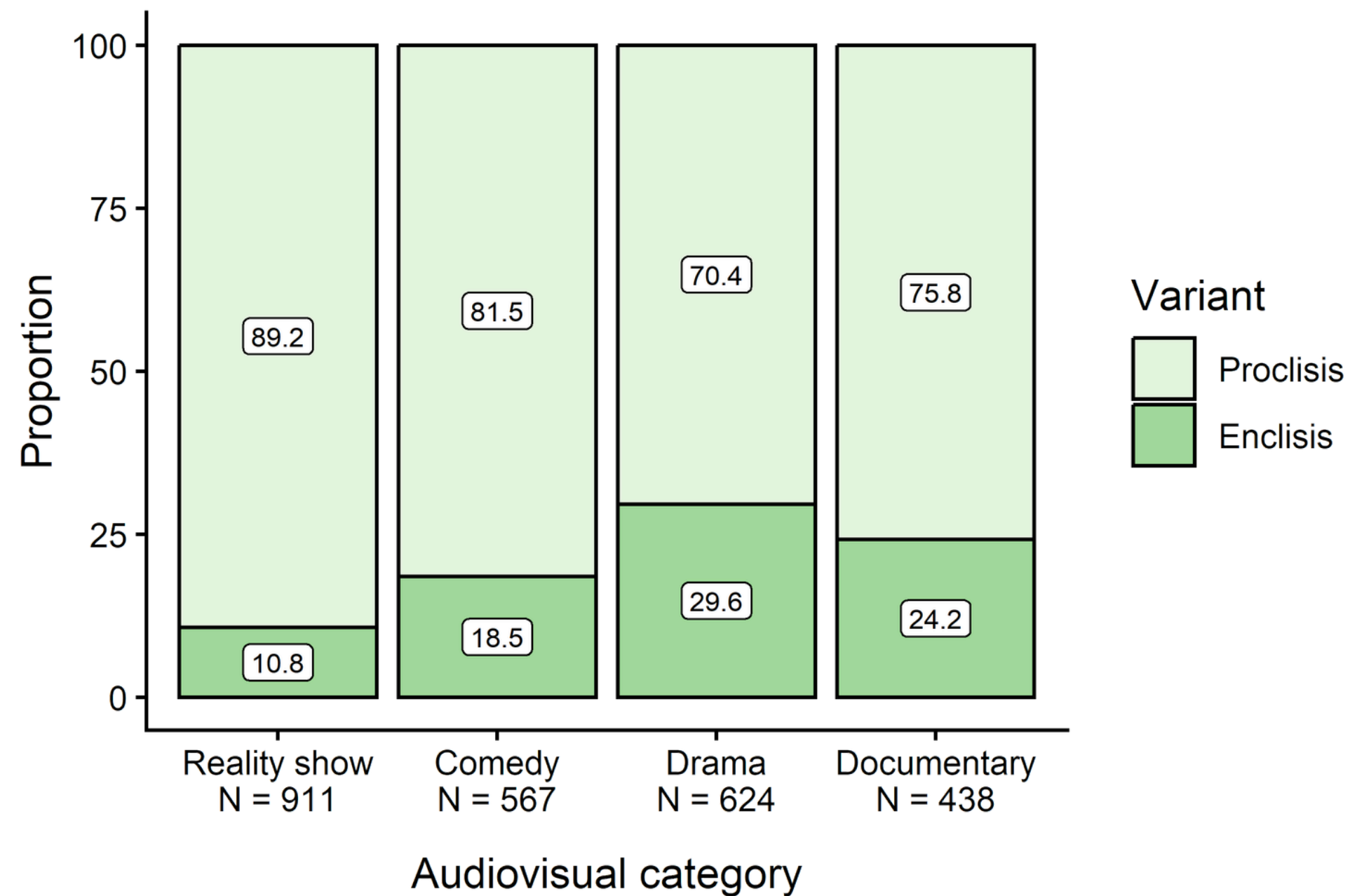
General distribution of third-person anaphoric direct object variants in *Netflix* subtitles (N = 1091)



Preliminary results

Figure 3

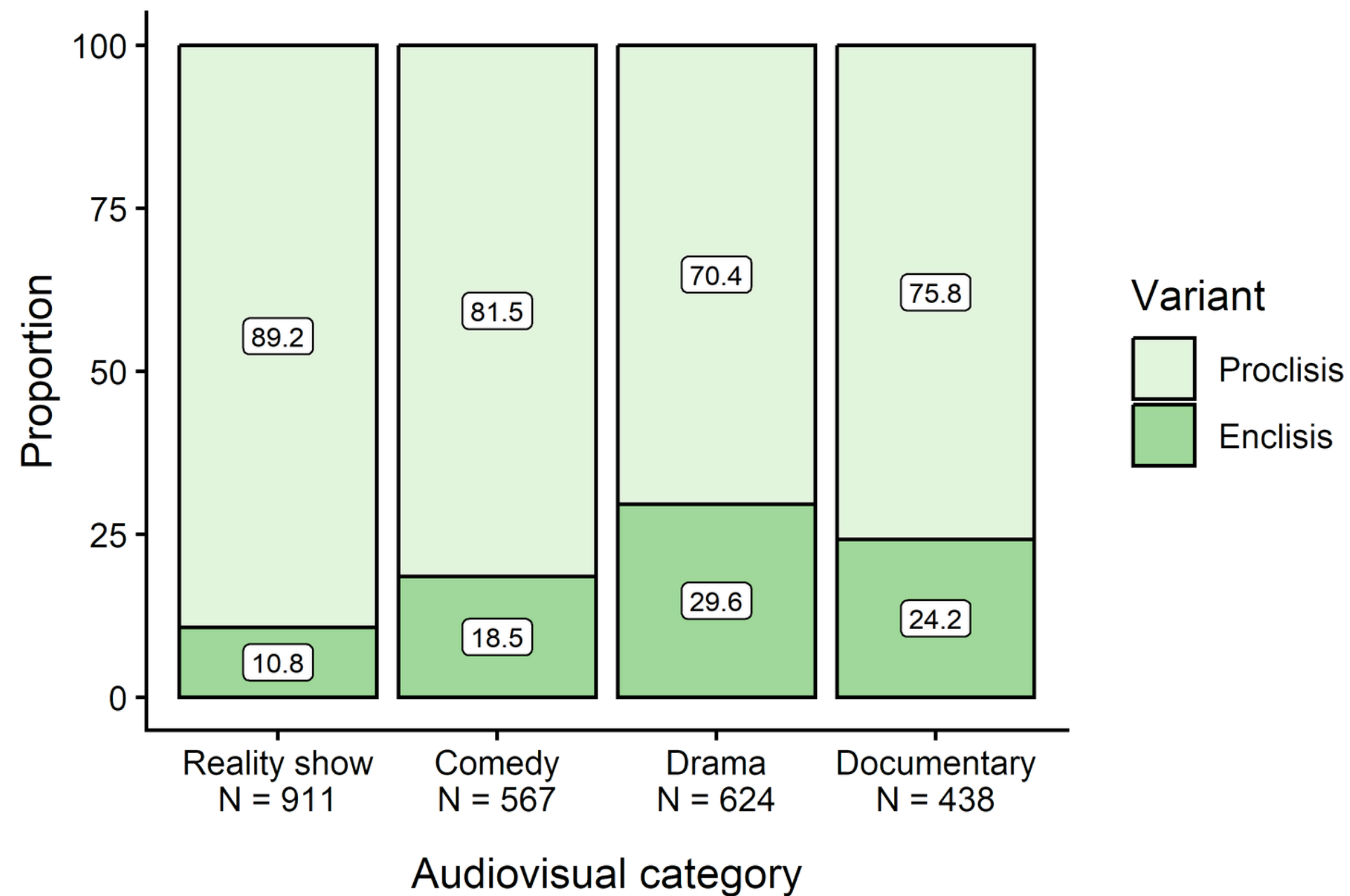
Distribution of pronominal placement variants in *Netflix* subtitles according to audiovisual category (N = 2540)



Preliminary results

Figure 3

Distribution of pronominal placement variants in *Netflix* subtitles according to audiovisual category (N = 2540)

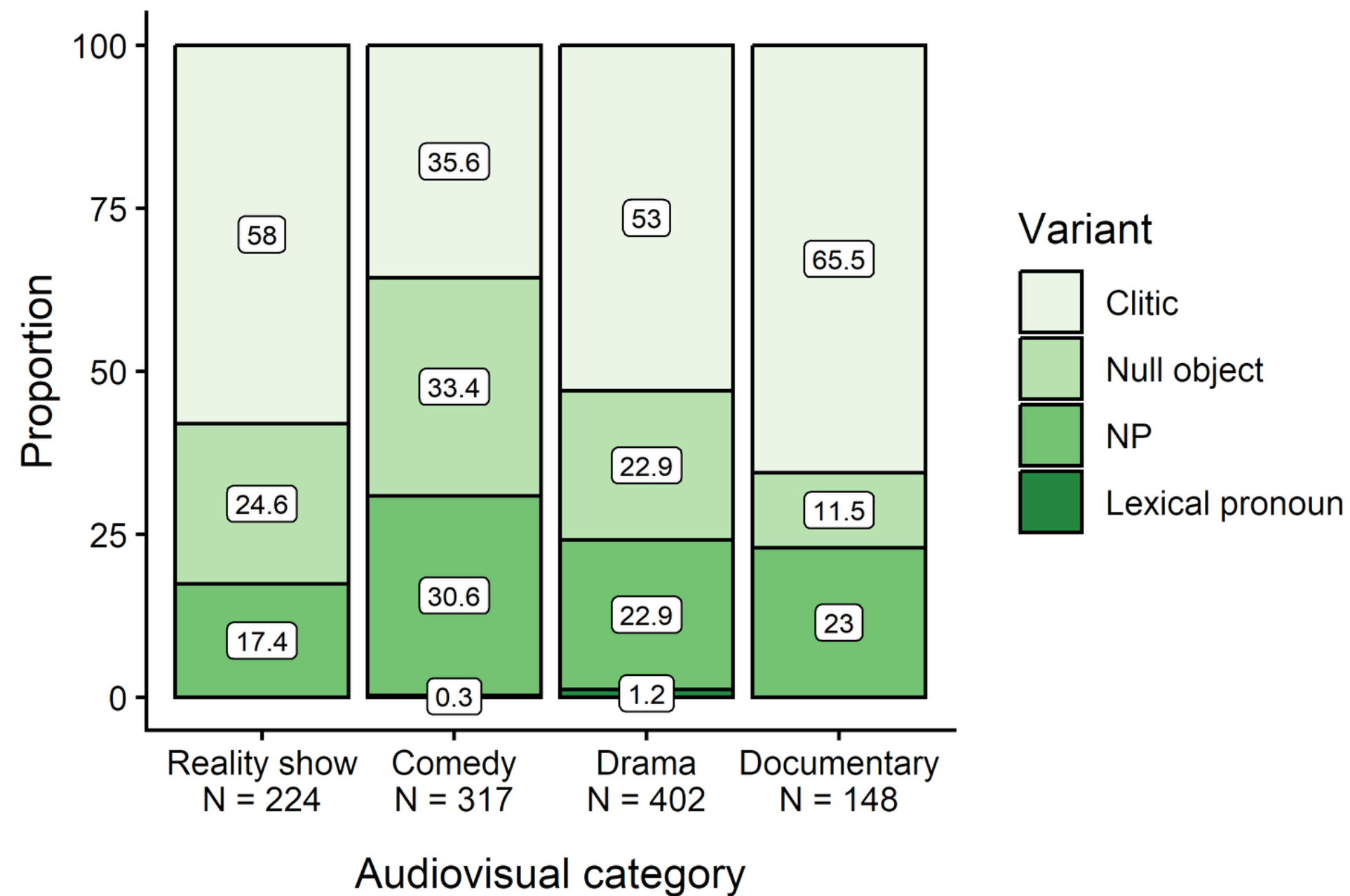


$p < 0.01$

Preliminary results

Figure 4

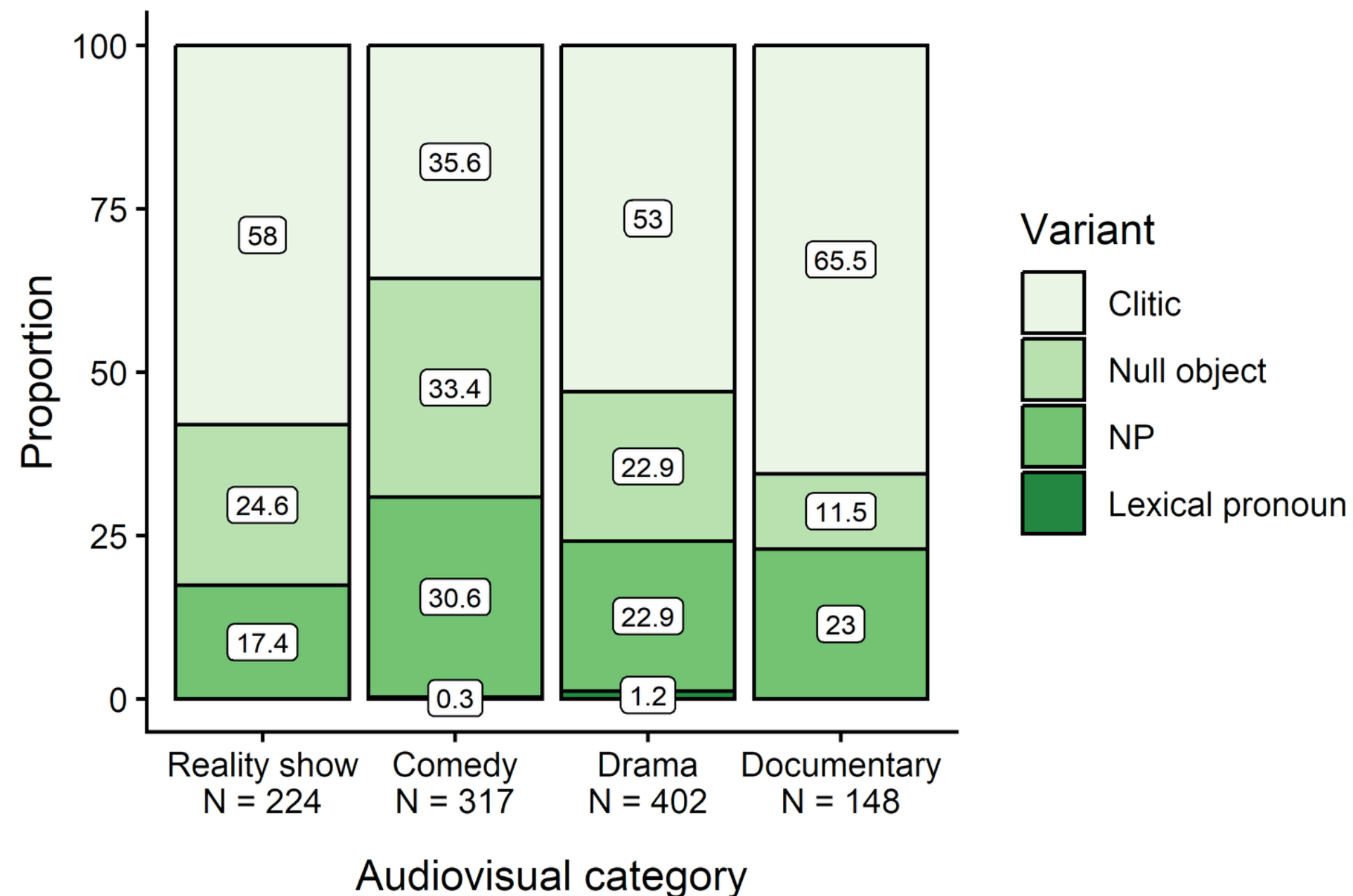
Distribution of third-person anaphoric direct object variants in Netflix subtitles according to audiovisual category (N = 1091)



Preliminary results

Figure 4

Distribution of third-person anaphoric direct object variants in Netflix subtitles according to audiovisual category (N = 1091)



$p < 0.01$



Future steps

- Complete the analysis and coding of the remaining linguistic and extralinguistic variables for both phenomena;



Future steps

- Complete the analysis and coding of the remaining linguistic and extralinguistic variables for both phenomena;
- Deepen the preliminary analyses and include logistic regression models to verify which variables actually exert an effect on the behavior of the phenomenon in the *corpus*;



Future steps

- Complete the analysis and coding of the remaining linguistic and extralinguistic variables for both phenomena;
- Deepen the preliminary analyses and include logistic regression models to verify which variables actually exert an effect on the behavior of the phenomenon in the *corpus*;
- Contrast the statistical data with the qualitative analysis of the Netflix subtitling manual, investigating how explicit guidelines for "clarity" and "adequacy" reflect linguistic norms and ideologies;



Future steps

- Complete the analysis and coding of the remaining linguistic and extralinguistic variables for both phenomena;
- Deepen the preliminary analyses and include logistic regression models to verify which variables actually exert an effect on the behavior of the phenomenon in the *corpus*;
- Contrast the statistical data with the qualitative analysis of the Netflix subtitling manual, investigating how explicit guidelines for "clarity" and "adequacy" reflect linguistic norms and ideologies;
- Compare the variation patterns with those of a third phenomenon: relativization strategies.

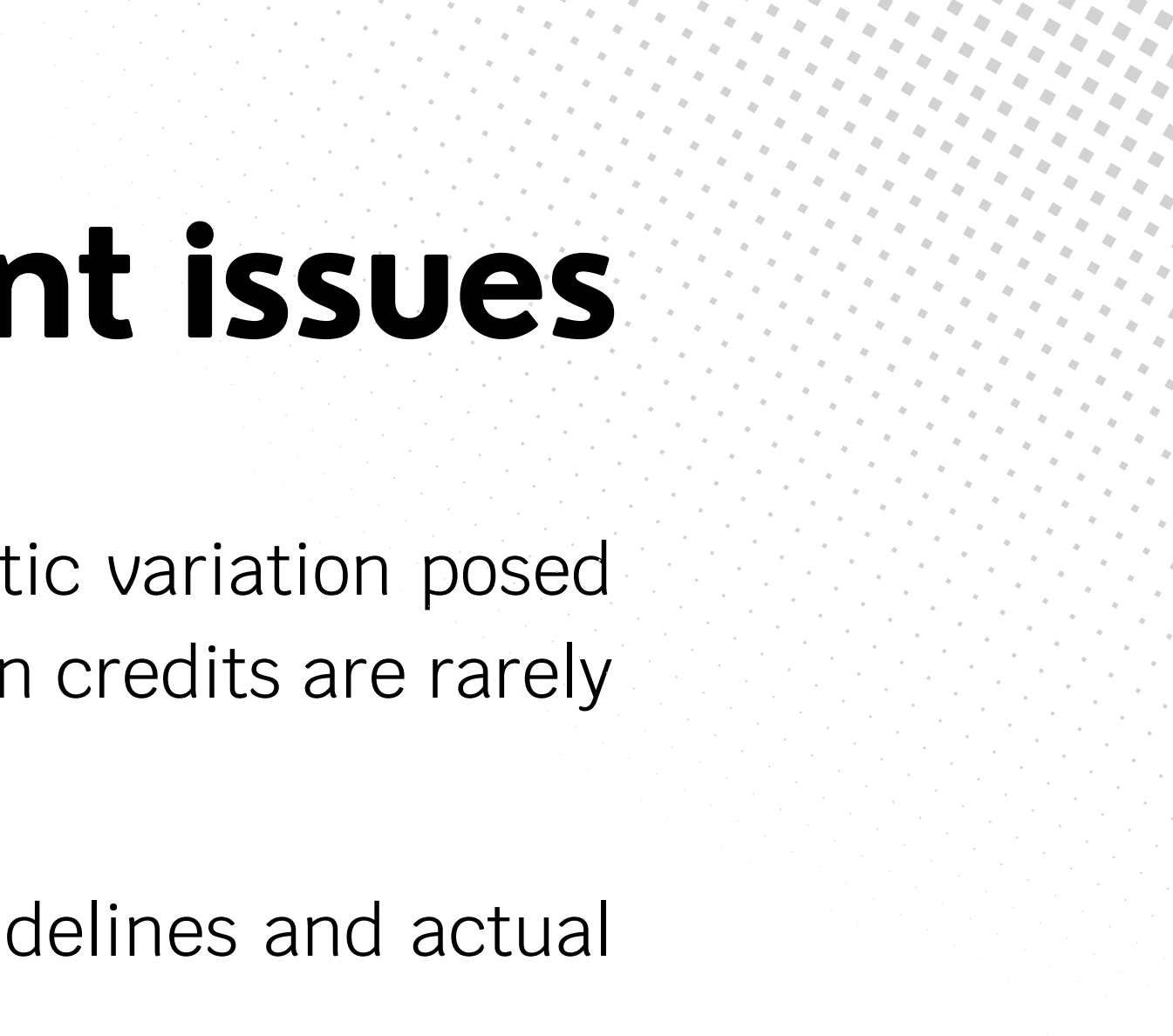


Current issues

- How to separate stylistic variation and the idiosyncratic variation posed by the subtitler, especially in a genre where translation credits are rarely given?



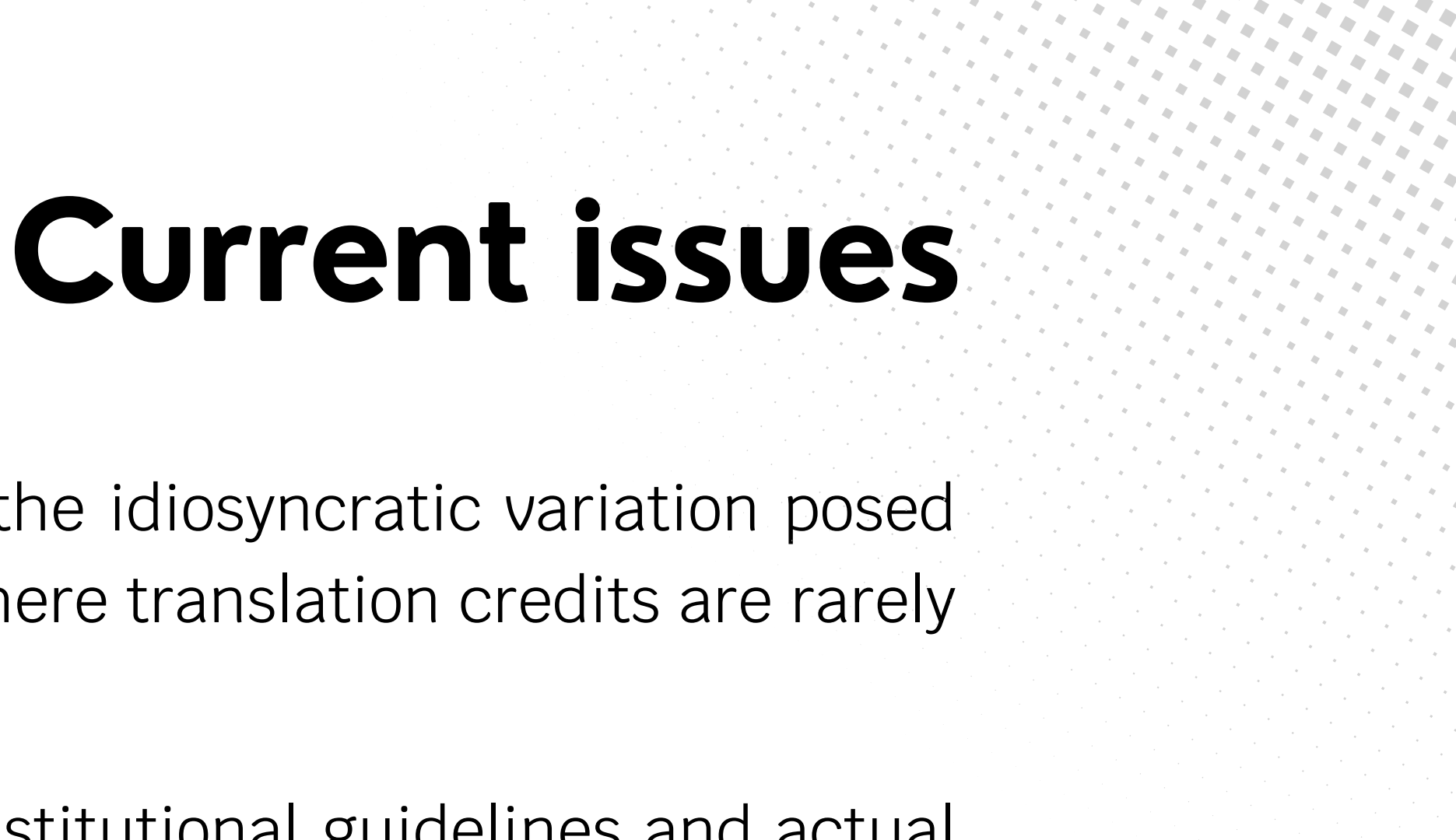
Current issues



- How to separate stylistic variation and the idiosyncratic variation posed by the subtitler, especially in a genre where translation credits are rarely given?
- How to evaluate the tension between institutional guidelines and actual translation practices?



Current issues



- How to separate stylistic variation and the idiosyncratic variation posed by the subtitler, especially in a genre where translation credits are rarely given?
- How to evaluate the tension between institutional guidelines and actual translation practices?
- When a genre that aims to emulate natural speech presents language closer to that of the written standard, are we facing a phenomenon of self-censorship generated by linguistic ideologies internalized by professionals, or an adaptive limit of the written modality itself?

References



Thank you!



livia.azevedo@estudante.ufscar.br