

Mixed Paradigms, Idiosyncratic Variation, and Deference in the Brazilian Portuguese Imperative

Candidate: Cecilia M. Ugartemendía

**Supervisors: Prof. Dr. Paul O'Neill
Prof. Dr. Johanna Wolf**



EUROPEAN PORTUGUESE - NON-DEFERENTIAL ADDRESS PARADIGM (TU)						
Pronoun	Nominative tu	Accusative te	Dative te	Reflexive te	Possessive teu(s)-tua(s)	+ preposition ti (<i>also contigo</i>)
Verb form	Positive imperative canta	Negative imperative não cantes	Present indicative canta	Present subjunctive cantes	Future indicative cantarás	Imperfect indicative cantavas

EUROPEAN PORTUGUESE - DEFERENTIAL ADDRESS PARADIGM (VOCÊ)						
Pronoun	Nominative você	Accusative o, no, a, na	Dative lhe	Reflexive se	Possessive seu(s) – sua(s)	+ preposition si (<i>also consigo</i>)
Verb form	Positive imperative cante	Negative imperative não cante	Present indicative canta	Present subjunctive cante	Future indicative cantará	Imperfect indicative cantava

THE PHENOMENON: VARIATION IN THE PRONOUN AND IMPERATIVE PARADIGM IN BRAZIL

- Historical loss of the *tu/você* deference distinction.
- Predominance of *você* in the Southeast (Lopes & Cavalcante 2011).
- *Você* as the preferred 2SG nominative, coexisting with etymological *tu* clitics (*te*, *teu*)
- → a **mixed paradigm**: forms from two historical sources within one paradigm.-

BRAZILIAN PARADIGM- NON-DEFERENTIAL ADDRESS MIXED PARADIGM (VOCÊ)						
Pronoun	Nominative você	Accusative Te (o, no, a, na)	Dative Te (lhe)	Reflexive se	Possessive seu(s) – sua(s)	+ preposition você
Verb form	Positive imperative CANTE - CANTA	Negative imperative NÃO CANTE – NÃO CANTA	Present indicative canta	Present subjunctive cante	Future indicative cantará	Imperfect indicative cantava

O senhor / a senhora: a new locus of deference — sharing the verbal paradigm with *você*.

EUROPEAN PORTUGUESE - DEFERENTIAL ADDRESS PARADIGM (VOCÊ)						
Pronoun	Nominative você	Accusative o, no, a, na (n)a	Dative lhe	Reflexive se	Possessive seu(s) – sua(s)	+ preposition si (<i>also consigo</i>)
Verb form	Positive imperative cante	Negative imperative não cante	Present indicative canta	Present subjunctive cante	Future indicative cantará	Imperfect indicative cantava

THE PHENOMENON: VARIATION IN THE PRONOUN AND IMPERATIVE PARADIGM IN BRAZIL

- Historical loss of the *tu/você* deference distinction.
- Predominance of *você* in the Southeast (Lopes & Cavalcante 2011).
- *Você* as the preferred 2SG nominative, coexisting with etymological *tu* clitics (*te*, *teu*)
- → a **mixed paradigm**: forms from two historical sources within one paradigm.-

BRAZILIAN PARADIGM- NON-DEFERENTIAL ADDRESS MIXED PARADIGM (VOCÊ)						
Pronoun	Nominative você	Accusative Te (o, no, a, na)	Dative Te (lhe)	Reflexive se	Possessive seu(s) – sua(s)	+ preposition você
Verb form	Positive imperative CANTE - CANTA	Negative imperative NÃO CANTE – NÃO CANTA	Present indicative canta	Present subjunctive cante	Future indicative cantará	Imperfect indicative cantava

The imperative in southeastern Brazil: what earlier studies show

Guy 1983; Sampaio 2001; Scherre 2002, 2004, 2007, 2008, 2012; Scherre et al. 2007; Scherre & Naro 2014; Barbosa et al. 2017; Lamberti & Schwenter 2018; Schwenter & Hoff 2023; O'Neill 2024; O'Neill & Ugartemendía 2025

(a) polarity

FAÇA (VOCÊ FORM) more likely in negative contexts

(b) immediacy

FAÇA (VOCÊ FORM) more likely in non-immediate contexts

(c) diatopy (region)

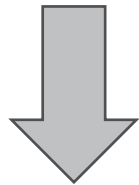
The true imperative (FAZ) predominates in the Southeast, Center-West and parts of the South; the suppletive (FAÇA) is more frequent in the Northeast

(d) formality

FAÇA (VOCÊ FORM) much more likely in formal contexts

+deference

FAÇA
(VOCÊ FORM)

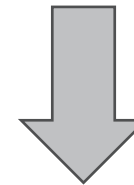


Senhora
(deferential)

immediate
affirmative

Non-Immediate
negative

+FAZ — **-FAÇA**
(TU FORM) (VOCÊ FORM)



você
(non deferential)

Methodology

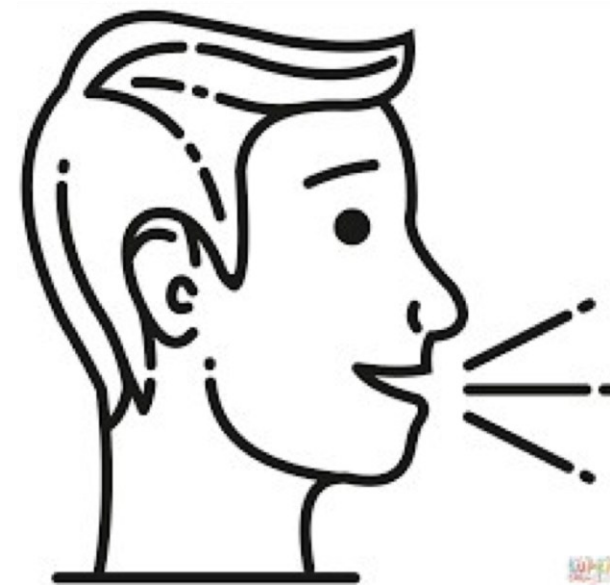
Elicitation via storyboards and flashcards;
controlled interactions.

Two interlocutor (context) conditions:

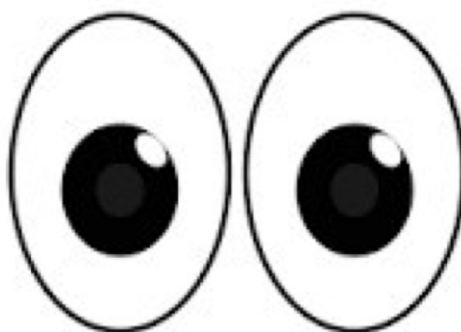
- *senhora* (deferential) vs. *você* (non-deferential)
- polarity (affirmative/negative) and immediacy (immediate/non-immediate)

– 10 speakers from Rio de Janeiro (19–21 yrs), all university students in related programs in the same University; 3 repetitions per speaker; 3 interviewers.





SEGUIR A PESSOA



VER ISSO



REPETITIONS

Storyboard: non-immediate

REPETITION A



Vai ao supermercado, **traz** 30 ovos, **faça** ovos para a Maria e **não esquece** de dar o remédio para ela às oito, **dá** às 8 horas o remédio.

*Go to the supermarket, **get** 30 eggs, **make** some eggs for Maria and **don't forget** to give her the medicine at eight o'clock, **give** her the medicine*

RED = etymologically derived from the paradigm TU

BLUE = etymologically derived from the paradigm VOCÊ

REPETITION B

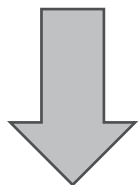


Ronaldo na semana que vem, **vai** ao supermercado, **traga** ovos, para as crianças você faz, **cê** faz ovos para elas comerem, e **não esquece** de dar o remédio da Ana todo o dia às 7. **Faça**... na semana que vem **faça** ovos nos jantares e **dê** o remédio todos os dias às 7.

*Ronaldo, next week, **go** to the supermarket, **get** some eggs, you can make them for the children, you make eggs for them to eat, and **don't forget** to give Ana her medicine every day at 7 o'clock. **Make**... next week, **make** eggs for dinner and **give** the medicine every day at 7 o'clock.*

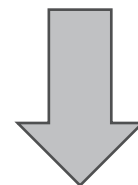


FAÇA



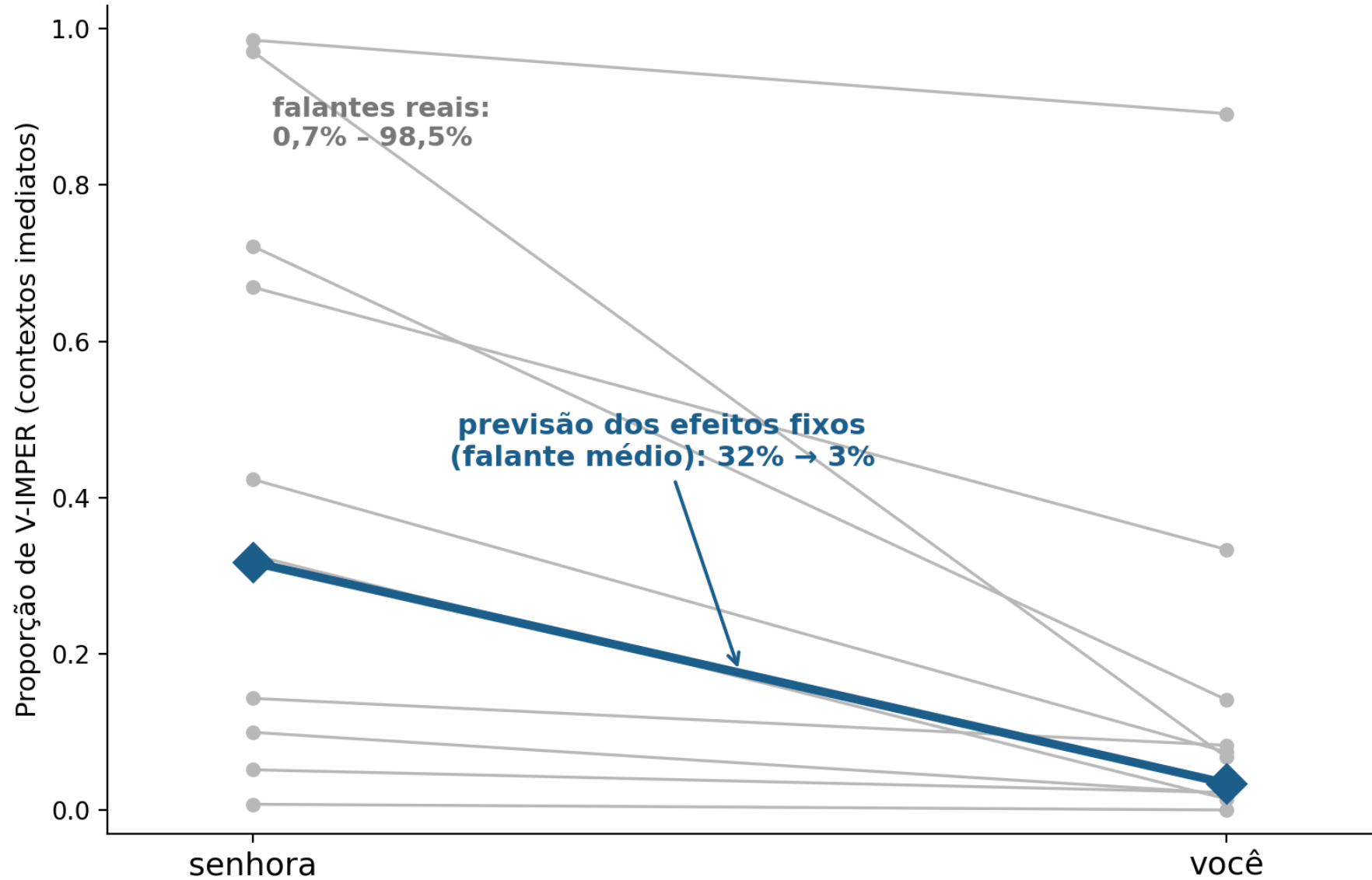
Senhora
(deferential)

+FAZ -FAÇA

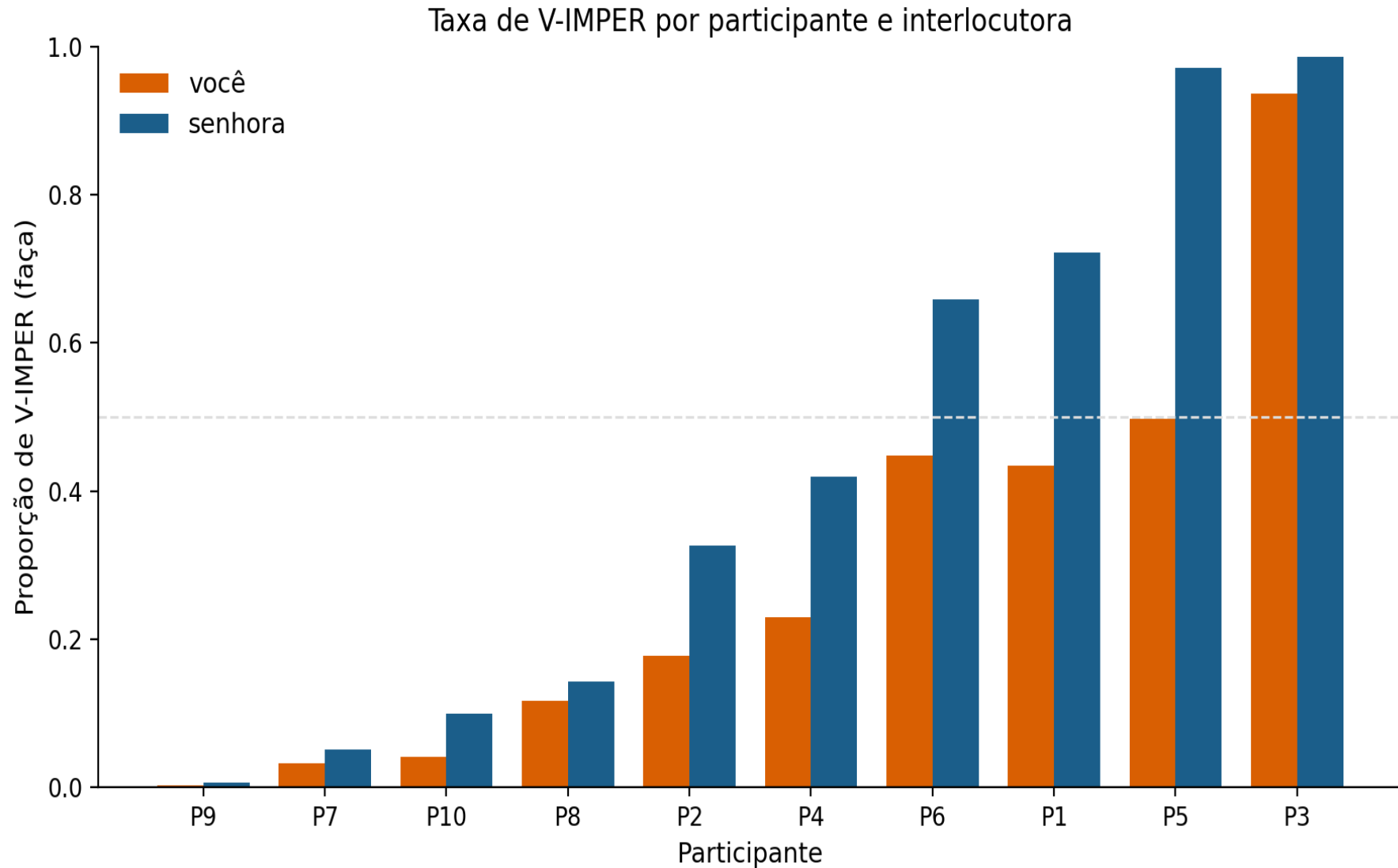


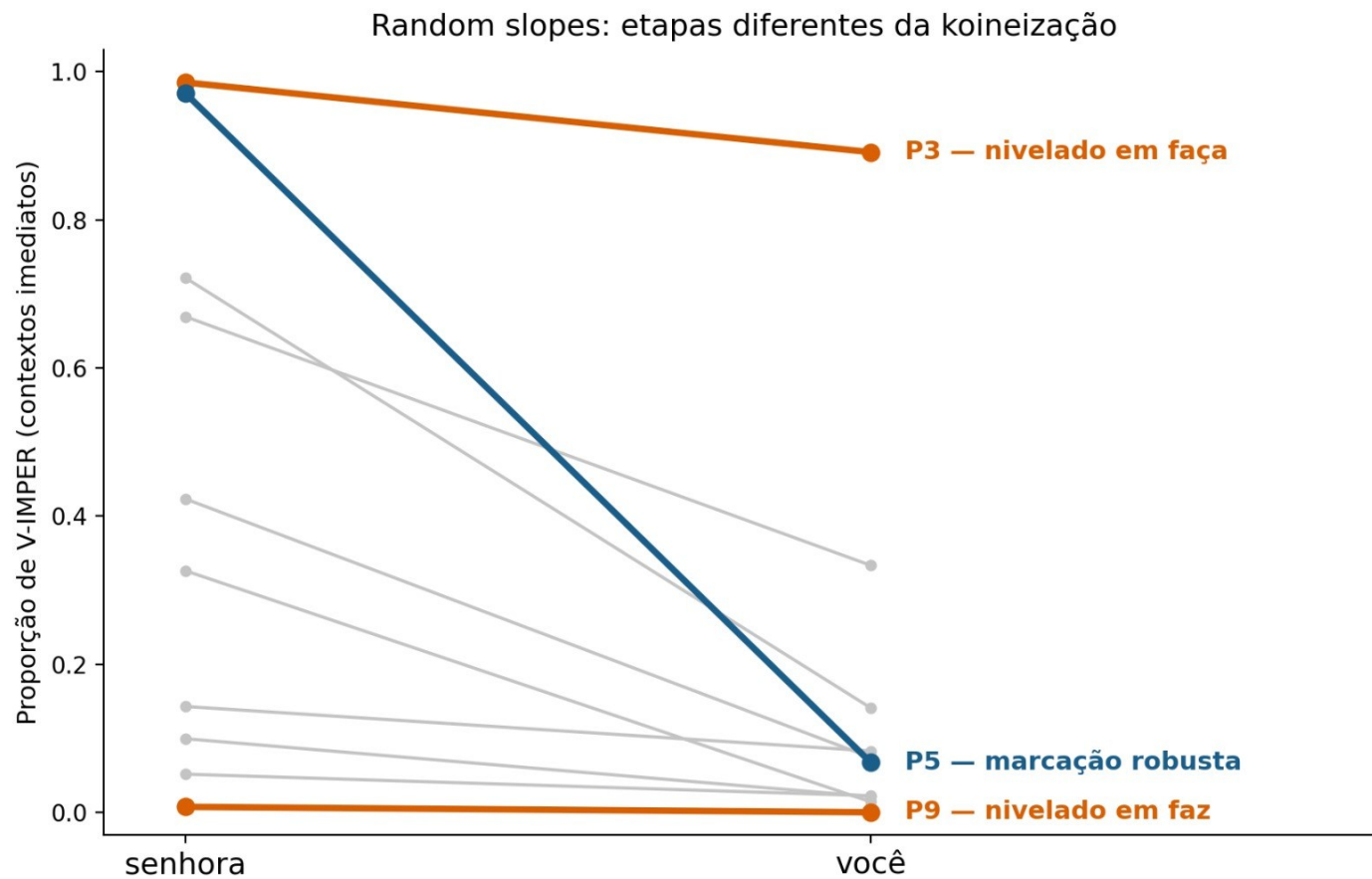
você
(non deferential)

Os efeitos fixos sozinhos: $R^2_m = 0,12$ — quase ninguém é o falante médio



Who is speaking matters more than who they're speaking to!





Speakers form profiles ranging from robust marking (P5) under conditioning to completed levelling (P9, P3).

Individual variation exceeds what the shared conditioning factors explain => and the conditioning factors themselves differ across speakers.

Conclusions

- 1. The interlocutor effect is real, but ...**
- 2. individuals vary idiosyncratically, and the conditioning factors influence them differently.**
- 3. Idiosyncratic variation should not be dismissed — for imperatives in Brazilian Portuguese, we should ask whether it is not mere noise to be discarded, but a property of the grammar in its current, not-(yet?)-focused state.**



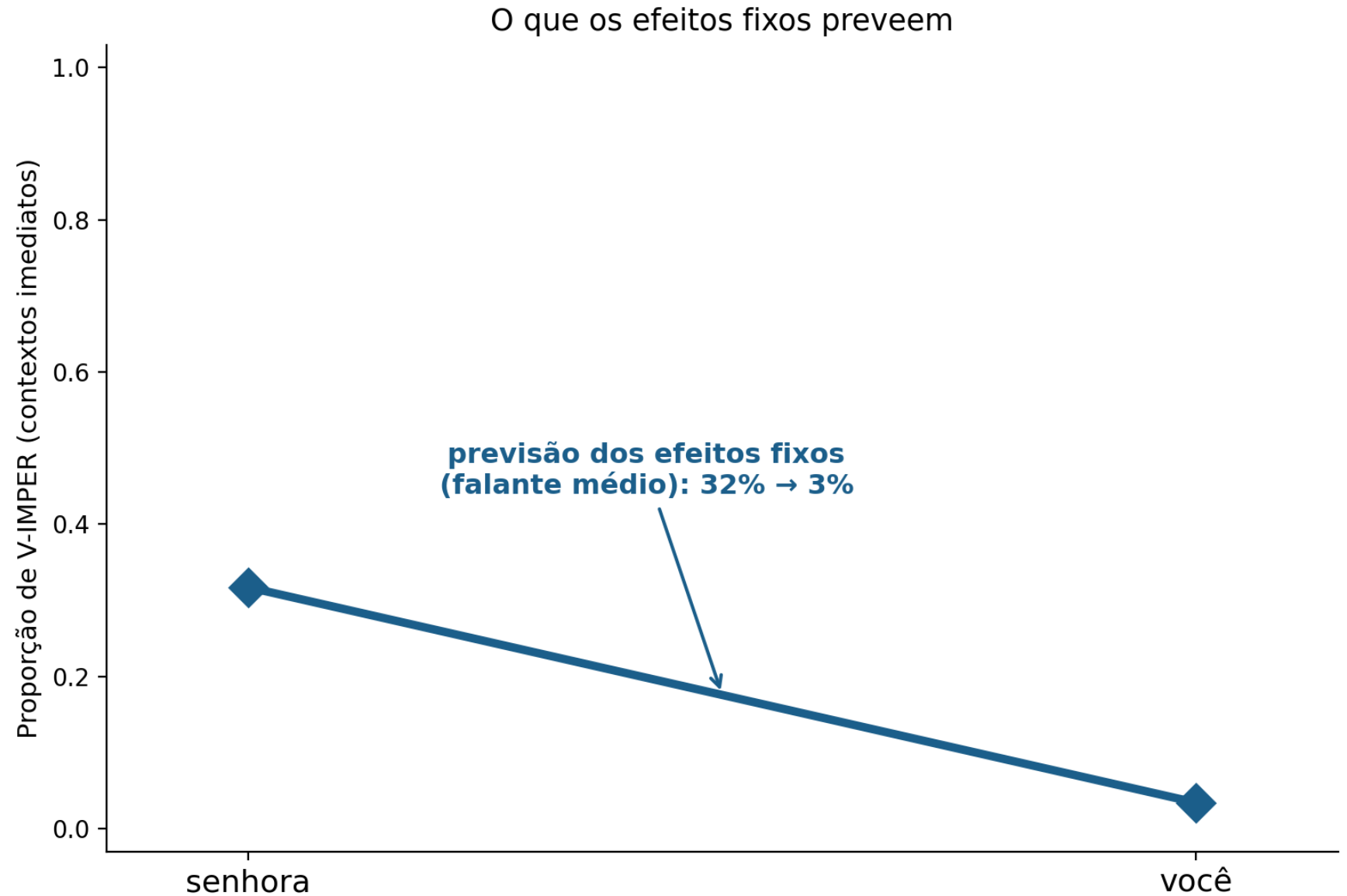
LUDWIG-
MAXIMILIANS-
UNIVERSITÄT
MÜNCHEN

Thank you!

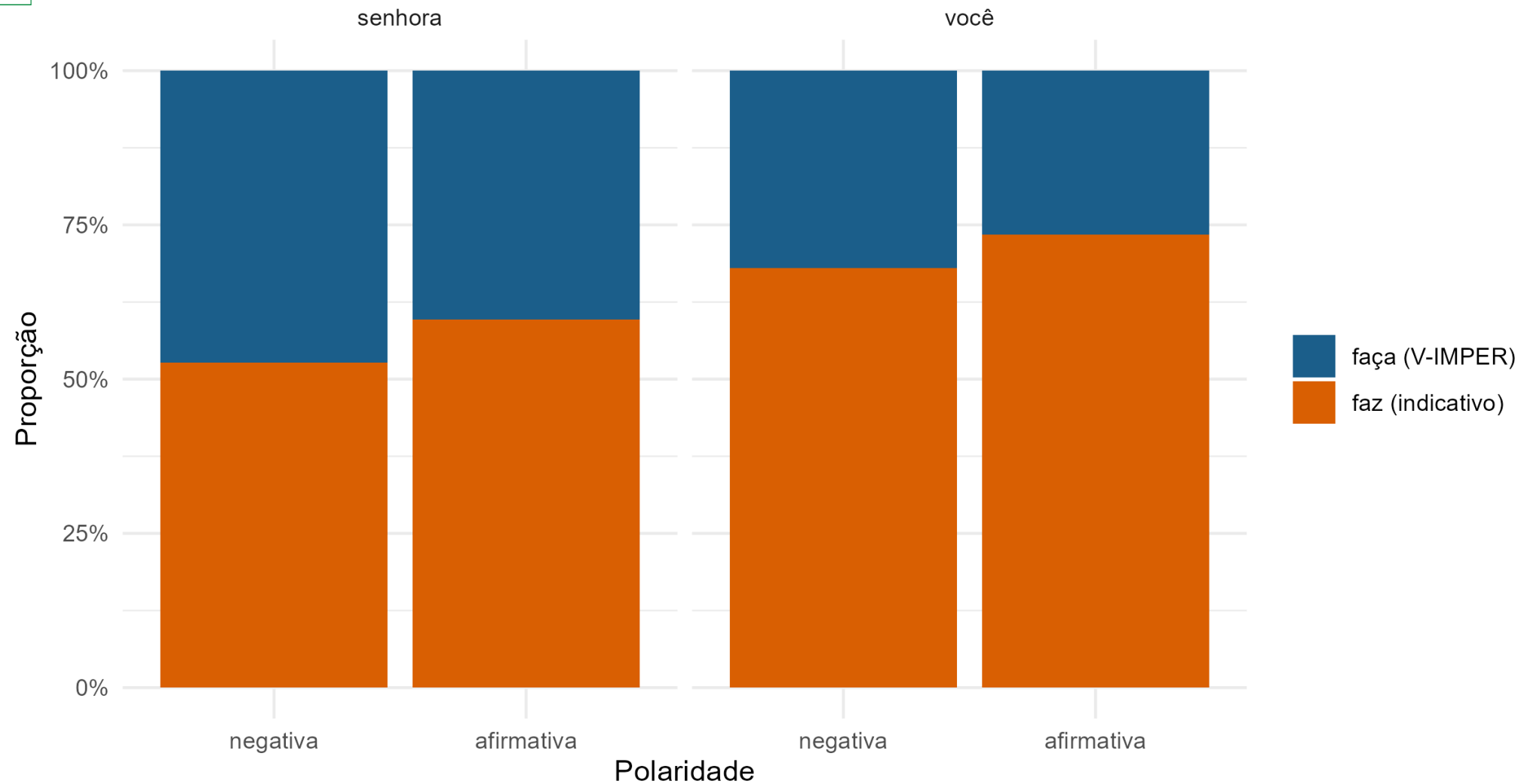
Muito obrigada

Vielen Dank!

The results



Formas por polaridade e interlocutora



What the random effects add

WHY A MIXED MODEL?

4,242 tokens — from only 10 speakers

The tokens are not independent.

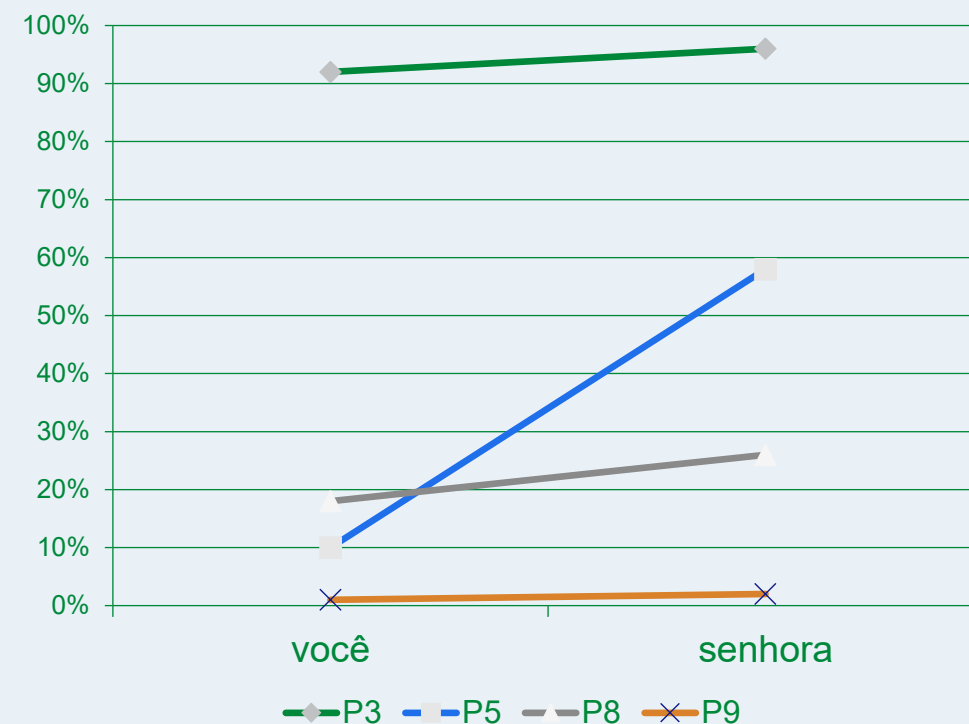
The mixed model gives each speaker:

- a random intercept = their own baseline
(some almost always faça, some almost never)
- a random slope = their own sensitivity
to the interlocutor (senhora vs. você)

**Spread between speakers (SD 2.91)
outweighs the interlocutor effect (2.59)**

P9 < 1% · P3 > 94%

ONE LINE = ONE SPEAKER



schematic · y-axis = % faça (V-IMPER)

Componente

% de la variación total

Hablante (+ lexema, como efectos aleatorios)

≈ **60 %**

Sin explicar (residual)

≈ 28 %

Interlocutor (*senhora / você*)

≈ 6 %

Inmediatez

≈ 5 %

Polaridade

≈ 0,2 % (prácticamente 0)

Total

100 %

Same data, two models — what random effects change

MODEL 1 · NO RANDOM EFFECTS

Classic logistic regression

Fixed effects only: interlocutor · polarity · immediacy

Treats all 4,242 tokens as independent

Interlocutor: $\beta = -1.38$ ($p < .001$)

senhora 40% vs. você 15% → odds × 4

All effects significant, in the expected direction.

The model looks perfect...

MODEL 2 · + RANDOM INTERCEPTS & SLOPE

Mixed-effects logistic regression (lme4)

+ random intercepts: speaker & lexeme (83)

+ random slope: interlocutor varies by speaker

The 4,242 tokens come from only 10 speakers

Interlocutor: $\beta = -2.59$ ($p < .001$)

senhora 32% vs. você 3% → odds × 13

By-speaker variance 8.45 > interlocutor effect 2.59

→ **who is speaking matters more than who they speak to**

Same data, two models — what random effects change

MODEL 1 · NO RANDOM EFFECTS

Classic logistic regression

Fixed effects only: interlocutor · polarity · immediacy

Treats all 4,242 tokens as independent

Interlocutor: $\beta = -1.38$ ($p < .001$)

senhora 40% vs. você 15% → odds × 4

All effects significant, in the expected direction.

The model looks perfect...

MODEL 2 · + RANDOM INTERCEPTS & SLOPE

Mixed-effects logistic regression (lme4)

+ random intercepts: speaker & lexeme (83)

+ random slope: interlocutor varies by speaker

The 4,242 tokens come from only 10 speakers

Interlocutor: $\beta = -2.59$ ($p < .001$)

senhora 32% vs. você 3% → odds × 13

By-speaker variance 8.45 > interlocutor effect 2.59

→ **who is speaking matters more than who they speak to**

THREE STAGES IN NEW DIALECT FORMATION

(TRUDGILL 1986, 2005, 2011; BRITAIN & TRUDGILL 2005)

Stage I – Adult migrants

High variability; minimal, inconsistent levelling.

Stage II – First native-born generation

Levelling increases — the most frequent, least marked and most transparent forms prevail.

Stage III – Later generations

The system stabilises; surviving variants are eliminated or reallocated to grammatical or social factors.

Underlying assumption:

Speakers accommodate until a focused, stable grammar emerges — **any variation not yet levelled should correlate with internal or external factors.**

Idiosyncratic variation

Non-systematic variation between and within individuals.

- **Importance:**

Challenges two key ideas of language:

- **Orderly heterogeneity** – the idea that variation is structured, coherent, and has indexical meaning.
- **Homogeneity assumption** – that sociologically similar communities are linguistically similar.

FICA EM CASA & COMPRE ONLINE

FAÇA SUPERMERCADO DE FORMA PRÁTICA E SEM SAIR DE CASA

separapramim.com.br

o seu shopping virtual

Desenvolvido por SGA SISTEMAS

The advertisement features a family (mother, father, and child) in a kitchen setting. The child is pointing at a laptop screen displaying a virtual grocery store interface. The mother is holding a red shopping basket, and the father is standing behind her. The text 'FICA EM CASA & COMPRE ONLINE' is overlaid on the top left. Below the image, a banner reads 'FAÇA SUPERMERCADO DE FORMA PRÁTICA E SEM SAIR DE CASA'. The website 'separa.com.br' is prominently displayed in the center, with 'separa' in a stylized font. To the right, it says 'o seu shopping virtual'. In the bottom right corner, it mentions 'Desenvolvido por SGA SISTEMAS' with a logo.