

# Revisiting Maxakalí nasal harmony: Evidence from loanwords and modeling

Jamie White (UCL)

Mário Coelho da Silva (UFMG)

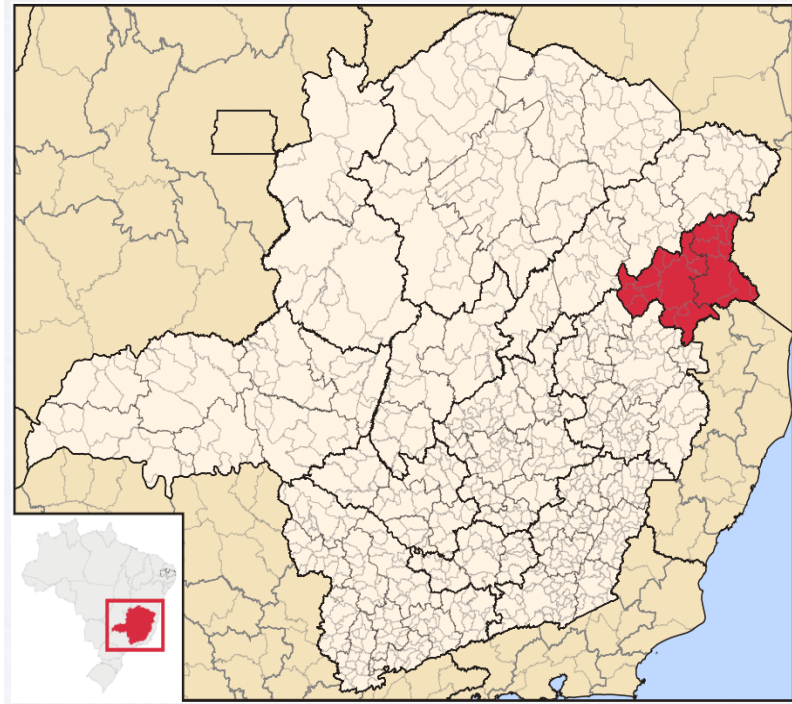
Andrew Nevins (UCL)

# Overview

- Maxakalí nasal harmony: **absolute** in native vocabulary
  - Difficult to determine underlying representations and grammatical mechanisms involved.
- Harmony is **variable** in loanwords (from Portuguese)
  - → Collect a corpus of loanwords and use a probabilistic model to discover which constraints best account for the data.
- **Main claims:**
  1. Nasality has preferred status in Maxakalí, especially in vowels and in prominent prosodic positions.
  2. Nasal harmony primarily enforced as local agreement in Maxakalí, with stronger agreement within syllables.

# Maxakalí

- Macro-Jê language of Brazil.
- ~2,000 speakers.



# Maxakalí

- Consonants:

	Bilabial	Dental	Palatal	Velar	Laryngeal
Voiceless	p	t	c	k	ʔ, h
Voiced	b / m	d / n	ʃ / ɲ	g / ŋ	

- Possible syllables: (C)V(C)
- CV'CVC is the most common word type.
- Stress is always final.
- Final codas are often lenited to a glide (which can be nasalised).

# Maxakalí nasal harmony

- Sequences of voiced segments must be entirely **nasal** or entirely **oral**.

## Acceptable

[ <b>n</b> ã <b>m</b> ĩŋ]	‘spirit’
[ <b>m</b> ãŋ <b>õ</b> n]	‘sun’
[ <b>k</b> a <b>d</b> o <b>p</b> ]	‘to mix in liquid’
[ <b>c</b> o <b>d</b> a <b>t</b> ]	‘soldier’

## Unacceptable

*[ <b>n</b> ã <b>b</b> ĩŋ]
*[ <b>b</b> aŋ <b>õ</b> n]
*[ <b>k</b> a <b>n</b> o <b>p</b> ]
*[ <b>c</b> õ <b>d</b> a <b>t</b> ]

- Voiceless** segments do not participate and they block.

## Acceptable

[ <b>k</b> o <b>c</b> ã <b>m</b> ]	‘fish hook’
[ <b>m</b> ũ <b>t</b> i <b>k</b> ]	‘with’

# Ambiguity in URs

What carries the nasal contrast in Maxakalí?

- Phonemic in both stops and vowels? (Gudschinsky, Popovich & Popovich 1970)
- Only in stops; nasal vowels derived? (Rodrigues 1981)
- Only in vowels, especially the stressed vowel? (cf. Walker 2003)

→ Writing system encodes the contrast on the vowel.

<mã> = [ mã ]

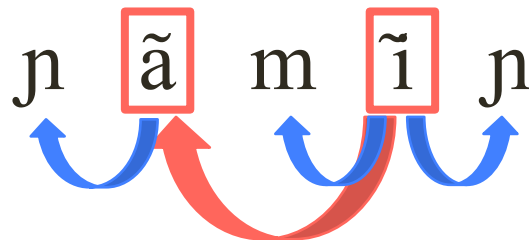
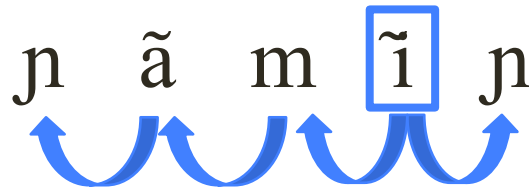
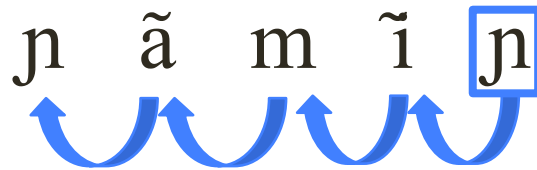
<ma> = [ ba ]

<manayit> = [ bada<sup>1</sup>ʃit ] < Brasilia (Port.)

# Ambiguity in harmony mechanism

How do speakers represent the nasal harmony pattern?

- Several accounts consistent with data. (e.g. Gudschinsky, Popovich, & Popovich 1970; Rodrigues 1981; Wetzels 2009; Silva & Nevins 2015)



# Loanwords

- Nasal harmony is not absolute in loanwords (from Portuguese).
  - Wetzels (2009): absolute in rimes, but not in onsets.

[dibãw̃] < Prt. [limãw̃] ‘lime’

- Variation also possible.

[pẽŋõŋ] ~ [peɟõŋ] < Prt. [feizãw̃] ‘bean’

## → Our approach

- Collect a corpus of loanword data.
- Use a MaxEnt model to investigate which grammar best accounts for the observed data.

# Data collection

- 18 native speakers of Maxakalí from Apne Yĩxux (the ‘Green Village’) in Minas Gerais, Brazil.
  - Balanced by gender, and across 3 age groups (15–29, 30–44, 45+).
- Speakers presented (audio only) with Portuguese word known to have an attested loanword in Maxakalí and asked to produce the Maxakalí version.
  - Productions recorded and transcribed phonetically.
- 774 word tokens recorded in total (96 types).
  - Average of 43 per speaker.

# Modeling

# MaxEnt models

- General statistical model used in several fields.
- Implemented here as a probabilistic model of the phonological grammar. (e.g. Goldwater & Johnson 2003, Wilson 2006, Hayes & Wilson 2008, White 2017)
- Specifically a type of Harmonic Grammar. (Legendre, Miyata, & Smolensky 1990)
  - Constraint-based.
  - Constraint have weights that are additive → model has the ‘ganging’ property.
- Provide model:
  - a set of input/output training data
  - a set of output candidates for each input
  - a set of constraints
  - violations

# MaxEnt learning

- The model's learning objective is to find the set of constraint weights that maximizes the likelihood of the observed data. (It is guaranteed to succeed!)
- Also, includes a (very weak) smoothing prior.
  - All else being equal, prefers weights to be near 0 and spread out among relevant constraints.
- After learning a grammar, the model outputs predicted probabilities for each candidate.
  - which we use to calculate the likelihood of the data, given the grammar.



# Output candidates

- Models considers and assigns probability to all relevant candidates:

[T $\tilde{V}$ N $\tilde{V}$ 'N $\tilde{V}$ N]

[T $\tilde{V}$ N $\tilde{V}$ 'N $\tilde{V}$ N]

[TVD $\tilde{V}$ 'N $\tilde{V}$ N]

[TVDV' $\tilde{N}$  $\tilde{V}$ N]

[TVDV'D $\tilde{V}$ N]

[TVNV'D $\tilde{V}$ N]

...

# Constraints in initial model

- 48 constraints included in initial model.
- **Markedness constraints**
  - Enforce restrictions on orality/nasality in the Maxakalí forms.
- **Faithfulness constraints**
  - Enforce maintaining the orality/nasality of the Portuguese source word.
  - General and more specific versions.

# Markedness constraints

## Local agreement

AGREE(nasal)

AGREE(nasal), voiced only

AGREE(nasal) within syllable

AGREE(nasal) within syllable, voiced only

AGREE(nasal) within rhyme

\*T $\tilde{V}$

\*#T $\tilde{V}$

\*T $\tilde{V}$ T

# Markedness constraints

## **Intersyllabic harmony**

ALIGN(nasal)-right

ALIGN(nasal)-left

# Markedness constraints

## Initial nasality preference

\*V<sub>ORAL</sub> in initial syllable

\*#V<sub>ORAL</sub>

# Faithfulness constraints

## General faithfulness

IDENT(nasal)

MAX(nasal)

IDENT-C(nasal)

MAX-C(nasal)

IDENT-V(nasal)

MAX-V(nasal)

# Faithfulness constraints

## **Positional faithfulness** (Beckman 1997)

IDENT(nasal) & MAX(nasal) in the following prominent positions:

- Onset
- Word-initial
- Maxakalí stressed syllable (i.e. final)
  - Onset of Mxk. stressed syllable
  - Vowel of Mxk. stressed syllable
- Portuguese stressed syllable
  - Onset of Prt. stressed syllable
  - Vowel of Prt. stressed syllable

# Faithfulness constraints

## **Faithfulness specific to place and manner**

IDENT(nasal) & MAX(nasal) specific to each place:

- Labial, Dental, Palatal, Velar

IDENT(nasal) specific to each manner:

- Plosives, fricatives, liquids, glides, rhotic

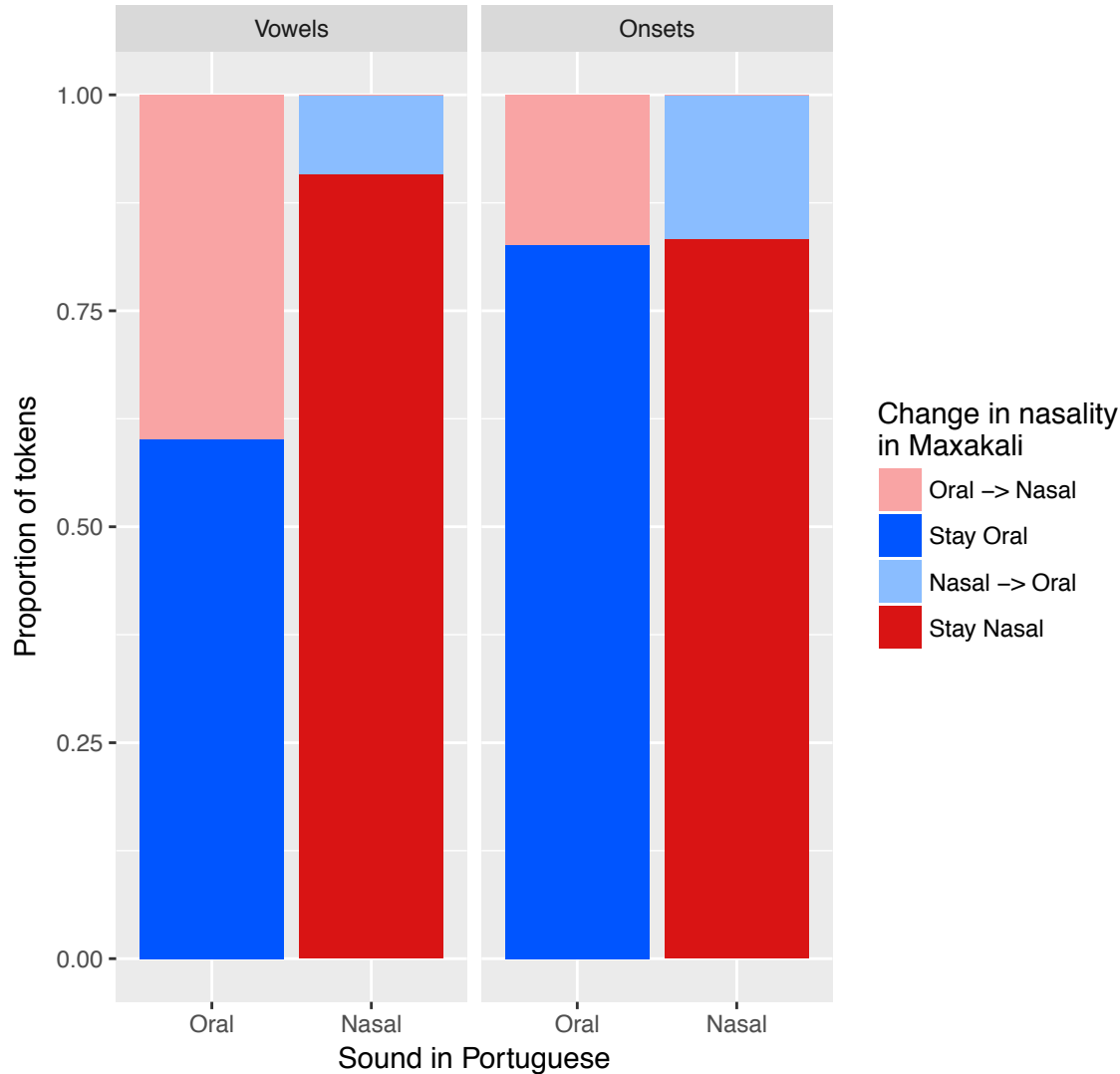
# Model comparison

- Following Hayes, Wilson & Shisko (2012), we used likelihood ratio tests to remove constraints not statistically warranted in the model.
  - (=same approach as when you do model comparison for, e.g., logistic regression)
- Backwards stepwise comparison:
  - Remove constraints one at a time, starting with the one resulting in the least improvement.
  - Stop when all remaining constraints significantly improve model fit.
- 48 constraints > 18 constraints in final model

# The final model and data patterns

# Overall changes for Vs and Cs

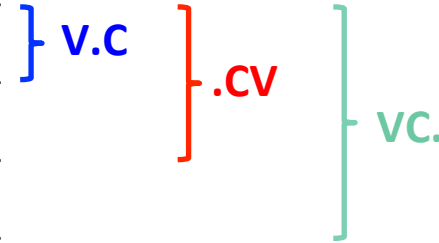
Overall changes in nasality from Portuguese to Maxakali (voiced sounds only)



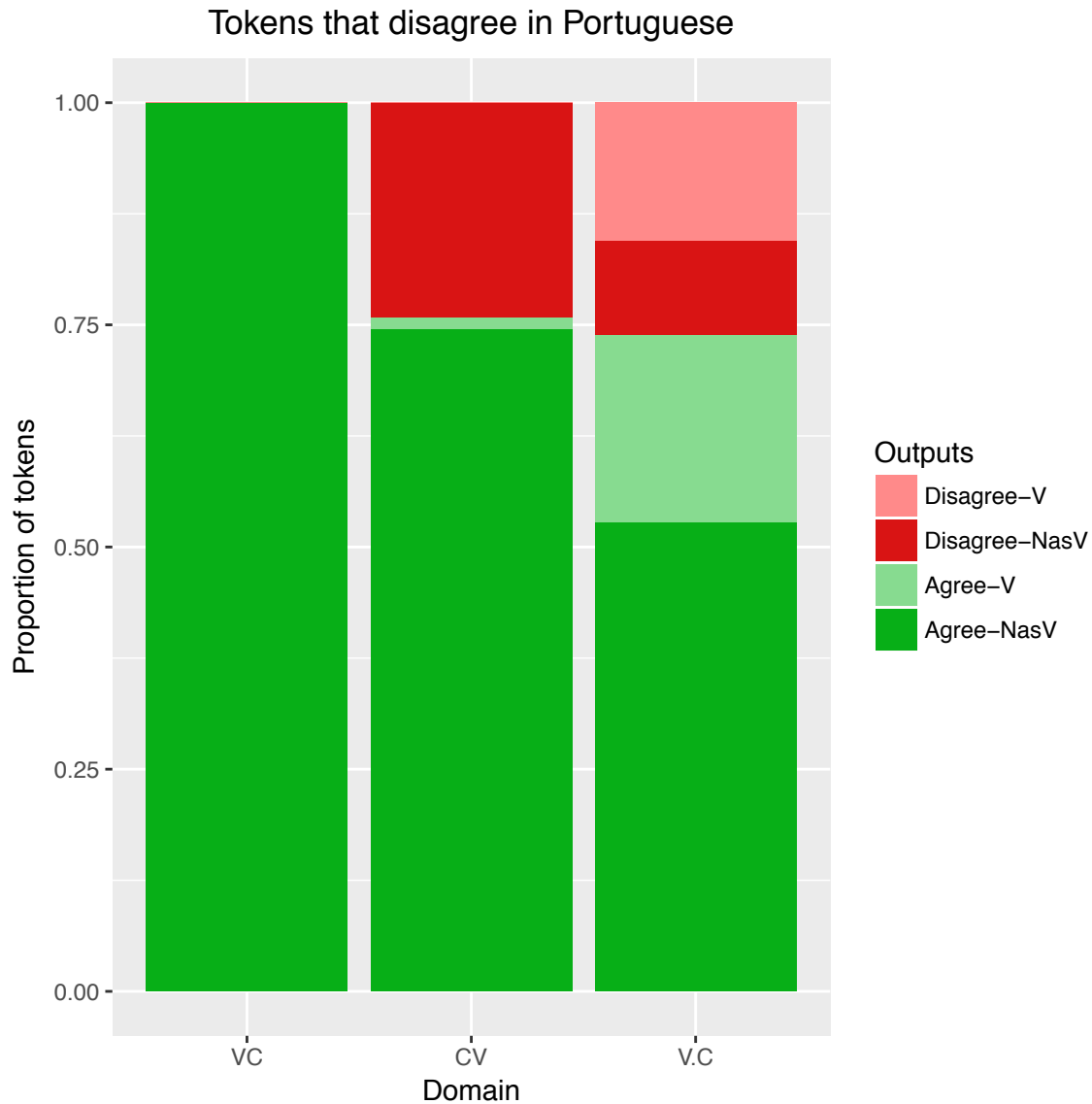
# Harmony constraints

## Local agreement

Constraint	Weight	p-value
AGREE(nasal) – voiced only	1.50	<.001
AGREE(nasal)-SYLL – voiced only	1.38	<.001
AGREE(nasal)-RIME	3.12	<.001



# Local agreement by domain



# Markedness constraints

## Local agreement

Constraint	Weight	p-value
AGREE(nasal) – voiced only	1.50	<.001
AGREE(nasal)-SYLL – voiced only	1.38	<.001
AGREE(nasal)-RIME	3.12	<.001
*T̃V	2.10	<.001



**Spontaneous denasalisation of vowels  
(otherwise very rare!)**

Prt. *cinquenta* [sĩ'kwẽta] > [ciku'ẽɔ]

Prt. *compadre* [kũ'padɾɪ] > [ku'paɔ]

Prt. *computador* [kõputa'doh] > [kuputa'do]

N.B. These cases tend to fall between two voiceless Cs and in word-initial position, but the model prefers general \*T̃V to \*#T̃V and \*T̃T.

# Markedness constraints

## Local agreement

Constraint	Weight	p-value
AGREE(nasal) – voiced only	1.50	<.001
AGREE(nasal)-SYLL – voiced only	1.38	<.001
AGREE(nasal)-RIME	3.12	<.001
*T <sup>~</sup> V	2.10	<.001

## Initial nasality preference

Constraint	Weight	p-value
*V <sub>ORAL</sub> – in syll 1	1.80	<.001
*#V <sub>ORAL</sub>	6.19	<.001

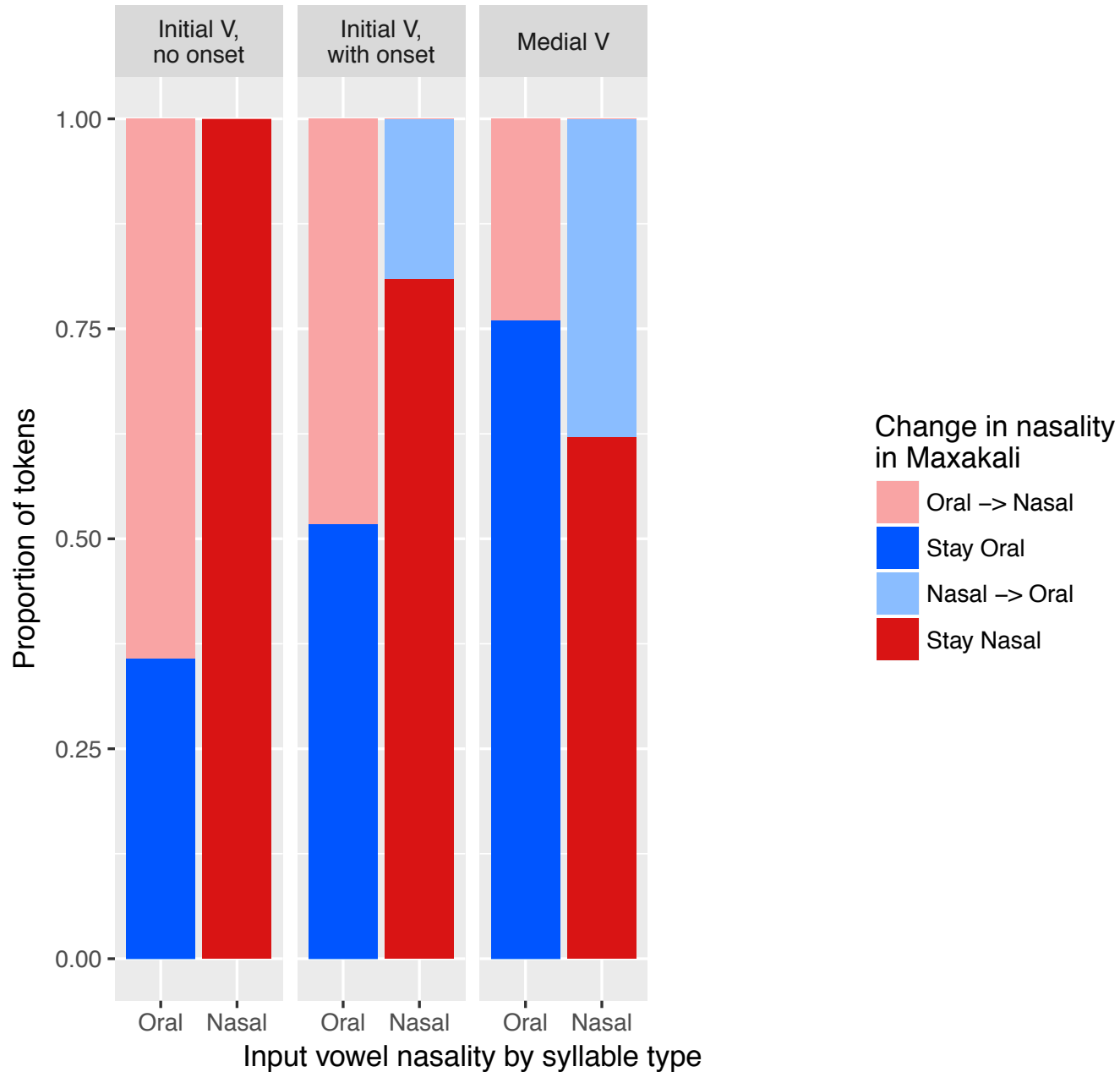
} #CV } #V

Frequent spontaneous nasalisation of initial vowels

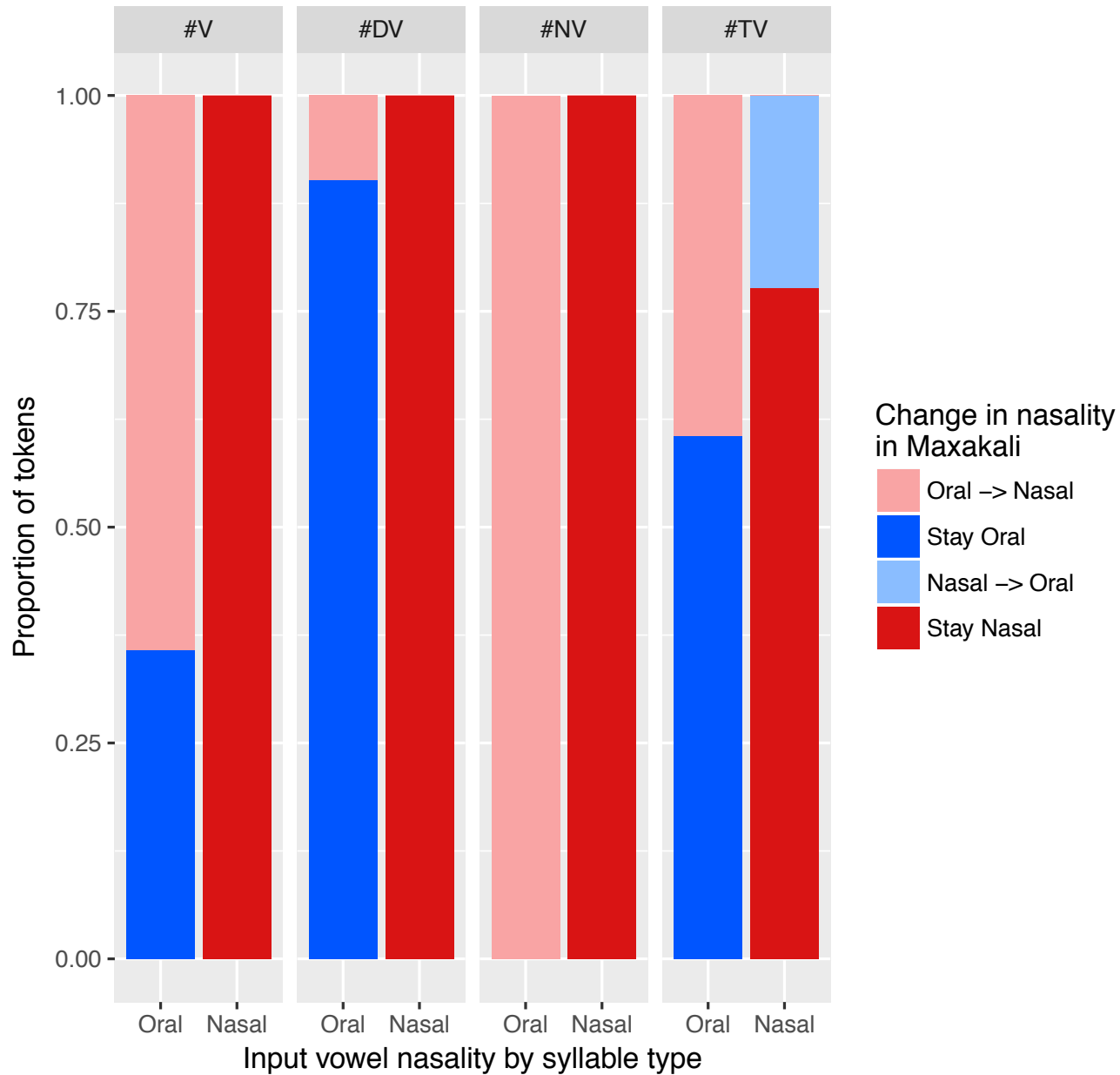
Prt. [a<sup>1</sup>gostu] > Mxk. [ã<sup>1</sup>guj]

Prt. [is<sup>1</sup>peɫu] > Mxk. [ĩj<sup>1</sup>pæj]

# Overall changes in nasality by vowel position



# Overall changes in vowel nasality for initial syllables



# Markedness constraints

## Local agreement

Constraint	Weight	p-value
AGREE(nasal) – voiced only	1.50	<.001
AGREE(nasal)-SYLL – voiced only	1.38	<.001
AGREE(nasal)-RIME	3.12	<.001
*T̃	2.10	<.001

## Initial nasality preference

Constraint	Weight	p-value
*V <sub>ORAL</sub> – in syll 1	1.80	<.001
*#V <sub>ORAL</sub>	6.19	<.001

## Intersyllabic harmony

Constraint	Weight	p-value
ALIGN(nasal)-RIGHT	0.65	<.001

Helpful for, e.g.:

Prt. *oliveira* [oli<sup>l</sup>verə] >

Mxk. [õñi<sup>l</sup>bɛ<sub>̃</sub>]

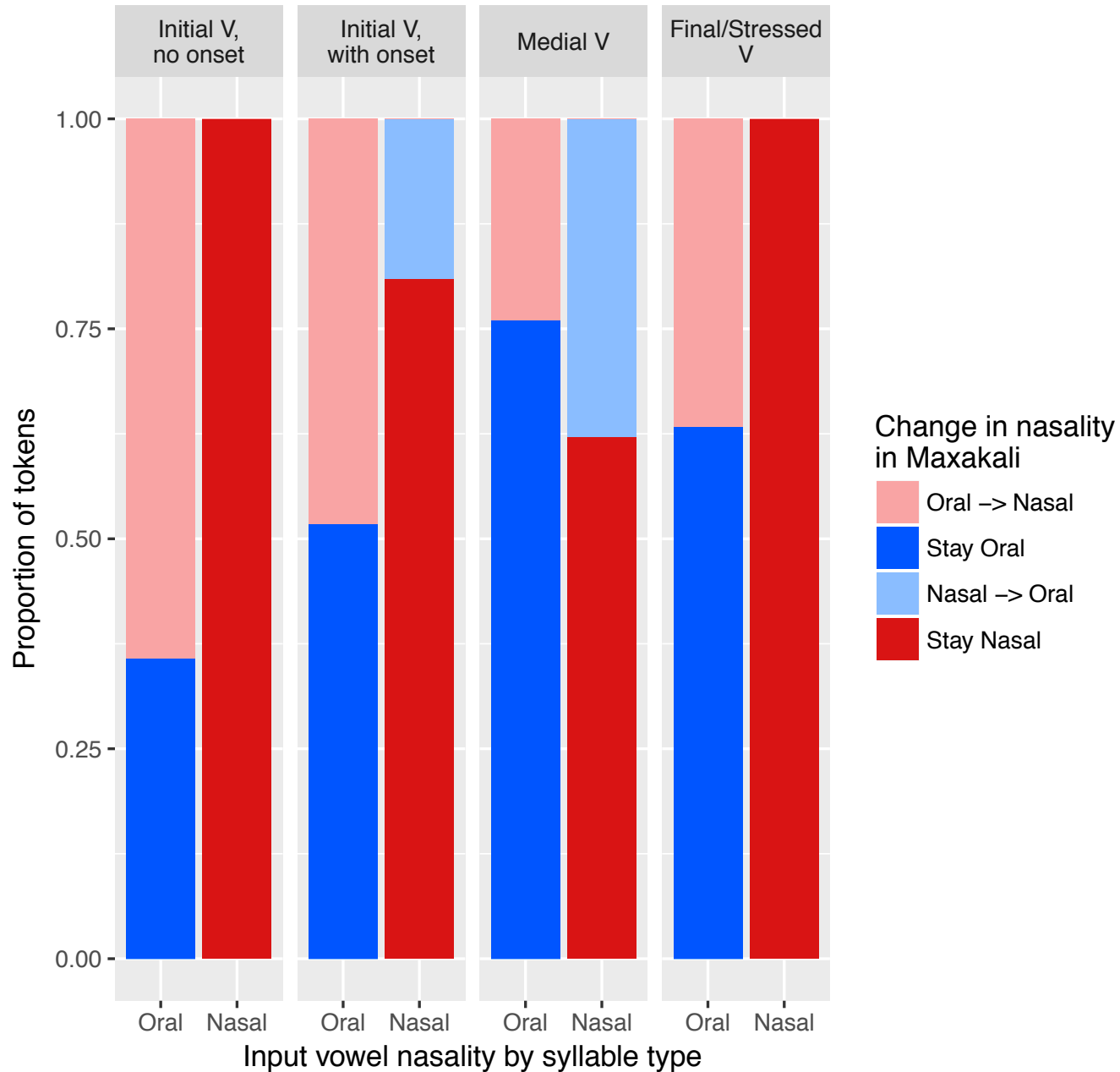
→ 0 weight for ALIGN-L !!

# Faithfulness constraints

## Maintaining nasality in vowels

Constraint	Weight	p-value	
MAX(nasal) – in vowels	2.04	<.001	
MAX(nasal) – in Mxk. stressed vowels	12.20	<.001	(final syllable)
MAX(nasal) – in Prt. stressed vowels	1.98	<.001	

# Overall changes in nasality by vowel position



# Faithfulness constraints

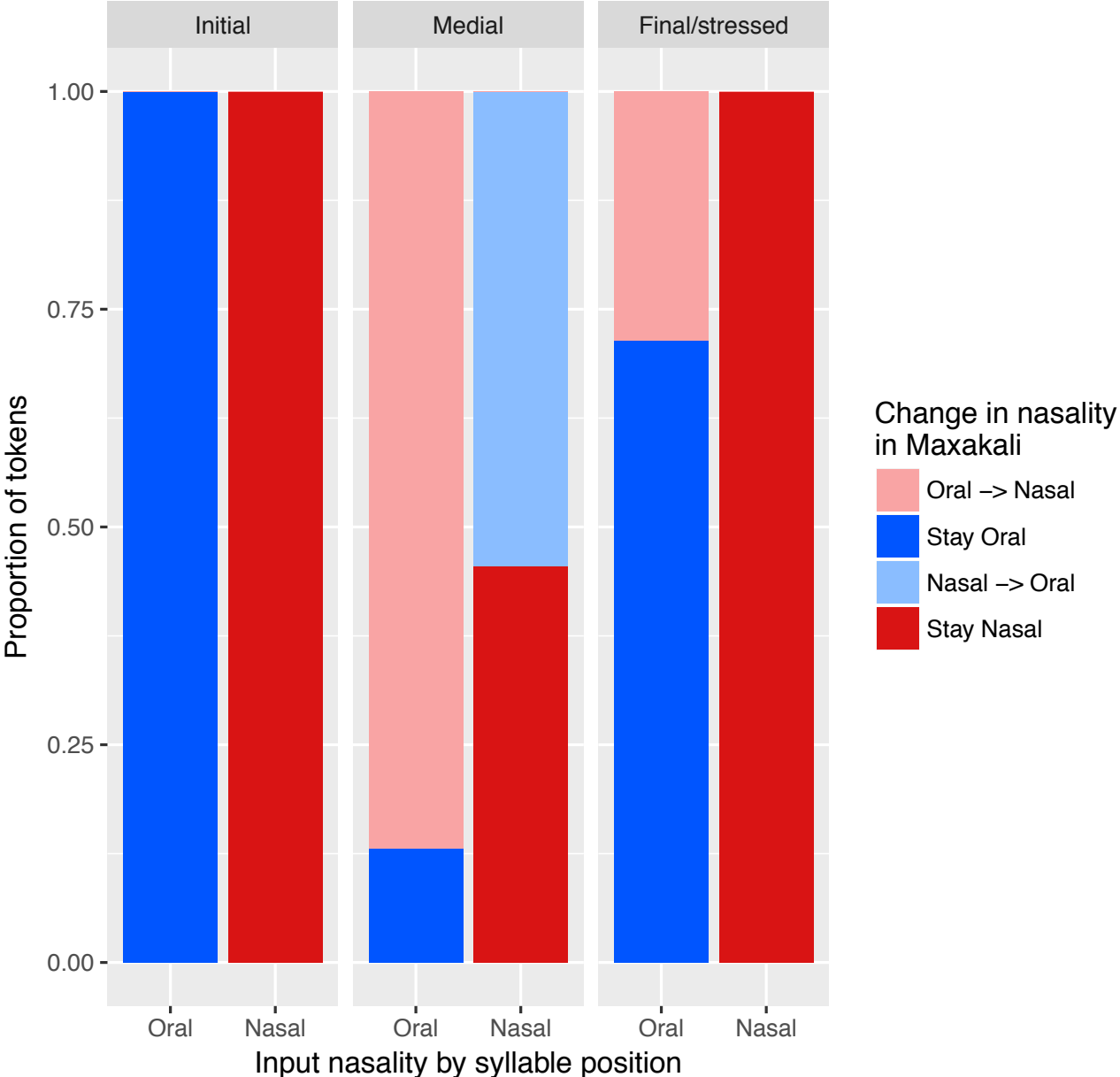
## Maintaining nasality in vowels

Constraint	Weight	p-value	
MAX(nasal) – in vowels	2.04	<.001	
MAX(nasal) – in Mxk. stressed vowels	12.20	<.001	(final syllable)
MAX(nasal) – in Prt. stressed vowels	1.98	<.001	

## Maintaining nasality/orality in Cs in prominent positions

Constraint	Weight	p-value	
IDENT(nasal) – in word-initial position	5.48	<.001	→ Prt. [bẽ <sup>h</sup> kʊ] >
IDENT(nasal) – in Mxk. stressed onsets	0.63	<.01	Mxk. [bãŋ]

# Overall changes in nasality for onsets by syllable position



# Faithfulness constraints

## Place-specific

Constraint	Weight	p-value
IDENT(nasal) – in dentals	1.20	<.001
IDENT(nasal) – in velars	4.46	<.001

→ \*ŋ in onsets

## Manner-specific

Constraint	Weight	p-value
IDENT(nasal) – in plosives	0.81	<.001
IDENT(nasal) – in fricatives	1.57	<.001
IDENT(nasal) – in glides	2.78	<.001
IDENT(nasal) – in rhotics	0.94	<.001

# Summary and conclusions

- Maxakalí speakers have a strong preference to maintain nasality.
  - Especially on vowels.
  - And especially on prominent prosodic positions (word-initially and in stressed syllables).
  - → Suggests that vowels play a larger role than consonants in the nasality contrast in Maxakalí. Consonants more likely to follow along.
- Maxakalí speakers like for word to start with nasality.
  - Will even spontaneously create it on vowels.

# Summary and conclusions

- Nasal harmony in Maxakalí appears to be driven primarily by **local agreement** constraints, causing nasality to propagate through a word, not intersyllabic harmony processes.
  - Greater agreement within syllable domains than across ('nasal concord' effect.)
  - No effect of R→L harmony observed – noteworthy since some previous analyses have posited it.
  - Small effect of L→R harmony.
- Loanword data and modeling allow us to explore these complex factors in ways not possible by relying on the native vocabulary alone.

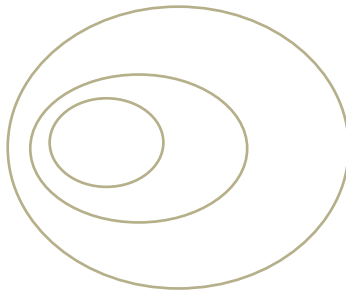
# Thank you

## Acknowledgments:

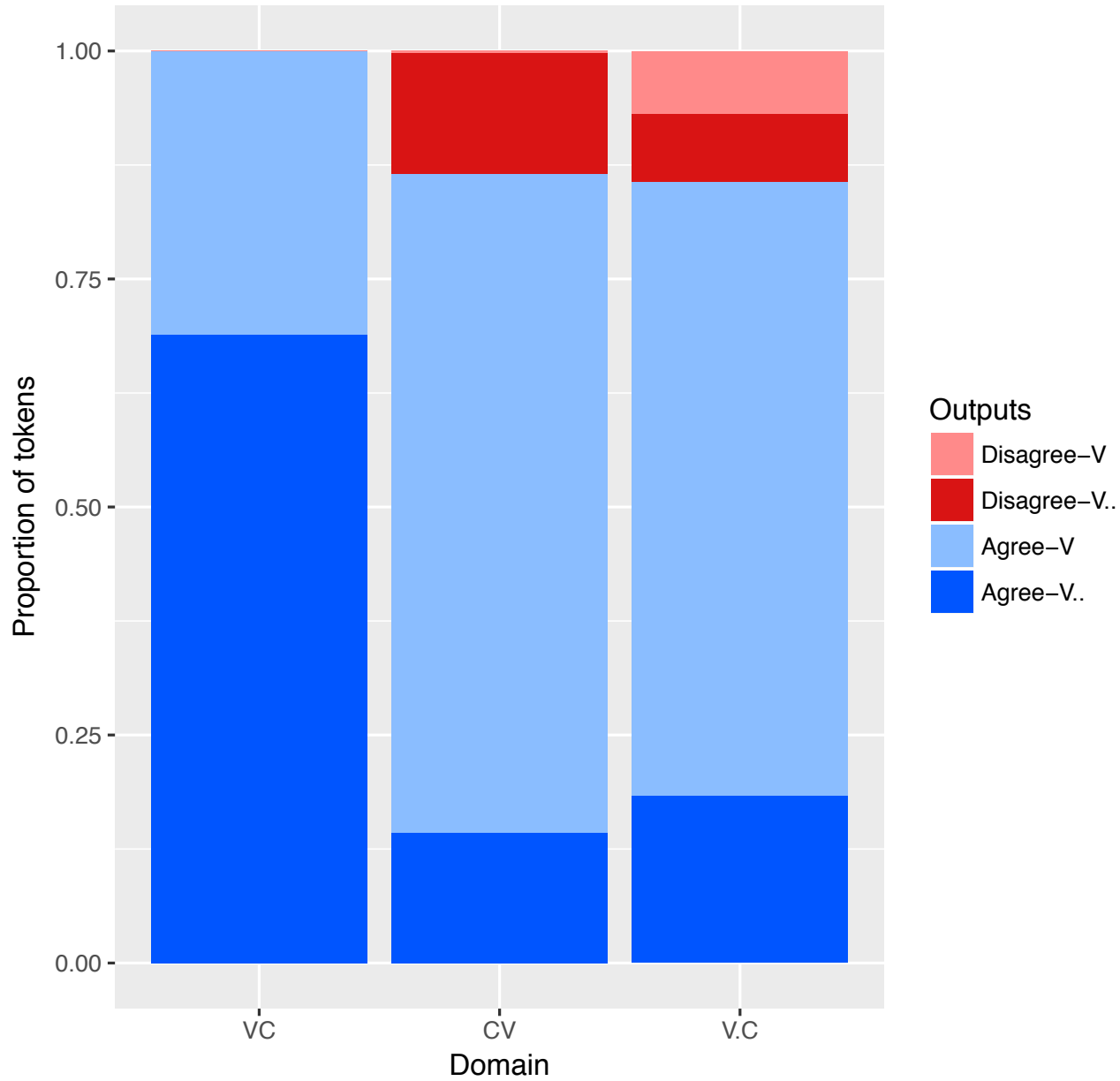
- UCL phonology group and participants at GLOW41.
- Special thank you to the speakers who provided data.

# Future work?

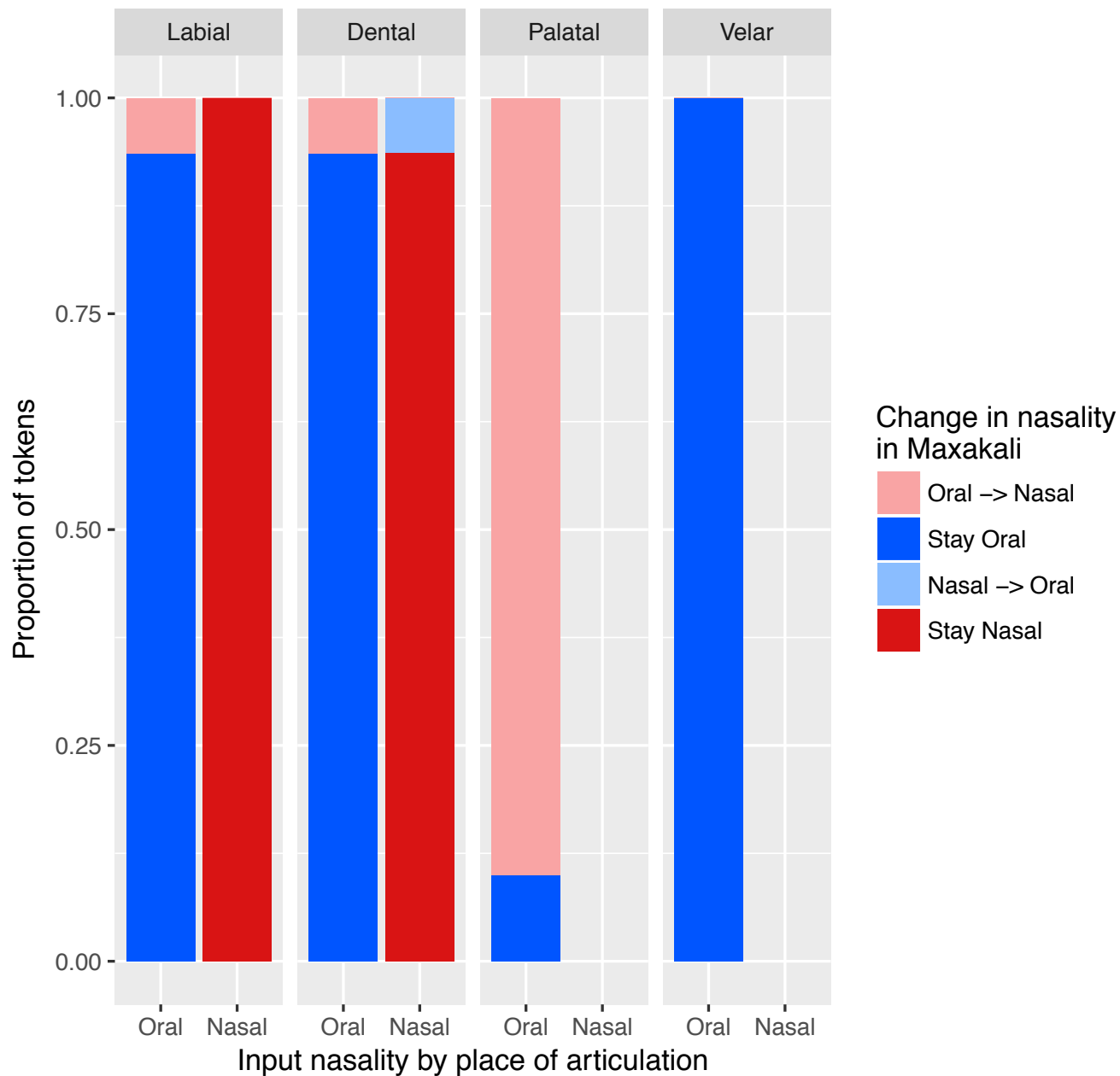
- Looking more at individual and generational patterns?
- Implicational nativisation effects: if x is adapted, then y is adapted? (Smith 2018)



# Tokens that agree in Portuguese



# Overall changes in nasality by place (onsets)



# Overall changes in nasality by manner (onsets)

