Evidence for Prosodic Phrase Marking in French Double Negation

An experimental approach

Jeremy Yeaton
Viviane Déprez, PhD

École normale supérieure de Paris
Rutgers University

ConSOLE XXVI: UCL, London
February 16, 2018

Nobody Say Nothing!

1. Negative **Concord**: A single negation reading of a sequence of Negative Concord Items (NCI)
   Nobody say anything
2. Negative **Discord**: Double negation
   Nobody say nothing $\rightarrow$ *Everybody say something*

But French can do both:
Personne ne dit rien.

Outline

Macro-Parametric Theory

Generative Theoretical Stance

NC languages are distinguished from DN ones by a macro-parameter *(Zanutini 1991, Haegeman 1995, Zeijlstra 2004)*

- French, Spanish, Catalan, etc.: **NC Languages**
- English, Dutch, German, etc.: **DN Languages**

Predictions:

- No real NC/DN ambiguity in languages
- DN Emerging in an NC language, or NC in a DN one would be a marked "anomaly", not part of the grammar

The Nature of NCI

NCI are **non-negative** expressions

NC reading from sentential negation (overt or covert):

- Indefinites *(Zeijlstra 2004, Chierchia 2013)*
  $\neg\exists x\exists y$[x said y]: Negated Indefinites
- Universals *(Giannakidou 2000, Shimoyama 2012)*
  $\forall x\forall y\neg$[x said y]: Negated Universals

Prediction: No DN readings

NCI are **negative** expressions

NC reading from Resumptive Quantification *(May 1989, de Swart & Sag 2002, Déprez 1997, 2000, 2011)*

- NO$< x, y >$ [x said y]: Resumptive Quantification (NC)
- NO$< x >$NO$< y >$ [x said y]: Scopal interaction (DN)

Prediction: Both NC & DN possible, but unclear why languages differ

The Nature of NCI

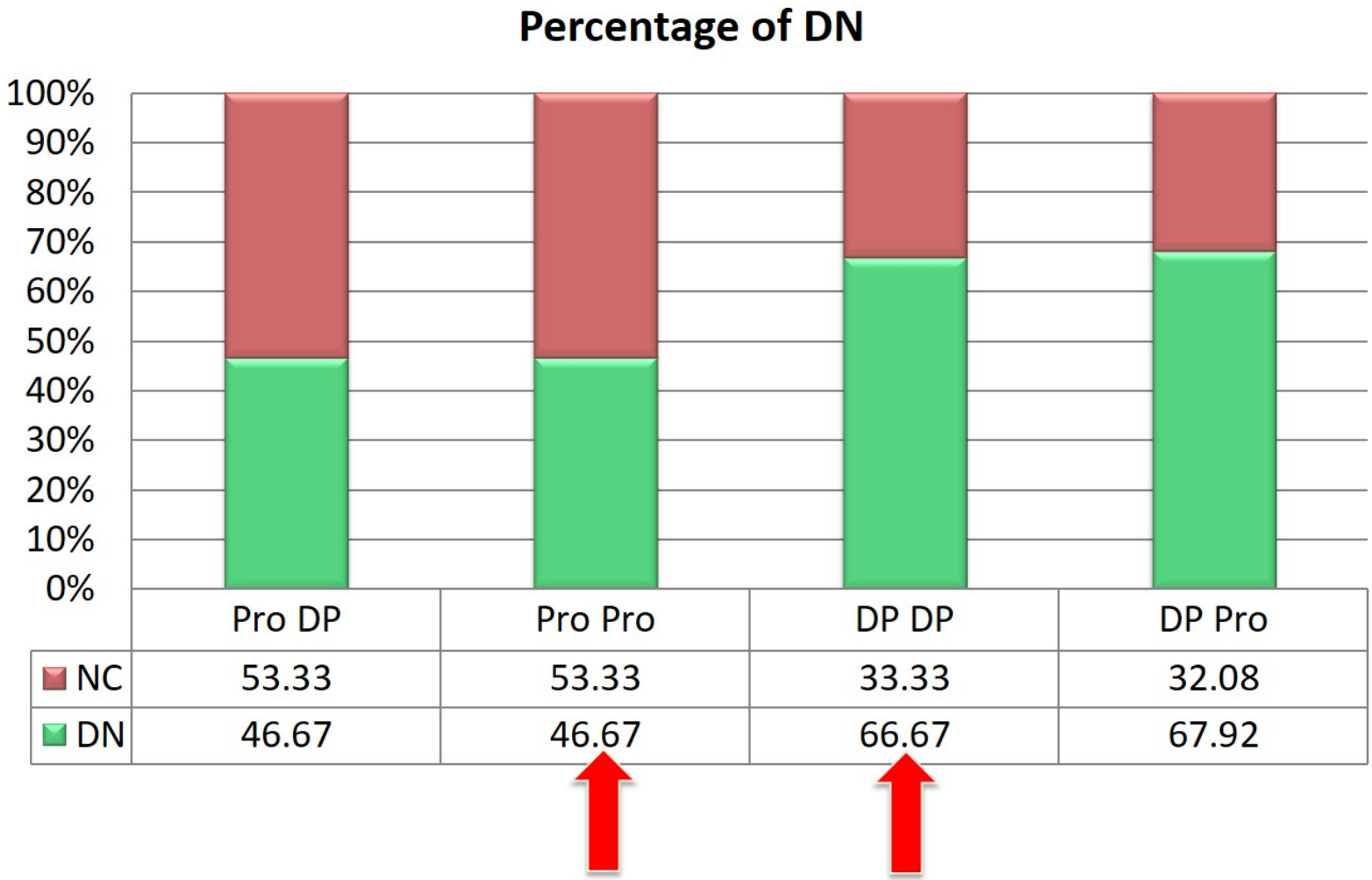
NCIs are **ambiguous** expressions (Micro-parametric approach)

- Lexical ambiguity *(Longobardi 1986)*
- Structural ambiguity *(Déprez 1997...2000)*
  - $[_{DP}...[_{NP}\,NCI]]$: non-negative
  - $[_{DP}\,NCI\,[_{NP}...]]$: negative

  Déprez 2011: NEG feature is interpretable at Phase Edge
  If DP = phase, NCI = negative
  1. At DP edge and
  2. At vP or TP/CP edge

  If DP $\neq$ phase, only (2) matters
- **Prediction:** NC & DN subject to structural conditions, DP internal & sentential, that can differ within and across languages

The Case of French



	Pro DP	Pro Pro	DP DP	DP Pro
NC	53.33	53.33	33.33	32.08
DN	46.67	46.67	66.67	67.92

Déprez et al, 2015, Picture choice task

DN/NC and Prosody

Afrikaans *(Huddlestone 2010)*:

Results supporting a prosodic distinction between DN/NC readings

Dutch *(Fonville & de Swart 2013)*:

Mixed results: no characteristic prosody, but some patterns more closely associated with one type of reading
Problem: Design assumed context = accessed reading, no verification

Catalan *(Prieto et al 2013, Tubau et al 2015)*:

Clear prosodic distinction in DN vs NC readings in answers to negative questions with NCI
Problem: In French, answers to negative questions with an NCI are not ambiguous: always DN

French Prosody

- No experimental evidence that prosody is a disambiguating factor in French, but some notes in the literature *(Corblin 1996, Tovena & Corblin 2008)*
- French uses phrasing to mark focus *(Féry 2000)*, indicated by duration, intensity, and tone
- In simple SVO sentences, *Major Prosodic Phrases (MaP)* are identified by a pitch movement, a lengthening, or a pause *(Avanzi et al, 2014)*

Research Question

1. Is prosody used in disambiguating French transitive sentences with two NCIs?
2. What are the prosodic indicators which are employed by speakers to mark these differences?

Predictions/ Hypotheses

1. Like other languages (e.g.: Spanish, Catalan), the DN reading in French will be prosodically marked.
2. Speakers will use a high pitch accent and extended duration to indicate this markedness.

General Idea

- Native French speakers were presented with simple, ambiguous transitive sentences with one (Control) or two (Critical) NCIs in context
- They were then asked to (at their own pace)
  1. Read the entirety silently for comprehension
  2. Read it aloud (as though to a child)
  3. Respond to T/F question
- Responses were recorded on an Asus Orion PRO gaming headset with a noise filtering microphone
- Recording took place at the *Laboratoire sur le Langage, le Cerveau, et la Cognition (L2C2)* in Bron, France
- Total experimental time $\leq$ 20 minutes

Some Examples

NC Context

1. *Dans notre famille, on est tous allergique à l'alcool :*
   (In our family, we are all allergic to alcohol)

DN Context

2. *Chez les jeunes, la consommation d'alcool est effrayante :*
   (Among the youth, the rate of alcohol consumption is frightening)

Ambiguous Critical Item

3. **Personne ne boit rien dans les soirées.**
   (Nobody drinks nothing/ anything at parties)

Interpretation Verification

4. *Ils ne boivent pas d'alcool.*
   (They don't drink alcohol)
   = **T** for NC interpretation
   = **F** or DN interpretation

Stimuli

- 40 total context/sentence pairs (8 items × 5 conditions):
  1. 8 × **Double Negative:** *Personne ne mange rien ici*
  2. 8 × **Negative Concord:** *Personne ne mange rien ici*
  3. 8 × **Negative Object:** *Marie ne mange rien ici*
  4. 8 × **Negative Subject:** *Personne ne mange mie ici*
  5. 8 × **Fillers**
- Pseudorandomized
- Same 8 frequent monosyllabic verbs
- Same number of syllables in target sentence
- Maximized sonorant use where possible
- Final PP to avoid sentence boundary L tone on object NCI

Participants

- 20 native French speakers (M=4)
- Age 18-45 (mostly students at University of Lyon)
- Representative of diverse regions of France
- All had a minimum of a university degree

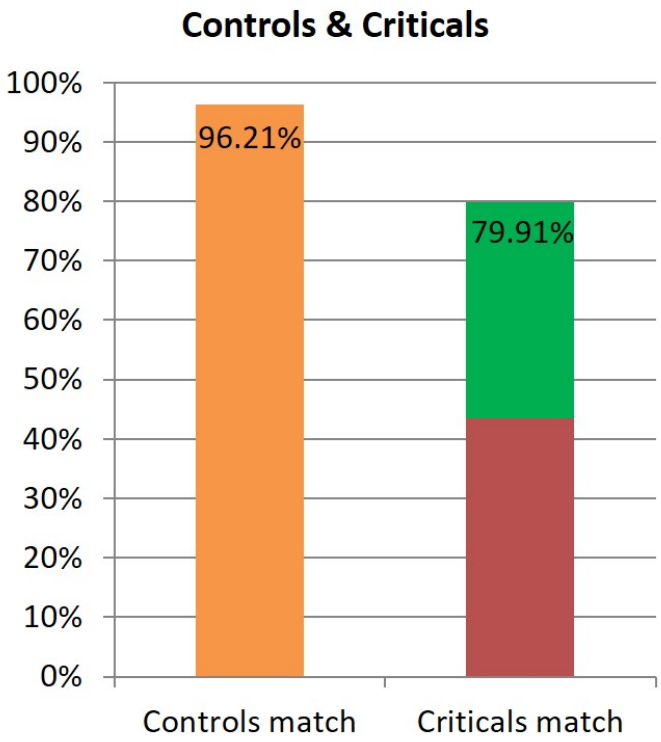
Data

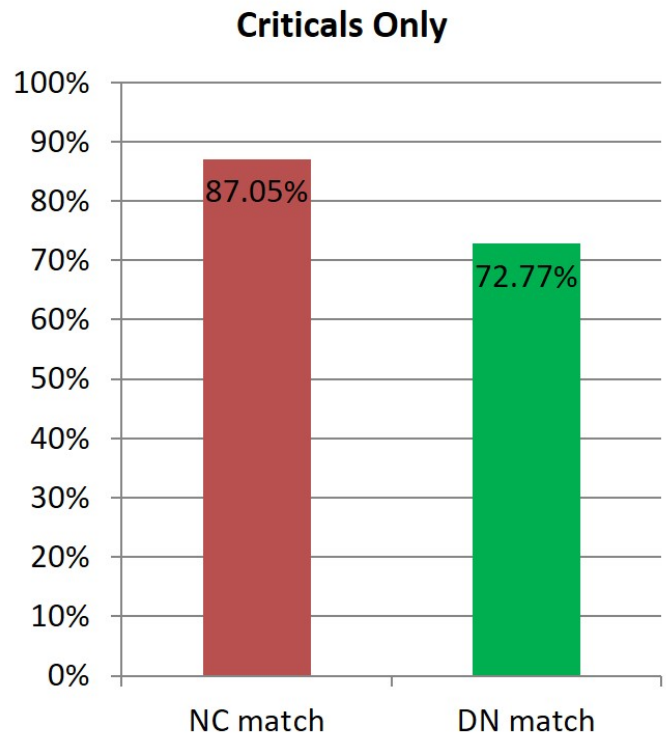
Condition	**Structure**	**Abbreviation**	**n**
Double Negation	NCI-NCI	DN	137
Negative Concord	NCI-NCI	NC	140
Subtotal Criticals			277
Single Negative Object	DP-NCI	NegOb	149
Single Negative Subject	NCI-DP	NegSub	149
Total			575

Analysis

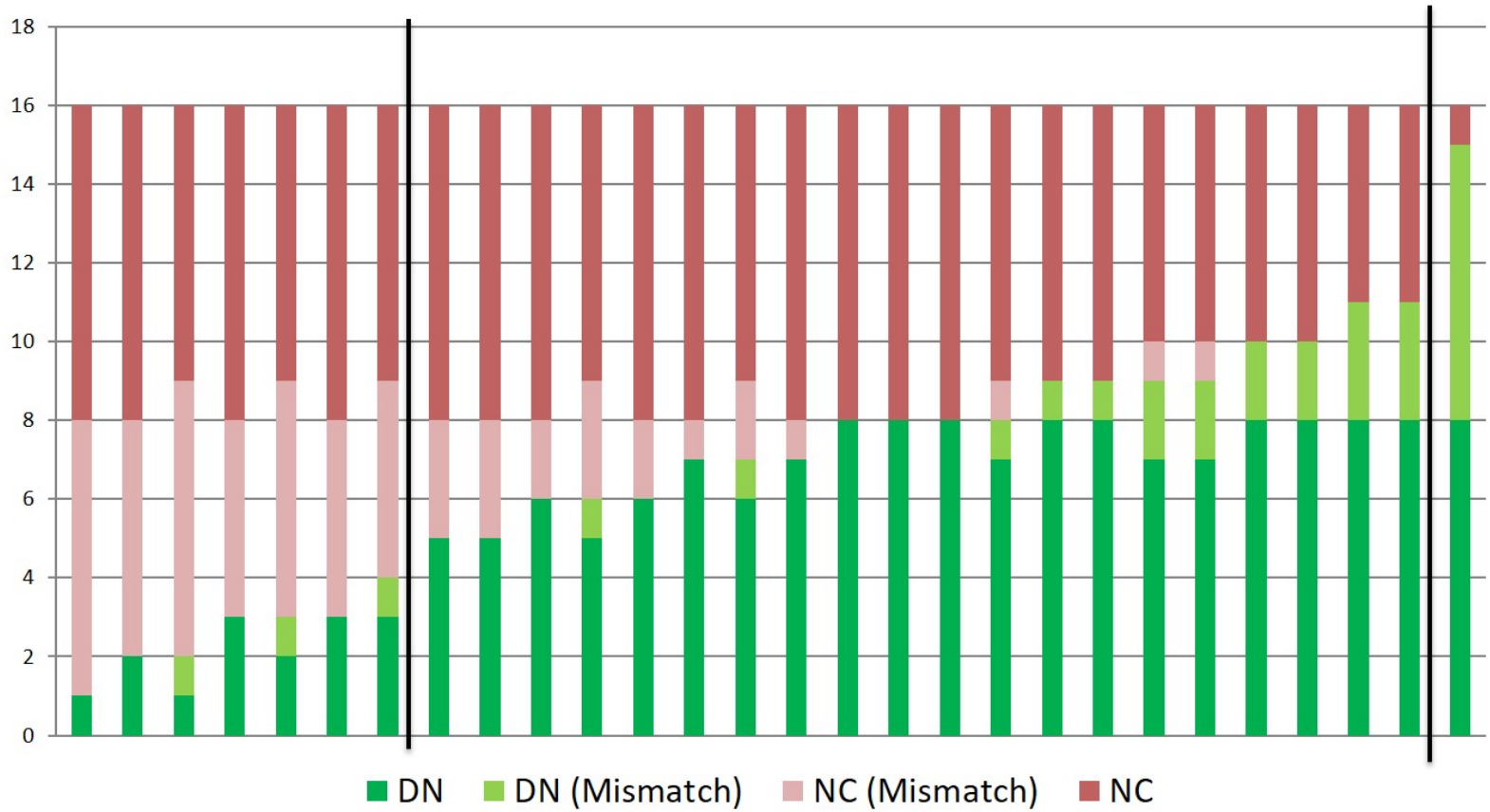
- Utterances were excised from context and text-aligned using **EasyAlign** *(J.-Ph. Goldman, 2011)* in **Praat** *(Boersma & Weenink, 2015)*
- Extracted for each syllable using **ProsodyPro** *(Xu, 2013)*:
  - Duration
  - Max F0
  - Min F0
  - 10 time-normalized F0 measurements
- Only the first 6 syllables are included: *per sonne ne [verb] rien PP*[1]
- F0 values were de-meaned
- Analysis performed in R (LM, LMEM)
- Removed 1,136/33,790 (3.4%) data points $\geq 3\sigma$ from μ

Behavioral Overview

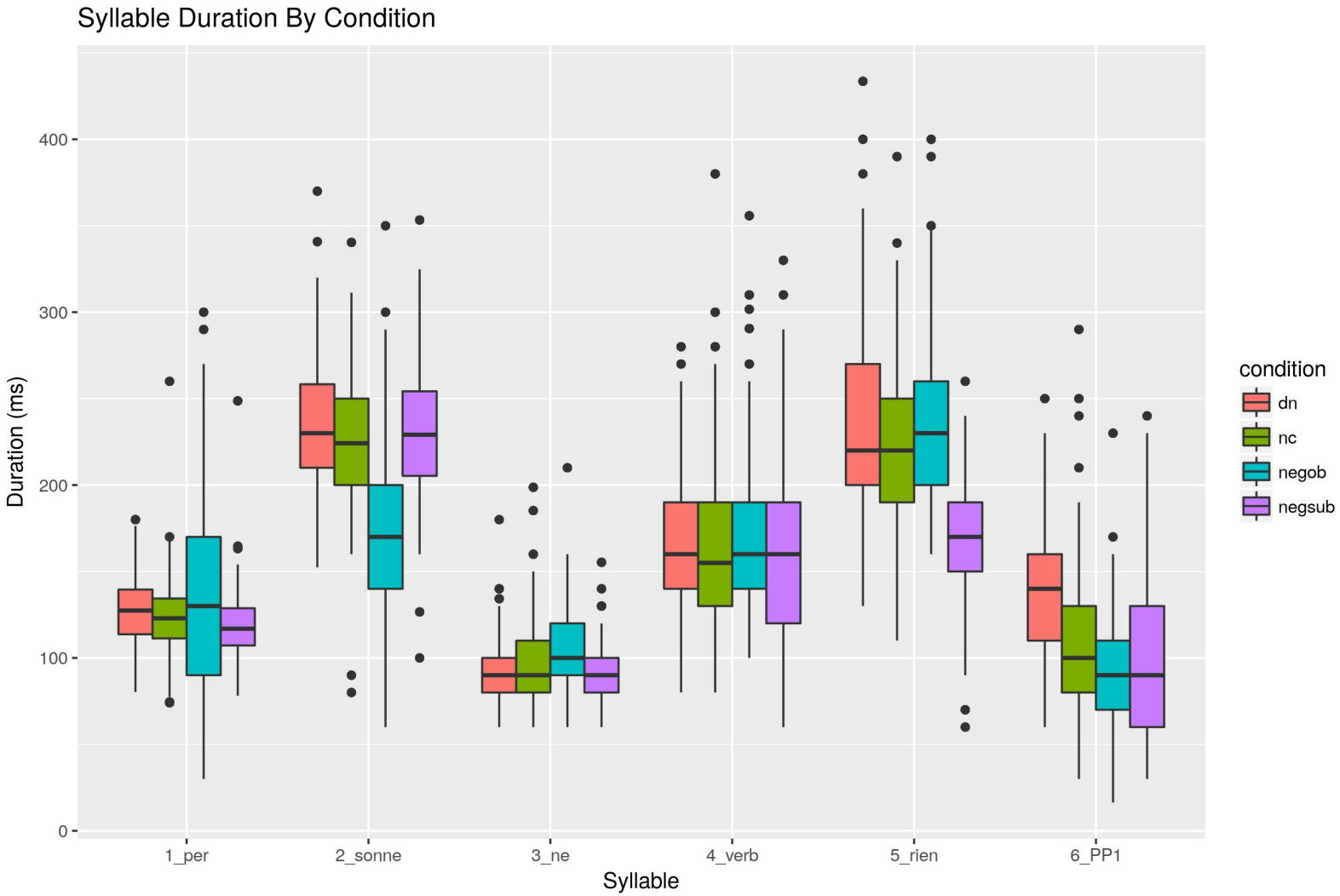




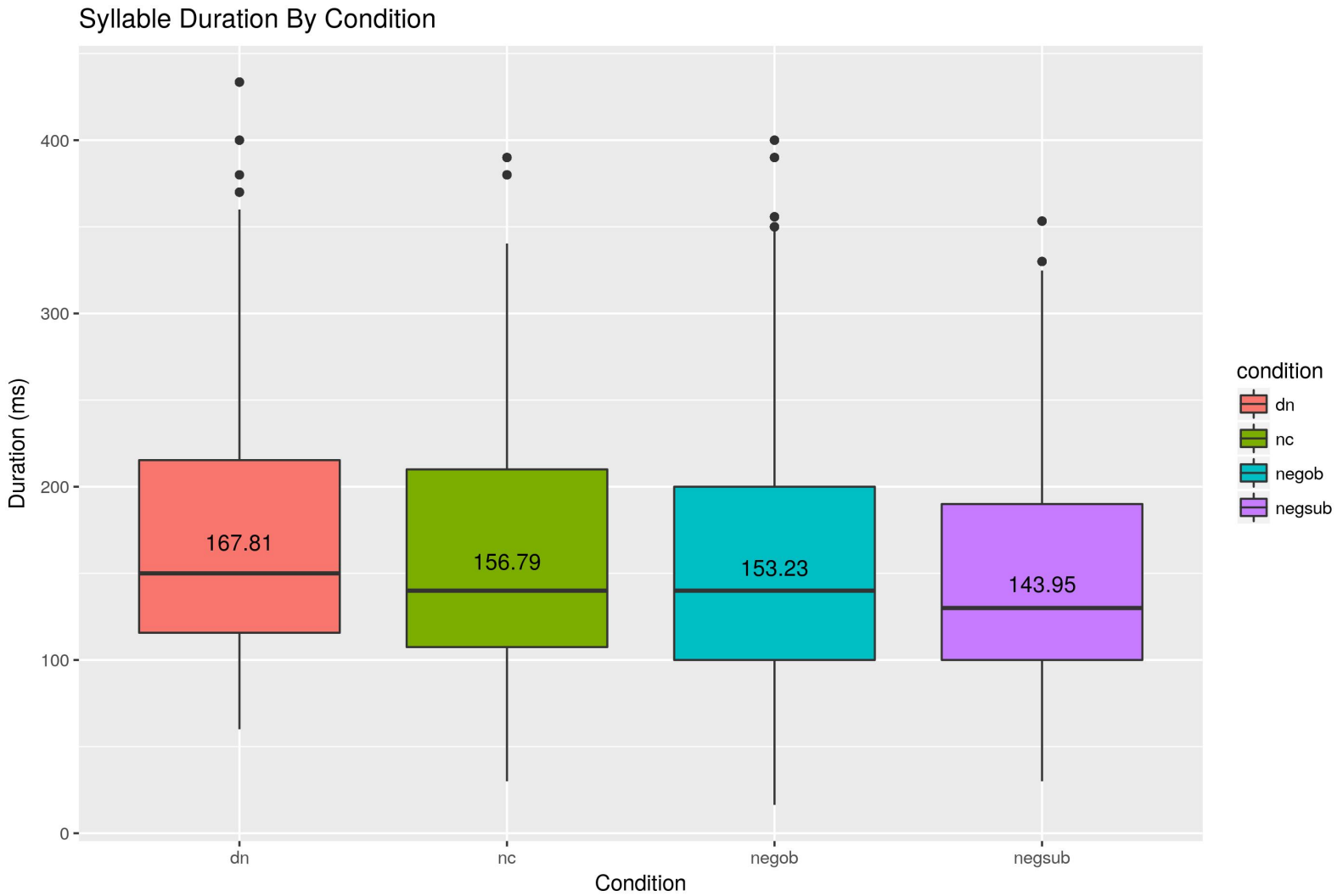
Behavioral Overview



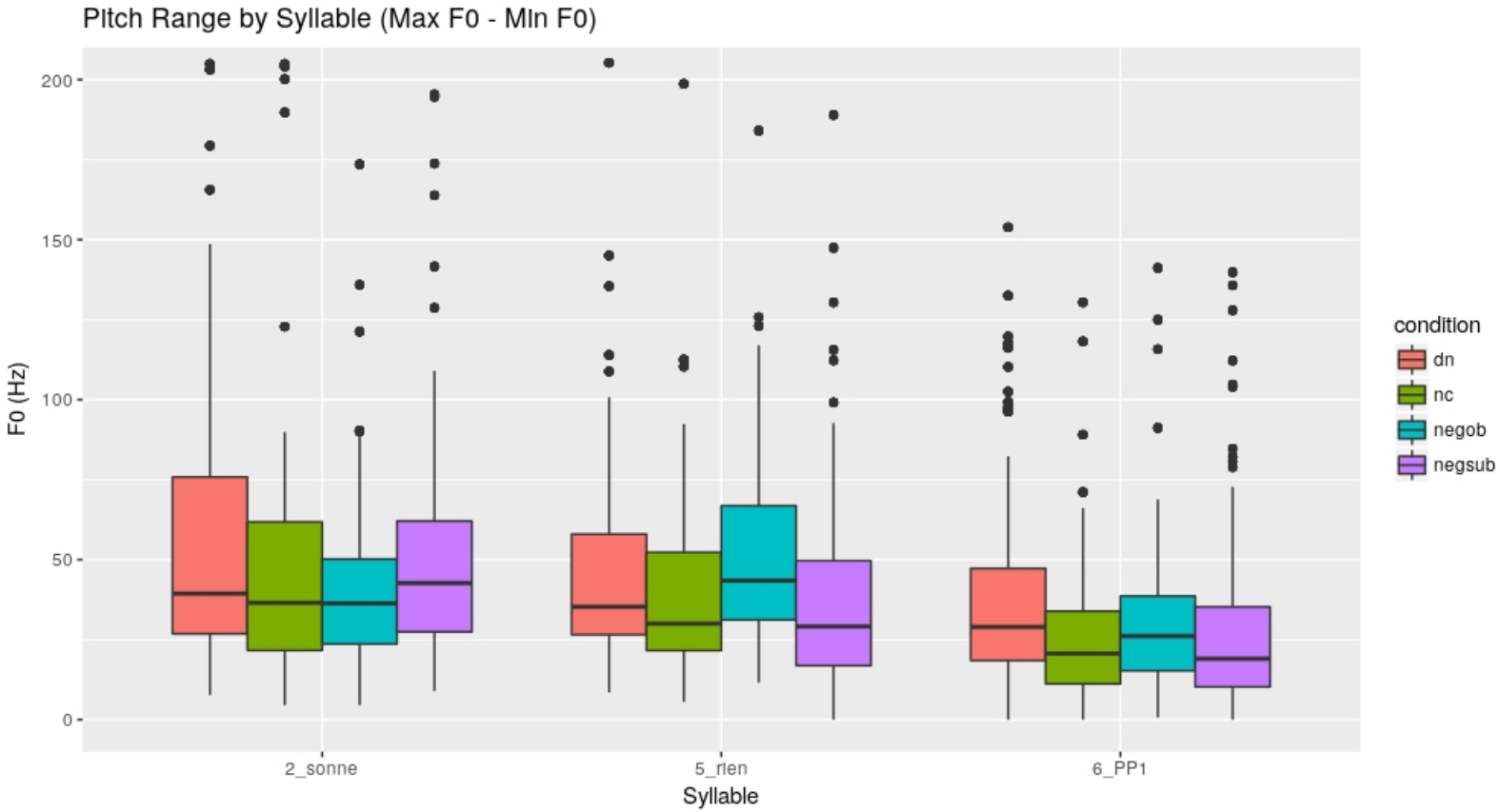
Duration


Syllable Duration By Condition
Duration (ms)
400
300
200
100
0
1_per
2_sonne
3_ne
4_verb
5_rien
6_PP1
Syllable
condition
dn
nc
negob
negsub

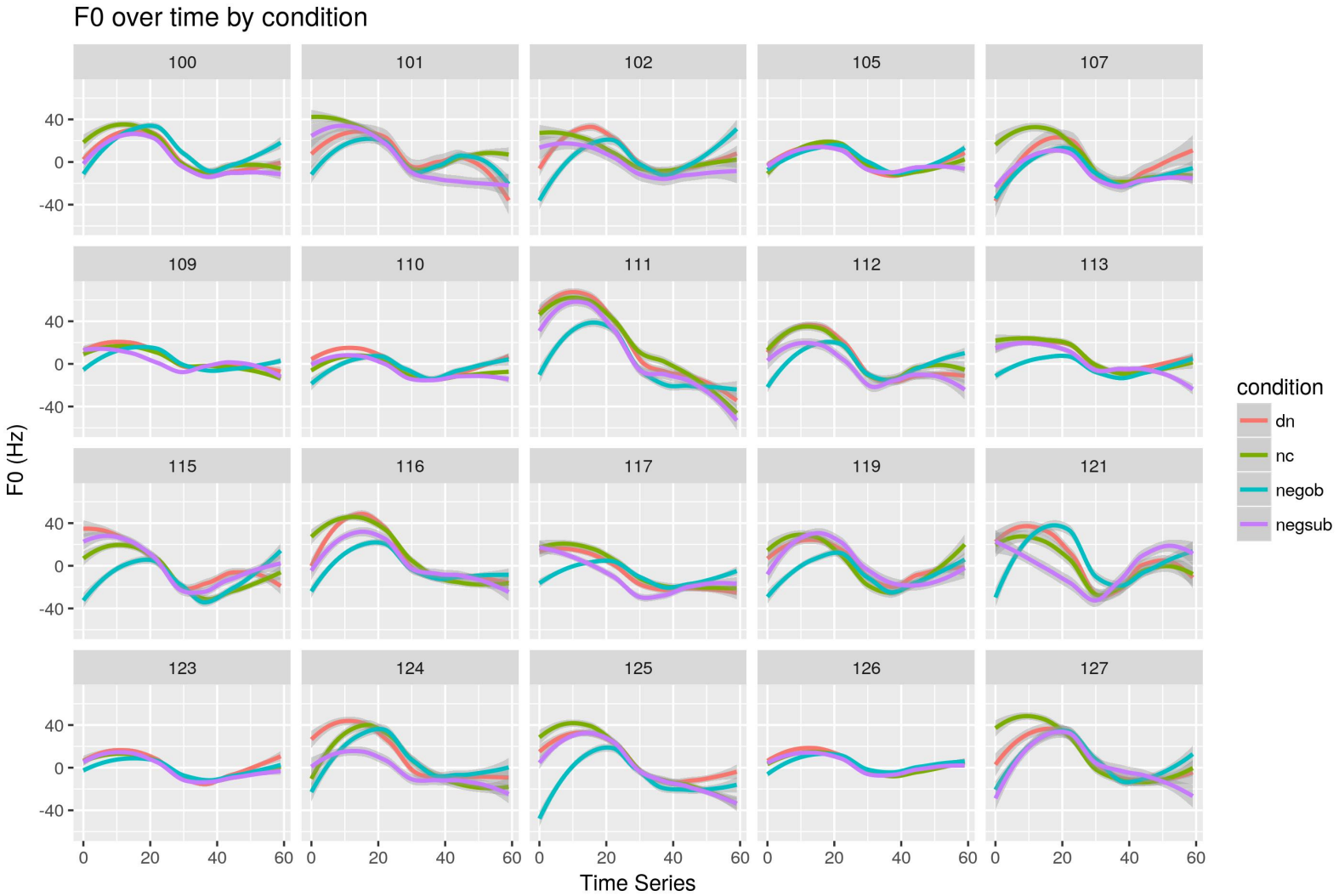
Duration


Syllable Duration By Condition
Duration (ms)
400
300
200
100
0
167.81
156.79
153.23
143.95
dn
nc
negob
negsub
Condition
condition
dn
nc
negob
negsub

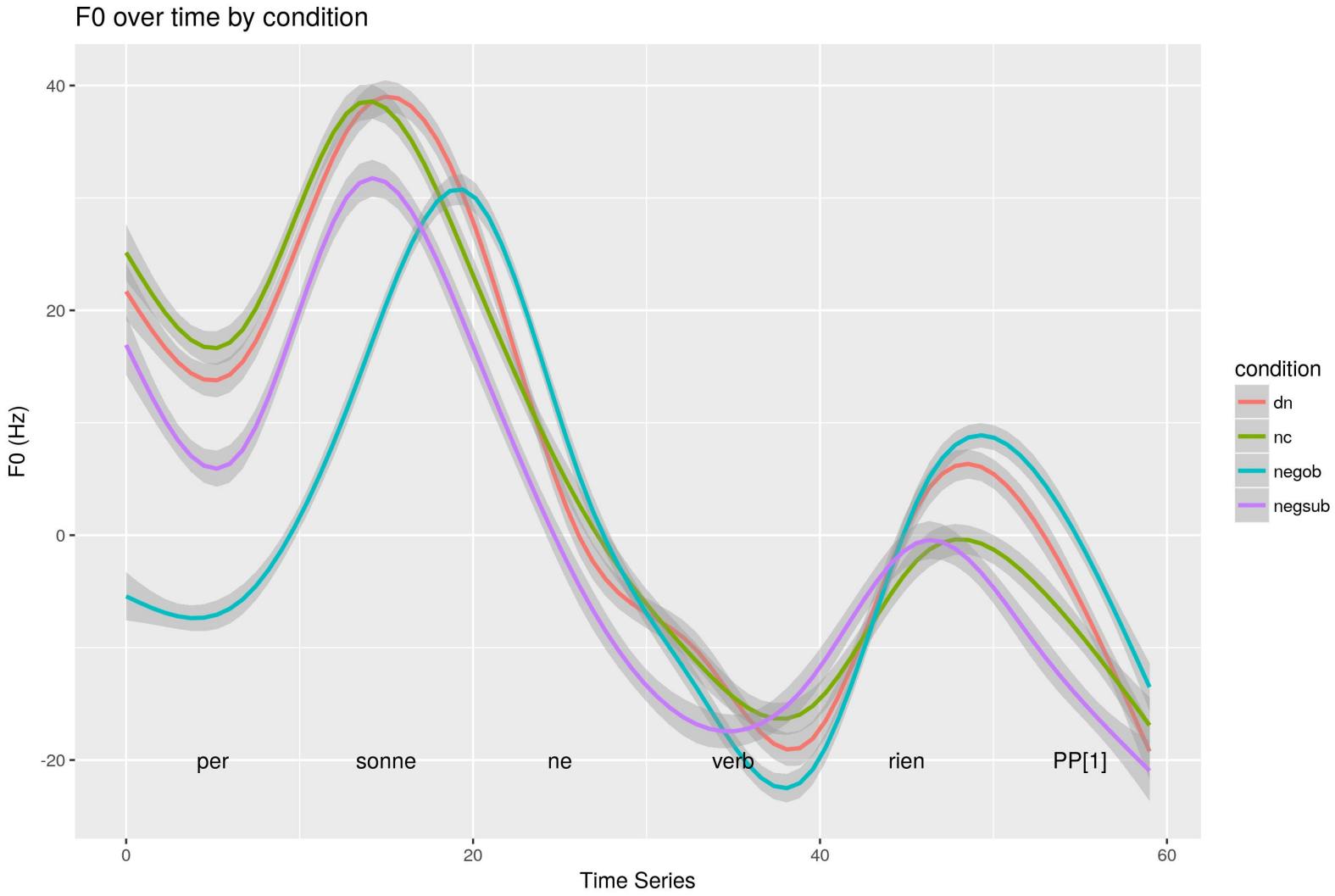
Range by Syllable



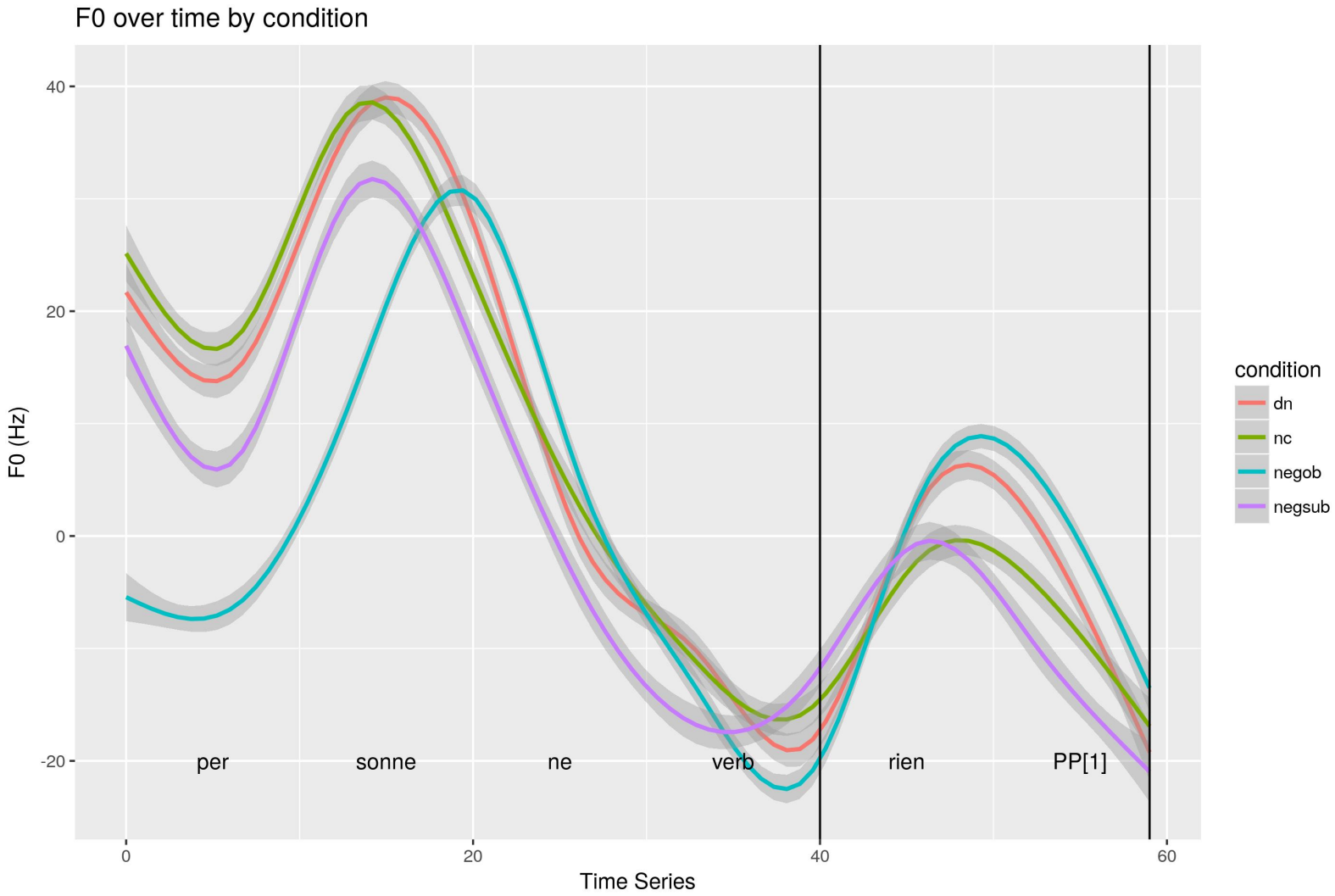
Pitch Contours by Subject



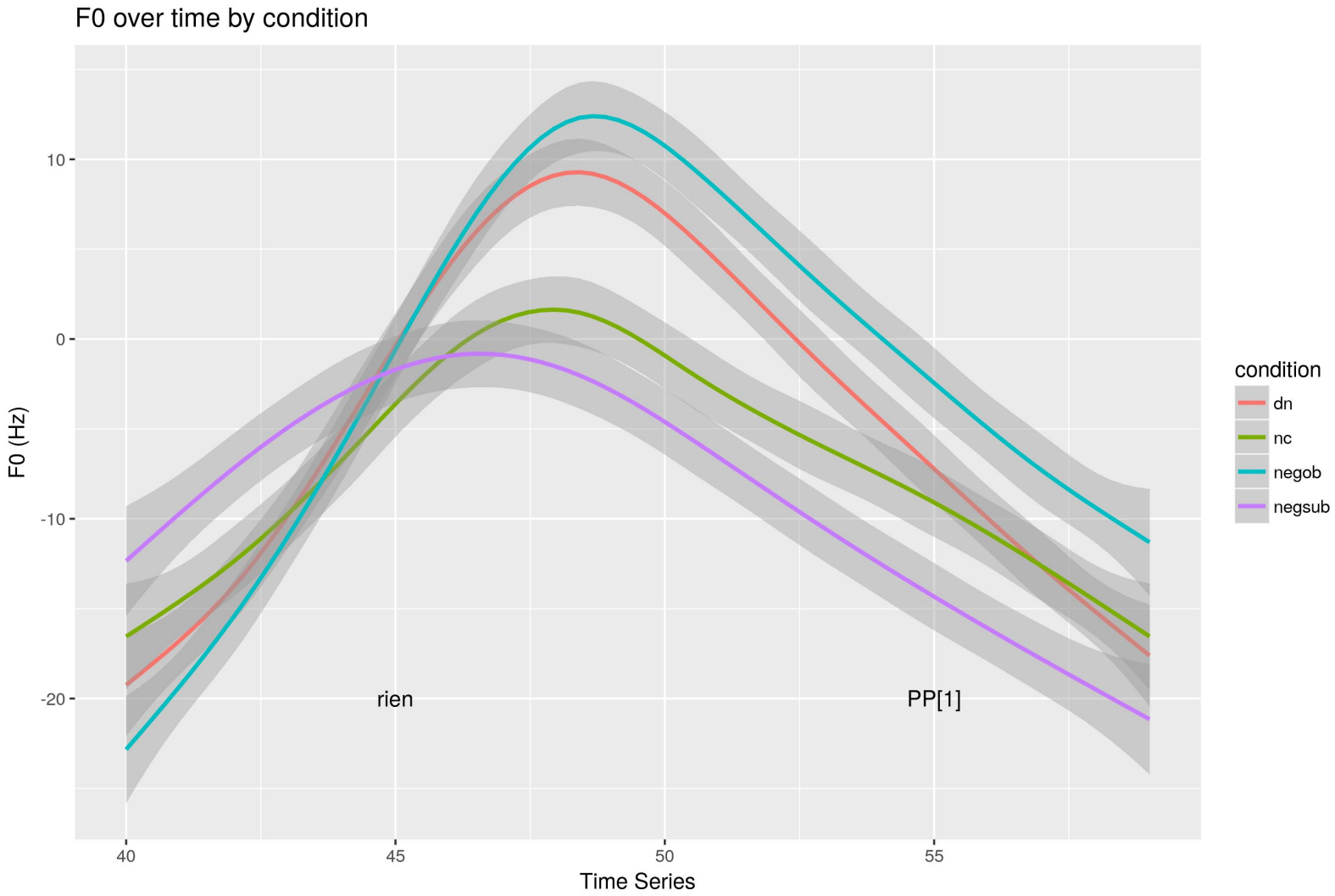
Aggregated Pitch Contours



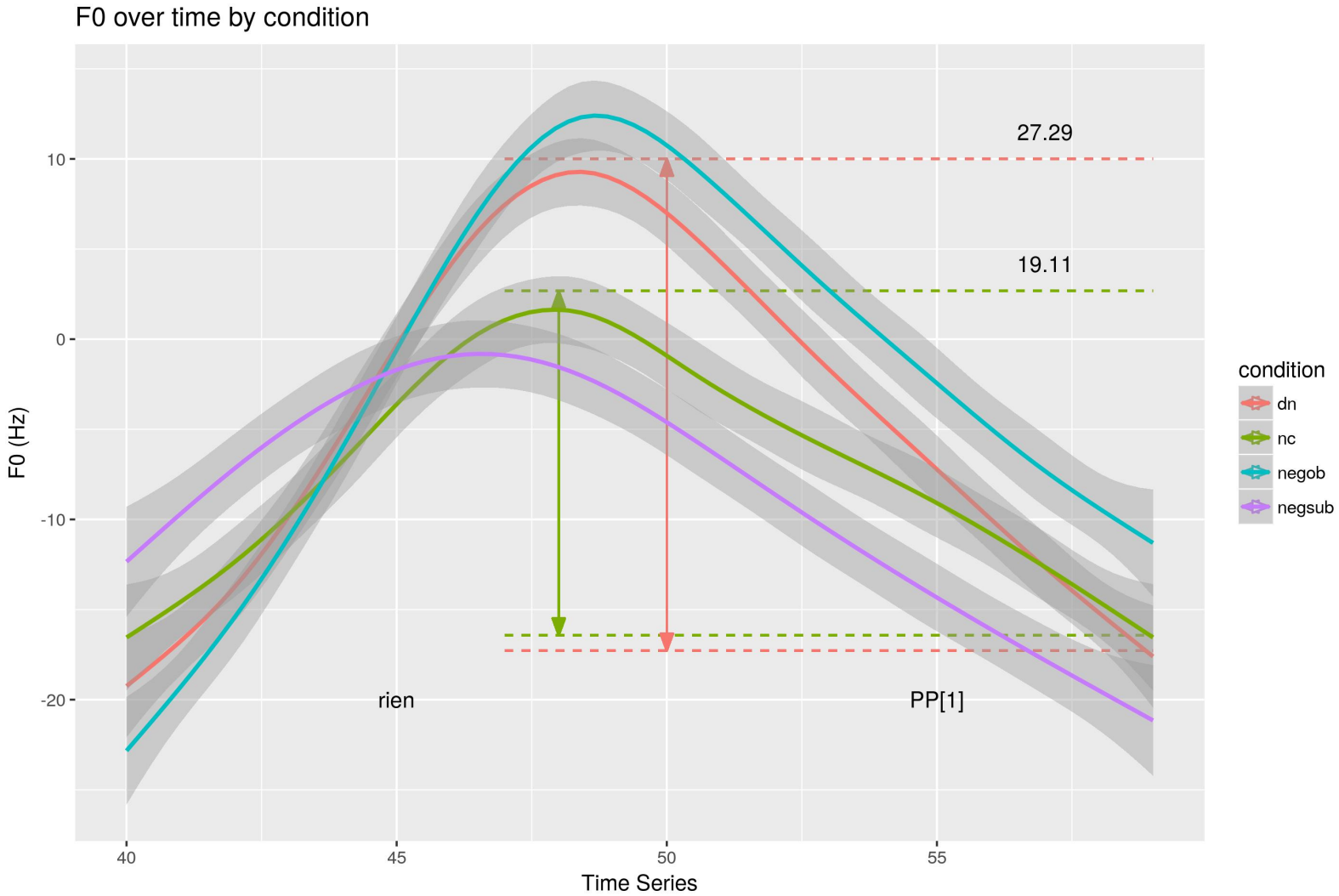
A Closer Look



A Closer Look



A Closer Look



NC Flattening

Linear Model:

$F0 \tilde{} \; timeseries + condition + time \times condition$

1. Interaction effect of *timeseries* × *condition*:
   NC $(t = -4.558, p = 5.25e-06)^{***}$
   NegSub $(t = 4.011, p = 6.11e-05)^{***}$
   NegOb $(t = 2.549, p = 0.0108)^{*}$
2. Main effect of *condition*:
   NC $(t = -5.077, p = 3.93e-07)^{***}$
   NegSub $(t = -4.771, p = 1.87e-06)^{***}$
   NegOb $(t = -2.006, p = 0.0449)^{*}$
3. Significant effect of *timeseries* across conditions

LM LMEM ANOVA

Back to the Research Questions

1. Is prosody used in disambiguating French transitive sentences with two NCIs?
2. What are the prosodic indicators which are employed by speakers to mark these differences?

Circling Back to the Predictions

1. Like other languages (e.g.: Spanish, Catalan), the DN reading in French will be prosodically marked.
2. Speakers will use a high pitch accent and extended duration to indicate this markedness.

Discussion: Duration

- Avanzi found that separate O phrasing is correlated with articulation rate (syllable duration)
- Extended duration in DN condition is encouraging for our hypothesis that NCI_2 is phrased separately

Discussion: F0

The F0 distinctions we see on the second NCI are realizations of phrasing and tone:

NC

- Focus on *personne*
- *rien* is phrased as part of VP
  L* H- L* H-/L- L%
  (([$_{DP}$**Personne**]$_{AP}$)([$_{VP}$ne V **rien**]$_{AP}$)...([$_{PP}$...PP...]$_{AP}$)$_{IP}$)

DN

- Focus on *personne*
- VP is "dephrased" *(Féry 2010)*
- Focus on *rien*, which forms its own phrase
  L* H- L* LH- L%
  (([$_{DP}$**Personne**]$_{AP}$) ne V ([$_{DP}$**rien**]$_{AP}$)...([$_{PP}$...PP...]$_{AP}$)$_{IP}$)

Conclusions

- The availability of both readings rules out a Macro-Parametric approach
- This could support either a Resumptive Quantification approach or a Micro-Parametric one
- The acoustic cues that we see in French to mark phrasing might be clues to Syntax:
  - **NC:** NCI_2 is phrased within the VP & has H- [$_{TP}$Personne [ne dit [$_{vP}$[$_{VP}$~~dit~~ rien]]]]
    NCI_2 remains inside VP, so its NEG *feature is not interpretable since it is not at an edge*
  - **DN:** NCI_2 forms its own prosodic phrase with LH- [$_{TP}$Personne [*ne dit* [$_{vP}$ rien [$_{VP}$ ~~rien~~]]]]
    NCI_2 is at vP edge where its NEG *feature is interpretable*

Outstanding Questions and Next Steps

- Are these differences actually perceptible to speakers?
- Might speakers emphasize these features more in a situation with less clear context?
- Could these strategies be employed with *less ambiguous* types of multiple NCI sequences to the same effect?
- How can we investigate the processing of these different readings with an ERP study? Does one have a higher cost?

Thank you for your attention!

Acknowledgements

Dr. Fanny Meunier, CNRS, L2C2
French Embassy in the U.S.
Rutgers Comparative and Experimental Linguistics Lab
Aresty Center for Undergraduate Research

jdyeaton27@gmail.com

R Output: Linear Model

```
Call:
lm(formula = demeaned_f0 ~ series * condition + series + condition,
    data = over46)

Residuals:
    Min      1Q  Median      3Q     Max
-87.291 -12.271   0.894  11.842  97.834

Coefficients:
                       Estimate Std. Error t value Pr(>|t|)
(Intercept)            128.7891     7.0734  18.207  < 2e-16 ***
series                  -2.4613     0.1332 -18.483  < 2e-16 ***
conditionnc            -50.4014     9.9274  -5.077 3.93e-07 ***
conditionnegob         -19.6224     9.7837  -2.006   0.0449 *
conditionnegsub        -47.1537     9.8829  -4.771 1.87e-06 ***
series:conditionnc       0.8520     0.1869   4.558 5.25e-06 ***
series:conditionnegob    0.4695     0.1842   2.549   0.0108 *
series:conditionnegsub   0.7461     0.1860   4.011 6.11e-05 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 20.76 on 7267 degrees of freedom
Multiple R-squared:  0.1558,  Adjusted R-squared:  0.155
F-statistic: 191.7 on 7 and 7267 DF,  p-value: < 2.2e-16
```

Back to results

R Output: LMEM

```
Linear mixed model fit by REML ['lmerMod']
Formula: demeaned_f0 ~ series * condition + condition + series + (1 +      condition | subj)
   Data: over46

REML criterion at convergence: 63525.4

Scaled residuals:
    Min      1Q  Median      3Q     Max
-4.7017 -0.5404 -0.0309  0.4849  5.4231

Random effects:
 Groups   Name            Variance Std.Dev. Corr
 subj     (Intercept)      60.61    7.785
          conditionnc      59.59    7.719   -0.15
          conditionnegob   83.66    9.147   -0.38  0.59
          conditionnegsub  49.58    7.041   -0.31  0.18  0.54
 Residual                 353.04   18.789
Number of obs: 7275, groups:  subj, 20

Fixed effects:
                      Estimate Std. Error t value
(Intercept)           130.2274     6.6366  19.622
series                 -2.4844     0.1206 -20.606
conditionnc           -52.0684     9.1518  -5.689
conditionnegob        -21.5778     9.0914  -2.373
conditionnegsub       -46.9842     9.0858  -5.171
series:conditionnc      0.8770     0.1692   5.182
series:conditionnegob   0.4982     0.1668   2.987
series:conditionnegsub  0.7371     0.1684   4.376

Correlation of Fixed Effects:
            (Intr) series cndtnnc cndtnngb cndtnngs srs:cndtnnc srs:cndtnngb
series      -0.963
```

R Output: ANOVA

```
Data: over46
Models:
full.lm: demeaned_f0 ~ series * condition + condition + series
ser_cond_slope.lmer: demeaned_f0 ~ series * condition + condition + series + (1 +
ser_cond_slope.lmer:     condition | subj)
                   Df   AIC   BIC logLik deviance  Chisq Chi Df Pr(>Chisq)
full.lm             9 64787 64849 -32385    64769
ser_cond_slope.lmer 19 63565 63696 -31763    63527 1242.2     10  < 2.2e-16 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
```

Back to results