# Every child matters:
# Statistical analysis with N-of-1

Giulia Calignano

12th May 2022

# outline

1. N = 1 it's a statistical thing
2. Longitudinal studies in Cognitive development
3. The present case
4. Hands on data (5)
5. GAMMs (4)
6. Open questions

# N-of-1, it's a thing

- An n-of-1 or single-subject study can be used to estimate mean trends over time
- Its target population consists of some larger set of time periods within a person.
- In psychology, *idiographic* (i.e., population-of-one) studies

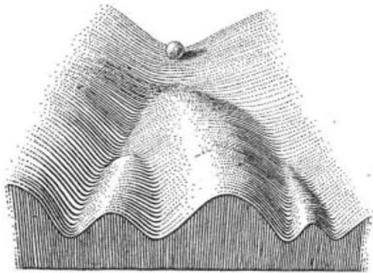You can think of an n-of-1 study as a sample from a population comprised entirely of events in one person's life:

**You are a world unto yourself [...].**
**In essence, "you" are a "y'all".**
Daza, 2019

# Longitudinal studies in Cognitive development



*C.H. Waddington (1940) – La metafora del paesaggio epigenetico*

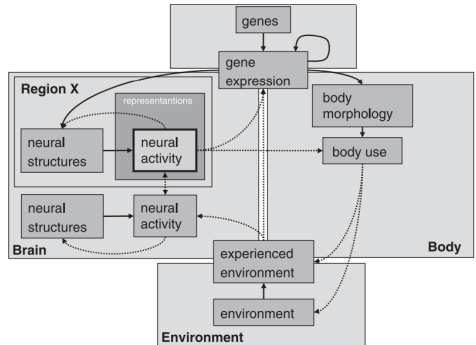Karmiloff-Smith, 1998
Westermann et al. 2007



**Figure 4** *The multiple interacting constraints shaping the construction of representations (neural activation patterns) in a specific cortical region (X). Region X here is not a primary sensory area so that the effects of environmental changes are mediated through other cortical regions. Representations can effect their own progressive change through multiple loops involving genes, other brain areas, the body and the environment. (Solid lines indicate a constraining relationship and dashed lines indicate the induction of change.)*

# The present case

**The early development of
lexical and post-lexical prominence in Italian. A case study**
PI: Cinzia Avesani, Claudio Zmarich
Center of National Research (CNR) of Padova
Mario Vayra
University of Bologna

# Corpus

- from 18 to 36 months of age
- **Modality**: repeated and spontaneous words; spontaneous sentences
- **Word types**: bi-, tri and quadrisyllabic words.
- **Stress on**: penultimate, ante-penultimate and final syllable
- **Sillable types**: CV and CVC
- **Segments**: V = all ; C = stops, fricatives, nasals, liquids, affricates One-, two- and multi-word utterances
- **Metrical positions**: unstressed (U) , stressed (S), prenuclerly accented (P) , nuclearly accented in intermediate/phonological (Nphp) and intonational phrases (N)

**TOTAL number of occurrences: 3454**

# Hands on data

- CV syllables
- Spontaneously produced
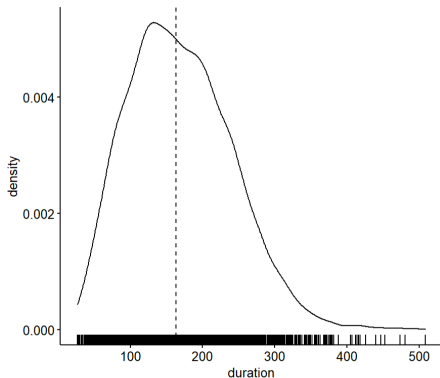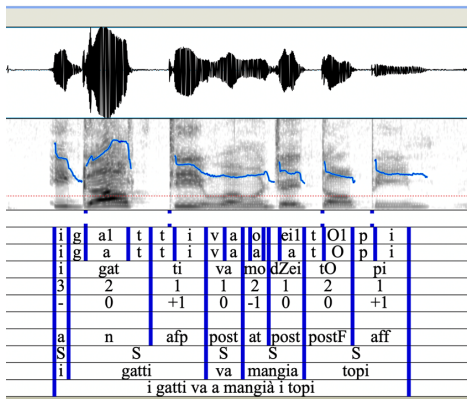- Position in word/sentence: final

**TOTAL SELECTED = 1156**

*Excluded from selection*:

Stressed/accented vowels in word-final position Unstressed vowels in word and sentence-final position

All child productions that did not conform to adult targets (e.g. due to stress inversions)

# Hands on data

**dependent variable**: syllables duration

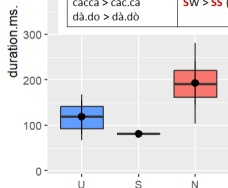# Descriptive stats: duration 18-30 month of age



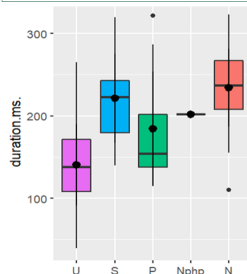| Trochaic bias (Gerken, 1991;1994) / Emergentist approach (Vihman 2018) | |
|---|---|
| tì.ti > tì.ti | **SW** > **WS** (22%) |
| tà.ta > tà.tà càcca > càc.cà dà.do > dà.dò | **SW** > **SS** (38%) |

**18**

**multi-word utterances with non-canonical order**
mamma palla damme; biberon fa mamma; limone ti do questo

**21**

**Multi-word utterances with canonical order – direct and indirect arguments**
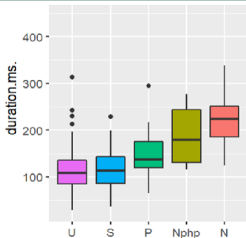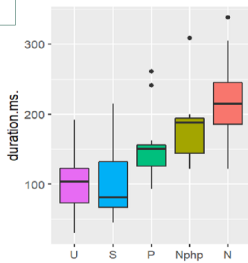salta la corda; la mamma cuce; chiama Ghighi con le scarpe; mette il vino nella bottiglia

**24**

*Sentence-initial Focus*
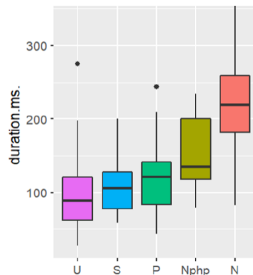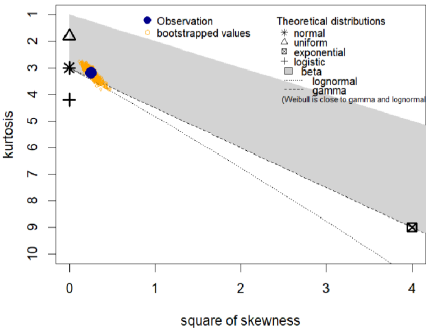:on la corda, salta; sotto la macchina, va; bambole, voglio

27

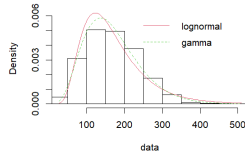*Sentence-final and sentence-initial Focus – postfocal accents*

30

36

# Distribution family
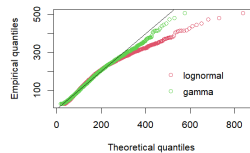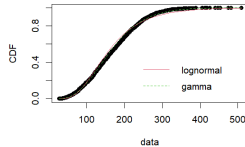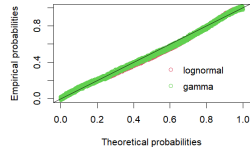
## Model comparison

Generalized Additive Mixed-effects Models (GAMMs)

**Pros2Eng:** categorical factor (5: U, S, P, Nphp, N)

```
m0: duration ~ s(Age, by = Pros2.Engl, k = 7)

m1: duration ~ Age * Pros2.Engl + s(Age, by = Pros2.Engl, k = 7)

Chi-square test of fREML scores
-----
  Model    Score Edf Difference    Df  p.value Sig.
1   m0 500.0548  11
2   m1 366.1923  20     133.863 9.000  < 2e-16  ***

AIC difference: 169.57, model m1 has lower AIC.
```

# Model check

Basis dimension (k) checking results. Low p-value (k-index<1) may
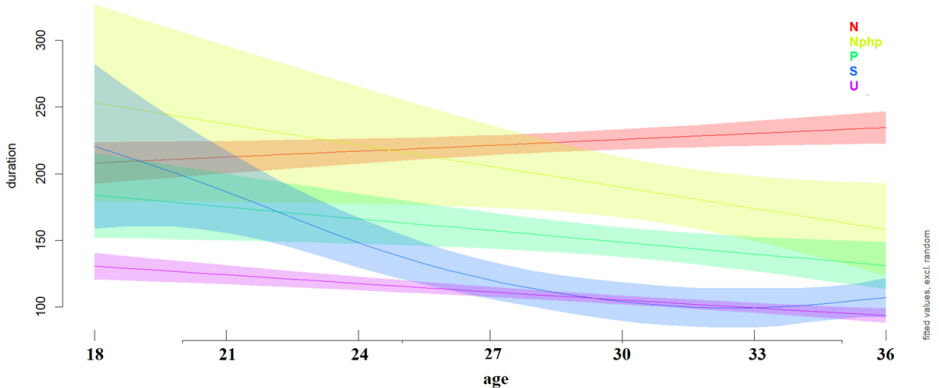indicate that k is too low, especially if edf is close to k'.

```
                      k'   edf k-index          p-value
s(Age):Pros2.EnglN   6.00 1.00     0.9 <0.000000000000002 ***
s(Age):Pros2.Englnphp 6.00 1.00   0.9             0.005 **
s(Age):Pros2.EnglP   6.00 1.00     0.9 <0.000000000000002 ***
s(Age):Pros2.EnglS   6.00 2.64     0.9             0.005 **
s(Age):Pros2.EnglU   6.00 1.15     0.9 <0.000000000000002 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```



ACF of resid(m1)
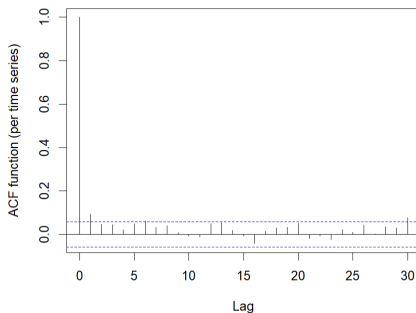
# Visualizing the Estimated effects

Would be insane to draft theoretical models of development an individual at a time?

May be less insane in a Bayesian perspective?

# References

Delignette-Muller, M. L., Dutang, C., Pouillot, R., Denis, J. B., & Delignette-Muller, M. M. L. (2015). Package 'fitdistrplus'. *Journal of Statistical Software,* 64(4), 1-34.

Karmiloff-Smith, A. (1998). Development itself is the key to understanding developmental disorders. *Trends in cognitive sciences*, 2(10), 389-398.

van Rij, J., Wieling, M., Baayen, R. H., van Rijn, H., & van Rij, M. J. (2017). Package 'itsadug'.

Westermann, G., Mareschal, D., Johnson, M. H., Sirois, S., Spratling, M. W., & Thomas, M. S. (2007). Neuroconstructivism. *Developmental science*, 10(1), 75-83.

Wickham, H. (2011). ggplot2. *Wiley interdisciplinary reviews: computational statistics*, 3(2), 180-185.

Wood, S. N. (2001). mgcv: GAMs and generalized ridge regression for R. R news, 1(2), 20-25.