The role of Phonological Neighborhood on error type: a Multinomial analysis

PSICOSTAT MEETINGS

OCTOBER 2024

Irene Bellin

Erica Iob

Irene Bellin
Speech Therapist
MCs in Cognitive Science
PhD student in Neuroscience
Padova Neuroscience Center - PNC
University of Padua, Italy

Erica Iob
Speech Therapist
MA in Linguistics
Future PhD student in Cognitive Computational Social Neuroscience - IMT Lucca, Italy

Prof. Eduardo Navarrete Sanchez
Associate Professor DPSS - University of Padua, Italy

Contents

Theoretical background

Linguistic studies on atypical populations have always provide useful insights for the comprehension of healthy functioning systems. The focus of the project is the analysis of ***Formal errors*** in aphasic speech, that seem to be the manifestation of an interaction between the lexical and phonological system of language processing (Gagnon et al., 1997).

Formal error:

A *formal error* is defined as the production of an existing word in a language lexicon which is phonologically related with the target (e.g. *hat* for "cat") (Blanken, 1990; Laganaro, 2013)

Theoretical background

Linguistic studies on atypical populations have always provide useful insights for the comprehension of healthy functioning systems. The focus of the project is the analysis of ***Formal errors*** in aphasic speech, that seem to be the manifestation of an interaction between the lexical and phonological system of language processing (Gagnon et al., 1997).

Psycholinguistic studies have identified the existence of **critical predictors** that influence speech accuracy that intrinsically belong to target words that are meant to say. **Phonological Neighbourhood Density** (PhND) is an index of phonological similarity of words in the lexicon. Unless words in dense PhND usually are named faster (Gordon, 2002), some studies on healthy participants reported the reverse pattern (Sadat et al., 2014; Vitevitch et al., 2016).

Experimental questions

a) Which psycholinguistic predictor influence speech **accuracy** in lexical production?

b) Do those predictors have an influence on **error type**? And, specifically, what is the influence on formal errors?

Participants

189 patients, mean age = 67 y. (SD = 15.2, range = 21 - 91)

Patients were tested in the Clinica Neurologica of Padua Hospital (data: 2007 to 2019)

Including criteria: a) Aphasia diagnosis; b) Italian native speakers;

Excluding criteria: a) Speech Apraxia diagnosis; b) Dysarthria diagnosis; c) Presence of neologistic jargon; d) Absence of paraphasias

Materials

Data consist in participants' responses at **Object naming** tests of several clinical evaluation batteries: AAT (Luzzati et al., 1996), ELLM (Allibrio et al., 2008), BNT (Kaplan et al., 2001), BADA (Miceli et al., 1991). N=**5498**

Participants were exposed to different number of items (mean of trials = 36; SD =39.2; range 12 - 268)

Predictors

Several psycholinguistic variables representing phonological and lexical features that notoriously affect lexical access, were controlled (values taken from *PhonItalia*, (Goslin et al., 2013)):

Length of the word (**Num Phonemes**); Lexical Frequency (a logarithmic transformation was applied, **Freq (log)**); Phonological Neighborhood Density (**PhND**); Frequency of the Phonological Neighborhood (**PhND_Freq**).

Predictors

Several psycholinguistic variables representing phonological and lexical features that notoriously affect lexical access, were controlled (values taken from *PhonItalia*, (Goslin et al., 2013)):

Length of the word (**Num Phonemes**); Lexical Frequency (a logarithmic transformation was applied, **Freq (log)**); Phonological Neighborhood Density (**PhND**); Frequency of the Phonological Neighborhood (**PhND_Freq**).

Procedure

Responses were categorized as *Correct* vs *Error*

Errors were in turn were categorized in error types: *Formal Errors; Semantic Errors; Non-Words; Non-Response; Unrelated lexical Errors; Others.*

Item analysis

a) Which psycholinguistic predictor influence speech **accuracy** in lexical production?

Item analysis

a) Which psycholinguistic predictor influence speech **accuracy** in lexical production?

For the first analysis we used a **binomial** logistic regression;

Item analysis

a) Which psycholinguistic predictor influence speech **accuracy** in lexical production?

For the first analysis we used a **binomial** logistic regression;

b) Do those predictors have an influence on **error type**? And, specifically, what is the influence on formal errors?

Item analysis

a) Which psycholinguistic predictor influence speech **accuracy** in lexical production?

For the first analysis we used a **binomial** logistic regression;

b) Do those predictors have an influence on **error type**? And, specifically, what is the influence on formal errors?

As for the second analysis, we performed **multinomial** logistic regression*;

Previous analysis

b) Do those predictors have an influence on **error type**? And, specifically, what is the influence on formal errors?

*Previous analysis: **binomial**

Different statistical models considering:

Formal Errors vs Semantic Errors

Formal Errors vs Unrelated Lexical Errors

Formal Errors vs Non-Response

Formal Errors vs Non-Word Errors

New Analysis: A multinomial analysis was performed in order to analyze responses within a single statistical model, without fitting different models for alternative response categories (Toffalini et al., 2020) and given the fact that the previous analysis was biased towards Formal Errors.

Item analysis

Items that are not present in *PhonItalia* (Goslin et al., 2013; N=65), items with syllable number greater than 4 (N=400), error type *Others* (only for (b) analysis; too heterogeneous) were excluded from the analysis.

Mixed Effects Models were conducted with Items and Participants as random factors. Analyses were performed using the **brms** package on the R software (Bürkner, 2017).

Results

Accuracy measure:

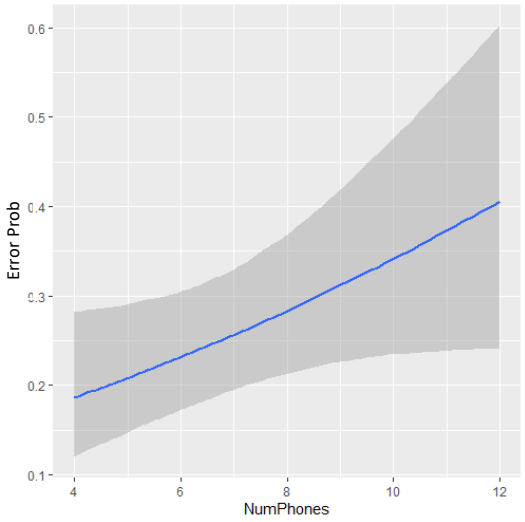
formula_Ac1=ErroreNum ~ NumPhones + logfqTot + Phon_N + Phon_N_MFreq + (1|Pacient)+(1|Item)

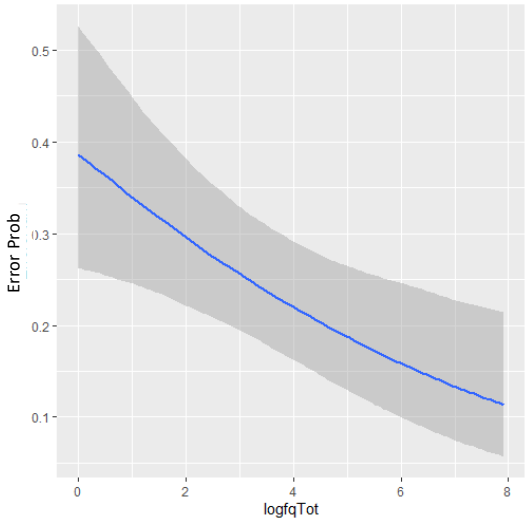
mod_Ac1=brm(formula=formula_Ac1, data=d2, family=bernoulli, chains=4)

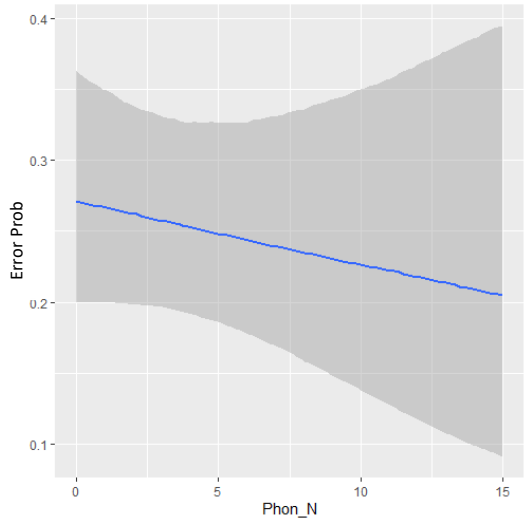
```
Population-Level Effects:
             Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
Intercept       -0.81      0.68    -2.14     0.57 1.00     2242     2641
NumPhones        0.14      0.07    -0.00     0.28 1.00     2268     2700
logfqTot        -0.20      0.07    -0.34    -0.06 1.00     2093     2734
Phon_N          -0.02      0.04    -0.10     0.05 1.00     2302     2740
Phon_N_MFreq    -0.27      0.09    -0.44    -0.10 1.00     1872     2598
```

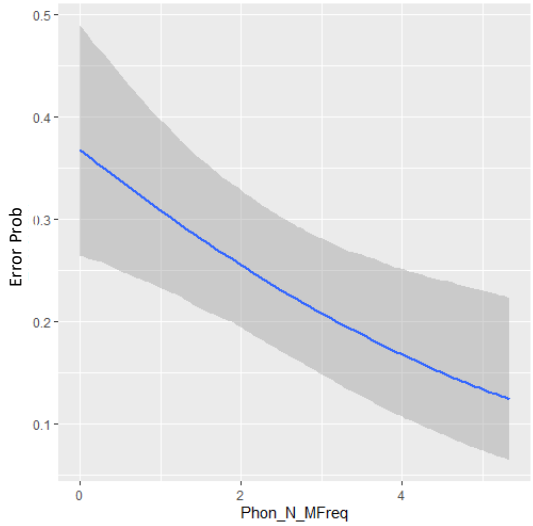
Results

Accuracy measure:









Results

Error type measure:

mod_Type1=brm(ErrorType2 ~ NumPhones + logfqTot + Phon_N + Phon_N_MFreq + (1|Pacient) + (1|Item), data=d3, family=categorical)

```
Population-Level Effects:
                            Estimate Est.Error l-95% CI u-95% CI
muResp2NR_Intercept            -4.26      1.01    -6.22    -2.29
muResp3Formal_Intercept        -4.84      0.94    -6.72    -3.02
muResp4Sem_Intercept           -2.78      1.30    -5.33    -0.22
muResp5NonWord_Intercept       -2.82      0.67    -4.11    -1.46
muResp6LexicalOthers_Intercept -3.50      1.24    -6.03    -1.15
muResp2NR_NumPhones             0.24      0.10     0.02     0.44
muResp2NR_logfqTot             -0.24      0.10    -0.45    -0.05
muResp2NR_Phon_N               -0.06      0.06    -0.17     0.04
muResp2NR_Phon_N_MFreq         -0.20      0.13    -0.46     0.05
muResp3Formal_NumPhones         0.14      0.09    -0.04     0.33
muResp3Formal_logfqTot         -0.03      0.09    -0.21     0.13
muResp3Formal_Phon_N            0.01      0.05    -0.08     0.10
muResp3Formal_Phon_N_MFreq     -0.17      0.11    -0.39     0.04
muResp4Sem_NumPhones            0.04      0.14    -0.24     0.31
muResp4Sem_logfqTot            -0.17      0.14    -0.44     0.11
muResp4Sem_Phon_N               0.06      0.07    -0.08     0.21
muResp4Sem_Phon_N_MFreq        -0.34      0.18    -0.69     0.01
muResp5NonWord_NumPhones        0.13      0.07    -0.01     0.26
muResp5NonWord_logfqTot        -0.19      0.07    -0.33    -0.06
muResp5NonWord_Phon_N          -0.05      0.03    -0.11     0.01
muResp5NonWord_Phon_N_MFreq    -0.20      0.08    -0.34    -0.04
muResp6LexicalOthers_NumPhones  0.00      0.13    -0.25     0.26
muResp6LexicalOthers_logfqTot   0.05      0.14    -0.21     0.33
muResp6LexicalOthers_Phon_N    -0.03      0.06    -0.15     0.10
muResp6LexicalOthers_Phon_N_MFreq -0.20   0.15    -0.50     0.08
```

Discussion

a. Role of Intercept: does it make sense to set *Correct Responses* as the Intercept? What are the theoretical implications? Considering a subset of only errors for (b) analysis, which type pf error should be set as the Intercept? Does it make sense to set *Formal Errors* as the Intercept? What are the theoretical implications?

Discussion

a. Role of Intercept: does it make sense to set *Correct Responses* as the Intercept? What are the theoretical implications? Considering a subset of only errors for (b) analysis, which type pf error should be set as the Intercept? Does it make sense to set *Formal Errors* as the Intercept? What are the theoretical implications?

b. Interpretation: how should we interpret results? Is it sufficient to consider Bayesian Credible Intervals?

Bibliography

Allibrio , G., Gori, A., Signorini, G., & Luzzatti, C.G. (2008). Esame del linguaggio al letto del malato (Empoli)

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. arXiv preprint arXiv:1506.04967.

Best, W. (1996). When Racquets Are Baskets But Baskets Are Biscuits, Where Do the Words Come From ? A Single Case Study of Formal Paraphasic Errors in Aphasia.

Blanken, G. (1990). Formal paraphasias: A single case study. Brain and Language, 38, 534-554.

Bürkner, P.-C. (2017). Brms: An R package for Bayesian multilevel models using stan. Journal of Statistical Software, 80, 1–28. https://doi.org/10.18637/jss.v080.i01

Gagnon, D.A., Schwartz, M.F., Martin, N.F., Dell, G.S., & Saffran, E.M. (1997). The Origins of Formal Paraphasias in Aphasics' Picture Naming. Brain and Language, 59, 450-472.

Gordon, J.K. (2002). Phonological neighborhood effects in aphasic speech errors: spontaneous and structured contexts. Brain and Language, 82, 113-145.

Goslin , J., Galluzzi, C., & Romani, C. (2013). PhonItalia : a phonological lexicon for Italian Behavior Research Methods

Kaplan, E., Goodglass , H., & Weintraub , S. (2001). Boston naming test . Pro ed

Laganaro, M., Chetelat-Mabillard, D., & Frauenfelder, U.H. (2013). Facilitatory and interfering effects of neighbourhood density on speech production: evidence from aphasic errors. Cognitive neuropsychology, 30 3, 127-46 .

Luzzatti C, Willmes K, De Bleser R. Aachener Aphasie Test: Versione Italiana (Seconda edizione). Firenze 1996; O.S. Organizzazioni Speciali

Miceli, G., Laudanna, A., Burani, C., & Campasso , R., Batteria per l'Analisi dei Deficit Afasici B.A.D.A., 1991; CEPSAG Università del Sacro Cuore

Pastore, M. (2015). Analisi dei dati in psicologia. Il Mulino.

Sadat, J., Martin, C. D., Costa, A., & Alario , F. X. (2014). Reconciling phonological neighborhood effects in speech production through single trial analysis. Cognitive psychology 68 , 33 58.

Toffalini, E., Mirandola, C., De Simone Irace, C., & Altoè, G. (2020). False memory for pictorial scripted material: the role of distinctiveness and negative emotion. *Cognition and Emotion*, *34*(7), 1489-1498

Vitevitch , M.S., & Luce, P.A. (2016). Phonological Neighborhood Effects in Spoken Word Perception and Production

Thank you.

Results - Supplementary

Error type measure:

We considered **only errors**!

mod_Type1=brm(ErrorType2 ~ NumPhones + logfqTot + Phon_N + Phon_N_MFreq + (1|Pacient) + (1|Item), data=d3, family=categorical)

```
Population-Level Effects:
                           Estimate Est.Error l-95% CI u-95% CI Rhat
muResp3Formal_Intercept       -1.42      1.13    -3.70     0.73 1.00
muResp4Sem_Intercept           0.23      1.60    -2.99     3.38 1.00
muResp5NonWord_Intercept       1.27      0.88    -0.44     3.04 1.00
muResp6LexicalOthers_Intercept -0.08      1.41    -2.93     2.64 1.00
muResp3Formal_NumPhones       -0.09      0.12    -0.31     0.14 1.00
muResp3Formal_logfqTot         0.12      0.12    -0.11     0.34 1.00
muResp3Formal_Phon_N           0.06      0.06    -0.06     0.17 1.00
muResp3Formal_Phon_N_MFreq    -0.03      0.14    -0.30     0.25 1.00
muResp4Sem_NumPhones          -0.16      0.17    -0.49     0.17 1.00
muResp4Sem_logfqTot            0.02      0.16    -0.29     0.34 1.00
muResp4Sem_Phon_N              0.10      0.09    -0.07     0.28 1.00
muResp4Sem_Phon_N_MFreq       -0.19      0.21    -0.60     0.20 1.00
muResp5NonWord_NumPhones      -0.14      0.09    -0.32     0.03 1.00
muResp5NonWord_logfqTot       -0.02      0.09    -0.20     0.16 1.00
muResp5NonWord_Phon_N         -0.02      0.05    -0.12     0.07 1.00
muResp5NonWord_Phon_N_MFreq   -0.03      0.11    -0.24     0.19 1.00
muResp6LexicalOthers_NumPhones -0.24     0.15    -0.52     0.06 1.00
muResp6LexicalOthers_logfqTot  0.19      0.15    -0.09     0.48 1.00
muResp6LexicalOthers_Phon_N    0.02      0.07    -0.12     0.16 1.00
muResp6LexicalOthers_Phon_N_MFreq -0.01  0.17    -0.34     0.31 1.00
```