Default and exceptional stress processing in Spanish as a testing ground for generative versus exemplar-based phonology models

evidence from ERPs

Karolina Broś, Martin Meyer & Volker Dellwo

University of Warsaw & University of Zurich

January 16th, 2019

Approaches to lexical storage

Approaches to lexical storage

1. Generativist models

- only unpredictable information that cannot be derived by rules is stored in the UR
- non-contrastive data and phonetic detail redundant for the processing of a given word are excluded
- by extension, predictable stress markers are excluded from the lexicon

Approaches to lexical storage

2. Usage-based models

- the theory of exemplars (Bybee, 2001, 2006): focus on the effects of frequency and other external factors on sound production and perception
- abandons fully abstract, phonemic representations of words or morphemes
- gradient, lexically diffuse differences in pronunciation are all stored in the mental lexicon as they are
- by extension, stress cannot be a derived or abstract category it is a bundle of acoustic and auditory features stored with each word represented in the exemplar cloud

Aim of the experiment

Put the two approaches to the test

Focus of the experiment

Spanish

Focus of the experiment

Spanish

- a language with variable stress

Focus of the experiment

Spanish

- a language with variable stress
- prevalence of one stress pattern over the others: partial stress predictability

Focus of the experiment

Spanish

- a language with variable stress
- prevalence of one stress pattern over the others: partial stress predictability
- over 64% (78.9%) of all Spanish words are stressed on the penultimate syllable (Morales-Front 2014, Quilis 1981)

Focus of the experiment

Spanish

- a language with variable stress
- prevalence of one stress pattern over the others: partial stress predictability
- over 64% (78.9%) of all Spanish words are stressed on the penultimate syllable (Morales-Front 2014, Quilis 1981)
- antepenults constitute merely 8% (or 2.76%): exceptional

Focus of the experiment

Spanish

- a language with variable stress
- prevalence of one stress pattern over the others: partial stress predictability
- over 64% (78.9%) of all Spanish words are stressed on the penultimate syllable (Morales-Front 2014, Quilis 1981)
- antepenults constitute merely 8% (or 2.76%): exceptional
- so: default penult pattern derivable by rules, with lexical exceptions

Spanish: assumptions

Spanish: assumptions

- minimal pairs, sensitivity to stress cues in perception

Spanish: assumptions

- minimal pairs, sensitivity to stress cues in perception
- Spanish people not 'stress-deaf' (Peperkamp et al. 2010)

Spanish: assumptions

- minimal pairs, sensitivity to stress cues in perception
- Spanish people not 'stress-deaf' (Peperkamp et al. 2010)

But: Is the default penultimate stress pattern processed differently than the exceptional antepenult?

Spanish: assumptions

- minimal pairs, sensitivity to stress cues in perception
- Spanish people not 'stress-deaf' (Peperkamp et al. 2010)

But: Is the default penultimate stress pattern processed differently than the exceptional antepenult?

Is the exceptional stress stored to facilitate word retrieval, as opposed to the default?

How to test this?

How to test this?

1. **Access to prelexical processing**

How to test this?

1. **Access to prelexical processing**

2. **Access to semantic activation**
(linking phonology with meaning)

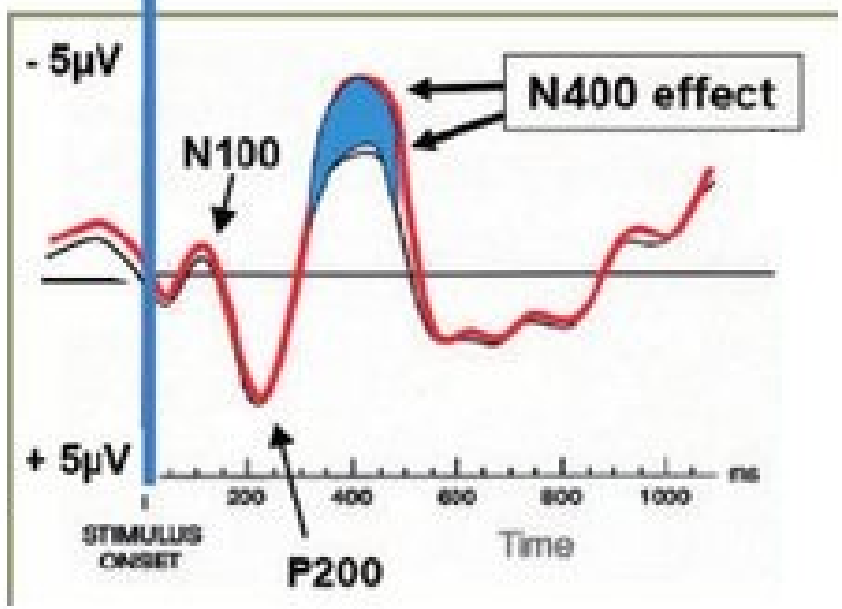
How to test this?

1. **Access to prelexical processing**

2. **Access to semantic activation**
(linking phonology with meaning)

3. **A paradigm evoking the N400 negativity effect**

John ate broccoli at dinner.
John ate democracy at dinner.



Stress perception studies using EEG

Knaus et al. (2007). The processing of word stress: EEG studies on task-related components.

Domahs et al. (2012). Stress 'deafness' in a language with fixed word stress: an ERP study on Polish.

Domahs et al. (2013). Processing (un)predictable word stress: ERP evidence from Turkish.

Molczanow et al. (2013). The lexical representation of word stress in Russian: Evidence from event-related potentials.

This experiment

This experiment

- 32 native speakers of Spanish (19 females) aged 19-32

This experiment

- 32 native speakers of Spanish (19 females) aged 19-32
- 240 stimuli containing correctly and incorrectly stressed words

This experiment

- 32 native speakers of Spanish (19 females) aged 19-32
- 240 stimuli containing correctly and incorrectly stressed words
- 60 penults and 60 antepenults with a CV.CV.CV structure

This experiment

- 32 native speakers of Spanish (19 females) aged 19-32
- 240 stimuli containing correctly and incorrectly stressed words
- 60 penults and 60 antepenults with a CV.CV.CV structure
- invariable carrier sentence

This experiment

- 32 native speakers of Spanish (19 females) aged 19-32
- 240 stimuli containing correctly and incorrectly stressed words
- 60 penults and 60 antepenults with a CV.CV.CV structure
- invariable carrier sentence
- words of matching frequencies

This experiment

- 32 native speakers of Spanish (19 females) aged 19-32
- 240 stimuli containing correctly and incorrectly stressed words
- 60 penults and 60 antepenults with a CV.CV.CV structure
- invariable carrier sentence
- words of matching frequencies
- controlled for phonological neighbourhood

Selection criteria

a) Proper names were excluded.

Selection criteria

a) Proper names were excluded.

b) Words were chosen so as not to become real Spanish words (lexical competitors) after the stress shift (i.e. after changing the stressed syllable).

Selection criteria

a) Proper names were excluded.

b) Words were chosen so as not to become real Spanish words (lexical competitors) after the stress shift (i.e. after changing the stressed syllable).

c) Words that have 10 or more phonological neighbours were excluded.

Selection criteria

a) Proper names were excluded.

b) Words were chosen so as not to become real Spanish words (lexical competitors) after the stress shift (i.e. after changing the stressed syllable).

c) Words that have 10 or more phonological neighbours were excluded.

d) Words which have a phonological neighbour of a higher frequency were excluded.

Selection criteria

a) Proper names were excluded.

b) Words were chosen so as not to become real Spanish words (lexical competitors) after the stress shift (i.e. after changing the stressed syllable).

c) Words that have 10 or more phonological neighbours were excluded.

d) Words which have a phonological neighbour of a higher frequency were excluded.

e) Words which have a phonological neighbour with the other stress pattern under investigation were excluded.

Stimuli

4 conditions:

seMAna (PUs – standard)

PAjaro (APUs – standard)

SEmana (PUd – deviant)

paJAro (APUd – deviant)

Stimuli

Pedro pronunció la palabra [target word] *otra vez*

Pablo pronunció la palabra [target word] *otra vez*

Dani pronunció la palabra [target word] *otra vez*

Lupe pronunció la palabra [target word] *otra vez*

Marta pronunció la palabra [target word] *otra vez*

Laura pronunció la palabra [target word] *otra vez*

Sonia pronunció la palabra [target word] *otra vez*

Hypotheses

- Incorrect stress will invoke a more robust negativity around 400 ms from the onset of the stimulus

Hypotheses

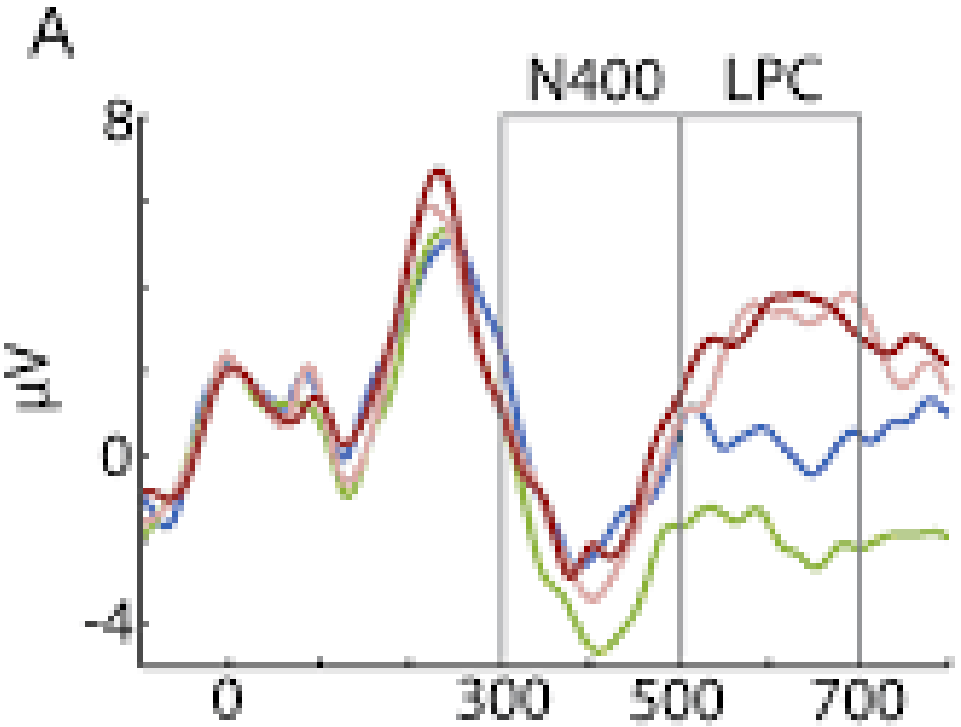
- Incorrect stress will invoke a more robust negativity around 400 ms from the onset of the stimulus
- A significant difference between the two stress patterns

Hypotheses

- Incorrect stress will invoke a more robust negativity around 400 ms from the onset of the stimulus
- A significant difference between the two stress patterns
- Possibly, task-related positivity (LPC)

Hypotheses

- Incorrect stress will invoke a more robust negativity around 400 ms from the onset of the stimulus
- A significant difference between the two stress patterns
- Possibly, task-related positivity (LPC)


A
N400
LPC
8
0
-4
μV
0
300
500
700

Hypotheses

- Explanation: N400 – response to a semantic violation

Hypotheses

- Explanation: N400 – response to a semantic violation
- If stress information is derived in online processing => no problems with procesing incorrect stress

Hypotheses

- Explanation: N400 – response to a semantic violation
- If stress information is derived in online processing =>
  no problems with procesing incorrect stress
- If stress information is stored =>
  mismatch between the memorised and the perceived word

Hypotheses

- Explanation: N400 – response to a semantic violation
- If stress information is derived in online processing =>
  no problems with procesing incorrect stress
- If stress information is stored =>
  mismatch between the memorised and the perceived word
- Conclusion:
  no difference in processing changes to penults and antepenults
  supports exemplar models

Hypotheses

- Explanation: N400 – response to a semantic violation
- If stress information is derived in online processing =>
  no problems with procesing incorrect stress
- If stress information is stored =>
  mismatch between the memorised and the perceived word
- Conclusion:
  no difference in processing changes to penults and antepenults
  supports exemplar models

  difficulty with antepenults but not penults supports
  the generative view

Results: accuracy scores

Results: accuracy scores

– threshold was 75% (ensure comprehension, SNR)

Results: accuracy scores

– threshold was 75% (ensure comprehension, SNR)
– 30 participants had an average of 9 misses in the experiment (230.8 correct answers out of 240)

Results: accuracy scores

– threshold was 75% (ensure comprehension, SNR)
– 30 participants had an average of 9 misses in the experiment (230.8 correct answers out of 240)
– correct PU and APU trials: 100-88.3% accuracy

Results: accuracy scores

– threshold was 75% (ensure comprehension, SNR)
– 30 participants had an average of 9 misses in the experiment (230.8 correct answers out of 240)
– correct PU and APU trials: 100-88.3% accuracy
– APUd caused problems (100-68.3%)

Results: accuracy scores

– threshold was 75% (ensure comprehension, SNR)
– 30 participants had an average of 9 misses in the experiment (230.8 correct answers out of 240)
– correct PU and APU trials: 100-88.3% accuracy
– APUd caused problems (100-68.3%)
– PUd were the second most difficult (100-81.6%)

Results: accuracy scores

– threshold was 75% (ensure comprehension, SNR)
– 30 participants had an average of 9 misses in the experiment (230.8 correct answers out of 240)
– correct PU and APU trials: 100-88.3% accuracy
– APUd caused problems (100-68.3%)
– PUd were the second most difficult (100-81.6%)
– statistics: significant effect of condition ($p = 0.0235$) but not stress pattern
– significant interaction ($p = 0.0108$)

Results: accuracy scores

– threshold was 75% (ensure comprehension, SNR)
– 30 participants had an average of 9 misses in the experiment (230.8 correct answers out of 240)
– correct PU and APU trials: 100-88.3% accuracy
– APUd caused problems (100-68.3%)
– PUd were the second most difficult (100-81.6%)
– statistics: significant effect of condition ($p = 0.0235$) but not stress pattern
– significant interaction ($p = 0.0108$)
– Bonferroni-corrected: significant difference between APUd and both APUs and PUs ($p = 0.002055$, $p = 0.000894$)

Results: accuracy scores

– threshold was 75% (ensure comprehension, SNR)
– 30 participants had an average of 9 misses in the experiment (230.8 correct answers out of 240)
– correct PU and APU trials: 100-88.3% accuracy
– APUd caused problems (100-68.3%)
– PUd were the second most difficult (100-81.6%)
– statistics: significant effect of condition ($p = 0.0235$) but not stress pattern
– significant interaction ($p = 0.0108$)
– Bonferroni-corrected: significant difference between APUd and both APUs and PUs ($p = 0.002055$, $p = 0.000894$)

APUd condition is especially difficult and caused most errors in stress correctness detection

Results: RTs

Results: RTs

– Mean RTs: 504 ms for APUs, 636 ms for APUd, 514 ms for PUs and 559 ms for PUd

Results: RTs

– Mean RTs: 504 ms for APUs, 636 ms for APUd, 514 ms for PUs and 559 ms for PUd
– difference in RTs (between standard and deviant) much greater in the case of the exceptional APU (132 ms) than in the case of the default PU (45 ms)

Results: RTs

– Mean RTs: 504 ms for APUs, 636 ms for APUd, 514 ms for PUs and 559 ms for PUd
– difference in RTs (between standard and deviant) much greater in the case of the exceptional APU (132 ms) than in the case of the default PU (45 ms)
– significant effect of condition ($F(3,78) = 4.415$, $p = 0.0064$)

Results: RTs

– Mean RTs: 504 ms for APUs, 636 ms for APUd, 514 ms for PUs and 559 ms for PUd
– difference in RTs (between standard and deviant) much greater in the case of the exceptional APU (132 ms) than in the case of the default PU (45 ms)
– significant effect of condition ($F(3,78) = 4.415$, $p = 0.0064$)
– Bonferroni-corrected: significant effect in APUd compared to APUs ($p = 0.0066$) and PUs ($p = 0.0155$)

Results: RTs

– Mean RTs: 504 ms for APUs, 636 ms for APUd, 514 ms for PUs and 559 ms for PUd
– difference in RTs (between standard and deviant) much greater in the case of the exceptional APU (132 ms) than in the case of the default PU (45 ms)
– significant effect of condition ($F(3,78) = 4.415$, $p = 0.0064$)
– Bonferroni-corrected: significant effect in APUd compared to APUs ($p = 0.0066$) and PUs ($p = 0.0155$)

Evidence for a significant difference in responses to deviants depending on the stress pattern

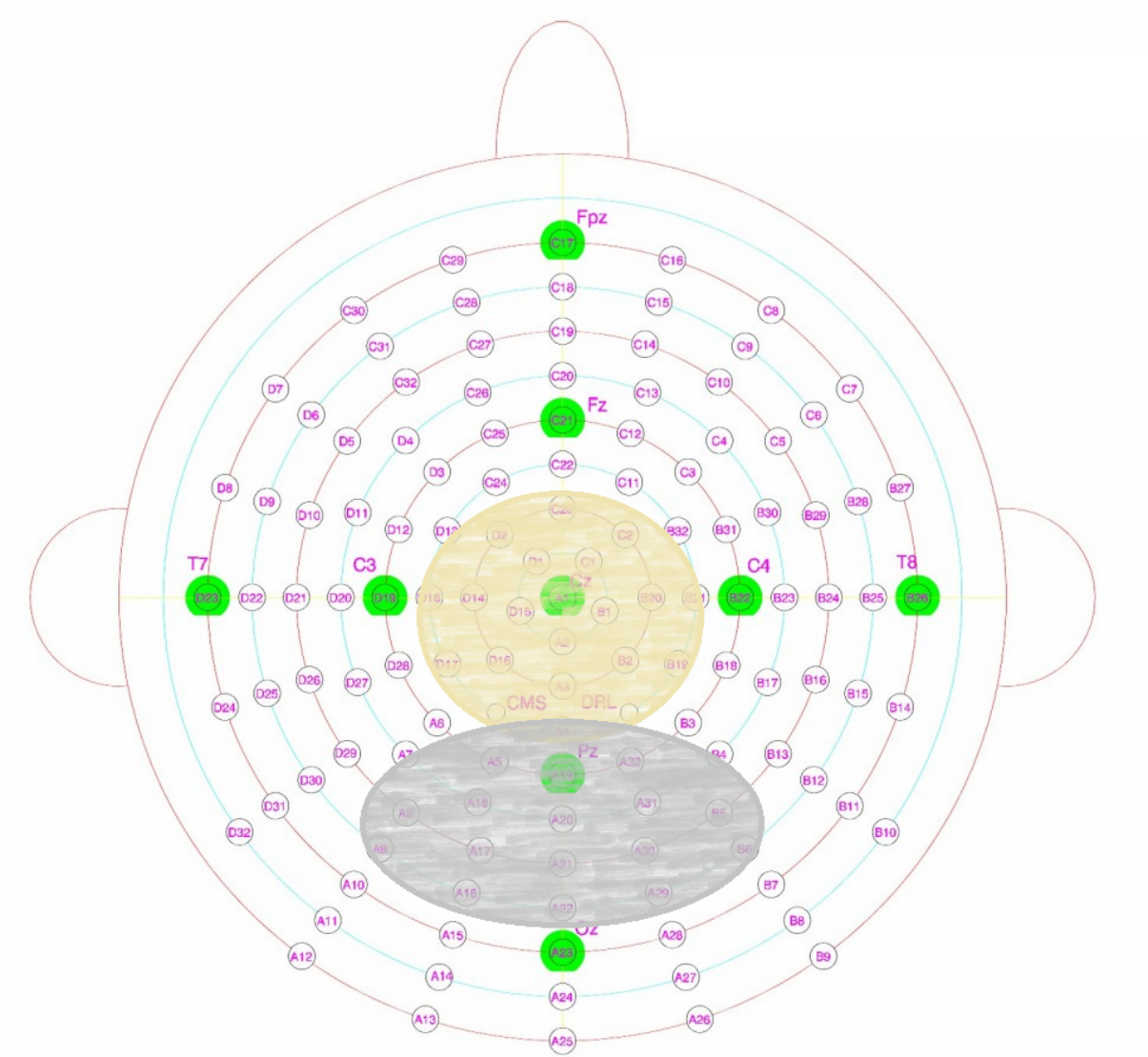
Results: RTs

– Mean RTs: 504 ms for APUs, 636 ms for APUd, 514 ms for PUs and 559 ms for PUd
– difference in RTs (between standard and deviant) much greater in the case of the exceptional APU (132 ms) than in the case of the default PU (45 ms)
– significant effect of condition ($F(3,78) = 4.415$, $p = 0.0064$)
– Bonferroni-corrected: significant effect in APUd compared to APUs ($p = 0.0066$) and PUs ($p = 0.0155$)

Evidence for a significant difference in responses to deviants depending on the stress pattern

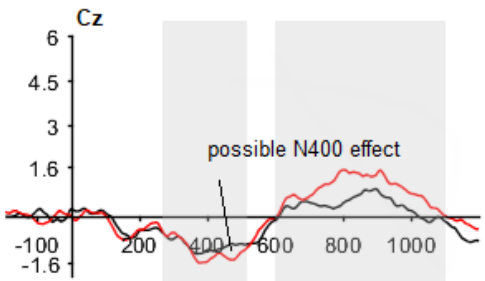
RT results match those of accuracy scores

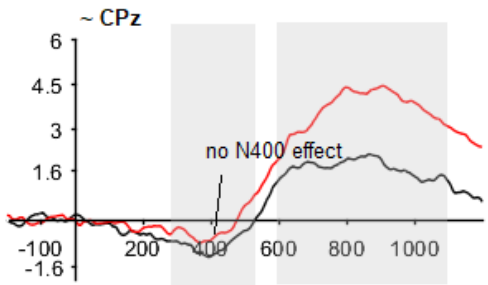
EEG results: Regions of interest (ROIs)

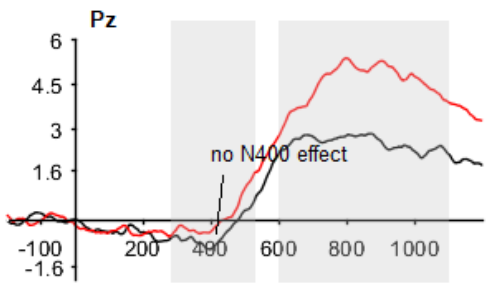


EEG results: Grand averages

EEG results: Grand averages





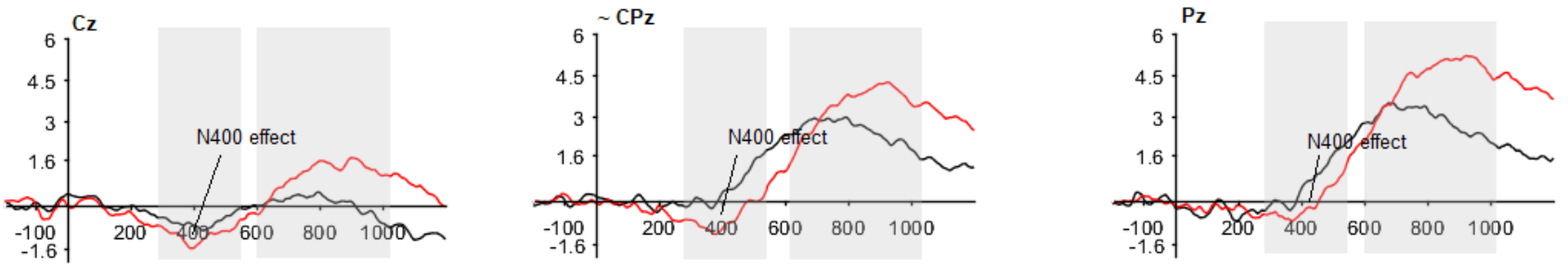


correctly stressed penults followed by correct response

incorrectly stressed penults followed by correct response

ANOVA results: no N400 effect was confirmed for the penults ($F(1,26) = 1.562$, $p = 0.222$). The hypothetical effect in the Cz electrode region was not confirmed statistically. In other regions, an opposite effect is seen instead: incorrect stress causes a less negative inflection in the 350-600 ms windows than correct stress (cf. antepenults).

EEG results: Grand averages



correctly stressed antepenults followed by correct response

incorrectly stressed antepenults followed by correct response

ANOVA results: main effect of condition (correct/incorrect) for the antepenults in the range of 350-600 ms from word onset ($F(1,26) = 20.38$, $p < 0.001$)

Statistical analysis

– Repeated measures ANOVAs

Statistical analysis

– Repeated measures ANOVAs
– Factors: stress (PU, APU), condition (standard, deviant), time window (TW, 350-600 ms, 600-950 ms) and region (frontal, central, parietal).

Statistical analysis

– Repeated measures ANOVAs
– Factors: stress (PU, APU), condition (standard, deviant), time window (TW, 350-600 ms, 600-950 ms) and region (frontal, central, parietal).
– No significant effect of stress nor condition in frontal electrodes

Statistical analysis

APU condition

main effect of condition ($F(1,26) = 20.38$, $p < 0.001$)
main effect of region ($F(1,26) = 30.36$, $p < 0.001$)
no interaction ($F(1,26) = 0.68$, $p = 0.417$)

Statistical analysis

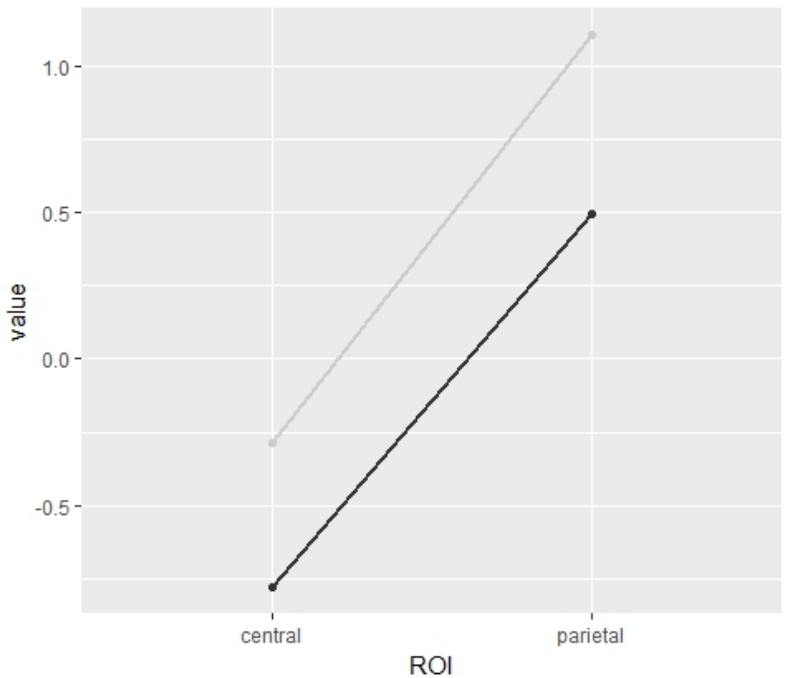
APU condition

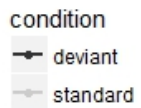
main effect of condition ($F(1,26) = 20.38$, $p < 0.001$)
main effect of region ($F(1,26) = 30.36$, $p < 0.001$)
no interaction ($F(1,26) = 0.68$, $p = 0.417$)

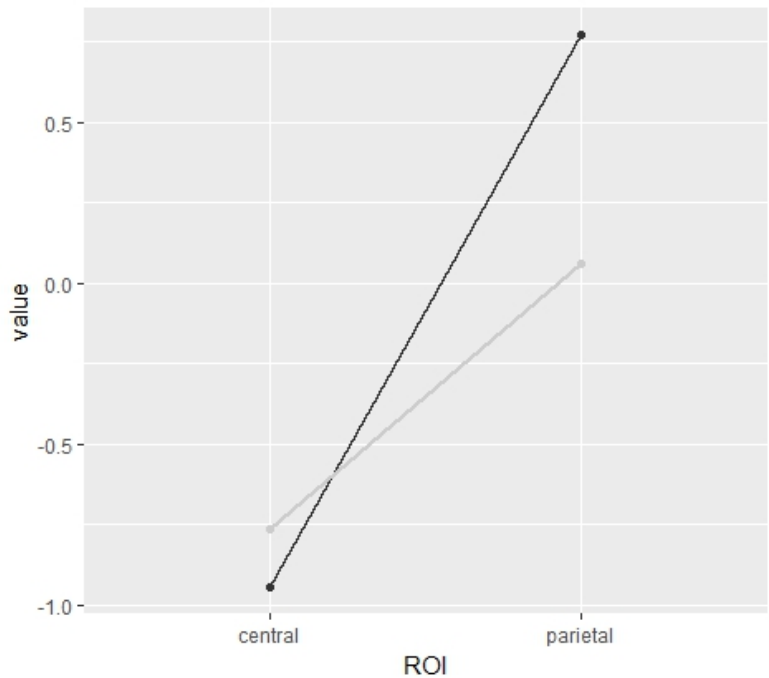
PU condition

no N400 effect ($F(1,26) = 1.562$, $p = 0.222$)
main effect of region ($F(1,26) = 23.63$, $p < 0.001$)
reverse interaction ($F(1,26) = 23.56$, $p < 0.001$)

Interaction plots







EEG results: Interpretation

PU words pronounced with the stress on the first syllable are **not in violation** of any expectation about the prominence of the second syllable

EEG results: Interpretation

PU words pronounced with the stress on the first syllable are **not in violation** of any expectation about the prominence of the second syllable

Hypothesis confirmed:

EEG results: Interpretation

PU words pronounced with the stress on the first syllable are **not in violation** of any expectation about the prominence of the second syllable

Hypothesis confirmed:

1. **Significant difference between penults and antepenults**

EEG results: Interpretation

PU words pronounced with the stress on the first syllable are **not in violation** of any expectation about the prominence of the second syllable

Hypothesis confirmed:

1. Significant difference between penults and antepenults

2. N400 effect only in the case of changes to the exceptional pattern

Results

General ANOVA on the centroparietal data:

– main effect of stress ($F(1,26) = 13.9$, $p < 0.001$)

– interaction between stress and cond ($F(1,26) = 12.88$, $p = 0.001$)

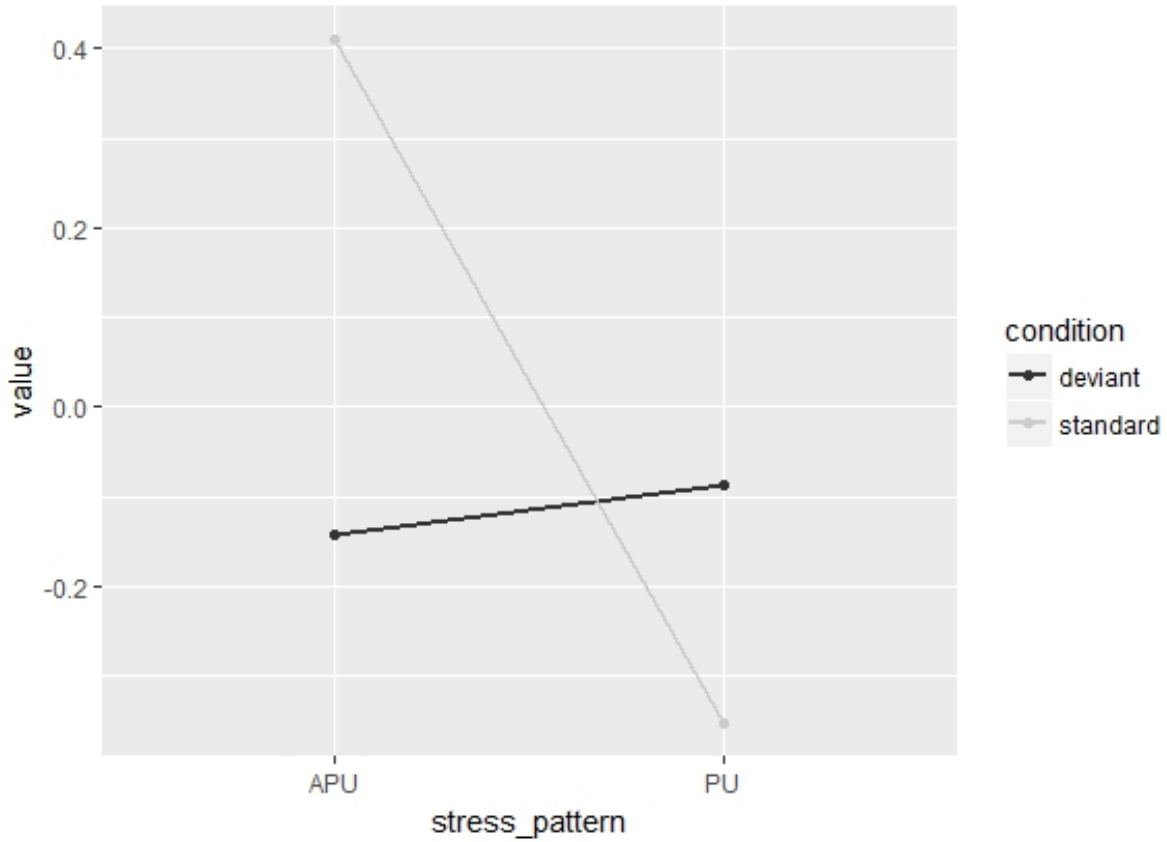
– **no main effect of condition** ($F(1,26) = 1.192$, $p = 0.285$)

Results

General ANOVA on the centroparietal data:
– main effect of stress ($F(1,26) = 13.9$, $p < 0.001$)
– interaction between stress and cond ($F(1,26) = 12.88$, $p = 0.001$)
– **no main effect of condition** ($F(1,26) = 1.192$, $p = 0.285$)

- **stress matters in the standard condition only**
- 'levelling' of the negativity effect between the two stresses

Interaction plot: N400 time window



The 'two-syllable' time window

The 'two-syllable' time window

- the stressed syllable is not reliably longer or of higher pitch

The 'two-syllable' time window

- the stressed syllable is not reliably longer or of higher pitch

APU words:

– mean F0 of the stressed antepenult is 222.9-224.0 Hz
– unstressed syllables have 267.5-264.2 Hz values
– **pitch is quite high at the beginning and steadily rising**
– duration: 187-196 ms in the stressed syllable
– falls to 151-153 ms in the unstressed one

The 'two-syllable' time window

- the stressed syllable is not reliably longer or of higher pitch

APU words:
– mean F0 of the stressed antepenult is 222.9-224.0 Hz
– unstressed syllables have 267.5-264.2 Hz values
– **pitch is quite high at the beginning and steadily rising**
– duration: 187-196 ms in the stressed syllable
– falls to 151-153 ms in the unstressed one

PU words:
– the second syllable is equally long or shorter than the first (182-193 ms vs. 190-200 ms)
– pitch is rising from 180 Hz to 200 Hz, never as high as in APUs
– **the rise is much greater in APU words (40 Hz)**

The 'two-syllable' time window

- **Upon hearing the first syllable Spanish speakers cannot determine whether this syllable is stressed**

The 'two-syllable' time window

- **Upon hearing the first syllable Spanish speakers cannot determine whether this syllable is stressed**

- Human perceptual system cannot distinguish differences in pitch below 3 semitones (12Hz/220Hz, Nooteboom 1997)

The 'two-syllable' time window

- **Upon hearing the first syllable Spanish speakers cannot determine whether this syllable is stressed**

- Human perceptual system cannot distinguish differences in pitch below 3 semitones (12Hz/220Hz, Nooteboom 1997)
- Confirmed by our data: no latency difference in electrophysiological response

LPC results

- significant effect of condition ($F(1,26) = 23.05$, $p < 0.001$)
- no effect of stress ($F(1,26) = 0.125$, $p = 0.726$)
- interaction between the two ($F(1,26) = 4.721$, $p = 0.039$)

LPC results

– significant effect of condition ($F(1,26) = 23.05$, $p < 0.001$)
– no effect of stress ($F(1,26) = 0.125$, $p = 0.726$)
– interaction between the two ($F(1,26) = 4.721$, $p = 0.039$)

- **correctness judgment occurs at this stage**

LPC results

– significant effect of condition ($F(1,26) = 23.05$, $p < 0.001$)
– no effect of stress ($F(1,26) = 0.125$, $p = 0.726$)
– interaction between the two ($F(1,26) = 4.721$, $p = 0.039$)

- **correctness judgment occurs at this stage**
- **prevalence of stress effects in the first TW points to the processing of prosody (stress pattern)**

LPC results

– significant effect of condition ($F(1,26) = 23.05$, $p < 0.001$)
– no effect of stress ($F(1,26) = 0.125$, $p = 0.726$)
– interaction between the two ($F(1,26) = 4.721$, $p = 0.039$)

- **correctness judgment occurs at this stage**
- **prevalence of stress effects in the first TW points to the processing of prosody (stress pattern)**
- **later on the hearer has to decide whether what (s)he heard was correct or incorrect: phonological-semantic integration must have taken place**

Conclusions

– **Penults** behave as true defaults whose underlying abstract representations are not indexed with stress information

Conclusions

– **Penults** behave as true defaults whose underlying abstract representations are not indexed with stress information
– the stress is inferred (or computed) from grammatical rules concerning default stress assignment

Conclusions

– **Penults** behave as true defaults whose underlying abstract representations are not indexed with stress information
– the stress is inferred (or computed) from grammatical rules concerning default stress assignment
– **Antepenults** must be stored together with the information concerning the syllable that is stressed

Conclusions

– **Penults** behave as true defaults whose underlying abstract representations are not indexed with stress information
– the stress is inferred (or computed) from grammatical rules concerning default stress assignment
– **Antepenults** must be stored together with the information concerning the syllable that is stressed
– deviation from this lexical stress is costly for the hearer

Conclusions

– **Penults** behave as true defaults whose underlying abstract representations are not indexed with stress information
– the stress is inferred (or computed) from grammatical rules concerning default stress assignment
– **Antepenults** must be stored together with the information concerning the syllable that is stressed
– deviation from this lexical stress is costly for the hearer

The data **support the generative phonology framework** which assumes that only unpredictable information is stored in the mental lexicon

Discussion

Grammatical operations, which translate acoustic detail and auditory cues into abstract features and constituents, cannot be limited to mere statistical inference

Discussion

Grammatical operations, which translate acoustic detail and auditory cues into abstract features and constituents, cannot be limited to mere statistical inference

Stress should be conceived of as an abstract category and disentangled from both segmental phonetic information and semantics

Discussion

Grammatical operations, which translate acoustic detail and auditory cues into abstract features and constituents, cannot be limited to mere statistical inference

Stress should be conceived of as an abstract category and disentangled from both segmental phonetic information and semantics

– Hearers respond to stress separately from the meaning of the word

Discussion

Grammatical operations, which translate acoustic detail and auditory cues into abstract features and constituents, cannot be limited to mere statistical inference

Stress should be conceived of as an abstract category and disentangled from both segmental phonetic information and semantics

– Hearers respond to stress separately from the meaning of the word
– Bottom-up speech perception approach (Norris et al. 2000)

Discussion

Grammatical operations, which translate acoustic detail and auditory cues into abstract features and constituents, cannot be limited to mere statistical inference

Stress should be conceived of as an abstract category and disentangled from both segmental phonetic information and semantics

– Hearers respond to stress separately from the meaning of the word
– Bottom-up speech perception approach (Norris et al. 2000)
– Top-down wrap-up, integration of prosody and semantics

Discussion

Grammatical operations, which translate acoustic detail and auditory cues into abstract features and constituents, cannot be limited to mere statistical inference

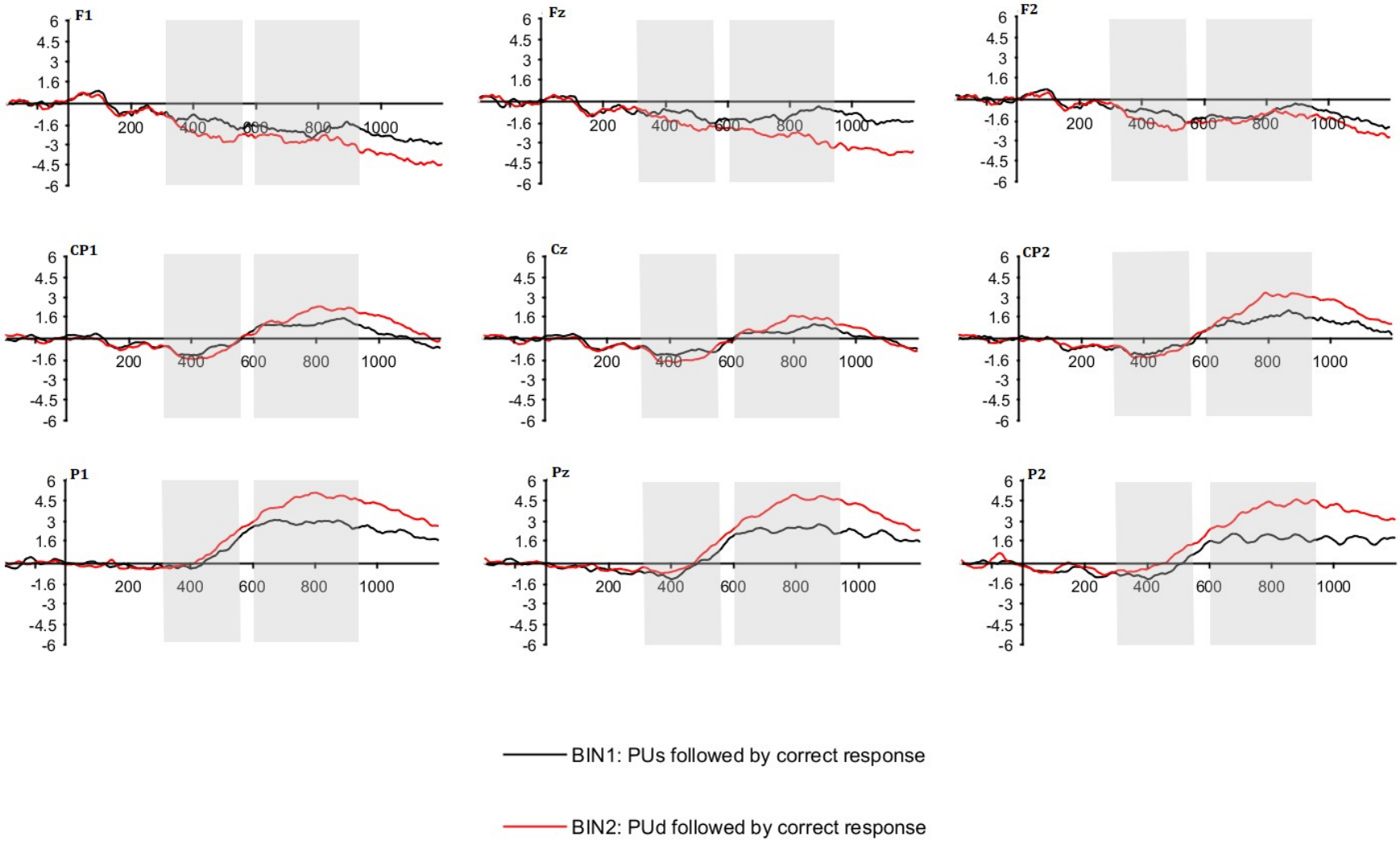
Stress should be conceived of as an abstract category and disentangled from both segmental phonetic information and semantics

– Hearers respond to stress separately from the meaning of the word
– Bottom-up speech perception approach (Norris et al. 2000)
– Top-down wrap-up, integration of prosody and semantics
– compatible with Poeppel et al.'s (2008) speech perception theory

Thank You!

Slides available at: *www.karolinabros.eu*

EEG results: Grand averages



EEG results: Grand averages