Units of phonological encoding revisited: evidence from asymmetrical clitics and lexicalized phrases

Marina Vigário[1], Sónia Frota[1], Joseph Butler[1] & Linda Wheeldon[2]

[1]University of Lisbon, [2]University of Birmingham

TIE 5 – Fifth European Conference on Tone and Intonation
University of Oxford, 6-8 September 2012

Partially funded by the project PTDC/CLE-LIN108722/2008

Background

A proper model of speech production requires the identification of the **basic unit of phonological encoding** (Levelt 1989; Wheeldon & Lahiri 1997, 2002; Levelt, Roelofs & Meyer 1999)

Unit of phonological encoding

The phonological form retrieved for production, after stored lexical representations are activated and assigned to positions in the syntactic representation of the utterance, before articulation.

Background

Sternberg et al. (1978) – production experiments with prepared lists of words varying in number and type of words > **latency of production [= time taken to start to speak] increases linearly with list length**

What is the relevant unit of retrieval and articulation?

→ Sternberg et al. (1978), Wheeldon & Lahiri (1997):

. **not stored lexical representations** – the same patterns are found for non-words

. **not syllables** – the *slope* of the latency function is similar for lists of matched bisyllabic words as for monosyllabic words (but: a small but significant diff. – see below)

(e.g. baby-rumble-market≈bay-rum-mark)

. **not syntactic words** – unstressed words do not affect the slope of the latency function

(e.g. bay-and-rum-and-mark ≈ bay-rum-mark)

→ Stenberg et al. (1978): the unit is the ***stress group***

[but the prosodic level was not controlled for: in lists, each PW is a phonological phrase and potentially an intonational phrase]

→ Levelt (1989): the unit of phonological encoding is the ***phonological word***

Background

- Wheeldon and colleagues conducted a series of experiments testing this claim, using the prepared speech paradigm (Sternberg et al. 1978, 1980) – Wheeldon & Lahiri 1997, 2002, 2011; Vogel & Wheeldon 2010)

→ Testing the effect of the number of *units* planned to be produced on the time taken to start to speak (production latency)

> variable types of *words*

> keeping constant

. number of syllables

. number of lexical items

. syntactic structure

Simple words	– DU, En, It
Complex words	– It
FuncWrds: (en)clitics	– DU, En
FuncWrds: PWs	– DU
Word+word compounds	– DU, It
2 content words phrases	– DU, En, It

DU=Dutch; En=English, It=Italian

Background

Wheeldon & Lahiri (1997) – Exp1 (prepared speech paradigm)

Clitic	Nonclitic	Control
2PW	3 PW	2PW
(Ik zoek **het**)ω(water)ω	(Ik zoek)ω(**vers**)ω (water)ω	(Ik zoek)ω(water)ω
I seek the water	*I seek fresh water*	*I seek water*

same number of sylls.; similar syntactic strucure; diff. number of PWs

1. **Visual display**:

het water	**vers** water	water
the water	*fresh water*	*water*

2. **Auditory question**:

Wat zoek je?	Wat zoek je?	Wat zoek je?
what do you seek?	*what do you seek?*	*what do you seek?*

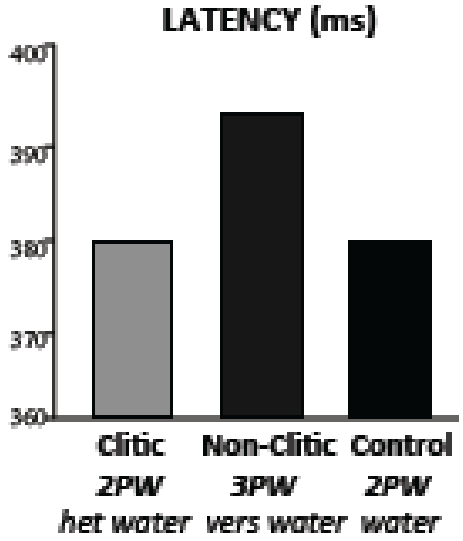
*3. **3 beeps** (variable time lags)*

4. **Prepared response** (as soon as the 3rd beep is heard)

(Ik zoek **het**)ω(water)ω	(Ik zoek)ω(**vers** water)ω	(Ik zoek)ω(water)ω
I seek the water	*I seek fresh water*	*I seek water*

Background

Wheeldon and colleagues found that production latencies are a function of the number of PW in the sentence to be produced



from Wheeldon & Lahiri (2011)

- ✓ latency of non-clitic condition (3PWs) significantly longer than clitic condition and controls (2PWs)
- ✓compounds also patterned like a single unit (DU, It)

→ W&L (2002) assumed that compounds display PW recursion and the **relevant unit is the recursive** PW

→ V&W (2010) assume that compounds form one Composite Group and the **relevant unit is the CG**

Background

- Sternberg et al. 1978 – **number of syllables** of the words to be produced did not affect production latency except for a small effect attributed to the size of the **first word**

- Resyllabification – the phonological restructurings that may occur across lexical boundaries (i.e., resyllabification) may suggest that the unit for articulation is larger than the syllable – the PW is the domain of resyllabification in Germanic languages, and this is compatible with the finding that the PW is the unit of articulation (Levelt 1989; Wheeldon & Lahiri 1997; Levelt, Roelofs & Meyer 1999)

Our goal

1. Replicate Wheeldon and colleagues experiments in another language – so far, only 3 languages tested; only Dutch for clitics **and** compounds

   European Portuguese (EP) – specific phonological properties, different from the languages previously examined

   strong cues for the prosodic word and the prosodic word group; large domain for resyllabification (Intonational Phrase); mixed rhythmic properties

2. Additional conditions: two types of compounds and two types of phonological clitics

EP prosodic phonology in a nutshell

PW – Prosodic Word (Vigário 1999, 2003)

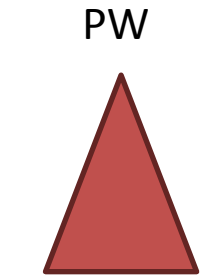
Evidence: word stress; segmental processes; phonotactic constraints; pitch accent distribution; emphatic stress; PW deletion processes

Units: **simple words;** most derived words; some prefixes and very few suffixes; internal members of word compounds; **host+enclitic**; **proclitic+host** (recursive PW)

e → ∅ /[… _]$_{PW}$

/bEb**e**/$_{PW}$ ‘drink’
/s**e**/$_{CL}$ ‘if’
/t**e**/$_{CL}$ ‘you-dat’

(bEb**e**)$_{PW}$	‘(he) drinks’
(bEb**e**)$_{PW}$ (Água)$_{PW}$	‘(he) drinks water’
(bEb[**j**]-o)$_{PW}$	‘drink-it’
(ouço-t**e**)$_{PW}$	‘(I) hear-you’
(s[**j**] (ouço)$_{PW}$)$_{PW}$	‘if (I) hear’

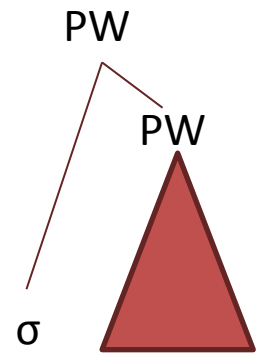


castanhas

‘chestnut’

contei-te

‘(I) told you’



com raiva

‘in rage’

EP prosodic phonology in a nutshell

PWG – Prosodic Word Group (Vigário 2003, 2010)

Evidence: rhythmic constraints on V deletion rules across PWs; focus accent distribution

Units: complex words with more than one PWG; morphosyntactic **compounds;** letter and letter-number combinations; highly frequent word combinations

Final V1 deletion blocking when V2 bears PWG stress ($[...V1]_{PW}['V2...]_{PW}]_{PWG}$)

PWG

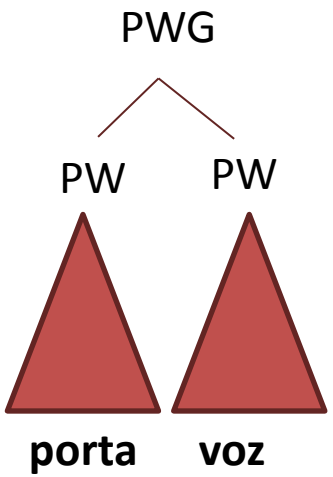
(o pOrt***a***-Óculos castAnho)ϕ

'the brown glassesbox'

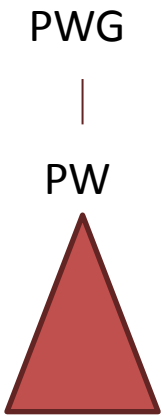
PWG PWG

(a quINt*a* Ordem dAda)ϕ

'the fifth order given'



'spokesman'



'chestnut'

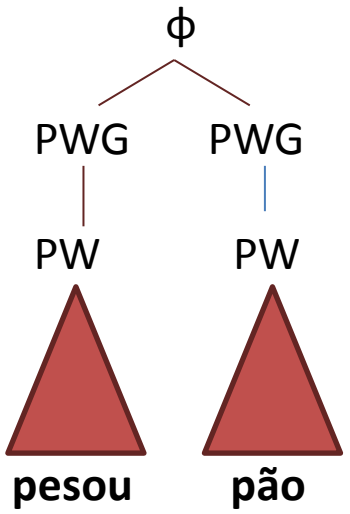
EP prosodic phonology in a nutshell

ϕ – Phonological Phrase (Frota 2000, 2009)

Evidence: stress clash resolution; rhythmic constraints on vowel sandhi, pitch accent distribution

Units: other **combinations of words; lexicalized phrases** (or phrasal compounds) usually like other word combinations within ϕ (not PWG)

BUT

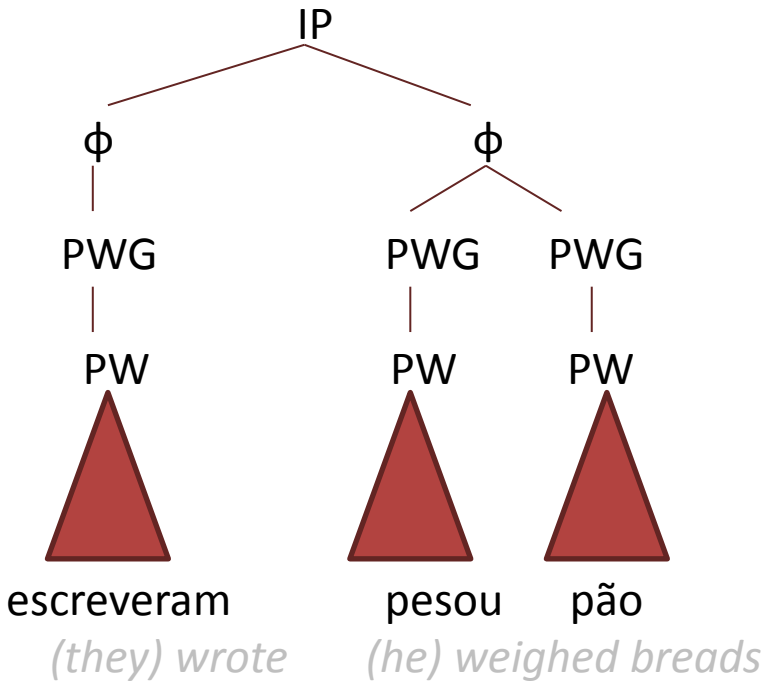


Lexicalized phrases :
variable degrees of lexicalization
- ✓semantic
- ~ syntactic (e.g. internal agr.)
- ~ phonological →

. 2 PWG	*verde-água*	lit. green-(like)water
. 2 PW	*grand[j]-área*	lit. big-area 'penalty area'
1 PWG	**lexicalized as a single PWG – idiosyncratic**	
. 1 PW	*pernalta*	lit. leg-long (class of birds)
1 PWG	**lexicalized as a single PW – idiosyncratic**	

EP prosodic phonology in a nutshell

IP – Intonational Phrase (Frota 2000, 2009)
Evidence: several segmental processes; **resyllabification**; domain for minimal tune; pre-boundary lenthening and pauses; tendency for unreduced forms of clitics at left edge
Units: adjacent ϕ within a root sentence; ϕs in a string not structurally attached to the sentence tree



Resyllabification – across the board within IP; blocked across IP

(a.**s** a.luna**s**)I (até onde sabemo**s**)I (obtiveram boa.**s** a.valiações)I
[z] [S] [S] [z] [S]
'the students, as far as we know, got good grades'

EP experiments – the materials

Two experiments

Exp1	CONDITIONS	Exp2
1. Simple words ((**c**astanhas)PW)PWG		1. Simple words ((**b**orboleta)PW)PWG
2. Compounds ((**p**orta)PW-(voz)PW)PWG		2. Compounds ((**b**ate)PW-(chapas)PW)PWG
3. Proclitics+host (((**c**om (raiva)PW)PW)PWG		3. Lexicalized phrases ((**b**olo)PW)PWG((podre)PW)PWG
4. Host+enclitics ((**c**ontei-te)PW)PWG		
5. Phrases with two lexical words ((**p**esou)PW)PWG((pão)PW)PWG		4. Phrases with two lexical words ((**b**ebe)PW)PWG ((sumos)PW)PWG
↑ carrier=4 syl + **test items=3 syl**		↑ carrier=4 syl + **test items=4 syl**

- Number and internal structure of syllables kept constant in each experiment
- Stimuli for comparison all starting in the same segment/major class, voicing
- Words and combinations of words matched for frequency – not high frequency →checked on a corpus of half a million words; but all well-known items of EP
- Some (inevitable) variation in syntactic structure

Predictions

Exp1

1. Simple words
   ((castanhas)PW)PWG
2. Compounds
   ((porta)PW-(voz)PW)PWG
3. Proclitics+host
   (((com (raiva)PW)PW)PWG
4. Host+enclitics
   ((contei-te)PW)PWG
5. Phrases with two lexical words
   ((pesou)PW)PWG((pão)PW)PWG

Exp2

1. Simple words
   ((borboleta)PW)PWG
2. Compounds
   ((bate)PW-(chapas)PW)PWG
3. Lexicalized phrases
   ((bolo)PW)PWG((podre)PW)PWG

4. Phrases with two lexical words
   ((bebe)PW)PWG ((sumos)PW)PWG

Prod latency is a function of number of	**Exp1**	**Exp2**
PW	1≈4 < 2≈5 (3 intermediate)	1 < 2≈3≈4
PWG	1≈2≈3≈4 < 5	1≈2 (3?)< 4
Morphosyntactic words	1≈2 < 3≈4≈5	1≈2 < 3≈4

Method

- Prepared speech paradigm (Wheeldon and colleagues)
- Speakers were asked to prepare the answer to the question “O que é que eles escreveram”? (‘what did they write?’) using “escreveram X” (‘(they)wrote X’), where X should correspond to the word(s) displayed on the screen. They were asked to respond as soon as possible after they heard 3 beeps:

  1st beep: 2 s after words display; 2nd beep 1 s;

  3rd beep: variable – 800ms, 1200ms, 1400ms
- Both experiments were performed by the same speaker in the same day, with an interval between them (order exps counterbalanced)
- A voice-key registered voice onset during sentence production; E-prime 2.0 was used; the whole experiment was sound recorded
- The stimuli were presented in 6 blocks, pseudo-randomized; block order was counterbalanced; each block was followed by a short pause

Method

- The experiments were preceded by a set of 6 practice trials
- For both experiments, 12 sets of comparable items were created for testing each condition, repeated three times

  →Exp1=180 trials (12x5x3); Exp2=144 trials (12x4x3)
- Subjects: native speakers of Standard EP, untrained, university educated, mean age 24 years
- The recorded responses were coded for errors, disfluencies, wrong prosodic patterns, etc.
  - Some speakers systematically used a non-neutral/unnatural prosodic pattern – they were excluded from the analysis
  - Only the data from speakers with natural neutral patterns were analysed

  → Exp1: 20 participants; Exp2: 18 participants

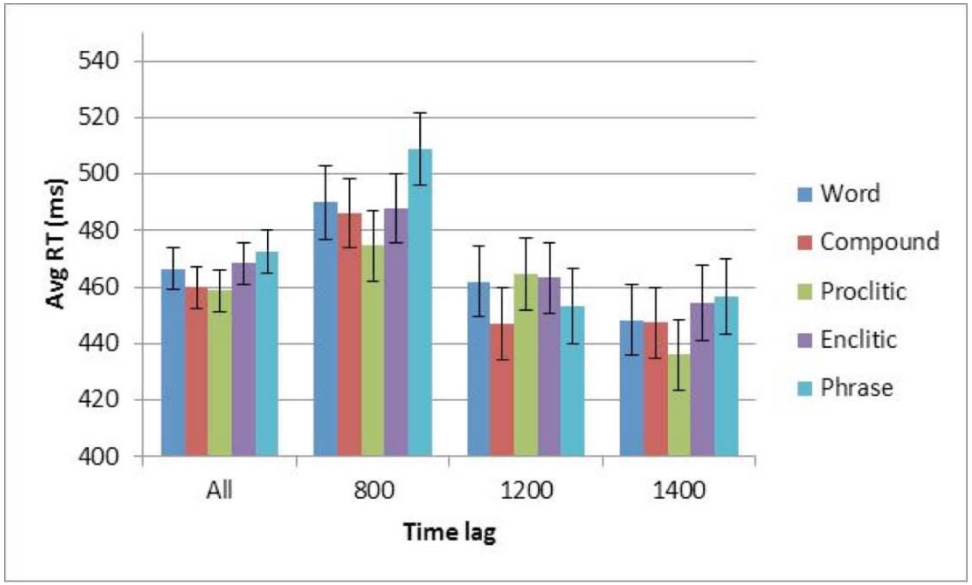
Results

- **Experiment 1**

Predictions

Morphosyntactic word:
word, compound <
proclitic, enclitic, phrase
PW: word, enclitic, proclitic
< compound, phrase
PWG: all (1) < phrase (2)

The data is most consistent with the **PWG** being the critical unit that determines latencies. The effect is most clear at lag 800 (but no interaction between lag and condition). Subjects are faster in the longer lags.



Repeated measures ANOVA: variables **condition** (W, C, Pro, E, Phr) and **time lag** (800, 1200, 1400)

A borderline effect of condition ($F1(4,76) = 2.19$, $p = .08$, $\eta^2 = .1$), with **PHRASES longer than the other conditions**. Pairwise comparisons revealed a significant difference between conditions Compound and Phrase, (mean diff = 9.96, $p < .05$), and Proclitic and Phrase (mean diff = 13.45, $p < .05$)

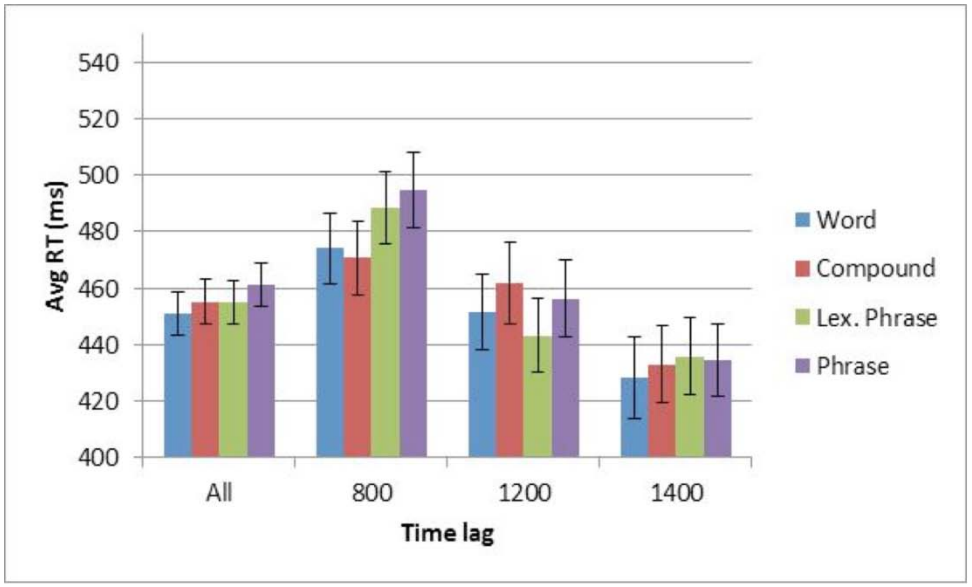
Results

- **Experiment 2**

Predictions

Morphosyntactic word:
word, compound <
lex. phrase, phrase
PW: word < all
PWG: word, compound <
phrase (lex. phrase?)

The data is most consistent with the **PWG** being the critical unit that determines latencies.



Repeated measures ANOVA: variables **condition** (W, C, Lex. P, Phr) and **time lag** (800, 1200, 1400)

A significant effect of condition ($F1(3,51) = 3.17$, $p < .05$, $\eta^2 = .16$), again with **PHRASES being longer**. Like in Experiment 1, the effect is most clearly seen at lag 800 (but no interaction between lag and condition). Subjects are faster in the longer lags.

Results

Experiments 1 and 2

- Phrase condition (2 PWGs) shows longer latencies.
- This pattern holds across all testing despite the variation found in the data.
- Overall the data is consistent with production latencies being a function of nº of PWG (and not PW or morphosyntactic words)
- These results provide initial evidence in favour of the PWG hypothesis.
- Why is the effect not stronger (like in Dutch)?

Results

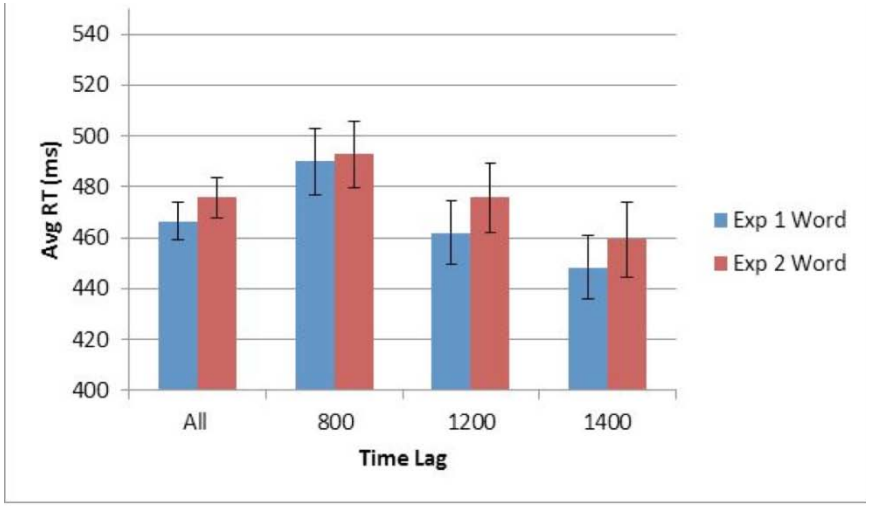
- **Number of syllables:** Comparing Exp1 and Exp2
  - The domain of resyllabification in EP is the IP. Could number of syllables of the whole IP matter in EP?
  - 3 conditions are identical in Exp1 and Exp2, but Exp1 one syllable less: Word, Compound, Phrase
  - Data from 14 subjects (those who performed well in *both* experiments)
  - Predictions:
    - If the whole sentence (IP) is a relevant domain, Exp1<Exp2 (Exp1, one syllable less)
    - If the relevant domain is the initial unit , Exp1≈ Exp2 (4 syllables)

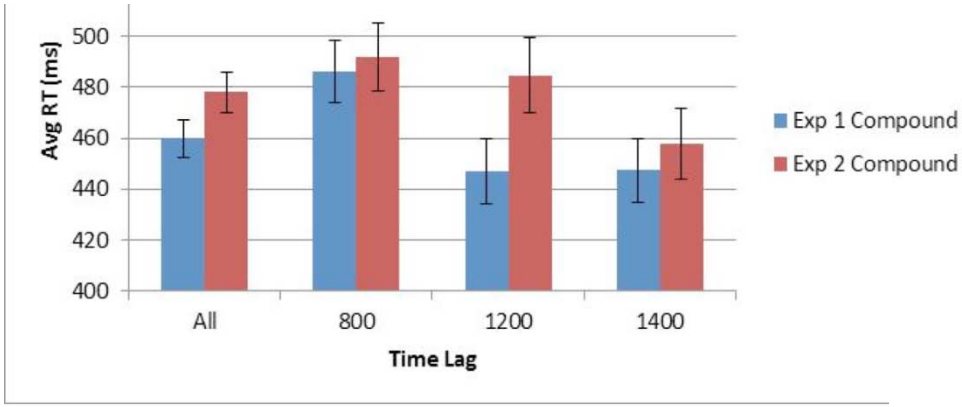
Results

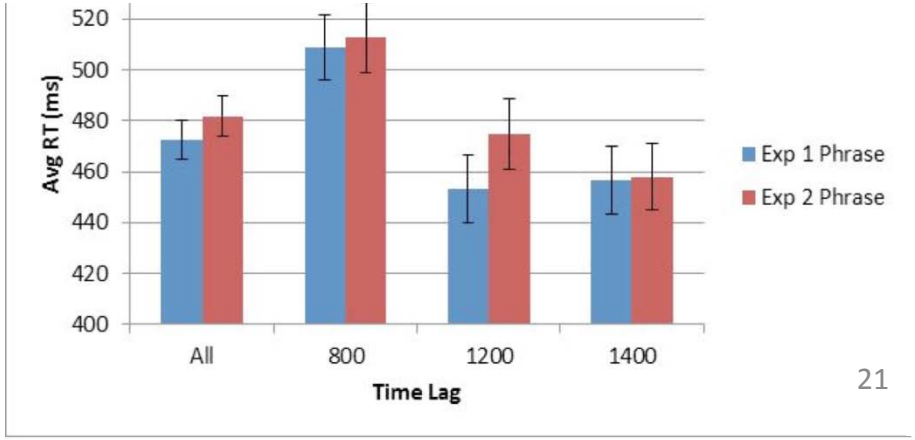
- Number of syllables

Consistent trend Exp1<Exp2
Results reach significance for Word and Compound ($p < .05$)

Results suggest that the whole sentence (IP) is a relevant domain for the nº of syllables effect in EP







NB. For Words and Compounds, the IP has 2 PWG; for Phrases it has 3 PWG

Discussion

- Results from both Exp1 and Exp2 are consistent with production latencies being a function of noº of PWG, providing initial evidence in favour of the PWG as the unit of phonological encoding in EP
- But the effect seems to be weaker in EP (and the variation in the data is high)…
- Methodological reasons?
  - More stimulus variation in our experiments than the W&L ones, making the task more complex
- Language specific properties?
  - Combination of word-based phonology with phrase-based phonology: IP is the domain for resyllabification; Mixed rhythm: combination of stress-timed and syllable timed

Discussion

- Effect of nº of syllables over the whole sentence (IP), suggests that the IP also plays a role: if domain of resyllabification is relevant for planning, when subjects have enough time (longer time lags) they may encode at the level of the syllable for the entire IP
- Future research
  - Testing the effect of higher level prosodic units in a prepared speech task (by varying the number of prosodic phrases, holding number of syllables constant)
  - Testing the syllable size effect on sequences with two IPs (by varying the nº of syllables in the first IP only, and in the second IP only)
- Initial results are promising and challenging, and suggest possible language specific effects. Much further research is needed!

Thank you!

Obrigada!

We are grateful to the research assistant Cátia Severino, who run most of the experiments

Susana Correia and Marisa Cruz for their collaboration in different phases of this research

And to our subjects...