

Background

Questions

Two measures have been shown to correlate with response times and neural activity:

Transition Probabilities between morphemes, indicating that the visual language processing system keeps track of local statistical relations of subparts of words. Thus far, this has mainly been studied with single derivational affixes. [1–15]

Relative Entropy between whole words, indicating that the visual language processing system keeps track of global comparisons between whole words with other whole words. [16–24]

Despite a large literature finding support for both measures, no study has thus far directly compared the two.

Methods

Stimuli

- 49 participants
- Native Spanish speakers
- Recruitment via Prolific

WUGará

+

bailabas

+

non-word

word

- Visual Lexical Decision Task
- 43 mono-syllabic verbal stems
- 43 mon-syllabic pseudo stems
- 13 suffixes

2 Suffixes							
	<u>stem</u>	<u>suffix 1</u>			<u>suffix 2</u>		
<u>Imperfect</u>	bail	+	aba	}	<u>sg.</u>	<u>pl.</u>	
	com	+	ía		1st	ø	mos
	part				2nd	s	n
				3rd	ø	n	
<u>Future</u>	bail	+	ar	}	1st	é	emos
	com	+	er		2nd	ás	án
	part	+	ir		3rd	á	án

3 Suffixes									
	<u>stem</u>		<u>suffix 1</u>		<u>suffix 2</u>			<u>suffix 3</u>	
							<u>sg.</u>	<u>pl.</u>	
<u>Future</u>	bail	+	ar	}	+ ía	+	1st	ø	mos
	com	+	er				2nd	s	n
	part	+	ir				3rd	ø	n

Control

- Trial Number
- Average bigram frequency of stimulus
- Log frequency of stimulus
- Experimental session

Control Model		
Variable	Estimate	P-val.
Intercept	1.109e+02	0.192
Trial Number	-8.978e-02	<2e-16 ***
Avg. Bigram Freq.	3.205e-07	0.587
Word Log. Freq.	-3.822e+01	<2e-16 ***
Session	4.884e+02	<2e-16 ***
AIC	304844.2	

Control Model		
Variable	Estimate	P-val.
Intercept	1.843e+02	0.044 *
Trial Number	-9.148e-02	<2e-16 ***
Avg. Bigram Freq.	1.253e-06	0.183
Word Log. Freq.	-4.648e+01	2.68e-15 ***
Session	4.767e+02	<2e-16 ***
AIC	132826	

Transition Probability

Calculation:

$$\text{Transition Probability (M}_1 \rightarrow \text{M}_2) = P(\text{M}_2 | \text{M}_1) = \frac{P(\text{M}_1 + \text{M}_2)}{P(\text{M}_1)}$$

↑ Transition Probability

Single Transition Probability Model:

Full Transition Probability Model:

↓ Response Time

√bail

aba

s

√bail

aba

s

Single Transition Probability Model		
Variable	Estimate	P-val.
Intercept	1.018e+02	0.23061
Trial Number	-8.980e-02	<2e-16 ***
Avg. Bigram Freq.	1.001e-07	0.86554
Word Log. Freq.	-7.087e+01	2.33e-09 ***
Session	4.884e+02	<2e-16 ***
TP	1.836e+02	0.00376 **
AIC	304827.7	

Full Transition Probability Model		
Variable	Estimate	P-val.
Intercept	-3.330e+01	0.7470
Trial Number	-8.983e-02	<2e-16 ***
Avg. Bigram Freq.	2.764e-07	0.6458
Word Log. Freq.	-4.573e+01	<2e-16 ***
Session	4.884e+02	<2e-16 ***
TP 1	5.077e+01	0.1640
TP 2	1.492e+02	0.0058 **
AIC	304822	

Single Transition Probability Model		
Variable	Estimate	P-val.
Intercept	1.795e+02	0.04962 *
Trial Number	-9.152e-02	<2e-16 ***
Avg. Bigram Freq.	9.988e-07	0.29377
Word Log. Freq.	-9.077e+01	0.00174 **
Session	4.767e+02	<2e-16 ***
TP	2.417e+02	0.11625
AIC	132813.6	

Full Transition Probability Model		
Variable	Estimate	P-val.
Intercept	8.310e+02	0.000219 ***
Trial Number	-9.116e-02	<2e-16 ***
Avg. Bigram Freq.	2.858e-06	0.005601 **
Word Log. Freq.	-4.620e+01	1.58e-11 ***
Session	4.767e+02	<2e-16 ***
TP 1	-2.824e+02	0.016869 *
TP 2	-5.680e+02	0.001975 **
TP 3	7.733e+01	0.288961
AIC	132781.3	

Relative Entropy

Calculation:

$$\text{Relative Entropy (stem)} = -\sum P(\text{stem+suffix}) \log_2 P(\text{stem+suffix})$$

↑ RE

Information Content

Processing Cost

↑ RT

Relative Entropy Model:

√bail

aba

s

High Entropy Verb (casar)

Class Average

Low Entropy Verb (bailar)

Relative Entropy Model		
Variable	Estimate	P-val.
Intercept	1.095e+02	0.198
Trial Number	-8.978e-02	<2e-16 ***
Avg. Bigram Freq.	3.058e-07	0.607
Word Log. Freq.	-3.817e+01	<2e-16 ***
Session	4.884e+02	<2e-16 ***
RE	2.480e+00	0.814
AIC	304839.6	

Relative Entropy Model		
Variable	Estimate	P-val.
Intercept	2.062e+02	0.0255 *
Trial Number	-9.148e-02	<2e-16 ***
Avg. Bigram Freq.	1.681e-06	0.0759
Word Log. Freq.	-4.728e+01	5.64e-16 ***
Session	4.767e+02	<2e-16 ***
RE	-4.075e+01	0.0165 *
AIC	132814.7	

AIC Model Comparison			
Full TP Model	Single TP Model	RE Model	Control Model
Single TP Model	RE Model	Control Model	
RE Model	Control Model		
Control Model			

AIC Model Comparison			
Full TP Model	Single TP Model	RE Model	Control Model
Single TP Model	RE Model	Control Model	
RE Model	Control Model		
Control Model			

Conclusions

Citations

Transition Probability (i.e. local morpheme to morpheme statistical relations) better predicts response time data than Relative Entropy (i.e. global word to word statistical relations). Models that incorporate multiple decomposition of affixes better predict response time data than those that treat an entire suffix cluster as a single suffix.

This supports theoretic and processing models that treat morphemes as obligatorily distinct units compared to models that treat whole words as the minimal unit of representation.

[1–15] Zweig & Pykkönen 2009; Vartiainen et al. 2009; Solomyak & Marantz 2009; Solomyak & Marantz 2010; Lehtonen et al. 2011; Lewis et al. 2011; Fruchter & Marantz 2015; Neophytou et al. 2018; Williams & Marantz 2018; Ohta et al. 2019; Wilder et al. 2019; Stockall et al. 2019; Oseki et al. 2020; Wray et al. 2022; Cayado et al. (in press); [16–24] Kostić et al. 2003; Baayen & Moscoso del Prado 2005; Milin et al. 2009a; Milin et al. 2009b; Nenadić et al. 2016; Hendrix et al. 2017; Filipović–Đurđević & Gatarić 2018; Filipović–Đurđević & Milin 2019; Nenadić et al. 2021;