Entropy's dilemma: Facilitatory and Inhibitory Effects of Information Entropy during Language Comprehension

Hossein Karimi (hokarimi@csuf.edu)[1], Jaden Zinn[2], Haomiao Li[2], Andrew Lei[1], Chidchanok Paircharoen[1]

[1]Department of Psychology, California State University, Fullerton, [2]Department of Psychology, Mississippi State University

Introduction: Based on contemporary theories of language processing, prior knowledge is constantly activated during language processing to generate predictions about the upcoming information (Kuperberg & Jaeger, 2016). Past research has primarily focused on measuring lexical predictability of *individual* words (e.g., *banana* in (1)), using measures such as cloze probability (Taylor, 1953), surprisal (Levy, 2008), and entropy reduction (Hale, 2003), which capture lexical predictability *once the current word is processed*. Unlike these measures, information entropy captures the *distribution of potential words* in a context (e.g., *banana*, *mango*, *orange* etc.) *before the current word is encountered*. However, information entropy may capture two distinct underlying processes, with potentially opposite effects. On the one hand, more lexical candidates that fit a sentential context may compete with one another, inhibiting lexical access (Dahan et al., 2001). On the other hand, more lexical candidates may lead to activation of more semantic features (e.g., [+*edible*], [+*fruit*], [+*tropical*]), facilitating lexical access (Brothers et al., 2023; Luke & Christianson, 2016). In two experiments, we investigated whether information entropy can exert both inhibitory and facilitatory effects. We calculated the sum of cosine semantic similarity between target words (e.g., *banana*), and the responses produced by each participant for each item (e.g., *mango*, *orange etc.*), which we dubbed *semantic entropy*. Then, using multiple regression, we examined the effect of information entropy when any effect of semantic entropy was controlled for. Crucially, because the semantic overlap component of information entropy was accounted for, any remaining effect of information entropy would reflect lexical competition, which we dubbed *lexical entropy*. **Method**: 169 native speakers of English completed two separate sessions 14-21 days apart. In the "cloze session", they provided up to 8 completions to 648 sentence fragments with varying contextual constraints as illustrated in (1), along with estimated probability for each word, as shown Table 1 (the target word *banana* and the following words were removed). Trial-level Shannon entropy values were calculated using normalized probabilities (by dividing the given probabilities by their sum). In the "reading session", participants read the same sentences in full while reading times were measured using eye tracking (Exp1; N=58) or self-paced reading (Exp2; N=111). The cloze session always came first so that the participants were not exposed to the target words. **Results & Discussion**: Manipulation checks showed expected effects of cloze (t=-3.30, p<.001) and contextual constraint (t=-4.57, p=.001). For the main analyses, we used mixed-effects regression models to estimate the effects of lexical and semantic entropy on reading times on the target words, and spillover regions (target+1 – target+5). Because semantic similarity and information entropy were moderately correlated (r = .41), we assessed all regression models for potential collinearity using the Variance Inflation Factor (VIF) and condition number (*kappa*). Both measured indicated low collinearity values (<2). Moreover, the coefficients for the single-predictor models retained their original signs. Target word frequency was a covariate in all models. In both experiments, we observed that while semantic entropy facilitated lexical processing, lexical entropy inhibited it, as shown in Table 2 and Figures 1 and 2. These results suggest that information entropy should be decomposed into semantic activation and lexical competition, corresponding to facilitation and inhibition during language comprehension, respectively. The results are most compatible with attractor dynamic models of lexical access that predict simultaneous facilitatory and inhibitory effects during lexical access based on semantic feature overlap and lexical competition (Mirman & Magnuson, 2008). Orthographic similarity had no effect on reading times, suggesting that word forms are not preactivated. EEG data has been collected from 60 participants on the same stimuli. The data will be analyzed to investigate the neural correlates of semantic and lexical entropy.

(1) Sample experimental item:

a) Low constraint: My picky younger brother refused to eat the ***banana*** at breakfast before the others.
b) Medium constraint: Because of fruit flies, he threw away the ***banana*** at breakfast before the others.
c) High constraint: The hungry monkey was trying to peel the ***banana*** at breakfast before the others.

Table 1. Sample responses for(1c)

	Word	Probability
word1	banana	95
word2	orange	60
word3	mango	50
word4	coconut	50
word5	…	…
word6		
word7		
word8		

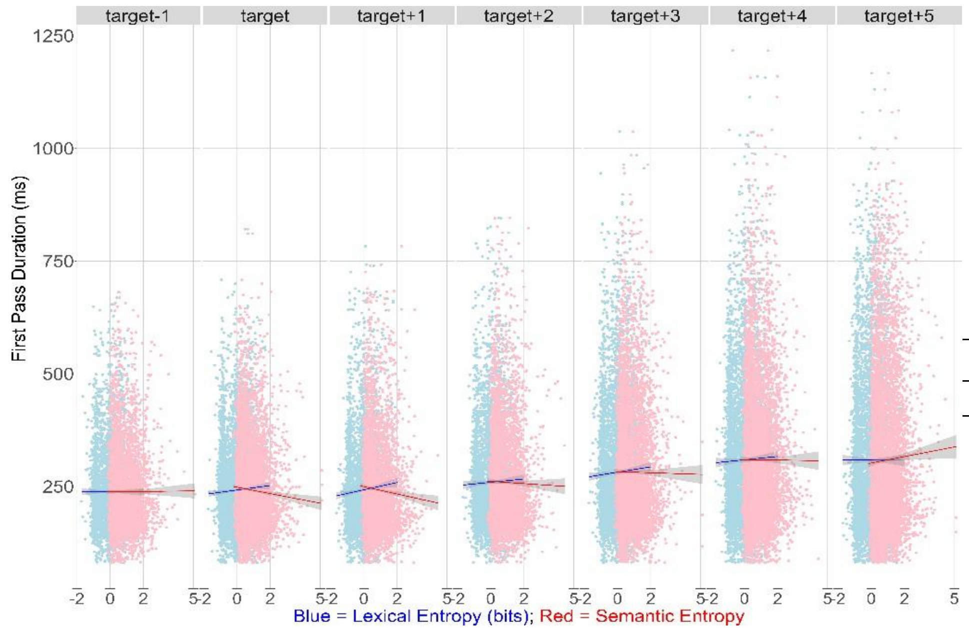

Figure 1. First-pass reading times as a function of Lexical and Semantic Entropy.

Table 2. Results on the target word ("banana")

Experiment	DV	*Predictor*	*t*	*p*
1 (Eye-Tracking)	Skipping Rate	Sem. Entropy	4.29	<.001
		Lex. Entropy	-2.11	.03
		Frequency	4.99	<.001
	First Fixation Duration	Sem. Entropy	-4.78	<.001
		Lex. Entropy	2.09	.04
		Frequency	-.61	.54
	First-Pass Time	Sem. Entropy	-4.04	<.001
		Lex. Entropy	2.58	.01
		Frequency	-2.50	.01
	Total Reading Time	Sem. Entropy	-3.64	<.001
		Lex. Entropy	2.72	.006
		Frequency	-.85	.39
	Regression Path Time	Sem. Entropy	-5.46	<.001
		Lex. Entropy	3.03	.003
		Frequency	-2.16	.03
2 (Self-paced reading)	Reaction time	Sem. Entropy	-1.32	.18
		Lex. Entropy	-1.00	.31
		Frequency	-.40	.68

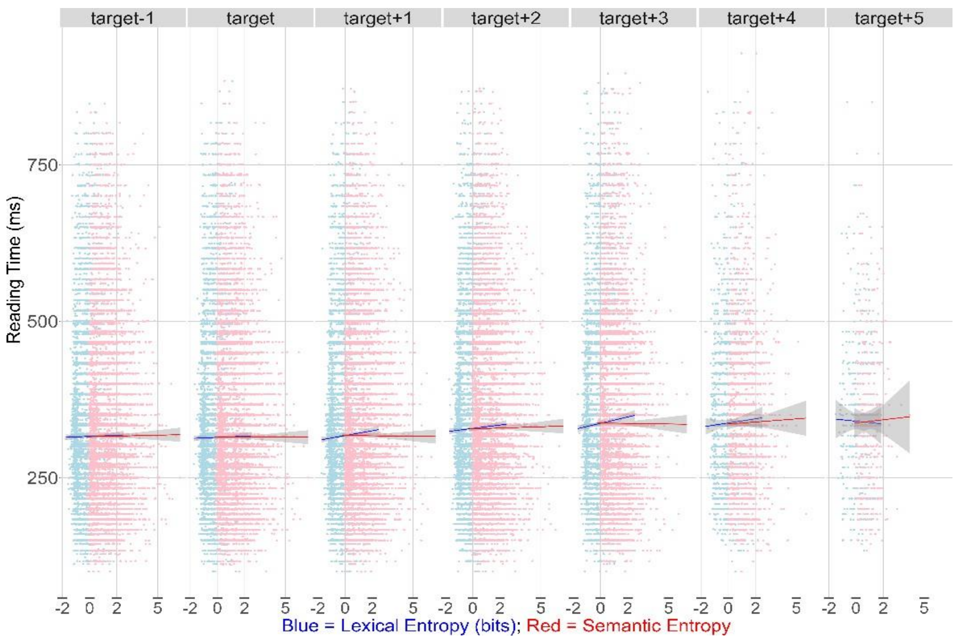

Figure 2. Self-paced reading times as a function of Lexical and Semantic Entropy.

Table 2. Results on the spillover region ("banana at breakfast before the others.")

Experiment	DV	*Predictor*	*t*	*p*
1 (Eye-Tracking)	First-Pass Time	Sem. Entropy	-2.95	.004
		Lex. Entropy	3.20	.002
		Frequency	-.79	.43
	Total Reading Time	Sem. Entropy	-3.98	<.001
		Lex. Entropy	2.28	.02
		Frequency	1.11	.26
	Regression Path Time	Sem. Entropy	-1.50	.13
		Lex. Entropy	.64	.52
		Frequency	1.83	.06
2 (Self-paced reading)	Reaction time	Sem. Entropy	-2.58	.01
		Lex. Entropy	2.12	.03
		Frequency	-4.01	<.001