Memory activation and interference model syntactic locality gradience differently

Marisa Ferrara Boston
Cornell University

AMLaP, University of York
September 6-8, 2010

Introduction

Is syntactic locality gradience attributable to different aspects of memory difficulty?

Memory difficulty can arise from activation, estimated by **string distance**, the Dependency Locality Theory (**DLT**) [3], or **retrieval activation** [9]. It can also arise from interference, estimated by **intervening nominals** or similarity-based interference (**SBI**) [9].

A computational model compares both types of memory factors to two locality phenomena: *wh*-island violations (WHVs) [11] and superiority violations (SUVs) [2].

Gradience in syntactic locality...

In WHVs, a *wh*-filler (who) is fronted across a *wh*-island (whether), leading to unacceptability (1).

(1) *Who did Diego find out whether they read the book?

In SUVs the filler (what) is fronted across a syntactically superior *wh*-phrase (who) (2).

(2) *Diego asked what who read.

Acceptability increases when the wh-filler (WHV1, SUV1, SUV2, SUV3), wh-island (WHV1, WHV2, WHV3), or wh-intervenor (SUV1, SUV2, SUV3) is changed. (Table 1).

Study	Measure	Conditions	Data
WHV1 [7]	RRT, **dismissed**	Who did Albert learn whether they dismissed...	-67ms
		Which employee did Albert learn whether they dismissed...	-78ms
		Who did Albert learn that they dismissed...	-88ms
WHV2 [12]	Acceptability	What do you wonder whether John bought?	-0.73
		What do you think that John bought?	0.38
		What wonders whether John bought a car?	0.71
		What thinks that John bought a car?	1.23
WHV3 [8]	Acceptability	What do you wonder who they caught him at by accident?	16
		What do you wonder if they caught him at by accident?	40
		What do you suppose that they caught him at by accident?	55
SUV1 [1]	RRT, **read**	Pat wondered what who read.	49ms
		Pat wondered what which student read.	33ms
		Pat wondered which book who read.	27ms
		Pat wondered which book which student read.	-5ms
SUV2 [6]	RRT, **signed**	Ashley disclosed what who signed...	49ms
		Ashley disclosed what which diplomat signed...	33ms
		Ashley disclosed which agreement who signed...	27ms
		Ashley disclosed which agreement which diplomat signed...	-4ms
SUV3 [4]	Syntactic judgment	Who persuaded who to visit you?	*
		Who did you persuade who to visit?	✓
		Who did you persuade her to visit?	✓
		Who did you persuade to visit who?	✓

Table 1: Experimental evidence for WHV and SUV gradience.

...is modeled...

A Nivre dependency parser [10] builds non-projective analyses of experimental sentences.

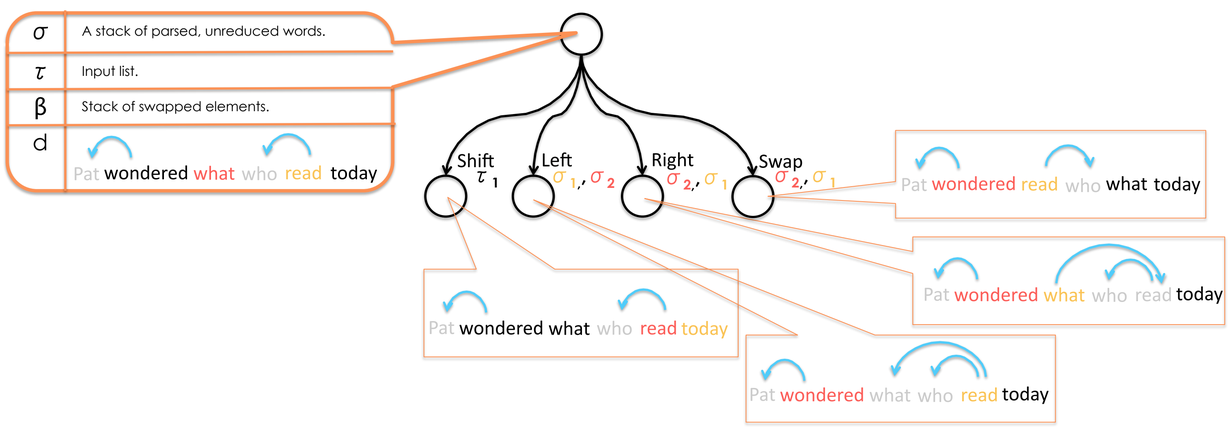

Figure 1: Parser states include σ, τ, β, and d. Shift, Left, Right, and Swap transitions lead to new states.

...probabilistically...

Probabilities based on aspects of the parser's memory state (Table 2) determine parser transitions.

Distance, DLT, and retrieval activation (RetAct) [9] represent activation. The types of Intervenors, SBI [9], and the conjunction of the two (Both) represent interference.

Feature	Feature Type	Includes
Distance	String position	$\sigma_1 - \sigma_2$
DLT	Count	$\text{intervenors}_{nom}(\sigma_2 ... \sigma_1)$
RetAct	Value	$\text{baselineActivation}(\sigma_2)$
Intervenors	Part-of-speech	$\text{intervenors}_{nom}(\sigma_2 ... \sigma_1)$
SBI	Value	$\text{interference}(\sigma_2)$
Both	Value::part-of-speech	$\text{interference}(\sigma_2)$::$\text{intervenors}_{nom}(\sigma_2 ... \sigma_1)$

Table 2: Memory-based probabilistic feature specifications.

Hypothesis: Increased surprisal at the verb indicates increased processing difficulty (⇑) integrating the *wh*-filler and verb across an island or intervenor.

...by activation or interference.

Surprisals from activation-based features pattern with WHV gradience (Table 3). Surprisals from interference-based features pattern with SUV gradience.

Condition	Human Data	Distance	DLT	RetAct	Intervenors	SBI	Both
WHV1-1	-67ms	0.284	1.743	1.853	1.333	1.104	0.700
WHV1-2	-78ms	0.284	1.611	1.853	1.333	1.104	1.095
WHV1-3	-88ms	0.284	1.244	1.853	1.333	1.104	1.095
WHV2-1	-0.73	0.614	1.786	2.391	1.769	1.590	0.978
WHV2-2	0.38	0.581	1.711	2.299	1.595	1.427	0.972
WHV2-3	0.71	0.630	1.129	1.208	1.269	1.270	0.695
WHV2-4	1.23	0.596	1.054	1.141	1.284	1.263	0.715
WHV3-1	16	0.549	1.608	2.274	1.573	1.490	1.144
WHV3-2	40	0.549	1.586	2.274	1.573	1.490	1.144
WHV3-3	55	0.569	1.004	2.387	1.679	1.440	1.165
SUV1-1	49ms	0.689	1.024	0.959	1.922	0.993	2.277
SUV1-2	33ms	0.725	1.588	1.366	1.665	1.284	2.134
SUV1-3	27ms	0.682	1.839	1.241	1.598	1.909	2.134
SUV1-4	-5ms	0.735	1.588	1.311	1.537	1.369	2.706
SUV2-1	49ms	0.872	2.220	1.913	2.467	2.062	2.277
SUV2-2	33ms	1.035	4.148	1.771	3.144	1.909	3.109
SUV2-3	27ms	0.968	2.339	1.784	3.018	1.833	2.134
SUV2-4	-4ms	0.843	5.510	1.644	2.727	1.706	2.706
SUV3-1	*	0.681	1.316	1.122	1.378	1.508	1.515
SUV3-2	✓	0.896	1.520	0.948	0.797	2.300	0.566
SUV3-3	✓	0.681	1.504	1.185	2.063	1.508	0.566
SUV3-4	✓	0.678	1.602	1.152	1.417	1.508	0.791

Table 3: Activation and interference model different aspects of syntactic locality.

Conclusion

The results from the computational model indicate that WHV and SUV gradience are attributable to different aspects of memory.

[1]Arnon, I., Snider, N., Hofmeister, P., Jaeger, T. F., & Sag, I. (To Appear). Cross-linguistic variation in a processing account: The case of multiple wh-questions. In *Proceedings of BLS 32*.
[2]Chomsky, N. (1973). Conditions on transformations. In S. Anderson & P. Kiparsky (Eds.), *A Festschrift for Morris Halle*.
[3]Gibson, E. (2000). Dependency locality theory: A distance-based theory of linguistic complexity. In A. Marantz, Y. Miyashita, & W. OÕNeil (Eds.), *Image, language, brain: Papers from the First Mind Articulation Symposium*.
[4]Haider, H. (2004). The superiority conspiracy: four constraints and a processing effect. In A. Stepanov, G. Fanselow, & R. Vogel (Eds.), *Minimality effects in syntax*.
[5]Hale, J. T. (2001). A probabilistic Earley parser as a psycholinguistic model. In *Proceedings of NAACL 2001*.
[6]Hofmeister, P. (2007). Retrievability and gradience in filler-gap dependencies. In *Proceedings of the 43rd Regional Meeting of the CLS*.
[7]Hofmeister, P., & Sag, I. A. (To Appear). Cognitive constraints and island effects. *Language*.
[8]Kluender, R., & Kutas, M. (1993). Subjacency as a processing phenomenon. *Language and Cognitive Processes* 8.
[9]Lewis, R., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science* 29.
[10]Nivre, J. (2009). Non-projective dependency parsing in expected linear time. *Proceedings of the ACL 2009*.
[11]Ross, J.R. (1967). *Constraints on variables in syntax*. MIT Dissertation.
[12]Sprouse, J., Wagers, M., & Phillips, C. (Submitted). A test of the relation between working memory capacity and syntactic island effects.

mfb74@cornell.edu
http://conf.ling.cornell.edu/Marisa/