

Lexical and Structural Constraints in Sentence Processing

Fatemeh Torabi Asr Spring 2011



+ Incrementality (so far)

Reading_time_effect(){

The readers/listeners has anticipations about what comes next;

if (continuation ~ anticipation)

Faster reading time; //structures already made ③

else

}

Slower reading time; //reanalysis required



+ Overview

Background

- Top-down vs. Bottom-up constraints in parsing
- Different parsing theories

Papers

- Effect of merely local syntactic coherence on sentence processing Tabor, Galantucci and Richardson. 2004
- The interaction of top-down and bottom-up statistics in the resolution of syntactic category ambiguity
 Gibson. 2005

Summary

The lawyer visited that cheep hotel to stay for the night.

6



The lawyer visited that cheep hotel to stay for the night.

7



The lawyer visited ... that cheep hotel ...





Local coherence accounts (LCA):

- Self-organized parsing
- Fixed-width buffers
- Firs-pass failure

Self-consistent parsing accounts (SCPA):

11

Full grammatical parsing

- Local coherence accounts (LCA):
 - Self-organized parsing
 - Fixed-width buffers
 - Firs-pass failure



- Local coherence accounts (LCA):
 - Self-organized parsing
 - Fixed-width buffers
 - Firs-pass failure



Local coherence accounts (LCA):

- Self-organized parsing
- Fixed-width buffers
- Firs-pass failure

The lawyer visited ... that cheep hotel ...

?

The lawyer visited that cheep hotel ...

DET ADJ N

- Self-consistent parsing accounts (SCPA):
 - Full grammatical parsing

- Self-consistent parsing accounts (SCPA):
 - Full grammatical parsing



- The goal is to examine these contrasting parsing accounts:
 - Partial parses are sometimes constructed (LCA)
 - Only parses consistent with the whole input are constructed (SCPA)

• Experiments:

- 1: Syntactic local coherence increases reading time (RT)
- 2: Semantic support for a local parse increases RT even more
- 3: People judge sentences with local coherence as ungrammatical

- Experiment 1: local coherence (syntactic] slows reading?
 - a. The coach smiled at the player tossed a Frisbee by the opposing team.
 - b. The coach smiled at the player who was tossed a Frisbee by the opposing team.





• Experiment 1: Items

-6	-	-5	-4		-3	-2	_	1	
The	c	oach	smiled		at	the	p	layer	
			0	1	2		3		Ambiguous
a.			tossed	a	frisbe	ee	by	(A/R)	
b.	who	was	tossed	а	frisbe	ee	by	(A/U)	Unambiguous
c.			thrown	а	frisbe	ee	by	(U/R)	
d.	who	was	thrown	a	frisbe	ee	by	(U/U)	

- Experiment 1: Procedure
 - Self-paced reading

The coach smiled at the player tossed a Frisbee by the ...

----- ----- ----- tossed ----- -----

• Experiment 1: Results



Fig. 1. Mean residual reading times from Experiment 1. Error bars show one standard error around each data point.

Experiment 1: Verification

- Argument: "throwing a Frisbee to a player" a typical cooccurrence?
- Answer: No significant difference in rating
 - (7) (a) Someone tossed a frisbee to the player. (5.83, SE = 0.14)
 - (b) Someone threw a frisbee to the player. (5.63, SE = 0.13)

- Experiment 1: Verification
 - Argument: "thrown" recipient interpretation more acceptable?
 - Answer: No significant difference between verb types.

- (8) (a) The player was tossed a frisbee by the opposing team. (Recipient/Ambiguous) (4.30, SE = 0.15)
 - (b) A frisbee was tossed to the player by the opposing team. (Theme/Ambiguous) (5.56, SE = 0.10)
 - (c) The player was thrown a frisbee by the opposing team. (Recipient/Unambiguous) (4.39, SE = 0.15)
 - (d) A frisbee was thrown to the player by the opposing team. (Theme/Unambiguous) (5.42, SE = 0.11)

Experiment 1: Verification

- Argument: "thrown" is biased towards recipient-extracted passive sense while "tossed" is biased toward active interpretation.
- Answer: Experiment 2 (let's keep ambiguous verbs the same between cases).

- Experiment 2: local coherence (syntactic + semantic) slows reading?
 - a. The bandit worried about the prisoner transported by the capricious guards.
 - b. The bandit worried about the gold transported by the capricious guards.







• Experiment 2: Items

-2	-1			0	1 Animate
the	prisoner			transported	the 🖊
the	prisoner	who	was	transported	the
the	gold			transported	the
the	gold	that	was	transported	the

Experiment 2: Results



Fig. 2. Mean residual reading times from Experiment 2. Error bars show one standard error around each data point.

Experiment 2: Verification

- Argument: "transporting the gold" a typical co-occurrence?
- Answer: No significant difference in (target) passive.

(11) The prisoner was transported. (Animate/Passive) (6.06, SE = 0.12)The prisoner transported something. (Animate/ Active) (6.03, SE = 0.12) The gold was transported. (Inanimate/Passive) (6.11, SE = 0.12)The gold transported something. (Inanimate/ Active) (2.16, SE = 0.14)

- Experiment 2: Verification
 - Argument: Any semantic interpretation of the locally parsed text?
 - Answer: "yes" answers to the second question increased with Reduction.
 - (12) The bandit worried about the prisoner (who was) transported the whole way.
 - (a) Was the prisoner transported? (prisoner = Theme of transporting)
 - (b) Did the prisoner transport something? (prisoner = Agent of transporting)

Experiment 3: Does grammaticality judgment interact with local coherence in a same way as reading time does?

The coach smiled at the player tossed a Frisbee



- Experiment 3: Items
 - 3.1: The coach smiled at the player (who was) tossed/thrown a Frisbee.
 - 3.2: The bandit talked remorsefully of the prisoner/gold (who/that was) transferred the whole way by the guards.

Experiment 3: Procedure

The coach smiled at the player tossed a Frisbee.
Grammatical? (yes/no)

Experiment 3: Result

Table 4

Mean rates of positive grammaticality judgments for the two subexperiments and the fillers in Experiment 3

(a) Subexperiment 3.1	
Ambiguous/Reduced	0.22 (0.040) 두
Ambiguous/Unreduced	0.72 (0.033)
Unambiguous/Reduced	0.39 (0.042)
Unambiguous/Unreduced	0.72 (0.038)
(b) Subexperiment 3.2	
Animate/Reduced	0.58 (0.045)
Animate/Unreduced	0.82 (0.030)
Inanimate/Reduced	0.75 (0.035)
Inanimate/Unreduced	0.84 (0.032)
(c) Fillers	
Grammatical	0.93 (0.012)
Ungrammatical	0.08 (0.013)

The values shown in parentheses are standard errors.

- Conclusions (with respect to incrementality):
 - Local parsing happens where local coherence exists.
 - Constructing a local parse inconsistent with the global parse suggests that preceding information is not maximally applied in parsing.
 - Ignoring information from previous steps is against incrementality.

Goals:

- Top-down & bottom-up information contribution
- Formulation suggestion (context-independent hypothesis)
- Against context-dependent hypothesis (Tabor et al 1997)
- Defend parallel processing
- Experiments:
 - 1 & 2 :Ambiguous words might increase RT independently from the context
 - 3: People keep track of alternative interpretations until it can be resolved

■ Same word in different environments (from Tabor et al. 1997)

- a. That cheap hotel was clean and comfortable to our surprise.
 - b. That cheap hotels were clean and comfortable surprised us.
- (2) a. The lawyer insisted that cheap hotel was clean and comfortable.
 - b. The lawyer insisted that cheap hotels were clean and comfortable.





- Different words in the same environment (from Tabor et al. 1997)
 - (3) a. The lawyer visited that cheap hotel to stay for the night.
 - b. The lawyer visited those cheap hotels to stay for the night.



The lawyer

visited that cheep hotel



Context-independent category-frequency hypothesis

P("that" being tagged as complementizer in this sentence) =

P("that" being a complementizer) *

Independent info

P(a complementizer following a verb)

Environmental info

- Experiment 1: "that" slows down reading independently from the context?
 - a. The lawyer for that skilled surgeon asked for a raise.
 - b. The lawyer for those skilled surgeons asked for a raise.



• Experiment 1: Items

- (11) a. Prep, that: The lawyer for that skilled surgeon asked for a raise.
 - b. Prep, those: The lawyer for those skilled surgeons asked for a raise.

Tested by Tabor 1997

- c. Verb, that: The lawyer visited that skilled surgeon before the hearings began.d. Verb, those: The lawyer visited those skilled
 - d. Verb, those: The lawyer visited those skilled surgeons before the hearings began.



• Experiment 1: Results



• Experiment 1: Results

- Significant determiner effect
- ✓ No significant category*determiner effect

Table 5	
Analysis of Variance results for Experiment 1	L

Source of variance	$\min F'$			
	df	min F' value		
Position 3 ("for"/"visited")				
Category	1,47	1.46		
Determiner	1,47	<1		
Category × Determiner	1,47	<1		
Position 4 ("that"/"those")				
Category	1,47	4.85*		
Determiner	1,47	1.26		
Category \times Determiner	1,47	<1		
Position 5 ("skilled")				
Category	1,47	4.19*		
Determiner	1,47	3.26		
Category × Determiner	1,47	<1		
Position 6 ("surgeon(s)")				
Category	1,47	<1		
Determiner	1,47	<1		
Category × Determiner	1,47	<1		
Position 7 ("asked"/"before")				
Category	1,47	2.31		
Determiner	1,47	3.77		
Category × Determiner	1,47	2.04		
Positions 8-end of sentence				
Category	1,47	1.89		
Determiner	1,47	<1		
Category × Determiner	1,47	1.80		
Positions 4–6 ("that/those skill				
Category	1,47	3.24		
Determiner	1,47	3.29		
Category × Determiner	1,47	<1		

- Experiment 1: Results
 - ? Small effect of determiner in verbal case
 - ? Verbal "those" condition read slower after critical region



• Experiment 2: Replication of experiment 1 to see...

- Even more significant slower reading time for "that" in verbal versions?
- Any similar pattern when comparing with "this"?

• Experiment 2: Items

- (12) a. Prep, that: The lawyer for that skilled surgeon asked for a raise.
 - b. Prep, those: The lawyer for those skilled surgeons asked for a raise.
 - c. Prep, this: The lawyer for this skilled surgeon asked for a raise.
 - d. Verb, that: The lawyer visited that skilled surgeon before the hearings began.
 - e. Verb, those: The lawyer visited those skilled surgeons before the hearings began.
 - f. Verb, this: The lawyer visited this skilled surgeon before the hearings began.

Experiment 2: Results

Table 7 Analysis of Variance results for Experiment 2

Source of variance	By participants			В	y items	min F'	
	df	F1 value	MSE	df	F2 value	df	min F' value
Position 3 ("for"/"visited"))						
Category	1,95	3.63	19669	1,35	3.97	1,95	1.89
Determiner	2,190	1.75	15381	2,70	1.64	2,198	<1
Category × Determiner	2,190	<1	20605	2,70	<1	2,198	<1
Position 4 ("that"/"those"/	"this")						
Category	1,95	53.3*	4432	1,35	39.5*	1,95	22.6*
Determiner	2,190	3.40*	3140	2,70	2.33	2,198	1.38
Category \times Determiner	2,190	1.44	3790	2,70	1.63	2,198	<1
Position 5 ("skilled")							
Category	1,95	6.97*	6050	1,35	8.52*	1,95	3.83
Determiner	2,190	17.41*	5939	2,70	22.54*	2,198	9.82*
Category imes Determiner	2,190	1.03	4966	2,70	1.09	2,198	<1
Position 6 ("surgeon(s)")							
Category	1,95	5.65*	4556	1,35	3.95	1,95	2.32
Determiner	2,190	5.84*	5418	2,70	3.90*	2,198	2.33
Category \times Determiner	2,190	<1	3851	2,70	<1	2,198	<1
Position 7 ("asked"/"before	e'')						
Category	1,95	2.70	5377	1,35	1.16	1,95	<1
Determiner	2,190	2.77	8866	2,70	2.70	2,198	1.36
$Category \times Determiner$	2,190	3.01	6575	2,70	2.26	2,198	1.29
Positions 8-end of sentence							
Category	1,95	<1	1441	1,35	<1	1,95	<1
Determiner	2,190	<1	1609	2,70	<1	2,198	<1
$Category \times Determiner$	2,190	1.93	1611	2,70	2.42	2,198	1.07
Positions 4–6 ("that/those/t	his skilled surg	geon(s)")					
Category	1,95	33.05*	2449	1,35	28.47*	1,95	L15.2*
Determiner	2,190	15.53*	2302	2,70	12.18*	2,198	6.82*
Category × Determiner	2,190	1.04	1940	2,70	<1	2,198	<1

Note: Significant effects are marked by asterisk.

- Experiment 3: sentence-initial effect of "that" due to parallel processing?
 - <u>That</u> experienced diplomats would be very helpful <u>made</u> the layer confident.

or a single arbitrary parse will be fast selected among possible ones, and only after reaching a disambiguating point in the sentence we observe a slower reading due to reanalysis?

• Experiment 3: Items

- (15) a. "that"-determiner: That experienced diplomat would be very helpful to the lawyer.
 - b. "those"-determiner: Those experienced diplomats would be very helpful to the lawyer.
 - c. "that"-complementizer: That experienced diplomats would be very helpful made the lawyer confident.





Experiment 3: Results

Longer reading time for "that" (even though it is shorter and more frequent)



Fig. 4. Residual reading times on the word "that"/"those" in sentence-initial context in Experiment 3.

- A flash back to Tabor etal 2004 (expl):
 - ''tossed'' was read slower
 - ✓ Even when the alternative tags were not possible



Fig. 1. Mean residual reading times from Experiment 1. Error bars show one standard error around each data point.

- A flash back to Tabor etal 2004 (exp2,3):
 - ? "transformed" with various behaviors!
 - ? A new component required in formulation (thematic plausibility)



Fig. 2. Mean residual reading times from Experiment 2. Error bars show one standard error around each data point.

- Conclusions (with respect to incrementality):
 - Context-independent category statistics affect parsing
 - Sentence-initial ambiguity effects suggest parallel processing
 - Parallelism is not in line with strict incrementality



Questions?

Fatemeh Torabi Asr Spring 2011