Repeated Measures ANOVA

Processing syntactic ambiguity & working memory capacity

Arina Banga & Tam Ho,

Methodology & Statistics, 14/03/07

Research questions

- Temporary syntactic ambiguity

  De krant publiceerde een column over een voetbalploeg uit de randgemeente en ***een over een basketploeg uit de hoofdstad…***

  - An article or a pronoun?
  - If *een* is a pronoun which noun phrase does it refer to?

    een column? (High attachment)
    een voetbalploeg? (Middle attachment)

- What decision do readers make? Why? And when?

Tuning hypothesis

- Tuning or exposure-based hypothesis

  "[T]he most frequently occuring resolution of an ambiguity is the resolution that people prefer." (Gibson & Schütze, 1999)

- Corpus analysis (English) by Desmet & Gibson (2003):

  - Pronoun "one" occurs more frequent in conjoined noun phrase constructions with *high* than middle attachment

  - In the absence of "one", *middle* atttachments are more frequent than high attachment

Desmet & Gibson (2003)

- Tested the hypothesis online
  - Experiment 1: English word-by-word self-paced reading experiment
  - Experiment 2: Dutch eye-tracking self-paced reading experiment (replicated)
- Prediction: readers' preference should reflect corpus frequency
  - Pronoun $\rightarrow$ high attachment
  - No pronoun $\rightarrow$ middle attachment

Experiment Design

- Crossed factors

	Attachment	
	High	Middle
Parallel NP	Cond1	Cond2
Pronoun	Cond3	Cond4

- 24 set of sentences; each set contained four versions

- *Latin square*: Participants saw only one of the four possible versions

Example

	Region 1	Region 2	Region 3	Region 4	Region 5	Region 6	Rest
	Subj + V	NP1	Prep + NP2	Prep + NP3	en + een/NP	Prep + NP	…
High-NP	*De krant publiceerde*	*een column*	*over een voetbalploeg*	*uit de randgemeente*	*en een artikel*	*over een basketploeg*	…
High-Pro	*De krant publiceerde*	*een column*	*over een voetbalploeg*	*uit de randgemeente*	*en een*	*over een basketploeg*	…
Middle-NP	*De krant publiceerde*	*een column*	*over een voetbalploeg*	*uit de randgemeente*	*en een basketploeg*	*uit de hoofdstad*	…
Middle-Pro	*De krant publiceerde*	*een column*	*over een voetbalploeg*	*uit de randgemeente*	*en een*	*uit de hoofdstad*	…

Results for Region 5

- No main effect of *disambiguation type* or *attachment site*
- Significant interaction between disambiguation type and attachment site ($F(1, 31) = 4.60$, $p < .05$)
  - *NP condition*: No difference in reading times for high and middle attachment ($F(1, 31) = 1.76$, $p = .20$)
  - *Pronoun condition*: Reading times for high and middle attachment differed significantly ($F(1, 31) = 9.59$, $p < .01$)

	High attachment	Middle attachment
Mean (ms)	282	329

Pronoun resolution

- Hemforth et al.'s (2000) *anaphoric binding hypothesis*:
  Parsers prefer to coindex pronouns with elements which belong to the main assertion of a sentence.

  De krant publiceerde ***een column over een voetbalploeg*** *uit de randgemeente en* ***een*** *over een basketploeg uit de hoofdstad…*

  While NP1 is the object of the verb, NP2 belongs only to the modifying PP.

- Also *predicate proximity hypothesis* (Gibson et al. 1996):
  Attachments that are structurally closer to verbs are favored

Locality constraint

- An interpretation associated with a local attachment is preferred over an interpretation with a less local attachment
  (Gibson & Pearlmutter, 1998)

  - Integration cost:

    The greater the distance between an incoming word and the dependent to which it attaches, the greater the integration cost

  - Memory cost

    The longer a predicted category must be kept in memory before being encountered, the greater is the cost for maintaining that prediction

Working memory

- From a *working memory* perspective,
  - ➔ attaching the pronoun to the higher NP involves more processing resources than attaching it to the middle NP
  - ➔ unless a high amount of processing resources is available (e.g. in readers with *high working memory span*)

- Desmet & Gibson did not control for working memory differences between their participants

- Introducing a *between-subject factor*
  - ➔ High vs. low working memory span
  - ➔ Will readers with *low working memory span* also show a high-attachment preference in the pronoun condition?

Repeated measures design

- Latin square design

	High-NP	High-Pro	Middle-NP	Middle-Pro
	Cond a	*Cond b*	*Cond c*	*Cond d*
List1	$Sent_{1a}$	$Sent_{2b}$	$Sent_{3c}$	$Sent_{4d}$
List2	$Sent_{2a}$	$Sent_{3b}$	$Sent_{4c}$	$Sent_{1d}$
List3	$Sent_{3a}$	$Sent_{4b}$	$Sent_{1c}$	$Sent_{2d}$
List4	$Sent_{4a}$	$Sent_{1b}$	$Sent_{2c}$	$Sent_{3d}$

- The four lists were distributed evenly over the two groups
- Why not testing four different groups?
  - ➔ Great variability in individual reading times

Determining span type

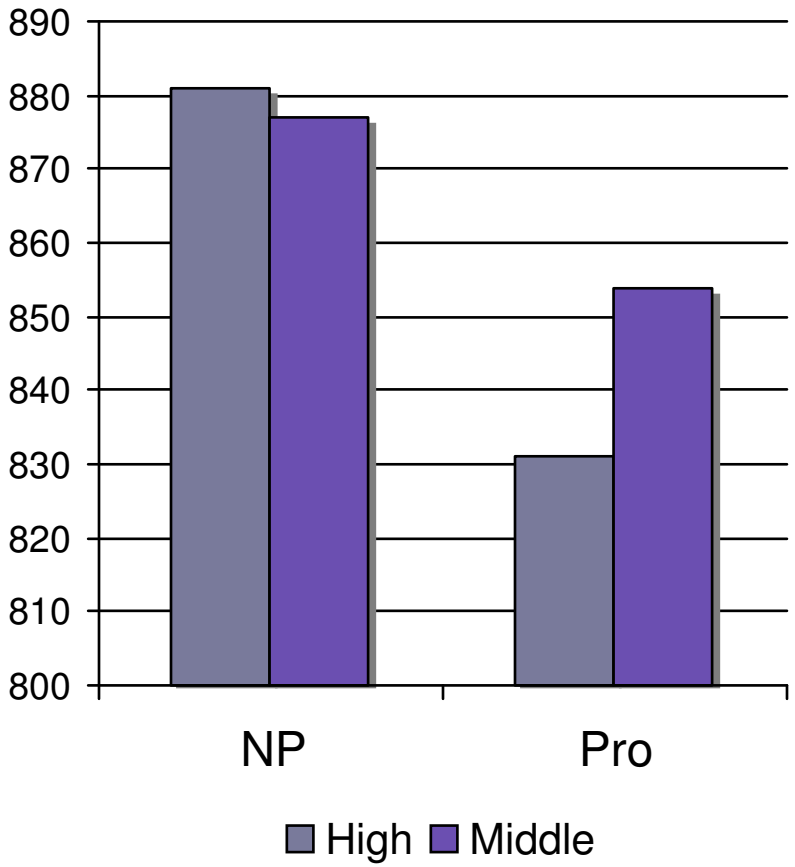
- Reading span task (Danemann & Carpenter 1980; Van den Noort & Haverkort (unpublished)
  - Participants had to read several sentences out loud & remember the final words
  - When signaled, the participants had to recall the memorized words in the order they appeared
  - The number of sentences (& words to be memorised) increased up to 6
  - Scores: Low span 2.0 - 2.5, Middle span 3.0 -3.5, High span 4.0 - 6.0
- Participants
  - 11 high spanners and 9 low spanners
  - All native speakers of Dutch
  - 18-28 years

Method & Procedure

- Word-by-word self-paced reading experiment
  - Each run started with a fixation point: **+**
  - The sentences appeared word by word in the middle of a computer screen
  - The participants pressed a key after finishing reading a word
  - The times between the key presses were recorded as the participants' reading times
  - Questions to some sentences were included to make sure that the participants read for comprehension

- For stimulus presentation & data collection the *E-prime* software was used

Results for Region 5

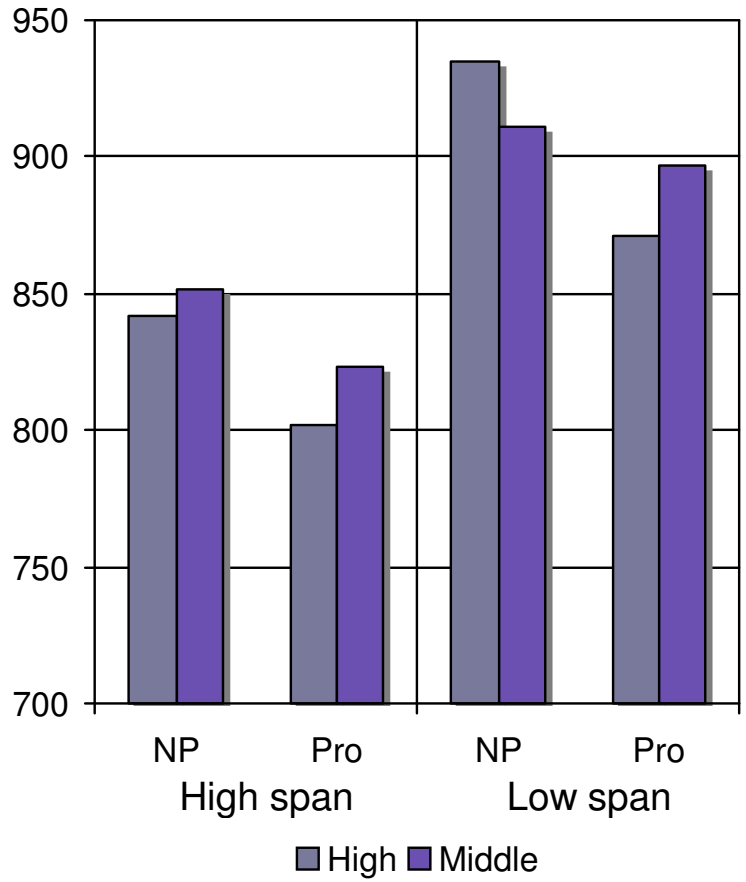


- Significant main effect in attachment site condition, $F(1, 17) = 4.55$, $p < .05$
- Non-significant interaction of attachment site and disambiguation type, $F(1, 17) = 2.52$, $p = .13$
- Desmet & Gibson (2003)

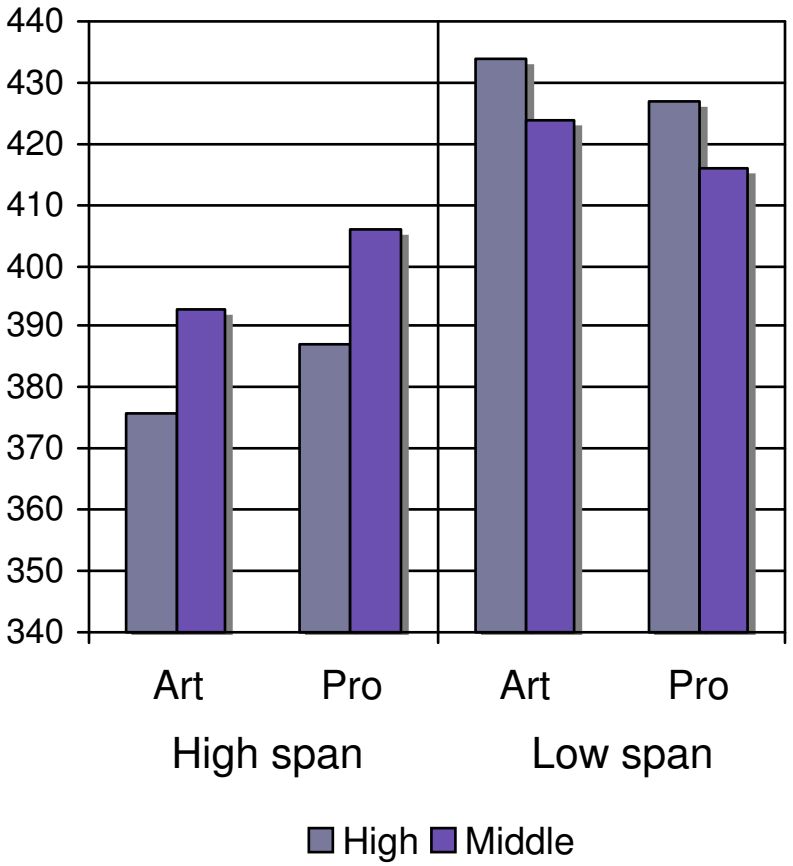
  Significant interaction between disambiguation type and attachment site ($F(1, 31) = 4.60$, $p < .05$) but no main effects

Region 5: high vs. low span readers

- No significant interaction with span
- The two groups show different patterns
- An analysis in terms of regions may be inadequate
- Reanalysis of region 5
  1. Conjunction *en* excluded
  2. Article and Pronoun compared
  3. Site of ambiguity resolution
  4. Two words after



Een - article vs. pronoun



- Significant interaction between disambiguation type and span, $F(1, 18) = 4.34$, $p < .05$

- *Post-hoc analyses*

  look at the two groups individually

  - *High spanners*

    significant main effect of disambiguation type, $F(1, 10) = 7.00$, $p < .05$ ⇒ sensitive to article/pronoun distinction

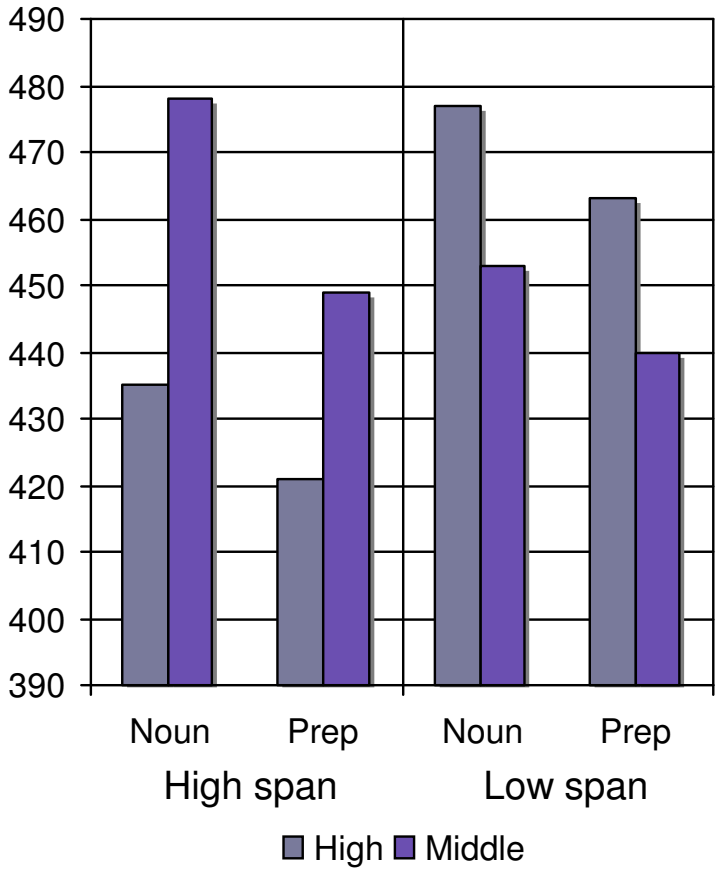
  - *Low spanners*

    no effects ⇒ article/pronoun distinction seems to be ignored

Ambiguity resolution

- Site of ambiguity resolution
  - Noun → *een* is an article
  - Preposition → *een* is a pronoun
  - Noun in the *parallel condition* and preposition in the *pronoun condition* compared

- A near significant interaction between span and the noun vs. preposition condition, $F(1, 18) = 3.90$, $p = .065$
  - Higher reading times for *noun* than *preposition* in both groups ⇒ perhaps *een* was expected to be a pronoun
  - No main effect of attachment site but pattern for the two groups is noticeably reverse



SPSS output

- High span readers

Tests of Within-Subjects Contrasts

Source	hm	nounprep	Type III Sum of Squares	df	Mean Square	F	Sig.
hm	Linear		4908,067	1	4908,067	1,192	,300
Error(hm)	Linear		41159,676	10	4115,968		
nounprep		Linear	14086,450	1	14086,450	2,354	,156
Error(nounprep)		Linear	59847,471	10	5984,747		
hm * nounprep	Linear	Linear	639,601	1	639,601	,920	,360
Error(hm*nounprep)	Linear	Linear	6949,056	10	694,906		

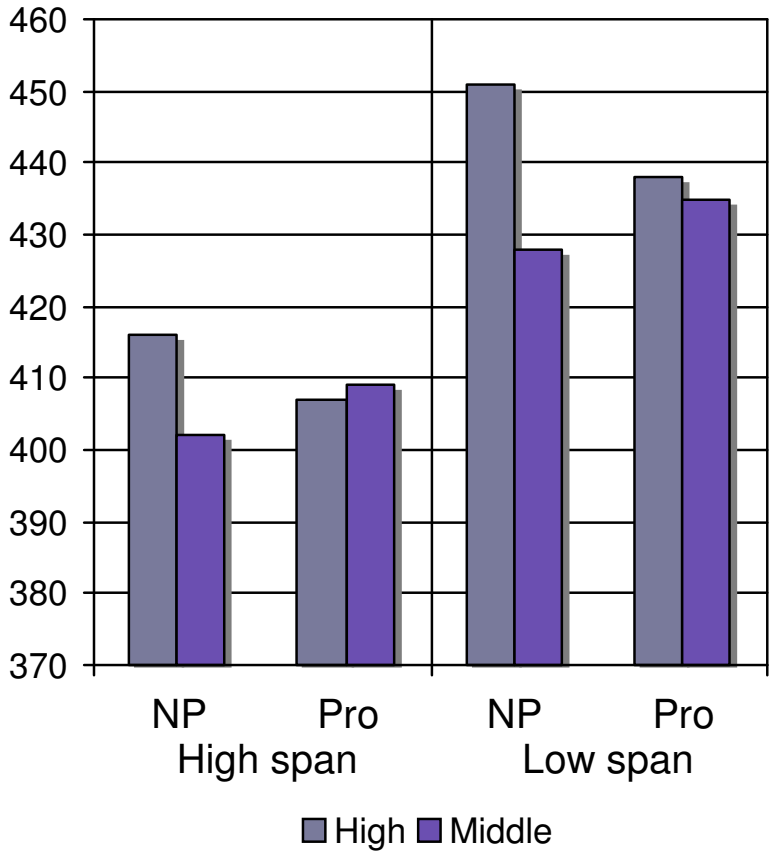
- Low span readers

Tests of Within-Subjects Contrasts

Source	hm	nounprep	Type III Sum of Squares	df	Mean Square	F	Sig.
hm	Linear		1474,439	1	1474,439	,788	,404
Error(hm)	Linear		13092,169	7	1870,310		
nounprep		Linear	4321,733	1	4321,733	2,885	,133
Error(nounprep)		Linear	10487,027	7	1498,147		
hm * nounprep	Linear	Linear	1,181	1	1,181	,001	,978
Error(hm*nounprep)	Linear	Linear	10196,882	7	1456,697		

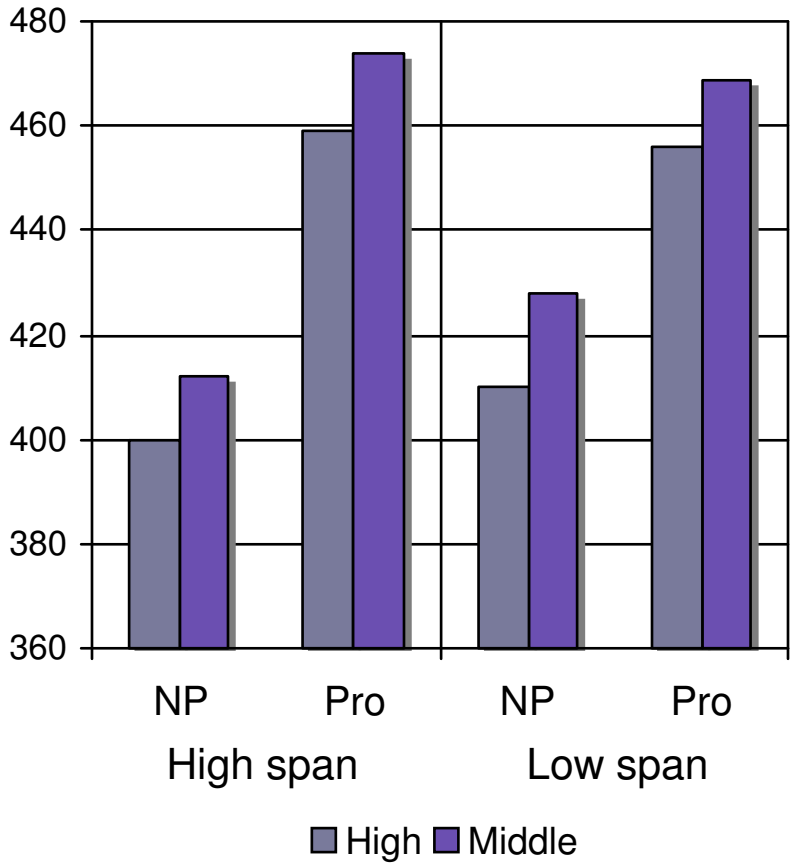
1st word after

- An interaction between attachment site and disambiguation type, $F(1, 18) = 2.05$, $p = .171$ (not significant)
  - *NP condition*: higher reading times for high attachment ⇒ reflects corpus frequency: middle attachment preference in the absence of a pronoun
  - *Pronoun condition*: reading times for high and middle attachment do not differ ⇒ no preference in terms of attachment yet



2nd word after

- Significant main effect of attachment site, $F(1, 18) = 13.33$, $p < .05$
  - Higher reading times for middle attachment in both groups ⇒ *predicate proximity constraint*
- A non-significant main effect of disambiguation type, $F(1, 18)$, $p = .141$
  - Higher reading times for pronoun condition in both groups ⇒ attaching pronoun requires additional processing recourses
- No interaction with Span



Language as a random effect

- *Random factor* : “If the number of possible levels of that factor greatly exceeds that of the number of levels included in the experiment” (Rietveld & van Hout 2005)

- For example: “Randomly” selected words from a dictionary or lexical data-base do not constitute the whole population of possible words

- Item variance can be controlled by experimental procedures such as *matching items* (Raaijmakers 1999)

- To make sure that the differences found in the experiment are due to the different conditions and not due to the differences between the test items themselves better do *item statistics*

Item statistics

- 24 items in each condition; 2 items were excluded
- Region 5
  - Desmet & Gibson (2003) found a significant interaction between attachment site and disambiguation type, $F(1, 23) = 4.76$, $p < .05$
  - Main effect of attachment site, $F(1, 21) = 3.11$, $p = .92$ (not significant)
  - No other effects were found when looking at the words individually
- Number of participants perhaps too low

Conclusion

- When analysing the sentences in terms of regions, experimental finding reflected corpus frequency
- But when the regions (region 5) were analysed word by word, an interaction with other constraints can be seen
  - *Span*: high spanners were sensitive to article/pronoun distinction $\Rightarrow$ perhaps an indication that they make use of both representations, while low spanners only use a single representation
  - *Predicate proximity*: decisions related to attachment site seems to depend on the distance of the NP or pronoun to the predicate as well

References

1. Clarck, H. H. (1973). The Language-as-Fixed-Effect Fallacy: A Critique of Language Statistics in Psychological Research. *Journal of Verbal Learning and Verbal Behavior, 12:4, 335-359.*
2. Daneman, M. & Carpenter P. A. (1980). Individual Differences in Working Memory and Reading. *Journal of Verbal Learning and Verbal Behavior, 19:4, 450-465.*
3. Desmet, T. & Gibson, E. (2003). Disambiguation preferences and corpus frequencies in noun phrase conjunction. *Journal of Memory and Language 49, 353-374*.
4. Gibson, E. & Pearlmutter, N. J. (1998). Constraints on sentence comprehension. *Trends in Cognitive Science 2:7, 262-268.*
5. Just, M.A. & Carpenter, P. A. (1992). A Capacity Theory of Comprehension: Individual Differences in Working Memory. *Psychological Review 99:1, 122-149.*
6. MacDonald, M. C. & Christiansen, M. H. (2002). Reassessing Working Memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review 109:1, 35-54*.
7. Raaijmakers, J. G. W., Schrijnemakers, J. M. C. & Gremmen, F. (1999). How to Deal with “The Language-as-Fixed-Effect Fallacy”: Common Misconceptions and Alternative Solutions. *Journal of Memory and Language 41, 416-426.*
8. Rietveld, T. & Van Hout, R. (2005). *Statistics in Language Research: Analysis of Variance.* Mouton, The Hague.
9. Waters, G. S. & Caplan, D. (1996). The Capacity Theory of Sentence Comprehension: Critique of Just and Carpenter (1992). *Psychological Review 103:4, 761-772.*