

# **Controlling Processing Factors in the Study of Subjacency**

**Philip Hofmeister, Ivan A. Sag, Neal Snider,  
Perry Rosenstein  
Stanford University**

- **What** did Sandy like \_\_\_ ?
- **What** did Sandy say she likes \_\_\_ ?
- **What** did Sandy say she thinks Kim likes \_\_\_ ?
- The person **who** we all said we wanted to meet \_\_\_ is...
- I met **more people** yesterday than anyone else in our group said they had been able to meet \_\_\_ in their time here.

# The Complex NP Constraint

(Ross 1967)

\*Which beverage did Leslie regret the fact that people like \_\_\_ ?

\*What did Lee say she met the woman who likes \_\_\_ ?

\*Which beverage can't Pat stand men who like \_\_\_ ?

**No element can be extracted from a 'complex NP', i.e. an NP consisting of an N, a clause, and perhaps other material.**

## Complex DP (with Relative Clause): (Szabolcsi 2002)

- \*Which kid must you call [the teacher who punished \_\_\_ ]?
- \*Where must you call [the teacher who put the book \_\_\_ ]?
- \*This kid, you must call [the teacher who punished \_\_\_ ].
- \*On the table, you must call [the teacher who put the book \_\_\_ ].
- \*What size shoes did you call [the man who wears \_\_\_ ]?
- \*How did you call [the man who behaved \_\_\_ ]?

## Complex DP (with complement clause): (Szabolcsi 2002)

- \*Which man did you hear [the rumor that my dog bit \_\_\_ ]?
- \*Where did you hear [the rumor that I put the book \_\_\_ ]?
- \*This man, I heard [the rumor that my dog bit \_\_\_ ].
- \*On the table, I heard [the rumor that you put the book \_\_\_ ].
- \*What size shoes did you hear [the rumor that I wear \_\_\_ ]?
- \*How did you hear [the rumor that I behaved \_\_\_ ]?

# Reanalysis as Subjacency

## Chomsky 1973, p. 73

No rule may move a phrase from position Y to position X (or conversely) in:

$$\dots X \dots [\alpha \dots [\beta \dots Y \dots] \dots] \dots X \dots ,$$

where  $\alpha$  and  $\beta$  are cyclic nodes.

(Cyclic nodes in English are S (IP) and NP (DP))

## N-Subjacency (Chomsky 1986)

$\beta$  is **n-subjacent** to  $\alpha$  iff there are fewer than  $n+1$  barriers for  $\beta$  that exclude  $\alpha$

**Subjacency:** If  $(\alpha_i, \alpha_{i+1})$  is a link of a chain, then  $\alpha_{i+1}$  is 1-subjacent to  $\alpha_i$ .

He is the man to whom I [\_\_\_ [wonder [\_\_\_ [who [\_\_\_ [knew [which book to [\_\_\_ [give \_\_\_ ]]]]]]]]]

He is the man to whom I [\_\_\_ [wonder [who John [\_\_\_ [told [which book to [\_\_\_ [give \_\_\_ ]]]]]]]]]

This is a paper [that we need to find [someone [who understands \_\_ ]]]. (Chung and McCloskey 1983)

This is a paper [that we need to find [someone [that we can intimidate with \_\_ ]]]. (Chung and McCloskey 1983)

It was Lucille that Mike went home [without criticizing [anyone who defended \_\_ ]]. (Postal 1998)

It was Lucille that Mike criticized [everyone who went home [without defending \_\_ ]]. (Postal 1998)

## Naturally Occurring Violations of Subjacency.

- Jim did this to an old 20-inch by 32-inch frame that I couldn't imagine ever finding [art to fit \_\_\_ ]. (Amy Dacyczyn. 1995. The tightwad gazette II. New York: Villard. 26. (<http://www.ling.upenn.edu/~beatrice/examples/>))

## Naturally Occurring Violations of Subjacency

Post 1: Incidentally, a friend of mine tried to get into contact with Aiwass a while ago, and while this failed, he did get the impression that Aiwass was not the HGA of Crowley by any means.

Post 2: What did he get [the impression that Aiwass was \_\_\_ ]? (<http://www.ling.upenn.edu/~beatrice/examples/>)

## Much Critical Island Data is Graded

Ross (1987), Kluender (1992, 1998, 2004), and others

\*Which rebel leader did you hear [Cheney's rumor [that the CIA assassinated \_\_ ]]?

??Which rebel leader did you hear [the rumor [that the CIA assassinated \_\_ ]]?

?Which rebel leader did you hear [a rumor [that the CIA assassinated \_\_ ]]?

Which rebel leader did you hear [rumors [that the CIA assassinated \_\_ ]]?

- Although the graded nature of many syntactic islands has been observed in the literature (e.g. by Chomsky 1986, Ross 1986, Kluender 1992, and others) the gradience has never been properly accounted for.
- A plausible reason for this is that the processing (performance) factors that interact with the grammar of filler-gap constructions have never been properly controlled for in the generative literature.

## Processing and Acceptability

- Intuitions of sentence well-formedness are intuitions of **acceptability**.
- Grammatical deviance can make sentences unacceptable.
- Processing difficulty can make grammatical sentences unacceptable.  
(Chomsky and Miller 1963, Kluender 1992+, Fanselow and Frisch 2004, ...)
- Judging acceptability is an experiment that needs to be done in a controlled fashion.

## Factors the Create Processing Difficulty in FGCs, Reducing Acceptability

The longer the distance between filler and gap, the harder the sentence is to process. (Gibson, Lewis, Vasisht,...)

Less accessible NPs intervening between filler and gap make sentences harder to process. (Gibson et al., Klender,...)

Less informative fillers (*who* vs. *which man*) make sentences harder to process (Sag et al. to appear, Hofmeister et al. in press, Hofmeister this session)

- These factors have been used to explain Superiority effects: **The Superiority Constraint (or the Minimal Link Condition, etc.)** are eliminated from the theory of **grammar**, according to Sag et al. and Hofmeister et al.
- They **derive ‘D-Linking’ effects from more general considerations** of contextual accommodation and differences of informativeness.

## Extraction from Nominal Complement Clauses

- He remembered {who/which country} Jill heard {rumors/the rumor/a rumor} that we had invaded due to increased political instability.
- He remembered which country Jill heard that we had invaded due to increased political instability.

Definiteness of the complex NP is also targeted as a factor in processing difficulty, because contrasts like the following have been noted in the literature:

- This is the paper that we really need to find a linguist who understands. >
- This is the paper that we really need to find the linguist who understands.

## Investigating Subjacency Effects with Reading-Time Studies

- Self-paced reading time (SPR) study.
- In SPRs, subjects read sentences at their own pace on a computer screen one word at a time.
- - - - - remembered - - - - - .
- Reading times for each word are then collected and analyzed as a measure of language processing difficulty with longer reading times signaling more difficulty.
- Here, we report only residual reading times, which factor out the effect of word length on reading time.
- Standard error bars are shown.

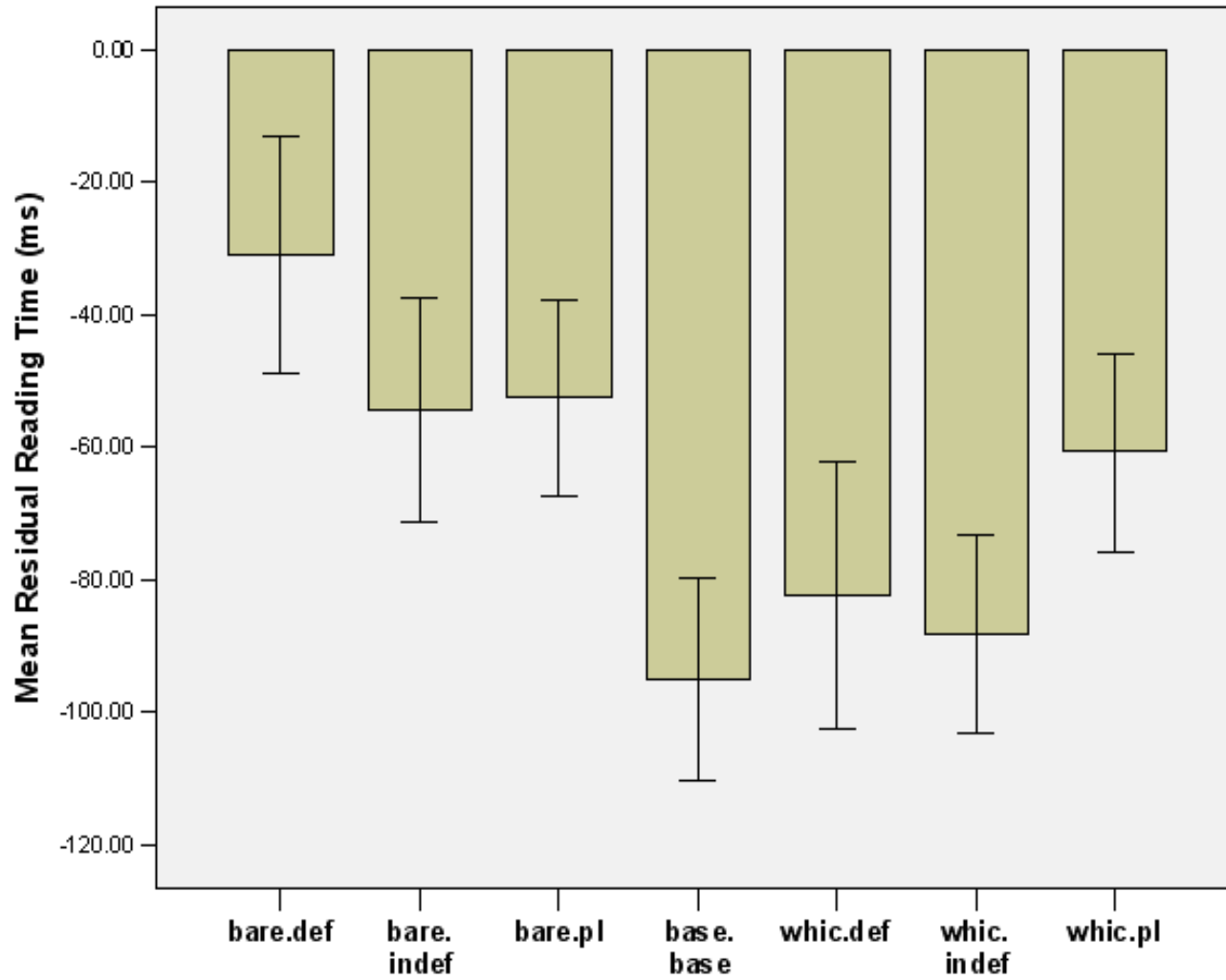
## Reading-Time Study of CNPC Violations

- 3 (definite/indefinite/plural) × 2 (which vs. bare) + 1 (baseline) design
- 36 items
- 72 fillers
- 25 subjects paid 15 dollars/session
- Each subject saw 1 condition for each item. (Latin Square Design)
- Stimuli were followed by a YES-NO comprehension question.

- He realized which prisoner Ashley countered the belief that we had interrogated without regard to international law. (WHICH-DEF)
- He realized which prisoner Ashley countered a belief that we had interrogated without regard to international law. (WHICH-INDEF)
- He realized which prisoner Ashley countered beliefs that we had interrogated without regard to international law. (WHICH-PLURAL)

- He realized who Ashley countered the belief that we had interrogated without regard to international law.  
(BARE-DEF)
- He realized who Ashley countered a belief that we had interrogated without regard to international law.  
(BARE-INDEF)
- He realized who Ashley countered beliefs that we had interrogated without regard to international law.  
(BARE-PLURAL)
- He realized who Ashley countered that we had interrogated without regard to international law. (BASELINE)

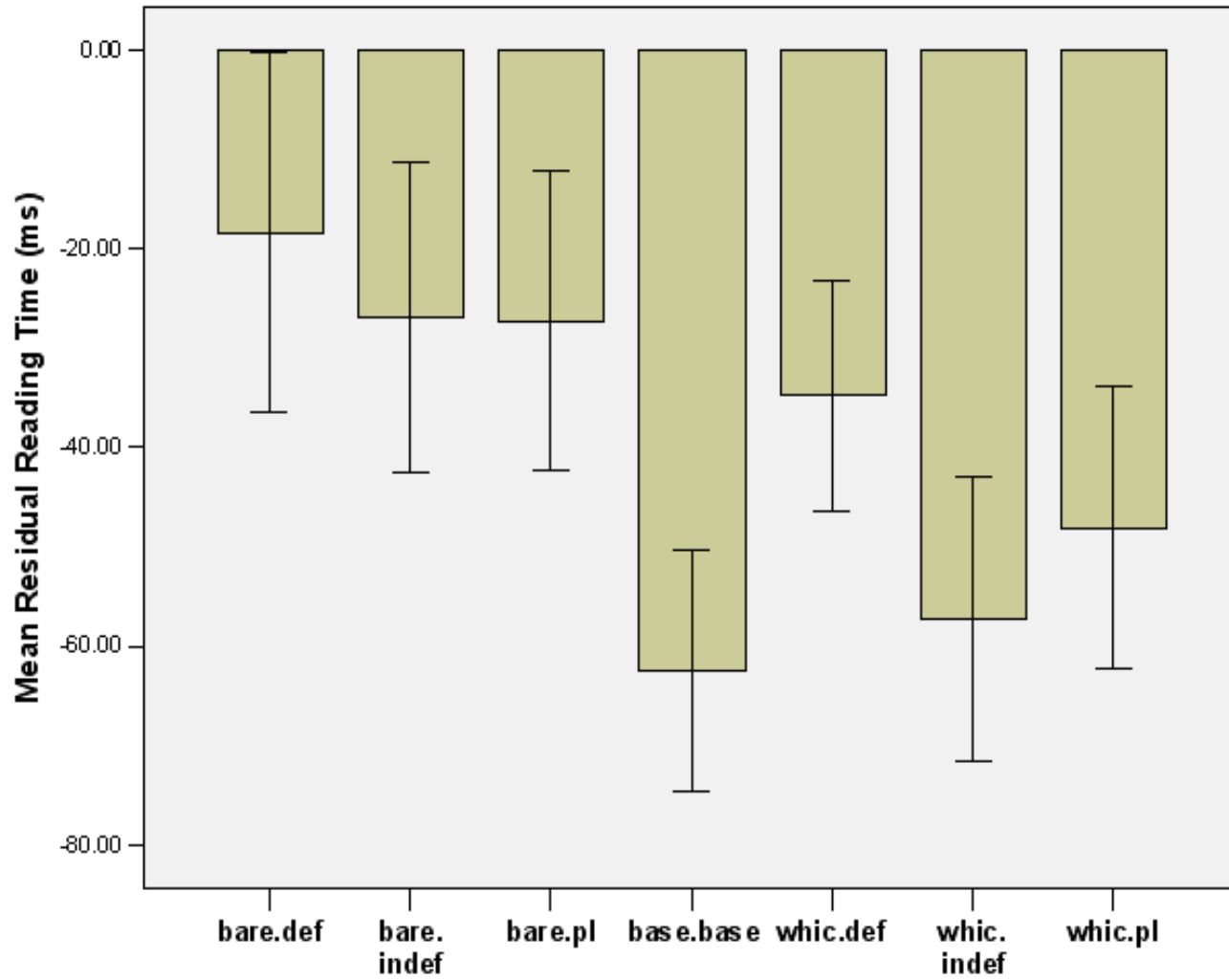
# Reading Times at Verb in CNPC Experiment 1



## Reading Times at Verb in CNPC Experiment 2

- Main effect of *wh*-type
- $F_1(1, 24) = 7.311, p = .012$
- $F_2(1, 35) = 8.494, p < .01$
- *which*-conditions not significantly different than baseline
- no significant effect of definiteness

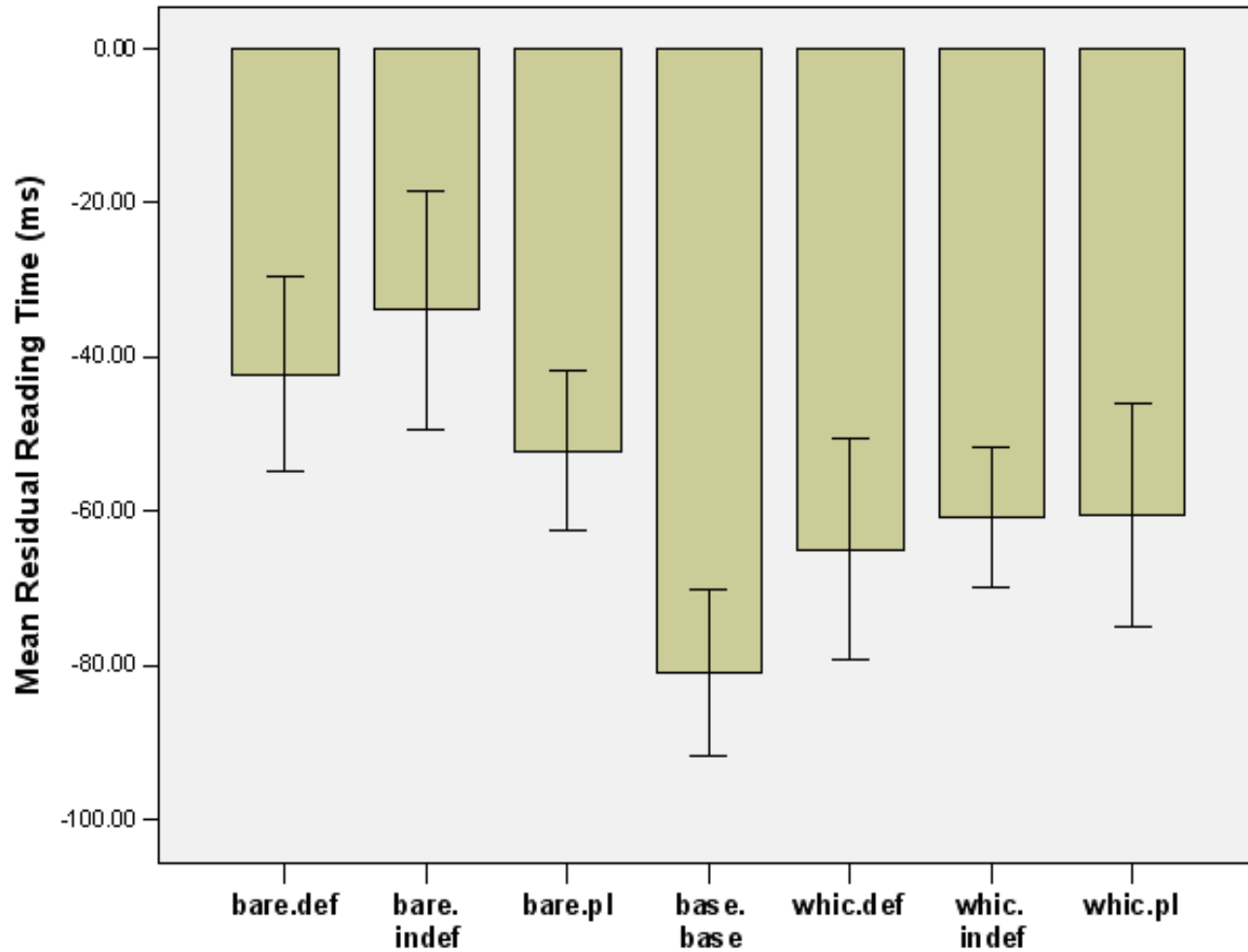
# Residual Reading Times One Word After Verb



## Reading times in spillover region

- Main effect of *wh*-type in spillover region, as well
- $F1(1, 24) = 7.37, p = .012$
- $F2(1,35) = 11.45, p = .002$
- No effect of definiteness
- Similar equivalence of baseline and *which*-conditions

# Residual Reading Times Two Words After Verb



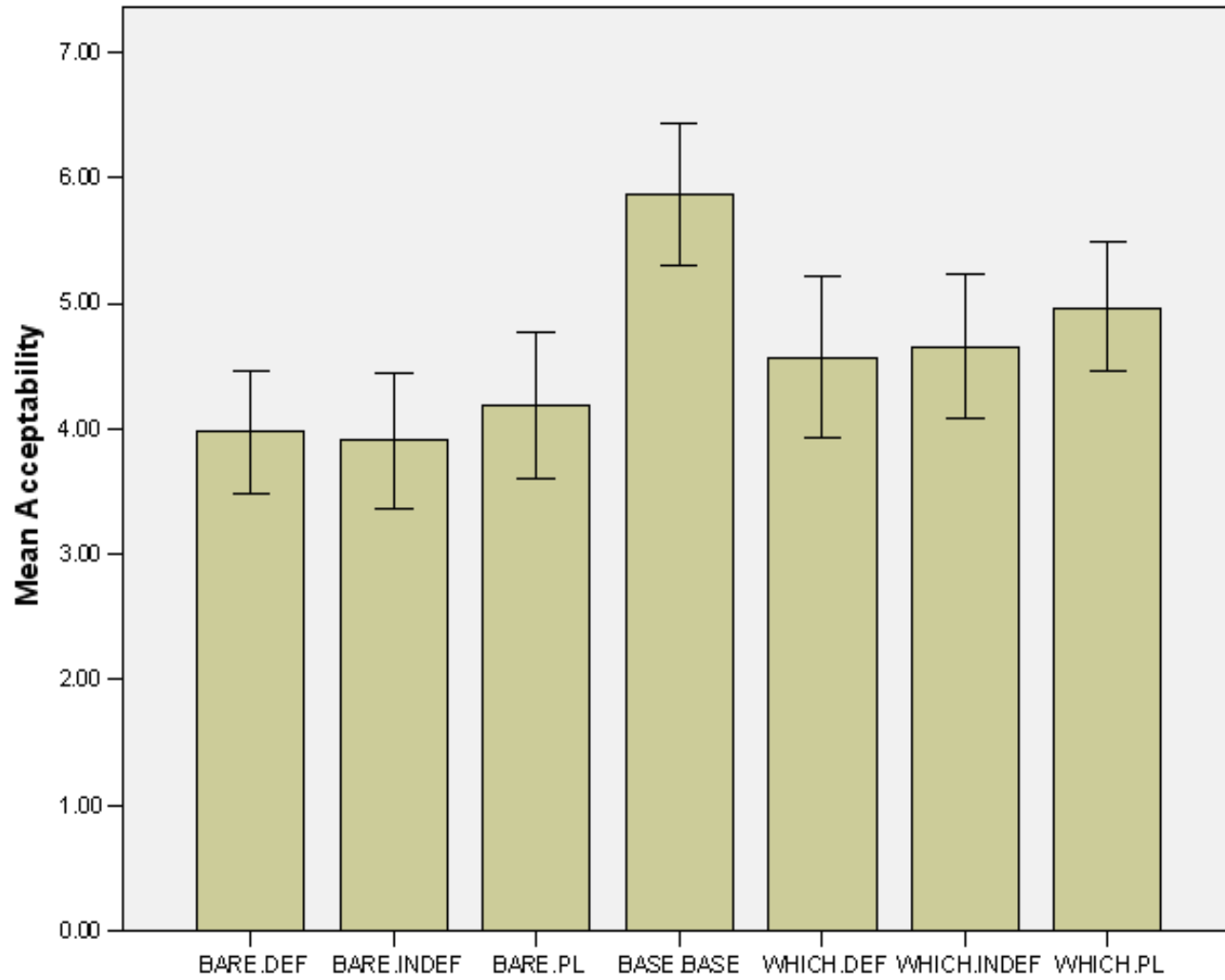
## Reading Time Experiments: Summary

- Clear effect of wh-type – informativeness of the wh-expression, both at verb and the spillover region.
- As the filler becomes more informative, the reading time approaches that of the baseline.
- Want to verify that this processing difficulty converges with (controlled) acceptability results

## Acceptability Judgments in CNPCs

- Same design and items as in reading-time study
- 16 subjects (different from 25 participants in reading-time study), given course credit for participation
- Sentences judged for naturalness on a scale of 0-7 to eliminate midpoint
- Error bars here represent 95% confidence intervals.

# Acceptability Judgment Results 1



## Acceptability Judgment Results 2

- Main effect of *wh*-type: which-NPs judged as being better than bare *wh*-items.
- $F(1,15) = 40.762, p < .001$
- $F(1,35) = 25.053, p < .001$
- Observed power = .998, i.e. .002 percent chance of a false positive
- Trend, but no significance, in definiteness: plurals judged better than definites and singular indefinites.
- Baseline significantly better than any island condition.

## Conclusions 1

- We have presented findings about the gradient nature of subjacency effects in English.
- We've also offered a framework in which subjacency effects in nominal complement clauses can be attributed to processing difficulty.
- This difficulty can easily be overcome by minimizing other sources of processing complexity.
- These subjacency effects are thus in large part captured by processing factors known to affect other filler-gap constructions.

## Conclusions 2

- It seems likely that competence grammar needs neither the CNPC nor the Subjacency Condition to rule out extraction from nominal complement clauses.
- These CNPC ‘violations’ are grammatical, but of reduced acceptability because of increased processing difficulty.
- This processing account better models the gradient nature of the phenomenon and its interaction with other factors.
- Our approach will let us build linguistic theories on a solid empirical base of results from diverse methodologies, including acceptability judgments, reading times, ERP activity (Kluender), and eye-tracking patterns (Tanenhaus et al.).

## Conclusions 3

- Our approach will also change the nature of competence theories because the data to be accounted for, once processing factors are controlled for, will be quite different.
- A processing account is more explanatory, deriving a gradient space of acceptability judgments from the very mechanisms that explain processing times.
- Here, we have reported on just one aspect of our broader efforts to understand the baseline processing difficulties in filler-gap constructions.
- We're also looking at different island phenomena and different experimental methodologies in our effort to recalibrate the space of data sets that island theorists address.

- The processing-based approach also provides new, more adequate ways that cross-linguistic variation can be modeled and quantified. (Arnon et al. in press)



## BUT: Cross-linguistic challenge

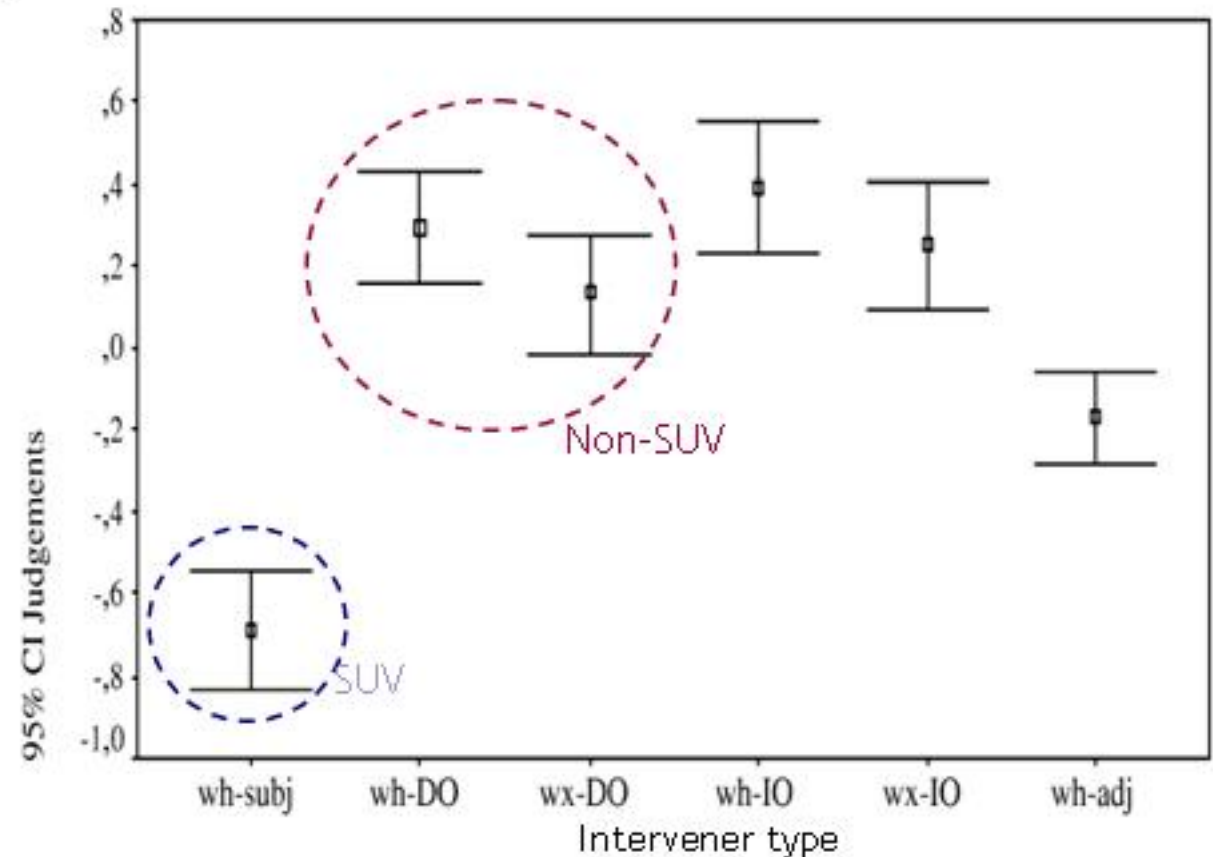
- If it's processing preferences, why the difference between languages?
- Re-evaluate the claims with more precise measures
- Review cross-linguistic data: Russian and German

# Superiority in German

Superiority has been argued not to exist in German (Grewendorf 1988, Müller 1991)

However, Featherston (2005) showed

- 1) SUVs are worse than non-SUVs
- 2) More accessible interveners improve acceptability  
( $F1=5.71$ ,  $p=0.025$ ;  
 $F2=51.65$ ,  $p<0.001$ )





# Superiority in Russian

- Argued not to exist in Russian
- Fedorenko and Gibson (2006) found no effect of wh-order (Superiority)
- Also no effect of filler or intervener accessibility

## The Competition Model (MacWhinney 1987)

- In this framework, the relative strength of surface cues like word order, case marking and subject-verb agreement is responsible for differences among languages in on-line and off-line sentence interpretation.

## The Strength of a Cue: 3 Factors

- Availability: the proportion of times that it is present.
- Reliability: the proportion of times where the cue marks the correct interpretation, when it is present.
- Cost: depends on the perceptual salience of the cue and the load it places on working memory.

## The Strength of Case Marking

- Availability: the proportion of times that the noun has unambiguous case marking.
- Reliability: the proportion of times a nominative-marked noun is the Agent of the sentence.
- The strongest cues lead to fastest reaction times and conflicting cues lead to inhibition and slowdown (Li, Bates & MacWhinney 1993).

## **Kempe and MacWhinney 1999**

- looked at the way that the availability of a cue is reflected in the processing benefits associated with it in on-line processing.
- Participants heard simple transitive sentences and had to identify the Agent as quickly as possible. While some sentences were ambiguous, others had various cues to the thematic assignment. The study manipulated the existence of cues like animacy, word order, and case marking.

## Kempe and MacWhinney 1999 Results

- cues that are more frequent, had a bigger benefit in on-line processing.
- corpus study of availability and reliability of animacy, case marking, and word order in German and Russian
- reliability is identical in the two languages
- case marking is less available in German than in Russian: there are more ambiguously marked nouns in German.

- Accordingly, reaction times were more speeded when case marking was added in Russian in comparison to German.
- The results were interpreted as showing that because case marking is more available in Russian, Russian speakers rely on it more in on-line processing.

## **Kempe and MacWhinney 1999**

Drawing on these findings, we suggest that

- the cost of general processing preferences is mediated by the availability of other cues in the languages.
- In a language where case marking is a highly available cue, speakers will rely on that cue and will be able to tolerate increased distance better.

- In other words, increased distance is costly across languages, as is lower accessibility, but the cost of violating those preferences is lower when other cues are highly available.
- In a fixed word order language, with no case marking cues, increased distance is very costly.
- In a language with case marking, distance is increasingly less costly, depending on the availability of other cues.

Can this kind of model explain the reported cross-linguistic differences in the processing of multiple wh-questions?

## The Case for Case

- One striking difference between the three languages is the availability of case marking:
- Case marking on nouns is not an available cue in English.
- Case marking exists in both German and Russian, but the case marking paradigms of nouns in German are more ambiguous than the ones in Russian (Kempe & MacWhinney 1999 for declaratives).

- Crucially, this also seems to hold for the availability of case marking in questions words.
- The morphological paradigm of German question words is more ambiguous than that of Russian.
- Three out of the seven German question words are ambiguous between nominative and accusative case, while only three out of ten Russian questions words are.

- We conducted a corpus study to test the hypothesis that German and Russian differ in the availability of case marking for question words.
- German: the (syntactically annotated) TIGER (v. 1) and NEGRA corpora, which consist of 50,000 sentences (900,000 tokens) and 10,000 sentences (176,000 tokens) of newspaper text, respectively.
- Russian: the dependency-parsed Uppsala corpus (Boguslavsky et al., 2002). The corpus consists of 17,772 sentences (256,034 tokens) of literary and informative text

- The availability of case marking for question words: the percentage of question words that were unambiguously marked as nominative or accusative out of the total number of question words.
- For German, only 11.3% of the question words in our sample are unambiguously case marked.
- In Russian, availability was three times higher: 34.8% of question words were unambiguously case marked.

- The differing effect of the processing preferences could be attributed to the differential availability of case marking as a cue.
- The effect of distance and accessibility is masked when case marking is a highly available (Russian).
- The effect is apparent when case marking is less available (German) and
- is the strongest when case marking is not an available cue (English).

- The gradient cross-linguistic difference is thus plausibly attributed to the different availability of case marking.

## Superiority Violations are Rare, but they Exist

- A: did you know that there are no licensing laws or sales taxes in andorra?

B: I did not. What did who bring back? [<http://gofree.indigo.ie/~rev/text/log/log046.html>]

- Although nothing on this planet (or any other) can compete with the utter horror that is cilantro! Where the heck did who the heck come up with adding that gaw-dawful weed to otherwise civilized hote-cue-zeen? [<http://www.scrappleface.com/MT/archives/001655.html>]

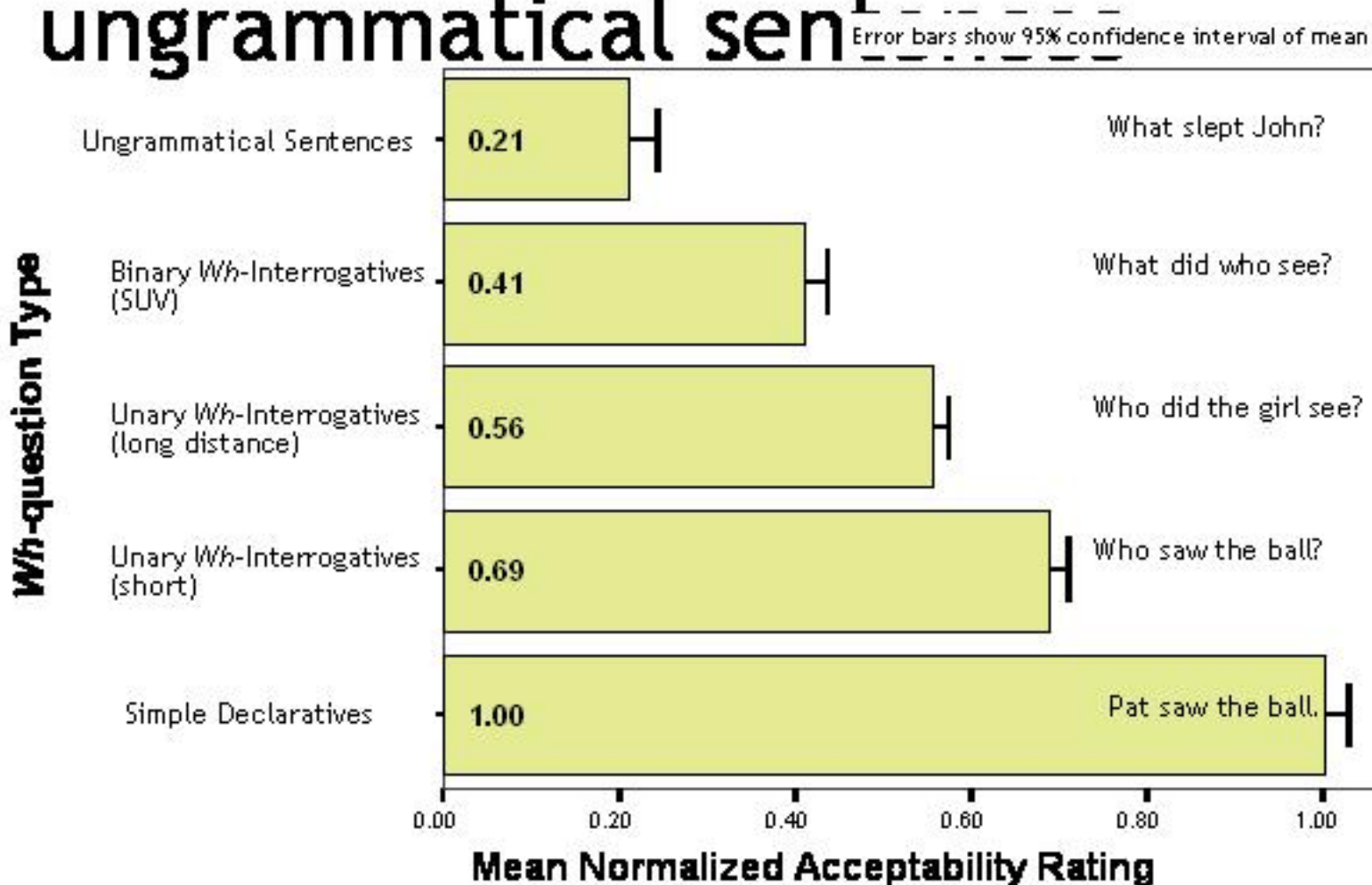
## SUVs Do Occur:

- We watch for a few more minutes as the recriminations begin: what did whom [*sic*] say and what did who hear? [*Tomato is Coming*, by Crispin Oduobuk]
- What, do you think this is a game? What rules should who follow? This shitsandwich is a reality – a competition for survival between all souls... [<http://www.cruel.com/discuss/viewTopic.php/83308>]

## SUVs Do Occur:

- Paige was silent on the phone for a moment. “Wait a minute...he asked you? He really asked you? You said yes? Oh my God!” Paige exclaimed.  
That got Jubilee’s attention. She turned around and looked at Paige. “What?! What did who ask her? What did she say yes to?” [<http://shifting-sands.alara.net/stories03/epin02.htm>]
- I must have missed something. What did who do to Pierre Salinger? [<http://www.freerepublic.com/forum/a3b1c8a4d1847.htm>]

# SUVs rated differently from ungrammatical sentences



## *Wh* Processing Hypotheses

- Gaps that are further from the filler are harder to process. (Gibson 1998, 2000)
- Less accessible fillers make the dependency harder to resolve.
- Less accessible intervenors make the dependency harder to resolve. (Warren and Gibson 2002)

## Summary

- We have cast cross-linguistic variation as an interaction of processing preferences and language specific features.
- Different costs are associated with dispreferred options, depending on the availability of other cues.
- The 'cost of dispreferred options is mediated by the strength of cues like word order, case marking, etc. that bias to the intended parse.
- With regard to multiple wh-questions, the different manifestation of ordering preferences across English, German, and Russian can be partially attributed to the different availability of case marking in those languages.
- Case marking is more available in these languages than in English. This explains why the effect of distance and

accessibility is most apparent in English, less so in German and not apparent in Russian.