Reflections on Superiority

The Stanford WH-Group
Inbal Arnon, Bruno Estigarribia, Philip Hofmeister, Florian Jaeger, Ivan A. Sag, and Neal Snider
Stanford University
Competence-Performance

- **Competence:** Our tacit, internalized knowledge of a language.

- **Performance:** The external evidence of language competence. Particular occasions of language usage, whose form is affected by factors other than competence.

- **Performance factors:** short term memory limitations, whether or not we have been drinking,...
Competence-Performance II

- **Competence:** A central repository of linguistic knowledge.

- **Performance:** Specialized mechanisms for comprehension and production of language.
Strong Competence Hypothesis

The process model must make direct use of the principles of grammar, as defined by the competence theory.
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 (Fanselow and Frisch to appear).
What Causes Low Acceptability: Grammar or Processing?

- The boy the cat the dog bit scratched started crying
- The correspondent everyone I met trusts is interviewing the president
- Gradience and amelioration by orthogonal factors suggests a performance account.
Superiority Effects

Who ___ saw what?

*What did who see ___ ?

Who ___ talked to who?

*Who did who talk to ___ ?
Superiority Condition (Chomsky 1973)

No rule can involve X,Y in the structure:

\[ \ldots X \ldots [ \ldots Z \ldots - WYV \ldots ] \ldots , \]

where the rule applies ambiguously to Z and Y,

and Z is superior to Y.
Universal Grammar

- Constraints of Competence Grammar are universal.
- Constraints of Competence Grammar are complicated.
- Constraints of Competence Grammar couldn’t be learned from experience.
- Constraints of Competence Grammar must be innate.
- Constraints of Competence Grammar seem specific to language.
- Constraints of Competence Grammar are part of the human biological endowment for language.
Superiority ‘Violations’

- Which newspaper did which student read __ ? [Karttunen 1977]

- I know what just about everybody was asked to do, but what did who (actually) do __ ? [Bolinger 1978]

- These Superiority ‘violations’ (henceforth SUVs) were subsequently explained away by Pesetsky (1987) as the result of a distinct grammatical mechanism which he called ‘D(iscourse)-Linking’.

- Who knows what who bought? (Kennedy 2005)
More SUVs:
The Principle of Minimal Compliance

- Who gave what to whom?
- What did who give to whom? (SUV)
- Who put what where?
- Where did who put what? (SUV)

Answer: Norvin Richard’s (1998) PMC:
‘You only have to pay the Superiorty Tax once’ (Pesetsky 2000).
Observation 3
Island violations with arguments in rhetorical questions in languages without overt wh-movement show Principle of Minimal Compliance effects.

Richards (1998) observes that under specific circumstances, island violations are judged better when a second wh-word is added to the construction. For example, following Richards’s assumption that Subjacency holds of both overt and covert movement, the contrast between (5a) and (5b) seems to indicate that adding a wh-word outside the island rectifies the island violation incurred by the in-situ (and covertly moved) wh-word.
(5) a.*What do you wonder whether John bought?

   b. Who wonders whether John bought what?

For Richards, this is a specific instance of a more general principle called the Principle of Minimal Compliance (PMC), in which constraint violations lose their effect if the very same constraint is respected elsewhere in the sentence. Of interest here is that island violations in MRQs demonstrate PMC effects in languages that allow MRQs, such as Japanese.
Known Processing Effects: the uncontrolled background

A wide array of experiments and experimental methodologies in psycholinguistics has confirmed the difficulty of processing long-distance dependencies.

Inside long-distance dependencies, reading times and response times to various tasks increase. (Wanner and Maratsos 1978; Chen et al. 2005).

Island constructions typically complicate these already considerable difficulties of dependency processing.
• For instance, many island constructions require processing numerous referential entities and relatively long-term storage of syntactic predictions for incomplete propositions, while simultaneously retaining the filler-phrase in memory and searching for the correct gap site which has no phonological manifestation.

• In this sense, syntactic islands are akin to the center-embedding examples discussed in Chomsky and Miller 1963 that represent the paradigmatic case of performance difficulties.
Factors that Create Processing Difficulty in Filler-Gap Constructions

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

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

Less informative fillers (who vs. which man) make sentences harder to process (Hofmeister 2007; Sag et al. to appear; Hofmeister et al. in press)
• Just as eliminating a clausal embedding in the center-embedding structures facilitates processing, a variety of factors in island constructions can be manipulated so as to maximize overall processing ease.

• Under such conditions, acceptability may rise to levels comparable to the acceptability of minimally different non-islands.

• E.g. reading-time evidence suggests that the usage of more informative filler-phrases facilitates the processing and acceptability of long-distance dependencies into syntactic islands.

• Can processing difficulty explain island effects without competence grammar constraints?
Superiority

- What did who like? (cf. Who liked what?)
- Can processing factors explain Superiority effects? (Sag et al. to appear; Hofmeister et al. in press)
Superiority Violations are Hard to Find, but they Exist!

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

- 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 gawdawful weed to otherwise civilized hote-cue-zeen? [http://www.scrappleface.com/MT/archives/001655.html]
More SUVs:

• 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]
What the fronted *wh*-expressions in these sentences have in common, however, is the fact that the initial *wh*-expression continues the topic thread and generally asks the more pertinent and relevant question in relation to the previous discourse.

You’re a complete mess... What did who do to you when you were a child?! 

You’re a complete mess... Who did what to you when you were a child?!
Accessibility of Filler x Intervener

Log-normalized acceptability rating

<table>
<thead>
<tr>
<th>Condition</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>WHICH-WHICH</td>
<td>0.52</td>
</tr>
<tr>
<td>BARE-WHICH</td>
<td>0.46</td>
</tr>
<tr>
<td>WHICH-BARE</td>
<td>0.37</td>
</tr>
<tr>
<td>BARE-BARE</td>
<td>0.38</td>
</tr>
</tbody>
</table>
Quick Summary: Superiority Effects

- Superiority effects are graded.
- General principles of processing may be able to explain superiority effects.
Cross-Linguistic Variation

But how can Superiority effects be explained in terms of processing? The Superiority Constraint isn’t universal. German doesn’t have it for example:

Was hat wer dem Patienten empfohlen?
what has who to.the patient recommended
‘What has who recommended to the patient?’
Cross-Linguistic Variation: German (Featherston 2005)

- Was hat welcher Zahnarzt dem Patienten empfohlen?
  what has which dentist to the patient recommended
  ‘What has which dentist recommended to the patient?’

- Was hat wer dem Patienten empfohlen?
  what has who to the patient recommended
  ‘What has who recommended to the patient?’
Russian (Fedorenko et al. 2005)

- Elena staralas’ razobrat’sia kto chto zakazal. Elena tried figure-out who what ordered ‘Elena tried to figure out who ordered what.’

- Elena staralas’ razobrat’sia chto kto zakazal. Elena tried figure-out who what ordered ‘Elena tried to figure out who ordered what.’
• Arnon et al. (in press), building on MacWhinney’s (1987) Competition Model, offer an account of these differences in terms of differences in the availability of case marking as a processing cue.

• Russian case morphology is sufficiently rich to enable speaker-hearers to overcome the processing difficulties imposed by less accessible fillers and interveners.

• English speaker/hearers have no such morphological cues, but are sometimes aided by discourse information, as we have seen.

• German is somewhere in between.
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.
Explaining Cross-Linguistic Variation

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.
Results

- 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.
Discussion

- 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).
Conclusion

The gradient cross-linguistic difference is thus plausibly attributed to the different availability of case marking.
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.
• The gradient cross-linguistic difference in Superiority effects is thus plausibly attributed to differences in the availability of case marking.
Conclusions

• Reading times and acceptability of English *wh*-questions are graded.

• There is considerable effect of orthogonal factors.

• The observed findings are in large part explained by appeal to processing factors known to affect other types of unbounded dependencies.

• The gradient, variable nature of these effects follows from a processing-based (performance) account (in line with standard competence/performance distinction).
• The generative grammar literature on *wh*-phenomena is based on highly questionable datasets.

• Understanding the various processing factors that affect acceptability judgments will change the nature of competence theories because the data sets left to be accounted 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.

• This approach holds promise for maximizing the explanation of linguistic phenomena, while at the same time minimizing grammar.
Collaborators:

- Philip Hofmeister
- Neal Snider
- Florian Jaeger
- Inbal Arnon
- Bruno Estigarribia

And Thanks to:

- Tom Wasow, Nick Romero, Perry Rosenstein, Anubha Kothari, Ted Gibson, Ev Fedorenko