Quantifying weak and strong crossover for *wh*-crossover and proper names

Summary. Strong and weak crossover have been studied for decades (Postal, 1971), yet there is little experimental work quantifying the relative severity of these violations. We develop a novel experiment which shows a significant difference in meaning acceptability between strong and weak crossover in English, favouring theories which distinguish the two. This experimental method also lets us address controversial cases of crossover where appeal to intuition has been insufficient, such as cataphora with proper names. We provide quantitative data showing that this displays a similar strong vs. weak crossover effect. More generally, our method provides a way to empirically probe cross-linguistic variation involving crossover phenomena, something which was long overdue.

Methodology. Classical, standalone binding / crossover sentences may be judged acceptable because binding is possible – or, because participants accommodate some other referent that the pronoun corefers with. This creates a confound for previous experiments (e.g. Kush 2013) using acceptability judgements. Our design capitalises on this ambiguity by testing two readings (co-indexations) of the same sentence: whether the pronoun corefers with a preceding distractor NP (*j*) or whether it is bound by the *wh*-word (*i*). This disentangles whether it is the structure of the sentence or its reading which causes the crossover violation.

Data (*wh*-crossover). We use a 2x3x2 design which compares the two orders *wh*…[gap]…pronoun (binding, B) and *wh*…pronoun…[gap] (crossover, CO) across three sentence types corresponding to strong (S), secondary strong (2S) and weak (W) crossover. Each sentence has two readings:

(1) *SB*: The teacher$_j$ wondered which$_i$ of the students _ enjoyed the essay topic they$_{i/j}$ had chosen.

(2) *2SB*: The teacher$_j$ couldn't decide which$_i$ student's poem topic _ frustrated them$_{i/j}$ the most.

(3) *WB*: The teacher$_j$ wondered which$_i$ of the students _ enjoyed their$_{i/j}$ project topic.

(4) *SCO*: The teacher$_j$ couldn't remember which$_i$ of the students they$_{i/j}$ said _ didn't need to hand in the essay.

(5) *2SCO*: The teacher$_j$ couldn't decide which$_i$ student's poem topic they$_{i/j}$ liked _ the most.

(6) *WCO*: The teacher$_j$ wondered which$_i$ student their$_{i/j}$ project topic frustrated _ the most.

We test each configuration in six lexical variants. We also test each item across masculine, feminine and singular *they* pronouns (see Bjorkman, 2017; Conrod, 2019; i.a.) for a total of 108 test items.

Response type piloting. We ran a pilot with 200 participants to compare two response types: (A) Present the target sentence and ask participants to rate two side-by-side paraphrases for the distractor NP and bound readings. (B) Present a context supporting one reading, then ask participants to rate the target sentence. Each used a 5-point Likert scale. Results trended in the same direction for both, but the paraphrase task (A) produced crisper results. Below, we present two experiments with the paraphrase task (see Fig. 5); next steps include a full replication with the context task (B).

Experiment 1. We recruited 144 self-reported native English speakers using Prolific (8 excluded, $n = 136$). Participants saw 6 target items corresponding to (1-6) and 6 fillers, in random order.

Results (*wh*-crossover). We fit an ordinal mixed effects model in R using `ordinal` (Christensen, 2019) with an interaction between gap/pronoun order and reading. Fig. 1 shows the model's proportions of ratings for each condition. We see little

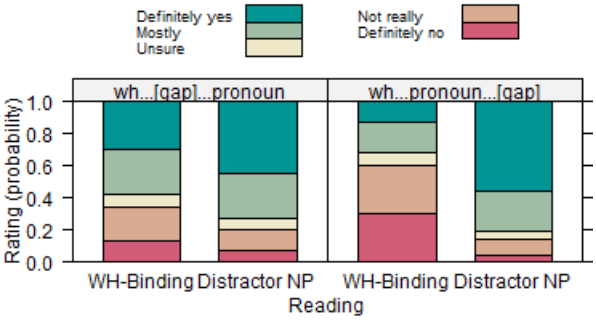

Fig. 1: Effect of gap/pronoun order

effect of reading alone, but a clear, significant effect of pronoun-before-gap on the bound reading (i.e. crossover vs. binding). We also see a significant positive effect of pronoun-before-gap with the distractor reading. This shows that it is the interpretation causing the low ratings, since the sentences are identical. We fit a second ordinal mixed effects model on just the bound reading of pronoun-before-gap (crossover) items to quantify the effect of strong vs. weak crossover, shown in Fig. 2. Notably, this effect is significant; weak crossover roughly doubles the likelihood of a high rating. Finally, results are comparable across pronoun gender, but singular *they* shows the least bias against bound readings compared to the distractor NP.

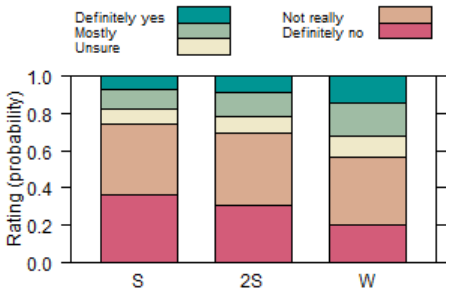

Fig. 2: Strong vs. weak crossover

Data (proper names). We use a 2x2 design which crosses proper name and pronoun order with strong and weak (possessive) configurations, balanced for pronoun gender. The acceptability of the *his*$_i$ reading in (10) is disputed (Chomsky, 1976; Lasnik and Stowell, 1991):

(7) The chef$_j$ knew that Daniel$_i$ was disappointed by the soup he$_{i/j}$ made.

(8) The chef$_j$ knew that Daniel$_i$'s soup had disappointed him$_{i/j}$.

(9) The chef$_j$ knew that he$_{i/j}$ was disappointed by the soup Daniel$_i$ made.

(10) The chef$_j$ knew that his$_{i/j}$ soup had disappointed Daniel$_i$.

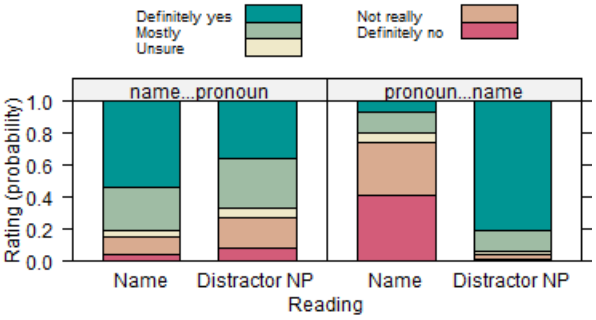

Fig. 3: Effect of name/pronoun order

Experiment 2. We recruited 48 native English speakers using Prolific (1 excluded, $n = 47$). Participants saw 6 target items and 6 fillers.

Results (proper names). We fit an ordinal mixed effects model with an interaction between name/pronoun order and reading, shown in Fig. 3. As above, we see no significant effect of the reading alone but a significant effect of pronoun-before-name on the name reading. We again see a significant effect of pronoun-before-name with the distractor reading, showing that only the cataphoric reading is dispreferred. We also see a significant effect of strong vs. weak, seen in Fig. 4.

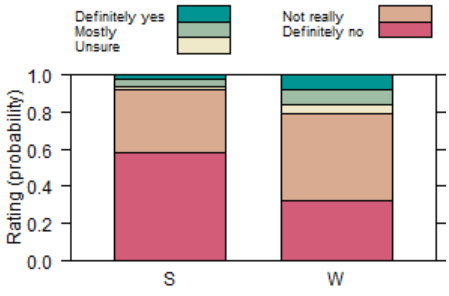

Fig. 4: Strong vs. weak in names

Conclusion. We present a novel experimental paradigm to measure strong and weak crossover. We find a significant difference in meaning availability between the two, contra Kush (2013) who did not find a difference using acceptability judgements. This supports theories which distinguish strong and weak crossover such as Koopman and Sportiche (1982), Safir (1984) or Ruys (2000), as opposed to unified accounts such as Reinhart (1983) or Safir (2004). We further find that proper names display a significant crossover effect similar to *wh*-crossover, supporting Rule I (Grodzinsky and Reinhart, 1993) and its derivatives. However, this is significantly less severe in weak configurations, suggesting that Rule I is not sufficient in these cases. This leaves an open theoretical question. More broadly, we propose a robust, adaptable methodology to test disputed cases of crossover, including relative clauses in English and French (Postal, 1993) and variation in weak crossover across languages (Bresnan, 1998; Lyu, 2017; i.a.).

Figures.

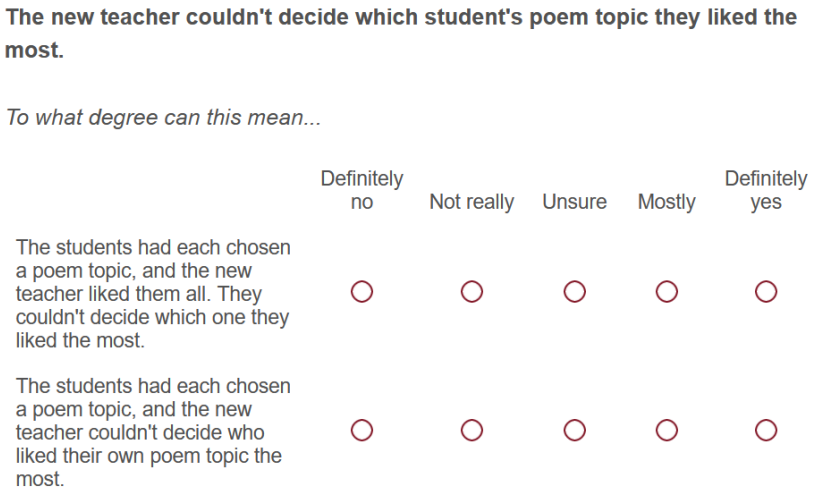

Fig. 5: Secondary strong crossover sentence in the "paraphrase" design (Experiment 1)

Parameter	Odds ratio	p-value
***Wh*-crossover**		
Distractor NP (reading)	–	$p = 0.14$
wh…pronoun…gap	0.33	$p < 0.05$
wh…pronoun…gap * Distractor NP	4.61	$p < 0.05$
Strong vs. weak	2.19	$p < 0.05$
Strong vs. secondary strong	–	$p = 0.30$
Proper names		
Distractor NP (reading)	–	$p = 0.08$
pronoun…name	0.06	$p < 0.05$
pronoun…name * Distractor NP	133.76	$p < 0.05$
Strong vs. weak	2.90	$p < 0.05$

Table 1: Model parameters for *wh*-crossover and proper names

References.

• Bjorkman, B. M. (2017). "Singular they and the syntactic representation of gender in English". *Glossa: a journal of general linguistics* 2.1 • Bresnan, J. (1998). "Morphology competes with syntax: Explaining typological variation in weak crossover effects". *Is the best good enough?: optimality and competition in syntax* • Chomsky, N. (1976). "Conditions on Rules of Grammar". *Linguistic Analysis* 2.4 •Christensen, R. H. B. (2019). *ordinal—Regression Models for Ordinal Data* • Conrod, K. (2019). "Pronouns raising and emerging". PhD Thesis • Grodzinsky, Y. and T. Reinhart (1993). "The Innateness of Binding and Coreference". *Linguistic Inquiry* 24.1 • Koopman, H. and D. Sportiche (1982). "Variables and the Bijection Principle". *Linguistic Review* 2.2 •Kush, D. (2013). "Respecting relations: Memory access and antecedent retrieval in incremental sentence processing". PhD Thesis. University of Maryland, College Park • Lasnik, H. and T. Stowell (1991). "Weakest crossover". *Linguistic Inquiry* 22.4 • Lyu, J. (2017). "Weak Crossover in Chinese—now you see it, now you don't". *Proceedings of the 17th Texas Linguistic Society* • Postal, P. M. (1971). *Cross-over phenomena* • Postal, P. M. (1993). "Remarks on weak crossover effects". *Linguistic Inquiry* 24.3 • Reinhart, T. (1983). *Anaphora and semantic interpretation*. Croom Helm linguistics series • Ruys, E. G. (2000). "Weak Crossover as a Scope Phenomenon". *Linguistic Inquiry* 31.3 • Safir, K. (1984). "Multiple variable binding". *Linguistic Inquiry* • Safir, K. (2004). *The Syntax of (In)dependence*