**Can slurs be used without being mentioned? Evidence from an inference judgement task**

**Background** Different types of content behave differently under ellipsis (see, e.g., Esipova 2019 for a brief overview and references therein). Thus, not-at-issue inferences that are an inextricable part of an item's lexical meaning can't be ignored in elliptical environments that require some form of identity with an antecedent instance of that item. This is true, e.g., for presuppositions of verbs encoding a stage of an event (*start*, *stop*, *continue*...), factive predicates (*know*, *regret*...), etc.:

(1) Pam stopped smoking, {but Kim didn't / and Kim did, too / and so did Kim}.
(i) $\rightarrow$ Pam used to smoke. (ii) $\rightarrow$ Kim used to smoke.

On the other end of the spectrum are pure expressives, which are always ignored under ellipsis:

(2) A: Did you bring a fucking gun to my house?
B: No, I didn't. / Yes, I did. / Yes, I did so. / Yes, I brought one.
(i) $\rightarrow$ A is experiencing strong emotions. (ii) $\not\rightarrow$ B is experiencing strong emotions.

Now, there are a few potentially relevant differences between these two cases: (i) the target inference in (1) is part of the lexical meaning of the head of the recovered VP, but the target inference in (2) is contributed by an adjunct inside an NP that is in turn inside the recovered VP or (contestably) targeted by *one*; (ii) the presupposition of *stop* is a precondition for the antecedent in (1) to make sense, but that's not the case for the contribution of *fucking* in (2); and perhaps most importantly, (iii) acts of producing expressives like *fucking* are purely performative, i.e., the speaker achieves their goal (here, expressing their emotions) by virtue of producing a certain form (use via mention), and there is no way to achieve this goal without performing this act (no use without mention).

**Question** Slurs, however, are a more complex case: (i) the "prejudice inference" is part of the lexical meaning of a slur, which can be the head of the antecedent constituent targeted by different types of ellipsis (like *stop*, unlike *fucking*); (ii) despite that, this inference is not crucial for the at-issue content of the sentence containing the slur to make sense (unlike *stop*, like *fucking*); (iii) slurs can be used performatively (use via mention) and can even have a performative effect of offense by virtue of being uttered in the absence of such intent on the speaker's part (mention without use), but it is unclear if the prejudice inference can be preserved if a slur is recovered but not uttered (use without mention). In this paper, I look at paradigms like (3) to assess the effect of different factors on the presence/strength of the prejudice inference and thus shed further light on the nature of this inference (the exchanges are set in a fictional universe where humans co-exist with centaurs, dwarves, elves, orcs, etc. and happen in the context of a criminal investigation):

(3) a. *Context: 'Tusky' is a slur for orcs.*
Detective: Did you see a tusky?
Witness: Yes. ('Bare') / Yes, I did. ('VPE') / Yes, I saw one. ('One') / Yes, I saw a tusky. ('Slur') / Yes, I saw an orc. ('Nonslur')

b. *Context: 'Tusky' is a slur for orcs. This slur can also be used as a verb meaning 'to crawl' (for any race), because orcs are stereotyped as living in caves and, thus, having to crawl through narrow spaces all the time. The detective is asking a question about a human.*
Detective: What happened next? Did he tusky under the table?
Witness: Yes. ('Bare') / Yes, he did. ('VPE') / Yes, he did so. ('So') / Yes, he tuskied under the table. ('Slur') / Yes, he crawled under the table. ('Nonslur')

*Question: How likely do you think that this witness is prejudiced against orcs?*

**Hypothesis** I hypothesized that the "prejudice likelihood" inferred from responses like those in (3) is gradient and is affected by several syntactic, semantic, and pragmatic factors: 1. Maintaining that slurs do have performative effects, I expected the likelihood to be highest when the witness utters the slur themselves ('Slur'), ostensibly both using and mentioning it. 2. I also expected the likeli-

hood to be lowest when the witness tacitly corrects the detective by using the neutral term instead ('Nonslur'), thus, neither mentioning nor using the slur and, furthermore, indirectly challenging the detective on their use of the slur in an attempt to minimize complicity (see, e.g., Cepollaro 2020 and references therein on unchallenged slurs). 'Nonslur' responses are, thus, expected to have lower prejudice likelihood ratings than any of the elliptical responses. 3. Finally, I hypothesized that the prejudice component of slurs is not exclusively performative, i.e., it does allow for use without mention. Thus, when the slur is obligatorily recovered, e.g., when it is the head of the constituent targeted by a proform requiring lexical identity with said head (more obviously in 'One' for nouns; less obviously in 'VPE' and 'So' for verbs), the prejudice likelihood is expected to be higher than in elliptical responses where the slur is not necessarily recovered ('Bare' and 'VPE' for nouns; 'Bare' for verbs). So, if all parts of the hypothesis are correct, we expect the following picture:

(4) Predicted prejudice likelihood ratings (from lowest to highest)
   a. Nouns: 'Nonslur' $<$ 'Bare'/'VPE' $<$ 'One' $<$ 'Slur'
   b. Verbs: 'Nonslur' $<$ 'Bare' $<$ 'VPE'/'So' $<$ 'Slur'

**Methods** The experiment involved 10 conditions (2 parts of speech, with 5 response types for each). Each participant saw 2 trials per condition and 2 attention checks (22 trials total); the trials looked similarly to (3). Participants assessed the prejudice likelihood by dragging a slider on a pseudo-continuous scale (mapped to 0–100) from 'Not at all likely' to 'Very likely'. Participants were recruited on Prolific (final $N$ = 128) and paid £1.25 for completing the task.

**Results** The results are visualized in Fig. 1. The statistically significant contrasts fully matched the prediction in (4a) for nouns, but only partially matched the prediction in (4b) for verbs:

(5) Statistically significant contrasts in prejudice likelihood ratings (from lowest to highest)
   a. Nouns: 'Nonslur' $<$ 'Bare'/'VPE' $<$ 'One' $<$ 'Slur'
   b. Verbs: 'Nonslur' $<$ 'Bare'/'VPE'/'So' $<$ 'Slur'

**Discussion** The results for noun slurs corroborate all parts of the original hypothesis, suggesting that slurs do make performative contributions (which is why actually saying a slur gives rise to a stronger effect than using it without mentioning), but are not exclusively performative (which is why the prejudice inference still persists to some extent if the slur is recovered, but not uttered). This calls for a hybrid analysis for slurs that doesn't reduce their prejudice component to just a presupposition (as in Schlenker 2007) or just a performative effect on the context (as in Potts 2007). The results for verb slurs corroborated parts 1 and 2, but not 3 of the hypothesis, possibly because: (i) in the absence of perfect English counterparts, verb slurs were harder to intuit about, so the data were more noisy, (ii) due to (i) and the less direct link between the meaning of a verb slur and the targeted group, the contrasts were overall less pronounced, and (iii) the identity requirements for verbs in VPE and *do so*-replacement are less clear than for nouns in *one*-replacement.

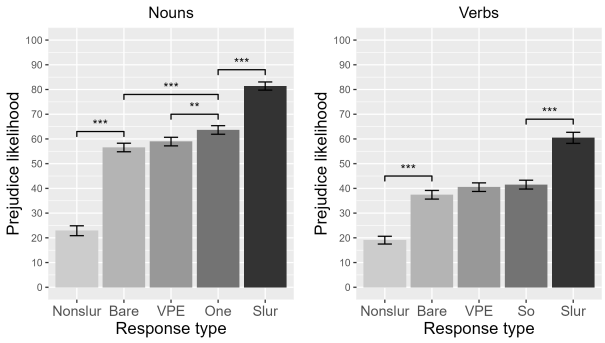


Fig. 1: Bar charts showing mean prejudice likelihood ratings of different types of responses to antecedent utterances with noun and verb slurs, with SE and key significant contrasts indicated.

**References** Cepollaro. 2020. isbnsearch.org/isbn/9781793610522 Esipova. 2019. ling.auf.net/lingbuzz/004676. Potts. 2007. doi.org/10.1515/TL.2007.011 Schlenker. 2007. doi.org/10.1515/TL.2007.017