Relying on a combination of corpus and experimental evidence, they authors argue that the status of the prejacent of evaluative adjective sentences is linked in a principled way to the question that the utterance is taken to address. Having individuated two crucial entailments of this type of sentences — the (truth of the) prejacent and the (truth of the) Generalization — the authors test three predictions of their proposal: (a) that the projectivity of the prejacent is influenced by the focus marking of the utterance; (b) that the stronger the inference to the truth of the generalization from the common ground, the less projective is the prejacent; (c) that the stronger the inference to the truth of the generalization from the common ground, the more at-issue is the prejacent. The first prediction is tested via a corpus study; the second and the third are tested via two experiments.

Overall, I find this submission quite strong. The topic is of high theoretical interest in the literature on projective content and secondary entailments; the methodology is sound and empirically rich; and the article is well written and well structured, even though I think the prose and the layout of the paper could use some reorganization. In sum, think this would make a good contribution to *Glossa.* I recommend that the article **be accepted with (moderate) revisions.** My main suggestion concerns the organizational aspects of the paper. In particular, the article in its current form is quite long. While part of this is justified by the need to spell out in detail all the steps of a variety of different methodologies — which the authors do quite well — I also think that the prose could be made a bit more tight, improving the readability of the article.

Here are some suggestions. Please keep in mind that these strictly reflect my own personal judgment, and are by no means *the only possible way* for shortening the article. I think it's more important that the authors find a way of tightening up the article than that they follow these specific recommendations to achieve this goal.

1- Do we really need a whole section for Experiment 1? As far as I understand, this study merely shows that not all prejacents of EAs project to the same degree. It's definitely an important starting point, but it seems to me that it could be presented as part of a background literature review, rather than as a stand alone study. After all, this experiment is setting up the stage by presenting a puzzle to Karttunen's original theory; it does not test the predictions of the novel account, which is outlined in Section 3. While I understand the importance of providing a detailed discussion, I think these findings could be condensed in a shorter background section that explains why K's theory falls short of explaining the behavior of evaluative adjective sentences. In my view, this would be sufficient to show the need for a novel account, allowing for a smooth transition to section 3.

2- I think Section 3 can be made more coincise. In particular, I am not sure I entirely follow why the authors linger on the discussion of prosody on page 15. To be sure, it's very true that prosodic factors represent a crucial window into the discourse status of any sentence. But, as the authors themselves admit, none of the studies manipulate prosody as a factor, and neither is prosody central too any of the three central theoretical predictions (except for (a); see below for further comments). Again, I think this digression takes away from the flow of the paper; perhaps it could be reduced to an extended footnote. If, by contrast, I am underestimating the importance of this part to the take away message of the paper, I would suggest that the authors be more explicit about the role of prosody in their theoretical claims.

3- While I think experiment 2 and 3 are convincing, I'm a little hesitant to say the same for the corpus study. To put it better: all the considerations and suggestions made by the authors are

reasonable; but as they admit, it is hard to collect definitive evidence without a proper comprehension and/or production study, one that ideally also takes prosodic factors into consideration . As it stands, I think that section 4.1, if properly reduced, could be fruitfully incorporated into a good "Avenue for further research" section. This would allow the authors to emphasize the contribution of the two experiments, while making the paper easier to follow.

Finally, on a conceptual level, I think the paper could benefit from more engagement with big picture issues about entailments and projective content. In particular, I was left wondering about the amenability of the proposed analysis to be extended to other cases of prejacents that have been previously claimed to be presupposed by the embedding verb. For example, how does this connect to the current debate on the status of the complements of emotive factives? (e.g., *I'm happy that p.*) Would a notion like the Generalization be helpful to shed light on these cases as well? It seems to me that, in principle, it could yield crucial insights, especially with respect to the puzzle of mood selection under such predicates (indicative vs. subjunctive in Spanish, Italian and Greek, see Giannakidou and Mari 2016). Based on these, it seems to me that the more focused the adjective, the more likely it is the prejacent to be in the indicative; and the more focused the prejacent, the more likely it is to be in the subjunctive. Of course, I do not expect the authors to engage with this issue specifIcally. But this is just an example of how the proposed theory could be extended to other domains. It would be good to walk away from the paper with a clearer sense of how the theory outlined in the article could be used to think about other domains. This would definitely make the paper stronger in my view.

Other comments:

- p.23: The authors write: "in losing her wallet, Katie was fortunate". I believe it should be *unfortunate* here; a minor typo which could severely impair the understanding of the experimental predictions! :-)

- Figure 7 is a little overinformative. Perhaps it can be broken down into two figures: one showing the average for the two conditions, and one showing the difference between the specific predicates.

- In general, I was a bit confused by the visualization of the data. While I do think that the differences between predicates are important, it would be useful to get a sense of the average across predicates for the manipulated condition. This was only shown for Experiment 3, though.