Review of Real-time lexical comprehension in young children learning American Sign Language
This is a great paper, and shows some important similarities in deaf and hearing ASL learners. I find fascinating the idea that because children learning ASL must rely on vision to process linguistic information and look at named objects, it is possible that the basic learning mechanisms, such as the allocation of visual attention might differ from hearing children learning spoken language in how they support lexical development (stated on page 5, bottom).

I think I just have some initial confusion about vagueness of terms such as "modality" and "visual attention". I understand the introduction starts with the big picture, yet some terms still need defining up front. For example, "processing efficiency" is mentioned on the first page, and it is not yet defined, and I have no idea what it refers to yet. I mention below some areas where the introduction needs more motivation (linking ideas). Also, I'm not clear on the *mechanisms* driving the predictions.

Moreover, the discussion feels rather short and amiss -- what are the implications for education, deaf children? Are they able they speak to the issue about ASL being is a natural language that is processed in the same way as spoken languages are? And that their study suggests common neural architectures for language processing and that ASL is not "bad" for the deaf child's brain?

1. First sentence of abstract is confusing because if children are interpreting *spoken* language, yet linguistic information drives rapid shifts in *visual* attention, then what are they looking at if they're interpreting *spoken* language.
2. And the last sentence of the abstract is confusing too because at this point it is so very vague to refer to "visual attention", at least inform the reader if you mean visual attention to look at a signer or to look at possible referents in the world (that may or may not be present).
3. The first paragraph of the introduction is a little vague, and I'm not sure how things link up – for example, the last sentence in this paragraph on line 34 – what does this have to do with eye movements, mentioned in the previous sentence?
4. Page 3, I'm not sure what you mean by "modality", given that this study is entirely in the visual modality: What do you mean by, "Do the findings from spoken language reflect language-general phenomena or are they specific to the auditory modality?" I'm guessing you are referring to looking at the referent before the full word is uttered, but that's a pretty superficial prediction about how processing speech and sign differ, given that the time course of speech and signs differ, besides modality. Basically, I just think that question isn't clear yet.
5. On p. 6 lines 24-36 there is a nice prediction (the second out of two) - possibly ASL learners, like spoken language users, would shift visual attention as soon as there's enough linguistic information. But I feel this begs qualification about basic human vision - is it possible that ASL learners could shift attention to the referent but nonetheless still perceive the signing in the parafovea(?) ... that's certainly plausible so that's a third prediction (or at least a qualification about the second prediction) and should be

mentioned... This also comes up on pg. 23 line 52-55 – maybe children can still see the sign in the parafovea while looking at the target object.

6. Page 6 line 24 - why are ASL LEARNERS compared to spoken language USERS? You've set up the hypothesis to sound like you are comparing ASL processing vs spoken language processing, to see if both show incremental patterns. I just wanted to point out that oddness, because that's not what you are doing.
7. The 2nd research question focuses on age-related differences, so I'd like some review of the literature on how age influences visual attention in hearing kids vs adults. The 4th research question focuses on "expressive vocabulary development" as well as age -- So what's the literature on how vocabulary size influences visual attention?
8. Measure of ASL expressive vocabulary size - it says "developed specifically for this project" so clearly some further adaptions/modifications were made on top of Anderson & Reilly which itself is a modification of the CDI for ASL users. So it needs more qualification, what was changed, why wasn't Anderson & Reilly sufficient by itself? How would others replicate this?
9. "Matched for visual salience" means what?
10. P. 10 line 26 "chosen based on naturalness" – how?
11. P. 10 line 33-36 - "minimal phonological overlap" really needs to be defined. How did each pair overlap? location, movement, handshape? Be specific. What was the definition of "minimal" - 1, 2, 3 parameters, etc.? Can they reference it with ASL-LEX?
12. Was the trial structure just like how it is presented for auditory LWL tasks (end of P. 10)? If not, explain the differences
13. What if children didn't orient to the signer during that 1s still frame? Throw out the trial? (ah it's explained later p. 13 line 44)
14. Computing measures of sign onset/offset by adults (explained on page 11). I would really like to see more information on this...how much variation was there among the 10 adult signers in figuring out sign onset? And really they also measured the time it took to press the button, so their definition of "sign onset" time is actually sign onset + button-pressing time.
15. How would the addition of auditory sensory experience change the way in which CODA children process ASL vs deaf children? They're both looking at the same language, so how could the simple fact that one group can hear (and sound isn't a part of this study) and the other cannot change how they process ASL visually (they're both natively exposed to ASL). I just want this to be clearer.
16. Why do adult signers show larger CI's for proportion of looking (Figure 2), compared to children? I'd expect it to be the opposite.
17. Pg 20 Line 54 - "like children learning spoken language, ASL learners improve...over the second and third years of their life..." but this age-based development is never really fleshed out in the introduction. Ditto for pg 22 line 29
18. Pg 23 line 19: "While the hearing children could use vision and hearing to process incoming information, this experience did not change the timing of gaze shifts during ASL comprehension as compared to their deaf peers." - I do not accept that...they could NOT use hearing to process incoming information, there was no voiceover or audio.

Why should they even expect a difference between deaf and CODA when viewing ASL? This part just needs more theoretical elaboration and motivation.

Minor:

1. Page 5, Are you sure to say that lexical decision tasks are not “on-Line”? I am quite sure they are called online, at least the cross-modal lexical decision task with priming effects is considered a measure of real time lexical access.
2. Why Figure 2 and 4 b/w but other figures are color?
3. P. 9 line 12-15 - "visual distractions" twice in one sentence
4. How far away was the baby from the screen?