

## 1 Question

Average ASR error rates show children are recognized worse. I ask where those extra errors happen.

The analysis groups ASR mistakes by word category: nouns, verbs, pronouns, adverbs, prepositions, and related groups.

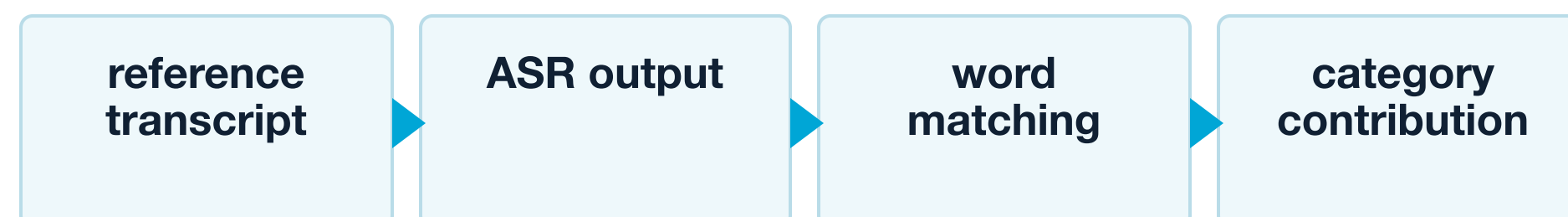
**Compare:** native Dutch children and teenagers in read and HMI speech.

**Measure:** which categories add most to the child higher error rate.

**Check:** whether category results come from specific words.

## 2 Data and method

Corpus	JASMIN-CGN
Groups	native Dutch children and native Dutch teenagers
Speakers	135 total: 72 children and 63 teenagers
Speech	read speech and human machine interaction speech
Reference words	116,687 read and 18,794 HMI
ASR systems	Google Chirp 2 and Whisper-large-v3
Scoring	WER from the existing ASR evaluation



The existing ASR scores give the overall error rate. I add word matching so each substituted or deleted reference word can be assigned to noun, verb, pronoun, or another category.

**category contribution** asks how many percentage points one word category adds to the error rate difference. A category matters when it is common, often wrong, or both.

## 3 Main result

HMI speech has higher absolute WER, but read speech has the larger child minus teenager difference.

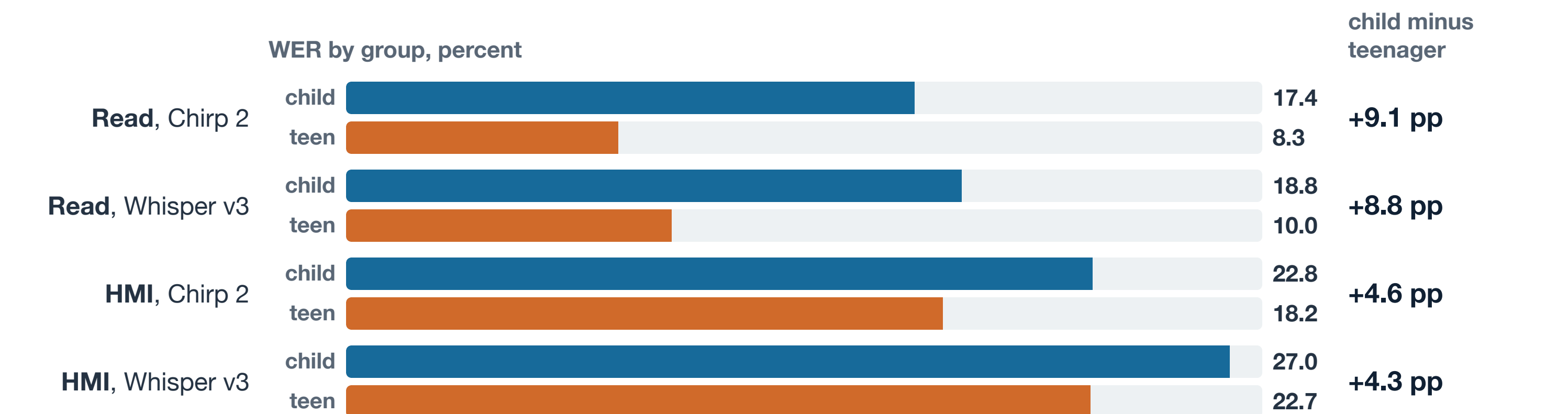
**Read speech about +9 pp**

larger child minus teenager WER difference, concentrated in nouns, verbs, and pronouns

**HMI speech about +4 to +5 pp**

higher absolute WER, but the group difference is smaller and spread over more categories

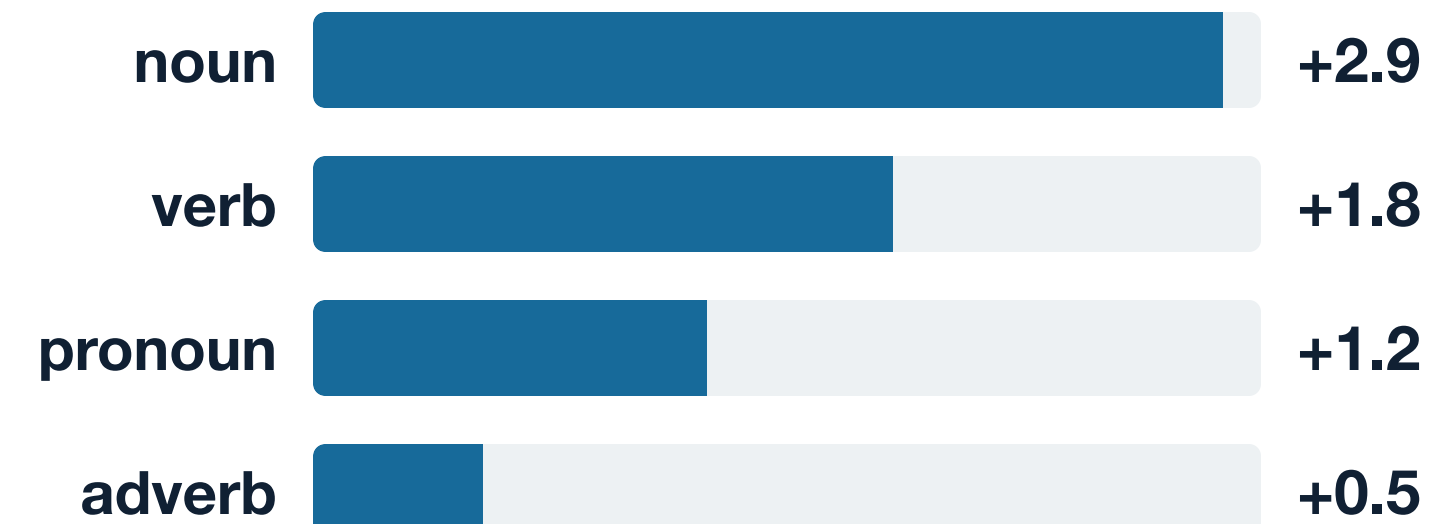
Overall error rate



Where the child higher difference appears: category contribution from substitutions and deletions, percentage points of all words. Values average Chirp 2 and Whisper v3.

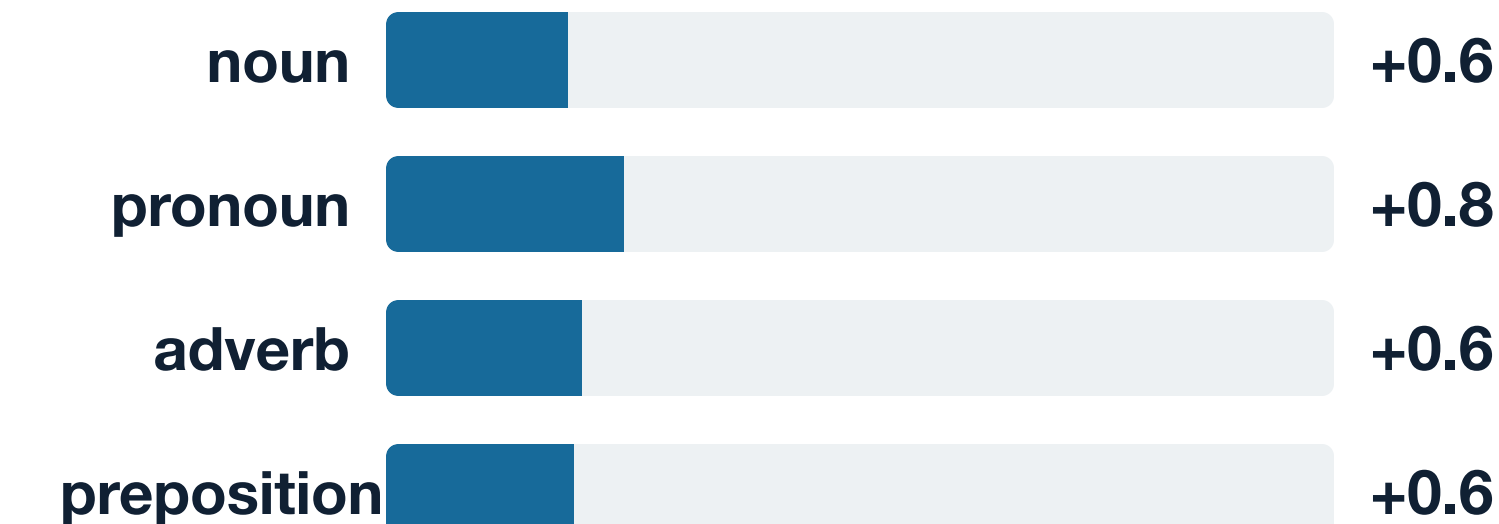
### Read speech

largest category contributions



### HMI speech

smaller and more distributed



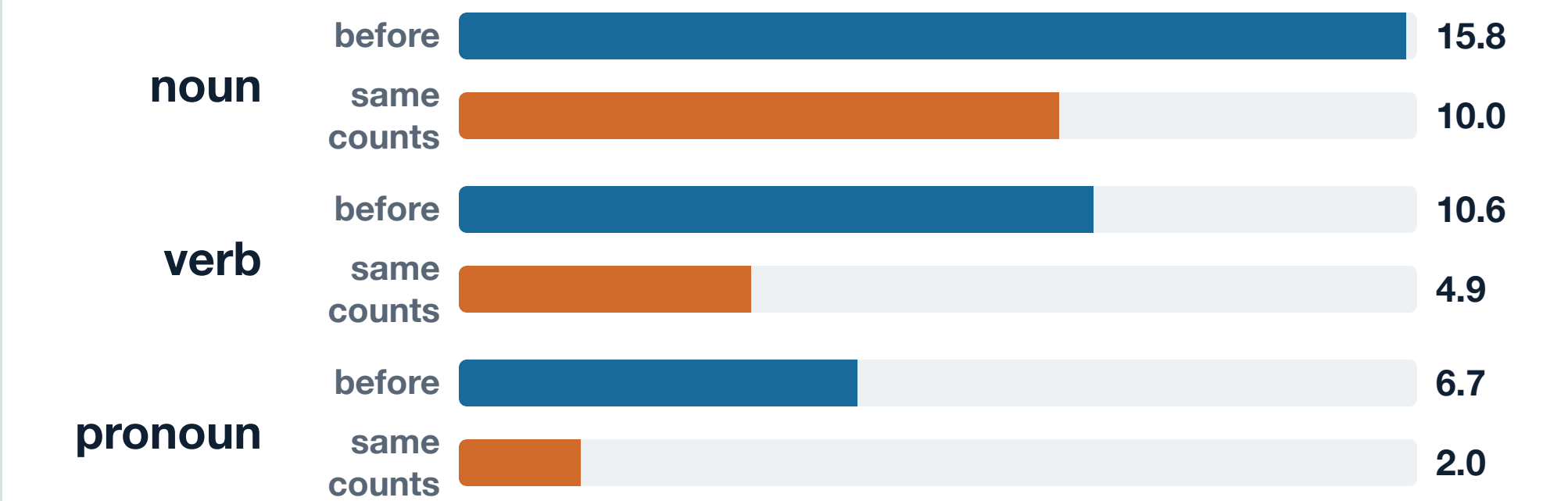
**The substitution and deletion part of the read speech difference is mostly in nouns, verbs, and pronouns. In HMI speech, no single word category dominates.**

**Interpretation:** word categories locate where the extra errors appear. The next checks test whether the read speech pattern comes from category counts, specific words, or both.

## 4 What remains with same category counts?

**Why this matters:** children and teenagers may not read sentences with the same counts of nouns, verbs, and pronouns. This control compares shared category count patterns, not the same exact words.

Read speech category error rate difference, percentage points



**With the same category counts, verb and pronoun differences shrink strongly. The noun error rate difference remains largest.**

## 5 Which words drive the read result?

These percentages are not word frequencies. They show which specific words increase the read speech difference for a category.



### Main conclusion

- Word categories locate the read speech difference.
- Specific words limit how far categories can be interpreted.
- Nouns and verbs mostly come from words found only in child read references.
- Pronouns are the clearest shared word category result.

Main caveat: the read speech comparison cannot fully separate speaker group from assigned reading text and vocabulary.