German-speaking 10-12 year-olds revise their noun predictions when reading prediction-inconsistent gender-marked articles

Adult language users leverage information provided by the sentence context to predict semantic [1] and, sometimes, form-related features [2] of upcoming linguistic input, such as a words grammatical gender [3]. Disconfirmed form predictions result in behavioral prenominal prediction mismatch effects where adult comprehenders slow down when reading prediction-inconsistent gender-marked articles [4]. However, it is not clear whether prenominal prediction mismatch effects index A), an early revision process that dampens the prediction of the noun, or B) a form-mismatch effect because the gender-marked article itself was predicted [5]. Children, too, use gender cues to infer upcoming referents, but the age at which children do so is language-specific: Spanish-speaking children use gender predictively at a much earlier age than children of Czech or German. In addition, whether children use gender predictively at all is determined by individual differences. However, predictive use of gender in children has exclusively been investigated by means of the visual-world eye-tracking technique, a method that often uses artificially prolonged auditory signals, failing to adequately capture the timing aspects associated with natural language processing [6]. Here, we ask two questions: First, do children show a prenominal prediction error cost, like adults do, under more natural, self-paced processing circumstances? Second, does the prediction mismatch effect reflect revision of the noun, or a prediction error associated with the form of the gender-marked article? If comprehenders use article gender to dampen their predictions of the upcoming noun, early gender mismatch should result in smaller predictability effects at the noun.

German-speaking 7-12 year-olds read prediction-matching or mismatching definite articles in a self-paced reading paradigm. All sentences (see (1) below) constrained expectations towards a German gender-marked noun and its definite article, while half of the sentences continued in an unpredictable noun from a different grammatical class and a spill-over region that continued the sentence plausibly. In order to allow for more time of prenominal gender mismatch to unfold, an article spill-over region was inserted. Prediction effects were measured in two critical regions, a prenominal and postnominal region, and we corrected the p-value threshold for statistical significance according to the amount of comparisons per region. Individual difference measures included age, processing speed (Digit Symbol test), vocabulary (PPVT) and executive control (verbal fluency S).

Figure 1 (left) shows reading times in the prenominal and postnominal regions. Gamma-link models fitted to untransformed reading times showed robust prediction-mismatch effects in the postnominal region, such that children read unpredictable nouns and their spill-over words more slowly than predictable ones (all p's < .01; see Table 1 for model syntax). However, there was no convincing evidence for a gender mismatch effect in the prenominal region: On the A+3 word, a numeric trend that indicated longer reading times for prediction-inconsistent words seemed to be masked by large individual differences (t = 1.83, p = .07). Consequently, we ran follow-up models which allowed prenominal mismatch effects on the A+3 word to vary with age, processing speed, and productive/receptive vocabulary, and we compared goodness of fit between models (see Table 2). The results showed that age was the best predictor of prenominal gender mismatch effects, such that older children showed larger prenominal mismatch effects (t = 2.23, p = .02; see Figure 1). Figure 1 additionally indicated that older children used prediction-inconsistent articles to revise their noun predictions: Following the early gender mismatch effect at the A+3 word, older children showed a smaller mismatch effect at the noun compared to younger children (cf. mid and right panels of Figure 1). Indeed, when we ran a follow-up model that allowed the interaction term between predictability and age to vary depending on word type (i.e., A+3 vs Noun, sum-coded with -1 and 1), a significant three-way interaction surfaced (t = 2.65, p = .01).

We found no consistent prenominal gender mismatch effect at the group level, suggesting that German-speaking children do not routinely use gender cues when processing language under more time-restricted conditions. However, older children in our sample did use gender cues, and they used them predictively to revise their expectations for the upcoming noun. Hence, our data attest to a trade-off between early gender mismatch at the article and late integration difficulty at the noun in older child readers.

(1)										
	Obwohl	sie	versprochen	hatte	es	nicht	zu	tun,	verrät	Annika …
	Even though	*she*	*promise*	*had*	*it*	*not*	*to*	*do,*	*reveals*	*Annika …*

	Prenominal Region				**Postnominal Region**			
	Article	A+1	A+2	A+3	Noun	N+1	N+2	N+3
Predictable:	**das**	bisher	gut	gehütete	**Geheimnis**	**ihrer**	**besten**	Freundin.
	the$_{Acc\text{-}Neut}$	*previously*	*well*	*kept*	*secret*	*of her*	*best*	*friend.*
Unpredictable:	**die**	bisher	gut	gehütete	**Position**	**ihrer**	**besten**	Freundin.
	the$_{Acc\text{-}Fem}$	*previously*	*well*	*kept*	*position*	*of her*	*best*	*friend.*

Table 1. Model syntax. We used gamma link models as they capture the natural RT distribution and avoid distortions introduced by log-transforming RTs [7].

glmer(RT ~ predictability + scale(trial) + scale(length) + scale(frequency) + scale(position) + (1+condition|subject) + (1|item), data=dat, family = Gamma(link = "log"), control = glmerControl(optimizer = "bobyqa"))

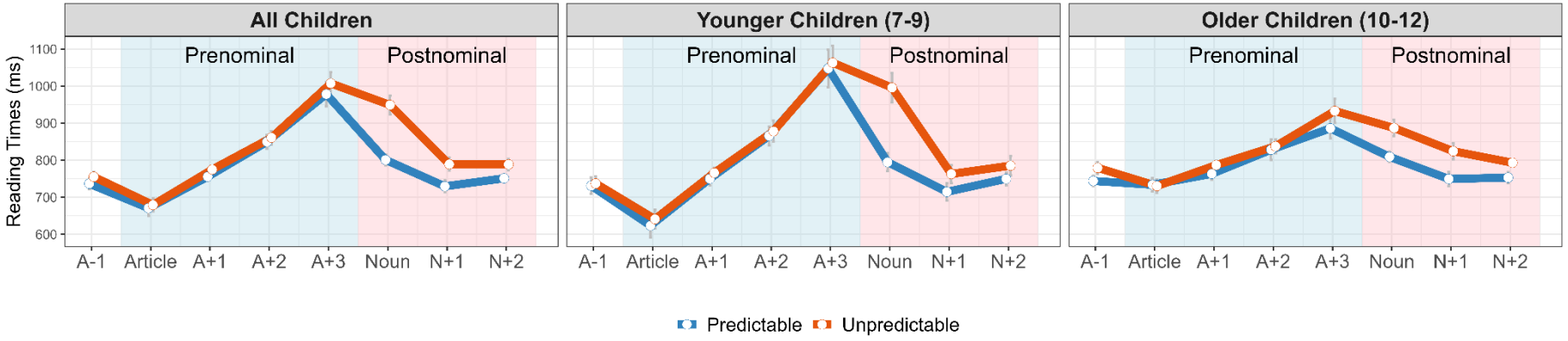
Table 2. Model syntax and improvement in fit for individual-difference models compared to base A+3 model. All ID models were fit on the exact same number of data points.

glmer(RT ~ predictability * scale(ID-Variable) + scale(trial) + scale(length) + scale(frequency) + scale(position) + (1+condition|subject) + (1|item), data=dat, family = Gamma(link = "log"), control = glmerControl(optimizer = "bobyqa")).

Model	AIC	Fit Improvement vs Base
Base (see Table 1)	22480	
Base * Age	**22472**	**8**
Base * Vocabulary	22482	-2
Base * ExecutiveCtrl	22479	1
Base * ProcessSpeed	22477	3

Figure 1

Reading times for the full sample of children (left panel), and split out by age group (mid and right panel), based on a median split of the continuous age variable.



[1] Federmeier, K. D., & Kutas, M. (1999). A rose by any other name: Long-term memory structure and sentence processing. JML, 41(4), 469-495. [2] Ito, A., Corley, M., Pickering, M. J., Martin, A. E., & Nieuwland, M. S. (2016). Predicting form and meaning: Evidence from brain potentials. JML, 86, 157–171. [3] Van Berkum, J. J., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: evidence from ERPs and reading times. JEP:LMC, 31(3), 443–467. [4] Haeuser, K. I., Kray, J., & Borovsky, A. (2022). Hedging bets in linguistic prediction: Younger and older adults vary in the breadth of predictive Processing. Collabra: Psychology, 8(1), 36945. [5] Fleur, D. S., Flecken, M., Rommers, J., & Nieuwland, M. S. (2020). Definitely saw it coming? The dual nature of the pre-nominal prediction effect. Cognition, 204, 104335. [6] Huettig, F., & Guerra, E. (2019). Effects of speech rate, preview time of visual context, and participant instructions reveal strong limits on prediction in language processing. Brain Research, 1706, 196–208. [7] *Lo, S., & Andrews, S. (2015). To transform or not to transform: Using generalized linear mixed models to analyse reaction time data.* Frontiers in Psychology, 6, 1171.