Supplementary materials

Behavioral data analyses

The statistical analysis of response time (RT) data was performed in R Core Team (2020). The RT analysis conformed to a 2 (Language: Polish [L1] vs. English [L2]) × 4 (Sentence type: semantically correct vs. semantically incorrect vs. stereotypically congruent vs. stereotypically incongruent sentences) × 2 (Gender: Females vs. Males) design, with Language and Sentence type as within-subject factors and Gender as a between-subject factor. The analysis was based on correct responses only. To reach normal distribution, the RT data was log-transformed. The RTs below 200 ms and above 1,500 ms as well as those falling outside the value of 1.5 interquartile range were discarded from further analyses, altogether resulting in the final rejection of 0.14% of behavioral data. The RT data were analyzed with linear mixed-effects models (Baayen et al., 2008; Jaeger et al., 2008; Barr et al., 2013; Barr, 2013), following the same procedures as for the ERP analysis (see the main manuscript).

Behavioral data: Response times (RTs)

The analysis of RT data showed a fixed effect of Sentence type, whereby semantically incorrect sentences were responded to faster than semantically correct, stereotypically congruent, and stereotypically incongruent sentences. Then, stereotypically incongruent sentences were responded to faster than semantically correct and stereotypically congruent sentences. Finally, stereotypically congruent sentences were responded to faster than semantically correct sentences (see **Table 1** and **Table 2** below**)**.

Table 1. Mean response times (with 95% confidence intervals)

	Sentence types	**Sentence types: Polish**	**Sentence type: English**
[1]	*M* = 675.4 ms [670.5, 680.5]	*M* = 668.3 ms [661.1, 675.9]	*M* = 682.2 ms [674.2, 689.4]
[2]	*M* = 600.1 ms [595.3, 605.8]	*M* = 592.4 ms [585.3, 599.3]	*M* = 609.2 ms [601.4, 617.6]
[3]	*M* = 658.9 ms [653.2, 663.0]	*M* = 654.3 ms [654.6, 656.9]	*M* = 661.5 ms [654.4, 668.0]
[4]	*M* = 650.3 ms [644.7, 655.1]	*M* = 649.5 ms [642.2, 656.3]	*M* = 650.8 ms [643.9, 658.3]
	Polish (L1)	**English (L2)**	
Females	*M* = 628.4 ms [623.6, 633.8]	*M* = 647.3 ms [642.1, 653.7]	
Males	*M* = 654.1 ms [648.7, 659.5]	*M* = 654.4 ms [648.3, 659.9]	

Sentence types: [1] – semantically correct; [2] – semantically incorrect; [3] – stereotypically congruent; [4] – stereotypically incongruent

There was also a Sentence type × Language interaction. *Post-hoc t*-tests showed the pattern mirroring the general Sentence type reported above for English (L2) sentences. In contrast, there was no difference in RTs between Polish (L1) stereotypically congruent and incongruent sentences (see **Table 1** and **Table 2** below**)**.

Finally, the analysis also yielded a Gender × Language interaction. *Post-hoc t*-tests showed faster RTs for Polish (L1) relative to English (L2) sentences in females, with no between-language difference in RTs in males (see **Table 1** and **Table 2** below**)**.

Table 2. The results of statistical analyses.

	Sentence types	**Sentence types: Polish**	**Sentence type: English**
[1] vs. [2]	*b* = .13, SE < .01, *t*(252) = 26.01, *p* < .001	*b* = .13, SE = .01, *t*(252) = 18.75, *p* < .001	*b* = .13, SE = . 01, *t*(252) = 18.04, *p* < .001
[1] vs. [3]	*b* = .02, SE < .01 *t*(252) = 4.99, *p* < .001	*b* = .02, SE = . 01, *t*(252) = 2.98, *p* < .001	*b* = .03, SE = . 01, *t*(252) = 4.06, *p* < .001
[1] vs. [4]	*b* = .04, SE < .01, *t*(252) = 7.99, *p* < .001	*b* = .03, SE = ., *t*(252) = 4.49, *p* < .001	*b* = .05, SE = . 01, *t*(252) = 6.78, *p* < .001
[2] vs. [3]	*b* = .10, SE < .01, *t*(252) = 21.20, *p* < .001	*b* = .11, SE = . 01, *t*(252) = 15.85, *p* < .001	*b* = .10, SE = . 01, *t*(252) = 14.16, *p* < .001
[2] vs. [4]	*b* = .09, SE < .01, *t*(252) = 18.01, *p* < .001	*b* = .10, SE = ., 01 *t*(252) = 14.28, *p* < .001	*b* = .08, SE = . 01, *t*(252) = 11.24, *p* < .001
[3] vs. [4]	*b* = .02, SE < .01, *t*(252) = 3.06, *p* = .002	*b* = .01, SE = . 01, *t*(252) = 1.53, *p* = .126	*b* = .02, SE = . 01, *t*(252) = 2.78, *p* = .005

	Females	**Males**
Polish (L1) vs. English (L2)	*b* = -.03, SE = .01, *t*(296) = -2.92, *p* = .004	*b* < .01, SE = .01, *t*(298) = .29, *p* = .771

Sentence types: [1] – semantically correct; [2] – semantically incorrect; [3] – stereotypically congruent; [4] – stereotypically incongruent

References:

Baayen, R.H., Davidson, D.J., & Bates, D.M. (2008). Mixed-effects modeling with crossed randocts for subjects and items. *Journal of Memory and Language, 59,* 390–412.

Barr, D.J. (2013). Random effects structure for testing interactions in linear mixed-effects models. *Frontiers in Psychology, 4.*

Barr, D.J., Levy, R., Scheepers, C., Tily, H.J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68.*

Jaeger, T.F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59,* 434–446.