

1 Supplementary Materials: Early language experience in a Tselal Mayan village

2 Marisa Casillas¹, Penelope Brown¹, & Stephen C. Levinson¹

3 ¹ Max Planck Institute for Psycholinguistics

4 Author Note

5 Correspondence concerning this article should be addressed to Marisa Casillas, P.O.

6 Box 310, 6500 AH Nijmegen, The Netherlands. E-mail: Marisa.Casillas@mpi.nl

7 Supplementary Materials: Early language experience in a Tselta Mayan village

8 Full model outputs

9 In the main text we only report *significant* effects on two speech environment variables:
10 TCDS min/hr and ODS min/hr. Here in the Supplementary Materials we give the full model
11 output tables for each analysis, including re-leveled versions of each model to show all three
12 of the two-way contrasts between the three-level time-of-day factor (i.e., morning vs. midday,
13 morning vs. afternoon, and midday vs. afternoon). We also include, for each of the measures,
14 a histogram showing how each variable is distributed (i.e., because they are non-normal
15 and/or zero-inflated) and a figure showing the distribution of model residuals. For every
16 negative binomial model, we also include the full model output table and residual plots for
17 matching gaussian mixed-effects regressions which uses a logged dependent measure. Such
18 gaussian models with logged measures are an alternative solution to analyzing non-normal
19 distributions sometimes used in psycholinguistics, but are not suitable for the current data
20 given how our speech environment measures are distributed, particularly in the randomly
21 sampled clips (see, e.g., Figures 1, 7, 10, 13, 19). Overall, however, the gaussian models show
22 a qualitatively similar pattern of results. None of the gaussian model results are presented in
23 the main text—only here as supplementary information.

24 How to interpret the model output

25 All models were run with the glmm-TMB library in R (Brooks et al., 2017a, 2017b).
26 Note that, in the negative binomial regressions, the dependent variables have been rounded
27 to the nearest integer (e.g., 3.2 minutes of TCDS per hour becomes 3 minutes per hour in
28 the model).

29 The predictors in the models are abbreviated as follows: tchiyr.std = centered,
30 standardized target child age in months; stthr.tri = the start time of the clip as either
31 morning, midday, or afternoon; hsz.std = centered, standardized household size of the target
32 child; nsk.std = centered, standardized number of speakers present in the clip,

33 `aclew_child_id` = the unique identifier for each child. The predictors are sometimes
34 combined in two-way interactions, as shown below with a “:” separator between predictor
35 names (e.g., `tchiyr.std:nsk.std` = a two-way interaction of target child age and number of
36 speakers present).

37 In each model output table, the “component” shows what kind of model the estimate
38 derives from (e.g., the zero-inflated models include both a conditional “cond” set of
39 predictors, random effects, and zero-inflation “zi” predictors). The “term” is the estimated
40 predictor. The “statistic” is the estimated z -statistic for each predictor’s effect. The other
41 labels are self-explanatory.

42 As more data are added to this corpus, the analyses will also be updated, as will this
43 supplementary model information, all of which will be available online at:

44 https://middycasillas.shinyapps.io/Tseltal_Child_Language_Environment/.

45 **Target-child-directed speech (TCDS)**

46 **Random clips.** TCDS rate in the random clips demonstrated a skewed distribution
47 with extra cases of zero Figure 1. We therefore modeled it using a zero-inflated negative
48 binomial mixed-effects regression in the main text: results for the two models demonstrating
49 all pairwise effects of time of day are shown in Table 1 and Table 2. The residuals for the
50 default model (Table 1) are shown in Figure 2.

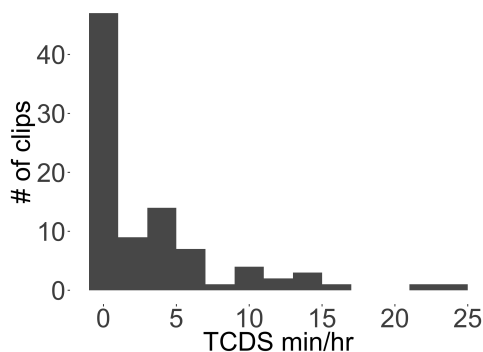


Figure 1. The distribution of TCDS rates found across the 90 random clips.

Table 1

Full output of the zero-inflated negative binomial mixed-effects regression of TCDS min/hr for the random sample, with midday as the reference level for time of day.

component	term	estimate	std.error	statistic	p.value
cond	(Intercept)	0.91	0.36	2.53	0.01
cond	tchiyр.std	0.60	0.36	1.68	0.09
cond	stthr.trimorning	0.83	0.40	2.09	0.04
cond	stthr.triafternoon	0.49	0.37	1.31	0.19
cond	hsz.std	0.01	0.22	0.04	0.97
cond	nsk.std	-0.12	0.16	-0.75	0.45
cond	tchiyр.std:stthr.trimorning	-0.28	0.39	-0.73	0.47
cond	tchiyр.std:stthr.triafternoon	-0.85	0.38	-2.26	0.02
cond	tchiyр.std:nsk.std	0.57	0.19	2.95	0.00
zi	(Intercept)	-57.43	15,426.18	0.00	1.00
zi	nsk.std	-55.68	15,691.06	0.00	1.00
random_effect	aclew_child_id	0.31	NA	NA	NA

Table 2

Model output of the zero-inflated negative binomial mixed-effects regression of TCDS min/hr for the random sample, with afternoon as the reference level for time of day.

component	term	estimate	std.error	statistic	p.value
cond	(Intercept)	1.40	0.22	6.47	0.00
cond	tchyr.std	-0.25	0.25	-1.02	0.31
cond	stthr.tri.amidday	-0.49	0.37	-1.31	0.19
cond	stthr.tri.amorning	0.34	0.27	1.26	0.21
cond	hsz.std	0.01	0.22	0.04	0.97
cond	nsk.std	-0.12	0.16	-0.75	0.45
cond	tchyr.std:stthr.tri.amidday	0.85	0.38	2.26	0.02
cond	tchyr.std:stthr.tri.amorning	0.57	0.30	1.90	0.06
cond	tchyr.std:nsk.std	0.57	0.19	2.95	0.00
zi	(Intercept)	-57.88	16,902.92	0.00	1.00
zi	nsk.std	-56.14	17,193.15	0.00	1.00
random_effect	aclew_child_id	0.31	NA	NA	NA

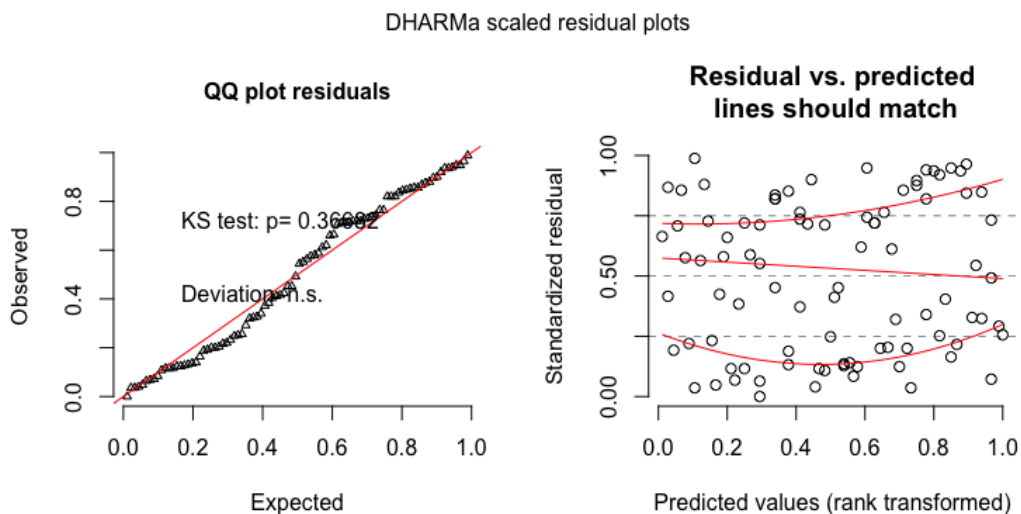


Figure 2. The model residuals from the zero-inflated negative binomial mixed-effects regression of TCDS min/hr for the random sample.

51 As an alternative analysis we generated parallel models of TCDS rate in the random
 52 clips using gaussian mixed-effects regression with logged values of TCDS: results for the two
 53 models demonstrating all pairwise effects of time of day are shown in Table 3 and Table 4.
 54 The residuals for the default gaussian model (Table 3) are shown in Figure 3.

Table 3

Full output of the gaussian mixed-effects regression of TCDS min/hr for the random sample, with midday as the reference level for time of day.

component	term	estimate	std.error	statistic	p.value
cond	(Intercept)	0.82	0.19	4.33	0.00
cond	tchiyр.std	0.54	0.22	2.42	0.02
cond	stthr.trimorning	0.50	0.25	2.02	0.04
cond	stthr.triafternoon	0.29	0.22	1.31	0.19
cond	hsz.std	-0.16	0.16	-0.99	0.32
cond	nsk.std	0.23	0.12	1.93	0.05
cond	tchiyр.std:stthr.trimorning	-0.17	0.27	-0.65	0.52
cond	tchiyр.std:stthr.triafternoon	-0.68	0.24	-2.85	0.00
cond	tchiyр.std:nsk.std	0.23	0.14	1.66	0.10
random_effect	aclew_child_id	0.21	NA	NA	NA
random_effect	Residual	0.84	NA	NA	NA

Table 4

Model output of the gaussian mixed-effects regression of TCDS min/hr for the random sample, with afternoon as the reference level for time of day.

component	term	estimate	std.error	statistic	p.value
cond	(Intercept)	1.11	0.15	7.55	0.00
cond	tchiyр.std	-0.14	0.18	-0.80	0.42
cond	stthr.tri.amidday	-0.29	0.22	-1.31	0.19
cond	stthr.tri.amorning	0.22	0.22	0.98	0.33
cond	hsz.std	-0.16	0.16	-0.99	0.32
cond	nsk.std	0.23	0.12	1.93	0.05
cond	tchiyр.std:stthr.tri.amidday	0.68	0.24	2.85	0.00
cond	tchiyр.std:stthr.tri.amorning	0.51	0.23	2.21	0.03
cond	tchiyр.std:nsk.std	0.23	0.14	1.66	0.10
random_effect	aclew_child_id	0.21	NA	NA	NA
random_effect	Residual	0.84	NA	NA	NA

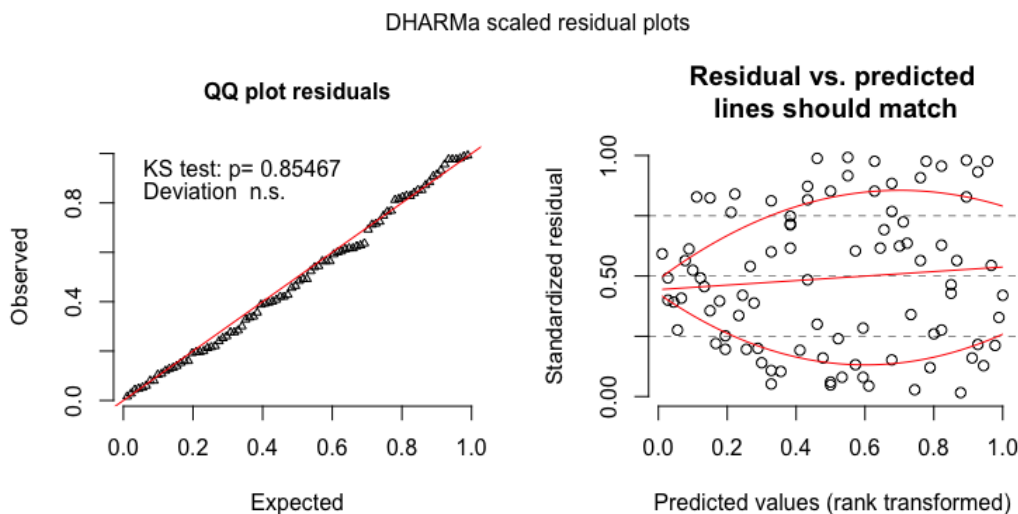


Figure 3. The model residuals from the gaussian mixed-effects regression of TCDS min/hr for the random sample.

55 **Turn-taking clips.** TCDS rate in the turn-taking clips demonstrated a slightly
 56 skewed, but unimodal distribution Figure 4. We therefore modeled it using a plain (i.e.,
 57 non-zero-inflated) negative binomial mixed-effects regression in the main text: results for the
 58 two models demonstrating all pairwise effects of time of day are shown in Table 5 and Table
 59 6. The residuals for the default model (Table 5) are shown in Figure 5.

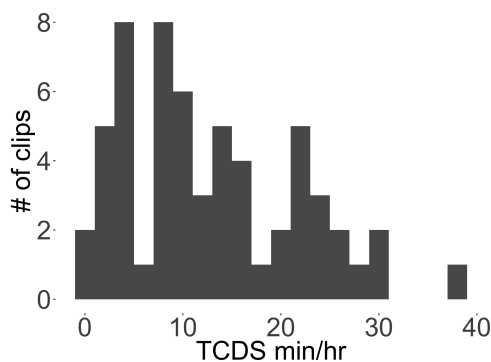


Figure 4. The distribution of TCDS rates found across the 59 turn-taking clips.

Table 5

Full output of the negative binomial mixed-effects regression of TCDS min/hr for the turn-taking sample, with midday as the reference level for time of day.

component	term	estimate	std.error	statistic	p.value
cond	(Intercept)	2.52	0.22	11.32	0.00
cond	tchiyр.std	0.08	0.21	0.38	0.70
cond	stthr.trimorning	0.14	0.29	0.48	0.63
cond	stthr.triafternoon	0.06	0.27	0.23	0.82
cond	hsz.std	0.12	0.14	0.86	0.39
cond	nsk.std	-0.13	0.10	-1.23	0.22
cond	tchiyр.std:stthr.trimorning	-0.13	0.29	-0.47	0.64
cond	tchiyр.std:stthr.triafternoon	0.00	0.24	0.01	1.00
cond	tchiyр.std:nsk.std	0.06	0.13	0.46	0.65
random_effect	aclew_child_id	0.19	NA	NA	NA

Table 6

Model output of the negative binomial mixed-effects regression of TCDS min/hr for the turn-taking sample, with afternoon as the reference level for time of day.

component	term	estimate	std.error	statistic	p.value
cond	(Intercept)	2.58	0.17	15.10	0.00
cond	tchiyр.std	0.08	0.19	0.44	0.66
cond	stthr.tri.amidday	-0.06	0.27	-0.23	0.82
cond	stthr.tri.amorning	0.08	0.22	0.34	0.74
cond	hsz.std	0.12	0.14	0.86	0.39
cond	nsk.std	-0.13	0.10	-1.23	0.22
cond	tchiyр.std:stthr.tri.amidday	0.00	0.24	-0.01	1.00
cond	tchiyр.std:stthr.tri.amorning	-0.14	0.26	-0.51	0.61
cond	tchiyр.std:nsk.std	0.06	0.13	0.46	0.65
random_effect	aclew_child_id	0.19	NA	NA	NA

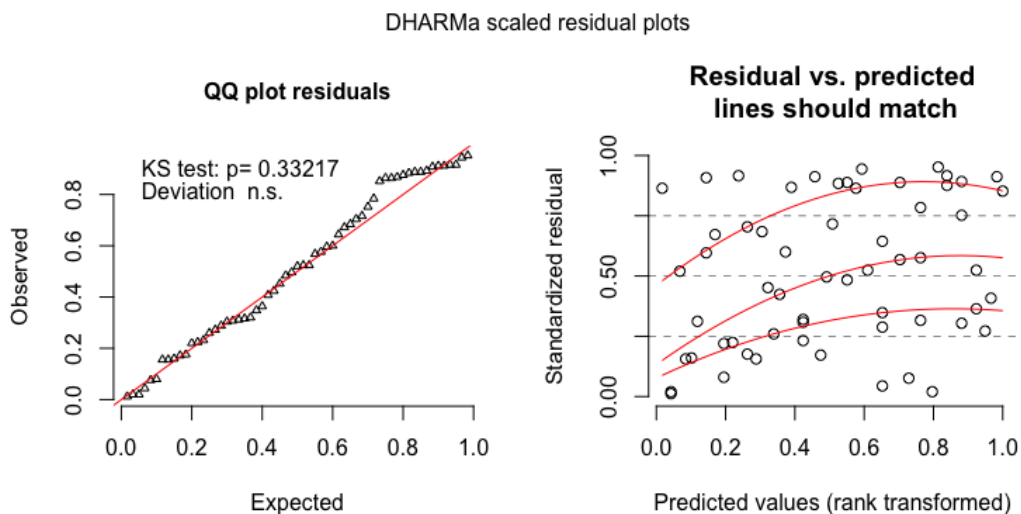


Figure 5. The model residuals from the negative binomial mixed-effects regression of TCDS min/hr for the turn-taking sample.

60 As an alternative analysis we generated parallel models of TCDS rate in the
 61 turn-taking clips using gaussian mixed-effects regression with logged values of TCDS: results
 62 for the two models demonstrating all pairwise effects of time of day are shown in Table 7 and
 63 Table 8. The residuals for the default gaussian model (Table 7) are shown in Figure 6.

Table 7

Full output of the gaussian mixed-effects regression of TCDS min/hr for the turn-taking sample, with midday as the reference level for time of day.

component	term	estimate	std.error	statistic	p.value
cond	(Intercept)	2.40	0.26	9.41	0.00
cond	tchiyр.std	0.09	0.23	0.37	0.71
cond	stthr.trimorning	0.13	0.34	0.38	0.70
cond	stthr.triafternoon	0.05	0.30	0.17	0.86
cond	hsz.std	0.13	0.15	0.89	0.37
cond	nsk.std	-0.14	0.12	-1.16	0.24
cond	tchiyр.std:stthr.trimorning	-0.17	0.32	-0.52	0.60
cond	tchiyр.std:stthr.triafternoon	0.04	0.27	0.15	0.88
cond	tchiyр.std:nsk.std	0.07	0.15	0.49	0.62
random_effect	aclew_child_id	0.22	NA	NA	NA
random_effect	Residual	0.71	NA	NA	NA

Table 8

Model output of the gaussian mixed-effects regression of TCDS min/hr for the turn-taking sample, with afternoon as the reference level for time of day.

component	term	estimate	std.error	statistic	p.value
cond	(Intercept)	2.46	0.18	13.76	0.00
cond	tchiyр.std	0.13	0.21	0.60	0.55
cond	stthr.tri.amidday	-0.05	0.30	-0.17	0.86
cond	stthr.tri.amorning	0.08	0.26	0.29	0.77
cond	hsz.std	0.13	0.15	0.89	0.37
cond	nsk.std	-0.14	0.12	-1.16	0.24
cond	tchiyр.std:stthr.tri.amidday	-0.04	0.27	-0.15	0.88
cond	tchiyр.std:stthr.tri.amorning	-0.21	0.29	-0.70	0.48
cond	tchiyр.std:nsk.std	0.07	0.15	0.49	0.62
random_effect	aclew_child_id	0.22	NA	NA	NA
random_effect	Residual	0.71	NA	NA	NA

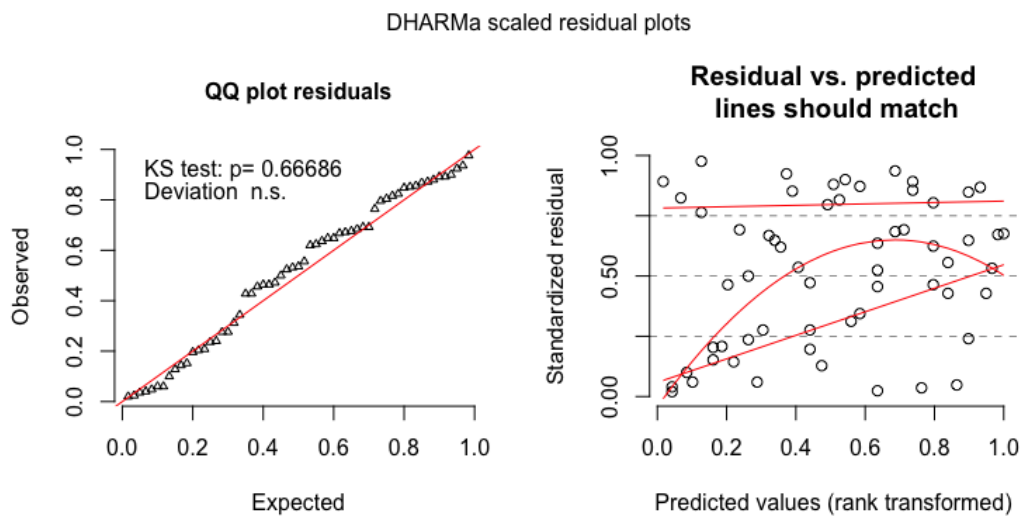


Figure 6. The model residuals from the gaussian mixed-effects regression of TCDS min/hr for the turn-taking sample.

64 Other-directed speech (ODS)

65 **Random clips.** ODS rate in the random clips demonstrated a skewed distribution
 66 with extra cases of zero Figure 7. We therefore modeled it using a zero-inflated negative
 67 binomial mixed-effects regression.in the main text: results for the two models demonstrating
 68 all pairwise effects of time of day are shown in Table 9 and Table 10. The residuals for the
 69 default model (Table 9) are shown in Figure 8.

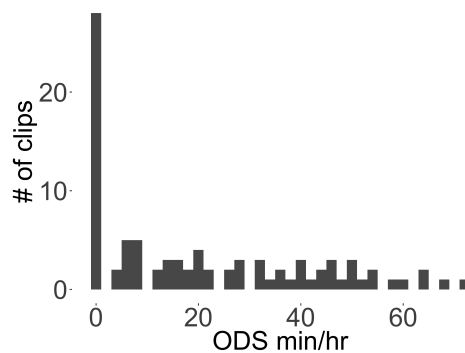


Figure 7. The distribution of ODS rates found across the 90 random clips.

Table 9

Full output of the zero-inflated negative binomial mixed-effects regression of ODS min/hr for the random sample, with midday as the reference level for time of day.

component	term	estimate	std.error	statistic	p.value
cond	(Intercept)	2.71	0.16	16.87	0.00
cond	tchiyр.std	-0.39	0.16	-2.43	0.02
cond	stthr.trimorning	0.45	0.18	2.49	0.01
cond	stthr.triafternoon	0.33	0.16	2.00	0.05
cond	hsz.std	-0.12	0.08	-1.52	0.13
cond	nsk.std	0.68	0.09	7.29	0.00
cond	tchiyр.std:stthr.trimorning	0.26	0.20	1.31	0.19
cond	tchiyр.std:stthr.triafternoon	0.42	0.17	2.42	0.02
cond	tchiyр.std:nsk.std	0.14	0.11	1.29	0.20
zi	(Intercept)	-51.51	13,502.22	0.00	1.00
zi	nsk.std	-55.02	13,734.07	0.00	1.00
random_effect	aclew_child_id	0.00	NA	NA	NA

Table 10

Model output of the zero-inflated negative binomial mixed-effects regression of ODS min/hr for the random sample, with afternoon as the reference level for time of day.

component	term	estimate	std.error	statistic	p.value
cond	(Intercept)	3.04	0.11	27.93	0.00
cond	tchiyр.std	0.03	0.10	0.32	0.75
cond	stthr.tri.amidday	-0.33	0.16	-2.00	0.05
cond	stthr.tri.amorning	0.12	0.15	0.83	0.41
cond	hsz.std	-0.12	0.08	-1.52	0.13
cond	nsk.std	0.68	0.09	7.29	0.00
cond	tchiyр.std:stthr.tri.amidday	-0.42	0.17	-2.42	0.02
cond	tchiyр.std:stthr.tri.amorning	-0.16	0.16	-0.98	0.33
cond	tchiyр.std:nsk.std	0.14	0.11	1.29	0.20
zi	(Intercept)	-50.05	10,018.85	0.00	1.00
zi	nsk.std	-53.54	10,190.89	0.00	1.00
random_effect	aclew_child_id	0.00	NA	NA	NA

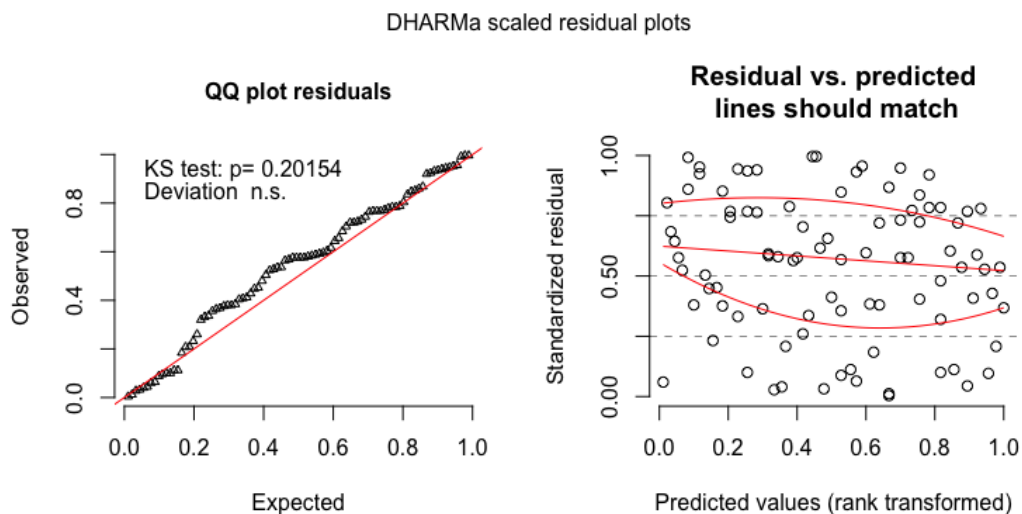


Figure 8. The model residuals from the zero-inflated negative binomial mixed-effects regression of ODS min/hr for the random sample.

70 As an alternative analysis we generated parallel models of ODS rate in the random
 71 clips using gaussian mixed-effects regression with logged values of ODS: results for the two
 72 models demonstrating all pairwise effects of time of day are shown in Table 11 and Table 12.
 73 The residuals for the default gaussian model (Table 11) are shown in Figure 9.

Table 11

Full output of the gaussian mixed-effects regression of ODS min/hr for the random sample, with midday as the reference level for time of day.

component	term	estimate	std.error	statistic	p.value
cond	(Intercept)	2.04	0.15	13.37	0.00
cond	tchiyр.std	-0.26	0.18	-1.49	0.14
cond	stthr.trimorning	0.23	0.21	1.09	0.28
cond	stthr.triafternoon	0.35	0.19	1.86	0.06
cond	hsz.std	-0.38	0.11	-3.37	0.00
cond	nsk.std	1.56	0.10	16.30	0.00
cond	tchiyр.std:stthr.trimorning	0.07	0.23	0.31	0.75
cond	tchiyр.std:stthr.triafternoon	0.43	0.20	2.08	0.04
cond	tchiyр.std:nsk.std	0.18	0.11	1.58	0.11
random_effect	aclew_child_id	0.00	NA	NA	NA
random_effect	Residual	0.73	NA	NA	NA

Table 12

Model output of the gaussian mixed-effects regression of ODS min/hr for the random sample, with afternoon as the reference level for time of day.

component	term	estimate	std.error	statistic	p.value
cond	(Intercept)	2.40	0.11	21.11	0.00
cond	tchiyр.std	0.16	0.13	1.22	0.22
cond	stthr.tri.amidday	-0.35	0.19	-1.86	0.06
cond	stthr.tri.amorning	-0.12	0.19	-0.64	0.52
cond	hsz.std	-0.38	0.11	-3.37	0.00
cond	nsk.std	1.56	0.10	16.30	0.00
cond	tchiyр.std:stthr.tri.amidday	-0.43	0.20	-2.08	0.04
cond	tchiyр.std:stthr.tri.amorning	-0.36	0.20	-1.82	0.07
cond	tchiyр.std:nsk.std	0.18	0.11	1.58	0.11
random_effect	aclew_child_id	0.00	NA	NA	NA
random_effect	Residual	0.73	NA	NA	NA

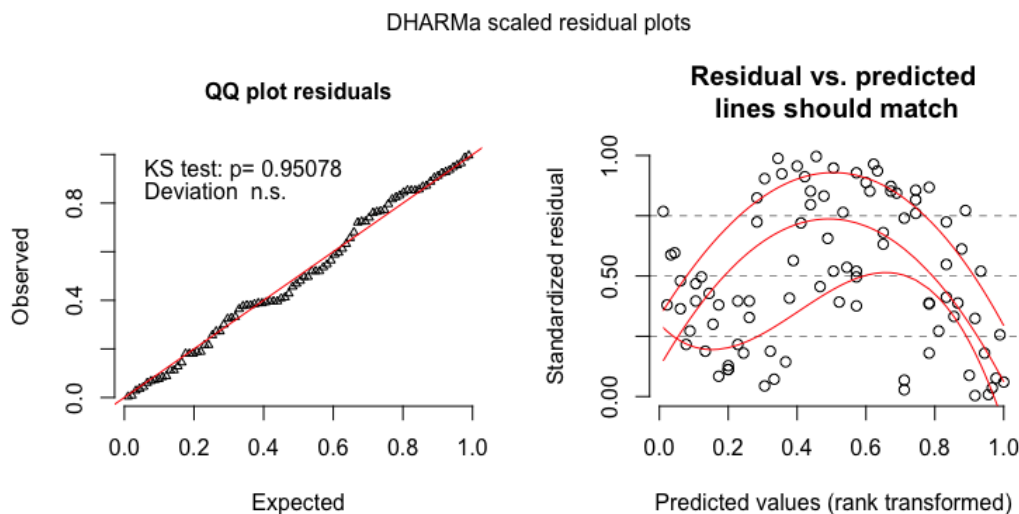


Figure 9. The model residuals from the gaussian mixed-effects regression of ODS min/hr for the random sample.

74 **Turn-taking clips.** ODS rate in the turn-taking clips demonstrated a skewed
 75 distribution with extra cases of zero Figure 10. We therefore modeled it using a zero-inflated
 76 negative binomial mixed-effects regression in the main text: results for the two models
 77 demonstrating all pairwise effects of time of day are shown in Table 13 and Table 14. The
 78 residuals for the default model (Table 13) are shown in Figure 11.

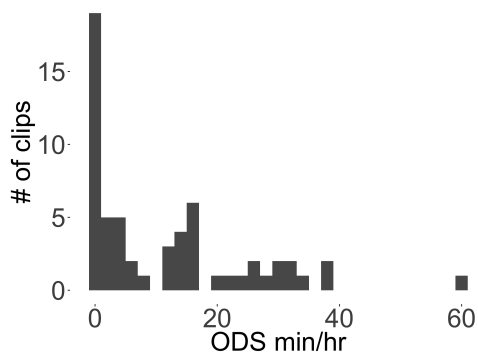


Figure 10. The distribution of ODS rates found across the 59 turn-taking clips.

Table 13

Full output of the negative binomial mixed-effects regression of ODS min/hr for the turn-taking sample, with morning as the reference level for time of day (note that most default models have midday as the reference level for time of day; the default model is changed here due to convergence issues).

component	term	estimate	std.error	statistic	p.value
cond	(Intercept)	2.64	0.16	16.02	0.00
cond	tchiyр.std	-0.80	0.23	-3.43	0.00
cond	stthr.tri.oafternoon	-0.61	0.25	-2.41	0.02
cond	stthr.tri.omidday	0.00	0.26	-0.01	0.99
cond	hsz.std	-0.18	0.09	-2.12	0.03
cond	nsk.std	0.63	0.10	6.44	0.00
cond	tchiyр.std:stthr.tri.oafternoon	0.48	0.29	1.62	0.11
cond	tchiyр.std:stthr.tri.omidday	0.54	0.30	1.77	0.08
cond	tchiyр.std:nsk.std	-0.01	0.14	-0.09	0.93
zi	(Intercept)	-31.97	11,304.01	0.00	1.00
zi	nsk.std	-31.33	11,122.86	0.00	1.00
random_effect	aclew_child_id	0.00	NA	NA	NA

Table 14

Model output of the negative binomial mixed-effects regression of ODS min/hr for the turn-taking sample, with afternoon as the reference level for time of day.

component	term	estimate	std.error	statistic	p.value
cond	(Intercept)	2.03	0.22	9.11	0.00
cond	tchyr.std	-0.33	0.25	-1.33	0.18
cond	stthr.tri.amidday	0.61	0.29	2.07	0.04
cond	stthr.tri.amorning	0.61	0.25	2.41	0.02
cond	hsz.std	-0.18	0.09	-2.12	0.03
cond	nsk.std	0.63	0.10	6.44	0.00
cond	tchyr.std:stthr.tri.amidday	0.06	0.31	0.20	0.84
cond	tchyr.std:stthr.tri.amorning	-0.48	0.29	-1.62	0.11
cond	tchyr.std:nsk.std	-0.01	0.14	-0.09	0.93
zi	(Intercept)	-32.22	12,257.76	0.00	1.00
zi	nsk.std	-31.58	12,061.33	0.00	1.00
random_effect	aclew_child_id	0.00	NA	NA	NA

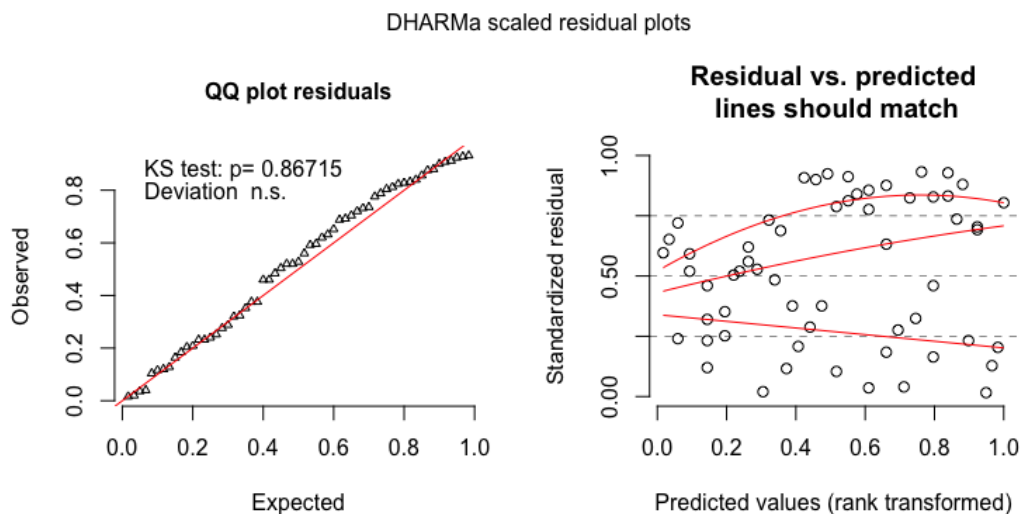


Figure 11. The model residuals from the zero-inflated negative binomial mixed-effects regression of ODS min/hr for the turn-taking sample.

79 As an alternative analysis we generated parallel models of ODS rate in the turn-taking
 80 clips using gaussian mixed-effects regression with logged values of ODS: results for the two
 81 models demonstrating all pairwise effects of time of day are shown in Table 15 and Table 16.
 82 The residuals for the default gaussian model (Table 15) are shown in Figure 12.

Table 15

Full output of the gaussian mixed-effects regression of ODS min/hr for the turn-taking sample, with midday as the reference level for time of day.

component	term	estimate	std.error	statistic	p.value
cond	(Intercept)	2.34	0.21	10.93	0.00
cond	tchyr.std	-0.15	0.20	-0.74	0.46
cond	stthr.trimorning	-0.25	0.28	-0.86	0.39
cond	stthr.triafternoon	-0.72	0.26	-2.77	0.01
cond	hsz.std	-0.27	0.12	-2.28	0.02
cond	nsk.std	1.09	0.11	10.02	0.00
cond	tchyr.std:stthr.trimorning	-0.25	0.29	-0.86	0.39
cond	tchyr.std:stthr.triafternoon	0.06	0.26	0.23	0.82
cond	tchyr.std:nsk.std	-0.08	0.14	-0.60	0.55
random_effect	aclew_child_id	0.00	NA	NA	NA
random_effect	Residual	0.69	NA	NA	NA

Table 16

Model output of the gaussian mixed-effects regression of ODS min/hr for the turn-taking sample, with afternoon as the reference level for time of day.

component	term	estimate	std.error	statistic	p.value
cond	(Intercept)	1.62	0.16	10.44	0.00
cond	tchiyр.std	-0.09	0.19	-0.49	0.62
cond	stthr.tri.amidday	0.72	0.26	2.77	0.01
cond	stthr.tri.amorning	0.47	0.24	1.94	0.05
cond	hsz.std	-0.27	0.12	-2.28	0.02
cond	nsk.std	1.09	0.11	10.02	0.00
cond	tchiyр.std:stthr.tri.amidday	-0.06	0.26	-0.23	0.82
cond	tchiyр.std:stthr.tri.amorning	-0.30	0.27	-1.13	0.26
cond	tchiyр.std:nsk.std	-0.08	0.14	-0.60	0.55
random_effect	aclew_child_id	0.00	NA	NA	NA
random_effect	Residual	0.69	NA	NA	NA

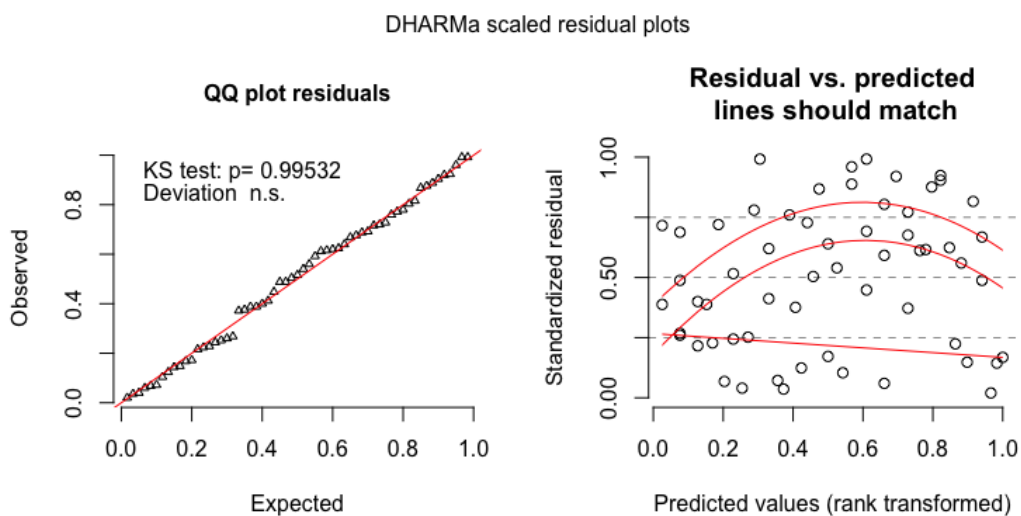


Figure 12. The model residuals from the gaussian mixed-effects regression of ODS min/hr for the turn-taking sample.

References

83

84 Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A.,

85 . . . Bolker, B. M. (2017a). glmmTMB balances speed and flexibility among packages

86 for zero-inflated generalized linear mixed modeling. *The R Journal*, 9, 378–400.

87 Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A.,

88 . . . Bolker, B. M. (2017b). Modeling zero-inflated count data with glmmTMB.

89 *bioRxiv*. <https://doi.org/10.1101/132753>