Supplementary materials

1. Additional instructions for coders

The coders were given several practical and more specific instructions to help them decide on the position of the answer point. These are given below.

Practical instructions

- Use the auditory information as well as the transcription (listen to one question at a time and immediately code that) and take into account intonation of the question (also in combination with context, see below).
- Take into account the participant's reaction; if they answer for example already during the question, it is clear that the answer point should be placed before that position.
- Take into account the context when determining the answer point. Listen to the interviews in the right order because many questions are asked based on the participant's previous reaction.
- When in doubt, listen to only part of the question and try to put yourself in the participant's position; would this information be enough to answer the question?

Specific instructions

- When in doubt about two consecutive words, take the latter as the answer word. However, when two words are clearly part of a compound (e.g., "Radboud University'), take the first of these as the answer word.
- In some cases, the interviewer first gives some information before actually asking the question. The answer word should always be in the question part.
- 'Or'-questions often allow the participant to start planning in the part before the 'or' because the latter part is often predictable from the first part, or the context.
- In the case of disfluencies (e.g., partial words), take the disfluency itself as the answer word if the word that was meant can be derived from the disfluency, otherwise, take the full word.
- Question words (like 'where', 'how long') often (but not always) indicate the expected answer immediately (in relation to the context).

2. ROIs in LMER

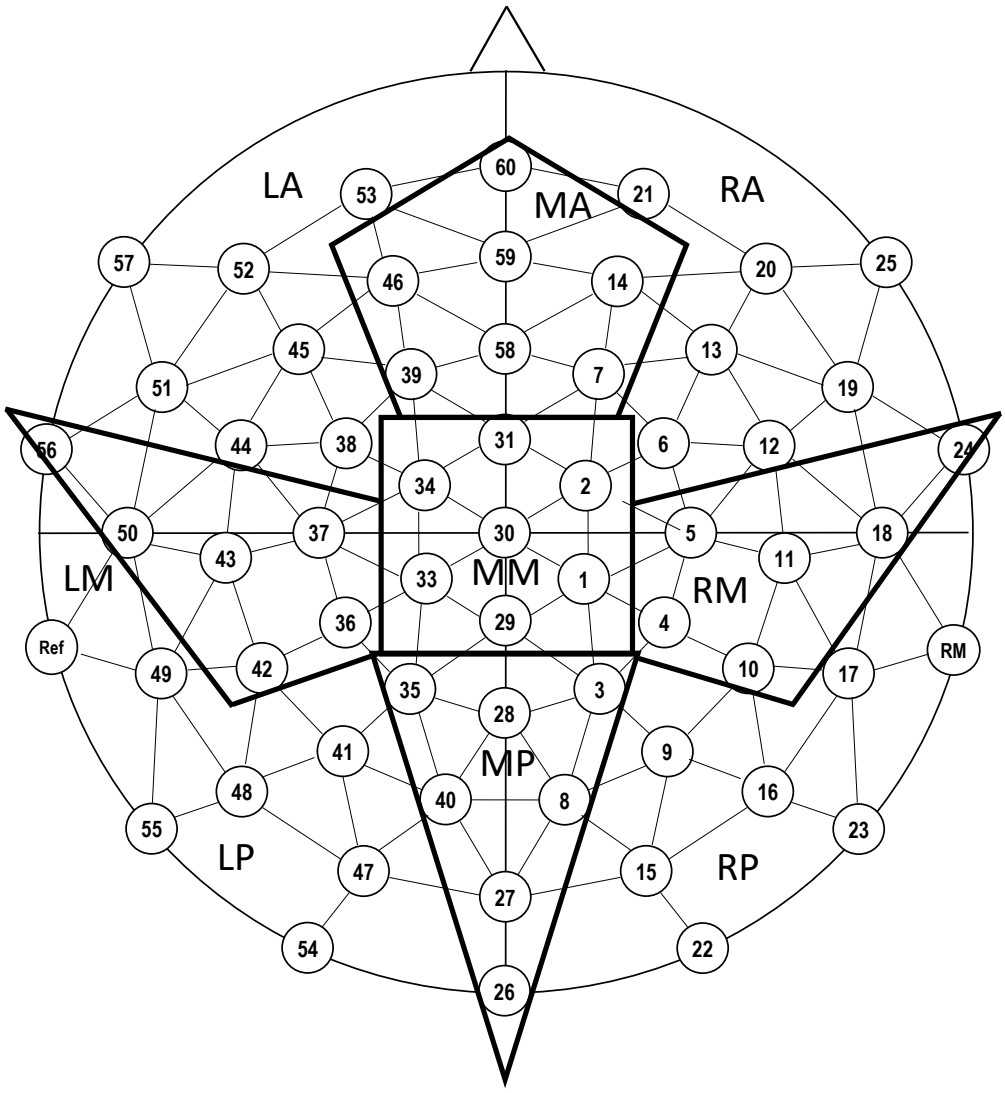

Figure S1. Division of electrodes over the 9 ROIs in the LMER EEG analyses. Each ROI is made up of 6 or 7 electrodes. The abbreviations indicate the names of the ROIs. The first letter indicates the level on the Anterior-Posterior factor (A: Anterior, M: Mid, P: Posterior) and the second letter indicates the level on the Left-Right factor (L: Left, M: Mid, R: Right).

3. Individual variability in response time

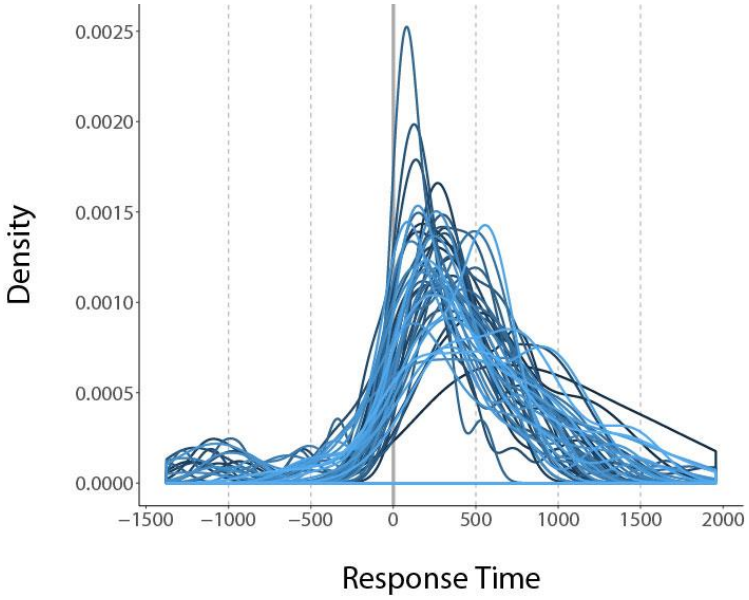

Figure S2. Density plots for response time (in milliseconds) for all participants separately. Every line represents a participant.

4. Distribution of question length

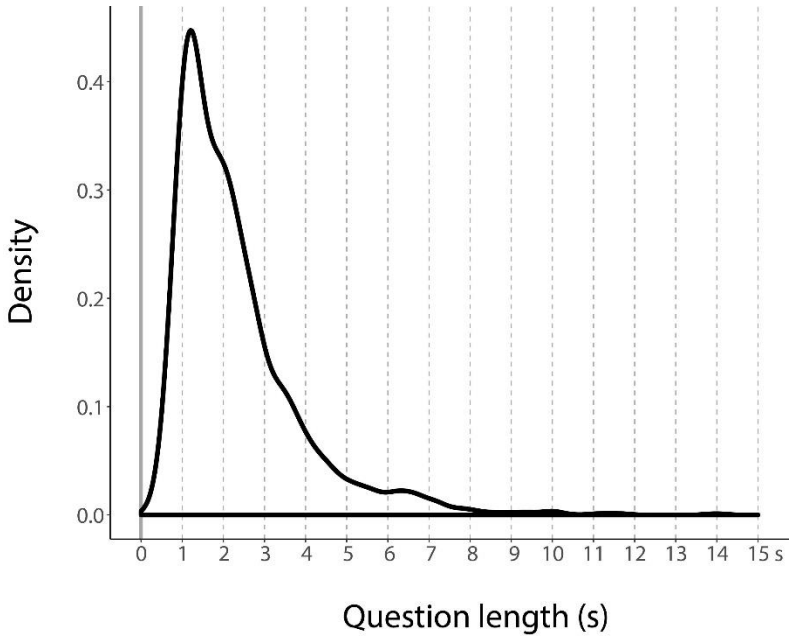

Figure S3. Density plot for question length (in seconds). The distribution is very right-skewed, with a mode just over one second, indicating relatively many short questions are present in the corpus.

5. Descriptive statistics agreed and non-scripted selections

The same descriptive statistics that were calculated on the full selection (see Figure 1 and section *3.1: Behavioral results*) were also calculated excluding questions on which the coders disagreed about the answer point ('agreed selection', 1091 questions) and excluding questions that were scripted ('non-scripted selection', 1052 questions). Table S1 shows mean, median, and estimated mode of Response Time, Planning Time, and normalized position within the question (relative to question onset) for both selections plus the full selection for comparison. The values are generally quite similar, except for the Response Times in the non-scripted selection, which are slightly later. This is to be expected because the scripted questions were all polar questions which required a 'yes' response, which is likely to be given very quickly.

Table S1. Descriptive statistics of general tendency for the full, agreed, and non-scripted selections on Response Time, Planning Time, and normalized position in the question.

		Full	**Agreed**	**Non-scripted**
Response time	*Mean*	380 ms	366 ms	408 ms
	Median	336 ms	320 ms	373 ms
	Mode	~ 200 ms	~ 200 ms	~ 250 ms
Planning time	*Mean*	1374 ms	1318 ms	1376 ms
	Median	1083 ms	1027 ms	1074 ms
	Mode	~ 700 ms	~ 700 ms	~ 750 ms
Normalized position in question	*Mean*	37.8%	37.7%	37.2%
	Median	36.2%	36.1%	34.8%
	Mode	~ 33%	~ 33%	~ 30%

6. Behavioral models

Table S2. Model for behavioral data.

Output behavioral model (Linear mixed model fit by maximum likelihood)

Formula

ResponseTime ~ PlanningTime_z + AnswerLength_z + QuestionType + QuestionOnset_to_AnswerPoint_z + FrequencyLog_z + WordType + Agreed + Scripted + PlanningTime_z * AnswerLength_z + PlanningTime_z * WordType + (PlanningTime_z + AnswerLength_z | Participant)

Control (lmerControl(optimizer = "bobyqa"))

AIC	BIC	logLik	deviance	df.resid
1299.8	1394.3	-631.9	1263.8	1390

Scaled Residuals

Min	1Q	Median	3Q	Max
-4.32	-0.50	-0.03	0.53	4.20

Random Effects

Groups	Name	Variance	Std.Dev	Corr	
Participant	Intercept	0.028	0.169		
	PlanningTime_z	0.007	0.082	0.29	
	AnswerLength_z	0.012	0.111	0.47	-0.29
Residual		0.127	0.357		

Fixed Effects

	Estimate	Std. Error	t-value
Intercept	0.396	0.030	13.126
PlanningTime_z	-0.138	0.019	-7.465
AnswerLength_z	0.158	0.022	7.170
QuestionType[polar]	-0.081	0.013	-6.307
QuestionOnset_to_AnswerPoint_z	0.013	0.011	1.166
FrequencyLog_z	0.008	0.012	0.635
WordType[content]	0.053	0.014	3.853
Agreed[yes]	-0.001	0.013	-0.080
Scripted[yes]	0.013	0.012	1.070
PlanningTime_z:AnswerLength_z	0.014	0.010	1.401
PlanningTime_z:QuestionType[polar]	0.006	0.013	0.508

Note. Number of observations = 1408; groups: Participant, 46

Table S3. Model for behavioral data after removal of planning times shorter than 400 ms

Output behavioral model; PlanningTime> 400 ms (Linear mixed model fit by maximum likelihood)

Formula

ResponseTime ~ PlanningTime_z + AnswerLength_z + QuestionType + QuestionOnset_to_AnswerPoint_z + FrequencyLog_z + WordType + Agreed + Scripted + PlanningTime_z * AnswerLength_z + PlanningTime_z * WordType + (PlanningTime_z + AnswerLength_z | Participant)

Control (lmerControl(optimizer = "bobyqa"))

AIC	BIC	logLik	deviance	df.resid
1269.8	1363.7	-616.9	1233.8	1346

Scaled Residuals

Min	1Q	Median	3Q	Max
-4.32	-0.50	-0.040	0.51	4.19

Random Effects

Groups	Name	Variance	Std.Dev	Corr	
Participant	Intercept	0.027	0.163		
	PlanningTime_z	0.007	0.082	0.36	
	AnswerLength_z	0.012	0.108	0.45	-.19
Residual		0.128	0.358		

Fixed Effects

	Estimate	Std. Error	t-value
Intercept	0.393	0.030	13.245
PlanningTime_z	-0.133	0.019	-7.089
AnswerLength_z	0.153	0.022	7.035
QuestionType[polar]	-0.087	0.013	-6.641
QuestionOnset_to_AnswerPoint_z	0.014	0.011	1.297
FrequencyLog_z	0.006	0.013	0.508
WordType[content]	0.049	0.014	3.421
Agreed[yes]	0.005	0.013	0.388
Scripted[yes]	0.014	0.013	1.154
PlanningTime_z:AnswerLength_z	0.015	0.010	1.442
PlanningTime_z:QuestionType[polar]	0.010	0.013	0.801

Note. Number of observations = 1364; groups: Participant, 46

7. Variability in ERPs

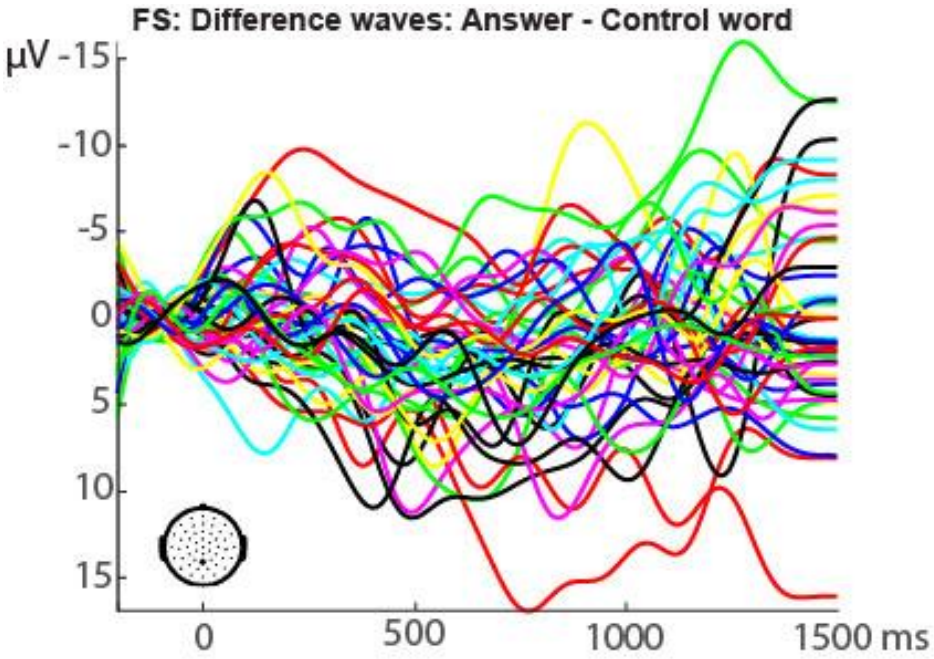

Figure S4. ERP difference waves for answer - control word per participant. Average ERP difference waves are displayed subtracting the ERPs for the control word from those for the answer word, for a representative electrode (see small head in bottom left). Every line displays averages for one participant.

8. ERP Models

Table S4. Model for ERP data for all 9 ROIs. Independent variable is the average voltage in a 400-800 ms window after critical (answer/control) word onset over all electrodes per ROI.

Output TFR model (Linear mixed model fit by maximum likelihood)

Formula

PowerLog ~ Condition*Ant-Post*Left-Right + FrequencyLog_z + WordType + QuestionOnset_to_Point_z + Agreed + Scripted + (Condition | Participant)

Control (lmerControl(optimizer = "bobyqa"))

AIC	BIC	logLik	deviance	df.resid
174691.2	174908.8	-87318.6	174637.2	23319

Scaled Residuals

Min	1Q	Median	3Q	Max
-9.06	-0.57	0.019	0.59	3.80

Random Effects

Groups	Name	Variance	Std.Dev	Corr
Participant	(Intercept)	2.567	1.602	
	Condition	0.991	0.996	0.07
Residual		102.921	10.145	

Fixed Effects

	Estimate	Std. Error	t-value
Intercept	-0.0904	0.2577	-0.3507
Condition[answer]	0.4042	0.1635	2.4714
Ant-Post[ant]	-0.3175	0.0940	-3.378
Ant-Post[mid]	0.1653	0.0939	1.7602
Left-Right[left]	0.0799	0.0939	0.8511
Left-Right[mid]	-0.1419	0.0939	-1.5106
FrequencyLog_z	-0.2693	0.0911	-2.9571
WordType[content]	0.2045	0.0940	2.1757
QuestionOnset_to_Point_z	0.2965	0.0716	4.1415
Agreed[yes]	0.4458	0.0849	5.2523
Scripted[yes]	-0.3161	0.0795	-3.9759
Condition[answer]:Ant-Post[ant]	-0.3151	0.0940	-3.3537
Condition[answer]:Ant-Post[mid]	0.0533	0.0939	0.5674
Condition[answer]:Left-Right[left]	-0.1170	0.0939	-1.2460
Condition[answer]:Left-Right[mid]	0.1319	0.0939	1.4044
Ant-Post[ant]:Left-Right[left]	-0.3286	0.1329	-2.4729
Ant-Post[mid]:Left-Right[left]	0.2156	0.1328	1.6231
Ant-Post[ant]:Left-Right[mid]	0.1105	0.1329	0.8316
Ant-Post[mid]:Left-Right[mid]	-0.4207	0.1328	-3.1672
Condition[answer]:Ant-Post[ant]:Left-Right[left]	-0.0625	0.1329	-0.4700
Condition[answer]:Ant-Post[mid]:Left-Right[left]	0.0128	0.1328	0.0965
Condition[answer]:Ant-Post[ant]:Left-Right[mid]	-0.0522	0.1329	-0.3929
Condition[answer]:Ant-Post[mid]:Left-Right[mid]	-0.0203	0.1328	-0.1530

Note. Number of observations = 23346*; groups: Participant, 46

Table S5. Model for ERP data for the Anterior ROIs. Independent variable is the average voltage in a 400-800 ms window after critical (answer/control) word onset in the electrodes over all electrodes per ROI (only including the 3 anterior ROIs).

Output ERP model (Linear mixed model fit by maximum likelihood)

Formula

Voltage ~ Condition + FrequencyLog_z + WordType + QuestionOnset_to_Point_z + Agreed + Scripted + Left-Right + (Condition | Participant)

Control (lmerControl(optimizer = "bobyqa"))

AIC	BIC	logLik	deviance	df.resid
59179.5	59269.9	-29576.8	59153.5	7744

Scaled Residuals

Min	1Q	Median	3Q	Max
-4.90	-0.57	0.010	0.59	3.23

Random Effects

Groups	Name	Variance	Std.Dev	Corr
Participant	(Intercept)	3.021	1.738	
	Condition	0.915	0.957	0.29
Residual		118.335	10.878	

Fixed Effects

	Estimate	Std. Error	t-value
Intercept	-0.3779	0.3196	-1.1830
Condition[answer]	0.0380	0.1939	0.1960
FrequencyLog_z	-0.3733	0.1691	2.2070
WordType[content]	0.1167	0.1745	0.6690
QuestionOnset_to_Point_z	0.1630	0.1327	1.2290
Agreed[yes]	0.4810	0.1574	3.0550
Scripted[yes]	-0.3379	0.1479	-2.2850
Left-Right[left]	-0.2510	0.1747	-1.4370
Left-Right[mid]	-0.0314	0.1746	-0.1800

Note. Number of observations = 7757; groups: Participant, 46

Table S6. Model for ERP data for the Mid ROIs. Independent variable is the average voltage in a 400-800 ms window after critical (answer/control) word onset in the electrodes over all electrodes per ROI (only including the 3 middle ROIs: LM, MM, RM, see Figure S1).

Output ERP model (Linear mixed model fit by maximum likelihood)

Formula

Voltage ~ Condition + FrequencyLog_z + WordType + QuestionOnset_to_Point_z + Agreed + Scripted + Left-Right + (Condition | Participant)

Control (lmerControl(optimizer = "bobyqa"))

AIC	BIC	logLik	deviance	df.resid
58388.3	58478.8	-29181.2	58362.3	7764

Scaled Residuals

Min	1Q	Median	3Q	Max
-5.90	-0.58	0.014	0.57	3.71

Random Effects

Groups	Name	Variance	Std.Dev	Corr
Participant	(Intercept)	2.107	1.452	
	Condition	0.913	0.956	0.11
Residual		104.867	10.240	

Fixed Effects

	Estimate	Std. Error	t-value
Intercept	0.0774	0.2793	0.2770
Condition[answer]	0.4469	0.1885	2.3710
FrequencyLog_z	-0.3240	0.1587	-2.0410
WordType[content]	0.1872	0.1639	1.1420
QuestionOnset_to_Point_z	0.3493	0.1247	2.8020
Agreed[yes]	0.4674	0.1481	3.1560
Scripted[yes]	-0.3057	0.1387	-2.2030
Left-Right[left]	0.2953	0.1642	1.7980
Left-Right[mid]	-0.5628	0.1642	-3.4270

Note. Number of observations = 7777; groups: Participant, 46

Table S7. Model for ERP data for the Posterior ROIs. Independent variable is the average voltage in a 400-800 ms window after critical (answer/control) word onset in the electrodes over all electrodes per ROI (only including the 3 posterior ROIs).

Output ERP model (Linear mixed model fit by maximum likelihood)

Formula

Voltage ~ Condition + FrequencyLog_z + WordType + QuestionOnset_to_Point_z + Agreed + Scripted + Left-Right + (Condition | Participant)

Control (lmerControl(optimizer = "bobyqa"))

AIC	BIC	logLik	deviance	df.resid
57075	57165.5	-28524.5	57049	7799

Scaled Residuals

Min	1Q	Median	3Q	Max
-9.70	-0.56	0.023	0.59	3.64

Random Effects

Groups	Name	Variance	Std.Dev	Corr
Participant	(Intercept)	3.153	1.776	
	Condition	0.733	0.856	-0.15
Residual		85.468	9.245	

Fixed Effects

	Estimate	Std. Error	t-value
Intercept	0.0200	0.3078	0.0650
Condition[answer]	0.7066	0.1693	4.1750
FrequencyLog_z	-0.1099	0.1429	-0.7690
WordType[content]	0.2893	0.1475	1.9620
QuestionOnset_to_Point_z	0.3622	0.1127	3.2160
Agreed[yes]	0.4119	0.1336	3.0830
Scripted[yes]	-0.2952	0.1251	-2.3610
Left-Right[left]	0.1917	0.1480	1.2950
Left-Right[mid]	0.1691	0.1479	1.1430

Note. Number of observations = 7812; groups: Participant, 46

9. TFR model

Table S8. Model for TFR data. Independent variable is the log of the average power in an 800-1000 ms window after critical (answer/control) word onset within a frequency range of 9-13 Hz over all electrodes per ROI.

Output TFR model (Linear mixed model fit by maximum likelihood)

Formula

PowerLog ~ Condition*Ant-Post*Left-Right + FrequencyLog_z + WordType + QuestionOnset_to_Point_z + Agreed + Scripted + (Condition | Participant)

Control (lmerControl(optimizer = "bobyqa"))

AIC	BIC	logLik	deviance	df.resid
39512.6	39723.2	-19729.3	39458.6	18018

Scaled Residuals

Min	1Q	Median	3Q	Max
-3.857	-0.682	-0.029	0.674	3.835

Random Effects

Groups	Name	Variance	Std.Dev	Corr
Participant	(Intercept)	0.305	0.553	
	Condition	0.006	0.076	-0.09
Residual		0.512	0.716	

Fixed Effects

	Estimate	Std. Error	t-value
Intercept	1.2030	0.0843	14.2760
Condition[answer]	-0.0082	0.0126	-0.6500
Ant-Post[ant]	-0.2306	0.0075	-30.6620
Ant-Post[mid]	-0.0264	0.0075	-3.5000
Left-Right[left]	-0.0407	0.0075	-5.4000
Left-Right[mid]	0.0357	0.0076	4.7220
FrequencyLog_z	0.0266	0.0074	3.5800
WordType[content]	0.0093	0.0160	0.5810
QuestionOnset_to_Point_z	-0.0074	0.0058	-1.2890
Agreed[yes]	0.0098	0.0068	1.4370
Scripted[yes]	-0.0360	0.0063	-5.7040
Condition[answer]:Ant-Post[ant]	0.0031	0.0075	0.4150
Condition[answer]:Ant-Post[mid]	-0.0060	0.0075	-0.7990
Condition[answer]:Left-Right[left]	-0.0027	0.0075	-0.3520
Condition[answer]:Left-Right[mid]	0.0000	0.0076	0.0030
Ant-Post[ant]:Left-Right[left]	0.0113	0.0106	1.0670
Ant-Post[mid]:Left-Right[left]	0.0025	0.0106	0.2370
Ant-Post[ant]:Left-Right[mid]	0.0059	0.0107	0.5560
Ant-Post[mid]:Left-Right[mid]	-0.0097	0.0107	-0.9110
Condition[answer]:Ant-Post[ant]:Left-Right[left]	-0.0020	0.0106	-0.1870
Condition[answer]:Ant-Post[mid]:Left-Right[left]	0.0022	0.0106	0.2100
Condition[answer]:Ant-Post[ant]:Left-Right[mid]	0.0082	0.0107	0.7690
Condition[answer]:Ant-Post[mid]:Left-Right[mid]	-0.0013	0.0107	-0.1250

Note. Number of observations = 18045*; groups: Participant, 46

*The number of observations differs for the ERP and TFR models because not all trials contain data in the given time window (since trials were cut off before speech onset). Given that the time window for TFRs occurs later than that for ERPs, the former model contains fewer observations.