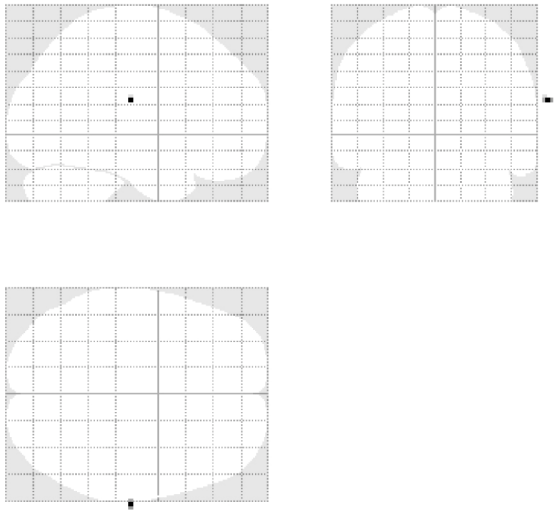
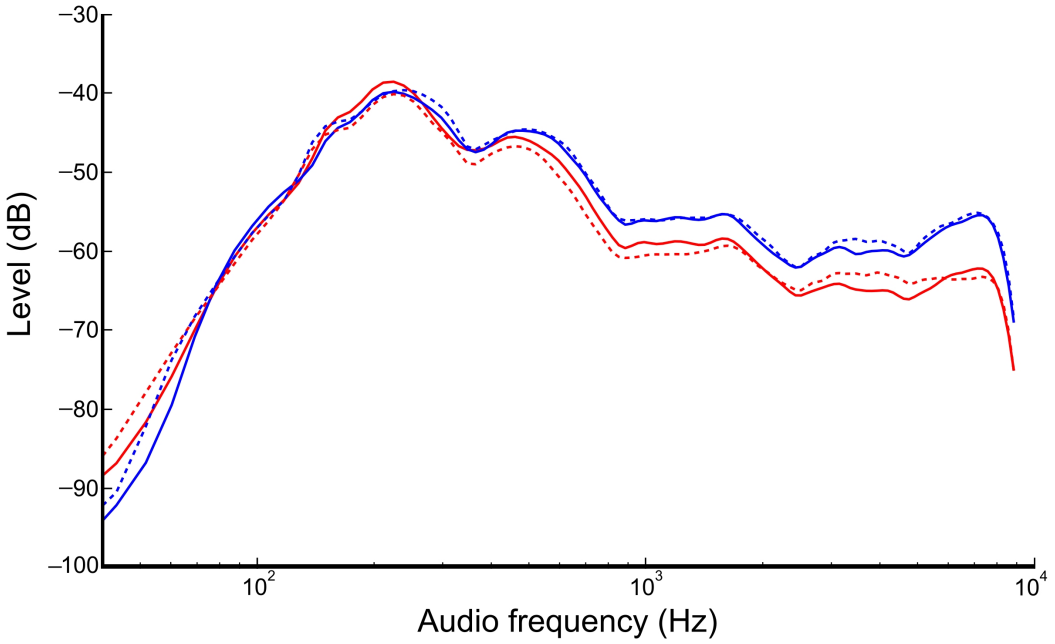


Supplementary Figure 1

Glassbrain projection of the group L30 > L960 functional localizer contrast.

The results are shown at a statistical significance threshold of $p < 0.001$ (uncorrected for multiple comparisons).

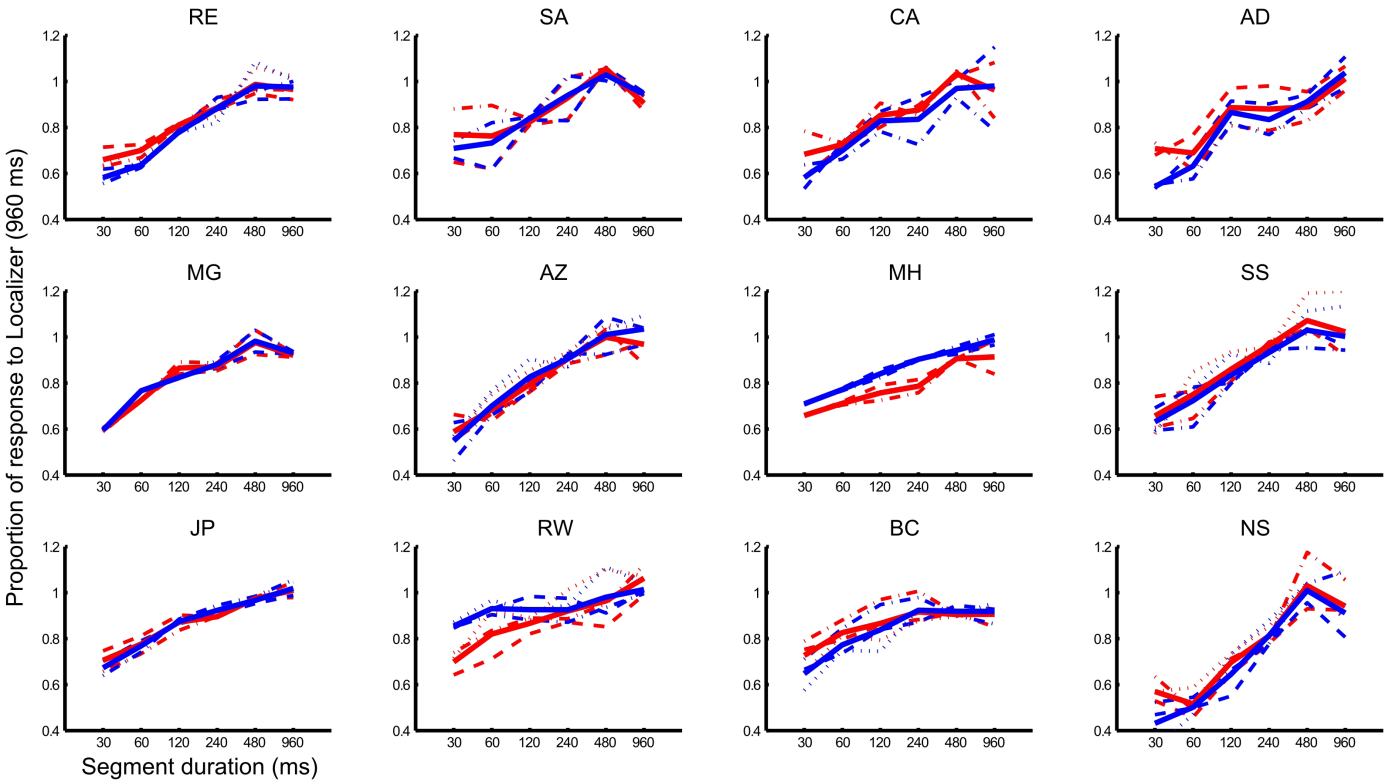


Supplementary Figure 2

Frequency power spectra for speech and modulation control sounds.

Frequency power spectra for speech (solid) and modulation control sounds (dashed) quilted with 30 ms (red) and 960 ms (blue) segment durations.

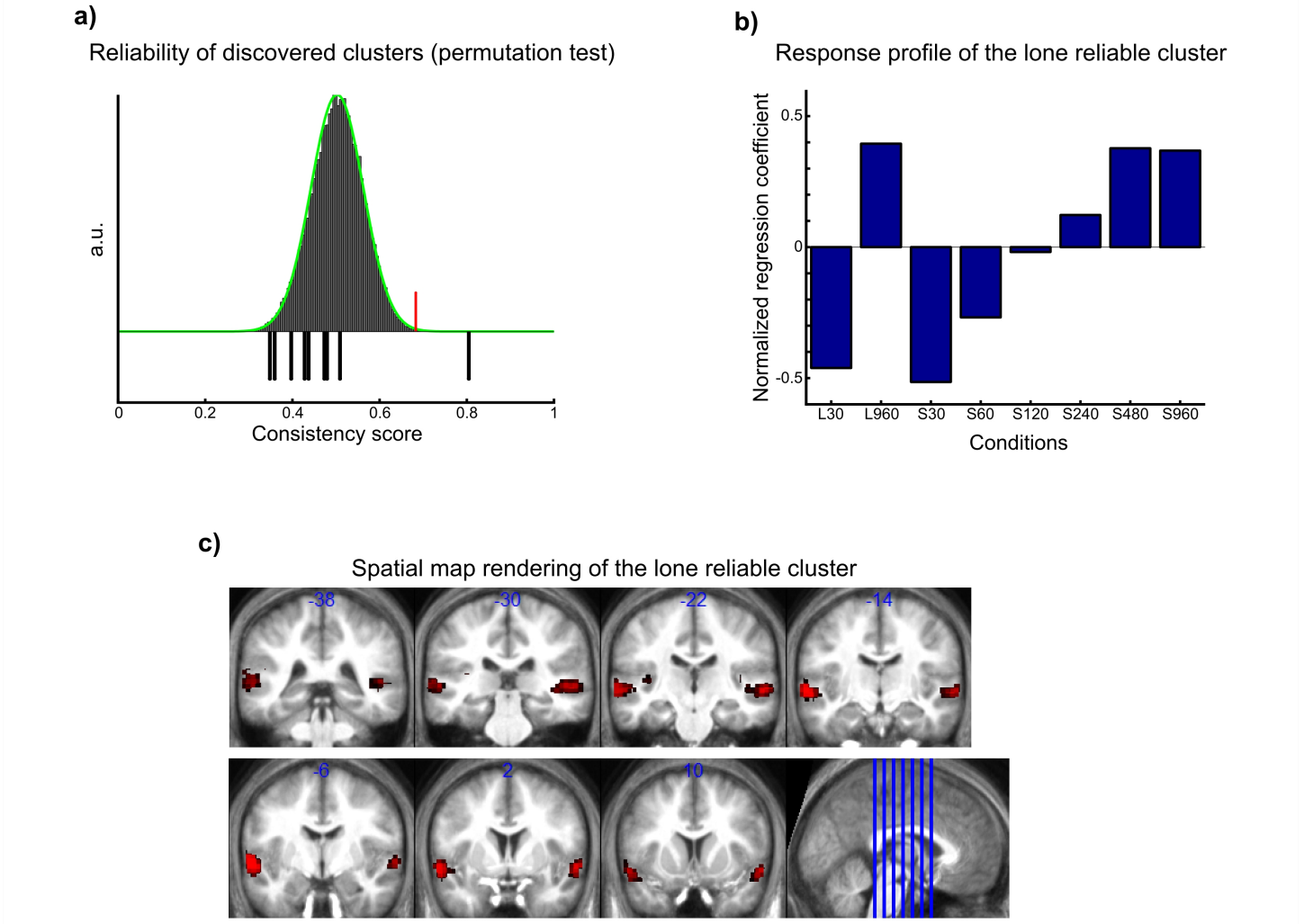
Nature Neuroscience: doi:10.1038/nn.4021



Supplementary Figure 3

Replicability across scanning sessions.

Replicability across scanning sessions for 12 participants who were scanned between two and four times. The graphs plot the BOLD response normalized by the response to the 960 ms functional localizer condition in each participant's individual fROI for the right and left hemispheric individual fROIs (red and blue, respectively) in the individual scanning sessions (dashed, dashed-dotted, dotted) that included the original speech quilts (i.e. not compressed speech quilts). The solid line plots the average across scanning sessions. Note that the majority of participants exhibits a plateau at around 480 ms segment duration.



Supplementary Figure 4

Clustering analysis of voxel response profiles[30].

a) A permutation test revealed that only the top-ranked cluster (out of nine) was statistically significant (red line segment). **b)** Profile of the top-ranked discovered cluster for the eight experimental conditions (L30, L960, S30 to S960). Note that the data were mean-centered, which is why the response profile is negative for the conditions yielding a low response. **c)** Rendering of this cluster on coronal cross-sections of our participants' average structural images (y = -38, -30, -22, -14, -6, 2, 10), thresholded at a voxel by functional system assignment probability of $r \geq 0.7$.

Supplementary Table S1

Information about participants. a) Participant IDs and respective scanning numbers. **b)** Participant IDs with respect to the various control experiments. Note that the missing participant numbers that do not appear in b) were scanned in additional control experiments that we do not report here; however, all scans of all participants contributed to the functional group ROI.

a)

ID	No.
JK	1
RE	2,16,32,40
SA	3,10,52
CA	4,15,20,45
AD	5,12,24,39
VN	6
DW	7,43
MG	8,17,26,49
AZ	9,19,30,44
MH	11,18,29
SS	13,21,(26)*,42
JP	14,28,31
RW	22,33,37
BC	23,35,38,41
NS	25,34,36,47
KD	(46,48)**
DC	50,(51)***

*: scanner malfunction
**: no reliable bilateral activation to Localizer contrast
***: trigger code malfunction

b)

Experiment	No.
ModCon	1,2,3,4,5,6,7,13,14
CoMod	21,22,25,28,29
NoiseVoc10	18,23,32,33,34
EnvSound	30,31,35,36,37
OrigMinChange	15,16,17,19,38,(48),(51)
CompDil	39,40,41,42,43,44,45,(46),47,49,50,52

The missing participant numbers were scanned in additional control experiments that we do not report here.

Supplementary Table S2

Speech rates of source materials for the quilts. Quilts were generated from 20 s excerpts of speakers reading a book, after breaths and pauses had been excised. Two such excerpts for each of the 4 female and 4 male speakers were analyzed. 2 s samples from the beginning (randomly selected to fall in the range of 0.5-6 seconds from the beginning of the excerpt), the middle (selected to fall between 7-12 s), and the end (15-19.5 seconds) of each 20 s excerpt were selected. For each of these 2 s samples, the syllable rate was extracted and is reported in the table. The results indicate that there were no major rate variations across the 20 s excerpts. Overall, the speakers generated the samples at the rates typical for speech (cross-linguistically), in the range of 5 syllables per second. Syllable rates were determined by a native German speaker (DP).

Speaker-sent	***Samp 1(syl/s)***	***Samp 2(syl/s)***	***Samp 3(syl/s)***	***Mean rate (syl/s)***
Elis-Ex1	4.75	5.0	5.5	5.08
Elis-Ex2	6.0	6.0	4.5	5.5
Kath-Ex1	6.2	5.5	6.0	5.9
Kath-Ex2	6.5	5.0	6.25	5.92
Katr-Ex1	5.5	5.0	4.0	4.83
Katr-Ex2	4.5	5.0	5.0	4.83
Ulli-Ex1	5.0	4.5	6.0	5.16
Ulli-Ex2	6.0	5.0	5.0	5.3
Andr-Ex1	5.0	5.5	5.5	5.3
Andr-Ex2	6.0	6.0	7.5	6.5
Pasc-Ex1	5.5	6.0	5.5	5.6
Pasc-Ex2	6.5	5.5	5.5	5.83
Mich-Ex1	5.5	6.0	5.5	5.6
Mich-Ex2	5.5	5.0	4.5	5.0
Tobi-Ex1	5.0	5.5	4.5	5.0
Tobi-Ex2	6.0	5.5	4.5	5.3
Mean syllable rate across speakers and items				***5.41***