Supplementary Information

Method

Sample size justification

Because this is the first study to examine individual relationships between children's language exposure and fMRI measures of language-related brain activation, effect size estimates were not available to inform sample size. However, the behavioral correlations between language input quantity/quality and children's language scores are typically moderate to strong ($0.4 < r < 0.6$) (Hirsh-Pasek et al., 2015; Hoff, 2003; Rowe, 2012; Weisleder & Fernald, 2013). For 80% power to detect such an effect in the expected direction at $\alpha = 0.05$, one would need to recruit 15-36 participants. Similarly, a majority of studies investigating correlations between behavioral measures and fMRI activation using appropriate *independent analyses* report correlations in the 0.5 to 0.7 range, with a median of 0.6 (Vul, Harris, Winkielman, & Pashler, 2009, Figure 5). Given that individual differences analyses (i.e., correlational analyses) have lower power than within-subjects analyses (i.e., condition differences), common sample-size planning tools for fMRI studies are not appropriate for the present power analysis. Instead, power curves specific to Pearson's correlations in the context of fMRI were consulted (Yarkoni & Braver, 2010). By these estimates, one would need to recruit 15-30 participants for the same parameters stated above. Combined, a sample size of 15-36 is recommended to find expected behavioral and neural effects. However, because of the likelihood of publication bias in previously reported effects (e.g., Anderson, Kelley, & Maxwell, 2017), we aimed for the highest end of this range (n = 36).

Statistical Analysis

All statistical analyses (with exception of whole-brain fMRI analyses) were performed in IBM SPSS Statistics version 24. The first approach was to conduct zero-order Pearson's correlations between children's assessment scores, SES demographics, and LENA measures of language exposure. Because all three variables were intercorrelated, we conducted linear regressions to determine which independent variables predicted unique variance in children's language scores, while controlling for the other independent variables. Finally, we conducted bootstrapped mediation with 10,000 iterations using the PROCESS macro for SPSS (Hayes, 2013) to determine whether language exposure mediated the relationship between SES and children's language scores. The same bootstrapping approach was applied to neural activation measures extracted from a region of interest (see main text).

Executive Functioning measure

In addition to the standardized assessments described in the main text, children also completed a non-standardized executive functioning (EF) task. Because EF relies on prefrontal regions adjacent to/overlapping with frontal language regions, EF was included to serve as a covariate/nuisance variable. The Hearts and Flowers version of the dots task (Davidson, Amso, Anderson, & Diamond, 2006) is commonly used to assess EF in both children and adults, because it requires all three EF dimensions (working memory, inhibition, and cognitive flexibility/switching) with simple instructions. Children rested their hands on a handlebar adjusted to finger-distance away from a touch screen computer and completed a practice run of quickly pressing on-screen buttons in this way. Then, a red heart or flower would appear on the right or left side of the screen, and children were instructed to press the button on the *same* side as a heart (congruent condition) and the button on the *opposite* side of a flower (incongruent condition). The task consisted of three consecutive blocks: a congruent block of 12 trials, an incongruent block of 12 trials, and a randomly mixed block (congruent and incongruent) of 49 trials. For all conditions, stimuli were displayed for 500 milliseconds (ms) with 1500 ms to respond and an interstimulus interval of 500 ms. Any response faster than 200 ms were considered to be anticipatory (Davidson et al., 2006) and excluded from analyses. Children received up to 12 practice trials before the congruent and incongruent blocks to ensure understanding of the rule. No practice was included before the mixed block, and thus the first trial was additionally excluded from analyses. The main outcome measures were the average accuracy across all trials in the mixed block and average reaction time (RT) in milliseconds across all correctly answered trials in the mixed block. Although rare, accuracy scores below 50% were not discarded because they could have been obtained by rule reversal, indicating an error in cognitive flexibility/switching.

Mean accuracy on the EF task was 72%, (*SD* = 22%) with a mean reaction time on correctly answered trials of 1140 ms (*SD* = 207). None of the fMRI analyses – including group mean task activation and whole brain correlates with LENA measures – changed with the inclusion of EF scores as a nuisance variable. This suggests that correlations between conversational turns and activation in Broca's area are not driven by differences in executive functioning.

References

Anderson, S. F., Kelley, K., & Maxwell, S. E. (2017). Sample-size planning for more accurate statistical power: A method adjusting sample effect sizes for publication bias and uncertainty. *Psychol Sci*, 956797617723724.

Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia, 44*(11), 2037-2078.

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York, NY: The Guilford Press.

Hirsh-Pasek, K., Adamson, L. B., Bakeman, R., Owen, M. T., Golinkoff , R. M., Pace, A., et al. (2015). The contribution of early communication quality to low-income children's language success. *Psychol Sci, 26*(7), 1071-1083.

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Dev, 74*(5), 1368-1378.

Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. *Child Dev, 83*(5), 1762-1774.

Vul, E., Harris, C., Winkielman, P., & Pashler, H. (2009). Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition. *Perspect Psychol Sci, 4*(3), 274-290.

Weisleder, A., & Fernald, A. (2013). Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychol Sci, 24*(11), 2143-2152.

Yarkoni, T., & Braver, T. S. (2010). Cognitive Neuroscience Approaches to Individual Differences in Working Memory and Executive Control: Conceptual and Methodological Issues. In A. Gruszka, G. Matthews & B. Szymura (Eds.), *Handbook of Individual Differences in Cognition: Attention, Memory, and Executive Control* (pp. 87-107). New York, NY: Springer New York.

Table S1. Group mean task activations for forward > backward speech

Z-value	MNI coordinates x	y	z	Anatomical Description
Left Hemisphere Cluster (3552 voxels)				
6.40	-52	-7	-12	Anterior superior temporal sulcus
6.39	-55	-39	4	Posterior superior temporal sulcus
6.11	-50	9	-22	Temporal pole
5.81	-57	-51	30	Supramarginal gyrus
Right Hemisphere Cluster (1418 voxels)				
5.93	54	-4	-13	Anterior superior temporal sulcus
5.47	51	13	-20	Temporal pole

Coordinates and anatomical descriptions of local peak activations for the forward > backward speech contrast, averaged over the entire sample (n = 36). Analyses revealed two significant clusters, one in each hemisphere, visualized in Figure 4.

Table S2. Correlation between conversational turns and forward > backward activation

Z-value	MNI coordinates x	y	z	Anatomical Description
4.59	-48	33	15	Left posterior pars triangularis
4.18	-56	33	10	Left anterior pars triangularis
3.55	-43	13	15	Left anterior pars opercularis

Coordinates and anatomical descriptions of local peak activations in a single cluster (Figure 5a, 766 voxels) exhibiting a significant correlation between the number of conversational turns children experienced per hour and activation in the forward > backward speech contrast.