Modular reconfiguration of an auditory control brain network supports adaptive listening behavior

Mohsen Alavash1,2, Sarah Tune1, and Jonas Oleser1,2

*Department of Psychology, University of Lübeck, 23562 Lübeck, Germany

Edited by Ranulfo Romo, Universidad Nacional Autonoma de Mexico, Mexico City, D.F., Mexico, and approved November 27, 2018 (received for review September 10, 2018)

Speech comprehension in noisy, multitalker situations poses a challenge. Successful behavioral adaptation to a listening challenge often requires stronger engagement of auditory spatial attention and context-dependent semantic predictions. Human listeners differ substantially in the degree to which they adapt behaviorally and can listen successfully under such circumstances. How cortical networks embody this adaptation, particularly at the individual level, is currently unknown. We here explain this adaptation from reconfiguration of brain networks for a challenging listening task (i.e., a linguistic variant of the Posner paradigm with concurrent speech) in an age-varying sample of n = 49 healthy adults undergoing resting-state and task fMRI. We here provide evidence for the hypothesis that more successful listeners exhibit stronger task-specific reconfiguration (hence, better adaptation) of brain networks. From rest to task, brain networks become reconfigured toward more localized cortical processing characterized by higher topological segregation. This reconfiguration is dominated by the functional division of an auditory and a cingulo-opercular module and the emergence of a conjoined auditory and ventral attention module along bilateral middle and posterior temporal cortices. Supporting our hypothesis, the degree to which modularity of this frontotemporal auditory control network is increased relative to resting state predicts individuals’ listening success in states of divided and selective attention. Our findings elucidate how fine-tuned cortical communication dynamics shape selection and comprehension of speech. Our results highlight modularity of the auditory control network as a key organizational principle in cortical implementation of auditory spatial attention in challenging listening situations.

Significance

How do brain networks shape our listening behavior? We here develop and test the hypothesis that, during challenging listening situations, intrinsic brain networks are reconfigured to adapt to the listening demands and thus, to enable successful listening. We find that, relative to a task-free resting state, networks of the listening brain show higher segregation of temporal auditory, ventral attention, and frontal control regions known to be involved in speech processing, sound localization, and effortful listening. Importantly, the relative change in modularity of this auditory control network predicts individuals’ listening success. Our findings shed light on how cortical communication dynamics shape selection and comprehension of speech in challenging listening situations and suggest modularity as the network principle of auditory attention.

Author contributions: M.A., S.T., and J.O. designed research; M.A. and S.T. performed research; M.A. and S.T. analyzed data; and M.A., S.T., and J.O. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

Data deposition: The complete dataset associated with this work, including raw data and connectivity maps, as well as code of network and statistical analyses are publicly available at https://osf.io/28r57/.

1M.A. and S.T. contributed equally to this work.
2To whom correspondence may be addressed. Email: mohsen.alavash@uni-luebeck.de or Jonas.Oleser@uni-luebeck.de.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1815321116/-/DCSupplemental.

Published online December 26, 2018.
general Listening task and individuals  
vol. 116  
performance. is the slope parameter estimate 
49) represent January 8, 2019boot- 
661 2.1,
0.001] (Fig. 2  
β 0.9; response speed: < 2.5,
0.001). Furthermore, as to be expected from 
49; 19
= 0.05; response speed: <
P P
P P 0.63,
P
P–
= 0.005, 
χ 1.1,
< Right

Two possible network reconfigurations (case 1 and case 2) charac-
~4 ~8 R 
runch 6 R
Alavash et al.

processing. within the green module, a new connection emerges (red link). Together, 
Modularity describes the segregation of nodes into relatively dense subsys-
tics (here shown in distinct colors), which are sparsely interconnected. Mean local efficiency is equivalent to global efficiency computed on the direct neighbors of each node, which is then averaged over all nodes. Tow-
ard higher functional segregation, the hypothetical baseline network loses the shortcut between the blue and green module (dashed link). Instead, within the green module, a new connection emerges (red link). Together, these form a segregated configuration tuned for a more localized cortical processing.

units of speech (44, 45) would lead to a reconfiguration toward lower segregation (more distributed cortical processing) (Fig. 1).

As such, we here capitalize on the degree and direction of the functional connectome reconfiguration as a proxy for an individual’s successful adaptation and thus, speech comprehension success in a challenging listening task (Fig. 24).

Results

We measured brain activity using fMRI in an age-varying sample of healthy adults (n = 49; 19–69 y old) and constructed large-scale cortical network models (SI Appendix, Fig. S1) during rest and while individuals performed a linguistic Posner task with concurrent speech. The task presented participants with two competing dichotomous sentences uttered by the same female speaker (Fig. 2A). Participants were probed on the last word (i.e., target) of one of these two sentences. Crucially, two visual cues were presented upfront. Either a spatial attention cue indicated the to be probed side, thus invoking selective attention, or it was uninformative, thus invoking divided attention. The second cue specified the semantic category of the sentence final words either very generally or more specifically, which allowed for more or less precise semantic prediction of the upcoming target word. Using (generalized) linear mixed effects models, we examined the influence of the listening cues and brain network reconfiguration on listening success, controlling for individuals’ age, head motion, and the probed side.

Informative Spatial and Semantic Cues Facilitate Speech Comprehension.
The analysis of listening performance using linear mixed effects models revealed an overall behavioral benefit from more informative cues. As expected, participants performed more accurately and faster under selective attention compared with divided attention [accuracy: odds ratio (OR) = 2.1, P < 0.001; response speed: β = 0.1, P < 0.001] (Fig. 2B, Left). Moreover, participants performed more accurately and faster when they were cued to the specific semantic category of the target word compared with a general category (accuracy: OR = 1.1, P < 0.05; response speed: β = 0.03, P < 0.01) (Fig. 2B, Right). Participants benefited notably more from the informative spatial cue than from the informative (i.e., specific) semantic cue. We did not find any evidence of an interactive effect of the two cues on behavior (likelihood ratio tests; accuracy: χ² = 0.005, P = 0.9; response speed: χ² = 2.5, P = 0.1).

Lending validity to our results, younger participants performed more accurately and faster than older participants as expected (main effect of age; accuracy: OR = 0.63, P < 0.0001; response speed: β = −0.03, P < 0.001). Furthermore, as to be expected from
the well-established right ear advantage for linguistic materials (46), participants were faster when probed on the right compared with the left ear (main effect of probe: \( \beta = 0.018, P < 0.01 \)) (SI Appendix, Table S1 has all model terms and estimates).

Higher Segregation of the Whole-Brain Network During Cued Speech Comprehension. A main question in this study was whether and how resting-state brain networks are reconfigured when listening challenges arise. We compared functional connectivity and three key topological network metrics between resting state and the listening task (SI Appendix has the definitions of brain metrics). We expected that resting-state brain networks would functionally reshape and shift either toward higher segregation (more localized cortical processing) or lower segregation (more distributed cortical processing) during the listening task (Fig. 1). Our network comparison revealed how the whole-brain network reconfigured in adaptation to cued speech comprehension (Fig. 3).

First, while mean within-module functional connectivity of the whole-brain network showed a significant increase from rest to task (Cohen’s \( d = 0.53, P < 0.01 \)) (Fig. 3A, Center), between-module connectivity showed the opposite effect (Cohen’s \( d = -0.72, P < 0.001 \)) (Fig. 3A, Right). Importantly, resting-state and listening task networks did not differ in their overall mean functional connectivity (\( P = 0.6 \)) (Fig. 3A, Left), which emphasizes the importance of modular reconfiguration (i.e., change in network topology) in brain network adaptation in contrast to a mere alteration in overall functional connectivity (change in correlation strength).

Second, during the listening task, functional segregation of the whole-brain network increased relative to resting state. This effect was consistently observed on the local, intermediate, and global scales of network topology, a reconfiguration regime in the direction of the hypothetical case 1 depicted in Fig. 1. More specifically, global network efficiency—a metric that has been proposed to measure the capacity of brain networks for parallel processing (47)—decreased from rest to task (Cohen’s \( d = -0.47, P < 0.001 \)) (Fig. 3B, Left). In addition, modularity of brain networks—a metric related to the decomposability of a network into subsystems (48)—increased under the listening task in contrast to resting state (Cohen’s \( d = 1, P < 0.001 \)) (Fig. 3B, Center). Across participants, network modularity was higher in 43 of 49 individuals, with a considerable degree of interindividual variability in its magnitude. This result complements the significant changes in within- and between-module connectivity described earlier.

Moreover, the whole-brain network during the task was characterized by an increase in its local efficiency relative to resting state (Cohen’s \( d = 0.66, P < 0.001 \)) (Fig. 3B, Right). Knowing that local efficiency is inversely related to the topological distance between nodes within local cliques or clusters (48), this effect is another indication of a reconfiguration toward higher functional segregation, reflecting more localized cortical processing. The above results were robust with respect to the choice of network density (SI Appendix, Fig. S2). Similar results were found when brain graphs were constructed (and compared with rest) separately per task block (SI Appendix, Fig. S3), which rules out any systematic difference between resting state and task due to the difference in scan duration (the duration of each block is approximately the duration of resting state).

In an additional analysis, we investigated the correlation of brain network measures between resting state and listening task state. Reassuringly, the brain network measures were all positively correlated across participants (SI Appendix, Fig. S4), supporting the idea that state- and trait-like determinants both contribute to individuals’ brain network configurations (28).

We next investigated whether reconfiguration of the whole-brain networks from the rest to task can account for the interindividual variability in listening success. To this end, using the same linear mixed effects model that revealed the beneficial effects of the listening cues on behavior, we tested whether the change in each whole-brain network metric (i.e., task minus rest) as well as its modulation by the different cue conditions could predict a listener’s speech comprehension success.

We found a significant interaction between the spatial cue and the change in individuals’ whole-brain network modularity in predicting accuracy (OR = 1.13, \( P < 0.01 \)). This interaction was driven by a weak positive relationship of modularity and accuracy during selective attention trials (OR = 1.17, \( P = 0.13 \)) and the absence of such a relationship during divided attention trials (OR = 1.02, \( P = 0.8 \)) (SI Appendix, Tables S1 and S2). This result provides initial evidence that, at the whole-brain level, higher network modularity relative to resting state coincides with successful implementation of selective attention to speech. Notably, when changes in other network measures were included in the model (i.e., change in local or global network efficiency instead of change in network modularity), no significant relationships were found. Also, the effect of the semantic cue on behavior was not significantly modulated by change in whole-brain network topology. More precisely, adding an interaction term between the semantic cue and change in brain network modularity did not significantly improve the model fit (likelihood ratio tests; accuracy: \( \chi^2 = 2.6, P = 0.1 \); response speed: \( \chi^2 = 3.8, P = 0.06 \)). There was also no significant interaction between age and change in network modularity in predicting behavior (likelihood ratio tests; accuracy: \( \chi^2 = 1, P = 0.3 \); response speed: \( \chi^2 = 3.1, P = 0.08 \)).

In the following sections, we will explore the cortical origin of the topological reconfiguration outlined above: first, at the whole-brain level and second, at a subnetwork level in more detail. In the final section of the results, we will illustrate the behavioral relevance of these reconfigurations at the individual level.

Reconfiguration of the Whole-Brain Network in Adaptation to Cued Speech Comprehension. Fig. 4 provides a comprehensive overview of group-level brain networks, functional connectivity, and the corresponding connectograms (circular diagrams) during resting state (Fig. 4A) and during the listening task (Fig. 4C). In Fig. 4, cortical regions defined based on ref. 49 are grouped and color coded according to their module membership determined using the graph-theoretical consensus community detection algorithm (48, 50) (SI Appendix has details).
When applied to the resting-state data in our sample, the algorithm decomposed the group-averaged whole-brain network into six modules (Fig. 4A). The network module with the highest number of nodes largely overlapped with the known default mode network (Fig. 4A, dark blue). At the center of our investigation was the second largest module, comprising mostly auditory nodes and a number of cingulo-opercular nodes (Fig. 4A, yellow). For simplicity, we will refer to this module as the auditory module. Additional network modules included (in the order of size) a module made up of visual nodes (Fig. 4A, cyan) and one encompassing mostly somatomotor nodes (Fig. 4A, light blue) as well as a dorsal attention module (Fig. 4A, green) and a frontoparietal module (Fig. 4A, red).

Fig. 4B shows how the configuration of the whole-brain network changed from rest to task. This network flow diagram illustrates whether and how cortical nodes belonging to a given network module during resting state (Fig. 4B, Left) changed their module membership during the listening task (Fig. 4B, Right). The streamlines tell us how the nodal composition of network modules changed, indicating a functional reconfiguration. According to the streamlines, the auditory module showed the most prominent reconfiguration (Fig. 4B, yellow module) (auditory), while the other modules underwent only moderate reconfigurations.

This reconfiguration of the auditory module can be summarized by two dominant nodal alterations. On the one hand, a nodal “branching” arises: a number of cingulate, frontal, temporal, and insular nodes change their module membership in adaptation to the listening task. On the other hand, a group of temporal and insular nodes merges with other temporal nodes from the default mode network module (Fig. 4B, Left, dark blue) and forms a new common module during the listening task (Fig. 4B, Right, yellow). Note that nodal merging should not be mistaken for functional integration. In a graph-theoretical sense, higher functional integration would be reflected by stronger between-module connectivity, which is opposite to what is found in this study (Fig. 3).

Interestingly, this latter group of temporal nodes includes posterior temporal nodes that are functionally associated with the ventral attention network (21, 49, 51). Considering the cortical...
surface maps, these alterations occur at the vicinity of the middle and posterior portions of the superior temporal lobes (note the dark blue regions in the right superior temporal sulcus (STS)/superior temporal gyrus in resting-state maps, which turn yellow in the listening-state maps).

In line with the changes described above, the reconfiguration of the auditory module is also observed as a change in the connection patterns shown by the connectograms (Fig. 4 A and C, circular diagrams). Relative to rest, tuning in to the listening task condenses the auditory module, and its nodes have fewer connections to other modules. In addition, a group of cingulate and frontal nodes from the auditory module at rest merges with several nodes from the frontoparietal/dorsal attention module and forms a new module during the listening task (Fig. 4 B, Right, dark brown and C).

Taken together, we have observed a reconfiguration of the whole-brain network that is dominated by modular branching and merging across auditory, cingulo-opercular, and ventral attention nodes (Fig. 4B).

Reconfiguration of a Frontotemporal Auditory Control Network in Adaptation to Cued Speech Comprehension. As outlined above, we had observed two prominent reconfigurations described in Fig. 4B, namely nodal branching and nodal merging. Accordingly, we identified the cortical origin of these nodal alterations based on the underlying functional parcellation (49). The identified nodes include auditory, ventral attention, and cingulo-opercular regions. We will refer to these frontotemporal cortical regions collectively as the auditory control network. This network encompasses 86 nodes (of 333 whole-brain nodes). According to Fig. 4B, this network is in fact a conjunction across all auditory, ventral attention, and cingulo-opercular nodes, irrespective of their module membership. For the purpose of a more transparent illustration, the cortical map of the auditory control network is visualized in Fig. 5A.

Similar to the whole-brain analysis (Fig. 4), the auditory control network was decomposed into modules by applying the consensus community detection algorithm again but this time, to the graph encompassing the 86 nodes. The result is visualized in Fig. 5A by grouping the nodes within the group-level functional connectivity matrix and the corresponding connectogram. During resting state, the consensus community detection algorithm had decomposed the auditory control network into four distinct modules that correspond well with their functional differentiation as in ref. 49. That is, auditory, ventral attention, and cingulo-opercular nodes form their own distinct homogenous modules.

During the listening task, however, the configuration of the auditory control network changed (Fig. 5B, Upper). According to the streamlines, the majority of auditory nodes (yellow) and posterior ventral attention nodes (orange) merged to form a common network module during the task (Fig. 5B, Right, first module from bottom). This reconfiguration co-occurred with an increase in functional connectivity within the auditory nodes and ventral attention nodes as well as between auditory and ventral attention nodes (Fig. 5B, Lower and SI Appendix, Fig. S5). The cortical and connectivity maps of the auditory–ventral attention module are shown in Fig. 5C, dark blue module.

This network module emerging during the listening task involves cortical regions mostly in the vicinity of the bilateral STS and Heschl’s gyrus as well as the right angular gyrus. Note also the increase in the functional connectivity across these nodes (Fig. 5C, connectivity matrix, dark blue module) and the increase in the number of connections within this module (Fig. 5C, connectogram, gray bars inside the dark blue module) (all relative to resting state). Collectively, this reconfiguration indicates a stronger functional cross-talk between auditory and posterior ventral attention regions during the listening task compared with resting state.

In contrast, frontal ventral attention nodes do not merge with other nodes and remain connected within a separate module during both resting state and task (Fig. 5B, Right, second module from top). These nodes overlap with cortical regions at the vicinity of inferior frontal gyrus (IFG) (Fig. 5C, light green module). Moreover, the majority of cingulo-opercular nodes (Fig. 5C, red) are grouped together during the listening task (Fig. 5C, light blue module). This co-occurs with a decrease in functional connectivity within the cingulo-opercular nodes as well as between auditory and cingulo-opercular nodes (Fig. 5B, Lower and SI Appendix, Fig. S5). This indicates higher functional segregation within the cingulo-opercular regions during the listening challenge relative to resting state.

Higher Modularity of the Auditory Control Network Predicts Speech Comprehension Success. To quantitatively assess the reconfiguration of the auditory control network described above, we compared functional connectivity and topology of this network between resting state and the listening task (Fig. 6A).

First, while mean overall connectivity and within-module connectivity of this network significantly increased from rest to task (mean connectivity: Cohen’s d = 0.60, P < 0.01; within-module connectivity: Cohen’s d = 0.7, P < 0.01), between-module connectivity showed the opposite effect (Cohen’s d = −1.15, P < 0.001) (Fig. 6A).

Second, functional segregation of the auditory control network was significantly increased from rest to task as manifested by a higher degree of modularity and mean local efficiency (modularity: Cohen’s d = 2.13, P < 0.001; local efficiency: Cohen’s d = 0.9, P < 0.001) (Fig. 6B).

Critically, (generalized) linear mixed effects analyses revealed significant interactions of the spatial cue and change in modularity of the auditory control network for both behavioral dependent measures. More precisely, the change in modularity of this network showed positive correlations with accuracy (Fig. 6C) and response speed (Fig. 6D) under both divided and selective attention trials. However, these brain–behavior correlations were stronger under selective (accuracy: OR = 1.37, P < 0.01; response speed: β = 0.04, P < 0.01) than under divided attention trials (accuracy: OR = 1.23, P < 0.05; response speed: β = 0.02, P < 0.05) (SI Appendix, Table S3 has all model terms and estimates). We did not find any evidence for an interaction between the semantic cue and change in modularity of the auditory control network in predicting behavior (likelihood ratio tests; accuracy: χ² = 0.3, P = 0.6; response speed: χ² = 0.4, P = 0.5). There was also no significant interaction between age and change in modularity of the auditory control network in predicting behavior (accuracy: χ² = 0.15, P = 0.7; response speed: χ² = 0.4, P = 0.5) (SI Appendix, Fig. S6).

Moreover, an additional linear mixed effects model with resting-state modularity and task modularity as two separate regressors revealed a significant main effect of the auditory control network modularity during the task (β = 0.05, P < 0.05) and an interaction between semantic cue and change in modularity of the auditory control network during the task with the spatial cue in predicting response speed (β = 0.01, P < 0.05) (SI Appendix, Table S4).

Apart from the nodal branching and merging in the auditory module, other reconfigurations, although less profound, could also be identified (Fig. 4B). These include modular reconfiguration of subnetworks encompassing (i) visual, dorsal attention, and cingulo-opercular nodes; (ii) somatomotor, auditory, and cingulo-opercular nodes; (iii) default mode, frontoparietal, and ventral attention nodes; and (iv) dorsal attention, frontoparietal, and cingulo-opercular nodes. To investigate whether these reconfigurations predict listening behavior, we performed analyses analogous to those undertaken for the auditory control network. That is, using four separate linear mixed effects models, we included the change (i.e., task minus rest) in network modularity of each of the four aforementioned subnetworks. None of these models were able to predict individuals’ listening behavior (accuracy or response speed). Thus, the prediction of speech comprehension success from modular reconfiguration of brain subnetworks was specific to the frontotemporal auditory control network (Fig. 6B).

Discussion

First, we have shown here how resting-state brain networks reshape in adaptation to a challenging listening task. This reconfiguration...
predominantly emerges from a frontotemporal auditory control network that shows higher modularity and local efficiency during the listening task, shifting the whole-brain connectome toward higher segregation (more localized cortical processing). Second, the degree to which modularity of this auditory control network increases relative to its resting-state baseline predicts an individual’s listening success. To illustrate, one SD change in modularity of the auditory control network increases relative to its resting-state baseline predicts an individual’s listening success. To illustrate, one SD change in modularity of the auditory control network increased a listener’s chance of successful listening by about 3% on average, and it made her response about 40 ms faster. While the effect of modularity on behavior was found in states of both divided and selective attention, it was stronger for selective attention trials.

**Optimal Brain Network Configuration for Successful Listening.** The functional connectome was reconfigured toward more localized cortical processing during the listening task as reflected by higher segregation of the whole-brain network. This functional specialization suggests that resolving the listening task mainly required attending to and processing the sentences as unconnected speech, arguably at the level of single words, which is often
Functional Segregation Within the Auditory Control Network Is Critical for Cued Speech Comprehension. In adaptation to the listening task, a segregated module emerged along bilateral STS and superior temporal gyrus with a posterior extension (Fig. 4A and C; cortical maps and Fig. 5C, dark blue module). This reconfiguration is in line with the proposal of a ventral stream in the functional neuroanatomy of speech processing (also referred to as the auditory what stream), which has been suggested to be involved in sound-to-meaning mapping during speech recognition tasks (52–56). In addition, the functional convergence of the auditory nodes with the posterior ventral attention nodes parallels the proposal that the posterior portion of the primary auditory cortex is involved in spatial separation of sound and speech (57–61).

Our connectomic approach thus extends the existing evidence by demonstrating an increased functional cross-talk between neural systems involved in speech processing and sound localization during a cued speech comprehension task. Notably, the functional significance of this task-driven network reconfiguration was corroborated by its impact on listening success, in particular in states of selective attention. During selective attention trials, we assume that an active attentional filter mechanism is being invoked to “tune out” the distracting sentence presented to the opposite ear and to selectively amplify the sentence presented to the cued ear. Thus, the stronger brain–behavior correlation under selective attention trials suggests that this flexible filter is implemented by the formation of a more segregated and specialized cortical module composed of auditory and ventral attention nodes.

Another key observation here was auditory regions in anterior and superior temporal cortices as well as anterior insula and cingulate cortex merging with the somatmotor module during the listening task (Fig. 4C, light blue). In parallel, the frontal ventral attention nodes in the vicinity of the IFG remained within the default mode module (Fig. 4C, dark blue). Within the auditory control network, these reconfigurations appeared as the formation of highly segregated modules (Fig. 5C; modules except the dark blue one). Importantly, the degree of the modular reconfiguration of cingulo-opercular and ventral attention nodes was predictive of listening success only when their network interactions were modeled in combination with the auditory nodes (i.e., the auditory control network) and not with somatomotor or default mode nodes.

Our findings concur conceptually with the earlier work by Vaden et al. (13), who showed that, during a word recognition experiment, regions within the cingulo-opercular network displayed elevated activation in response to speech in noise, and regions along the ventral attention network showed the opposite transitions between trials. Moreover, the authors found that cingulo-opercular regions exhibited higher functional connectivity relative to silence periods. Our results refine their findings and provide a clear distinction between cingulo-opercular and ventral attention nodes in their communication with auditory nodes during cued speech processing (Fig. 5B). Our findings support the hypothesis that, during a listening challenge, cingulo-opercular regions form an autonomous “core” control system that monitors effortful listening performance (62–64), while ventral attention nodes implement attentional filtering of the irrelevant speech (8).

Our connectomic approach can be distinguished from previous activation studies in two aspects. First, we view the neural systems involved in challenging speech comprehension as a large-scale dynamic network with configuration that changes from idling or resting-state baseline to a high-duty listening task state. This allowed us to precisely delineate the cortical network embodiment of behavioral adaptation to a listening challenge.

Second, our graph-theoretical module identification unveiled a more complete picture (i.e., nodes, connections, and functional boundaries) of cortical subsystems supporting listening success. This allowed us to identify and characterize an auditory control network with modular organization that revealed how functional coordination within and between auditory, cingulo-opercular, and ventral attention regions controls the selection and processing of speech during a listening challenge.

-associated with cortical regions localized to auditory, cingulo-opercular, and inferior frontal regions under adverse listening situations (1, 13, 43).

However, segregated topology is not necessarily the optimal brain network configuration to resolve every difficult speech comprehension task. Instead, we argue that the most favorable configuration depends on the nature and instruction of the task, in particular on the level of linguistic and semantic information that needs to be integrated to resolve the task (cf. refs. 41, 44, and 45). Indeed, previous studies in the contexts of working memory and complex reasoning have found a shift of the functional connectome toward higher network integration (33, 36). Along these lines of research, our findings emphasize the importance of the degree (rather than the direction) of modular reconfiguration of brain networks in predicting individuals’ behavior.
Is Modularity the Network Principle of Auditory Attention? The main finding of this study is that higher modularity of the auditory control network relative to resting state predicts individuals’ listening success in states of both divided and selective attention. Among the three network metrics, only modularity linked the reconfiguration of brain networks with this behavioral utilization of a spatial cue. This suggests that auditory spatial attention in adverse listening situations is implemented across cortex by a fine-tuned communication within and between neural assemblies across auditory, ventral attention, and cingulo-opercular regions. This is in line with recent studies showing that higher modularity of brain networks supports auditory perception or decision making (65, 66).

Inherent in the principle of hierarchy, modularity allows formation of complex architectures that facilitate functional specificity, robustness, and behavioral adaptation (67, 68). Our results suggest that individuals’ ability in modifying intrinsic brain network modules is crucial for their attention to speech and ultimately, successful adaptation to a listening challenge. Notably, our results help us better understand individual differences in speech comprehension at the systems level, with implications for hearing assistive devices and rehabilitation strategies.

Limitations. The logic of modular reconfiguration of an auditory control brain network in individual listeners implies that it is possible to get reliable individual-specific estimates of network modularity. Recently, Gordon et al. (69) demonstrated that reliability of functional connectivity differs greatly across individuals, measures of interest, and anatomical locations. Important to this study, they showed that network modularity achieved an average difference of <3% from the split data sample with only 10 min of data, while other network measures required longer scan duration. It is possible that the relatively short duration of the resting-state scan in this study may be hiding some significant individual variability. Future studies will be important to recover such important variation by dense sampling of individual brains (70, 71).

Knowing that auditory spatiotemporal attention in adverse listening situations offers an indirect measure of neural activations, it remains unknown how far our findings generalize to networks constructed based on brain neurophysiological responses recorded during resting state and listening (72). In this respect, an important question is how network coupling between neural oscillations relates to listening behavior (73, 74). In addition, although the sentences used in our linguistic Posner task were carefully designed to experimentally control individuals’ use of spatial and semantic cues during listening, their trial-by-trial presentation makes the listening task less naturalistic. Future studies will be important to investigate the brain network correlates of listening success during attention to continuous speech. Lastly, in this study, the brain networks were constructed based on correlations between regional time series, which disregard directionality in the functional influences that regions may have on one another. Therefore, investigating causal relationships across cortical regions within the auditory control network would help us better understand the (bottom-up vs. top-down) functional cross-talks between auditory, ventral attention, and cingulo-opercular regions during challenging speech comprehension.

Conclusion. In sum, our findings suggest that behavioral adaptation of a listener hinges on task-specific reconfiguration of the functional connectome at the level of network modules. We conclude that the communication dynamics within the fronto-temporal auditory control network are closely linked with the successful deployment of auditory spatial attention in challenging listening situations.

Materials and Methods

Participants. Seventy-one healthy adult participants were invited from a larger cohort in which we investigate the neurocognitive mechanisms underlying adaptive listening in middle-aged and older adults [a large-scale study entitled “The listening challenge: How ageing brain adapt (AUDADAPT)” (75)]. All participants were right-handed [assessed by a translated version of the Edinburgh Handedness Inventory (75)] and had normal or corrected-to-normal vision. None of the participants had any history of neurological or psychiatric disorders. However, three participants had to be excluded from the study due to an incidental diagnosis made by an anesthesiologist based on the acquired structural MRI. In addition, 15 participants were excluded because of incomplete measurements due to technical issues, claustrophobia, or task performance below chance level. Four additional participants completed the experiment but were removed from data analysis due to excessive head motion inside the scanner (i.e., scan-to-scan movement >1.5 mm of translation or 1.3° of rotation; rms of framewise displacement >0.5 mm). Accordingly, 49 participants were included in the main analysis (mean age of 45.6 y old; age range = 19–69 y old; 37 females). All procedures were in accordance with the Declaration of Helsinki and approved by the local ethics committee of the University of Lübeck. All participants gave written informed consent and were financially compensated (86/h).

Procedure. Each imaging session consisted of seven fMRI runs (Fig. 2A): eyes-open resting state (~8 min) followed by six runs in which participants performed a challenging speech comprehension task (~9 min each) (SI Appendix has a detailed description of the task). A structural scan (~5 min) was acquired at the end of the imaging session (SI Appendix has MRI acquisition parameters).

Preprocessing. During the resting state, 480 functional volumes were acquired, and during each run of the listening task, 540 volumes were acquired. To allow for signal equilibrium, the first 10 volumes of resting state and each run of the task were removed. Preprocessing steps were undertaken in SPM12. First, the functional images were spatially realigned to correct for head motion using a least square approach and the six rigid body affine transformations. Second, the functional volumes were corrected for slice timing to adjust for differences in image acquisition times between slices. Subsequently, functional images were coregistered to each individual’s structural image. Third, the normalization parameters from the subject’s native space to the standard space (Montreal Neurological Institute and Hospital (MNI) coordinate system) were obtained via the unified segmentation of their T1 images (76). The normalization parameters were then applied to each of the echo-planar image (EPI) volumes. No spatial smoothing was applied on the volumes to avoid potential artificial distance-dependent correlations between voxels’ BOLD signals (77, 78).

Motion Artifact Removal. In recent years, the potentially confounding impact of head motion or nonneural physiological trends on the temporal correlations between BOLD signals has been raised by multiple studies (79, 80). Knowing that older participants often make larger within-scan movements than younger participants (28, 81), the issue of motion is particularly important in this study, where the age of the participants spans across a wide range. Thus, we accounted for motion artifacts in three stages of our analysis. First, we used the denoising method proposed by ref. 82 that decomposes brain signals in wavelet space, identifies and removes trains of spikes that likely reflect motion artifacts, and reconstructs denoised signals from the remaining wavelets. Second, we implemented a nuisance regression to suppress nonneural and motion artifacts as explained in the next section. Third, we used rms of framewise displacement as a regressor of no interest in the mixed effects analyses (Statistical Analysis).

Cortical Parcellation. Cortical nodes were defined using a previously established functional parcellation (49), which has shown relatively higher accuracy compared with other parcellations (83). This parcellation is based on surface-based resting-state functional connectivity boundary maps and encompasses 333 cortical areas. We used this template to estimate mean BOLD signals across voxels within each cortical region per participant.

Nuisance Regression. To minimize the effects of spurious temporal correlations induced by physiological and movement artifacts, a general linear model was constructed to regress out white matter and cerebrospinal fluid time series together with the six rigid body movement parameters and their first-order temporal derivatives. Subsequently, the residual time series obtained from this procedure were concatenated across the six runs of the listening task for each participant. The cortical time series of the resting state and listening task were further used for functional connectivity and network analysis (SI Appendix, Fig. S1). Note that global signal was not regressed out from the regional time series, as it is still an open question in the field whether global signal regression in fact regresses out (84–86). In this study, wavelet despiking of motion artifacts allowed us to accommodate the spatial and temporal heterogeneity of motion artifacts. Importantly, our main results point to the reconfiguration of a frontotemporal auditory control network, which does not overlap with the mainly sensorimotor disruption map of the global signal (84).
Generalized Linear Mixed Effects Models. Brain–behavior relationship was investigated within a linear mixed effects analysis framework. To this end, either the single-trial behavioral measures (accuracy or response speed) across all participants was treated as the dependent variable. The main predictors in the model were the spatial and semantic cues, each at two levels (divided vs. selective and general vs. specific, respectively). Since in this study, the main question is whether reconfiguration of resting-state brain networks in adaptation to the listening challenge relates to listening success, we used the difference between a given network parameter across resting state and the listening task (i.e., task minus rest) as the third main predictor. The linear mixed effects analysis framework allowed us to control for other variables that entered as regressors of no interest in the model. These included age, head motion (i.e., difference in rms of framewise displacement between resting state and task), and side probed (left or right). Mixed effects analyses were performed in R (R Core Team) using the packages lme4 and sjPlot.

Model Selection. To estimate the best fit of the linear mixed effects models, we followed an iterative model fitting procedure, which started with an intercept-only null model (7). Fixed effects terms were added in a stepwise procedure, and the change of model fit (performed using maximum likelihood estimation) was assessed using likelihood ratio tests. Deviation coding was used for categorical predictors. All continuous variables were Z scored. Selection of terms was a binary process, which collapsed out a linear effects model with the logit as link function. In the case of response speed, the distribution of the data did not significantly deviate from normal distribution (Shapiro–Wilk test P value = 0.7). Thus, in this case, we used linear mixed effects models with the underlying distribution set to Gaussian. For both models predicting accuracy or response speed, we report P values for individual fixed-effect terms that were derived using the Satterthwaite approximation for degrees of freedom. As a measure of effects size, for the model predicting correctness, we report OR, and for response speed, we report the regression coefficient (β).

Data Visualization. Brain surfaces were visualized using the Connectome Workbench. Connectograms were visualized using the Brain Data Viewer. Network flow diagrams were visualized in MapEquation.

Data Availability. The complete dataset associated with this work, including raw data and connectivity maps, as well as code of network and statistical analyses are publicly available at https://osf.io/2br57/

ACKNOWLEDGMENTS. Anne Herrmann, Susanne Shellebach, Franziska Scharata, and Leonhard Waschke helped acquire, manage, and archive the data. Research was supported by the European Research Council Consolidator Grant AUDADAPT 646696 to (J.O.).

1. Olesen J, Wise RJ, Dresner MA, Scott SK (2007) Functional integration across brain regions improves speech perception under adverse listening conditions. J Neurosci 27: 2283–2289.
2. Davis MH, Ford MA, Kherif F, Johnsrude IS (2011) Does semantic context benefit speech perception? a new tool to detect “top-down” processes? Psychol Bull 137:699–709.
3. Wöstemann M, Herrmann B, Maess B, Olesen J (2016) Spatiotemporal dynamics of auditory attention synchronize with speech. Proc Natl Acad Sci USA 113:3873–3878.
4. Dai L, Best V, Shinn-Cunningham BG (2018) Sensorineural hearing loss degrades biorhythms and hemispheric asymmetry. J Cogn Neurosci 30:2530–2541.
5. Finn ES, et al. (2015) Functional connectome fingerprinting: Identifying individuals from their brain connectivity. Nat Neurosci 18:1204–1211.
6. Kidd GR, Watson CS, Gygi B (2007) Individual differences in auditory abilities. J Acoust Soc Am 121:1539–1558.
7. Power JD, et al. (2011) Functional network organization of the human brain. Proc Natl Acad Sci USA 108:18101–18106.
8. Fedorenko E, Thompson-Schill SL (2014) Reworking the language network. Trends Cogn Sci 18:120–126.
9. Hagoort P (2014) Nodes and networks in the neural architecture for language: Broca’s region and beyond. Curr Opin Neurobiol 28:136–141.
10. Fujimoto S, Horwitz B, Simonny K (2015) The functional connectivity of speech control. PLoS Biol 13:e1002209.
11. Chai LR, Mattar MG, Blank IA, Fedorenko E, Basser DT (2016) Functional network dynamics of the language system. Cereb Cortex 26:4148–4159.
12. Price C, Thierry G, Grifﬁths T (2000) Speech-speciﬁc auditory processing: Where is it? Trends Cogn Sci 4:271–276.
13. Raichle ME, et al. (2001) A default mode of brain function. Proc Natl Acad Sci USA 98: 676–682.
14. Dosenbach NU, et al. (2007) Distinct brain networks for adaptive and stable task control in humans. Proc Natl Acad Sci USA 104:11073–11078.
15. Power JD, et al. (2011) Functional network organization of the human brain. Neuroinformatics 7:665–678.
16. Shen K, Hutchison RM, Bearden C, Mcintosh AR (2015) Network structure shapes spontaneous functional connectivity dynamics. J Neurosci 35:5579–5588.
17. Laumann TO, et al. (2017) On the stability of BOLD fMRI correlations. Cereb Cortex 27: 4719–4732.
18. Knyazeva MG, et al. (2014) EEG-based connectivity analysis of the human brain. Neuro Image 96:594–607.
19. Price C, Thierry G, Grifﬁths T (2000) Speech-speciﬁc auditory processing: Where is it? Trends Cogn Sci 4:271–276.
20. Raichle ME, et al. (2001) A default mode of brain function. Proc Natl Acad Sci USA 98: 676–682.
21. Dosenbach NU, et al. (2007) Distinct brain networks for adaptive and stable task control in humans. Proc Natl Acad Sci USA 104:11073–11078.
22. Power JD, et al. (2011) Functional network organization of the human brain. Neuroinformatics 7:665–678.
23. Shen K, Hutchison RM, Bearden C, Mcintosh AR (2015) Network structure shapes spontaneous functional connectivity dynamics. J Neurosci 35:5579–5588.
24. Laumann TO, et al. (2017) On the stability of BOLD fMRI correlations. Cereb Cortex 27: 4719–4732.
25. Mueller S, et al. (2013) Individual variability in functional connectivity architecture of the human brain. Neuron 77:586–595.
26. Cole MW, Bassett DS, Power JD, Braver TS, Petersen SE (2014) Intrinsic and task-evoked network architectures of the human brain. Neuron 83:238–251.
27. Finn ES, et al. (2015) Functional connectome fingerprinting: Identifying individuals and networks using patterns of brain connectivity. Nat Neurosci 18:1664–1671.
28. Georgii L, Rubinow M, Cavanagh MR, Henson RN (2015) State and trait components of functional connectivity: Individual differences vary with mental state. J Neurosci 35:13949–13961.
29. Gratton C, Laumann TO, Gordon EM, Adeyemo B, Petersen SE (2016) Evidence for two independent factors that modify brain networks to meet task goals. Cell Rep 17: 1276–1288.

30. Gordon EM, Laumann TO, Adeyemo B, Petersen SE (2017) Individual variability of the system-level organization of the human brain. Cereb Cortex 27:386–399.

31. Gordon EM, et al. (2017) Individual-specific features of brain systems identified with resting state functional correlations. Neuroimage 146:918–939.

32. Spadone S, et al. (2015) Dynamic reorganization of human resting-state networks during visuospatial attention. Proc Natl Acad Sci USA 112:8112–8117.

33. Oaten JD, D’Esposito M (2016) The segregation and integration of distinct brain networks and their relationship to cognition. J Neurosci 36:12083–12094.

34. Schultz DH, Cole MW (2016) Higher intelligence is associated with less task-related brain network reconfiguration. J Neurosci 36:8551–8561.

35. Bolt T, Nor J, Sj, Rubini M, Uddin LQ (2017) Correspondence between evoked and intrinsic functional brain network configurations. Hum Brain Map 38:1992–2007.

36. Hearne LJ, Cocchi L, Zalesky A, Mattingley JB (2017) Reconfiguration of brain network architectures between resting-state and complexity-dependent cognitive reasoning. J Neurosci 37:387–399.

37. Gratton C, et al. (2018) Functional brain networks are dominated by stable group and individual factors, not cognitive or demographic variation. Neuro 98:439–452.e5.

38. Tomasi D, Volkow ND (2012) Laterality patterns of brain functional connectivity: Gender effects. Front Psychol 3:27.

39. Muller AM, Meyer M (2014) Language in the brain at rest: New insights from resting state data and graph theoretical analysis. Front Hum Brain 8:228.

40. Tavor I, et al. (2016) Task-free fMRI predicts individual differences in brain activity during task performance. Science 352:216–222.

41. Peele JE (2012) The hemispheric lateralization of speech processing depends on what “speech” is: A hierarchical perspective. Front Hum Neurosci 6:309.

42. Sporns O (2013) Network attributes for segregation and integration in the human brain. Cereb Cortex 23:162–171.

43. DeWitt I, Rauschecker JP (2012) Phoneme and word recognition in the auditory ventral stream. Proc Natl Acad Sci USA 109:E505–E514.

44. Lerner Y, Honey CJ, Silbert LJ, Hasson U (2011) Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. J Neurosci 31:2906–2915.

45. de Koninck Y, et al. (2009) The hierarchical cortical organization of human speech processing. J Neurosci 33:6539–6557.

46. Hugdahl K (2003) The Asymmetrical Brain (Cambridge, MA), pp 441–475.

47. Bullmore E, Sporns O (2009) Complex brain networks: Graph theoretical analysis of brain connectivity. Annu Rev Neurosci 32:183–200.

48. Rubinov M, Sporns O (2010) Complex network measures of brain connectivity: Uses and interpretations. Neuroimage 52:1059–1069.

49. Muller AM, Meyer M (2014) Language in the brain at rest: New insights from resting state data and graph theoretical analysis. Front Hum Brain 8:228.

50. Corbetta M, Patel G, Shuman GL (2008) The reorienting of the human brain: from environment to theory of mind. Neuro 58:306–324.

51. Alain C, Arnott SR, Hevenor S, Grafton ST, Grady CL (2001) “What” and “where” in the human auditory cortext. Proc Natl Acad Sci USA 98:12301–12306.

52. Alain C, Arnott SR, Hevenor S, Grafton ST, Grady CL (2001) “What” and “where” in the human auditory cortext. Proc Natl Acad Sci USA 98:12301–12306.

53. Alain C, Arnott SR, Hevenor S, Grafton ST, Grady CL (2001) “What” and “where” in the human auditory cortext. Proc Natl Acad Sci USA 98:12301–12306.

54. Alain C, Arnott SR, Hevenor S, Grafton ST, Grady CL (2001) “What” and “where” in the human auditory cortext. Proc Natl Acad Sci USA 98:12301–12306.

55. Alain C, Arnott SR, Hevenor S, Grafton ST, Grady CL (2001) “What” and “where” in the human auditory cortext. Proc Natl Acad Sci USA 98:12301–12306.

56. Alain C, Arnott SR, Hevenor S, Grafton ST, Grady CL (2001) “What” and “where” in the human auditory cortext. Proc Natl Acad Sci USA 98:12301–12306.

57. Alain C, Arnott SR, Hevenor S, Grafton ST, Grady CL (2001) “What” and “where” in the human auditory cortext. Proc Natl Acad Sci USA 98:12301–12306.

58. Alain C, Arnott SR, Hevenor S, Grafton ST, Grady CL (2001) “What” and “where” in the human auditory cortext. Proc Natl Acad Sci USA 98:12301–12306.

59. Alain C, Arnott SR, Hevenor S, Grafton ST, Grady CL (2001) “What” and “where” in the human auditory cortext. Proc Natl Acad Sci USA 98:12301–12306.

60. Alain C, Arnott SR, Hevenor S, Grafton ST, Grady CL (2001) “What” and “where” in the human auditory cortext. Proc Natl Acad Sci USA 98:12301–12306.

61. Alain C, Arnott SR, Hevenor S, Grafton ST, Grady CL (2001) “What” and “where” in the human auditory cortext. Proc Natl Acad Sci USA 98:12301–12306.

62. Alain C, Arnott SR, Hevenor S, Grafton ST, Grady CL (2001) “What” and “where” in the human auditory cortext. Proc Natl Acad Sci USA 98:12301–12306.

63. Alain C, Arnott SR, Hevenor S, Grafton ST, Grady CL (2001) “What” and “where” in the human auditory cortext. Proc Natl Acad Sci USA 98:12301–12306.

64. Alain C, Arnott SR, Hevenor S, Grafton ST, Grady CL (2001) “What” and “where” in the human auditory cortext. Proc Natl Acad Sci USA 98:12301–12306.

65. Alain C, Arnott SR, Hevenor S, Grafton ST, Grady CL (2001) “What” and “where” in the human auditory cortext. Proc Natl Acad Sci USA 98:12301–12306.

66. Alain C, Arnott SR, Hevenor S, Grafton ST, Grady CL (2001) “What” and “where” in the human auditory cortext. Proc Natl Acad Sci USA 98:12301–12306.