



BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

Causal contributions of the domain-general (Multiple Demand) and the language-selective brain networks to perceptual and semantic challenges in speech comprehension

Lucy J. MacGregor^{1*}, Rebecca A. Gilbert¹, Zuzanna Balewski², Daniel J. Mitchell¹, Sharon W. Erzinçlioğlu¹, Jennifer M. Rodd³, John Duncan¹, Evelina Fedorenko⁴, Matthew H. Davis¹

¹ Medical Research Council Cognition and Brain Sciences Unit, University of Cambridge

² University of California at Berkeley

³ University College London

⁴ Massachusetts Institute of Technology

*Lucy.MacGregor@mrc-cbu.cam.ac.uk

Acknowledgements

We would like to thank Rahel Schumacher and Vitor Zimmerer for discussion and Peter Watson for statistical advice. We also thank two anonymous reviewers for helpful comments on earlier versions of the manuscript.

Conflict of interest

Authors report no conflict of interest

Funding Sources

This work was supported by Medical Research Council Grants MC_UU_00005/5 to MHD and MC_UU_0005/6 to JD; by NIH Awards R01-DC016950 and R01-DC016607 to EF; and by research funds from the McGovern Institute for Brain Research to EF. JR is funded by the Economic and Social Research Council Grant ES/S009752/1

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

Abstract

Listening to spoken language engages domain-general Multiple Demand (MD, fronto-parietal) regions of the human brain, in addition to domain-selective (fronto-temporal) language regions, particularly when comprehension is challenging. However, there is limited evidence that the MD network makes a functional contribution to core aspects of understanding language. In a behavioural study of volunteers (n=19) with chronic brain lesions, but without aphasia, we assessed the causal role of these networks in perceiving, comprehending and adapting to spoken sentences made more challenging by acoustic-degradation or lexico-semantic ambiguity. We measured perception of and adaptation to acoustically-degraded (noise-vocoded) sentences with a word report task before and after training. Participants with greater damage to MD but not language regions required more vocoder channels to achieve 50% word report, indicating impaired perception. Perception improved following training, reflecting adaptation to acoustic degradation, but adaptation was unrelated to lesion location or extent. Comprehension of spoken sentences with semantically ambiguous words was measured with a sentence coherence judgement task. Accuracy was high and unaffected by lesion location or extent. Adaptation to semantic ambiguity was measured in a subsequent word association task, which showed that availability of lower-frequency meanings of ambiguous words increased following their comprehension (word-meaning priming). Word-meaning priming was reduced for participants with greater damage to language but not MD regions. Language and MD networks make dissociable contributions to challenging speech comprehension: using recent experience to update word meaning preferences depends on language-selective regions, whereas the domain-general MD network plays a causal role in reporting words from degraded speech.

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

Key words

Speech perception, language comprehension, semantic ambiguity, lesion, adaptation, learning, Multiple Demand system

Introduction

During speech comprehension, listeners are continually challenged by various aspects of the input, which leads to uncertainty at multiple levels of the linguistic hierarchy. For example, acoustic challenges arise when speech is quiet, in an unfamiliar accent, produced by a young child who has not yet mastered articulation, or otherwise degraded. In such cases, perception of the individual phonemes and lexical forms is more uncertain. Linguistic challenges arise when there is lexical-semantic or syntactic ambiguity, or complexity from low-frequency words or constructions, such that the intended meaning is unclear. To resolve these uncertainties during speech comprehension, listeners make use of diverse sources of information (Altmann & Kamide, 1999; Cutler, Dahan, & van Donselaar, 1997; Garrod & Pickering, 2004; Hagoort, Hald, Bastiaansen, & Petersson, 2004; Münster & Knoeferle, 2018; Özyürek, 2014; Van Berkum, 2009; Zhang, Frassinelli, Tuomainen, Skipper, & Vigliocco, 2021). Furthermore, listeners learn in response to their experiences: they show perceptual and semantic adaptation such that improvements in the perception and comprehension of different types of challenging speech can be observed over time (Davis, Johnsruide, Hervais-Adelman, Taylor, & McGettigan, 2005; Rodd, Cutrin, Kirsch, Millar, & Davis, 2013). In this paper, we consider the potential functional contributions of two distinct groups of cortical brain regions – the domain-selective language network and domain-general Multiple Demand (MD) network - to successful perception and comprehension of different types of challenging speech, and to subsequent perceptual and semantic adaptation.

Role of language-selective versus domain-general (Multiple Demand) regions in language comprehension

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

The language-selective network is a set of left-lateralised frontal and temporal regions that respond reliably to linguistic stimuli with different input modalities, languages and tasks (e.g., Binder et al., 1997; Blank, Kanwisher, & Fedorenko, 2014; Fedorenko, Duncan, & Kanwisher, 2012; Fedorenko, Hsieh, Nieto-Castanon, Whitfield-Gabrieli, & Kanwisher, 2010; MacSweeney et al., 2002; Mahowald & Fedorenko, 2016; Mineroff, Blank, Mahowald, & Fedorenko, 2018; Paunov, Blank, & Fedorenko, 2019; Scott, Galle, & Fedorenko, 2017; for a review, see Fedorenko, 2014), but not to non-linguistic stimuli such as music, mathematical expressions, or computer code (Fedorenko, Behr, & Kanwisher, 2011; Ivanova et al., 2020; Monti, Parsons, & Osherson, 2012). These regions are functionally connected (Braga, DiNicola, Becker, & Buckner, 2020), and show correlated response profiles at rest and during naturalistic listening (Blank et al., 2014; Braga et al., 2020; Mineroff et al., 2018; Paunov et al., 2019), leading to their characterisation as a functionally coherent network. Lesion studies show that damage to, or degeneration of, this network leads to impairments in language function (E. Bates et al., 2003; Mesulam et al., 2014; Mirman et al., 2015; Mirman & Thye, 2018; Turken & Dronkers, 2011), but does not cause deficits in other cognitive domains (Apperly, Samson, Carroll, Hussain, & Humphreys, 2006; Fedorenko & Varley, 2016; Ivanova et al., 2021; Polk & Kertesz, 1993; Varley, Klessinger, Romanowski, & Siegal, 2005; Varley & Siegal, 2000; Varley, Siegal, & Want, 2001) indicating a necessary and selective role of the network in language comprehension.

Sometimes, linguistic stimuli also activate a set of bilateral frontal, parietal, cingular and opercular regions (for a large scale fMRI investigation and relevant discussion, see Diachek, Blank, Siegelman, Affourtit, & Fedorenko, 2020), which together form the Multiple Demand (MD) network (Duncan, 2010b, 2013). This network is domain-general, responding during diverse demanding tasks (Duncan & Owen, 2000; Fedorenko et al., 2012; Fedorenko, Duncan, & Kanwisher, 2013; Hugdahl, Raichle, Mitra, & Specht, 2015; Shashidhara,

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

Mitchell, Erez, & Duncan, 2019) and has been linked to cognitive constructs such as executive control, working memory, selective attention, and fluid intelligence (Assem, Blank, Mineroff, Ademoglu, & Fedorenko, 2020; Cole & Schneider, 2007; Duncan & Owen, 2000; Vincent, Kahn, Snyder, Raichle, & Buckner, 2008; Woolgar, Duncan, Manes, & Fedorenko, 2018). Regions of the MD network show strongly synchronized activity, and fluctuation patterns that dissociate sharply from those of the language network (Blank et al., 2014; Mineroff et al., 2018; Paunov et al., 2019). Moreover, damage to the MD network leads to patterns of cognitive impairment that differ from those observed in cases of language network damage (Duncan, 2010a; Fedorenko & Varley, 2016; Woolgar et al., 2018; Woolgar et al., 2010), confirming a functional dissociation between the two networks (see Fedorenko & Blank, 2020 for a review focusing on the dissociation between subregions of Broca's area).

Recently, it has been argued that the MD network does not play a functional role in language comprehension (Blank & Fedorenko, 2017; Diachek et al., 2020; Shain, Blank, van Schijndel, Schuler, & Fedorenko, 2020; Wehbe et al., 2021; for reviews, see Campbell & Tyler, 2018; Fedorenko & Shain, 2021). That is, activation of MD regions does not reflect core cognitive operations that are essential to language comprehension such as perceiving word forms and accessing word meanings. Instead, it is proposed that activation of MD regions reflects a general increase in effort, which is imposed by task demands in particular, or in some cases even mis-localisation of language-selective activity because of the proximity of the two systems in some parts of the brain (e.g., in the inferior frontal gyrus (IFG), Fedorenko et al., 2012; see Quillen, Yen, & Wilson, 2021 for evidence that increased linguistic and non-linguistic task demands matched on difficulty, differentially recruit language-selective versus domain-general regions).

However, existing evidence that domain-general MD regions do not contribute to language comprehension is limited in two ways. First, relevant studies have typically drawn

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

conclusions about function based on the magnitude of neural activity (e.g., the BOLD fMRI response). The strongest *causal* inference about the necessity (and selectivity) of brain regions for particular cognitive processes comes from approaches that transiently disrupt neural functioning in the healthy brain (e.g., TMS) and measure the effects on behaviour, or from cases of acquired brain damage, either in case studies or multi-patient lesion-symptom mapping investigations that exploit inter-individual variability in behavioural and neural profiles to link specific brain systems to behavioural outcomes (Halai, Woollams, & Lambon Ralph, 2017). A recent lesion study found that the extent of damage to the MD network predicted deficits in fluid intelligence; in contrast, MD lesions did not predict remaining deficits in verbal fluency after the influence of fluid intelligence was removed, which instead were predicted by damage to the language-selective network (Woolgar et al., 2018), in line with the dissociation discussed above. These results provide convincing evidence that the MD network but not the language network contributes to fluid intelligence, but suggested that the MD network contribution did not extend to language function. However, given that language function was assessed with a verbal fluency task—an elicited production paradigm that relies on a host of diverse cognitive operations—the question of whether the MD network causally contributes to specific aspects of language comprehension remains unanswered.

A second limitation of previous studies is their focus on the comprehension of clearly perceptible and relatively unambiguous language, whereas naturalistic speech comprehension typically involves dealing with noise and uncertainty in the input. For example, speech may be in an unfamiliar accent, or contain disfluencies and mispronunciations; there may be background speech, other sounds or distractions; and the words and syntax may be ambiguous or uncommon. These features can make identifying words and inferring meaning – core computations of comprehension – more difficult (for a review of different types of

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

challenges to speech comprehension, see Johnsrude & Rodd, 2015). It therefore remains a possibility that the MD network is functionally critical for successful comprehension in these more challenging listening situations (Diachek et al., 2020).

Challenges to speech perception, comprehension and adaptation

Here, we focus on two challenges, which arise from acoustic degradation and from lexical-semantic ambiguity. Acoustically degraded speech makes word recognition less accurate (Mattys, Davis, Bradlow, & Scott, 2012), reduces perceived clarity (Sohoglu, Peelle, Carlyon, & Davis, 2014), increases listening effort (Pichora-Fuller et al., 2016; Wild et al., 2012), and encourages listeners to utilise informative semantic contextual cues (Davis, Ford, Kherif, & Johnsrude, 2011; Miller, Heise, & Lichten, 1951; Obleser, Wise, Dresner, & Scott, 2007; Rysop, Schmitt, Obleser, & Hartwigsen, 2021) and other forms of prior knowledge (Miller & Isard, 1963; Sohoglu et al., 2014; Sumbly & Pollack, 1954). Acoustically degraded speech engages brain regions that plausibly fall within the MD network including parts of the premotor, motor, opercular and insular cortex (Davis & Johnsrude, 2003; Du, Buchsbaum, Grady, & Alain, 2014, 2016; Erb, Henry, Eisner, & Obleser, 2013; Evans & Davis, 2015; Hardy et al., 2018; Hervais-Adelman, Carlyon, Johnsrude, & Davis, 2012; Vaden, Kuchinsky, Ahlstrom, Dubno, & Eckert, 2015; Vaden et al., 2013; Wild et al., 2012), as well in the angular gyrus (Rysop et al., 2021) and inferior frontal gyrus (Davis et al., 2011; Davis & Johnsrude, 2003). Furthermore, disruption of premotor regions either by TMS (D'Ausilio et al., 2009; D'Ausilio, Bufalari, Salmas, & Fadiga, 2012; Meister, Wilson, Deblieck, Wu, & Iacoboni, 2007) or following lesions (Moineau, Dronkers, & Bates, 2005; for a review, see Pulvermuller & Fadiga, 2010) has been shown to impair perception of degraded speech. However, given that the MD network was not explicitly defined in previous studies, the

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

functional contribution of the MD network to acoustically degraded speech perception remains untested.

Lexical-semantic ambiguity (for a review, see Rodd, 2018) challenges comprehension because of the competition between alternative meanings of a single word form during meaning access (Rayner & Duffy, 1986; Rodd, Gaskell, & Marslen-Wilson, 2002; Seidenberg, Tanenhaus, Leiman, & Bienkowski, 1982; Swinney, 1979), and because costly reinterpretation is sometimes required (Blott, Rodd, Ferreira, & Warren, 2021; Duffy, Morris, & Rayner, 1988; Rodd, Johnsrude, & Davis, 2010, 2012). Domain-general cognitive operations may be useful in responding to the challenge, as evidenced by the positive relationship between an individual's success in semantic ambiguity resolution and executive functioning skill (Gernsbacher & Faust, 1991; Gernsbacher, Varner, & Faust, 1990; Khanna & Boland, 2010) and dual task studies showing that performance on non-linguistic visual tasks is impaired during semantic reinterpretation (Rodd et al., 2010), but these domain-general operations may be plausibly generated by either language-selective or domain-general cortical regions.

Functional imaging studies show that semantic ambiguity resolution engages left-lateralised frontal and temporal brain regions typical of the language-selective network, specifically posterior parts of middle and inferior temporal lobe, anterior temporal lobe, and the posterior IFG (Bilenko, Grindrod, Myers, & Blumstein, 2009; Musz & Thompson-Schill, 2017; Rodd, Davis, & Johnsrude, 2005; Vitello, Warren, Devlin, & Rodd, 2014; Zempleni, Renken, Hoeks, Hoogduin, & Stowe, 2007; for a review, see Rodd, 2020). The possibility that the IFG in particular plays a causal role is supported by the observation that individuals with Broca's aphasia have difficulties in using context to access subordinate word meanings (Hagoort, 1993; Swaab, Brown, & Hagoort, 1997; Swinney, Zurif, & Nicol, 1989), although

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

patients in these studies were selected based on their language profile rather than lesion location.

Although subregions within the IFG form part of the language-selective network, as discussed above, there are also subregions that fall within the domain-general MD network (e.g., Fedorenko & Blank, 2020). Indeed IFG recruitment during ambiguity resolution has been typically accounted for by invoking domain-general constructs of cognitive control or conflict resolution (Novick, Trueswell, & Thompson-Schill, 2005; Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997) which resolve competition between alternative meanings of ambiguous words (Musz & Thompson-Schill, 2017). Currently, the heterogeneity of the IFG makes activations within this region difficult to interpret functionally, without careful anatomical identification of relevant components (Tahmasebi et al., 2012).

A range of studies show that listeners can adapt to the challenges of perceiving and comprehending acoustically degraded or semantically ambiguous sentences. Listeners' perception of degraded speech improves spontaneously over time with repeated exposure (Davis et al., 2005; Guediche, Blumstein, Fiez, & Holt, 2014; Hervais-Adelman, Davis, Johnsrude, & Carlyon, 2008; Loebach & Pisoni, 2008; Sohoglu & Davis, 2016; Stacey & Summerfield, 2008), so long as attention is directed to speech (Huyck & Johnsrude, 2012). This perceptual adaptation is facilitated by visual/auditory feedback presented concurrently or in advance (Wild et al., 2012), generalises across talkers (Huyck, Smith, Hawkins, & Johnsrude, 2017), is supported by lexical-level information such that learning through exposure to pseudo-words is less effective than with real words (although in some cases, learning with pseudowords is possible, Hervais-Adelman et al., 2008), but does not additionally benefit from sentence-level semantic information (learning was as effective with meaningless syntactic prose; Davis et al., 2005).

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

Regarding adaptation to ambiguous words, research has shown that accessing a less frequent (subordinate) meaning of an ambiguous word is easier following exposure to the same meaning of an ambiguous word in a prime sentence. Although the cognitive operations underpinning this so-called word meaning priming effect remain somewhat underspecified, the effect can be described as a form of longer-term lexico-semantic learning since it can be observed tens of minutes or even hours after initial exposure, or perhaps longer if adaptation is consolidated by sleep (Betts, Gilbert, Cai, Okedara, & Rodd, 2018; Gaskell, Cairney, & Rodd, 2019; Rodd et al., 2013).

The current study

In the current study, we ask whether speech perception and comprehension in different challenging circumstances, and adaptation in response to these challenges, depend on the MD network or the language-selective network. To do this, we investigated the impact of lesions to these networks, on behavioural measures of speech perception, comprehension and adaptation. Stimuli, data and analysis code are available at <https://osf.io/fm67z/>.

We recruited participants (n=19) on the basis of having long-standing lesions that either (1) had substantial overlap with the domain-selective language network, (2) had substantial overlap with the domain-general Multiple Demand (MD) network or (3) had overlap with neither language nor MD network. The participants performed behavioural tasks to assess the immediate effects and longer-term consequences of two types of listening challenge. For the first challenge (acoustic-phonetic), we measured perception of noise-vocoded spoken sentences in a word report task. Adaptation to this type of acoustic degradation was assessed in subsequent word report task following a period of training. For the second challenge (lexico-semantic), we measured comprehension of spoken sentences that included low-frequency meaning of semantically ambiguous words, in a sentence

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

coherence judgement task. Adaptation to semantic ambiguity was assessed in a word association task to measure the consequences of experience with the lower-frequency meanings for subsequent meaning access. Whilst all cognitive tasks will require the contribution of some general cognitive operations (e.g. attention, working memory), our tasks were chosen to be simple enough for participants to perform thus minimising the demands on such general cognitive operations, in the absence of acoustic-degradation or lexico-semantic ambiguity. These tasks are made more difficult by challenges to perceptual processes (e.g. acoustically-degraded speech), or semantic processes (e.g. lexico-semantic ambiguity) that are a central part of language function. We acknowledge that perceptual and semantic challenges to language function may have secondary impacts on domain-general functions (e.g. due to increased working memory demand, or a requirement that listeners use sentence context to support processing). However, we expect the same sorts of additional domain-general operations to apply both to degraded speech perception and semantic ambiguity comprehension, as well as to adaptation to these challenges. Thus, if brain lesions have a dissociable impact on accommodating these different challenges, then this would suggest a causal contribution to a specific aspect of language functioning rather than a contribution to domain general processes.

Planned analyses focused on comparing behavioural performance measures across the three participant groups. However, the aetiologies (stroke, tumour excision etc.) that lead to brain lesions do not respect functional boundaries of the two networks of interest, and therefore our primary analyses treated lesion volume in each network as a graded rather than a categorical factor. We report group identities of the participants in the demographics table (Table 1), lesion maps (Figure 1) and in the data plots since this was the basis of our participant recruitment and so that interested readers can observe how individual participants

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

– defined on the basis of lesion location – perform our various tasks. Group analyses are included in the Supplementary Information.

For each task, behavioural performance measures were associated with lesion location and extent by performing correlational analyses using probabilistic functional activation atlases (e.g., Woolgar et al., 2018). Finally, across-task analyses assessed potential dissociations between the contributions of these two networks for accommodating and adapting to different sources of listening challenge during speech comprehension.

Methods

Participants

Twenty-one right-handed native English speakers were recruited from the Cambridge Cognitive Neuroscience Research Panel (CCNRP), a database of volunteers who have suffered a brain lesion and have expressed interest in taking part in research. Participants were invited to take part in the current research on the basis that they had chronic lesions (minimum time since injury of 3 years) to cortical areas falling predominantly in the language or Multiple Demand (MD) networks (or lesions in other areas for control participants), but without knowledge of their behavioural profiles. Thus, volunteers were not recruited on the basis of a known language impairment or aphasia diagnosis. The two networks were broadly defined based on previous functional imaging data from typical volunteers (described below), and linked to lesions traced on anatomical MRI scans for CCNRP volunteers. Participants gave written informed consent under the approval of the Cambridge Psychology Research Ethics Committee. Data from two participants were not included in the final analyses of either task (one participant was unable to complete either task due to fatigue and hearing difficulties; a second failed to complete the semantic ambiguity experimental tasks and also had difficulties accurately reporting back the words they heard in the degraded speech task,

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

achieving only 68% word report accuracy for the clear speech condition across pre- and post-training test sessions, see task details below).

The remaining 19 participants (8 female, mean age 61 years, range 37 – 75 years) had brain lesions caused by tumour excision (n=8), stroke (haemorrhagic: n=6, ischaemic: n=1), or a combination of these (tumour excision and haemorrhagic stroke: n=1) with other causes being abscess excision (n=1) or resection because of epileptic seizures (n=1), and one of unknown cause (n=1). Individual participant characteristics are detailed in Table 1. Two participants contributed data to just one of the tasks (md6 was excluded from the degraded speech experiment for not completing the task; md10 was excluded from the semantic ambiguity analyses for giving multiple responses during the word association task). Thus, data from 18 participants were included for each of the experiments analysed separately (see below) and from 17 participants for the cross-experimental analyses.

The National Adult Reading Test (NART) (Nelson, 1982) was used to estimate premorbid IQ. The Test of Reception of Grammar (TROG-2, Bishop, 2003) was used as a background assessment of linguistic competence.

Table 1

Participant profiles including demographics, lesion details, IQ based on the NART, and TROG2 scores. Participant md6 was excluded from the degraded speech task analyses; participant md10 was excluded from the semantic ambiguity task analyses.

Participant	Group	Age at test	Sex	Lesion aetiology	Lesion location	Premorbid IQ (from NART)	TROG2 Score	Total Lesion Volume (cm ³)
lang1	LANG	50	F	Tumour	L posterior temporal, posterior corpus	108	4	9.49

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

					callosum, posterior			
					cingulum			
lang2	LANG	75	F	Stroke (ischaemic)	L frontal and anterior	86	8	4.06
					insular cortex			
lang3	LANG	58	F	Stroke (haemorrhagic)	R fronto-temporo- parietal and anterior	123	8	39.70
					thalamus			
lang4	LANG	65	F	Tumour	L temporal	117	6	1.90
lang5	LANG	50	M	Other (abscess removal)	L temporal	101	4	9.71
lang6	LANG	56	M	Tumour	R inferior parietal/temporal	112	2	2.38
md1	MD	58	M	Stroke (ischaemic)	L occipito-parietal	97	5	8.21
md2	MD	63	F	Tumour	L superior parietal lobe	123	7	7.06
md3	MD	64	M	Other (unknown)	R fronto-temporo- parietal surrounding insular cortex. Anterior branch of internal capsule	115	5	8.59
md4	MD	53	F	Tumour	R superior frontal	106	4	14.68
md5	MD	69	M	Stroke (haemorrhagic)	L frontal	101	4	13.15
md6*	MD	61	M	Tumour + Stroke (haemorrhagic)	R posterior frontal and some anterior/medial, extending to post- central gyrus/parietal areas	123	3	31.27
md7	MD	52	F	Tumour	Bifrontal	112	6	9.02
md8	MD	74	M	Tumour	R frontal	120	6	7.68
md9	MD	67	M	Tumour	L anterior frontal	126	8	4.05
md10*	MD	81	M	Stroke (haemorrhagic)	L temporo-occipital	81	4	3.26
other1	OTHER	58	M	Other (resection for epilepsy)	R temporal	NA	6	7.84

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

other2	OTHER	60	M	Stroke (haemorrhagic)	R basal ganglia (putamen + caudate + thalamus) and internal capsule dorsal anterior insula	97	7	5.37
other3	OTHER	37	F	Stroke (haemorrhagic)	L frontal	105	5	2.52

Lesion analysis

Lesion analysis followed procedures developed in previous research (Woolgar et al., 2018). Each participant had a structural MRI image (T1-weighted Spoiled Gradient Echo (SPGR) MRI scans with 1x1x1mm resolution) which included lesion tracing as part of previous participation in the CCNRP. From these images, we estimated the volume of lesion that overlapped with the language network, the Multiple Demand network or elsewhere. The two networks were defined from probabilistic fMRI activation maps constructed from large numbers of healthy participants (Language: n=220, MD: n=63) who performed tasks developed to localise language processing and domain-general executive processing (see Blank et al., 2014; Fedorenko, 2014; Fedorenko et al., 2013; Mahowald & Fedorenko, 2016). The activation maps for the language network contrasted data from participants reading or listening to sentences versus lists of pseudowords (neural responses in the language network are modality-independent; Fedorenko et al., 2010; Scott et al., 2017); those for the Multiple Demand network contrasted data from participants performing a hard versus easy visuo-spatial working memory task (remembering 8 vs. 4 locations, respectively, in a 3 X 4 grid). The visuo-spatial task captures all major components of the MD network defined by overlap of multiple demands (Assem, Glasser, Van Essen, & Duncan, 2020). Furthermore, defining the network with a non-auditory, non-language task, makes relating the impact of damage to the network on spoken language functions potentially more noteworthy (than using a task that targets auditory or language processing). Each individual participant's activation map for the relevant contrast (sentences > pseudowords, hard > easy spatial working memory) was thresholded at a $p < .001$ uncorrected level, binarised and normalised before the resulting images were combined in template space. Thus, the

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

language and MD networks are functionally defined for each individual separately before being combined (for discussion of the benefits of using an individual-subject approach see Fedorenko, 2021). The resulting probabilistic activation overlap maps (Figure 1A) contain information in each voxel about the proportion of participants who show a significant effect (at $p < .001$) for the contrast of interest. Following Woolgar et al. (2018), we thresholded the probabilistic map for each network at 5%, thus retaining voxels in which activation was present for at least 5% of the contributing participants.

We then calculated the lesion volume falling into each network (defined in the probabilistic map) for each of the 19 participants (Figure 1B). Participants were initially assigned to one of three broad groups (LANG, MD or OTHER) based on the proportion and volume of their lesions falling into language and Multiple Demand regions as well as the overall proportion of each network that was damaged (Figure 1C; see Supplementary Information for further details of group assignment). However, since assignment of participants to groups is based on arbitrary lesion volume thresholds and because the group allocation for several participants was not clear-cut (e.g. lang2, lang3, md5 in Figure 1B) our main analyses correlate behavioural performance measures with lesion volume in the two key networks, thereby avoiding these arbitrary choices. We detail the group assignments in describing the participants and results so that the interested reader can track information about individual participants. Group analyses are included in the Supplementary Information.

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

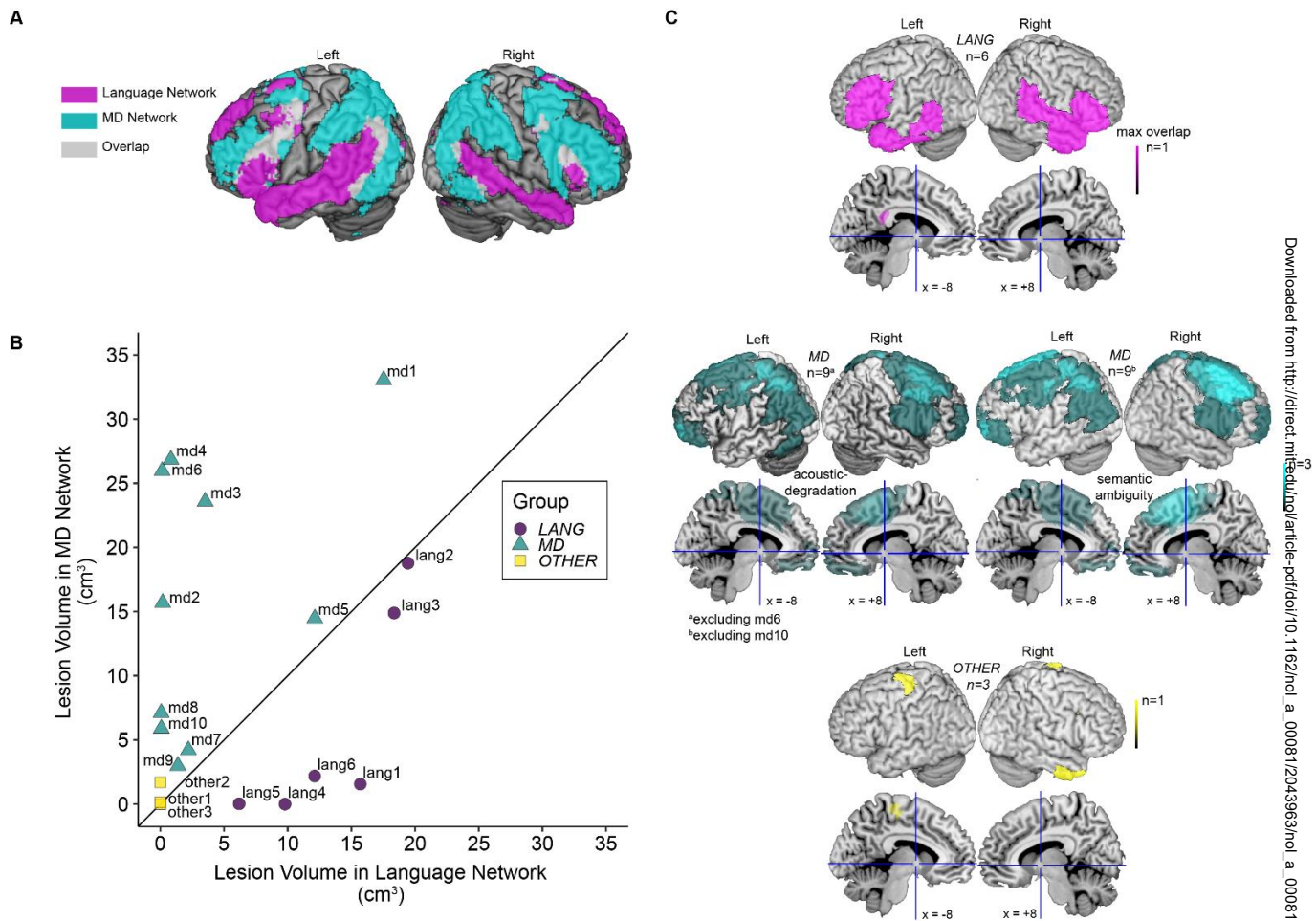


Figure 1

A. The language and MD networks against which we compared participants' lesions. The images show probabilistic activation maps of the language network and the MD network based on fMRI data from large numbers of neurotypical participants (language: $n=220$; MD: $n=63$), which have been thresholded to show regions active in at least 5% of participants during the relevant functional task and plotted onto a volume rendering of the brain. **B.** Volume of lesion falling into each network for each of the 19 participants in the present study. Solid line depicts an equal volume of each network affected by the lesion. Different colours/shapes indicate assignment of the participants into the LANGUAGE (LANG), MD and OTHER Groups upon which recruitment was based (for categorical group analyses, see

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

Supplementary Information). **C. Lesion overlap across participants depicted on volume renderings of the brain and on midline sagittal slices viewed as if from the left or right (Montreal Neurological Institute (MNI) space x coordinates of -8 and $+8$, cross-hairs show the location of $y = 0$, $z = 0$ in these slices). Images are shown separately for participants originally assigned to each of the three groups (see *Supplementary Information* for group analyses). Two participants assigned to the MD group (*md6*, *md10*) contributed data to tasks for only one type of challenge and therefore images are shown separately for the two challenge types. Brighter colours reflect greater lesion overlap across participants.**

Statistical analysis

Analyses were performed using R statistical software (version 3.6.1, R Core R Development Core Team, 2019). For each task, the primary analyses assessed whether more extensive damage to the language and MD networks were associated with more impaired performance on the behavioural tasks, with one-tailed Pearson's r correlation coefficients. We compared the strength of different correlations within task (i.e. comparing the impact of damage to language and MD networks on a given behavioural measure) and between task (i.e. comparing the impact of damage to a given network on different behavioural measures) with two-tailed Meng's z tests (Meng, Rosenthal, & Rubin, 1992) using the 'cocor' package (Diedenhofen & Musch, 2015). The between-task comparisons focused on the 17 participants for whom we had data for both the degraded speech and the lexical-semantic ambiguity tasks.

Challenge 1. Acoustically-degraded speech perception and adaptation

The first challenge increased speech comprehension difficulty at the acoustic-phonetic level of the input, by acoustic degradation of spoken sentences with noise vocoding (Shannon, Zeng, Kamath, Wyganski, & Ekelid, 1995). Noise vocoding reduces the spectral detail in the speech signal but retains the slow amplitude modulations which approximately

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

reflect syllabic units, and the broad-band spectral changes that convey speech content. These low frequency modulations and broadband spectral modulations have been shown to be most important for accurate speech perception (Elliott & Theunissen, 2009; Shannon et al., 1995). We selected the particular numbers of channels in the vocoder based on previous research, which established that intelligibility (as measured by word report: how many words of the sentence a participant can accurately report) increases with the logarithmic increase in number of channels (McGettigan, Rosen, & Scott, 2014). In healthy adults with good hearing, for short sentences of 6-13 words long, intelligibility is near 100% for 16-channel vocoded speech, near 0% for 1-channel vocoded speech, and at an intermediate level for 4-channel vocoded speech (Peelle, Gross, & Davis, 2013). We assessed speech perception in terms of the logarithmic number of channels estimated as required to achieve 50% word report accuracy of these sentences and assessed adaptation by comparing performance before and after a training period.

Stimuli

The stimuli for the degraded speech task were forty declarative sentences, varying in length (6-13 words, $M=9$, $SD = 2.45$) and duration (1.14 to 3.79 seconds, $M= 2.12$, $SD = 0.60$), which were selected from coherent low ambiguity sentences used in previous studies (Davis et al., 2011).

Sentences were recorded by a female native speaker of British English and digitised at a sampling rate of 22050Hz. We created three degraded versions of the sentences, of decreasing intelligibility, using 16-, 8- and 4- channels in the vocoder. To do this, the frequency range 50-8000Hz was divided into 16, 8, or 4 logarithmically spaced frequency channels. Each channel was low pass filtered at 30Hz and half-wave rectified to produce an amplitude envelope for each channel, which was then applied to white noise that was filtered

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

in the same frequency band. Finally, the channels were recombined to create the noise-vocoded version of the sentence.

The 40 sentences were grouped into 8 sets of 5 sentences such that each set contained 45 words in total and were expected (based on previous word report data) to be approximately equally intelligible. Each participant heard all 8 sentence sets, but assignment of sets to the different levels of degradation (clear, 16-, 8-, 4-channel vocoded) and to the pre- and post-training test (described below) was counterbalanced across participants.

Procedure

The experiment started with 4 practice trials to familiarise the participants with the stimuli and the word report task. Participants listened to 4 different sentences (not included in the test set) at increasing levels of degradation (clear, 16, 8, 4) and after each sentence had to repeat the sentence or as many words from the sentence as possible in the correct order. The experiment then followed a test-train-test format (c.f. Sohoglu & Davis, 2016) with the 40 experimental sentences (8 sets of 5 sentences; see stimuli for details of assignment of the sentences to the pre-test and post-test and to the different levels of degradation). In the initial test, participants listened to 20 of the sentences, 5 at each level of degradation (clear, 16, 8, 4; order randomised uniquely for each participant) and performed the word report task. This was followed by a training period in which participants listened passively to the same 20 sentences, each repeated 4 times at decreasing levels of degradation, whilst the written text of the sentence was presented visually on a computer screen. Following the training, participants listened to the other (previously unheard) 20 sentences, 5 at each level of degradation and again performed the word report task.

Data processing and analysis

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

For each participant, we calculated the number and proportion of words correctly reported for each sentence at each level of degradation (clear, 16-, 8-, 4-channel vocoded) and for the pre- and post-training test. Words were scored correct only if there was a perfect match with the spoken word from the sentence (morphological variants were scored as incorrect, but homonyms, even if semantically anomalous, were scored correct). Words reported in the correct order were scored correct even if intervening words were absent or incorrectly reported, but scored as incorrect if they were reported in the wrong order. To verify that decreasing the number of vocoded channels increased the challenge of speech perception and that training facilitated perception, we analysed differences in proportion of words correctly reported between the sentences with different numbers of channels (clear, 16, 8, 4) and pre- and post-test sentences, with a logistic mixed effects model (GLMM) using the ‘lme4’ package (D. Bates, Machler, Bolker, & Walker, 2015) . The model had a single categorical fixed effects predictor for Training (pre-test or post-test) with deviation coding defining one planned contrast: pre-test = -1/2 versus post-test = 1/2. There was also a continuous fixed effect predictor of Log2Channels. The final model contained a by-subject random intercept, by-subject slopes with Training and Log2Channel and a by-item random intercept for sentence.

To quantify the relationship between acoustic degradation and speech perception performance in single participants we also fit a logistic psychometric function to the word report accuracy data separately for each participant, for averaged data, and for pre- and post-training tests separately using the ‘quickpsy’ package (Linares & Lopez-Moliner, 2006). The parameters of the logistic function were estimated using direct maximisation of the likelihood with the following equation:

$$(f(x : \alpha, \beta, \gamma, \lambda)) = \gamma + \frac{1 - \gamma - \lambda}{1 + e - (\frac{x}{\alpha})\beta}$$

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

During the fitting, we treated clear speech as equivalent to 32-channel vocoded speech and converted the number of channels vocoded at each level of degradation into their log equivalents (χ). From each fit, we obtained alpha (α), the number of channels estimated to give 50% accuracy on the word report task. This value, referred to as “threshold number of channels” was used for the subsequent analyses of the impact of lesion on performance (c.f. McGettigan et al., 2014). Lower alpha values indicate that fewer channels were required to reach this threshold, and thus reflect better performance or more accurate perception. Beta (β) corresponds to the slope or steepness of the curve. Gamma (γ) is the guess rate, which was fixed to 0 for this open set speech task. Lambda (λ) is the lapse rate, or expected proportion of errors as the number of channels reaches the highest levels. Lambda represents the upper horizontal asymptote and was fixed at 1 minus the proportion of correct word report observed for clear speech for each participant separately. This was required as some participants did not achieve 100% word report for clear speech.

Challenge 2. Semantically-ambiguous speech comprehension and adaptation

The second challenge increased speech comprehension difficulty at the lexical-semantic level, by the inclusion of semantically ambiguous words, in sentence contexts that in most cases supported the lower frequency meaning. We assessed speech comprehension in terms of the speed and accuracy of judging the coherence of these sentences, which were interspersed with sentences without ambiguities and anomalous sentences. The coherence judgement task appeared well-suited for assessing competence at semantic ambiguity resolution for several reasons. First, to respond accurately listeners must understand the whole sentence and not just identify one (or more) unusual words. For example, a sentence might initially make sense but then become anomalous only at the end (“*It was a rainy day and the family were thinking to the banana*”) or might initially seem odd but would

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

eventually make sense (“*It was a terrible hand and the gambler was right to sit it out*”). Secondly, since most of the meanings that we used in the sentences were the less frequent meanings, accurate performance relies on listeners utilising contextual cues to select the appropriate meaning rather than the higher-frequency, more accessible meaning. The use of lower-frequency word meanings also maximised our chance to observe word meaning priming effects, as described below. Thirdly, participants make a speeded judgement giving a continuous measure of performance in addition to accuracy.

To assess the increase in availability of low frequency word meanings in response to experience, we measured changes to meaning preferences in a word association task. This task provides a direct measure of how participants interpret ambiguous word forms in the absence of any sentence context. Specifically, using two counterbalanced sentence sets, we measured the increase in proportion of word association responses that were consistent with the (low frequency) meaning used in the sentence context for ambiguous words that had been heard (primed) compared to those that had not (unprimed). Counterbalanced assignment of sentences to primed and unprimed conditions for different participants ensured that differences in meaning frequency or dominance did not confound assessment of the word-meaning priming effect (for further discussion of word-meaning priming, see Rodd et al., 2013).

Stimuli

The stimuli for the coherence judgement task were one hundred and twenty declarative sentences, selected from two previous studies (Davis et al., 2011; Rodd et al., 2005). Of these, 40 were high-ambiguity coherent sentences, 40—low-ambiguity coherent sentences and 40—anomalous sentences. The high-ambiguity sentences each contained 2 ambiguous words that were disambiguated within the sentence (e.g., “*The PITCH of the*

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

NOTE was extremely high"; the ambiguous words were not repeated across the set of 40 sentences). Prior dominance ratings (Gilbert & Rodd, 2022) indicated that in most of the sentences, the context biased the interpretation of the ambiguous words towards their subordinate (less frequent) meanings (mean dominance = 0.31; SD = 0.25). The low-ambiguity sentences were matched with the high-ambiguity sentences across the set for number of words, number of syllables, syntactic structure and naturalness but contained words with minimal ambiguity (e.g. "*The pattern on the rug was quite complex*"). These 80 coherent sentences were separated into two lists (List A and List B), each containing 20 high-ambiguity and 20 low-ambiguity sentences. Participants were presented with sentences from either list (List A or List B) and thus were exposed to half of the ambiguous words in this part of the experiment. Each list also contained all 40 anomalous sentences (i.e. the same sentences were presented to all participants) which had been created from the low-ambiguity sentences by randomly substituting content words matched for syntactic class, frequency of occurrence and numbers of syllables (e.g., "*There were tweezers and novices in her listener heat*"). Thus, the anomalous sentences had identical phonological, lexical and syntactic properties but lacked coherent meaning (see Table 2 for psycholinguistic properties of the 3 sentence types).

Table 2*Descriptive characteristics of the three sentence types.*

Sentence type	N	Mean (range) number of words	Mean (range) number of syllables	Mean (range) duration in seconds	Mean (range) naturalness rating on 9-point scale	Mean (range) imageability rating on 9-point scale
High ambiguity	40	9.6 (6-18)	11.8 (7 – 22)	2.2 (1.1-4.0)	6.1 (3.6-7.9)	5.4 (2.1-9.0)
Low ambiguity	40	9.6 (6-18)	12.2 (8 – 23)	2.2 (1.6-4.3)	6.4 (3.4-7.9)	5.0 (2.0-8.0)
Anomalous	40	9.0 (6-13)	11.6 (6 – 20)	2.3 (1.3 – 3.5)		

The stimuli for the word association task were the 80 ambiguous target words from the 40 high-ambiguity sentences. Given that participants had only heard half of the high-ambiguity sentences in the sentence coherence judgement task (List A or List B), for 40 of the ambiguous words, the subordinate meaning was primed (previously heard in a supportive sentence context) and for the other 40, the subordinate meaning was not primed.

Sentences and single words were recorded individually by a male native speaker of British English (MHD) and sentences were equated for RMS amplitude across conditions.

Procedure

The task consisted of two phases. In the first phase, participants listened to 80 sentences (20 high-ambiguity, 20 low-ambiguity, 40 anomalous) and had to judge as quickly and as accurately as possible the coherence of each sentence. They indicated their response by pressing a green button if the sentence made sense and a red button if it did not.

Participants were given examples (not included in the test set) to encourage them to listen to the sentence in its entirety before making the judgment.

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

Following the coherence judgement task, participants completed other behavioural tasks (not relevant to the current investigation) for 20-30 minutes before moving to the second phase: a word association task. In this phase, participants heard 80 ambiguous words presented in isolation, of which half had been presented in phase 1 (primed) and half were new (unprimed; counterbalanced across participants). For each word, participants had to repeat it and then say the first related word that came to mind. Responses were audio recorded and later coded as consistent with the subordinate meaning (e.g., “NOTE-music”) or inconsistent with the subordinate meaning (“NOTE-write”).

Data processing and analysis

There were 1,440 experimental trials (18 participants X 80 items). We excluded trials with very fast responses (more than 300 ms before the offset of the sentence), which were assumed to arise from accidental key presses or anticipatory responses. This resulted in the exclusion of 2 anomalous sentence trials and 1 low-ambiguity sentence trial.

For each participant, we first assessed whether they could discriminate the coherent sentences (high-ambiguity and low-ambiguity) from the incoherent sentences better than would be expected by chance, by calculating *d*-prime values for the high-ambiguity and low-ambiguity sentences separately:

$$d_{prime} = z(Hits) - z(False Alarms)$$

Hits correspond to the proportion of coherent sentences correctly judged as coherent. False alarms correspond to the proportion of incoherent sentences incorrectly judged as coherent. To allow for calculation of the z-scores, hit rates of 1 were adjusted by $1 - \frac{1}{2N}$ (i.e., to a value of 0.975) and false alarm rates of 0 were adjusted by $\frac{1}{2N}$ (Macmillan & Kaplan, 1985, i.e. to a value of 0.0125).

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

The false alarm rate was necessarily identical for high-ambiguity and low-ambiguity conditions (since we only included a single set of incoherent sentences), differences in accuracy between the high-ambiguity and low-ambiguity sentences can be assessed using error rates when participants judged these coherent sentences to be anomalous. Therefore, for the main accuracy analyses we excluded the 40 anomalous sentence trials, leaving 719 trials (1 trial was excluded based on a fast RT; see above).

The response time analyses focused on ambiguous and unambiguous sentence trials. Of the 720 total experimental trials (18 participants X 40 items), we excluded trials incorrectly judged as incoherent (23 trials: 14 ambiguous, 9 unambiguous). For exclusions of trials based on response times, we followed the general principle of minimal trimming with model criticism (Baayen & Milin, 2010). We excluded trials with very fast response times (less than 300 ms before offset; as for the accuracy analysis), which were assumed to reflect accidental key presses (1 trial) as well as trials with very slow response times (3 trials with responses longer than 4,000 ms after sentence offset) because we were interested in speeded responses. Further exclusions were considered after first determining whether any transformation of the dependent variable was required to meet assumptions of the LME models, of homogeneity of residual variance and normally distributed residuals. Model diagnostic plots (quantile-quantile and histogram plots of the residuals) for the raw, log₁₀-transformed and inverse transformed response time data showed that log₁₀ transformation best met the assumptions. Examination of the plots for outliers indicated that no further trimming was necessary, thus there were 693 correctly judged coherent trials included in the analyses.

We analysed differences in accuracy and response times between the high-ambiguity and low-ambiguity sentence trials with a logistic mixed effects model (GLMM; accuracy) or a linear mixed effects model (LMM; log-10 response times) using the 'lme4' package (D.

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

Bates et al., 2015). The models had a single categorical fixed effects predictor for Sentence Type (High-ambiguity or Low-ambiguity) with deviation coding defining one planned contrast: High-ambiguity = 1/2 versus Low-ambiguity = -1/2. The final models each contained a by-subject and by-item random intercept.

The correlational analyses used the model residuals (comparing predictions to the data) to estimate the ambiguity response time effect (difference between responses for high-ambiguity and low-ambiguity sentence trials) for each participant. A positive residual difference indicates that the participant's ambiguity effect was larger than predicted by the model (response times were slower than estimated for the high-ambiguity condition and/or faster than estimated for the low-ambiguity condition). A negative residual difference means that their response time effect was smaller than predicted by the model (response times were faster than estimated for the high-ambiguity condition and/or slower than estimated for the low-ambiguity condition).

For the word association task, each response was independently coded for consistency with the subordinate meaning used in the priming sentence by two of the authors (LM and ZB), who were blind to the experimental condition (primed/unprimed) of the responses. For example, the word "ball" came from the sentence "*The ball was organised by the pupils to celebrate the end of term*", so responses such as "party" and "dance" were coded as consistent whereas responses such as "kick" and "round" were coded as inconsistent. The consistency scores for the unprimed words give a baseline measure of the preference for the dominant meaning. Response codes from the first author were used with the exception of one participant for whom data were lost and only the codings from the second rater were available; inter-rater reliability for the remainder of the responses was high (94% agreement from 1,360 responses, Cohen's Kappa = 0.862).

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

We analysed differences in the proportions of responses consistent with the subordinate meaning between primed and unprimed words (word meaning priming) with a logistic mixed effects model (GLMM) with a categorical fixed effect predictor for Priming Type (Primed or Unprimed) with deviation coding defining one planned contrast: Primed = 1/2 versus Unprimed = -1/2. There was also a continuous fixed effect predictor of Meaning Dominance (Gilbert & Rodd, 2022) and the associated interactions. The final model contained a by-subject and by-item random intercept and a by-subject random slope for Dominance.

In the main correlational analyses we used the model residuals (comparing predictions to the data) to estimate word priming effects (difference between response values for primed and unprimed words) for each participant. A positive residual difference indicates that the participant's priming effect was larger than predicted by the model (proportion of responses consistent with the subordinate meaning was underestimated for the primed condition and/or overestimated for the unprimed condition). A negative residual difference means that their priming effect was smaller than predicted by the model (proportion of responses consistent with the subordinate meaning was overestimated for the primed condition and/or underestimated for the unprimed condition).

Results

Challenge 1. Acoustically-degraded speech perception and adaptation

Word Report Task

Figure 2A shows the mean proportion of words correctly reported for speech with different numbers of channels, for the pre- and post-training tests. Word report accuracy was near ceiling (100%) for the clear speech reflecting the participants' ability to perform the

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

task, but was close to floor levels for the 4-channel vocoded condition, reflecting the challenge of the acoustic degradation. The mixed effect model confirmed that speech perception accuracy increased as the Log2 number of channels increased (model coefficient: $\beta = 3.199$, $SE = 0.236$, $z = 13.556$, $p < .0001$). Accuracy was greater following training (model coefficient: $\beta = 1.262$, $SE = 0.392$, $z = 3.217$, $p = .001$), showing that participants were able to learn. There was no interaction between the level of degradation and training.

The outputs of fitting the data with a logistic psychometric function are shown in Figure 2B. Analyses to assess the impact of lesions on performance used the threshold number of channels (the estimated number of channels required for 50% word report accuracy) with lower values reflecting better perception (fewer channels needed to reach 50% accuracy). Figure 2C shows the mean performance before and after training for the group and for individual participants.

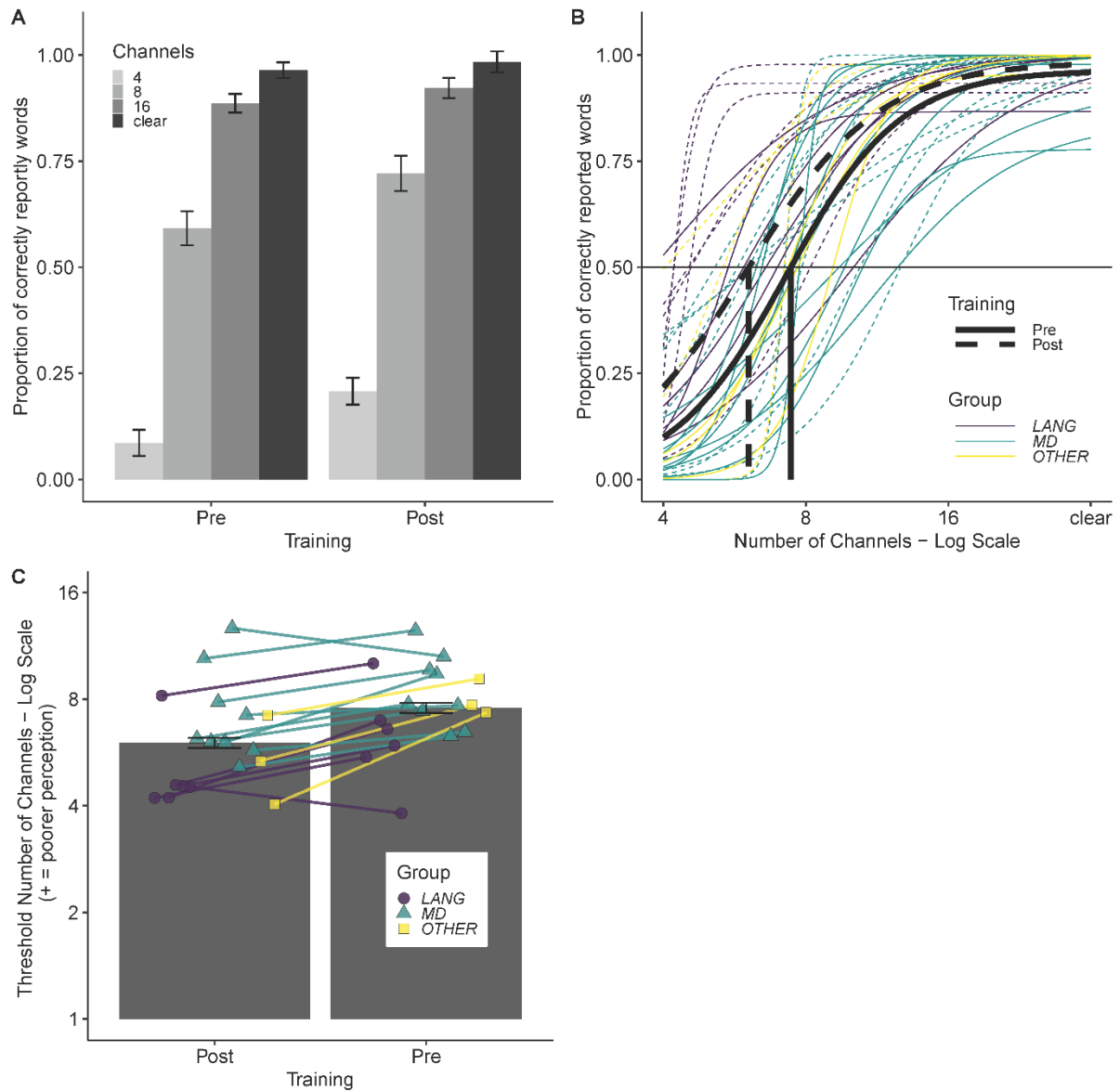


Figure 2

A. Word report accuracy scores for different levels of degradation and the Pre- and Post-Training tests separately. Bars show mean values across all 18 participants and error bars show ± 1 SEM, adjusted to remove between-subject variance (Morey, 2008). **B.** Psychometric logistic function fits separately for the pre- (solid) and post- (dashed) training data for the mean across all 18 participants (black colour) and each participant separately (coloured by

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

Group). The horizontal line indicates the 50% word report accuracy threshold. Vertical lines indicate the estimated threshold number of channels corresponding to the 50% word report accuracy threshold for the mean fits across all 18 participants. C. Estimated threshold number of channels (log scale) required for 50% accuracy in the word report task for the Pre- and Post-Training tests separately. Bars show mean values across all 18 participants and error bars show ± 1 SEM, adjusted to remove between-subject variance (Morey, 2008). Individual participant values are overlaid (colour and shape reflect participant Group; see Supplementary Information).

Figure 3 shows the relationship between degraded speech perception performance and the extent and location of lesions. Correlational analyses showed that the mean threshold number of channels across pre- and post-training tests positively correlated with damage to the MD network ($r = .427, p = .039$) but not with damage to the Language network ($r = -0.152, p = .727$), or with total damage ($r = 0.216, p = .194$). Comparisons of these correlations demonstrated that poorer speech perception was numerically more strongly predicted by damage to the MD network than to the Language network, although this did not reach the $p < .05$ threshold of statistical significance ($z = -1.954, p = .051$). There was no evidence for MD network damage being more predictive of speech perception than total damage ($z = -1.122, p = .262$). There were no correlations between perceptual learning of degraded speech (i.e. change in threshold from pre- to post-training) and the volume of brain damage in the language network ($r = -0.123, p = .727$), MD network ($r = 0.003, p = .504$) or total damage ($r = -0.006, p = .491$). There was also no evidence that MD damage was more predictive of degraded speech perception than of degraded speech adaptation ($z = -1.403, p = .161$).

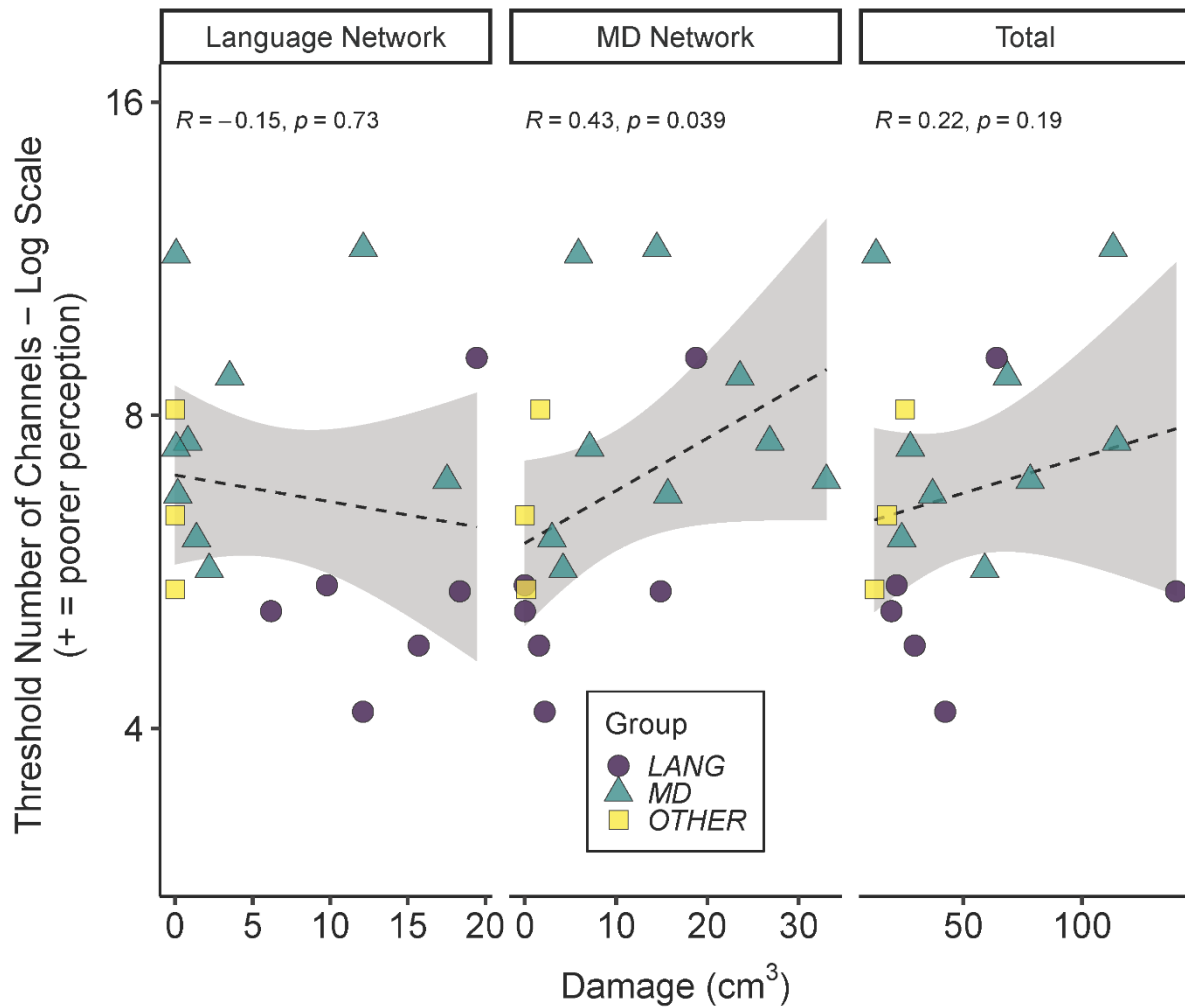


Figure 3

Individual participant data for estimated threshold number of channels (log scale) required for 50% word report accuracy for the mean of Pre- and Post-Training tests plotted separately against damage to the language network, to the MD network and total damage (colour and shape reflect participant Group; see Supplementary Information). Higher threshold number of channels indicates worse speech perception performance. The dashed line shows the linear best fit and grey shaded areas show 95% confidence intervals.

Challenge 2. Semantically-ambiguous speech comprehension and adaptation

Sentence Coherence Judgment Task

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

All participants showed d-prime values substantially above 0 indicating successful discrimination of the incoherent sentences from both the high-ambiguity (mean = 3.66, SD = 0.25, range = 3.0 – 4.20) and the low-ambiguity sentences (mean = 3.75, SD = 0.37, range = 2.72 – 4.20). Figure 4 shows the mean error rates (Figure 4a) and the mean response times (Figure 4B) for the different Sentence Types and for individual participants. Across all participants the proportions of correct responses were near ceiling (mean = 0.97, SD = 0.03, range = 0.92 – 1.0). The mixed effect model showed no effect of Sentence Type on accuracy (model coefficient: $\beta = -0.440$, SE = 0.538, $z = -0.817$, $p = .414$) and hence we have no evidence that sentences containing ambiguous words were less well understood.

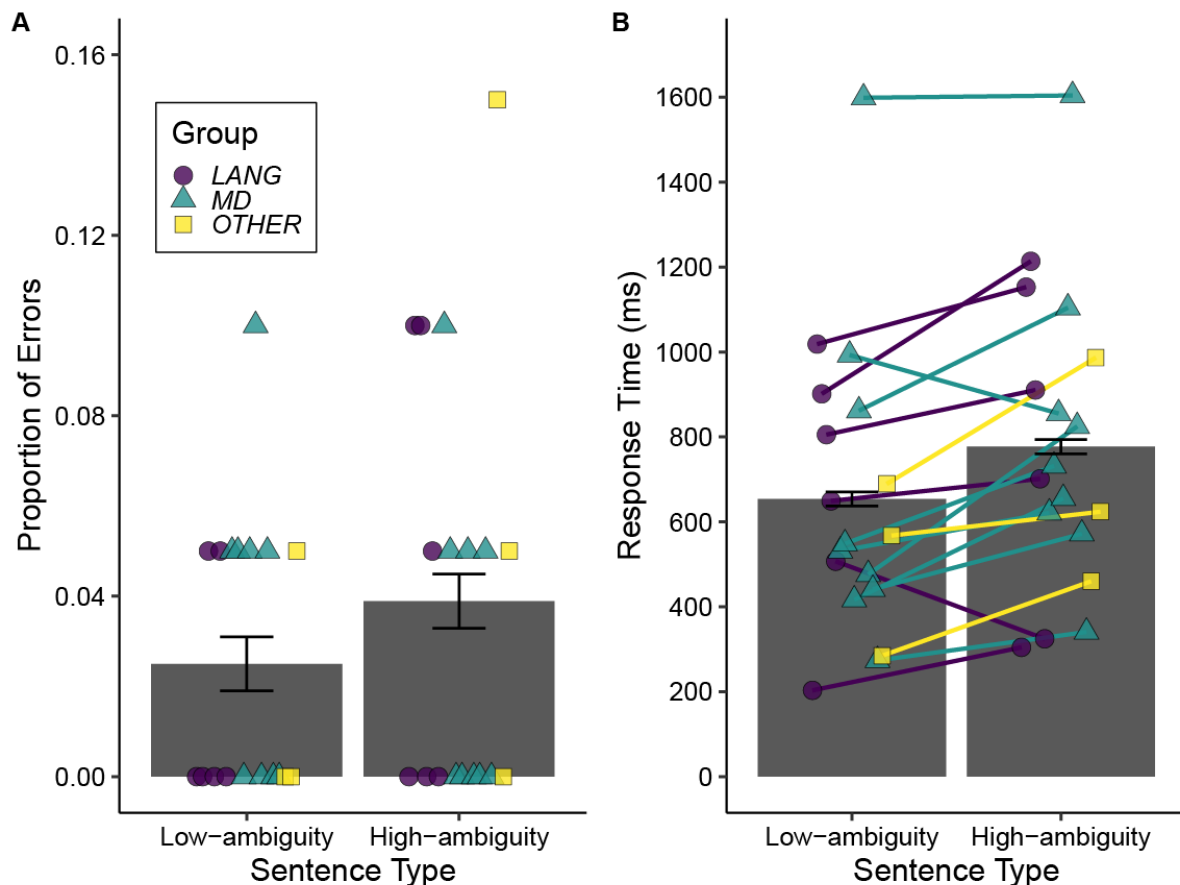


Figure 4

A. Proportion of errors and B. Response times measured from sentence offset for coherence judgements to low-ambiguity and high-ambiguity sentences. In each case, bars reflect the

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

mean values across all 18 participants and error bars show ± 1 SEM, adjusted to remove between-subject variance (Morey, 2008). Individual participant values are plotted (colour and shape reflect participant Group; see Supplementary Information).

The mixed effect model for response times confirmed that high-ambiguity sentences were responded to more slowly than low-ambiguity sentences ($\beta = 0.043$, $SE = 0.019$, $t(2.319) = 2.319$, $p = .023$), showing that ambiguous words increased the challenge of sentence comprehension. However, there was no correlation between individual ambiguity response time effects (model residual difference measure) and extent of damage to the language network ($r = -0.102$, $p = .656$), the MD network ($r = -0.038$, $p = .559$) or overall damage ($r = 0.076$, $p = .383$).

Word Association Task

We excluded primed trials corresponding to words from sentences that were responded to incorrectly in the coherence judgement task. This resulted in exclusion of 28 trials (words from 14 ambiguous sentences: 1.94% of data) across the 18 participants, leaving 1,412 observations. For unprimed words (i.e., for ambiguous words that were not presented to participants in the coherence judgement task), the mean proportion of responses (across items and participants) that were consistent with the subordinate meaning of the word was 0.29 ($SD = 0.09$). This value, which gives a baseline measure of the preference for the dominant meaning, indicates that the sentence-primed meanings were indeed the subordinate or less preferred meanings (note that the value is similar to the one derived from an existing database (Gilbert & Rodd, 2022; see section on Stimuli above). Figure 5 shows the mean proportion of responses consistent with the subordinate meaning for primed and unprimed words.

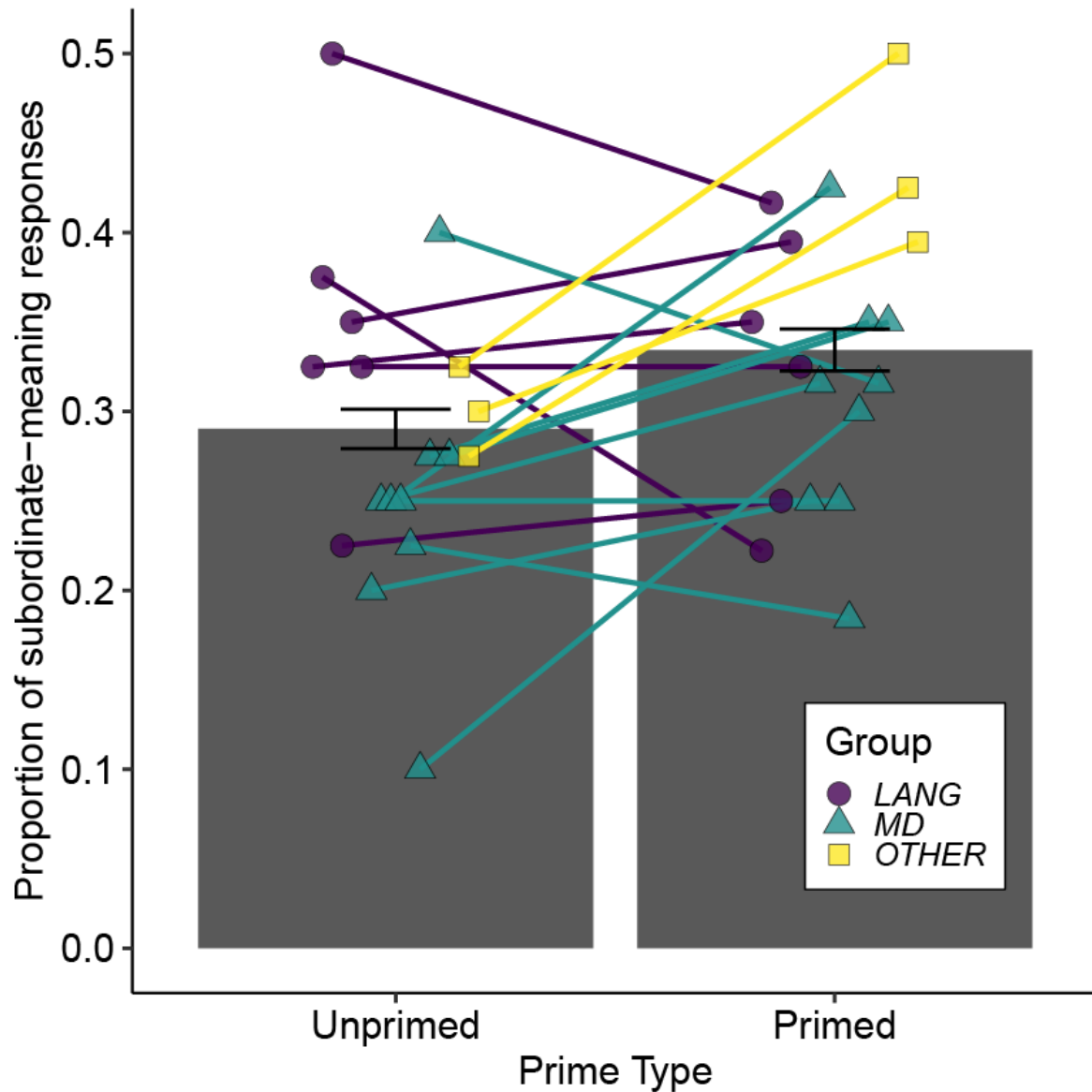


Figure 5 Proportion of responses consistent with the subordinate meaning for ambiguous words in the Unprimed and Primed conditions. Bars reflect the mean values across all 18 participants and error bars show \pm SEM, adjusted to remove between-subject variance. Individual participant values are plotted (colour and shape reflect participant Group; see Supplementary Information).

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

We observed a main effect of Priming ($\beta = 0.352$, $SE = 0.137$, $z = 2.565$, $p = .010$), which reflects a higher proportion of responses consistent with the subordinate meaning for the primed compared to unprimed words. This finding demonstrates a change in word meaning preferences in response to recent experience of sentences containing ambiguous words. We also observed a main effect of meaning Dominance of the word ($\beta = 1.112$, $SE = 0.111$, $z = 10.054$, $p < .0001$), reflecting an increase in proportion of responses consistent with the subordinate meaning as the dominance of that meaning increased (became less subordinate and closer in frequency to the alternative dominant meaning). There was an interaction between Dominance and Priming ($\beta = -0.300$, $SE = 0.171$, $z = -2.124$, $p = .034$), reflecting a stronger Priming effect for meanings that were more subordinate.

Correlational analyses revealed a negative relationship between individual word meaning priming effects (model residual difference measure) and the extent of damage to the language network ($r = -0.659$, $p = .001$) but not the MD network ($r = -0.035$, $p = .446$) or total damage ($r = -0.180$, $p = .237$). Comparisons between these correlations showed that word meaning priming was more strongly predicted by damage to the language network than to the MD network ($z = -2.182$, $p = .0291$); a comparison with the correlation with total damage did not reach the conventional $p < .05$ threshold ($z = -1.863$, $p = .062$). There was also evidence that damage to the language network was more predictive of individual participants' word meaning priming than the ambiguity response time effect ($z = -2.6523$, $p = .008$). Figure 6 shows scatter plots of the correlations between word-meaning priming and damage to language, MD Networks, and total damage. A further correlational analysis showed no relationship between participants' ambiguity response time effect and their word meaning priming effect (both measured using the model residuals: $r = -0.26$, $p = .298$, 2-tailed), suggesting that reduced word meaning priming effect could not be simply explained as due to poorer comprehension.

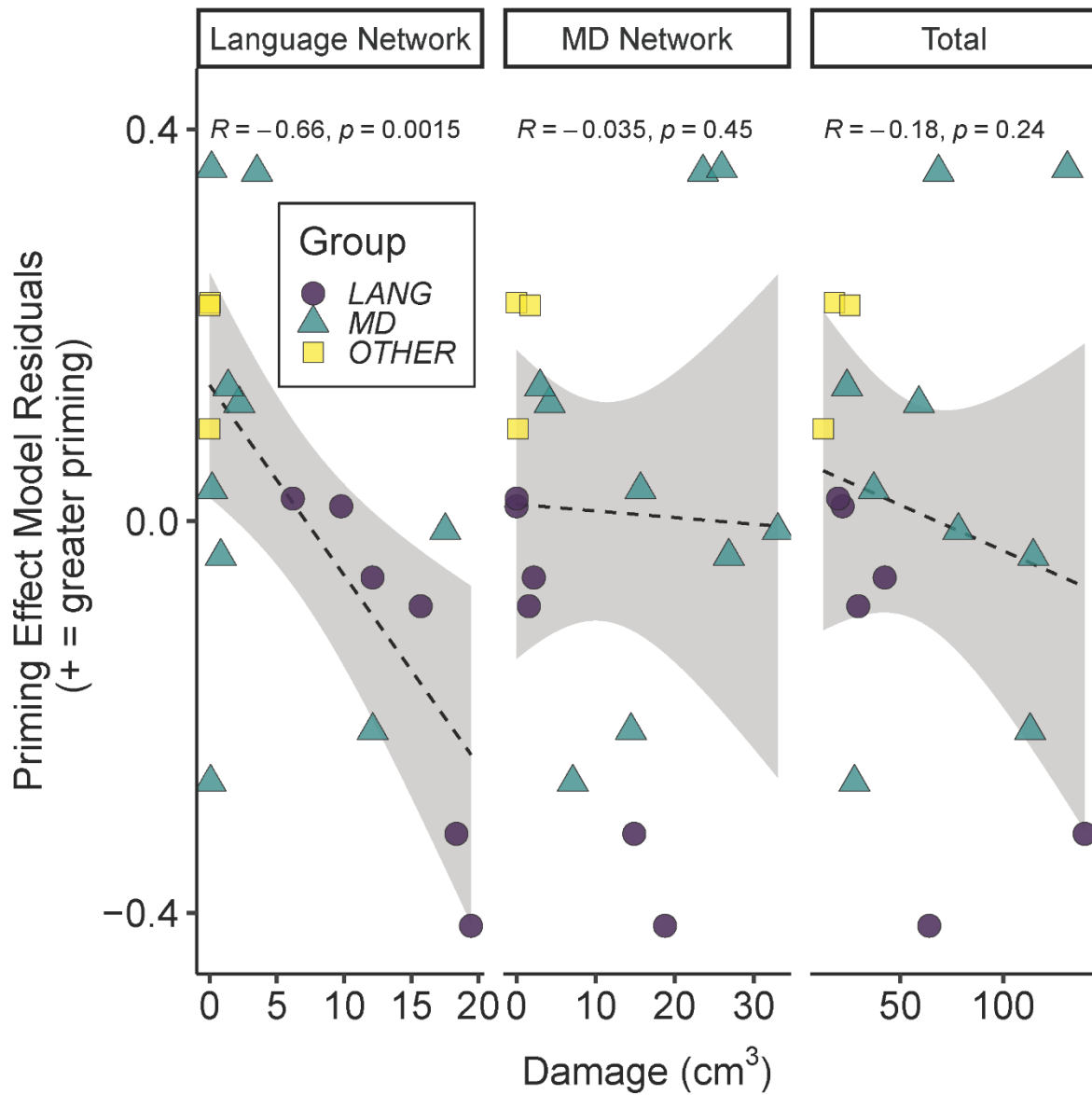


Figure 6

Individual participant data for word meaning priming effects estimated from the model residuals, plotted separately against damage to the language network, to the MD network and total damage (colour and shape reflect participant Group; see Supplementary Information). The dashed line shows the linear best fit and grey shaded areas show 95% confidence intervals.

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

Dissociations between challenges to speech perception and comprehension

Table 3 summarise the correlations between the lesion volume in each network and measures of acoustically-degraded speech perception, semantically-ambiguous speech comprehension and associated adaptation to each of the challenges. Table 3A displays lesion-behaviour correlations from the 18 participants tested for each challenge. Where significant lesion-behaviour correlations are observed, additional comparisons between correlations with behaviour for lesions to the two networks and within each challenge type are also shown. Table 3B displays lesion-behaviour correlations for the 17 participants for whom we have data for both types of challenge. This table shows comparisons between correlations for the two types of challenge.

Table 3

Summary of correlations between the extent of damage to the two networks of interest (MD network and language network) and performance on the tasks measuring speech perception/comprehension and adaptation (labelled (1) – (4)) for the two types of challenge (acoustic degradation and semantic ambiguity). Where significant lesion-behaviour correlations are observed, additional comparisons of the strength of these correlations are also shown (NA indicates that no comparison was performed since neither of the correlations to be compared was different from zero). Correlations are shown for data from the 18 participants for each type of challenge and associated within-challenge (across-lesion, across-task) comparisons. These are the results reported in the sections 4.1 and 4.2. Note that there are two participants (md6, md10) who performed tasks for only one challenge type. Across-challenge comparisons are show for data from the 17 participants for whom we have data for both types of challenge. Significant lesion-behaviour correlations (in bold) are shown between the MD network and Task (1): acoustically-degraded speech perception

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

(turquoise) and between the language network and Task (4): adaptation to semantic ambiguity (word meaning priming; purple). Note that in the case of Task (1) the finding that an increase in lesion volume is associated with worse behavioural outcomes is reflected in a positive correlation since higher perceptual thresholds indicate worse perception. In Task (4) the finding that an increase in lesion volume is associated with worse behavioural outcomes is reflected in a negative correlation since less word meaning priming indicates less adaptation.

CHALLENGE	TASK	LESION		
		MD network	language network	MD vs. language network
Acoustic degradation	(1) Perception	$r = 0.427, p = .039$	$r = -0.152, p = .727$	$z = -1.954, p = .051$
	n=18^a			
	(2) Adaptation	$r = .003, p = .504$	$r = -0.123, p = .313$	NA
	(1) vs. (2)	$z = -1.403, p = .161$	NA	
Semantic ambiguity	(3) Comprehension	$r = -0.038, p = .559$	$r = -.102, p = .656$	NA
	n=18^b			
	(4) Adaptation	$r = -0.035, p = .446$	$r = -0.659, p = .001$	$z = -2.182, p = .0291$
	(3) vs. (4)	NA	$z = -2.652, p = .008$	
Acoustic degradation vs. Semantic ambiguity	(1) vs. (3)	$z = 1.644, p = .100$	NA	
	n=17^{a,b}			
	(2) vs. (4)	NA	$z = -1.678, p = .093$	
	(1) vs. (4)	$z = -1.042, r = .300$	$z = -2.160, r = .031$	
	(3) vs. (2)	NA	NA	

^aexcluding md6

^bexcluding md10

As detailed in the task-specific results (summarised in Table 3), acoustically-degraded speech perception was predicted by the degree of MD network damage and this correlation was in the opposite direction to, but not significantly stronger than, the correlation with language network damage volume (see section 4.1.1). Conversely semantically-ambiguous

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

speech adaptation (word meaning priming) was predicted by language network damage and this correlation was significantly stronger than the non-significant correlation with MD network damage volume (section 4.2.2). This double dissociation provides evidence for causal associations between the integrity of the MD network and abilities at degraded speech perception and between the language network and word meaning priming. Further comparisons of the strength of correlations between tasks within the same type of challenge (acoustic degradation or semantic ambiguity) showed that damage to the language network was more predictive of impaired word meaning priming than of comprehension of ambiguous sentences. This finding provides support for a specific contribution of the language network to adaptation that is independent to its role in comprehension (at least as measured here), which was shown to be largely independent of language or MD network lesions. However, there was no evidence that damage to the MD network was more strongly predictive of degraded speech perception than of degraded speech adaptation.

To further explore the specificity of the contribution of the MD and language networks to acoustically degraded speech perception and word meaning priming respectively, we also compared the strength of correlations between tasks using data from the 17 participants who performed both the acoustic degradation and the semantic ambiguity tasks (Table 3). There was no evidence that damage to the MD Network was more predictive of degraded speech perception than of word meaning priming ($z = -1.042$, $p = .300$), or of comprehension of ambiguous sentences ($z = 1.644$, $p = .100$). Damage to the language network was more predictive of impaired word meaning priming than of degraded speech perception ($z = -2.160$, $p = .031$), although the comparison with degraded speech adaptation did not reach the $p < .05$ threshold ($z = -1.678$, $p = .093$).

Discussion

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

We report two main findings. First, we show that damage to the domain-general MD network, but not the language-selective network, causes significant impairments to the perception of acoustically degraded speech. Word report accuracy for noise-vocoded sentences decreased as the number of channels in the vocoder decreased reflecting an increased challenge to speech perception. The degree of perceptual impairment (i.e. the number of channels required for 50% correct word report) depends on the extent of damage to the MD network, but not damage to the language network (Figure 3; Table 3). Word recognition improved following a period of training, reflecting adaptation or perceptual learning for this form of acoustic degradation, but the degree of learning was not reliably predicted by lesion location or extent.

In contrast to these results with acoustically challenging speech, we found no evidence that semantically-challenging speech comprehension was dependent on the MD system: all participants were highly accurate in judging the coherence of sentences and were no less accurate when the sentences contained ambiguous words, indicating an intact ability to access the typically less frequent (subordinate) word meanings used in our high-ambiguity sentences. Although participants were slower to make judgements for sentences which include ambiguous words, reflecting more effortful comprehension when words have multiple meanings, there was no significant association between response time slowing for ambiguous sentences and the extent of damage to the MD or language networks.

Our second main finding is that despite accurate comprehension of semantically ambiguous speech, damage to the language, but not to the MD network caused a significant reduction in updating of word meaning preferences following recent linguistic experience. As shown in previous studies of individuals without brain lesions, our participants were (as a group) more likely to generate a word associate related to the less frequent meaning of an ambiguous word when they had encountered this meaning in an earlier sentence (word

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

meaning priming, as reported by Rodd et al., 2013). However, the magnitude of this word meaning priming effect was predicted by the extent of damage to the language, but not the MD network (Figure 6; Table 3), a dissociation that was supported by a statistically significant difference between the strength of these two correlations. The reduction in word meaning priming was not explained by sentence comprehension difficulties as there was no correlation between the magnitude of word meaning priming and increased response times when judging the coherence of sentences containing semantically ambiguous words. Furthermore, across-task comparisons showed that the damage to the language network was more predictive of impaired word meaning priming than impaired comprehension of ambiguous sentences or impaired perception of acoustically degraded speech.

Below, we discuss our two main findings in greater detail. First, we discuss possible cognitive operations performed by the MD network that are required for the perception of acoustically degraded speech. We then turn to the linguistic challenge of resolving lexical-semantic ambiguity. We discuss the functional contribution of the language-selective network in adaptation such that low frequency meanings of semantically ambiguous words become more accessible following recent exposure. In a final section we consider the dissociation between these different challenges to speech processing and explore implications for the neural basis of speech perception, comprehension, and adaptation.

The MD network makes a causal contribution to perception of acoustically degraded speech

Recently it has been argued that the MD network does not play a functional role in language comprehension (Blank & Fedorenko, 2017; Diachek et al., 2020; Shain et al., 2020; Wehbe et al., 2021; for reviews, see Fedorenko, 2014; Campbell & Tyler, 2018). According to such an account, activations observed during language comprehension within the MD

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

network could reflect a generic increase in effortful processing or contributions to specific task demands (such as decision making), rather than computations essential for language comprehension (e.g. identifying words; accessing word meanings). However, this line of research left open the possibility that MD contributions may be necessary when speech is acoustically degraded and challenging to perceive (Diachek et al., 2020). Here we provide novel evidence that the MD network indeed makes a causal contribution to perception of acoustically degraded speech by assessing the impact of damage to MD regions on performance in a word report task that indexes cognitive operations required for word identification.

Previous fMRI studies have shown that listening to acoustically challenging speech is associated with an increase in activation in prefrontal and motor regions that plausibly fall within the MD network (Adank, 2012; Davis & Johnsrude, 2003; Du et al., 2016; Erb et al., 2013; Hardy et al., 2018; Hervais-Adelman et al., 2012; Rysop et al., 2021; Vaden et al., 2015; Vaden et al., 2013; Wild et al., 2012). However, these studies did not explicitly define MD regions, or test the necessity of MD contributions and hence this association has not been firmly established. A substantial advance, then, comes from our finding that neural integrity of the MD network supports more successful word report for degraded speech, which allows us to conclude a causal role of MD regions in degraded speech perception.

The MD network has previously been linked to a diverse range of domain-general cognitive constructs, including executive control, working memory, and fluid intelligence. These constructs may reflect a combination of different cognitive operations including setting and monitoring of task goals, directing attention, and the storage, maintenance, integration and inhibition of information across different time scales. It is therefore of interest to consider which of these operations, performed by the MD network, might be critical for the perception of acoustically degraded speech. For example, focused attention may be particularly

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

important when the identities of specific phonemes or words are uncertain. Monitoring may be important for tracking the accuracy of phoneme perception and word recognition over time.

Future work can tease apart these possible distinct cognitive operations, either by focusing on potential contribution of distinct sub-networks within the broader MD network, or by exploring correlations between these other functions of MD networks and perception of degraded speech. Given strong evidence of inter-regional correlations during naturalistic listening paradigms (Assem, Blank, et al., 2020; Blank et al., 2014; Mineroff et al., 2018; Paunov et al., 2019), we here treated the MD network as a functionally integrated system. However, other research concerned with domain general cognitive processes has proposed that the MD network consists of at least two interconnected, but distinct sub-networks (one comprising lateral frontal and parietal areas, and the other—cingular and opercular areas), which may contribute differently to cognition (Dosenbach, Fair, Cohen, Schlaggar, & Petersen, 2008; Dosenbach et al., 2007; Nomura et al., 2010). In the context of effortful speech comprehension, Peelle (2018) proposes a three-way distinction between fronto-parietal, premotor and cingular-opercular contributions to attention, working memory and performance monitoring processes respectively. Consistent with the proposed role of cingular-opercular regions are data showing that activation is associated with better word recognition on subsequent trials (Vaden et al., 2015; Vaden et al., 2013), which may reflect mechanisms for tracking the accuracy of phoneme perception and word recognition over time. Although the present data cannot adjudicate between bi- or tripartite views, further research using similar methods and data from a larger number of individuals could potentially dissociate the effect of lesions of these three sub-networks and establish underlying mechanisms. For example, we might predict that focal damage to cingular-opercular regions would result in a greater impairment in degraded speech perception when perceptual

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

difficulty varies from trial-to-trial compared to cases in which trial difficulty is grouped into blocks.

Replicating a range of previous behavioural findings (Davis et al., 2005; Hervais-Adelman et al., 2008; Huyck & Johnsrude, 2012; Loebach & Pisoni, 2008; Peelle & Wingfield, 2005; Sohoglu & Davis, 2016), we showed that listeners adapt to acoustically degraded speech over time. This finding extends earlier observations of perceptual learning to individuals with lesions to language-selective and domain-general regions. We found no evidence that damage to either MD or language-selective networks led to reduced perceptual learning, and hence cannot make causal claims about the contribution of either network to this form of learning. Future studies using similar methods would benefit from a larger number of participants, with more variable and more extensive lesions.

The language-selective network makes a causal contribution to adaptation to semantically ambiguous speech

Semantically ambiguous words introduce a substantial challenge to speech comprehension because of the need to engage competition processes to select between alternative meanings and the cognitive cost of reinterpretation when initial selection fails (Rodd et al., 2002; Rodd et al., 2010, 2012). The presence of two or more ambiguous words in each of the high ambiguity sentences used in our study made comprehension especially challenging. Nonetheless, we observed that comprehension – indicated by judging high ambiguity sentences to be coherent – was ultimately successful (although slower than for low ambiguity sentences) and that accuracy in judging coherence did not differ between high and low ambiguity sentences. Neither response time differences nor the accuracy of coherence judgements were associated with the degree of damage to MD or language-selective brain

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

networks. Thus, our study does not provide evidence for a specific causal role of either of these brain networks for comprehension of sentences containing ambiguous words.

Despite intact comprehension of sentences containing semantically ambiguous words, we observed differential effects of lesion location and extent on learning mechanisms involved in adapting lexical-semantic processing after successful disambiguation. Previous research has established that recent exposure to low-frequency (subordinate) meanings of ambiguous words in a sentence context facilitates subsequent meaning access and selection of those meanings, a process termed word-meaning priming (Betts et al., 2018; Gaskell et al., 2019; Gilbert, Davis, Gaskell, & Rodd, 2018; Rodd et al., 2016; Rodd et al., 2013). Previous functional imaging studies have not studied neural activity associated with word-meaning priming and hence the present results make a novel contribution to understanding the neural basis of this adaptation process. We here replicated the standard word meaning priming effect for the group of participants tested overall, but showed that the magnitude of the priming effect was significantly reduced by damage to the language-selective but not the MD network.

There is substantial anatomical overlap between the language network shown here to be critical for updating of word-meaning preferences following successful disambiguation, and the fronto-temporal brain regions previously shown to respond to semantic ambiguity resolution (Bilenko et al., 2009; Musz & Thompson-Schill, 2017; Rodd et al., 2005; Vitello et al., 2014; Zempleni et al., 2007; for a review, see Rodd, 2020), consistent with shared neural resources between semantic comprehension and subsequent adaptation. The absence of an effect of language-network damage on immediate comprehension, coupled with the observed impact of language-network damage on semantic adaptation may therefore reflect the relatively high functioning of the volunteers, the limited severity of the language lesions, and/or the relative insensitivity of the comprehension task to distinguishing between these

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

relatively unimpaired volunteers. It also remains possible that particular subregions of the language network are differentially important for immediate comprehension compared to subsequent adaptation. These issues could be explored in future work, which would clearly benefit from larger numbers of participants. A larger sample would also allow the use of alternative methods such as Voxel-based Lesion-Symptom Mapping to localise function more specifically within the network, for example by contrasting frontal and temporal lobe lesions.

One striking illustration of the longevity of learning is that word-meaning priming has previously been observed 24 hours after a single exposure to an ambiguous word; especially if there is an intervening period of sleep (Gaskell et al., 2019). This latter finding, in combination with a wider literature on the role of consolidation processes that facilitate the acquisition of new lexical knowledge (Dumay & Gaskell, 2007; Gaskell & Dumay, 2003; Tamminen, Payne, Stickgold, Wamsley, & Gaskell, 2010), led Gaskell et al. (2019) to suggest that word meaning priming may involve a two-stage complementary systems account of learning (McClelland, 2013), as proposed for the acquisition of novel words (Davis & Gaskell, 2009). According to this account, short-term learning arises from hippocampally-mediated binding of associations between words in the sentences, while these short-term changes are consolidated into long-term changes to word meaning preferences after sleep.

The present study constrains these complementary systems accounts of learning by revealing a causal contribution of language-selective cortical regions even for short-term adaptation of familiar word meanings. Future work could further consider the interaction of hippocampal and cortical regions in the learning and maintenance of meaning preferences over different time scales and the relationship between learning novel vocabulary and updating of existing lexical semantic knowledge (for recent meta-analyses of word form learning and consolidation, see, Schimke, Angwin, Cheng, & Copland, 2021; Tagarelli, Shattuck, Turkeltaub, & Ullman, 2019). We note that previous research has shown that

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

individuals with aphasia (identified behaviourally) can learn novel vocabulary but that learning is highly variable (Kelly et al., 2009; Tuomiranta et al. 2011; 2014). The present work similarly shows variability in the impact of cortical lesions on adapting the meanings of familiar words. However, our participants were not recruited on the basis of language impairment and retained good comprehension both on a standardised measure of sentence comprehension (TROG2) and on the experimental measure of ambiguity resolution tested here. It might be that individuals with more extensive lesions to language selective cortex, or more focal lesions of posterior temporal and inferior frontal regions that contribute to ambiguity resolution would show a greater impairment to comprehension. Such a finding would suggest that a common set of cortical regions support comprehension and learning of ambiguous words in sentences. Further refinement of lesion definitions and tests of larger samples of individuals could also provide more detailed anatomical evidence concerning the relative contributions of language-selective and/or domain-general sub-regions of the IFG to semantic ambiguity resolution. These regions lie in close proximity and thus may appear to overlap in group studies (Fedorenko & Blank, 2020), which may explain why they have not been dissociated in previous imaging research on semantic ambiguity resolution.

Neural dissociation of different challenges to speech perception and comprehension

Taken together, our findings provide a double dissociation indicating independent functional contributions of the MD and language-selective networks to responding to and adapting to different types of difficult-to-understand sentences. Specifically, we show that the challenge of perceiving acoustically degraded sentences (measured in terms of word report accuracy) is causally linked to the degree of damage to the MD network but not to the language-selective network (although the comparison of correlations was not statistically significant; see Table 3). Conversely, the challenge of post-comprehension adaptation to

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

semantically ambiguous words in sentences (measured in terms of word meaning priming) causally depends on the integrity of the language-selective but not the MD network; moreover, in this case there is a reliable difference between the significant (language) and null (MD) correlations.

Here we tested a limited set of challenges to speech comprehension, thus we cannot make general statements concerning dissociable contributions made by each of these cortical networks to all forms of perceptual or semantic challenge. However, our data provide initial evidence for the task-specificity of causal contributions. Focusing first on the effect of language network damage, the correlation between language lesion volume and word meaning priming was significantly different from the null correlation between lesion volume and coherence judgment response times for ambiguous sentences, indicating a greater sensitivity to the integrity of the language network for adaptation compared to initial comprehension of lexico-semantic ambiguity. It is important to note that our results are based on data from individuals without aphasia in whom lesions extended to a maximum of only ~11 % of the language cortex (Supplementary Information). It is likely that more extensive lesions, or indeed more sensitive tests, would detect a contribution of the language network to comprehension. Furthermore, the reliable correlation between lesion volume and word-meaning priming could also be dissociated from the (null) effect of lesions on degraded speech perception, suggesting that the integrity of the language network is more important for lexico-semantic than for acoustic or perceptual challenges. We note, however, that our definition of the language network was derived from studies using both written and spoken language and hence likely excluded early auditory processing stages. It would be of interest to explore lesion definitions based on localising the speech perception system which might reveal other systems that causally support abilities at degraded speech perception.

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

Equivalent across-task comparisons of the effect of MD network damage did not reach statistical significance. Despite a reliable correlation between MD lesion volume and impaired perception of acoustically degraded speech, this effect could not be clearly dissociated from the null effects of lesion volume on tasks involving semantic processing, or perceptual and semantic adaptation. We therefore cannot draw strong conclusions about the specificity of the contribution of the MD network to degraded speech perception based on the current data.

For instance, it remains possible that the MD contribution is a secondary consequence of an increase in demands on domain-general cognition (e.g. working memory) that would affect all aspects of language functioning, but are emphasised by the word report task for degraded speech. While we had expected that additional demands on domain-general operations would be equivalent across each of our listening challenges and tasks, we do not have evidence that this was the case. It might be that more careful titration of task difficulty will be required if we are to demonstrate that network-specific lesions impair specific aspect(s) of language function (perception vs. comprehension vs. adaptation to challenge types) rather than apparently specific effects being mediated by domain general processes.

Relatedly, a reviewer raised the possibility that unmeasured auditory and cognitive impairments may have had consequences for participants' task performances. Such impairments could affect performance even when participants achieved a high level of accuracy such as reporting words for clear speech or distinguishing coherent and anomalous sentences. Further, they could potentially account for the poorer task performance when challenge was introduced. This concern is based on previous research that demonstrates a variety of higher-level cognitive consequences of hearing impairment on tasks requiring speech perception and comprehension (for a review, see Humes et al., 2012). For example, a participant with a hearing impairment may experience increased demands on domain-general

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

functions depending on the task, and the extent to which they can manage these demands may therefore depend on cognitive abilities. As above, we had no reason to suppose differential effects of any auditory or cognitive impairments on our different linguistic tasks, but we acknowledge the limitation and suggest that future studies should measure auditory and cognitive abilities more broadly alongside linguistic measures of interest.

Further studies exploring a wider range of challenges to speech comprehension and with larger samples of participants might specify the causal contributions identified here in more detail. For example, we might assess lesion correlates of perception and adaptation for other forms of perceptual challenges to speech comprehension such as those arising when speech is in background noise, or speech sounds are perceptually ambiguous (see Mattys et al., 2012 for a review of these listening challenges). Future studies might also consider whether other forms of semantic, syntactic or lexical challenge to comprehension are also causally associated with the integrity of the MD or language networks. In this way, building on the current methods and findings, one could map the hierarchy of cognitive processes involved in speech perception and comprehension onto specific brain regions that support them. However, as mentioned above (sections 5.1 and 5.2), larger samples of patients will be needed if we are to conduct more anatomically specific analyses at the level of individual voxels (e.g. using Voxel-based Lesion-Symptom Mapping) or functional subregions within the larger networks studied here.

Conclusions

Speech comprehension in naturalistic situations requires listeners to accommodate and learn in response to a range of perceptual and semantic challenges that make spoken sentences more difficult to recognise and understand. Behavioural data from individuals with lesions to language-selective and domain-general MD networks demonstrate different

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

functional contributions of these two networks depending on the source of the listening challenge. In particular, the MD network appears to be necessary for the perception of acoustically degraded speech, whereas using recent experience to update meaning preferences for ambiguous words appears to depend on anatomically distinct, fronto-temporal regions argued to form a specialised language network.

In this work we considered two specific challenges, but future work should consider whether differences in the ways in which acoustic degradation and lexical-semantic ambiguity engage and depend on the domain-general MD network and domain-selective language network translate to other perceptual and semantic challenges and to more naturalistic speech processing. For example, speech perception must be resilient in the face of unfamiliar accents, mispronunciations, and competing sounds. Comprehension processes must accommodate multiple forms of syntactically or semantically complex and ambiguous speech. Many of these are situations in which activation of inferior frontal regions has been observed (Blanco-Elorrieta, Gwilliams, Marantz, & Pylkkanen, 2021; Boudewyn et al., 2015; January, Trueswell, & Thompson-Schill, 2009; Kuperberg et al., 2003; Novais-Santos et al., 2007) and an attribution to domain-general MD processing has sometimes been made. However, given the evidence for functionally distinct language-selective and domain-general sub-regions lying in close proximity within the IFG, and the individual variability in their precise locations (Fedorenko & Blank, 2020), such conclusions may be premature. Further studies of individuals with focal lesions can be used to determine whether accommodating these other perceptual and semantic challenges to speech processing similarly depends on the integrity of domain-general or language-selective brain regions. These perceptual and semantic challenges are common for the noisy and ambiguous spoken language that listeners perceive and comprehend every day.

References

- Adank, P. (2012). The neural bases of difficult speech comprehension and speech production: Two Activation Likelihood Estimation (ALE) meta-analyses. *Brain and Language*, *122*(1), 42-54.
- Altmann, G. T., & Kamide, Y. (1999). Incremental interpretation at verbs: restricting the domain of subsequent reference. *Cognition*, *73*(3), 247-264.
- Apperly, I. A., Samson, D., Carroll, N., Hussain, S., & Humphreys, G. (2006). Intact first- and second-order false belief reasoning in a patient with severely impaired grammar. *Social Neuroscience*, *1*(3-4), 334-348.
- Assem, M., Blank, I. A., Mineroff, Z., Ademoglu, A., & Fedorenko, E. (2020). Activity in the fronto-parietal multiple-demand network is robustly associated with individual differences in working memory and fluid intelligence. *Cortex*, *131*, 1-16.
- Assem, M., Glasser, M. F., Van Essen, D. C., & Duncan, J. (2020). A Domain-General Cognitive Core Defined in Multimodally Parcellated Human Cortex. *Cerebral Cortex*, *30*(8), 4361-4380.
- Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research*, *3*(2), 12-28.
- Bates, D., Machler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1-48.
- Bates, E., Wilson, S. M., Saygin, A. P., Dick, F., Sereno, M. I., Knight, R. T., et al. (2003). Voxel-based lesion-symptom mapping. *Nature Neuroscience*, *6*(5), 448-450.
- Betts, H. N., Gilbert, R. A., Cai, Z. G., Okedara, Z. B., & Rodd, J. M. (2018). Retuning of lexical-semantic representations: Repetition and spacing effects in word-meaning priming. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *44*(7), 1130-1150.

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

- Bilenko, N. Y., Grindrod, C. M., Myers, E. B., & Blumstein, S. E. (2009). Neural correlates of semantic competition during processing of ambiguous words. *Journal of Cognitive Neuroscience*, *21*(5), 960-975.
- Binder, J. R., Frost, J. A., Hammeke, T. A., Cox, R. W., Rao, S. M., & Prieto, T. (1997). Human Brain Language Areas Identified by Functional Magnetic Resonance Imaging. *The Journal of Neuroscience*, *17*(1), 353-362.
- Bishop, D. (2003). Test for reception of grammar: TROG-2. London: PsychCorp, Pearson.
- Blanco-Elorrieta, E., Gwilliams, L., Marantz, A., & Pylkkanen, L. (2021). Adaptation to mispronounced speech: evidence for a prefrontal-cortex repair mechanism. *Scientific Reports*, *11*(1), 97.
- Blank, I. A., & Fedorenko, E. (2017). Domain-General Brain Regions Do Not Track Linguistic Input as Closely as Language-Selective Regions. *Journal of Neuroscience*, *37*(41), 9999-10011.
- Blank, I. A., Kanwisher, N., & Fedorenko, E. (2014). A functional dissociation between language and multiple-demand systems revealed in patterns of BOLD signal fluctuations. *Journal of Neurophysiology*, *112*(5), 1105-1118.
- Blott, L. M., Rodd, J. M., Ferreira, F., & Warren, J. E. (2021). Recovery from misinterpretations during online sentence processing. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *47*(6), 968-997.
- Boudewyn, M. A., Long, D. L., Traxler, M. J., Lesh, T. A., Dave, S., Mangun, G. R., et al. (2015). Sensitivity to Referential Ambiguity in Discourse: The Role of Attention, Working Memory, and Verbal Ability. *Journal of Cognitive Neuroscience*, *27*(12), 2309-2323.

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

- Braga, R. M., DiNicola, L. M., Becker, H. C., & Buckner, R. L. (2020). Situating the left-lateralized language network in the broader organization of multiple specialized large-scale distributed networks. *Journal of Neurophysiology*, *124*(5), 1415-1448.
- Campbell, K. L., & Tyler, L. K. (2018). Language-related domain-specific and domain-general systems in the human brain. *Current Opinions in Behavioral Sciences*, *21*, 132-137.
- Cole, M., & Schneider, W. (2007). The cognitive control network: Integrated cortical regions with dissociable functions. *NeuroImage*, *37*, 343-360.
- Cutler, A., Dahan, D., & van Donselaar, W. (1997). Prosody in the Comprehension of Spoken Language: A Literature Review. *Language and Speech*, *40*(2), 141-201.
- D'Ausilio, A., Pulvermüller, F., Salmas, P., Bufalari, I., Begliomini, C., & Fadiga, L. (2009). The Motor Somatotopy of Speech Perception. *Current Biology*, *19*(5), 381-385.
- D'Ausilio, A., Bufalari, I., Salmas, P., & Fadiga, L. (2012). The role of the motor system in discriminating normal and degraded speech sounds. *Cortex*, *48*(7), 882-887.
- Davis, M. H., Ford, M. A., Kherif, F., & Johnsrude, I. S. (2011). Does semantic context benefit speech understanding through "top-down" processes? Evidence from time-resolved sparse fMRI. *Journal of Cognitive Neuroscience*, *23*(12), 3914-3932.
- Davis, M. H., & Gaskell, M. G. (2009). A complementary systems account of word learning: neural and behavioural evidence. *Philosophical Transactions of the Royal Society of London B Biological Sciences*, *364*(1536), 3773-3800.
- Davis, M. H., & Johnsrude, I. S. (2003). Hierarchical processing in spoken language comprehension. *Journal of Neuroscience*, *23*(8), 3423-3431.
- Davis, M. H., Johnsrude, I. S., Hervais-Adelman, A., Taylor, K., & McGettigan, C. (2005). Lexical information drives perceptual learning of distorted speech: evidence from the

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

- comprehension of noise-vocoded sentences. *Journal of Experimental Psychology: General*, *134*(2), 222-241.
- Diachek, E., Blank, I. A., Siegelman, M., Affourtit, J., & Fedorenko, E. (2020). The Domain-General Multiple Demand (MD) Network Does Not Support Core Aspects of Language Comprehension: A Large-Scale fMRI Investigation. *Journal of Neuroscience*, *40*(23), 4536-4550.
- Diedenhofen, B., & Musch, J. (2015). cocor: A comprehensive solution for the statistical comparison of correlations. *PLoS One*, *10*(4), e0121945.
- Dosenbach, N. U., Fair, D. A., Cohen, A. L., Schlaggar, B. L., & Petersen, S. E. (2008). A dual-networks architecture of top-down control. *Trends in Cognitive Sciences*, *12*(3), 99-105.
- Dosenbach, N. U., Fair, D. A., Miezin, F. M., Cohen, A. L., Wenger, K. K., Dosenbach, R. A., et al. (2007). Distinct brain networks for adaptive and stable task control in humans. *Proceedings of the National Academy of Sciences USA*, *104*(26), 11073-11078.
- Du, Y., Buchsbaum, B. R., Grady, C. L., & Alain, C. (2014). Noise differentially impacts phoneme representations in the auditory and speech motor systems. *Proceedings of the National Academy of Sciences USA*, *111*(19), 7126-7131.
- Du, Y., Buchsbaum, B. R., Grady, C. L., & Alain, C. (2016). Increased activity in frontal motor cortex compensates impaired speech perception in older adults. *Nature Communications*, *7*, 12241.
- Duffy, S. A., Morris, R. K., & Rayner, K. (1988). Lexical Ambiguity and Fixation Times in Reading. *Journal of Memory and Language*, *27*(4), 429-446.
- Dumay, N., & Gaskell, M. G. (2007). Sleep-associated changes in the mental representation of spoken words. *Psychological Science*, *18*(1), 35-39.

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

- Duncan, J. (2010a). *How Intelligence Happens*: Yale University Press.
- Duncan, J. (2010b). The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour. *Trends in Cognitive Sciences*, 14(4), 172-179.
- Duncan, J. (2013). The structure of cognition: attentional episodes in mind and brain. *Neuron*, 80(1), 35-50.
- Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neuroscience*, 23(10), 475-483.
- Elliott, T. M., & Theunissen, F. E. (2009). The Modulation Transfer Function for Speech Intelligibility. *PLOS Computational Biology*, 5(3), e1000302.
- Erb, J., Henry, M. J., Eisner, F., & Obleser, J. (2013). The brain dynamics of rapid perceptual adaptation to adverse listening conditions. *Journal of Neuroscience*, 33(26), 10688-10697.
- Evans, S., & Davis, M. H. (2015). Hierarchical Organization of Auditory and Motor Representations in Speech Perception: Evidence from Searchlight Similarity Analysis. *Cerebral Cortex*, 25(12), 4772-4788.
- Fedorenko, E. (2014). The role of domain-general cognitive control in language comprehension. *Frontiers in Psychology*, 5, 335.
- Fedorenko, E. (2021). The early origins and the growing popularity of the individual-subject analytic approach in human neuroscience. *Current Opinion in Behavioral Sciences*, 40, 105-112.
- Fedorenko, E., Behr, M. K., & Kanwisher, N. (2011). Functional specificity for high-level linguistic processing in the human brain. *Proceedings of the National Academy of Sciences USA*, 108(39), 16428-16433.
- Fedorenko, E., & Blank, I. A. (2020). Broca's Area Is Not a Natural Kind. *Trends in Cognitive Sciences*, 24(4), 270-284.

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

- Fedorenko, E., Duncan, J., & Kanwisher, N. (2012). Language-selective and domain-general regions lie side by side within Broca's area. *Current Biology*, *22*(21), 2059-2062.
- Fedorenko, E., Duncan, J., & Kanwisher, N. (2013). Broad domain generality in focal regions of frontal and parietal cortex. *Proceedings of the National Academy of Sciences USA*, *110*(41), 16616-16621.
- Fedorenko, E., Hsieh, P. J., Nieto-Castanon, A., Whitfield-Gabrieli, S., & Kanwisher, N. (2010). New method for fMRI investigations of language: defining ROIs functionally in individual subjects. *Journal of Neurophysiology*, *104*(2), 1177-1194.
- Fedorenko, E., & Shain, C. (2021). Local implementation of general computations: The case of human language comprehension. *Current Directions in Psychological Science*.
- Fedorenko, E., & Varley, R. A. (2016). Language and thought are not the same thing: evidence from neuroimaging and neurological patients. *Annals of the New York Academy of Sciences*, *1369*(1), 132-153.
- Garrod, S., & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive Sciences*, *8*(1), 8-11.
- Gaskell, M. G., Cairney, S. A., & Rodd, J. M. (2019). Contextual priming of word meanings is stabilized over sleep. *Cognition*, *182*, 109-126.
- Gaskell, M. G., & Dumay, N. (2003). Lexical competition and the acquisition of novel words. *Cognition*, *89*(2), 105-132.
- Gernsbacher, M. A., & Faust, M. E. (1991). The mechanism of suppression: a component of general comprehension skill. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *17*(2), 245-262.
- Gernsbacher, M. A., Varner, K. R., & Faust, M. E. (1990). Investigating differences in general comprehension skill. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *16*(3), 430-445.

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

- Gilbert, R. A., Davis, M. H., Gaskell, M. G., & Rodd, J. M. (2018). Listeners and readers generalize their experience with word meanings across modalities. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 44(10), 1533-1561.
- Gilbert, R. A., & Rodd, J. M. (2022). Dominance Norms and Data for Spoken Ambiguous Words in British English. *Journal of Cognition*, 5(1), 4.
- Guediche, S., Blumstein, S. E., Fiez, J. A., & Holt, L. L. (2014). Speech perception under adverse conditions: insights from behavioral, computational, and neuroscience research. *Frontiers in Systems Neuroscience*, 7, 126-126.
- Hagoort, P. (1993). Impairments of Lexical Semantic Processing in Aphasia - Evidence from the Processing of Lexical Ambiguities. *Brain and Language*, 45(2), 189-232.
- Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science*, 304(5669), 438-441.
- Halai, A. D., Woollams, A. M., & Lambon Ralph, M. A. (2017). Using principal component analysis to capture individual differences within a unified neuropsychological model of chronic post-stroke aphasia: Revealing the unique neural correlates of speech fluency, phonology and semantics. *Cortex*, 86, 275-289.
- Hardy, C. J. D., Marshall, C. R., Bond, R. L., Russell, L. L., Dick, K., Ariti, C., et al. (2018). Retained capacity for perceptual learning of degraded speech in primary progressive aphasia and Alzheimer's disease. *Alzheimers Research Therapy*, 10(1), 70.
- Hervais-Adelman, A., Carlyon, R. P., Johnsrude, I. S., & Davis, M. H. (2012). Brain regions recruited for the effortful comprehension of noise-vocoded words. *Language and Cognitive Processes*, 27(7-8), 1145-1166.

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

- Hervais-Adelman, A., Davis, M. H., Johnsrude, I. S., & Carlyon, R. P. (2008). Perceptual learning of noise vocoded words: effects of feedback and lexicality. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(2), 460-474.
- Hugdahl, K., Raichle, M. E., Mitra, A., & Specht, K. (2015). On the existence of a generalized non-specific task-dependent network. *Frontiers in Human Neuroscience*, *9*, 430-430.
- Humes, L. E., Dubno, J. R., Gordon-Salant, S., Lister, J. J., Cacace, A. T., Cruickshanks, K. J., et al. (2012). Central presbycusis: a review and evaluation of the evidence. *Journal of the American Academy of Audiology*, *23*(8), 635-666.
- Huyck, J. J., & Johnsrude, I. S. (2012). Rapid perceptual learning of noise-vocoded speech requires attention. *Journal of the Acoustical Society of America*, *131*(3), EL236-242.
- Huyck, J. J., Smith, R. H., Hawkins, S., & Johnsrude, I. S. (2017). Generalization of Perceptual Learning of Degraded Speech Across Talkers. *Journal of Speech, Language and Hearing Research*, *60*(11), 3334-3341.
- Ivanova, A. A., Mineroff, Z., Zimmerer, V., Kanwisher, N., Varley, R., & Fedorenko, E. (2021). The Language Network Is Recruited but Not Required for Nonverbal Event Semantics. *Neurobiology of Language*, *2*(2), 176-201.
- Ivanova, A. A., Srikant, S., Sueoka, Y., Kean, H. H., Dhamala, R., O'Reilly, U. M., et al. (2020). Comprehension of computer code relies primarily on domain-general executive brain regions. *Elife*, *9*.
- January, D., Trueswell, J. C., & Thompson-Schill, S. L. (2009). Co-localization of Stroop and Syntactic Ambiguity Resolution in Broca's Area: Implications for the Neural Basis of Sentence Processing. *Journal of Cognitive Neuroscience*, *21*(12), 2434-2444.

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

- Johnsrude, I. S., & Rodd, J. R. (2015). Factors that increase processing demands when listening to speech. In G. Hickok & S. Small (Eds.), *Neurobiology of Language* (pp. 491-502): Academic Press.
- Khanna, M. M., & Boland, J. E. (2010). Children's use of language context in lexical ambiguity resolution. *Quarterly Journal of Experimental Psychology*, *63*(1), 160-193.
- Kuperberg, G. R., Holcomb, P. J., Sitnikova, T., Greve, D., Dale, A. M., & Caplan, D. (2003). Distinct patterns of neural modulation during the processing of conceptual and syntactic anomalies. *Journal of Cognitive Neuroscience*, *15*(2), 272-293.
- Linares, D., & Lopez-Moliner, J. (2006). quickpsy: An R package to fit psychometric functions for multiple groups. *The R Journal*, *8*(1), 122-131.
- Loebach, J. L., & Pisoni, D. B. (2008). Perceptual learning of spectrally degraded speech and environmental sounds. *Journal of the Acoustical Society of America*, *123*(2), 1126-1139.
- Macmillan, N. A., & Kaplan, H. L. (1985). Detection theory analysis of group data: Estimating sensitivity from average hit and false-alarm rates. *Psychological Bulletin*, *98*(1), 185-199.
- MacSweeney, M., Woll, B., Campbell, R., McGuire, P. K., David, A. S., Williams, S. C. R., et al. (2002). Neural systems underlying British Sign Language and audio-visual English processing in native users. *Brain*, *125*(7), 1583-1593.
- Mahowald, K., & Fedorenko, E. (2016). Reliable individual-level neural markers of high-level language processing: A necessary precursor for relating neural variability to behavioral and genetic variability. *Neuroimage*, *139*, 74-93.
- Mattys, S. L., Davis, M. H., Bradlow, A. R., & Scott, S. K. (2012). Speech recognition in adverse conditions: A review. *Language and Cognitive Processes*, *27*(7-8), 953-978.

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

- McClelland, J. L. (2013). Incorporating Rapid Neocortical Learning of New Schema-Consistent Information Into Complementary Learning Systems Theory. *Journal of Experimental Psychology: General*, *142*(4), 1190-1210.
- McGettigan, C., Rosen, S., & Scott, S. K. (2014). Lexico-semantic and acoustic-phonetic processes in the perception of noise-vocoded speech: implications for cochlear implantation. *Frontiers in Systems Neuroscience*, *8*, 18.
- Meister, I. G., Wilson, S. M., Deblieck, C., Wu, A. D., & Iacoboni, M. (2007). The essential role of premotor cortex in speech perception. *Current Biology*, *17*(19), 1692-1696.
- Meng, X. L., Rosenthal, R., & Rubin, D. B. (1992). Comparing Correlated Correlation-Coefficients. *Psychological Bulletin*, *111*(1), 172-175.
- Mesulam, M. M., Rogalski, E. J., Wieneke, C., Hurley, R. S., Geula, C., Bigio, E. H., et al. (2014). Primary progressive aphasia and the evolving neurology of the language network. *Nature Reviews Neurology*, *10*(10), 554-569.
- Miller, G. A., Heise, G. A., & Lichten, W. (1951). The intelligibility of speech as a function of the context of the test materials. *Journal of Experimental Psychology: Human Perception and Performance*, *41*, 329-335.
- Miller, G. A., & Isard, S. (1963). Some perceptual consequences of linguistic rules. *Journal of Verbal Learning and Verbal Behavior*, *2*(3), 217-228.
- Mineroff, Z., Blank, I. A., Mahowald, K., & Fedorenko, E. (2018). A robust dissociation among the language, multiple demand, and default mode networks: Evidence from inter-region correlations in effect size. *Neuropsychologia*, *119*, 501-511.
- Mirman, D., Chen, Q., Zhang, Y. S., Wang, Z., Faseyitan, O. K., Coslett, H. B., et al. (2015). Neural organization of spoken language revealed by lesion-symptom mapping. *Nature Communications*, *6*.

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

- Mirman, D., & Thye, M. (2018). Uncovering the Neuroanatomy of Core Language Systems Using Lesion-Symptom Mapping. *Current Directions in Psychological Science*, 27(6), 455-461.
- Moineau, S., Dronkers, N. F., & Bates, E. (2005). Exploring the processing continuum of single-word comprehension in aphasia. *Journal of Speech, Language and Hearing Research*, 48(4), 884-896.
- Monti, M. M., Parsons, L. M., & Osherson, D. N. (2012). Thought beyond language: neural dissociation of algebra and natural language. *Psychological Science*, 23(8), 914-922.
- Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorial in Quantitative Methods for Psychology*, 4(2), 61-64.
- Münster, K., & Knoeferle, P. (2018). Extending Situated Language Comprehension (Accounts) with Speaker and Comprehender Characteristics: Toward Socially Situated Interpretation. *Frontiers in Psychology*, 8(2267).
- Musz, E., & Thompson-Schill, S. L. (2017). Tracking competition and cognitive control during language comprehension with multi-voxel pattern analysis. *Brain and Language*, 165, 21-32.
- Nelson, H. E. (1982). National Adult Reading Test (NART): For the assessment of premorbid intelligence in patients with demntia: Test manual. Windsor: NFER-Nelson.
- Nomura, E. M., Gratton, C., Visser, R. M., Kayser, A., Perez, F., & D'Esposito, M. (2010). Double dissociation of two cognitive control networks in patients with focal brain lesions. *Proceedings of the National Academy of Sciences USA*, 107(26), 12017-12022.
- Novais-Santos, S., Gee, J., Shah, M., Troiani, V., Work, M., & Grossman, M. (2007). Resolving sentence ambiguity with planning and working memory resources: Evidence from fMRI. *Neuroimage*, 37(1), 361-378.

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

- Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2005). Cognitive control and parsing: Reexamining the role of Broca's area in sentence comprehension. *Cognitive Affective and Behavioral Neuroscience*, 5(3), 263-281.
- Obleser, J., Wise, R. J., Dresner, M. A., & Scott, S. K. (2007). Functional integration across brain regions improves speech perception under adverse listening conditions. *Journal of Neuroscience*, 27(9), 2283-2289.
- Özyürek, A. (2014). Hearing and seeing meaning in speech and gesture: insights from brain and behaviour. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1651), 20130296.
- Paunov, A. M., Blank, I. A., & Fedorenko, E. (2019). Functionally distinct language and Theory of Mind networks are synchronized at rest and during language comprehension. *Journal of Neurophysiology*, 121(4), 1244-1265.
- Peelle, J. E., Gross, J., & Davis, M. H. (2013). Phase-locked responses to speech in human auditory cortex are enhanced during comprehension. *Cerebral Cortex*, 23(6), 1378-1387.
- Peelle, J. E., & Wingfield, A. (2005). Dissociations in perceptual learning revealed by adult age differences in adaptation to time-compressed speech. *Journal of Experimental Psychology: Human Perception and Performance*, 31(6), 1315-1330.
- Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W., Humes, L. E., et al. (2016). Hearing Impairment and Cognitive Energy: The Framework for Understanding Effortful Listening (FUEL). *Ear and Hearing*, 37 Suppl 1, 5s-27s.
- Polk, M., & Kertesz, A. (1993). Music and Language in Degenerative Disease of the Brain. *Brain and Cognition*, 22(1), 98-117.
- Pulvermuller, F., & Fadiga, L. (2010). Active perception: sensorimotor circuits as a cortical basis for language. *Nature Reviews Neuroscience*, 11(5), 351-360.

- Quillen, I. A., Yen, M., & Wilson, S. M. (2021). Distinct Neural Correlates of Linguistic and Non-Linguistic Demand. *Neurobiology of Language*, 1-24.
- R Development Core Team. (2019). R: A language and environment for statistical computing (Version 3.6.1). Vienna, Austria.
- Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: effects of word frequency, verb complexity, and lexical ambiguity. *Memory and Cognition*, 14(3), 191-201.
- Rodd, J. M. (2018). Lexical ambiguity. In S.-A. Rueschemeyer & M. G. Gaskell (Eds.), *Oxford Handbook of Psycholinguistics* (2nd ed., pp. 96-117). Oxford: OUP Oxford.
- Rodd, J. M. (2020). Settling Into Semantic Space: An Ambiguity-Focused Account of Word-Meaning Access. *Perspectives in Psychological Science*, 15(2), 411-427.
- Rodd, J. M., Cai, Z. G. G., Betts, H. N., Hanby, B., Hutchinson, C., & Adler, A. (2016). The impact of recent and long-term experience on access to word meanings: Evidence from large-scale internet-based experiments. *Journal of Memory and Language*, 87, 16-37.
- Rodd, J. M., Cutrin, B. L., Kirsch, H., Millar, A., & Davis, M. H. (2013). Long-term priming of the meanings of ambiguous words. *Journal of Memory and Language*, 68(2), 180-198.
- Rodd, J. M., Davis, M. H., & Johnsrude, I. S. (2005). The neural mechanisms of speech comprehension: fMRI studies of semantic ambiguity. *Cerebral Cortex*, 15(8), 1261-1269.
- Rodd, J. M., Gaskell, G., & Marslen-Wilson, W. D. (2002). Making sense of semantic ambiguity: Semantic competition in lexical access. *Journal of Memory and Language*, 46(2), 245-266.

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

- Rodd, J. M., Johnsrude, I. S., & Davis, M. H. (2010). The role of domain-general frontal systems in language comprehension: evidence from dual-task interference and semantic ambiguity. *Brain Lang*, *115*(3), 182-188.
- Rodd, J. M., Johnsrude, I. S., & Davis, M. H. (2012). Dissociating frontotemporal contributions to semantic ambiguity resolution in spoken sentences. *Cereb Cortex*, *22*(8), 1761-1773.
- Rysop, A. U., Schmitt, L. M., Obleser, J., & Hartwigsen, G. (2021). Neural modelling of the semantic predictability gain under challenging listening conditions. *Hum Brain Mapp*, *42*(1), 110-127.
- Schimke, E. A. E., Angwin, A. J., Cheng, B. B. Y., & Copland, D. A. (2021). The effect of sleep on novel word learning in healthy adults: A systematic review and meta-analysis. *Psychonomic Bulletin and Review*, *28*(6), 1811-1838.
- Scott, T. L., Gallee, J., & Fedorenko, E. (2017). A new fun and robust version of an fMRI localizer for the frontotemporal language system. *Cognitive Neuroscience*, *8*(3), 167-176.
- Seidenberg, M. S., Tanenhaus, M. K., Leiman, J. M., & Bienkowski, M. (1982). Automatic access of the meanings of ambiguous words in context: some limitations of knowledge-based processing. *Cognitive Psychology*, *14*, 489-537.
- Shain, C., Blank, I. A., van Schijndel, M., Schuler, W., & Fedorenko, E. (2020). fMRI reveals language-specific predictive coding during naturalistic sentence comprehension. *Neuropsychologia*, *138*, 107307.
- Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science*, *270*(5234), 303-304.

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

- Shashidhara, S., Mitchell, D. J., Erez, Y., & Duncan, J. (2019). Progressive Recruitment of the Frontoparietal Multiple-demand System with Increased Task Complexity, Time Pressure, and Reward. *Journal of Cognitive Neuroscience*, *31*(11), 1617-1630.
- Sohoglu, E., & Davis, M. H. (2016). Perceptual learning of degraded speech by minimizing prediction error. *Proc Natl Acad Sci U S A*, *113*(12), E1747-1756.
- Sohoglu, E., Peelle, J. E., Carlyon, R. P., & Davis, M. H. (2014). Top-down influences of written text on perceived clarity of degraded speech. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(1), 186-199.
- Stacey, P. C., & Summerfield, A. Q. (2008). Comparison of Word-, Sentence-, and Phoneme-Based Training Strategies in Improving the Perception of Spectrally Distorted Speech. *Journal of Speech, Language, and Hearing Research*, *51*(2), 526-538.
- Sumbly, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *Journal of the Acoustical Society of America*, *26*, 212-215.
- Swaab, T. Y., Brown, C., & Hagoort, P. (1997). ERP evidence for delayed lexical selection in Broca's aphasia. *Journal of Psychophysiology*, *11*(4), 379-380.
- Swinney, D. A. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior*, *18*, 645-659.
- Swinney, D. A., Zurif, E., & Nicol, J. (1989). The Effects of Focal Brain Damage on Sentence Processing: an examination of the neurological organization of a mental module. *Journal of Cognitive Neuroscience*, *1*(1), 25-37.
- Tagarelli, K. M., Shattuck, K. F., Turkeltaub, P. E., & Ullman, M. T. (2019). Language learning in the adult brain: A neuroanatomical meta-analysis of lexical and grammatical learning. *Neuroimage*, *193*, 178-200.

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

- Tahmasebi, A. M., Davis, M. H., Wild, C. J., Rodd, J. M., Hakyemez, H., Abolmaesumi, P., et al. (2012). Is the link between anatomical structure and function equally strong at all cognitive levels of processing? *Cereb Cortex*, 22(7), 1593-1603.
- Tamminen, J., Payne, J. D., Stickgold, R., Wamsley, E. J., & Gaskell, M. G. (2010). Sleep Spindle Activity is Associated with the Integration of New Memories and Existing Knowledge. *Journal of Neuroscience*, 30(43), 14356-14360.
- Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences USA*, 94(26), 14792-14797.
- Turken, A. U., & Dronkers, N. F. (2011). The neural architecture of the language comprehension network: converging evidence from lesion and connectivity analyses. *Frontiers in Systems Neuroscience*, 5, 1.
- Vaden, K. I., Kuchinsky, S. E., Ahlstrom, J. B., Dubno, J. R., & Eckert, M. A. (2015). Cortical Activity Predicts Which Older Adults Recognize Speech in Noise and When. *Journal of Neuroscience*, 35(9), 3929-3937.
- Vaden, K. I., Kuchinsky, S. E., Cuta, S. L., Ahlstrom, J. B., Dubno, J. R., & Eckert, M. A. (2013). The cingulo-opercular network provides word-recognition benefit. *Journal of Neuroscience*, 33(48), 18979-18986.
- Van Berkum, J. J. (2009). The neuropragmatics of 'simple' utterance comprehension: An ERP review. In K. Y. U. Sauerland (Ed.), *Semantics and pragmatics: From experiment to theory* (pp. 276-316). Basingstoke: Palgrave Macmillan.
- Varley, R. A., Klessinger, N. J., Romanowski, C. A., & Siegal, M. (2005). Agrammatic but numerate. *Proc Natl Acad Sci U S A*, 102(9), 3519-3524.

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

- Varley, R. A., & Siegal, M. (2000). Evidence for cognition without grammar from causal reasoning and 'theory of mind' in an agrammatic aphasic patient. *Current Biology*, *10*(12), 723-726.
- Varley, R. A., Siegal, M., & Want, S. C. (2001). Severe impairment in grammar does not preclude theory of mind. *Neurocase*, *7*(6), 489-493.
- Vincent, J. L., Kahn, I., Snyder, A. Z., Raichle, M. E., & Buckner, R. L. (2008). Evidence for a frontoparietal control system revealed by intrinsic functional connectivity. *Journal of Neurophysiology*, *100*(6), 3328-3342.
- Vitello, S., Warren, J. E., Devlin, J. T., & Rodd, J. M. (2014). Roles of frontal and temporal regions in reinterpreting semantically ambiguous sentences. *Frontiers in Human Neuroscience*, *8*.
- Wehbe, L., Blank, I. A., Shain, C., Futrell, R., Levy, R., von der Malsburg, T., et al. (2021). Incremental language comprehension difficulty predicts activity in the language network but not the multiple demand network. *Cerebral Cortex*.
- Wild, C. J., Yusuf, A., Wilson, D. E., Peelle, J. E., Davis, M. H., & Johnsrude, I. S. (2012). Effortful listening: the processing of degraded speech depends critically on attention. *Journal of Neuroscience*, *32*(40), 14010-14021.
- Woolgar, A., Duncan, J., Manes, F., & Fedorenko, E. (2018). The multiple-demand system but not the language system supports fluid intelligence. *Nat Hum Behav*, *2*(3), 200-204.
- Woolgar, A., Parr, A., Cusack, R., Thompson, R., Nimmo-Smith, I., Torralva, T., et al. (2010). Fluid intelligence loss linked to restricted regions of damage within frontal and parietal cortex. *Proceedings of the National Academy of Sciences USA*, *107*(33), 14899-14902.

BRAIN NETWORKS FOR CHALLENGES TO SPEECH COMPREHENSION

Zempleni, M. Z., Renken, R., Hoeks, J. C. J., Hoogduin, J. M., & Stowe, L. A. (2007).

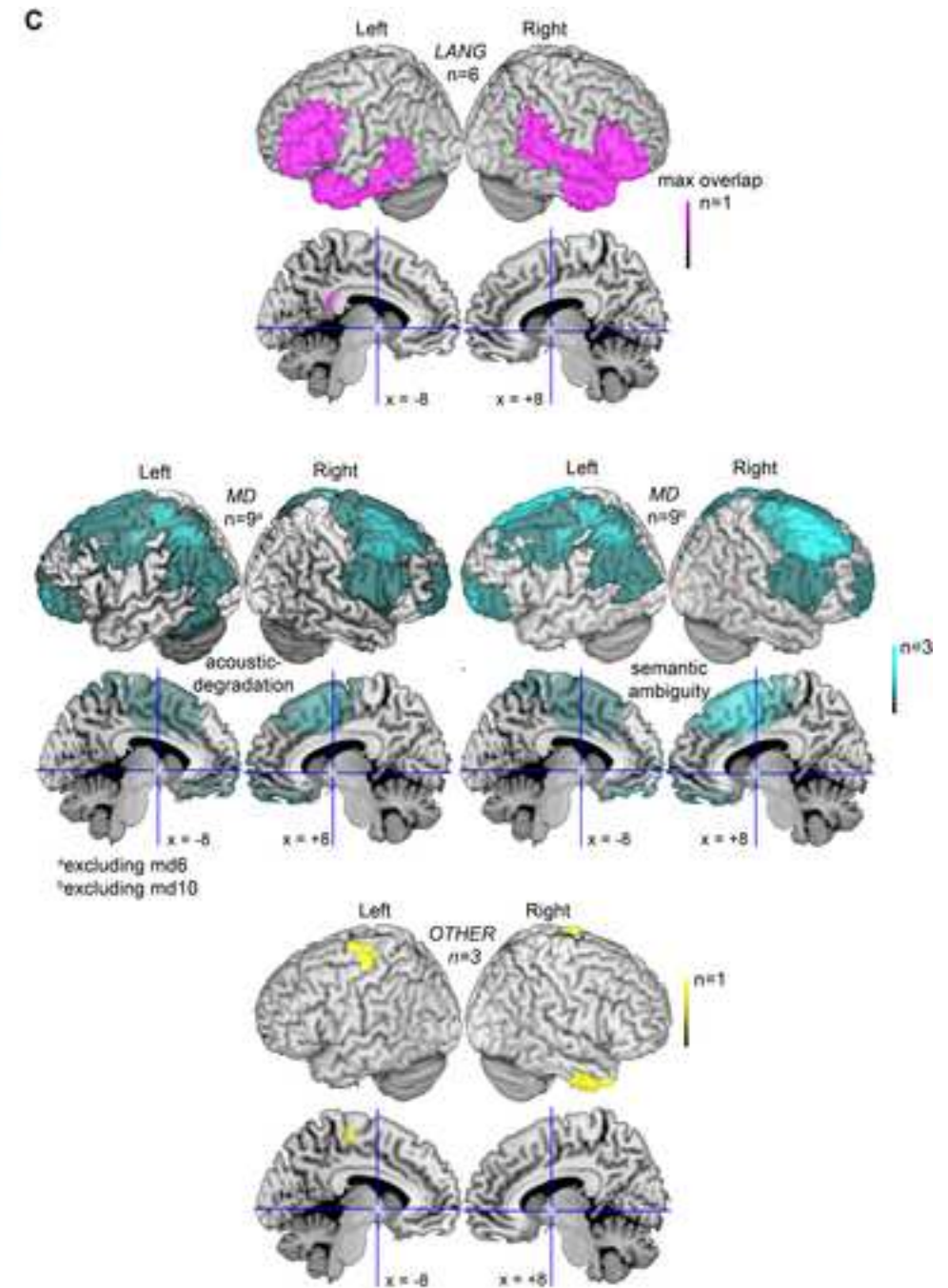
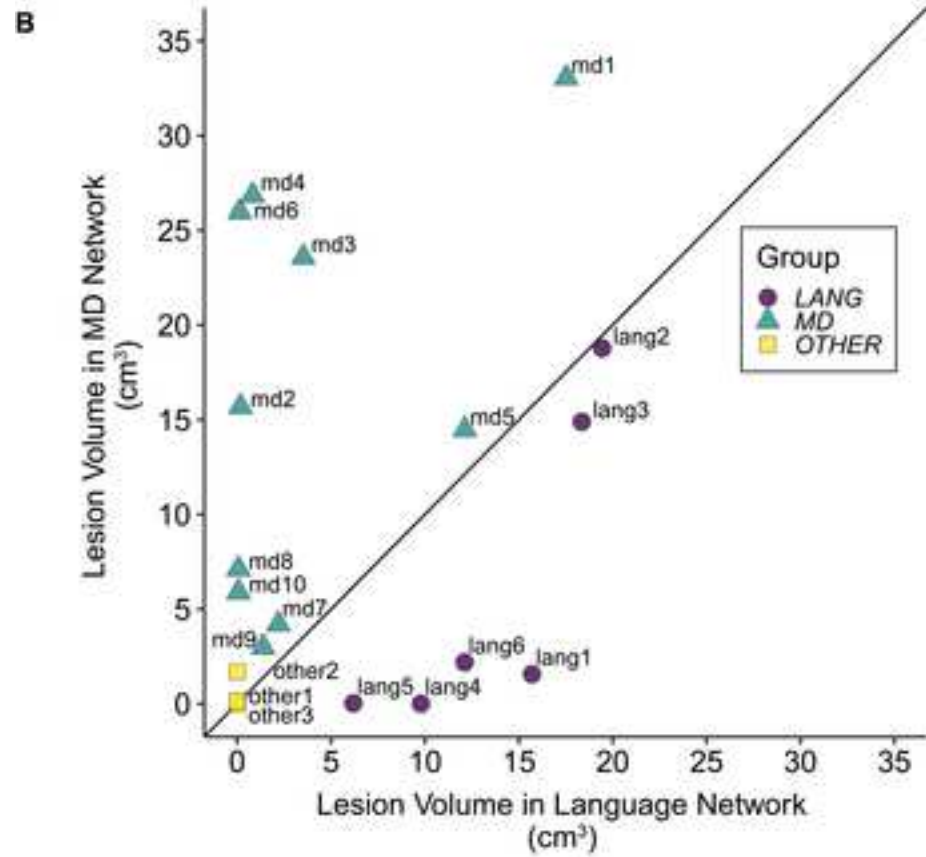
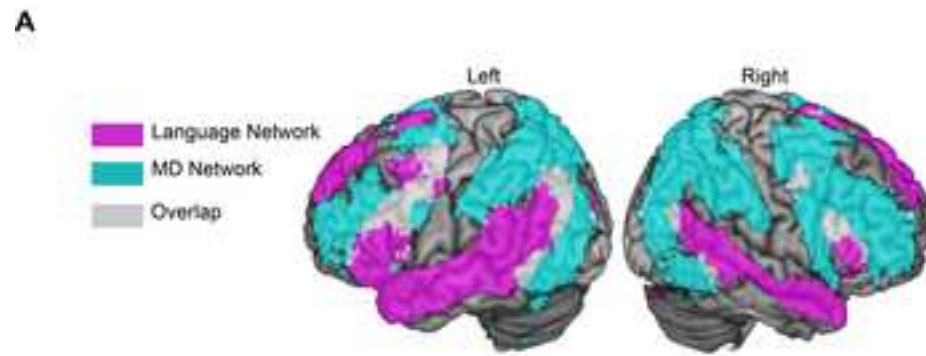
Semantic ambiguity processing in sentence context: Evidence from event-related fMRI. *Neuroimage*, 34(3), 1270-1279.

Zhang, Y., Frassinelli, D., Tuomainen, J., Skipper, J. I., & Vigliocco, G. (2021). More than

words: word predictability, prosody, gesture and mouth movements in natural

language comprehension. *Proceedings of the Royal Society B: Biological Sciences*,

288(1955), 20210500.



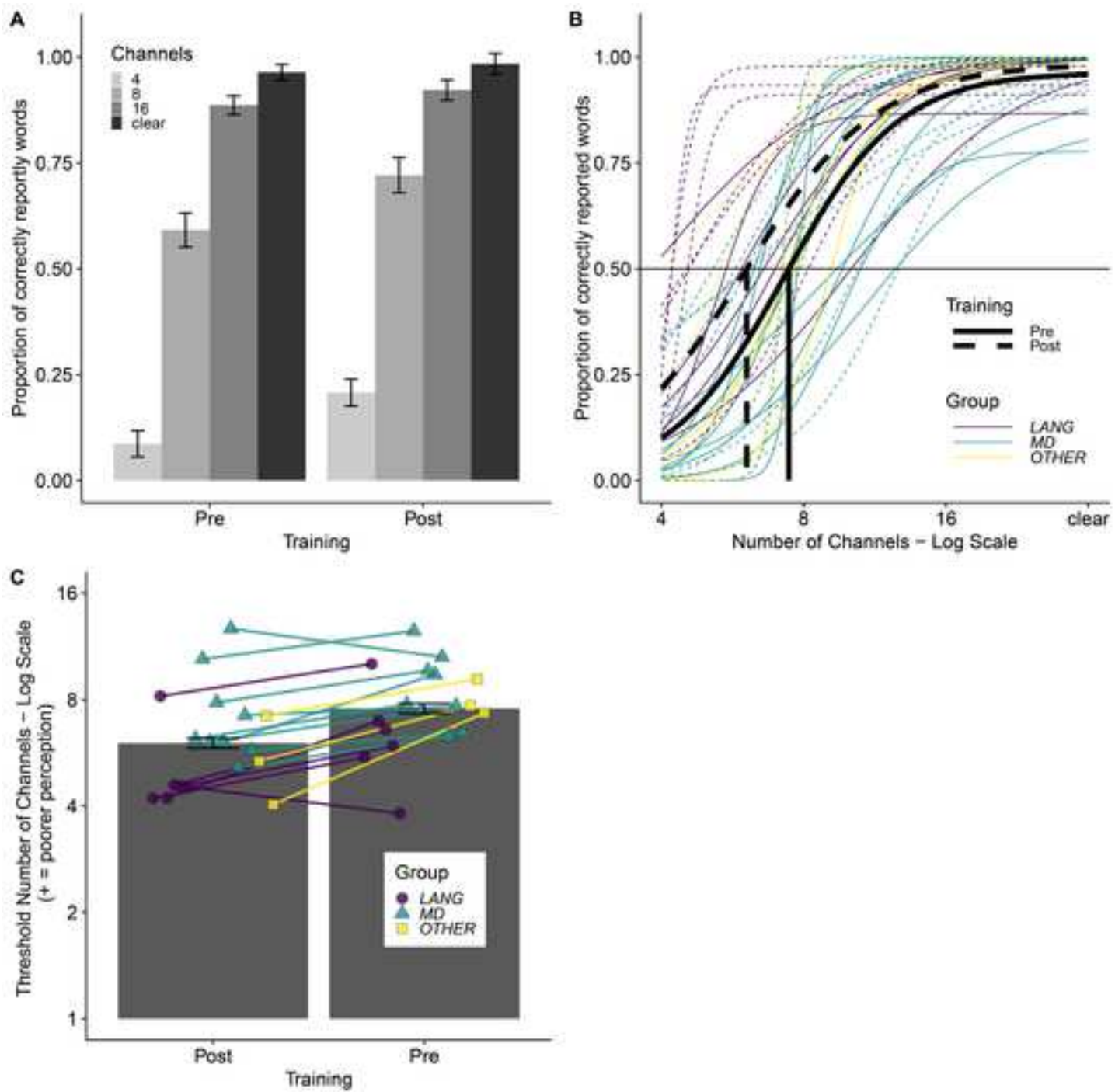
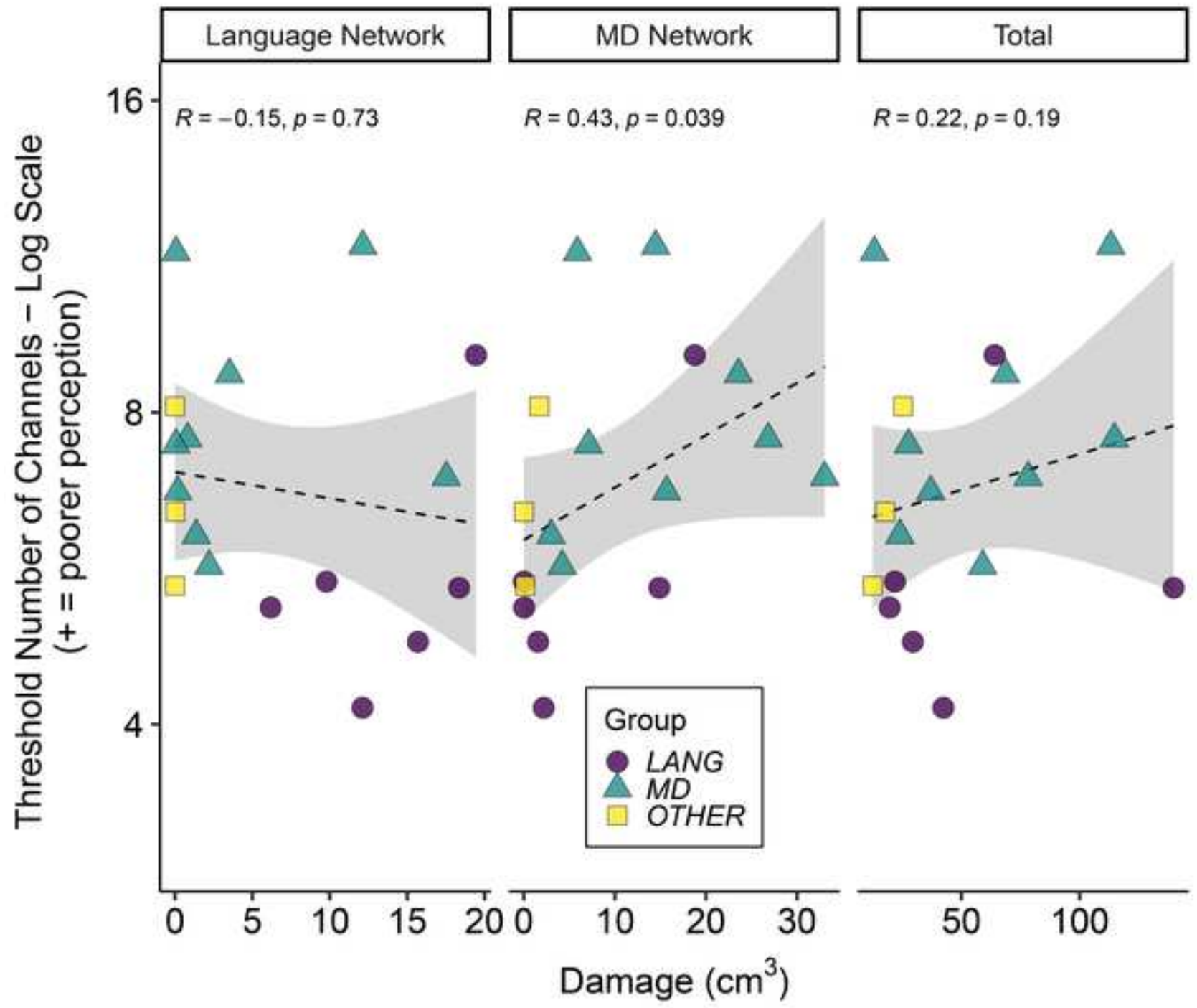
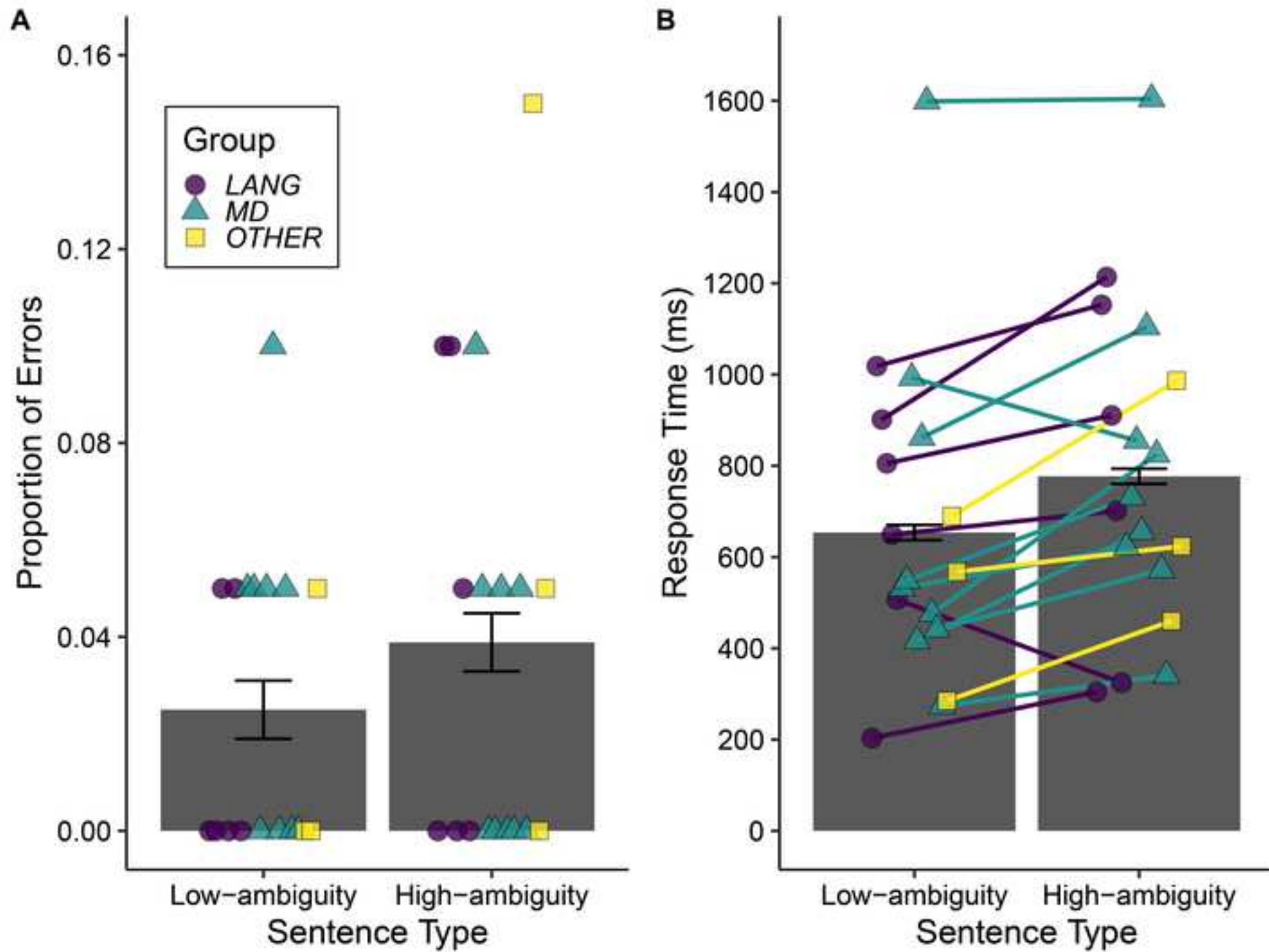


Figure 3

[Click here to access/download;Figure;Fig3-ScatterPlot_WordReport_Threshold_by_Damage.tif](#)





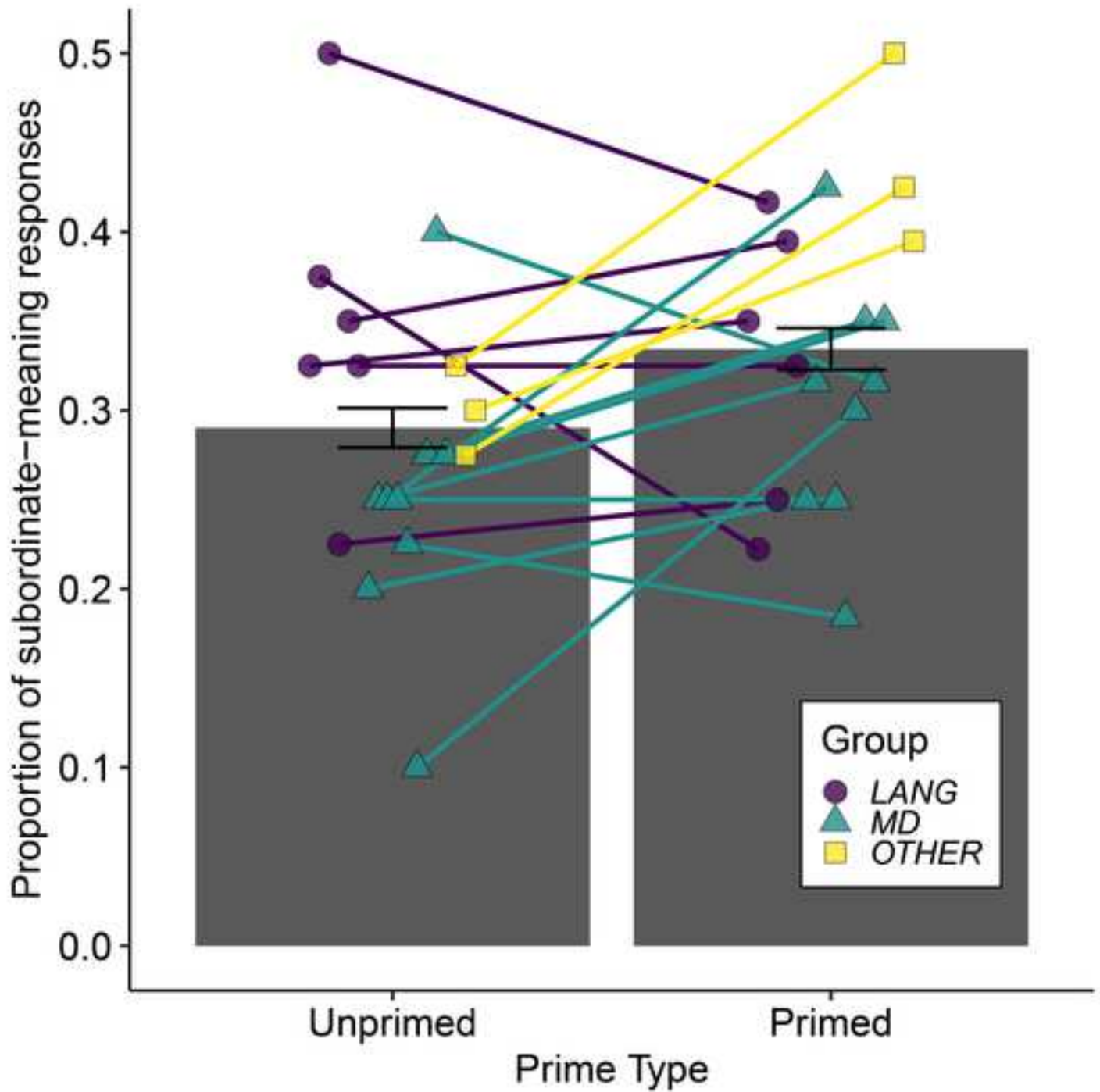


Figure 6

