



The Neural Basis of Metaphor Comprehension: Evidence from Left Hemisphere Degeneration

Nathaniel Klooster^{1,2}, Marguerite McQuire¹, Murray Grossman^{1,2},
Corey McMillan^{1,2}, Anjan Chatterjee^{1,3,4}, and Eileen Cardillo^{1,4}

¹Department of Neurology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

²Penn Frontotemporal Degeneration Center, University of Pennsylvania, Philadelphia, PA, USA

³Moss Rehabilitation Research Institute, Elkins Park, PA, USA

⁴Penn Center for Neuroaesthetics, University of Pennsylvania, Philadelphia, PA, USA

Keywords: figurative language, neurodegeneration, left hemisphere, LpMTG, LIFG

ABSTRACT

Despite the ubiquity of metaphor in cognition and communication, it is absent from standard clinical assessments of language, and the neural systems that support metaphor processing are debated. Previous research shows that patients with focal brain lesions can display selective impairments in processing metaphor, suggesting that figurative language abilities may be disproportionately vulnerable to brain injury. We hypothesized that metaphor processing is especially vulnerable to neurodegenerative disease, and that the left hemisphere is critical for normal metaphor processing. To evaluate these hypotheses, we tested metaphor comprehension in patients with left-hemisphere neurodegeneration, and in demographically matched healthy comparison participants. Stimuli consisted of moderately familiar metaphors and closely matched literal sentences sharing the same source term (e.g., *The interview was a painful crawl / The infant's motion was a crawl*). Written sentences were presented, followed by four modifier-noun answer choices (one target and three foils). Healthy controls, though reliably better at literal than metaphor trials, comprehended both sentence conditions well. By contrast, participants with left-hemisphere neurodegeneration performed disproportionately poorly on metaphor comprehension. Anatomical analyses show relationships between metaphor accuracy and patient atrophy in the left middle and superior temporal gyri, and the left inferior frontal gyrus, areas that have been implicated in supporting metaphor comprehension in previous imaging research. The behavioral results also suggest deficits of metaphor comprehension may be a sensitive measure of cognitive dysfunction in some forms of neurodegenerative disease.

INTRODUCTION

Metaphor, the description of one idea in terms of another, plays an important role in cognition and communication. The use of metaphor is pervasive, accounting for more than 13% of words in written discourse and more than 6% of spoken language (Steen et al., 2010). Metaphor strongly influences people's thinking in fundamental ways. The metaphors of "shaping" or "framing" are often used to describe the influence of metaphor on thinking (e.g., Lakoff & Johnson, 1980). Meta-analyses show that metaphorical language is more persuasive than comparable literal language (Sopory & Dillard, 2002; Van Stee, 2018). The presence or absence of metaphor when describing an issue affects how people think about topics such as cancer (e.g., Gibbs & Franks, 2002; Hauser & Schwarz, 2015), crime (Thibodeau & Boroditsky, 2011, 2013), significant personal relationships (Lee & Schwarz, 2014; Robins & Mayer, 2000), the acceptability of cognitive enhancement

Citation: Klooster, N., McQuire, M., Grossman, M., McMillan, C., Chatterjee, A., & Cardillo, E. (2020). The neural basis of metaphor comprehension: Evidence from left hemisphere degeneration. *Neurobiology of Language*, 1(4), 474–491. https://doi.org/10.1162/nol_a_00022

DOI:
https://doi.org/10.1162/nol_a_00022

Supporting Information:
https://doi.org/10.1162/nol_a_00022

Received: 26 April 2019
Accepted: 14 June 2020

Competing Interests: The authors have declared that no competing interests exist.

Corresponding Author:
Nathaniel Klooster
klooster@penmedicine.upenn.edu

Handling Editor:
Steven Small

Copyright: © 2020 Massachusetts Institute of Technology. Published under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

(Conrad, Humphries, & Chatterjee, 2019), the urgency of climate change (Flusberg, Matlock, & Thibodeau, 2017; Nerlich & Jaspal, 2012), and the brilliance of an idea (Elmore & Luna-Lucero, 2017).

Beyond influencing how people think, metaphor can affect how people act. Metaphor can drive attention (Matlock, 2004), affecting what information is attended to in social environments (Bowes & Katz, 2015). The choice of metaphor can affect how communities choose to handle public health issues (Barry, Brescoll, Brownell, & Schlesinger, 2009) and how patients are routed through health care facilities (Hilligoss, 2014). Because metaphorical language can be persuasive and can influence behavior, metaphors are widely used in didactic contexts (see the review in Saban, 2006), and in the teaching of medicine (Van Rijn-Van Tongeren, 1997), and in science specifically (Beger & Jäkel, 2015). Metaphor use is pervasive, it frames thinking, and it affects human behavior.

The neural bases of metaphor are debated. Early studies of metaphor comprehension focused on possible hemispheric differences, with some highlighting a privileged role for the right hemisphere (Bottini et al., 1994; Brownell, Simpson, Bihrlé, Potter, & Gardner, 1990; Winner & Gardner, 1977). The right hemisphere hypothesis for metaphor argues that the right hemisphere plays a specific and necessary role in supporting the ability to understand metaphor. The right hemisphere is thought to be especially important in the comprehension of novel metaphors (Bohrn, Altmann, & Jacobs, 2012; Mashal & Faust, 2009; Mashal, Faust, & Hendler, 2005).

Over time, however, the role of the left hemisphere in metaphor comprehension has begun to be considered more seriously. Functional imaging studies indicate left hemisphere regions support metaphor comprehension in addition to right hemisphere regions (Bambini, Gentili, Ricciardi, Bertinetto, & Pietrini 2011; Cardillo, Watson, Schmidt, Kranjec, & Chatterjee, 2012; Chen, Widick, & Chatterjee, 2008; Lee & Dapretto, 2006; Obert et al., 2014; Schmidt & Seger, 2009; Yang, Fuller, Khodaparast, & Krawczyk, 2010), and sometimes perhaps exclusively (Diaz, Barrett, & Hogstrom, 2011; Rapp, Leube, Erb, Grodd, & Kircher, 2004, 2007). Meta-analyses of the functional imaging literature also suggest that metaphor comprehension is a bilaterally mediated process (Bohrn et al., 2012; Rapp, Mutschler, & Erb, 2012; Reyes-Aguilar, Valles-Capetillo, & Giordano, 2018; Yang, 2014). These studies indicate that the bilateral inferior frontal gyri and the left posterior middle temporal gyrus (LpMTG) specifically are activated during metaphor comprehension.

Neuroimaging studies have their inferential limitations. They are limited in making claims of a causal role for brain regions underlying a cognitive function. Patient studies, while more difficult to conduct, offer an important constraint on theorizing based solely on imaging studies (Fellows et al., 2005). Studies with focal lesion patients find evidence of patients with selective metaphor deficits despite normal literal sentence comprehension following left-sided damage (Cardillo, McQuire, & Chatterjee, 2018; Gagnon, Goulet, Giroux, & Joannette, 2003; Ianni, Cardillo, McQuire, & Chatterjee, 2014; Tompkins, 1990). These studies implicate the left frontal and posterior temporal cortices in metaphor comprehension (Cardillo et al., 2018; Zaidel, Kasher, Soroker, & Batori, 2002). The traditional view that the right hemisphere plays a privileged role in metaphor comprehension is increasingly difficult to justify, though it remains a staple of clinical and basic neuroscience teaching.

The neural bases for metaphor comprehension are relevant to models of language processing, but also for adequately addressing the therapeutic needs of clinical populations. Studies of diverse patient populations indicate metaphor processing is frequently impacted by brain injury or disease. Following a traumatic brain injury, patients displayed disruptions with metaphor processing (Yang et al., 2010). Studies also report impairments in metaphor comprehension in

Parkinson's disease (Fernandino et al., 2013; Monetta & Pell, 2007) and Alzheimer's disease (Amanzio, Geminiani, Leotta, & Cappa, 2008; Papagno, 2001; Roncero & de Almeida, 2014; Winner & Gardner, 1977). Figurative language abilities more broadly are impaired in mild cognitive impairment (Cardoso, Silva, Maroco, de Mendonça, & Guerreiro, 2014) and in Alzheimer's disease (Papagno, Lucchelli, Muggia, & Rizzo, 2003). These studies raise the possibility that metaphor comprehension is an especially fragile linguistic ability. In contrast to stable, focal lesions, neurodegenerative diseases have neural consequences and cognitive impairments that change gradually over time. Metaphor processing is a complex cognitive ability that requires contributions from many subdomains of cognition. Degenerative damage in distinct neural systems may lead to a common cognitive impairment. For instance, the effects of small decrements in multiple aspects of cognition could aggregate to a failure in metaphor comprehension. The complexity of metaphor processing and its resulting fragility in the face of distributed brain damage raise the possibility that metaphor comprehension is a sensitive measure of cognitive dysfunction in neurodegenerative disease.

The Current Study

We hypothesized that metaphor processing relies on an intact left hemisphere and that metaphor comprehension is vulnerable to neurodegenerative disease. To evaluate these hypotheses, we tested metaphor comprehension with a rigorously controlled task in a group of patients with left hemisphere neurodegeneration (LHND), and in demographically matched healthy control (HC) participants. While many patient studies of metaphor processing are strictly behavioral, the current study related behavior to structural MRI. Patients first underwent structural MRI scans and later completed the metaphor task. Task performance was related to cortical thickness values. We predicted that patients would be especially impaired on metaphor trials (compared to HCs and compared to their own performance on literal trials) and that these impairments would be associated with patterns of left temporal lobe atrophy in the patient group. Based on results of functional imaging studies of metaphor processing using stimuli from the same set as used here (Cardillo et al., 2012), we assessed whether patterns of atrophy in three regions of interest (ROI) frequently engaged by metaphor tasks—the left inferior frontal gyrus (LIFG), the right inferior frontal gyrus (RIFG), and the LpMTG—would relate to metaphor task performance.

MATERIALS AND METHODS

Metaphor Task

Target sentences

Stimuli included 32 matched metaphor–literal sentence pairs taken from the published stimulus sets of Cardillo, Schmidt, Kranjec, & Chatterjee, 2010; Cardillo, Watson, & Chatterjee, 2016). All 64 sentences were in the form “The X was a Y,” where Y was the shared word or phrase in the literal–metaphor match (hereafter, *source term*). The source term of the sentence, Y, was either an entity noun (e.g., *The relay was a sprint race / The math test was an intelligence race*) or an event noun (*The interruption was a loud knock / His emails were an insistent knock*). Source terms also always included sensorimotor features—half auditory (e.g., *knock*) and half motion (e.g., *race*). See Table 1 for examples.

Stochastic Optimization of Stimuli software (Armstrong, Watson, & Plaut, 2012) was used to ensure that metaphor and literal sentences were matched on many of the published norms for the items: the number of characters, words, and content words; and the average frequency and concreteness of their content words; as well as their valence (% positive), an online

Table 1. Stimuli examples

Sentence	Type	Example	Target	Foil 1	Foil 2	Foil 3
Metaphor	Entity-Auditory	Her chores were a sad tune.	gloomy routine	funeral hymn	playful exercise	prison system
	Entity-Motion	The summer romance was a merry-go-round.	dizzying delight	amusement ride	serious punishment	brick fireplace
	Event-Auditory	The cool breeze was a lullaby.	calming weather	reassuring melody	uncomfortable temperature	cracked basin
	Event-Motion	The prize money was a lift.	financial assistance	small elevator	economic burden	delicious apple
Literal	Entity-Auditory	The jingle was a happy tune.	catchy song	loose change	radio static	flower vase
	Entity-Motion	The construction was a new merry-go-round.	colorful carousel	real estate	moldy dungeon	grocery cart
	Event-Auditory	The child's favorite was a lullaby.	soothing song	biological offspring	screaming match	cracked knuckles
	Event-Motion	The bed was a heavy lift.	weighty mattress	striped sofa	lightweight frame	full trashcan

measure of semantic processing difficulty (valence judgment reaction time), and familiarity (all p 's > 0.09). Values for these variables were taken from published norms (Cardillo et al., 2010; Cardillo et al., 2016). Overall, metaphors were moderately familiar (4.6 / 7; 1 = *very unfamiliar*, 7 = *very familiar*) and were well-understood (Interpretability $M = 0.91$, $SD = 0.09$). Metaphors were less imageable than their literal counterparts ($p < 0.01$) and, as intended, more figurative ($p < 0.01$). These differences and equivalences held true whether collapsing across variables of noninterest (modality, source term), or whether distinguishing items further by these characteristics. For item properties, see Supplementary Table 2 in the online supporting information located at https://www.mitpressjournals.org/doi/suppl/10.1162/nol_a_00022.

Answer choices

Each target sentence was accompanied by four possible answers, a correct target and three foils. Each answer choice consisted of a modifier (adjective or noun) and a noun. Foils for the metaphor stimuli (Table 1) were (1) the literal meaning of the sentence, (2) the opposite of the metaphorical meaning of the sentence, and (3) an unrelated answer. Foils for the literal sentences were (1) a category associate of the agent of the sentence not implied by the sentence, (2) the opposite of the literal meaning of the sentence, and (3) an unrelated answer. In this way, answer choices were designed to be informative of the nature of comprehension difficulty on incorrect trials. Answer choices were matched on average frequency, concreteness, and valence. For full materials used in this study see Supplementary Table 1 in the online supporting information.

Task procedure

For both groups, stimuli were presented visually on a laptop using E-Prime 2.0 software. On every trial, a sentence was presented at the top of the screen. When the participant indicated they were done reading the target sentence, the sentence remained on the screen and answer choices were presented below it, randomized to one of four quadrants in the lower half of the screen. Participants were instructed to choose the answer that best matched the meaning of the sentence and to guess if unsure.

Patients read the sentences and indicated to a researcher (orally or by pointing) which answer they thought best matched the meaning of the sentence. To limit demands on working memory for patients, the research assistant recorded the patient's answer and advanced the trial. HCs controlled the testing laptop and made their responses without assistance.

Piloting Process

Ten adults free of neurological disease or brain injury and with overall cognition and verbal intelligence within the normal range ($Mean = 27.9 \pm 2.0$, Mini-Mental State Examination [MMSE; Folstein, Folstein, & McHugh, 1975]; $Mean = 113.6 \pm 10.1$, American National Adult Reading Test [AMNART; Nelson & O'Connell, 1978]) were recruited from a database of healthy older adults to serve as pilot participants for the metaphor and literal multiple-choice stimuli. Participants were paid \$15/hr and gave informed consent for their participation in accordance with procedures of the University of Pennsylvania Institutional Review Board (protocol #806447). Pilot participants were native English speakers and were matched to the patient group on age ($M = 69.5 \pm 7.5$) and education ($M = 15.3 \pm 3.0$).

Pilot participants completed the task as described above. Item accuracy analysis revealed seven of the 64 items for which the correct answer was selected $\leq 60\%$ of the time. Based on the foils most commonly chosen for these items, target answers and/or foils were revised to disambiguate the answer choices. Target sentences were not altered. Patients and a new group of HCs were tested on these revised stimuli.

Participants

Healthy comparison participants

Nineteen HCs were recruited from a database of healthy older adults to participate in the study. HCs were paid \$15/hr and gave informed consent in accordance with procedures of the University of Pennsylvania Institutional Review Board (protocol #806447). HCs were native English speakers, matched to the patient group on age ($M = 65.2 \pm 10.2$) and education ($M = 15.4 \pm 2.6$). Neuropsychological testing confirmed they scored within the normal range on the MMSE ($M = 28.8 \pm 1.6$), indicating normal cognition, and on the AMNART ($M = 117.83 \pm 8.20$), indicating normal verbal intelligence.

Patients

Thirteen patients were recruited from a Frontotemporal Degeneration Center to participate in the study. All patients were diagnosed with logopenic-variant primary progressive aphasia (lvPPA) at the time of testing, according to established criteria (Gorno-Tempini et al., 2011), and confirmed through a consensus process. Since this initial diagnosis, on reassessment three were diagnosed with progressive supranuclear palsy, one with Alzheimer's disease, one with behavioral-variant fronto-temporal dementia, and one with nonfluent PPA. These were clinical diagnoses based only on phenotype. The remaining patients retained a lvPPA diagnosis at the time of publication. Patients and HC participants were matched on age and education. Patients displayed mild impairment on the MMSE ($M = 24.62$, Tables 2 and 3) and as a group performed significantly worse than HCs (Table 2). Patients were paid \$15/hr and gave informed consent in accordance with procedures of the University of Pennsylvania Institutional Review Board (protocol #806447).

Table 2. Demographics

Group	Age	Chronicity	Education	MMSE
LHND ($n = 13$)	63.31 (± 6.9)	3.69 (± 1.49)	15.62 (± 2.9)	24.62 (± 4.5)**
HC ($n = 19$)	65.21 (± 10.2)		15.39 (± 2.6)	28.58 (± 1.6)**

Note. ** HC MMSE > LHND ($p < 0.01$). LHND = left-hemisphere neurodegeneration, HC = healthy control, MMSE = Mini-Mental State Examination.

Anatomical Methods

T1 Whole-brain imaging

High-resolution T1-weighted MPRAGE structural scans were acquired for all but one patient (excluded due to claustrophobia and anxiety even following sedation) and 35 healthy controls comparable to the patient group (Age: $M = 67.37$, $p < 0.15$; Education: $M = 16.03$, $p < 0.65$; 10 males). MRI data was acquired on a 3T Siemens Tim Trio scanner with an 8-channel head coil, with $T = 1,620$ ms, $T_2 = 3.09$ ms, flip angle = 15° , 192×256 matrix, and 1 mm^3 voxels. T1-weighted MRI images were then preprocessed to compute cortical thickness using Advanced Normalization Tools (ANTs; Tustison et al., 2014). Briefly, each individual dataset was deformed using a symmetric and diffeomorphic registration routine using ANTs to register each volume to a standard local template space in a canonical stereotactic coordinate system. ANTs provide a highly accurate registration routine using symmetric and topology-preserving diffeomorphic deformations to minimize bias toward the reference space and to capture the deformation necessary to aggregate images in a common space. Then, we used N4 bias correction to minimize heterogeneity (Tustison et al., 2010) and the ANTs Atropos tool to segment images into six tissue classes (cortex, white matter, cerebrospinal fluid, subcortical grey structures, brainstem, and cerebellum) using template-based priors, and to generate probability maps of each tissue. Voxel-wise cortical thickness was measured in millimeters (mm) from the pial surface and then transformed into Montreal Neurological Institute (MNI) space, smoothed using a 2 sigma full-width half-maximum Gaussian kernel, and downsampled to 2 mm isotropic voxels.

We conducted several types of anatomic analyses. The first targeted specific ROIs based on activation patterns in previous fMRI studies of metaphor comprehension. The second examined single-subject atrophy patterns in a subgroup of patients who show metaphor impairment. Exploratory supplemental analyses probed brain-behavior relationships of task performance within areas of degeneration as identified by a group mask.

To characterize brain areas where the patients as a group displayed significant neural degeneration relative to matched controls, a patient atrophy mask (Figure 1 and Supplementary Table 3 in the online supporting information) was calculated through nonparametric permutation-based analyses with threshold-free cluster enhancement (TFCE; Smith & Nichols, 2009) with the randomize tool in FSL (<http://fsl.fmrib.ox.ac.uk/fsl/fslwiki>). Cortical thickness was compared across groups. To depict atrophic areas that show maximal overlap in the group, clusters that meet a conservative threshold of $p < 0.005$ (family-wise error corrected with TFCE), and contain a minimum of 200 adjacent voxels are reported.

When a more liberal threshold of $p < 0.05$ is applied, areas of variability driven by smaller subsets of participants are evident (Supplementary Figure 1 in the online supporting information). Patients varied widely in the extent of atrophy present in regions outside the left temporal lobe.

To specifically probe ROIs implicated as critical nodes in the metaphor comprehension network, individual variability in patient performance was related to variability in cortical thickness in three ROIs: the LIFG, the RIFG, and the LpTMG temporal gyrus. ROIs were generated by creating 10 mm

Table 3. Patient neuropsychological profile

Subject	LHND1	LHND2	LHND3	LHND4	LHND5	LHND6	LHND7	LHND8	LHND9	LHND10	LHND11	LHND12	LHND13	Mean
PVLT Recall	9	8	3	5	4	4	2	0	8	3	8	8	0	4.77
PPT Words	24	25	25	25	21	16	25	23	23	25	13	22	24	22.38
PPT Pictures	26	26	24	25	23	19	26	22	24	26	24	24	24	24.10
Naming	93.33	86.67	90	40	60	36.67	81.25	18.75	93.33	93.75	96.88	78.13	–	57.9
Animal Fluency	19	19	10	11	1	6	11	5	13	10	16	7	6	10.31
“f” fluency	13	15	7	7	3	10	7	5	4	17	9	4	4	8.08
Trails A Time	31	35	46	45	–	89	37	30	48	130	34	54	66	126.23
Trails B Time	80	72	300	130	–	282	211	166	228	300	182	181	–	193.82
PVLT1	5	4	1	1	3	1	1	0	4	4	4	5	0	2.54
Digit Span Forward	3	3	5	5	5	5	1	2	2	4	4	10	2	3.92
Digit Span Backward	5	5	4	3	0	6	2	2	3	4	4	7	1	3.54
MMSE	29	27	23	28	25	19	21	26	28	20	15	29	30	24.61

Note. LHND = left-hemisphere neurodegeneration, PVLT = Philadelphia Verbal Learning Test (Libon, Mattson, Glosser, & Kaplan, 1996), PPT = Pyramids and Palm Trees (Howard & Patterson, 1992), MMSE = Mini-Mental State Exam.

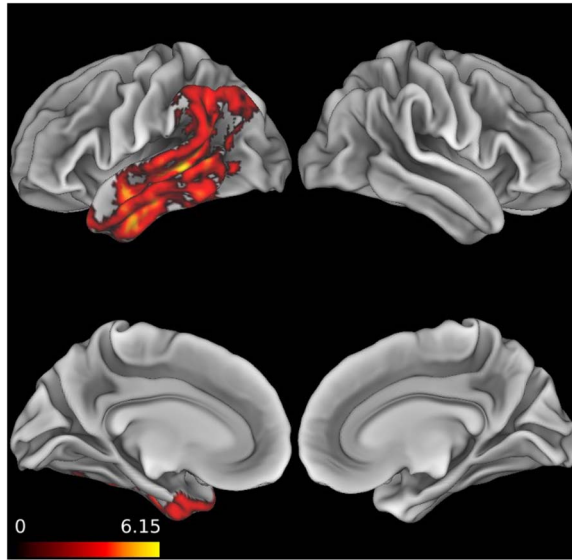


Figure 1. Patient whole brain atrophy. Pattern of cortical thinning in patients compared to healthy comparisons (significant at $p < 0.005$, family-wise error corrected with threshold-free cluster enhancement).

radius spheres around the peak coordinates in MNI space reported for each of these clusters in an fMRI study of metaphor comprehension using similar stimuli (LIFG = $-50, 29, -1$; RIFG = $50, 26, 5$; LpMTG = $-62, -50, -9$; Cardillo et al., 2012). Performance across literal and metaphor conditions was related to cortical thickness in these ROIs while controlling for age, sex, and years of education.

To characterize individual subject patterns of reduced cortical thickness in three individuals from the metaphor-impaired subgroup (see Neuroanatomical Results), individualized heatmaps of Z-transformed cortical thickness relative to 156 demographically comparable healthy adults with a self-reported negative psychiatric and neurological history were generated. These HCs were additionally screened as cognitively normal using a >27 (out of 30) score on the MMSE. To generate these heatmaps, the mean and *SD* for each voxel in template space for the control cohort were calculated. Then, Z scores of each patient's whole brain cortical thickness map relative to the HC means and *SD*s were generated. Individual heatmaps were then masked by the patient atrophy mask and can be interpreted as an LHND-specific Z map of age, sex, and education appropriate patterns of reduced cortical thickness.

Supplemental exploratory analyses relate task performance to cortical atrophy in the patient group as a whole. The randomize tool in the FMRIB Software Library was used to run regression analyses between the behavior of interest and patient cortical thickness values. Nonparametric permutations ($n = 10,000$) were run for each score of interest. Clusters that met a height threshold of $p < 0.05$ uncorrected with TFCE and a minimum of 25 adjacent voxels are reported. A 25-voxel threshold was chosen.

Statistical Methods

Linear mixed-effects models (LMEMs) were used to analyze the influence of variables of interest on metaphor task performance. The lme4 package (Bates, Maechler, Bolker, & Walker, 2015) in R (Version 3.3.1) was used for the following analyses. A theoretically motivated model was built for each analysis.

RESULTS

Behavioral Results

Analysis of HC performance revealed 1 of the 19 participants performed more than 2.5 *SDs* below the group mean (overall accuracy = 42.2% correct; 46.9% literal, 37.5% metaphor). This participant's data was removed from further analysis.

HC participants responded to two of the items with low accuracy, indicating that these items were statistical outliers (for both items, only 8/18 participants answered correctly placing these items more than 2.5 *SDs* below the other items' mean accuracy). These two items were removed from further analysis. The inclusion of these items, or of the outlier participant described above, did not affect the patterns of results described below.

To examine group differences, LMEMs were used to analyze the relationships between group (patients, HC), figurativeness (literal, metaphor), and accuracy. As random effects, participant and item were included as intercepts, with a by-subject random slope included for the effect of figurativeness and a by-group random slope included for item. These analyses (Tables 4 and 5 and Figure 2) revealed a fixed effect of figurativeness, with literal sentences receiving more accurate responses than metaphors ($p < 0.05$, Cohen's $d = 0.23$), a fixed effect of group, with HCs outperforming LHND patients ($p < 0.01$, $d = 0.48$), and an interaction between group and figurativeness with the LHND group performing especially poorly on the metaphor stimuli ($p < 0.05$, $d = 0.34$). The difference between auditory (75.96%) and motion (79.55%) accuracy was not significant. There were no interactions of modality with group or figurativeness. There was no significant difference between event (80.00%) and entity (76.88%) accuracy, and there were no interactions of source with group or figurativeness.

Error analysis

To illuminate the nature of comprehension failures, the proportion of each foil type selected on incorrect trials was calculated. For metaphor trials, HCs (87%) and LHND (78%) showed a strong literal bias, most often choosing the foil that provided the literal meaning to the target metaphor rather than the correct metaphorical meaning. For literal trials, HCs (48%, 48%) and LHND (46% vs. 38%) showed an even split between the semantic associate and the opposite meaning foils.

Patients

Single case analyses

To examine different patterns of impaired comprehension, individual patient performance was compared to the HC group using LMEMs. Subgroups based on behavior were then examined for different patterns of atrophy. Accuracy was modeled as a function of group identity (patient or HC), figurativeness (literal vs. metaphor trials), and education, and an interaction between

Table 4. Percent accuracy by group and figurativeness

	Literal	Metaphor	Mean
LHND ($n = 13$)	75.0	50.51	63.90
HC ($n = 18$)	93.58	84.26	89.10
Mean	85.79	70.75	

Note. LHND = left-hemisphere neurodegeneration, HC = healthy control.

Table 5. Fixed effects

Parameter	Estimate	Standard Error	<i>t</i> value
Intercept	0.533194	0.125328	4.254***
Group (LHND)	-0.189489	0.054223	-3.495**
Education	0.026019	0.007747	3.359**
Figurativeness (Metaphor)	-0.093171	0.040388	-2.307*
Group (LHND) × Figurativeness (Metaphor)	-0.136316	0.057948	-2.352*

Note. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. LHND = left-hemisphere neurodegeneration.

figurativeness and group, with the random effects of subject and item. The Holm method was used to correct for multiple comparisons.

These analyses revealed that four patients (LHND2, LHND3, LHND4, LHND11) performed indistinguishably from HCs overall and comprehended the literal and metaphor stimuli alike (p 's > 0.05). Five patients displayed a general sentence-comprehension impairment (LHND5, LHND6, LHND8, LHND10, LHND13), with impaired performance on the task overall and on the literal stimuli (all p 's < 0.0001). For four patients, there was an interaction between group (patient or HC) and figurativeness (all p 's < 0.05), showing evidence for a disproportionate metaphor impairment, with worse metaphor comprehension performance than predicted by their literal sentence performance (LHND1, LHND7, LHND9, LHND12; see Table 6). LHND7 could not complete scanning due to anxiety. While all were diagnosed with lvPPA at the time

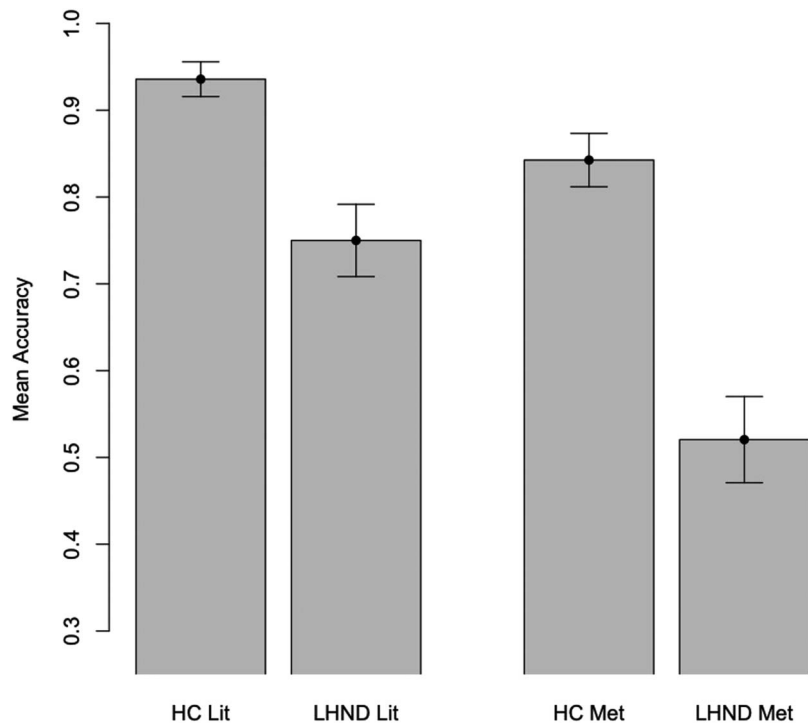


Figure 2. Metaphor task performance. Accuracy on literal (Lit) and metaphor (Met) trials by group showing significant effects of group (HC > LHND, $p < 0.01$), figurativeness (literal accuracy > metaphor accuracy, $p < 0.05$), and their interaction (with LHND especially impaired on metaphor accuracy, $p < 0.05$). LHND = left-hemisphere neurodegeneration, HC = healthy control.

Table 6. Single cases with metaphor impairment

Patient	Literal (32 max)	Literal <i>t</i> value	Literal <i>p</i> value	Metaphor (30 max)	Metaphor <i>t</i> value	Metaphor <i>p</i> value	Cohen's <i>d</i>
LHND1	30	0.03	0.488	17	-1.77	0.048	0.894
LHND7	30	0.03	0.488	17	-1.77	0.048	0.894
LHND9	26	-2.14	0.024	13	-2.62	0.009	0.909
LHND12	28	-1.05	0.154	16	-1.98	0.032	0.794

Note. *p*-values are one-tailed and compared to the control group's performance using the Crawford-Howell test (Crawford & Howell, 1998) for case-control comparisons. They provide a point estimate of the abnormality of each patient's score. Cohen's *d* values are the effect sizes for the differences between conditions for each patient. LHND = left-hemisphere neurodegeneration.

of testing, LHND7 was subsequently diagnosed with Alzheimer's disease, while the other patients retained a diagnosis of lvPPA at the time of publication.

Neuropsychological measures and performance

In the patient group, task performance was related to performance on standard neuropsychological assessments (Table 3). Item accuracy was modeled as a function of the fixed effects of figurativeness and education, with random intercepts of subject and item and a by-subject random slope for figurativeness. As the current study is not powered to consider all the neuropsychological tests' influence on task performance in the same model, the neuropsychological tests were examined one at a time for their ability to predict task performance by evaluating whether their inclusion improved model fit.

These analyses revealed that while many tests were related to performance on the literal sentences, including MMSE ($p < 0.05$, $d = 0.0695$), reverse Digit Span ($p < 0.05$, $d = 0.1505$), Naming ($p < 0.05$, $d = 0.2314$), Animals ($p < 0.005$, $d = 0.0790$), Philadelphia Verbal Learning Test (Libon, Mattson, Glosser, & Kaplan, 1996) recall ($p < 0.05$, $d = 0.1005$), Pyramids and Palm Trees (PPT; Howard & Patterson, 1992) picture ($p < 0.05$, $d = 0.1408$), PPT word ($p < 0.001$, $d = 0.1165$), and Complex Figure Test copy ($p < 0.01$, $d = 0.1141$), only lexical fluency was related to metaphor accuracy ($p < 0.01$, $d = 0.0570$).

Neuroanatomical Results

ROI analyses

Relationships between cortical thickness and literal and metaphor performance were investigated in three ROIs (LIFG, RIFG, LpMTG) motivated by previous fMRI results indicating that these regions are engaged in comprehension of metaphors from the same stimulus set as used in this study. Patient cortical thickness did not differ from that of HCs in the RIFG ($p > 0.91$, $d = 0.036$), while the LIFG showed a trend and a much larger effect size ($p < 0.11$, $d = 0.591$). Cortical thickness in HCs did not differ between the RIFG and the LIFG ($p < 0.37$), while there was a significant difference in these areas in the patient group ($t = 6.48$, $p < 0.0001$, $d = 0.322$), with the LIFG showing reduced thickness compared to the RIFG. No relationship was seen between overall performance or literal performance and integrity of these areas in the patients (p 's > 0.7). On metaphor comprehension, LIFG thickness was significantly associated with accuracy ($p < 0.05$, $d = 0.368$), while RIFG thickness showed a trend ($p > 0.18$).

Patients showed significant atrophy relative to HCs in the LpMTG ($p < 0.01$, $d = 1.174$). For overall task performance, the LpMTG showed a trend toward being associated with accuracy

($p < 0.098$). Atrophy in the LpMTG did not relate to literal accuracy ($p > 0.3$). On metaphor trials (Figure 3), LpMTG thickness was significantly related to accuracy ($p < 0.01$, $d = 0.730$).

Single subject analyses

Figure 4 and Supplementary Table 3 in the online supporting information depict single subject atrophy heatmaps for the three participants with a disproportionate metaphor deficit and available neuroimaging.

Supplemental analyses: VBM within areas of degeneration

Exploratory voxel-based morphometry regressions were run relating patient performance on literal and metaphor trials to the cortical thickness values within their atrophy mask (Figure 1). For literal stimuli (Supplementary Figure 2 and Supplementary Table 4), atrophy in the superior temporal gyrus and the angular gyrus related to poorer accuracy. For metaphor stimuli (Supplementary Figure 3 and Supplementary Table 4), atrophy in the fusiform gyrus and the middle temporal gyrus related to poorer accuracy.

DISCUSSION

Metaphor is used pervasively in communication and in cognition, shaping thought and influencing behavior. How the brain mediates comprehension of metaphor is debated. To investigate the neural bases of metaphor comprehension, we tested patients with neurodegenerative disease affecting the left hemisphere and a group of demographically matched HC participants on their ability to resolve the meaning of metaphoric and matched literal sentences. To summarize our results before discussing them in detail, our patients performed poorly on metaphor trials compared to HCs and compared

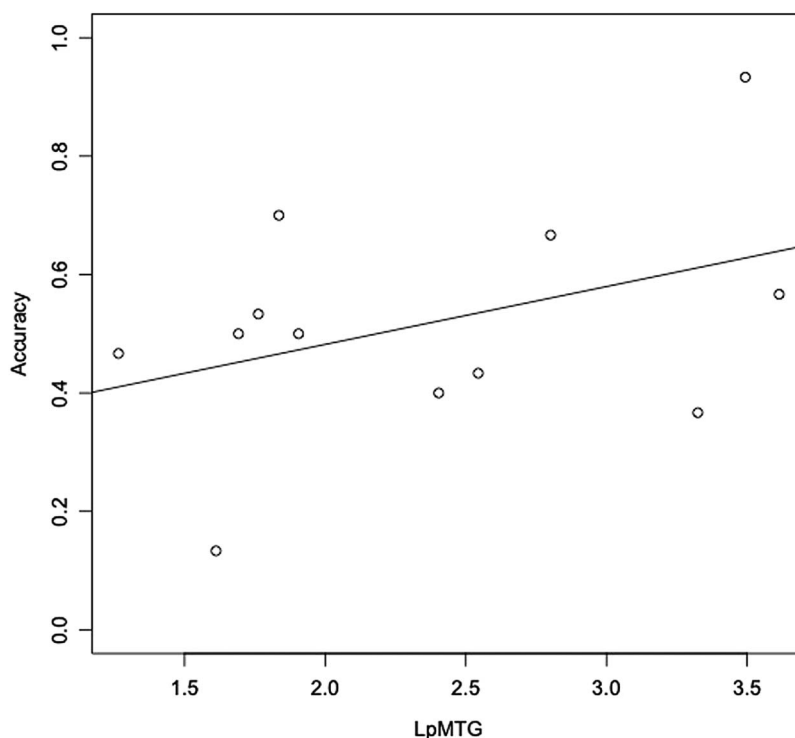


Figure 3. Metaphor trial accuracy in patients as a function of LpMTG cortical thickness (mm). LpMTG = left posterior middle temporal gyrus.

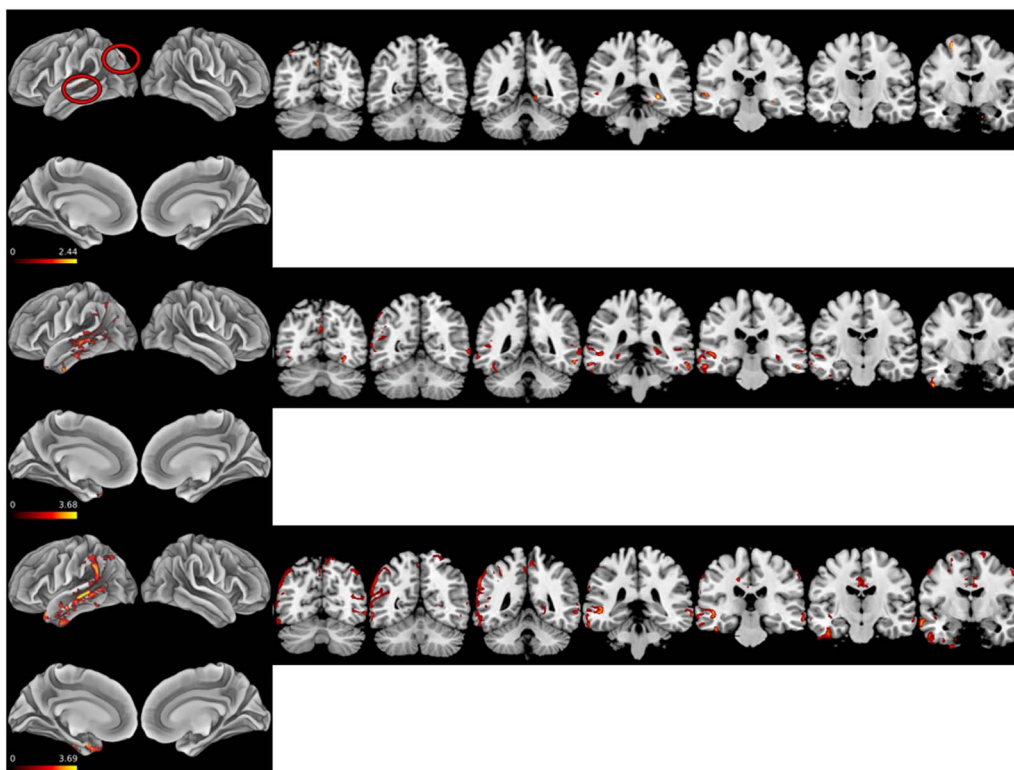


Figure 4. Z score heat maps for three patients with disproportionate metaphor impairment. Top: LHND1 Middle: LHND9. Bottom: LHND12. LHND = left-hemisphere neurodegeneration.

to their own performance on literal trials. These results point to the important role that the left hemisphere plays in understanding metaphor. Thinning of LIFG, left middle temporal gyrus, and left superior temporal gyrus, areas implicated in previous work as supporting metaphor comprehension, was associated with poor metaphor performance in some of our patients here. The results also show that metaphor processing can be disproportionately impaired compared to literal sentence comprehension, suggesting that metaphor comprehension deficits could be a sensitive measure of cognitive change in neurodegenerative disease, revealing impairments before literal language is impacted.

The variability in previously proposed neural substrates for metaphor comprehension present in the literature may have risen from relative lack of control of psycholinguistic properties of sentences used to test brain-behavior relationships (Cardillo et al., 2010; Citron & Goldberg, 2014; Schmidt, Kranjec, Cardillo, & Chatterjee, 2010). By measuring and balancing literal and metaphor stimuli on various properties impacting comprehension difficulty—number of characters, words, and content words, average frequency and concreteness of content words, and familiarity, valence, and a measure of semantic processing speed of sentences—the current study attempted to avoid these difficulties and provide a more balanced contrast between literal and metaphor sentences.

The patients' impaired metaphor task performance cannot be explained as following from general cognitive impairment. While analyses relating patient task performance to their neuropsychological test performance are exploratory given the relatively small sample size, patients' neuropsychological profiles were mostly not associated with metaphor deficits. The patients' performance on literal trials related to many neuropsychological test scores, including tests of

overall cognition, working memory, episodic memory, and semantic memory. It makes sense that patients struggling to understand simple literal sentences would display impairments in multiple cognitive domains. The lack of clear relationships observed here between performance on metaphor trials and other aspects of cognition as measured through psychometric testing has been noted in other patient studies of metaphor (Amanzio et al., 2008). This observation affirms that traditional neuropsychological tests do not adequately capture metaphoric language deficits. We agree with the suggestion (Rapp & Wild, 2011) that nonliteral language should be included in standard neuropsychological assessment batteries.

The observation raises the question of why the patients struggle on metaphor trials. One possibility is that the inhibitory demands of metaphor comprehension are difficult for patients. To resolve the meaning of a metaphor, the literal sense of the sentence or features of the source term must be inhibited in order to identify the correct figurative sense of the word and sentence (Gernsbacher & Robertson, 1999; Papagno, 2001). The error analysis suggests that this inhibition is difficult for healthy participants. On incorrect trials, the literal sense was most often endorsed by HCs. For patients too, the literal foil was most often chosen on incorrect trials. The patients failed more often here than HCs, suggesting greater difficulty with the inhibitory demands of metaphor trials. A failure of inhibitory control also accords with the observation that lexical fluency—an index of executive function—was correlated in this population with metaphor comprehension accuracy. Successfully resolving the meaning of a metaphorical sentence also requires greater flexibility as features and properties of one domain are applied and compared to another. Cognitive flexibility is often diminished with general atrophy (Eslinger, Moore, Anderson, & Grossman, 2011; Kehagia, Barker, & Robbins, 2010; Swartz, Stuss, Gao, & Black, 2008) and this association may be contributing to the patients' deficits.

Our results provide important evidence for left-hemisphere mediation of metaphor. In this study, patients with neurodegeneration restricted to the left hemisphere displayed deficits in understanding moderately familiar metaphors. Functional imaging evidence also implicates areas in the left hemisphere used by the healthy brain in metaphor comprehension (Bohrn et al., 2012; Rapp et al., 2012; Yang, 2014). The current study and other patient studies (Cardillo et al., 2018; Gagnon et al., 2003; Ianni et al., 2014; Tompkins, 1990) provide evidence for the importance of left-hemisphere regions for normal metaphor comprehension. Our anatomic analyses included a targeted ROI analysis, an analysis within areas of group degeneration, and a subgroup brain-behavior analysis of patients based on their patterns of performance. For the ROI analysis, metaphor comprehension related to thickness of LpMTG and LIFG.

The left temporal lobe and LIFG are linked to semantic ambiguity resolution in literal language (Zempleni, Renken, Hoeks, Hoogduin, & Stowe, 2007; Davis et al., 2007). When words have multiple senses, or when multiple concepts can resolve the meaning of an ambiguous sentence, these left hemisphere structures are implicated in successful meaning resolution. This ability shares similarities to the challenge of resolving the meaning of a metaphor, a sentence with multiple possible meanings to be resolved, and indeed similar brain regions are implicated in supporting these abilities.

The LpMTG and the LIFG were related to metaphor comprehension deficits, consistent with previous neuroimaging studies of metaphor (Cardillo et al., 2012; Chen et al., 2008; Lee & Dapretto, 2006; Schmidt & Seger, 2009; Yang et al., 2010) and suggesting an important role for these regions. While such a relationship was absent in RIFG, this may be because the current patient sample was not suitable for testing its role, as the patients show atrophy in LIFG but not RIFG. The LpMTG has been linked to demands of semantic processing more generally (e.g., Noonan, Jefferies, Visser, & Lambon Ralph, 2013).

The single-subject atrophy maps document the relationship between the left temporal lobe and metaphor trial accuracy broadly. All three show relationships between LpMTG integrity and metaphor trial performance. Interestingly, all three display a relationship between medial temporal lobe atrophy and metaphor performance. Metaphor task performance may be a sensitive measure of cognitive decline in neurodegenerative diseases that affect the medial temporal lobe such as Alzheimer's disease.

When considering anatomy and neuropsychological profile and their relation to successful metaphor comprehension, there may be "many routes to failure." As the network of brain regions that support metaphor processing is delineated, it is likely that disruptions to any nodes of this network, or the connections between them, could lead to impairment. Metaphor processing is a complex cognitive ability that depends on many subdomains of cognition including semantic memory, working memory, (semantic) executive demands, inhibition, abstract thinking, and cognitive flexibility. It is possible that subtle disruptions to any of these subdomains can lead to impaired metaphor comprehension.

Limitations of the current study include the sample sizes of the participant groups. The current results should be replicated with larger groups of participants. A priori power analyses were not completed. We tested as many patients as possible. Given the relatively small sample size in the current study and the difficulty of recruiting such patients, we elected to limit our ROI analyses to a few strongly motivated areas. Analyses relating patient task performance to neuropsychological assessment should be examined in larger populations. HC participants were closely matched to the patients demographically, and their behavioral data was normally distributed, but a larger comparison group would strengthen the confidence in our findings. Despite these limitations, robust group differences were observed, and strong relationships were detailed between anatomy and metaphor task performance.

The current study provides evidence that metaphor processing can be disproportionately impaired compared to literal sentence comprehension. The evaluation of metaphor processing may provide a more sensitive assessment of the earliest cognitive changes in neurodegenerative disease. Tests of metaphor may reveal impairments before literal language is impacted. While the cognitive consequences of focal lesions are observed rapidly, neurodegenerative disease processes are progressive and are often quite subtle in the earliest stages. More sensitive measures of cognitive change are needed as screening instruments and as outcome measures in the study of neurodegenerative disease. Metaphor processing, with its distributed neural support and cognitive complexity may be sensitive to early anatomical and cognitive changes.

In conclusion, we show that patients can display a disproportionate deficit in metaphor comprehension, compared to their own performance on literal sentence comprehension and to HC performance. The left hemisphere plays an important role in metaphor processing.

FUNDING INFORMATION

Anjan Chatterjee, NIH, Award ID: R01DC012511. Nathaniel Bloem Klooster, NIH, Award ID: T32 HD071844.

AUTHOR CONTRIBUTIONS

Nathaniel Klooster analyzed the data, created figures, and wrote the manuscript. Marguerite McQuire designed the study, created stimuli, and collected data. Murray Grossman oversaw experimental and neuropsychological data collection and contributed to manuscript writing.

Corey McMillan created figures and contributed to manuscript writing. Anjan Chatterjee designed the study and contributed to manuscript writing. Eileen Cardillo designed the study, created stimuli, and contributed to manuscript writing.

REFERENCES

- Amanzio, M., Geminiani, G., Leotta, D., & Cappa, S. (2008). Metaphor comprehension in Alzheimer's disease: Novelty matters. *Brain and Language*, *107*(1), 1–10. DOI: <https://doi.org/10.1016/j.bandl.2007.08.003>, PMID: 17897706
- Armstrong, B. C., Watson, C. E., & Plaut, D. C. (2012). SOS! An algorithm and software for the stochastic optimization of stimuli. *Behavior Research Methods*, *44*(3), 675–705. DOI: <https://doi.org/10.3758/s13428-011-0182-9>, PMID: 22351612
- Bambini, V., Gentili, C., Ricciardi, E., Bertinetto, P. M., & Pietrini, P. (2011). Decomposing metaphor processing at the cognitive and neural level through functional magnetic resonance imaging. *Brain Research Bulletin*, *86*(3–4), 203–216. DOI: <https://doi.org/10.1016/j.brainresbull.2011.07.015>, PMID: 21803125
- Barry, C. L., Brescoll, V. L., Brownell, K. D., & Schlesinger, M. (2009). Obesity metaphors: How beliefs about the causes of obesity affect support for public policy. *The Milbank Quarterly*, *87*(1), 7–47. DOI: <https://doi.org/10.1111/j.1468-0009.2009.00546.x>, PMID: 19298414, PMCID: PMC2879183
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. DOI: <https://doi.org/10.18637/jss.v067.i01>
- Beger, A., & Jäkel, O. (2015). The cognitive role of metaphor in teaching science: Examples from physics, chemistry, biology, psychology and philosophy. *Philosophical Inquiries*, *3*(1), 89–112. DOI: <https://doi.org/10.4454/philing.v3i1.116>
- Bohrn, I. C., Altmann, U., & Jacobs, A. M. (2012). Looking at the brains behind figurative language—A quantitative meta-analysis of neuroimaging studies on metaphor, idiom, and irony processing. *Neuropsychologia*, *50*(11), 2669–2683. DOI: <https://doi.org/10.1016/j.neuropsychologia.2012.07.021>, PMID: 22824234
- Bottini, G., Corcoran, R., Sterzi, R., Paulesu, E., Schenone, P., Scarpa, P., ... Frith, D. (1994). The role of the right hemisphere in the interpretation of figurative aspects of language: A positron emission tomography activation study. *Brain*, *117*(6), 1241–1253. DOI: <https://doi.org/10.1093/brain/117.6.1241>, PMID: 7820563
- Bowes, A., & Katz, A. (2015). Metaphor creates intimacy and temporarily enhances theory of mind. *Memory & Cognition*, *43*(6), 953–963. DOI: <https://doi.org/10.3758/s13421-015-0508-4>, PMID: 25762491
- Brownell, H. H., Simpson, T. L., Bihrl, A. M., Potter, H. H., & Gardner, H. (1990). Appreciation of metaphoric alternative word meanings by left and right brain-damaged patients. *Neuropsychologia*, *28*(4), 375–383. DOI: [https://doi.org/10.1016/0028-3932\(90\)90063-T](https://doi.org/10.1016/0028-3932(90)90063-T), PMID: 1692980
- Cardillo, E. R., McQuire, M., & Chatterjee, A. (2018). Selective metaphor impairments after left, not right, hemisphere injury. *Frontiers in Psychology*, *9*, 2308. DOI: <https://doi.org/10.3389/fpsyg.2018.02308>, PMID: 30559690, PMCID: PMC6286990
- Cardillo, E. R., Schmidt, G. L., Kranjec, A., & Chatterjee, A. (2010). Stimulus design is an obstacle course: 560 matched literal and metaphorical sentences for testing neural hypotheses about metaphor. *Behavior Research Methods*, *42*(3), 651–664. DOI: <https://doi.org/10.3758/BRM.42.3.651>, PMID: 20805587, PMCID: PMC2952404
- Cardillo, E. R., Watson, C., & Chatterjee, A. (2016). Stimulus needs are a moving target: 240 additional matched literal and metaphorical sentences for testing neural hypotheses about metaphor. *Behavior Research Methods*, *49*(2), 471–483. DOI: <https://doi.org/10.3758/s13428-016-0717-1>, PMID: 26956680
- Cardillo, E. R., Watson, C. E., Schmidt, G. L., Kranjec, A., & Chatterjee, A. (2012). From novel to familiar: Tuning the brain for metaphors. *NeuroImage*, *59*(4), 3212–3221. DOI: <https://doi.org/10.1016/j.neuroimage.2011.11.079>, PMID: 22155328, PMCID: PMC3288556
- Cardoso, S., Silva, D., Maroco, J., de Mendonça, A., & Guerreiro, M. (2014). Non-literal language deficits in mild cognitive impairment. *Psychogeriatrics*, *14*(4), 222–228. DOI: <https://doi.org/10.1111/psyg.12101>, PMID: 25495083
- Chen, E., Widick, P., & Chatterjee, A. (2008). Functional-anatomical organization of predicate metaphor processing. *Brain and Language*, *107*(3), 194–202. DOI: <https://doi.org/10.1016/j.bandl.2008.06.007>, PMID: 18692890, PMCID: PMC2638768
- Citron, F. M., & Goldberg, A. E. (2014). Metaphorical sentences are more emotionally engaging than their literal counterparts. *Journal of Cognitive Neuroscience*, *26*(11), 2585–2595. DOI: https://doi.org/10.1162/jocn_a_00654, PMID: 24800628
- Conrad, E. C., Humphries, H., & Chatterjee, A. (2019). Attitudes toward cognitive enhancement: The role of metaphor and context. *AJOB Neuroscience*, *10*(1), 35–47. DOI: <https://doi.org/10.1080/21507740.2019.1595771>, PMID: 31070552
- Crawford, J. R., & Howell, D. C. (1998). Comparing an individual's test score against norms derived from small samples. *The Clinical Neuropsychologist*, *12*(4), 482–486. DOI: <https://doi.org/10.1076/clin.12.4.482.7241>
- Davis, M. H., Coleman, M. R., Absalom, A. R., Rodd, J. M., Johnsrude, I. S., Matta, B. F., ... Menon, D. K. (2007). Dissociating speech perception and comprehension at reduced levels of awareness. *Proceedings of the National Academy of Sciences*, *104*(41), 16032–16037. DOI: <https://doi.org/10.1073/pnas.0701309104>, PMID: 17938125, PMCID: PMC2042157
- Diaz, M. T., Barrett, K. T., & Hogstrom, L. J. (2011). The influence of sentence novelty and figurativeness on brain activity. *Neuropsychologia*, *49*, 320–330. DOI: <https://doi.org/10.1016/j.neuropsychologia.2010.12.004>, PMID: 21146553, PMCID: PMC3034783
- Elmore, K. C., & Luna-Lucero, M. (2017). Light bulbs or seeds? How metaphors for ideas influence judgments about genius. *Social Psychological and Personality Science*, *8*(2), 200–208. DOI: <https://doi.org/10.1177/1948550616667611>
- Eslinger, P. J., Moore, P., Anderson, C., & Grossman, M. (2011). Social cognition, executive functioning, and neuroimaging correlates of empathic deficits in frontotemporal dementia. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *23*(1), 74–82. DOI: <https://doi.org/10.1176/appi.neuropsych.23.1.74>, PMID: 21304142, PMCID: PMC3641646
- Fellows, L. K., Heberlein, A. S., Morales, D. A., Shivde, G., Waller, S., & Wu, D. H. (2005). Method matters: An empirical study of impact in cognitive neuroscience. *Journal of Cognitive Neuroscience*, *17*(6), 850–858. DOI: <https://doi.org/10.1162/0898929054021139>, PMID: 15969904
- Fernandino, L., Conant, L. L., Binder, J. R., Blindauer, K., Hiner, B., Spangler, K., ... Desai, R. H. (2013). Where is the action? Action sentence processing in Parkinson's disease. *Neuropsychologia*,

- 51(8), 1510–1517. DOI: <https://doi.org/10.1016/j.neuropsychologia.2013.04.008>
- Flusberg, S. J., Matlock, T., & Thibodeau, P. H. (2017). Metaphors for the war (or race) against climate change. *Environmental Communication*, 11(6), 769–783. DOI: <https://doi.org/10.1080/17524032.2017.1289111>
- Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). “Mini-mental state”: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12(3), 189–198. DOI: [https://doi.org/10.1016/0022-3956\(75\)90026-6](https://doi.org/10.1016/0022-3956(75)90026-6), PMID: 1202204
- Gagnon, L., Goulet, P., Giroux, F., & Joannette, Y. (2003). Processing of metaphoric and non-metaphoric alternative meanings of words after right- and left-hemispheric lesion. *Brain and Language*, 87(2), 217–226. DOI: [https://doi.org/10.1016/S0093-934X\(03\)00057-9](https://doi.org/10.1016/S0093-934X(03)00057-9), PMID: 14585291
- Gernsbacher, M. A., & Robertson, R. R. (1999). The role of suppression in figurative language comprehension. *Journal of Pragmatics*, 31(12), 1619–1630. DOI: [https://doi.org/10.1016/S0378-2166\(99\)00007-7](https://doi.org/10.1016/S0378-2166(99)00007-7), PMID: 25520540, PMCID: PMC4266403
- Gibbs, R. W., Jr., & Franks, H. (2002). Embodied metaphor in women’s narratives about their experiences with cancer. *Health Communication*, 14(2), 139–165. DOI: https://doi.org/10.1207/S15327027HC1402_1, PMID: 12046796
- Gorno-Tempini, M. L., Hillis, A. E., Weintraub, S., Kertesz, A., Mendez, M., Cappa, S. F., ... Manes, F. (2011). Classification of primary progressive aphasia and its variants. *Neurology*, 76(11), 1006–1014. DOI: <https://doi.org/10.1212/WNL.0b013e31821103e6>, PMID: 21325651, PMCID: PMC3059138
- Hauser, D. J., & Schwarz, N. (2015). The war on prevention: Bellicose cancer metaphors hurt (some) prevention intentions. *Personality and Social Psychology Bulletin*, 41(1), 66–77. DOI: <https://doi.org/10.1177/0146167214557006>, PMID: 25352114
- Hilligoss, B. (2014). Selling patients and other metaphors: A discourse analysis of the interpretive frames that shape emergency department admission handoffs. *Social Science & Medicine*, 102, 119–128. DOI: <https://doi.org/10.1016/j.socscimed.2013.11.034>, PMID: 24565149
- Howard, D., & Patterson, K. (1992). *Pyramid and palm trees: A test of semantic access from pictures and words*. Bury St. Edmunds, UK: Thames Valley Test Company.
- Ianni, G. R., Cardillo, E. R., McQuire, M., & Chatterjee, A. (2014). Flying under the radar: Figurative language impairments in focal lesion patients. *Frontiers in Human Neuroscience*, 8, 871. DOI: <https://doi.org/10.3389/fnhum.2014.00871>, PMID: 25404906, PMCID: PMC4217389
- Kehagia, A. A., Barker, R. A., & Robbins, T. W. (2010). Neuropsychological and clinical heterogeneity of cognitive impairment and dementia in patients with Parkinson’s disease. *The Lancet Neurology*, 9(12), 1200–1213. DOI: [https://doi.org/10.1016/S1474-4422\(10\)70212-X](https://doi.org/10.1016/S1474-4422(10)70212-X), PMID: 20880750
- Lakoff, G., & Johnson, M. (1980). The metaphorical structure of the human conceptual system. *Cognitive Science*, 4(2), 195–208. DOI: https://doi.org/10.1207/s15516709cog0402_4
- Lee, S. S., & Dapretto, M. (2006). Metaphorical vs. literal word meanings: fMRI evidence against a selective role of the right hemisphere. *NeuroImage*, 29, 536–544. DOI: <https://doi.org/10.1016/j.neuroimage.2005.08.003>, PMID: 16165371
- Lee, S. W., & Schwarz, N. (2014). Framing love: When it hurts to think we were made for each other. *Journal of Experimental Social Psychology*, 54, 61–67. DOI: <https://doi.org/10.1016/j.jesp.2014.04.007>
- Libon, D. J., Mattson, R. E., Glosser, G., & Kaplan, E. 1996. A nine-word dementia version of the California Verbal Learning Test. *The Clinical Neuropsychologist*, 10, 237–244. DOI: <https://doi.org/10.1080/13854049608406686>
- Mashal, N., & Faust, M. (2009). Conventionalisation of novel metaphors: A shift in hemispheric asymmetry. *Laterality*, 14(6), 573–589. DOI: <https://doi.org/10.1080/13576500902734645>, PMID: 19253086
- Mashal, N., Faust, M., & Hendler, T. (2005). The role of the right hemisphere in processing nonsalient metaphorical meanings: Application of principal components analysis to fMRI data. *Neuropsychologia*, 43(14), 2084–2100. DOI: <https://doi.org/10.1016/j.neuropsychologia.2005.03.019>, PMID: 16243053
- Matlock, T. (2004). Fictive motion as cognitive simulation. *Memory & Cognition*, 32(8), 1389–1400. DOI: <https://doi.org/10.3758/BF03206329>, PMID: 15900932
- Monetta, L., & Pell, M. D. (2007). Effects of verbal working memory deficits on metaphor comprehension in patients with Parkinson’s disease. *Brain and Language*, 101(1), 80–89. DOI: <https://doi.org/10.1016/j.bandl.2006.06.007>, PMID: 16875726
- Nelson, H. E., & O’Connell, A. (1978). Dementia: The estimation of premorbid intelligence levels using the New Adult Reading Test. *Cortex*, 14(2), 234–244. DOI: [https://doi.org/10.1016/S0010-9452\(78\)80049-5](https://doi.org/10.1016/S0010-9452(78)80049-5), PMID: 679704
- Nerlich, B., & Jaspal, R. (2012). Metaphors we die by? Geoenvironment, metaphors, and the argument from catastrophe. *Metaphor and Symbol*, 27(2), 131–147. DOI: <https://doi.org/10.1080/10926488.2012.665795>
- Noonan, K. A., Jefferies, E., Visser, M., & Lambon Ralph, M. A. (2013). Going beyond inferior prefrontal involvement in semantic control: Evidence for the additional contribution of dorsal angular gyrus and posterior middle temporal cortex. *Journal of Cognitive Neuroscience*, 25(11), 1824–1850. DOI: https://doi.org/10.1162/jocn_a_00442, PMID: 23859646
- Obert, A., Gierski, F., Calmus, A., Portefaix, C., Declercq, C., Pierot, L., & Caillies, S. (2014). Differential bilateral involvement of the parietal gyrus during predicative metaphor processing: An auditory fMRI study. *Brain and language*, 137, 112–119. DOI: <https://doi.org/10.1016/j.bandl.2014.08.002>, PMID: 25193417
- Papagno, C. (2001). Comprehension of metaphors and idioms in patients with Alzheimer’s disease: A longitudinal study. *Brain*, 124(7), 1450–1460. DOI: <https://doi.org/10.1093/brain/124.7.1450>, PMID: 11408339
- Papagno, C., Lucchelli, F., Muggia, S., & Rizzo, S. (2003). Idiom comprehension in Alzheimer’s disease: The role of the central executive. *Brain*, 126(11), 2419–2430. DOI: <https://doi.org/10.1093/brain/awg243>, PMID: 12902312
- Rapp, A. M., Leube, D. T., Erb, M., Grodd, W., & Kircher, T. T. (2004). Neural correlates of metaphor processing. *Brain Research*, 20, 395–402. DOI: <https://doi.org/10.1016/j.cogbrainres.2004.03.017>, PMID: 15268917
- Rapp, A. M., Leube, D. T., Erb, M., Grodd, W., & Kircher, T. T. (2007). Laterality in metaphor processing: Lack of evidence from functional magnetic resonance imaging for the right hemisphere theory. *Brain and Language*, 100, 142–149. DOI: <https://doi.org/10.1016/j.bandl.2006.04.004>, PMID: 16677700
- Rapp, A. M., Mutschler, D. E., & Erb, M. (2012). Where in the brain is nonliteral language? A coordinate-based meta-analysis of functional magnetic resonance imaging studies. *NeuroImage*, 63(1), 600–610. DOI: <https://doi.org/10.1016/j.neuroimage.2012.06.022>, PMID: 22759997
- Rapp, A. M., & Wild, B. (2011). Nonliteral language in Alzheimer dementia: A review. *Journal of the International Neuropsychological Society*, 17(2), 207–218. DOI: <https://doi.org/10.1017/S1355617710001682>, PMID: 21241530

- Reyes-Aguilar, A., Valles-Capetillo, E., & Giordano, M. (2018). A quantitative meta-analysis of neuroimaging studies of pragmatic language comprehension: In search of a universal neural substrate. *Neuroscience*, *395*, 60–88. DOI: <https://doi.org/10.1016/j.neuroscience.2018.10.043>, PMID: 30414881
- Robins, S., & Mayer, R.E. (2000). The metaphor framing effect: Metaphorical reasoning about text-based dilemmas. *Discourse Processes*, *30*, 57–86. DOI: https://doi.org/10.1207/S15326950dp3001_03
- Roncero, C., & de Almeida, R. G. (2014). The importance of being apt: Metaphor comprehension in Alzheimer's disease. *Frontiers in Human Neuroscience*, *8*, 973. DOI: <https://doi.org/10.3389/fnhum.2014.00973>, PMID: 25520642, PMCID: PMC4251318
- Saban, A. (2006). Functions of metaphor in teaching and teacher education: A review essay. *Teaching Education*, *17*(4), 299–315. DOI: <https://doi.org/10.1080/10476210601017386>
- Schmidt, G. L., Kranjec, A., Cardillo, E. R., & Chatterjee, A. (2010). Beyond laterality: A critical assessment of research on the neural basis of metaphor. *Journal of the International Neuropsychological Society*, *16*(1), 1–5. DOI: <https://doi.org/10.1017/S1355617709990543>, PMID: 19765354, PMCID: PMC2926935
- Schmidt, G. L., & Seger, C. A. (2009). Neural correlates of metaphor processing: The roles of figurativeness, familiarity and difficulty. *Brain and Cognition*, *71*(3), 375–386. DOI: <https://doi.org/10.1016/j.bandc.2009.06.001>, PMID: 19586700, PMCID: PMC2783884
- Smith, S. M., & Nichols, T. E. (2009). Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. *NeuroImage*, *44*(1), 83–98. DOI: <https://doi.org/10.1016/j.neuroimage.2008.03.061>, PMID: 18501637
- Sopory, P., & Dillard, J. P. (2002). The persuasive effects of metaphor: A meta-analysis. *Human Communication Research*, *28*(3), 382–419. DOI: <https://doi.org/10.1111/j.1468-2958.2002.tb00813.x>
- Steen, G. J., Dorst, L., Herrmann, B., Kaal, A., Krennmayr, T., & Pasma, T. (2010). *A method for linguistic metaphor identification: From MIP to MIPVU*. Amsterdam/Philadelphia: John Benjamins. DOI: <https://doi.org/10.1075/celcr.14>
- Swartz, R. H., Stuss, D. T., Gao, F., & Black, S. E. (2008). Independent cognitive effects of atrophy and diffuse subcortical and thalamo-cortical cerebrovascular disease in dementia. *Stroke*, *39*(3), 822–830. DOI: <https://doi.org/10.1161/STROKEAHA.107.491936>, PMID: 18258840
- Thibodeau, P. H., & Boroditsky, L. (2011). Metaphors we think with: The role of metaphor in reasoning. *PLOS One*, *6*(2), e16782. DOI: <https://doi.org/10.1371/journal.pone.0016782>, PMID: 21373643, PMCID: PMC3044156
- Thibodeau, P. H., & Boroditsky, L. (2013). Natural language metaphors covertly influence reasoning. *PLOS One*, *8*(1), e52961. DOI: <https://doi.org/10.1371/journal.pone.0052961>, PMID: 23301009, PMCID: PMC3534638
- Tompkins, C. (1990). Knowledge and strategies for processing lexical metaphor after right or left hemisphere brain damage. *Journal of Speech, Language, and Hearing Research*, *33*(2), 307–316. DOI: <https://doi.org/10.1044/jshr.3302.307>
- Tustison, N. J., Avants, B. B., Cook, P. A., Zheng, Y., Egan, A., Yushkevich, P. A., & Gee, J. C. (2010). N4ITK: Improved N3 bias correction. *IEEE Transactions on Medical Imaging*, *29*(6), 1310–1320. DOI: <https://doi.org/10.1109/TMI.2010.2046908>, PMID: 20378467, PMCID: PMC3071855
- Tustison, N. J., Cook, P. A., Klein, A., Song, G., Das, S. R., Duda, J. T., ... Avants, B. B. (2014). Large-scale evaluation of ANTs and FreeSurfer cortical thickness measurements. *NeuroImage*, *99*, 166–179. DOI: <https://doi.org/10.1016/j.neuroimage.2014.05.044>, PMID: 24879923
- Van Rijn-van Tongeren, G. W. (1997). *Metaphors in medical texts* (Vol. 8). Amsterdam: Rodopi.
- Van Stee, S. K. (2018). Meta-analysis of the persuasive effects of metaphorical vs. literal messages. *Communication Studies*, *69*(5), 1–22. DOI: <https://doi.org/10.1080/10510974.2018.1457553>
- Winner, E., & Gardner, H. (1977). The comprehension of metaphor in brain-damaged patients. *Brain*, *100*(4), 717–729. DOI: <https://doi.org/10.1093/brain/100.4.717>, PMID: 608117
- Yang, F. G., Fuller, J., Khodaparast, N., & Krawczyk, D. C. (2010). Figurative language processing after traumatic brain injury in adults: A preliminary study. *Neuropsychologia*, *48*(7), 1923–1929. DOI: <https://doi.org/10.1016/j.neuropsychologia.2010.03.011>, PMID: 20230844
- Yang, J. (2014). The role of the right hemisphere in metaphor comprehension: A meta-analysis of functional magnetic resonance imaging studies. *Human Brain Mapping*, *35*(1), 107–122. DOI: <https://doi.org/10.1002/hbm.22160>, PMID: 22936560, PMCID: PMC6868953
- Zaidel, E., Kasher, A., Soroker, N., & Batori, G. (2002). Effects of right and left hemisphere damage on performance of the "Right Hemisphere Communication Battery." *Brain and Language*, *80*(3), 510–535. DOI: <https://doi.org/10.1006/brln.2001.2612>, PMID: 11896655
- Zempleni, M. Z., Renken, R., Hoeks, J. C., Hoogduin, J. M., & Stowe, L. A. (2007). Semantic ambiguity processing in sentence context: Evidence from event-related fMRI. *NeuroImage*, *34*(3), 1270–1279. DOI: <https://doi.org/10.1016/j.neuroimage.2006.09.048>, PMID: 17142061