Current Neurology and Neuroscience Reports (2019) 19: 53
https://doi.org/10.1007/s11910-019-0971-6

BEHAVIOR (H. S. KIRSHNER, SECTION EDITOR)

Multivariate Approaches to Understanding Aphasia and its Neural Substrates

Stephen M. Wilson[1] · William D. Hula[2,3]

Published online: 27 June 2019

Abstract

Purpose of Review Aphasia is often characterized in terms of subtype and severity, yet these constructs have limited explanatory power, because aphasia is inherently multifactorial both in its neural substrates and in its symptomatology. The purpose of this review is to survey current and emerging multivariate approaches to understanding aphasia.

Recent Findings Techniques such as factor analysis and principal component analysis have been used to define latent underlying factors that can account for performance on batteries of speech and language tests, and for characteristics of spontaneous speech production. Multivariate lesion-symptom mapping has been shown to outperform univariate approaches to lesion-symptom mapping for identifying brain regions where damage is associated with specific speech and language deficits. It is increasingly clear that structural damage results in functional changes in wider neural networks, which mediate speech and language outcomes.

Summary Multivariate statistical approaches are essential for understanding the complex relationships between the neural substrates of aphasia, and resultant profiles of speech and language function.

Keywords Aphasia · Multivariate · Factor analysis · Principal components analysis · Neural substrates · Multivariate lesion-symptom mapping

Introduction

In clinical practice, aphasia is most often characterized in terms of subtype and severity. Approximately ten subtypes are commonly recognized—Broca's aphasia, Wernicke's aphasia, conduction aphasia, and so on—which were defined by nineteenth century pioneers [1, 2] and refined by aphasiologists of the Boston school [3–5]. The symptomatology of each subtype is generally explained in terms damage to centers and/or connections of a cognitive-neuroanatomical model. Another salient dimension on which aphasias vary is severity, and the idea that a single functional capacity explains many aspects of aphasic behavior also has a rich history [6–9].

While concepts of subtype and severity have great clinical utility, neither of these constructs are particularly faithful to the underlying neural substrates of aphasia. Over the last few decades, cognitive neuroscientists have developed detailed models of how language processing is organized in the brain [10–12]. While many specifics continue to be debated, a consensus has emerged that language function is the product of numerous interacting regions and networks, some of which are specialized for particular linguistic domains (e.g., phonology, semantics, syntax), and others of which support functions that are relevant to multiple cognitive domains (e.g., attention networks, cingulo-opercular network) [13–15].

Aphasia is caused by damage to these networks, whether by ischemic or hemorrhagic stroke, neurodegeneration, surgical resection, or other etiologies. With the possible partial exception of neurodegeneration [16], the neurological events causing the damage do not generally respect the functional architecture of the networks and regions that are impacted. Aphasia then involves varying degrees of damage to a varying

This article is part of the Topical Collection on *Behavior*

✉ Stephen M. Wilson
stephen.m.wilson@vanderbilt.edu

[1] Department of Hearing and Speech Sciences, Vanderbilt University Medical Center, Nashville, TN, USA

[2] Audiology and Speech Pathology Program, VA Pittsburgh Healthcare System, Pittsburgh, PA, USA

[3] Department of Communication Science and Disorders, University of Pittsburgh, Pittsburgh, PA, USA

number of language and domain-general networks, and consequently, linguistic domains (e.g., phonology, semantics, syntax) are often impaired to differing degrees. Moreover, surviving brain regions are recruited to support compensatory processes [17], even in neurodegenerative aphasias [18]. This situation is not well captured simply by characterizing patients in terms of aphasia subtype or severity.

In short, aphasia is inherently multifactorial both in its neural substrates and in its symptomatology. Progress in understanding aphasia will depend on conceptual and statistical approaches that incorporate this multidimensionality. The purpose of this brief review is to survey multivariate approaches to understanding aphasia, with a focus on current and emerging techniques and perspectives.

Multivariate analysis of speech and language deficits

The related techniques of common factor analysis (FA) and principal components analysis (PCA) are among the most commonly and productively applied multivariate analysis techniques in studies of speech and language behavior in aphasia. The earliest applications of factor analysis to aphasic language [19, 20] produced similar results, if diverging interpretations, with Jones and Wepman emphasizing multidimensionality corresponding to input and output modality [19] and Schuell and colleagues highlighting the positive correlations among the factors and the idea of a unidimensional underlying language deficit [20]. There have been many subsequent applications of FA and PCA in the development or analysis of language assessment batteries [5, 21–26], and these have generally substantiated the presence of both a general severity factor and multiple other factors that essentially align with input and output modality. Notably, a very similar pattern was found in factor analysis of item-level self-reports of communicative function by persons with aphasia [27]. Although most investigators have followed Jones' and Wepman's emphasis on distinct domains or modalities of language performance, it should be noted that neither FA nor PCA methods alone can resolve the issue, in part because results and conclusions will depend on the particular sample of behaviors chosen for analysis [24]. As noted by Goodglass and Kaplan [5], correlations among subtests may be due to a common cognitive basis or the spatial contiguity of lesions to brain areas that support those functions.

More recent applications of PCA in aphasia [28, 29•, 30–33] have demonstrated the utility of this approach for defining a small number of components (typically two to four in this literature) as linear combinations of a larger number of test scores. These components have been given neurolinguistic interpretations based on the variables contributing to them, and score estimates derived from them have then used with voxel-based lesion-symptom mapping (VLSM) or related techniques to investigate hypotheses about the neural correlates of behavior. Some of these more recent studies have included new indicator variables motivated by modern psycholinguistic theories of language performance in aphasia [e.g., 34] and have expanded what is known about the dimensions along which aphasic language performance systematically varies. Butler and colleagues [28] analyzed data from 31 individuals with aphasia on 17 tasks including picture naming, word and nonword repetition, semantic and phonological judgments, word and sentence comprehension, and visuospatial problem solving. They identified three components—phonology, semantics, and cognition—all of which correlated to varying degrees with a general severity factor. A subsequent study [33] on the same dataset added four indicator variables indexing connected language production, resulting in an additional component related to speech quantity. Mirman and colleagues [29•, 30] used a similar approach on a larger dataset of 99 individuals and also identified four components, but the nature of the components was somewhat different: speech recognition, speech production, semantic recognition and semantic production errors. In addition to the differences in samples size, the discrepancies between these two sets of results may reflect differences in the sets of measures used: for instance, Butler and colleagues did not include indices of semantic or phonological naming errors, while Mirman and colleagues did not include measures of connected speech or visuospatial function. Thus, in addition to the frequently observed distinction between input and output, these analyses have supported the claim that people with aphasia meaningfully vary along at least semi-independent dimensions of semantic and phonological/motor speech ability.

Another line of FA studies has focused on indicator variables derived from spontaneous speech samples. In their update of an earlier study [35] limited by methodological issues, Vermeulen and colleagues [36] analyzed 17 quantitative connected speech variables (e.g., speech rate, mean length of utterance, number of semantic paraphasias, seconds of incomprehensible speech) and a naming test score collected from 122 individuals with aphasia. They identified five factors: syntactic ability, phonological paraphasia, neologistic paraphasia, articulatory impairment, and vocabulary. In a more recent study, Casilio and colleagues [37] took a different approach, deriving perceptual ratings of 27 features of connected speech samples from 24 patients with aphasia. Following a detailed investigation of the inter-rater reliability of these ratings and consideration of their concurrent validity with more objective measures derived from the same samples, FA was used to define four factors: paraphasia, logopenia, agrammatism, and motor speech. In comparing their results to Vermeulen and colleagues', Casilio and colleagues [37] noted close correspondence between three pairs of factors across the two analyses: agrammatism with syntactic ability,

motor speech with articulatory impairment, and logopenia with vocabulary, while the paraphasia factor was loosely related to the remaining two factors in Vermeulen and colleagues' analysis. While the convergence of these two studies is encouraging, it is worth noting that the factors they identified contrast with those arising from PCA applied to batteries of tests described above [28, 29•, 30, 33]. Because formal tests and connected speech analysis provide complementary but consistent information about patients' speech and language function [39, 40], future studies that combine these two types of data, ideally with large numbers of patients of diverse etiologies, may be able to resolve these discrepancies and identify stable and consistent factors.

In contrast to the above studies, which have all identified multidimensional components or factors that are theoretically motivated or interpretable, another noteworthy line of FA studies provides a potentially useful counterpoint. Caplan and colleagues' investigations of syntactic processing in aphasia [38–40] have consistently found support for a unidimensional model of performance on various sentence comprehension tasks, suggesting that distinctions in sentence structure motivated by linguistic theory do not account for meaningful variance. The strength of this conclusion is moderated by the relatively small sample sizes used and the problems noted above in interpreting a strong unidimensional factor in the presence of lesions determined by spatial rather than functional architecture.

In reviewing this literature, consideration of two issues may be useful to investigators seeking to apply these methods moving forward. One is the distinction between FA and PCA, which have both different goals and distinct estimation procedures [41, 42]. PCA is a data reduction technique that aims to represent all of the variance in the observed indicator variables in as small a number of weighted linear combinations (components) as possible. Because it analyzes all of the variance in the data, PCA, like regression analysis, carries the assumption that each of the observed variables is measured without error. By contrast, FA seeks to explain the covariances among the observed variables in terms of latent factors that are assumed to be the underlying causes of those variables. For this reason, FA is the preferred technique when the goal of the analysis is to describe a theoretical latent structure underlying a set of observed variables [41, 43]. Furthermore, because FA analyzes only the variance that is shared among the indicator variables, rather than the total variance, FA models typically obtain lower loadings and communalities (proportions of indicator variable variance attributable to factors) than PCA models of the same data. While these differences may be small when communalities are high, they increase as the communalities and the number of indicators per factor decrease [41]. These FA-PCA differences are important because they can affect the structure of the model, estimates of the factor or component scores, and the replicability of solutions across samples [41, 42].

Second, most FA studies of aphasia have used exploratory factor analysis (EFA), as opposed to confirmatory factor analysis (CFA) models. There are several complicating issues that arise in using EFA. One issue concerns the need to determine the number of factors or components to extract from the data. In every case cited above, investigators used the Kaiser-Guttman criterion, which requires an eigenvalue > 1, indicating that all retained components explained more variance than any single observed variable. This criterion is recognized to both over- and underestimate the number of factors or components, depending on the context, and alternative methods that employ statistical tests of model-data fit or parallel analysis, which bases the decision on comparison of the empirical eigenvalues to those obtained from random data, are currently preferred [41, 42]. A second issue is that users of EFA must choose a rotation of the factors or components in the multidimensional space derived from the analysis to interpret. Rotations are typically chosen to maximize the approximation to a simple structure, i.e., one in which each variable loads on only one factor or component. While many rotation methods exist, one of the most important decisions an analyst must make is whether to choose a rotation in which the factors are permitted to correlate with one another or one in which they are constrained to be independent. The differences in loadings may be small in some cases [e.g., 30], but these differences may have larger consequences for factor or component scores in the presence of substantial inter-factor correlations. In other cases, orthogonal rotations may be preferred for practical reasons (e.g., maximizing uncorrelated predictive variance [e.g., 28]). However, exclusive use of orthogonal rotations can obscure important features of the data, such as simple structure or substantively interpretable correlations between factors and the potential existence of higher-order factors [41, 42]. The most frequently used approach in the aphasia studies cited above (PCA followed by orthogonal varimax rotation) is likely sub-optimal for defining the structure underlying a set of observed variables [42]. Furthermore, over-reliance on exploratory approaches may limit progress because of the complicating issues discussed above.

Given the amount of existing literature on aphasic language behavior and the substantial agreement across studies, more frequent application of explicitly confirmatory modeling approaches could be productive. Advantages of these approaches include a priori statistical identification of the factors, thus obviating the need for rotation, and greater emphasis on absolute and relative tests of model fit. For example, Swinburn and colleagues [25] used a sequence of nested CFA models and likelihood ratio tests to demonstrate that a model based on the subscale structure of the Comprehensive Aphasia Test fit the data better than either a unidimensional

model or one based the results of a prior exploratory PCA. CFA models have also been used to identify optimal indicators of lexical diversity in aphasic discourse [44] and to investigate substantive hypotheses about sentence processing [38].

We conclude this section by describing a particularly innovative application of confirmatory latent trait modeling. Walker, Hickok, and Fridriksson [45•] analyzed a large set of responses to the Philadelphia Naming Test (PNT) using a hybrid multinomial processing tree/item response theory (IRT) model designed to predict error types at the item level as a function of several latent traits motivated by current theories of word production. Briefly, multinomial processing trees are probabilistic models that explain categorical outcomes as a series of branching binary decisions [46, 47]. While productive, these models typically do not account for heterogeneity among persons or items. IRT is a modeling approach that predicts categorical responses as a logistic function of continuous latent variables, and has been used for decades in education [48] and more recently in medicine and rehabilitation [49] to construct behavioral tests and measure individual differences between both persons and items. While IRT models have been used previously to derive estimates of overall naming ability in aphasia [50, 51], the important innovations of Walker and colleagues' approach are that it has the potential to provide psychometrically robust measurement of latent constructs (for both individuals and groups) motivated by current theories of word production, and to provide powerful means of testing those theories against empirical data.

Multivariate analysis of the neural substrates of aphasia

For over 150 years, studies of individuals with aphasia have been used to infer relationships between damage to specific brain regions and the emergence of specific speech and language deficits [1]. This became increasingly feasible with the advent of CT and MRI [52], and over the last two decades, the majority of such studies have used voxel-by-voxel techniques such as VLSM [53, 54] and voxel-based morphometry (VBM) [55]. These are mass univariate approaches in which statistical tests are carried out at each voxel to quantify the extent to which damage to that voxel is associated with the speech/language deficit in question. Numerous voxelwise studies have systematically investigated the neural correlates of numerous speech/language variables reflecting multiple language domains [56, 57].

However, the mass univariate approach has two fundamental limitations. First, by analyzing each voxel in isolation, it fails to take into account spatial contingencies between the lesion status of voxels, which arise from the underlying pathological processes [58–61, 62•]. For instance, in a stroke cohort, if the middle temporal gyrus (MTG) is damaged, then it is highly likely that the adjacent superior temporal gyrus (STG) is damaged too, because MTG damage is usually caused by middle cerebral artery strokes, and to be large enough to reach the MTG, they would typically impact the STG along the way. These spatial contingencies are systematic and pervasive, and have been shown to lead to substantial displacement of the brain regions that are inferred to be critical for any given behavior [60, 63]. Most VLSM studies include lesion size as a covariate, which somewhat ameliorates the problem by accounting in a gross way for damage to other brain regions [64], but this is far from a complete fix [62•]. Another partial solution is to carry out post-hoc multiple regression analyses in which several brain regions are included as independent variables to elucidate whether their contributions are independent or not [54, 65]. This approach, however, is limited in scope in that only a handful of regions can feasibly be investigated.

The second fundamental limitation of the mass univariate approach is that it fails to consider the potential impact of disconnections caused by lesions [61, 62•]. For example, damage to white matter tracts involved in language processing could result in similar deficits to damage of the cortical nodes that communicate via those tracts [66]. Some researchers have argued that disruption of functional connectivity caused by damage to white matter is in fact the primary driver of language and other deficits after stroke [67, 68•].

Multivariate approaches to analyzing neural data have the potential to overcome these two limitations of the mass univariate approach, by providing a principled framework for inferring relationships between brain damage and speech-language outcomes that allows for contingencies and interactions between multiple brain regions and networks. Traditional multivariate methods such as multiple regression are unlikely to lead to much headway, because of the very large number of brain regions (or voxels) that need to be considered [69]. Rather, machine learning methods such as support vector machines (SVM) [70], support vector regression (SVR) [71], random forests [72], Gaussian process model regression [73], and Bayesian networks [74] have been proposed to investigate brain-behavior relationships in aphasia.

A complete description and comparison of all of these methods is beyond the scope of this review, so we will limit ourselves to describing the basic concepts behind SVM and SVR, which are the most commonly used machine learning methods in neuroimaging. In brief, individual patients' brains are conceptualized as points in an n-dimensional space, where n may represent the number of brain regions considered, the number of voxels, or some set of features that are derived from brain images. Consider first an application in which we wish to predict a symptom or syndrome based on imaging, such as, for example, agrammatism. An optimal separating hyperplane is identified such that, as far as feasible without overfitting, patients with agrammatism lie on one side of the hyperplane,

and those without agrammatism lie on the other side [75]. For any new case, it is possible to predict whether or not they will exhibit agrammatism based on which side of the hyperplane they lie on. Non-linear kernels can be used so that the hyperplane need not be linear. Moreover, the hyperplane can be projected back onto brain regions, voxels, or features, in order to determine which brain locations are most important for prediction. Finally, by considering not just which side of the hyperplane a new case lies on, but their distance from the hyperplane, we can predict not just binary diagnoses, but any continuous behavioral variable; this is termed support vector regression (SVR). Two technical considerations are noteworthy. First, the success or failure of SVM and SVR can be strongly influenced by selection and tuning of *hyperparameters*, such as *C*, which controls the tradeoff between separability and stability in fitting the hyperplane. Many other machine learning methods also involve hyperparameters. The choice and/or optimization of hyperparameters has been received scant attention in the aphasia literature to date [70, 71]. Second, it is important to carefully design cross-validation schemes such that models are trained only on training data, and tested on "held out" data that has not contributed in any way to constructing the model [72].

In an early multivariate study, SVMs were used to differentiate patients with three different variants of primary progressive aphasia from neurologically normal controls, as well as from each other, based on structural imaging [70]. The average accuracy for pairwise discrimination between groups was over 92%, suggesting that it is usually possible to correctly diagnose individual patients based on MRI alone. Of note, the authors computed images of lateralized atrophy by subtracting estimated gray matter densities between the hemispheres; in this way, they ensured that the SVMs would have direct access to information about patterns of lateralized atrophy known to be diagnostic for PPA. In another study, Gaussian process model regression was used to predict a speech production measure in post-stroke aphasia based on structural imaging, demographic data, and time [73]. Using atlas-based regions of interest, the authors were able to account for 59% of the variance in the speech production measure, which was significantly greater than the variance that could be accounted for based on left hemisphere lesion volume alone (47%). Other studies have predicted Western Aphasia Battery (WAB) aphasia type from structural imaging [76], WAB subscores from structural imaging [77], good versus poor recovery from subacute functional imaging [78], and a range of language scores from multimodal structural and functional imaging [72].

Other studies have focused on identifying the patterns of brain damage that are associated with different syndromes, symptoms, or outcomes; this is sometimes termed multivariate lesion-symptom mapping (MLSM) [30, 32, 71, 79–81, 82•, 83]. These studies have identified brain regions associated with WAB subscores [32, 79], accuracy on the PNT [32, 82•, 83], semantic and phonological errors on the PNT [32, 71], components derived from rotated PCA of batteries of language measures [30, 32, 80], as well as many more specific measures such as pseudoword repetition [81] and apraxia of speech [32]. Some of these studies have focused on damage to specific brain regions, while others have also considered disconnection of networks [32, 72, 77]. Generally, the brain regions and networks associated with the language functions studied have been consistent with findings from functional neuroimaging and other cognitive neuroscience methodologies. Some highlights include a dissociation between speech recognition and speech production processes, which were localized to the superior temporal gyrus and supramarginal gyrus respectively [30], and localization of sentence comprehension to posterior temporal cortex [32].

Several studies have compared MLSM to VLSM based on data from individuals with aphasia [30, 32, 71, 82•, 83], and/or simulated datasets in which real lesion data is linked to simulated behavioral data based on hypothetical lesion-behavior relationships, and models are constructed with the goal of recovering the known relationships [60, 71, 82•]. In synthetic data in which damage to multiple regions may contribute to behavioral outcomes, multivariate approaches have been shown to outperform univariate approaches in identifying all relevant regions [60, 71, 82•]. When there are multiple regions that contribute to a behavioral outcome, univariate approaches may end up identifying the region that lies between them, even though that region may play no actual role in the function [60]. In contrast to studies of synthetic data which have indicated significant advantages of MLSM, analyses of the same real aphasia datasets with VLSM and MLSM have revealed surprisingly similar maps. Mirman and colleagues [30] performed MLSM using SVR [71] to identify brain regions associated with semantic recognition, semantic errors, speech production and speech recognition, and compared the regions they had identified for these same components in a previous VLSM study [29•]. The Dice coefficients of similarity for the four maps ranged from 0.92 to 0.98, reflecting a striking degree of similarity. Qualitative comparisons between VLSM and MLSM maps in other studies are also remarkably similar [32, 71, 82•, 83]. It is not yet clear why the clear theoretical advantages of MLSM do not translate into dramatically different findings for the speech/language variables that are typically investigated.

Multivariate analyses of the neural substrates of aphasia need not be limited to structural imaging data. Recently, there has been a growing appreciation of the necessity of considering structural and functional changes

in tandem. A combined fMRI/VBM study showed that syntactic deficits are associated with left frontal atrophy in PPA, but that the left frontal atrophy impacts the functionality of a wider perisylvian network, ultimately affecting syntax [84]. In a similar vein, an analysis using joint independent component analysis showed that left temporo-parietal damage in post-stroke aphasia was associated with reduced activity in several nodes of the canonical semantic network [85]. Several studies in post-stroke aphasia have indicated that patterns of right hemisphere recruitment are dependent on the location and extent of left hemisphere damage, and that right hemisphere activation may or may not be positively associated with language outcomes depending on the precise nature of left hemisphere damage [86–88].

Especially noteworthy are a recent series of studies by Corbetta and colleagues that have employed highly innovative multivariate analyses to investigate relationships between structural damage, functional connectivity, and behavior in a large longitudinal series of stroke patients. These studies are not focused on aphasia specifically, but language is considered along with other domains commonly affected in stroke (motor, memory, attention, visual). In terms of behavior, these studies have emphasized the surprising extent to which deficits are correlated across domains (e.g. language and memory), as well as within each domain (e.g. expressive and receptive function within the language domain) [9, 67, 89]. Corbetta and colleagues argue that the covariance structure of deficits follows from two facts about stroke: first, strokes frequently damage white matter, impairing communication between multiple brain regions, and second, strokes have physiological effects that extend far beyond the lesion site. They have characterized these effects primarily using measures of functional connectivity derived from resting state fMRI. Most saliently, they showed that reduced interhemispheric connectivity between homotopic regions is highly predictive of deficits in higher functions, including language [68•]. Interestingly, increased intrahemispheric connectivity between left hemisphere regions was also predictive of better language outcomes, while increased intrahemispheric connectivity between right hemisphere regions was associated with worse language outcomes. Reduced interhemispheric connectivity was associated with decreased segregation of networks that are normally anticorrelated (the dorsal attention network and the default mode network). Re-establishment of integration within networks and segregation between networks was associated with recovery from aphasia and other deficits [90]. This work converges with other MLSM studies that have shown how aphasia severity is determined by structural measures of network integrity [91, 92].

Conclusion

We have surveyed multivariate approaches to understanding aphasia, both from a behavioral perspective and from a neural perspective. Considering the clear-cut multidimensionality of aphasia and its neural underpinnings, it is no surprise that multidimensional approaches have been growing in prevalence and prominence over the last few years. Disentangling multiple behavioral and neural variables requires large datasets, so it is encouraging that the field of language neuroscience is becoming increasingly collaborative, with the emergence of a number of large scale projects that will generate datasets of the size and substance that will form a sound basis for future progress [32, 93–95].

Acknowledgments The authors thank Andrew T. DeMarco for helpful comments on an earlier version of this manuscript.

Compliance with Ethical Standards

Conflict of Interest Stephen M. Wilson declares no potential conflicts of interest. William D. Hula reports grants from NIDCD (R01 DC013270, R21 DC016080), during the conduct of the study.

Human and Animal Rights and Informed Consent This article does not contain any studies with human or animal subjects performed by any of the authors.

References

Papers of particular interest, published recently, have been highlighted as:
- Of importance

1. Wernicke C. Der Aphasische Symptomencomplex. Breslau: Cohn and Weigert; 1874.
2. Lichtheim L. On aphasia. Brain. 1885;7:433–84.
3. Geschwind N. Disconnexion syndromes in animals and man. I Brain. 1965;88:237–94.
4. Geschwind N. Disconnexion syndromes in animals and man. II Brain. 1965;88:585–644.
5. Goodglass H, Kaplan E. The assessment of aphasia and related disorders. 1st ed. Philadelphia: Lea and Febiger; 1972.
6. Marie P. La troisième circonvolution frontal gauche ne joue aucun rôle spécial dans la fonction du langage. Sem Médicale. 1906;26:241–7.
7. Goldstein K. Language and language disturbances. New York: Grune and Stratton; 1948.
8. Schuell H, Jenkins JJ, Jimenez-Pabon E. Aphasia in adults. New York: Harper & Row; 1964.
9. Corbetta M, Siegel JS, Shulman GL. On the low dimensionality of behavioral deficits and alterations of brain network connectivity after focal injury. Cortex. 2018;107:229–37.
10. Hickok G, Poeppel D. The cortical organization of speech processing. Nat Rev Neurosci. 2007;8:393–402.
11. Price CJ. A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. NeuroImage. 2012;62:816–47.

12. Wilson SM, Bautista A, McCarron A. Convergence of spoken and written language processing in the superior temporal sulcus. NeuroImage. 2018;171:62–74.
13. Fedorenko E, Thompson-Schill SL. Reworking the language network. Trends Cogn Sci. 2014;18:120–6.
14. Geranmayeh F, Brownsett SLE, Wise RJS. Task-induced brain activity in aphasic stroke patients: what is driving recovery? Brain. 2014;137:2632–48.
15. Yen M, DeMarco AT, Wilson SM. Adaptive paradigms for mapping phonological regions in individual participants. NeuroImage. 2019;189:368–79.
16. Seeley WW, Crawford RK, Zhou J, Miller BL, Greicius MD. Neurodegenerative diseases target large-scale human brain networks. Neuron. 2009;62:42–52.
17. Turkeltaub PE, Messing S, Norise C, Hamilton RH. Are networks for residual language function and recovery consistent across aphasic patients? Neurology. 2011;76:1726–34.
18. Wilson SM, Brambati SM, Henry RG, Handwerker DA, Agosta F, Miller BL, et al. The neural basis of surface dyslexia in semantic dementia. Brain. 2009;132:71–86.
19. Jones LV, Wepman JM. Dimensions of language performance in aphasia. J Speech Hear Res. 1961;4:220–32.
20. Schuell H, Jenkins JJ, Carroll JB. A factor analysis of the Minnesota Test for Differential Diagnosis of Aphasia. J Speech Hear Res. 1962;5:349–69.
21. Kertesz A, Phipps J. Numerical taxonomy of aphasia. Brain Lang. 1977;4:1–10.
22. Clark C, Crockett DJ, Klonoff H. Factor analysis of the Porch Index of Communication Ability. Brain Lang. 1979;7:1–7.
23. Clark C, Crockett DJ, Klonoff H. Empirically derived groups in the assessment of recovery from aphasia. Brain Lang. 1979;7:240–51.
24. Hanson WR, Riege WH, Metter EJ, Inman VW. Factor-derived categories of chronic aphasia. Brain Lang. 1982;15:369–80.
25. Swinburn K, Porter G, Howard D. Comprehensive Aphasia Test. Hove: Psychology Press; 2004.
26. Goodglass H, Kaplan E. The assessment of aphasia and related disorders. 2nd ed. Philadelphia: Lea and Febiger; 1983.
27. Hula WD, Doyle PJ, Stone CA, Austermann Hula SN, Kellough S, Wambaugh JL, et al. The Aphasia Communication Outcome Measure (ACOM): dimensionality, item bank calibration, and initial validation. J Speech Lang Hear Res. 2015;58:906–19.
28. Butler RA, Ralph MAL, Woollams AM. Capturing multidimensionality in stroke aphasia: mapping principal behavioural components to neural structures. Brain. 2014;137:3248–66.
29• Mirman D, Chen Q, Zhang Y, Wang Z, Faseyitan OK, Coslett HB, et al. Neural organization of spoken language revealed by lesion-symptom mapping. Nat Commun. 2015;6:6762 **Applied PCA to a battery of language measures to identify four underlying components, and then used lesion-symptom mapping to show that each component was associated with damage to specific brain regions.**
30. Mirman D, Zhang Y, Wang Z, Coslett HB, Schwartz MF. The ins and outs of meaning: behavioral and neuroanatomical dissociation of semantically-driven word retrieval and multimodal semantic recognition in aphasia. Neuropsychologia. 2015;76:208–19.
31. Fridriksson J, Yourganov G, Bonilha L, Basilakos A, Den Ouden D-B, Rorden C. Revealing the dual streams of speech processing. Proc Natl Acad Sci U S A. 2016;113:15108–13.
32. Fridriksson J, den Ouden D-B, Hillis AE, Hickok G, Rorden C, Basilakos A, et al. Anatomy of aphasia revisited. Brain. 2018;141:848–62.
33. Halai AD, Woollams AM, Ralph MAL. 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. 2017;86:275–89.
34. Schwartz MF, Dell GS, Martin N, Gahl S, Sobel P. A case-series test of the interactive two-step model of lexical access: evidence from picture naming. J Mem Lang. 2006;54:228–64.
35. Wagenaar E, Snow C, Prins R. Spontaneous speech of aphasic patients: a psycholinguistic analysis. Brain Lang. 1975;2:281–303.
36. Vermeulen J, Bastiaanse R, Van Wageningen B. Spontaneous speech in aphasia: a correlational study. Brain Lang. 1989;36:252–74.
37. Casilio M, Rising K, Beeson PM, Bunton K, Wilson SM. Auditory-perceptual rating of connected speech in aphasia. Am J Speech Lang Pathol. 2019;28:550–68.
38. DeDe G, Caplan D. Factor analysis of aphasic syntactic comprehension disorders. Aphasiology. 2006;20:123–35.
39. Caplan D, Waters G, DeDe G, Michaud J, Reddy A. A study of syntactic processing in aphasia I: behavioral (psycholinguistic) aspects. Brain Lang. 2007;101:103–50.
40. Caplan D, Michaud J, Hufford R. Dissociations and associations of performance in syntactic comprehension in aphasia and their implications for the nature of aphasic deficits. Brain Lang. 2013;127:21–33.
41. Floyd FJ, Widaman KF. Factor analysis in the development and refinement of clinical assessment instruments. Psychol Assess. 1995;7:286–99.
42. Fabrigar LR, Wegener DT, MacCallum RC, Strahan EJ. Evaluating the use of exploratory factor analysis in psychological research. Psychol Methods. 1999;4:272–99.
43. Tabachnick BG, Fidell LS. Using multivariate statistics. 6th ed. Boston: Pearson; 2007.
44. Fergadiotis G, Wright HH, West TM. Measuring lexical diversity in narrative discourse of people with aphasia. Am J Speech Lang Pathol. 2013;22:S397–408.
45•. Walker GM, Hickok G, Fridriksson J. A cognitive psychometric model for assessment of picture naming abilities in aphasia. Psychol Assess. 2018;30:809–26 **Developed an innovative multidimensional cognitive psychometric model of picture-naming with theoretical and clinical utility, and provided preliminary validity evidence.**
46. Batchelder WH. Multinomial processing tree models and psychological assessment. Psychol Assess. 1998;10:331–44.
47. Erdfelder E, Auer T-S, Hilbig BE, Aßfalg A, Moshagen M, Nadarevic L. Multinomial processing tree models: a review of the literature. J Psychol. 2009;217:108–24.
48. Lord FM, Novick MR, Birnbaum A. Statistical theories of mental test scores. Oxford: Addison-Wesley; 1968.
49. Edelen MO, Reeve BB. Applying item response theory (IRT) modeling to questionnaire development, evaluation, and refinement. Qual Life Res. 2007;16:5–18.
50. del Toro CM, Bislick LP, Comer M, Velozo C, Romero S, Rothi LJG, et al. Development of a short form of the Boston Naming Test for individuals with aphasia. J Speech Lang Hear Res. 2011;54:1089–100.
51. Fergadiotis G, Kellough S, Hula WD. Item response theory modeling of the Philadelphia Naming Test. J Speech Lang Hear Res. 2015;58:865–77.
52. Mohr JP. Broca's area and Broca's aphasia. In: Whitaker H, Whitaker H, editors. Studies in neurolinguistics, vol. 1. New York: Academic Press; 1976.
53. Damasio H, Grabowski TJ, Tranel D, Hichwa RD, Damasio AR. A neural basis for lexical retrieval. Nature. 1996;380:499–505.
54. Bates E, Wilson SM, Saygin AP, Dick F, Sereno MI, Knight RT, et al. Voxel-based lesion-symptom mapping. Nat Neurosci. 2003;6:448–50.
55. Ashburner J, Friston KJ. Voxel-based morphometry–the methods. NeuroImage. 2000;11:805–21.
56. Schwartz MF, Kimberg DY, Walker GM, Faseyitan O, Brecher A, Dell GS, et al. Anterior temporal involvement in semantic word

retrieval: voxel-based lesion-symptom mapping evidence from aphasia. Brain. 2009;132:3411–27.
57. Wilson SM, Henry ML, Besbris M, Ogar JM, Dronkers NF, Jarrold W, et al. Connected speech production in three variants of primary progressive aphasia. Brain. 2010;133:2069–88.
58. Rorden C, Karnath HO. Using human brain lesions to infer function: a relic from a past era in the fMRI age? Nat Rev Neurosci. 2004;5:812–9.
59. Husain M, Nachev P. Space and the parietal cortex. Trends Cogn Sci. 2007;11:30–6.
60. Mah Y-H, Husain M, Rees G, Nachev P. Human brain lesion-deficit inference remapped. Brain. 2014;37:2522–31.
61. Herbet G, Lafargue G, Duffau H. Rethinking voxel-wise lesion-deficit analysis: a new challenge for computational neuropsychology. Cortex. 2015;64:413–6.
62• Xu T, Jha A, Nachev P. The dimensionalities of lesion-deficit mapping. Neuropsychologia. 2018;115:134–41 **Explained the negative consequences of univariate approaches to lesion-symptom mapping ignoring patterns of lesion distribution and distal effects of lesions, and advocated multivariate alternatives and systematic evaluation with synthetic data.**
63. Inoue K, Madhyastha T, Rudrauf D, Mehta S, Grabowski T. What affects detectability of lesion-deficit relationships in lesion studies? NeuroImage Clin. 2014;6:388–97.
64. Sperber C, Karnath H-O. Impact of correction factors in human brain lesion-behavior inference. Hum Brain Mapp. 2017;38:1692–701.
65. Wilson SM, Lam D, Babiak MC, Perry DW, Shih T, Hess CP, et al. Transient aphasias after left hemisphere resective surgery. J Neurosurg. 2015;123:581–93.
66. Bonilha L, Fridriksson J. Subcortical damage and white matter disconnection associated with non-fluent speech. Brain. 2009;132:e108.
67. Corbetta M, Ramsey L, Callejas A, Baldassarre A, Hacker CD, Siegel JS, et al. Common behavioral clusters and subcortical anatomy in stroke. Neuron. 2015;85:927–41.
68• Siegel JS, Ramsey LE, Snyder AZ, Metcalf NV, Chacko RV, Weinberger K, et al. Disruptions of network connectivity predict impairment in multiple behavioral domains after stroke. Proc Natl Acad Sci U S A. 2016;113:E4367–76 **Many post-stroke deficits, including aphasia, can be predicted based on measure of functional connectivity, especially reduced interhemispheric communication**.
69. Caplan D, Waters G, Kennedy D, Alpert N, Makris N, Dede G, et al. A study of syntactic processing in aphasia II: neurological aspects. Brain Lang. 2007;101:151–77.
70. Wilson SM, Ogar JM, Laluz V, Growdon M, Jang J, Glenn S, et al. Automated MRI-based classification of primary progressive aphasia variants. NeuroImage. 2009;47:1558–67.
71. Zhang Y, Kimberg DY, Coslett HB, Schwartz MF, Wang Z. Multivariate lesion-symptom mapping using support vector regression. Hum Brain Mapp. 2014;35:5861–76.
72. Pustina D, Coslett HB, Ungar L, Faseyitan OK, Medaglia JD, Avants B, et al. Enhanced estimations of post-stroke aphasia severity using stacked multimodal predictions. Hum Brain Mapp. 2017;38:5603–15.
73. Hope TMH, Seghier ML, Leff AP, Price CJ. Predicting outcome and recovery after stroke with lesions extracted from MRI images. NeuroImage Clin. 2013;2:424–33.
74. Chen R, Hillis AE, Pawlak M, Herskovits EH. Voxelwise Bayesian lesion-deficit analysis. NeuroImage. 2008;40:1633–42.
75. Vapnik V. Statistical learning theory. New York: Wiley; 1998.
76. Yourganov G, Smith KG, Fridriksson J, Rorden C. Predicting aphasia type from brain damage measured with structural MRI. Cortex. 2015;73:203–15.
77. Yourganov G, Fridriksson J, Rorden C, Gleichgerrcht E, Bonilha L. Multivariate connectome-based symptom mapping in post-stroke patients: networks supporting language and speech. J Neurosci. 2016;36:6668–79.
78. Saur D, Ronneberger O, Kümmerer D, Mader I, Weiller C, Klöppel S. Early functional magnetic resonance imaging activations predict language outcome after stroke. Brain. 2010;133:1252–64.
79. Xing S, Lacey EH, Skipper-Kallal LM, Jiang X, Harris-Love ML, Zeng J, et al. Right hemisphere grey matter structure and language outcomes in chronic left hemisphere stroke. Brain. 2015;139:227–41.
80. Lacey EH, Skipper-Kallal LM, Xing S, Fama ME, Turkeltaub PE. Mapping common aphasia assessments to underlying cognitive processes and their neural substrates. Neurorehabil Neural Repair. 2017;31:442–50.
81. DeMarco AT, Turkeltaub PE. A multivariate lesion symptom mapping toolbox and examination of lesion-volume biases and correction methods in lesion-symptom mapping. Hum Brain Mapp. 2018;39:4169–82.
82• Pustina D, Avants B, Faseyitan OK, Medaglia JD, Coslett HB. Improved accuracy of lesion to symptom mapping with multivariate sparse canonical correlations. Neuropsychologia. 2018;115:154–66 **Described a new approach to multivariate lesion-symptom mapping, and demonstrated that it outperformed a univariate approach on simulated and real aphasia data.**
83. Thye M, Mirman D. Relative contributions of lesion location and lesion size to predictions of varied language deficits in post-stroke aphasia. NeuroImage Clin. 2018;20:1129–38.
84. Wilson SM, DeMarco AT, Henry ML, Gesierich B, Babiak M, Miller BL, et al. Variable disruption of a syntactic processing network in primary progressive aphasia. Brain. 2016;139:2994–3006.
85. Griffis JC, Nenert R, Allendorfer JB, Szaflarski JP. Linking left hemispheric tissue preservation to fMRI language task activation in chronic stroke patients. Cortex. 2017;96:1–18.
86. Griffis JC, Nenert R, Allendorfer JB, Vannest J, Holland S, Dietz A, et al. The canonical semantic network supports residual language function in chronic post-stroke aphasia. Hum Brain Mapp. 2017;38:1636–58.
87. Skipper-Kallal LM, Lacey EH, Xing S, Turkeltaub PE. Functional activation independently contributes to naming ability and relates to lesion site in post-stroke aphasia: lesion site, activity, and naming in stroke. Hum Brain Mapp. 2017;38:2051–66.
88. Skipper-Kallal LM, Lacey EH, Xing S, Turkeltaub PE. Right hemisphere remapping of naming functions depends on lesion size and location in poststroke aphasia. Neural Plast. 2017;2017:8740353.
89. Ramsey LE, Siegel JS, Lang CE, Strube M, Shulman GL, Corbetta M. Behavioural clusters and predictors of performance during recovery from stroke. Nat Hum Behav. 2017;1:0038.
90. Siegel JS, Seitzman BA, Ramsey LE, Ortega M, Gordon EM, Dosenbach NUF, et al. Re-emergence of modular brain networks in stroke recovery. Cortex. 2018;101:44–59.
91. Bonilha L, Rorden C, Fridriksson J. Assessing the clinical effect of residual cortical disconnection after ischemic strokes. Stroke. 2014;45:988–93.
92. Marebwa BK, Fridriksson J, Yourganov G, Feenaughty L, Rorden C, Bonilha L. Chronic post-stroke aphasia severity is determined by fragmentation of residual white matter networks. Sci Rep. 2017;7:8188.
93. Seghier ML, Patel E, Prejawa S, Ramsden S, Selmer A, Lim L, et al. The PLORAS database: a data repository for predicting language outcome and recovery after stroke. NeuroImage. 2016;124:1208–12.
94. MacWhinney B, Fromm D. AphasiaBank as BigData. Semin Speech Lang. 2016;37:10–22.

95. Lukic S, Barbieri E, Wang X, Caplan D, Kiran S, Rapp B, et al. Right hemisphere grey matter volume and language functions in stroke aphasia. Neural Plast. 2017;2017:5601509.