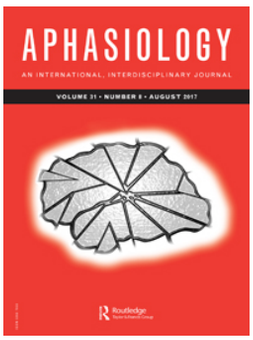

Aphasiology

ISSN: 0268-7038 (Print) 1464-5041 (Online) Journal homepage: http://www.tandfonline.com/loi/paph20

Rapid recovery from aphasia after infarction of Wernicke's area

Stephanie A. Yagata, Melodie Yen, Angelica McCarron, Alexa Bautista , Genevieve Lamair-Orosco & Stephen M. Wilson

To cite this article: Stephanie A. Yagata, Melodie Yen, Angelica McCarron, Alexa Bautista , Genevieve Lamair-Orosco & Stephen M. Wilson (2017) Rapid recovery from aphasia after infarction of Wernicke's area, Aphasiology, 31:8, 951-980, DOI: 10.1080/02687038.2016.1225276

To link to this article: http://dx.doi.org/10.1080/02687038.2016.1225276

APHASIOLOGY, 2017
VOL. 31, NO. 8, 951–980
https://doi.org/10.1080/02687038.2016.1225276

Routledge
Taylor & Francis Group

Rapid recovery from aphasia after infarction of Wernicke's area

Stephanie A. Yagata[a], Melodie Yen[a,b], Angelica McCarron[a], Alexa Bautista [a], Genevieve Lamair-Orosco[a] and Stephen M. Wilson [a,b,c]

[a]Department of Speech, Language, and Hearing Sciences, University of Arizona, Tucson, AZ, USA; [b]Department of Linguistics, University of Arizona, Tucson, AZ, USA; [c]Department of Neurology, University of Arizona, Tucson, AZ, USA

ABSTRACT

Background: Aphasia following infarction of Wernicke's area typically resolves to some extent over time. The nature of this recovery process and its time course have not been characterised in detail, especially in the acute/subacute period.

Aims: The goal of this study was to document recovery after infarction of Wernicke's area in detail in the first 3 months after stroke. Specifically, we aimed to address two questions about language recovery. First, which impaired language domains improve over time and which do not? Second, what is the time course of recovery?

Methods & Procedures: We used quantitative analysis of connected speech and a brief aphasia battery to document language recovery in two individuals with aphasia following infarction of the posterior superior temporal gyrus (STG). Speech samples were acquired daily between 2 and 16 days post stroke, and also at 1 month and 3 months. Speech samples were transcribed and coded using the CHAT system in order to quantify multiple language domains. A brief aphasia battery was also administered at a subset of five time points during the 3 months.

Outcomes & Results: Both patients showed substantial recovery of language function over this time period. Most, but not all, language domains showed improvements, including fluency, lexical access, phonological retrieval and encoding, and syntactic complexity. The time course of recovery was logarithmic, with the greatest gains taking place early in the course of recovery.

Conclusions: There is considerable potential for amelioration of language deficits when damage is relatively circumscribed to the posterior STG. Quantitative analysis of connected speech samples proved to be an effective, albeit time consuming, approach to tracking day-by-day recovery in the acute/subacute post-stroke period.

ARTICLE HISTORY
Received 12 April 2016
Accepted 3 August 2016

KEYWORDS
Stroke; aphasia; recovery; connected speech; Wernicke's area

Introduction

Wernicke's area, in the left posterior perisylvian region, is a key brain area for language (Wernicke, 1874). Indeed, many authors consider Wernicke's area to be the most fundamental and indispensable language region of the brain (Marie, 1906; Penfield & Roberts,

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

1959). Wernicke's area is generally defined as including the posterior superior temporal gyrus (STG), and depending on the definition, may include the adjacent supramarginal gyrus (SMG), middle temporal gyrus (MTG), and/or angular gyrus (Bogen & Bogen, 1976). In the acute stage, damage to Wernicke's area is associated with Wernicke's aphasia, characterised by impaired comprehension and repetition, and fluent paraphasic speech (Hillis et al., 2001).

Despite the apparently critical role of Wernicke's area in language processing, patients with infarction or resection of the posterior STG and adjacent structures often show considerable recovery of language function (Kertesz, Lau, & Polk, 1993; Laska, Hellblom, Murray, Kahan, & Von Arbin, 2001; Penfield & Roberts, 1959; Selnes, Knopman, Niccum, Rubens, & Larson, 1983; Selnes, Niccum, Knopman, & Rubens, 1984; Weiller et al., 1995). Recovery is highly variable (Kertesz & McCabe, 1977) and the extent of recovery depends on whether adjacent structures such as the SMG, angular gyrus, and MTG are involved (Kertesz & Benson, 1970; Kertesz et al., 1993; Metter et al., 1990; Naeser, Helm-Estabrooks, Haas, Auerbach, & Srinivasan, 1987; Selnes et al., 1983).

The nature and time course of the recovery process after Wernicke's area infarction have not been characterised in detail. Previous studies have investigated just one or two language measures such as single word comprehension (Selnes et al., 1984), sentence comprehension (Selnes et al., 1983), comprehension in general (Kertesz et al., 1993; Naeser et al., 1987), jargon in speech production (Kertesz & Benson, 1970), or overall aphasia severity (Kertesz et al., 1993; Kertesz & McCabe, 1977; Laska et al., 2001). Moreover, most studies have first assessed patients at an initial time point of 2 weeks to 1 month, meaning that recovery taking place prior to that point has not been documented. This early period constitutes an important missing piece of the recovery process, because in many cases, the greatest gains take place in the first few weeks after stroke (Basso, 1992; Pedersen, Stig Jørgensen, Nakayama, Raaschou, & Olsen, 1995).

However, there are practical challenges to studying recovery of language function in the immediate aftermath of a stroke. Patients' and their families' lives have just been turned upside down, and it is typically a traumatic and confusing time. Patients are transferred from emergency departments to intensive care units to acute care units. Some may be discharged home rather abruptly, within a day or two, if there are no physical deficits, while others are transferred to inpatient rehabilitation facilities. Patients with comprehension deficits are often not testable on any aphasia battery that requires following instructions and providing specific responses. Even if comprehension is spared, it is often not feasible to repeatedly administer a comprehensive aphasia battery such as the Boston Diagnostic Aphasia Examination (Goodglass, Kaplan, & Barresi, 2001). On the other hand, briefer aphasia screening tools that have been developed (e.g., Azuar et al., 2013; Flamand-Roze et al., 2011) do not quantify language in sufficient detail to fully capture the nature of recovery patterns.

An alternative approach is to assess recovery of language function through quantitative analysis of connected speech (Holland et al., 1985; MacWhinney, Fromm, Forbes, & Holland, 2011; Saffran, Berndt, & Schwartz, 1989; Wilson et al., 2010). Connected speech samples can potentially be obtained through conversation at the bedside in the early days following a stroke, and quantitative analysis of connected speech can provide rich information about multiple language domains. In a single individual with global aphasia, Holland et al. (1985) used connected speech samples obtained daily to document

recovery from global to anomic aphasia over a period of 2 weeks, suggesting that this approach can be an effective method for simultaneously collecting information about changes in multiple language domains at frequent time points. Several other studies have documented recovery from aphasia later in the subacute period based in part on longitudinal analyses of connected speech (Gandour, Marshall, Kim, & Neuburger, 1991; Kohn & Smith, 1994; Marshall, 1982).

The goal of the present study was to document recovery after infarction of Wernicke's area in detail in the first 3 months after stroke. Specifically, we aimed to address two questions about language recovery. First, which impaired language domains improve over time and which do not? Second, what is the time course of recovery?

Methods

We studied two individuals with aphasia following acute left middle cerebral artery stroke. We collected 5-minute conversational samples of connected speech daily between 2 days and 16 days post stroke, and then at 1 month and 3 months post stroke. We transcribed and coded each sample using CHAT (MacWhinney, 2000, 2012; MacWhinney et al., 2011) in order to analyse the speech samples for multiple measures of expressive language. We supplemented our analyses of connected speech with a brief aphasia battery at a subset of five time points during the 3 months.

Participants

Two individuals with aphasia due to acute left hemisphere ischemic stroke to Wernicke's area participated in the study. Beginning in the summer of 2014, we began to screen all patients seen by the stroke service at the University of Arizona Medical Center, in order to identify individuals with new onset aphasia for longitudinal investigation. These two patients were the first two patients recruited, and both had focal damage to Wernicke's area. Both were native English speakers, and neither had a history of any pre-existing impairments of speech or language.

Connected speech samples were also obtained from six healthy native English speakers in their 70s (four female; mean age 73.0 years; range 70–78 years), in order to establish normative ranges for connected speech measures. These participants were recruited after attending a talk on language and the brain at a community centre. They were all right-handed, native speakers of English, and neurologically normal. Their scores on the Mini Mental State Examination (Folstein, Folstein, & McHugh, 1975) ranged from 29 to 30.

The study was approved by the University of Arizona Institutional Review Board. All participants provided informed consent for the study and were modestly compensated for their participation.

Participant JI

JI (not his real initials), a right-handed male, experienced a stroke in the summer of 2014 and was first approached 2 days after the stroke. He was 76 years old. Prior to his stroke,

he had been working as a security guard. He had completed high school and 1 year of university-level education. He had no previous neurological history.

His family awoke to a commotion at 6:00 a.m., and observed JI fall. His family reported that he was unable to "find words". He was brought to the emergency department by his family at 6:45 a.m. Nursing staff described "garbled speech upon arrival". By 7:30 a.m., JI was reported to be following basic commands, but "only able to answer yes and no to questions". He was taken for clinical magnetic resonance imaging (MRI) at approximately 8:00 a.m. At 11:45 a.m., his physician noted the following in regard to the onset of JI's symptoms: "I heard him say, '2 a.m.', '5 a.m.', and '2–1'; he is unable to verify time of onset by nodding or writing". No other observations were made about JI's speech/language until the next afternoon, at which time his speech–language pathologist reported that he was reliably able to produce "yes", "no", and "let's try it". In the morning, 2 days post stroke, the same speech–language pathologist noted improved expressive communication, but observed that he correctly named only one of the three items on his breakfast tray.

We first approached JI 2 days post stroke. We explained the study to him using simplified language (e.g., active sentence structures; high-frequency, early-acquired vocabulary), and we ascertained that he understood the required components of informed consent by talking with him about the study, answering his questions, and asking him questions to verify comprehension of key concepts, e.g., "Can you stop at any time?". He consented to participate in the study, and a conversational speech sample and brief aphasia battery (see later) were acquired. At this time, JI presented with severe conduction aphasia, based on our clinical impression. He was fluent, with pervasive phonemic paraphasias, neologisms, word-finding difficulties, and attempts to self-correct that were mostly unsuccessful. Confrontation naming and repetition were severely impaired. Single word comprehension was largely preserved, but sentence comprehension was poor, likely secondary to verbal working memory deficits.

JI remained in acute care for 4 days after which he was transferred to an inpatient rehabilitation facility for 8 days. Following inpatient rehabilitation, he was discharged home. We obtained connected speech samples every day between 2 days and 16 days post stroke, and at 30 and 110 days post stroke. Brief aphasia batteries were obtained at five time points. A summary of the data acquired is shown in Table 1. JI continued to exhibit conduction aphasia over the course of the study, with severity decreasing markedly over time.

JI did not receive speech–language therapy for aphasia during acute care. However, he received treatment for aphasia from a speech–language pathologist during inpatient rehabilitation one to two times daily, targeting word retrieval strategies. Upon discharge home, he did not receive continued speech–language therapy.

On the morning of JI's stroke, MRI was acquired according to a standard stroke protocol, including DWI, FLAIR, T1-weighted, and GRE. Arterial spin labelling perfusion imaging was attempted, but was not completed due to lack of patient cooperation. At JI's 1-month follow-up visit, we acquired three-dimensional T1-weighted structural images, FLAIR, and DTI on a Siemens 3 T Skyra scanner.

All available images were coregistered to one another using SPM5 (Friston, 2007). The T1-MPRAGE image was transformed to MNI space using Unified Segmentation in SPM5 and the same transformation was applied to all other images. The lesion was

Table 1. Data acquired and connected speech sample characteristics.

	Time (days post stroke)	Location	Duration of patient speech analysed (seconds)	Words	Utterances	Aphasia battery	Imaging
JI	0 (stroke)	Intensive care/MRI					DWI
	2	Acute care	99	134	37	Yes	
	3	Acute care	253	312	68		
	4	Acute care	244	377	70		
	5	Acute care	275	692	106		
	6	Inpatient rehabilitation	248	433	92		
	7	Inpatient rehabilitation	320	562	106		
	8	Inpatient rehabilitation	246	394	78		
	9	Inpatient rehabilitation	331	622	91	Yes	
	10	Inpatient rehabilitation	284	517	90		
	11	Inpatient rehabilitation	263	534	76		
	12	Inpatient rehabilitation	351	738	86		
	13	Inpatient rehabilitation	263	584	90		
	14	Home	271	511	58		
	15	Home	278	694	72		
	16	Home	261	585	79	Yes	
	30	Outpatient MRI	219	459	62	Yes	T1-MPRAGE, DTI
	110	Home	245	560	76	Yes	
VG	−347	Intensive care/MRI					DTI
	−344	Acute care/MRI					T1-SPACE
	0 (stroke)	Intensive care/MRI					DWI, FLAIR
	2	Acute care	33	32	9	Yes	
	3	Acute care	220	227	38		
	4	Acute care	256	266	34		
	5	Acute care	260	168	31		
	6	Acute care	217	210	35		
	8	Inpatient rehabilitation	164	173	33		
	9	Inpatient rehabilitation	196	228	37	Yes	
	10	Inpatient rehabilitation	256	251	35		
	11	Inpatient rehabilitation	253	126	22		
	12	Inpatient rehabilitation	261	283	41		
	13	Inpatient rehabilitation	264	210	24		
	14	Inpatient rehabilitation	247	349	45		
	15	Inpatient rehabilitation	240	276	51		
	16	Inpatient rehabilitation	152	265	32	Yes	
	41	Home	158	333	55	Yes	
	92	Home	218	487	81	Yes	

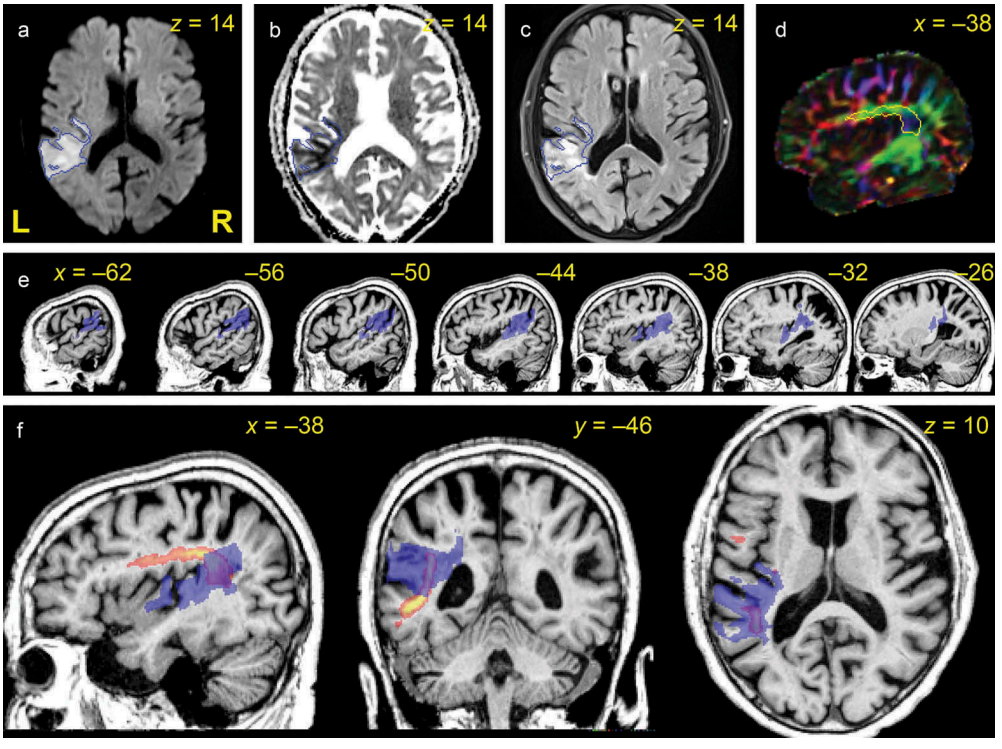

Figure 1. JI neuroimaging and lesion mapping. (a) lesion outlined on acute DWI (mean diffusivity); (b) lesion outlined on acute DWI (apparent diffusion coefficient); (c) lesion outlined on FLAIR at 1 month post-stroke time point; (d) arcuate fasciculus identified on DTI at 1 month post-stroke time point; (e) lesion shown on a series of sagittal images; (f) lesion shown on sagittal, coronal, and axial images with respect to the arcuate fasciculus, which is entirely disconnected.

drawn manually using ITK-SNAP 3.2.0 (Yushkevich et al., 2006), based on acute DWI (mean diffusivity, Figure 1(a) and apparent diffusion coefficient, Figure 1(b)) and 1-month FLAIR images (Figure 1(c)). The arcuate fasciculus was tracked using FSL 5.0 (Behrens, Berg, Jbabdi, Rushworth, & Woolrich, 2007), based on two masks: one in the anterior–posteriorly oriented fibres of the superior longitudinal fasciculus and the other in the dorsal–ventrally oriented fibres behind the Sylvian fissure (Figure 1(d)), following previously described procedures (Galantucci et al., 2011).

JI's lesion, shown in a series of sagittal images in Figure 1(e), was primarily concentrated in the posterior STG extending into the SMG and the angular gyrus. The lesion extended to the posterior insula. The ascending part of the posterior arcuate fasciculus was entirely destroyed by the lesion (Figure 1(f)).

Participant VG

VG (not her real initials), a right-handed female, experienced a stroke in the summer of 2014, and was first approached 1 day after the stroke. She was 85 years old. VG was a retired operating room nurse; and prior to her stroke, she had been an active member in her local senior centre. Her neurological history was significant for a previous right

hemisphere stroke in the basal ganglia and centrum semiovale in the summer of 2013. This previous stroke had not resulted in any speech or language deficits.

VG presented to the emergency department at approximately 6:30 a.m. on the day of her stroke after a neighbour found her at home "unable to speak clearly". The same neighbour had seen her the previous afternoon at which time she had not shown any signs or symptoms of stroke. Upon admission, nursing staff described her speech as "garbled". Around 7:00 a.m., a physician noted that she was able to answer some yes/no questions, but that she was not a reliable historian. VG was taken for clinical MRI at approximately 9:00 a.m. Nursing staff noted that at 10:30 a.m. she was "verbalizing understanding"; but around 4:00 p.m., her case manager reported that she was "unable to verbalize".

The next morning, one of the hospital's speech–language pathologists described VG as "aphasic, with comprehension appearing more intact than expression", and noted that her language was characterised by "frequent paraphasias and perseverations". The speech–language pathologist also noted some dysarthria. We approached VG later that day, at which time she was largely non-verbal, answering only "yes" or "no" to questions. At that point, we judged her comprehension too impaired for her to provide informed consent.

We approached her again the next day (2 days post stroke), at which point her comprehension had improved somewhat. As described earlier for JI, we explained the study to VG and talked with her about it until we were satisfied that she understood the required components of informed consent. She provided consent and we acquired a conversational sample and a brief aphasia battery (see later). At this time, VG presented with moderate Wernicke's aphasia, based on our clinical impression. She was fluent, but her connected speech was characterised by pervasive word-finding difficulties and paragrammatism, some jargon, and occasional semantic and phonemic paraphasias. Confrontation naming and repetition were severely impaired. Her comprehension was impaired at the single word and sentence level. She also exhibited moderate dysarthria of mixed spastic/hypokinetic type.

VG remained in acute care for 6 days following her stroke, and was transferred to an inpatient rehabilitation facility on day 7. She remained in inpatient rehabilitation for 16 days, at which time she was discharged home. We obtained connected speech samples every day between 2 days and 16 days post, except for day 7 on which she was exhausted due to the transfer to the rehabilitation facility. We also obtained samples at 41 and 92 days post stroke. Brief aphasia batteries were obtained at five time points. A summary of the data acquired is shown in Table 1. Over the course of the study, VG's language impairment evolved from Wernicke's aphasia to conduction aphasia, as her comprehension and word finding improved. Her dysarthria gradually reduced in severity from moderate to mild–moderate by 3 months post stroke.

VG was evaluated but not treated for aphasia during acute care, and she received speech–language therapy for aphasia one to two times daily during her period of inpatient rehabilitation, focusing on word retrieval strategies. After she returned home, VG received occupational therapy, but no speech–language therapy.

On the morning of her stroke, MR images were acquired according to a standard stroke protocol, including DWI, FLAIR, T1-weighted, T2-weighted, and GRE. Dynamic susceptibility contrast perfusion imaging was performed but was non-diagnostic due

to a technical issue. Although there were no three-dimensional structural images or DTI acquired at this time, these images were available from approximately 1 year prior, having been acquired after the previous right subcortical stroke. The T1-weighted structural image was of low quality due to motion artefact; however, it was adequate for visualisation of key structures.

The images were processed as described for JI, except that the acute FLAIR was used along with DWI for lesion delineation, since the acute imaging was acquired sufficiently late that the lesion appeared hyperintense on FLAIR. Lesion delineation is shown in Figure 2(a–c) and fibre tracking in Figure 2(d).

VG's lesion, shown in a series of sagittal slices in Figure 2(e), encompassed much of the left posterior STG and posterior superior temporal sulcus (STS) as well as portions of the MTG and SMG. The lesion extended medially, effectively dissecting the ascending part of the posterior arcuate fasciculus (Figure 1(f)). There were also lesions to the left basal ganglia and posterior insula, which extended into external capsule fibres. In combination with the previous lesion of the right basal ganglia, these lesions were presumably responsible for VG's motor speech deficits.

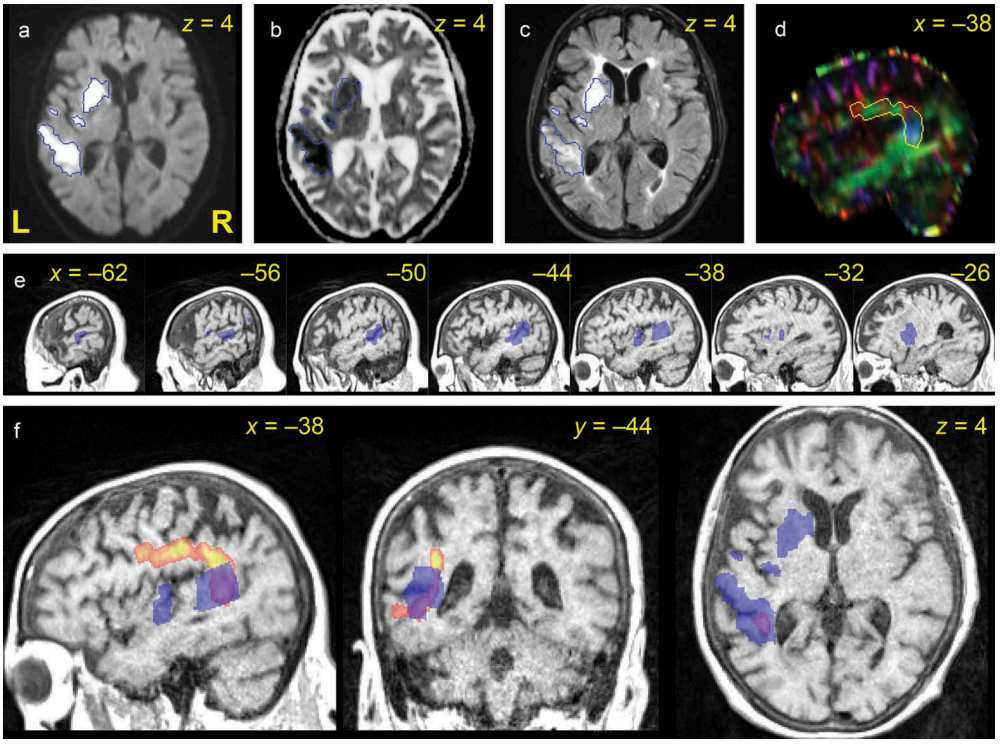

Figure 2. VG neuroimaging and lesion mapping. (a) lesion outlined on acute DWI (mean diffusivity); (b) lesion outlined on acute DWI (apparent diffusion coefficient); (c) lesion outlined on acute FLAIR; (d) arcuate fasciculus identified on DTI; (e) lesion shown on a series of sagittal images; (f) lesion shown on sagittal, coronal, and axial images with respect to the arcuate fasciculus, which is entirely disconnected.

Connected speech samples

Elicitation and recording of connected speech samples

Connected speech samples were elicited by the first author (SAY, a second-year masters student in Speech–Language Pathology) based on predetermined conversational prompts (Table 2). A different prompt was used each session, in order to elicit more natural conversational speech (rather than repeating the same questions, day after day). When necessary, more than one prompt was used in a session.

We aimed to collect at least 5 minutes of conversation, although a few sessions, including both patients' 2 days post time points, and some of the later time points that included aphasia batteries, were somewhat shorter than this because we wanted to minimise our imposition on the patients. When more than 5 minutes were acquired, only the first 5 minutes were analysed. From this 5-minute sample (or less, if less than 5 minutes was acquired), the sample duration (Table 1) was calculated as the total time that the patient was speaking (including pauses). Excluded from the sample duration were examiner speech, pauses between examiner speech and patient speech, patients' direct responses to questions, and any interruptions.

All conversations were recorded using a Sanken lavalier microphone (COS-11D) and a Canon camcorder (Vixia HF S20).

Transcription and coding of connected speech samples

Transcription and coding of the connected speech samples followed the CHAT system (MacWhinney, 2000, 2012), including newer codes that were introduced for quantifying aphasic speech (MacWhinney et al., 2011). CHAT provides a standardised way to represent naturalistic speech samples in aphasia with the means to capture phonological, lexical, and morphosyntactic disturbances. Speech samples were transcribed using EUDICO Linguistic Annotator (ELAN; Wittenburg, Brugman, Russel, Klassmann, & Sloetjes, 2006) in order to encode precise timing information. All speech samples were transcribed in standard English orthography except for the use of International Phonetic

Table 2. Prompts for elicitation of connected speech samples.

"Tell me about …"	what happened when you had your stroke
	what you did yesterday
	the best trip you ever took
	the worst trip you ever took
	your favourite holiday as a child
	a happy childhood memory
	your worst childhood memory
	how you met your husband/wife/partner
	when you got married
	when you had your first child
	your first job
	when you retired
	a time you were really scared/embarrassed/angry
Example prompts, if no response in approximately 10 seconds:	"Do you remember a trip you took?"
	"Tell me about your favourite Thanksgiving/ Christmas".
	"Tell me about your wedding".

Alphabet (IPA) to transcribe phonemic paraphasias and neologisms. Each sample was transcribed and coded by either SAY, MY (a first-year Ph.D. student in Linguistics) or GL-O (a fourth-year undergraduate major in Speech, Language and Hearing Sciences), all of whom were trained in the lab of SMW, a researcher with experience in the analysis of connected speech in aphasia (Wilson et al., 2010). All samples were then comprehensively checked and edited by SAY to maximise consistency of transcription and coding.

Utterance boundaries were determined based on the principles described by Saffran et al. (1989). Errors were coded at the word level and/or at the utterance level. Word-level errors were classified as phonological errors, semantic errors, neologisms, formal lexical device errors, morphological errors, or omissions (MacWhinney, 2000, 2012). Word-level errors were counted regardless of whether or not they were retraced and/or corrected; and if errors were repeated, they were counted each time. Five utterance-level error codes defined in the CHAT system were used to code grammatical errors (agrammatic or paragrammatic), empty speech, circumlocution, jargon, and perseverative utterances (the last two of these categories were infrequent so were not used in our analyses). We defined three additional codes at the utterance level: semantically anomalous utterances, non-sentence utterances (e.g., isolated noun phrases), and embeddings (see definitions later). Utterances could meet the criteria for multiple utterance-level codes. To avoid double-counting errors, the utterance-level codes for grammatical errors and semantically anomalous utterances were not used when the errors in question could be captured by word-level codes.

After transcription and coding, each morpheme was tagged for part of speech using the program *mor* (MacWhinney, 2000, 2012), then manually disambiguated.

We analysed the coded and tagged transcriptions by calculating measures (described later) to capture change over time in four broad language domains: (1) speech rate and speech sound errors; (2) lexical content; (3) morphosyntax and complexity; and (4) disruptions to fluency. An in-house MATLAB program was used to derive all measures from the coded transcripts and to carry out statistical analyses.

Speech rate and speech sound errors

Speech rate in words per minute was calculated as the total number of words produced divided by the time spent speaking. All real words and neologisms were counted, whereas fillers, false starts, unintelligible speech, words that were subsequently retraced, and direct responses to examiner questions were excluded. The time spent speaking was the duration of the speech sample, minus examiner speech, pauses between examiner speech and patient speech, direct responses to examiner questions, and interruptions.

Phonological errors were classified as phonemic paraphasias or neologisms. Phonemic paraphasias were defined as a substitution, omission, or deletion of the onset, nucleus, or coda in one syllable of the target word (MacWhinney, 2000, 2012). Neologisms were defined as words with unknown targets, or when the target was known, words that contained more than one substitution, omission, or deletion. In calculating the measure of phonological errors, neologisms were each weighted as two as they contained at least two errors. This measure, along with many other measures described later, was calculated per 100 words in order to normalise for sample length.

One patient, VG, had significant dysarthria that changed across the 3-month period of the study. Dysarthric errors were not individually coded in the samples. Instead, to

capture change in motor speech over time, five graduate students in speech–language pathology listened to each of VG's speech samples presented in a pseudorandomised order and rated them for dysarthric features. The students were unfamiliar with VG prior to rating the samples. They were presented only with audio recordings, and not with video, in order to avoid visual cues as to the patient's stage of recovery. The five students rated each sample for seven features: prosody, rate, resonance, articulation, voice quality, respiration, and overall intelligibility. Each feature was rated on a 7-point scale: 0 = no impairment; 1 = mild; 2 = mild–moderate; 3 = moderate; 4 = moderate–severe; 5 = severe; 6 = mute. The features and rating scale were loosely based on Darley, Aronson, and Brown (1969) and Duffy (2013). All students had completed coursework on motor speech disorders in which numerous dysarthric speech samples had been rated on similar scales, and all had some experience evaluating clients with motor speech disorders. A mean dysarthria score for each sample was computed by averaging scores across all five students and all seven features.

Lexical content

A composite measure of impaired lexical access was calculated by adding the counts of word-level semantic errors, neologisms with unknown targets, incomplete utterances (those that trailed off, almost always due to word-finding problems), utterances containing empty speech, utterances containing circumlocution, and semantically anomalous utterances. The utterance-level code for semantically anomalous utterances was used when utterances were semantically inappropriate for the context, but where the semantic anomaly could not be pinned down to any word-level problem.

The proportion of closed class words was calculated by dividing the number of closed class words by the total number of open and closed class words. Similarly, the proportion of verbs was calculated by dividing the number of verbs by the number of verbs and nouns. These proportions are indirect measures of agrammatism (which results in reduced proportions of closed class words and verbs) and empty speech (which results in elevated proportions of closed class words and verbs) (Wilson et al., 2010).

Morphosyntax and complexity

Mean length of utterance in words, a commonly used metric of syntactic complexity, was calculated by dividing the total number of words in each sample (excluding fillers, false starts, unintelligible speech, and retraced words) by the number of utterances.

Bound morphemes were counted as a measure of morphological complexity, including inflectional and derivational morphemes. Incorrectly used bound morphemes were included in the count as they demonstrated attempts to use morphology.

Utterances containing embeddings were coded as a measure of syntactic complexity. Embeddings were required to contain a subject or a finite verb form. If an utterance contained multiple embeddings, each was counted. Embeddings were counted even if utterances were not grammatical, but the defining criteria for embedding needed to be clearly present.

A composite measure of morphosyntactic errors was calculated by adding the counts of word-level morphological errors, errors in formal lexical devices, omitted words (where these could be determined with reasonable confidence), utterances that were

not complete sentences (e.g., noun phrases, verb phrases, etc.), and agrammatic or paragrammatic utterances. Because the utterance-level code for agrammatic/paragrammatic speech was used only when utterances contained more than one word-level morphosyntactic error, it was given double weight in the composite measure.

Disruptions to fluency

In this category, we included three further phenomena that contribute to an overall impression of non-fluent speech: filled pauses (e.g., "uh", "um", and "er"), false starts, and retracings. False starts were defined as partial words, i.e., words that were abandoned, usually after just one or a few phonemes had been produced. Retracings were defined as sequences of one or more complete words (i.e., not solely a filler or a false start), which were made redundant by subsequent repetitions, amendments, elaborations, or alternative expressions.

Brief aphasia battery

At a subset of five time points (Table 1), each patient's speech and language was further assessed using a brief in-house aphasia battery, which was designed to be quicker to administer than commonly used batteries and so more appropriate for the acute and subacute time points.

The battery quantified expressive and receptive function through five language tasks: (1) confrontation naming; (2) single word comprehension; (3) sentence comprehension; (4) repetition of words and sentences; and (5) reading aloud of words and sentences. The confrontation naming portion consisted of 12 items, varying in word length and frequency, with high name agreement and monomorphemic targets. Pictures were colourised versions of Snodgrass and Vanderwert's line drawings (Rossion & Pourtois, 2004). The single word comprehension task comprised 16 items, 4-alternative forced choice auditory word–picture matching. There were 12 items with semantic, phonemic, and unrelated foils, and 4 difficult low-frequency items (e.g., "asparagus", "artichoke") with 3 closely related semantic foils each. The sentence comprehension task consisted of 12 items, with 4-alternative forced choice sentence–picture matching. The stimuli were similar to those created for a previous study (Wilson et al., 2010), and varied in terms of length and syntactic complexity. The repetition and reading tasks involved single words of varying length, frequency, regularity, and morphological complexity (40%), sentences of varying length and word frequency including items loaded on function words (40%) and pseudowords of varying length (20%).

The battery was administered by SAY, AM (a third-year undergraduate major in Speech, Language, and Hearing Sciences) or AB (a fourth-year undergraduate major in Speech, Language, and Hearing Sciences), all of whom were trained by SMW, who designed the battery. All sessions were videotaped, then transcribed and scored by AM or AB. Correct responses scored one point each, except where they were delayed, self-corrected, or where repetitions were requested, in which case they scored half a point (as in the Comprehensive Aphasia Test; Swinburn, Porter, & Howard, 2004).

Statistical analysis

To identify overall patterns of change and the relationships of the language variables to one another, principal components analysis (PCA) was performed based on correlations (i.e., the variables were normalised because they had different scales). The dysarthria variable was not included for VG in order to make her PCA results comparable to those for JI (but including the dysarthria variable did not significantly change the results).

For each language variable described earlier, and for each patient, we assessed whether the variable showed a significant change over time. Linear models of language measures by time generally resulted in skewed residuals supporting the necessity of log-transforming the independent variable. Moreover, previous research has shown that recovery from aphasia in the acute/subacute stages after a stroke is characterised by a "decelerating" curve, with the greatest gains made early, and the pace of recovery slowing over time (Basso, 1992; Pedersen et al., 1995). Therefore, we assumed a logarithmic time course; and for each variable in each patient, we fit the equation $y = a.\log(t) + b$, where y is the language variable in question, t is time (days past stroke), and a and b are coefficients.

Since the direction of change was assumed to be positive for measures associated with successful language use (e.g., mean length of utterance) and negative for measures measuring erroneous language use (e.g., phonological errors), we computed one-tailed p-values accordingly for most measures. Two-tailed p-values were computed for the proportion of closed class words, and the proportion of verbs, since these quantities can be perturbed in either direction in aphasia (Wilson et al., 2010).

Because we investigated the time course of numerous language variables in two different patients, the issue of multiple comparisons must be considered. For JI, 11 of 17 statistical tests were significant at $p \leq 0.05$. With a false discovery rate of $q = 0.1$ (Benjamini & Hochberg, 1995), all 11 would remain significant. For VG, 9 of 18 statistical tests were significant at $p \leq 0.05$. With a false discovery rate of $q = 0.1$, all nine would remain significant. Because all p-values ≤0.05 remain significant, for the sake of simplicity, we present uncorrected p-values for the remainder of the paper.

Correlation coefficients are described as little to none ($r < 0.1$), weak ($0.1 \leq r < 0.3$), moderate ($0.3 \leq r < 0.5$), or strong ($r \geq 0.5$), independent of statistical significance (Cohen, 1988).

Results

Principal components analysis

All changes over time in all language measures are summarised in Table 3.

PCA was performed for each patient to identify overall patterns of change and the relationships between the language variables. For each patient, scree plots revealed that the first principal component accounted for a large portion of the variance: 48% in JI and 40% in VG.

For each patient, the first principal component tracked recovery over time with a logarithmic time course (JI: strong correlation between first principal component score and log-transformed time, $r = 0.81$, $p < 0.0001$, Figure 3(a); VG: strong correlation,

Table 3. Changes in language measures.

	Patient			
	JI		VG	
Measure	*r*	*p*	*r*	*p*
First principal component	0.81	<0.0001***	0.66	0.0029**
Speech rate and speech sound errors				
Words per minute	0.64	0.0030**	0.73	0.0006***
Phonological errors (phw)	−0.72	0.0005***	−0.39	0.067
Dysarthria	*N*/A		−0.55	0.0129*
Lexical content				
Closed class words (proportion)	−0.77	0.0003***	0.52	0.037*
Verbs (proportion)	−0.49	0.048*	0.36	0.17
Impaired lexical access (phw)	−0.83	<0.0001***	−0.43	0.048
Morphosyntax and complexity				
Mean length of utterance (words)	0.65	0.0024**	0.28	0.15
Bound morphemes (phw)	0.37	0.070	0.19	0.24
Embeddings (phw)	0.60	0.0058**	0.54	0.0153*
Morphosyntactic errors (phw)	−0.08	0.38	−0.42	0.0537
Other disruptions to fluency				
Filled pauses (phw)	−0.34	0.091	−0.47	0.0321*
False starts (phw)	−0.22	0.20	−0.32	0.11
Retracings (phw)	−0.58	0.007**	−0.29	0.14
Aphasia battery				
Confrontation naming	0.96	0.0045**	0.93	0.012*
Single word comprehension	0.71	0.090	0.34	0.29
Sentence comprehension	0.51	0.19	0.85	0.034*
Repetition	0.98	0.0014**	0.93	0.010*
Reading aloud	0.94	0.0079**	0.73	0.082

*$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$. phw = per 100 words.

$r = 0.66$, $p = 0.0029$, Figure 3(b)). The factor loadings of these first principal components were similar for the two patients, with a few notable exceptions (Figure 3(c)). In general, the first principal components were positively weighted on variables associated with successful language use and negatively weighted on variables measuring erroneous language use. The exceptions to this general pattern were the proportions of closed class words and verbs, which changed in opposite directions in the two patients, and morphosyntactic errors, which changed over time only in VG and not in JI.

The second principal components did not change over time in either patient (JI: $p = 0.57$; VG: $p = 0.96$), suggesting that only the first principal components were meaningful.

Speech rate and speech sound errors

Speech rate in words per minute increased over time for both patients (JI: strong correlation, $r = 0.64$, $p = 0.0030$, Figure 4(a); VG: strong correlation, $r = 0.73$, $p = 0.0006$, Figure 4(b)). These increases in words per minute reflected more rapid production of individual words, decreased pauses within and between utterances, and especially for JI, decreases in the number of retracings (see later) across the recovery period. By 3 months post stroke, both patients' speech rates were within in the control range, which is indicated by the grey-shaded area in Figure 4 and subsequent figures.

The prevalence of phonological errors decreased over time in both patients (JI: strong correlation, $r = -0.72$, $p = 0.0005$, Figure 4(c); VG: moderate marginally significant correlation, $r = -0.39$, $p = 0.067$, Figure 4(d)). Examples are provided later. Earlier samples

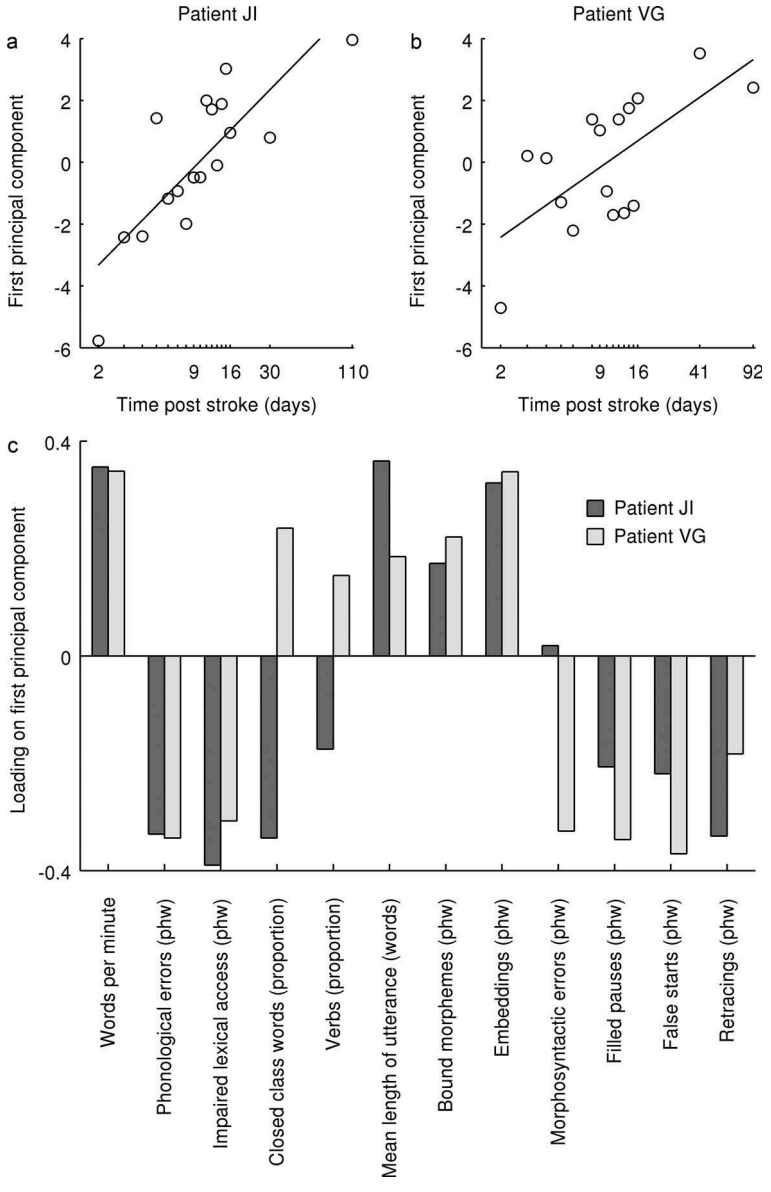

Figure 3. Principal components analysis (PCA) of language measures over time. (a) Change over time in the first principal component in JI. (b) Change over time in the first principal component in VG. (c) Loading of 12 language variables on the first principal component for each of the two patients.

were characterised by the presence of pervasive phonemic paraphasias and neologisms (1a–c, 2a,b), which reduced in frequency and in terms of deviance from intended targets over the course of recovery (1d,e, 2c–e). The control participants made negligible numbers of phonological errors (two of the six controls made a single error each).

In this paper, the following conventions apply to all examples: (i) examples are brief excerpts from each patient's speech samples, selected as representative of the patterns

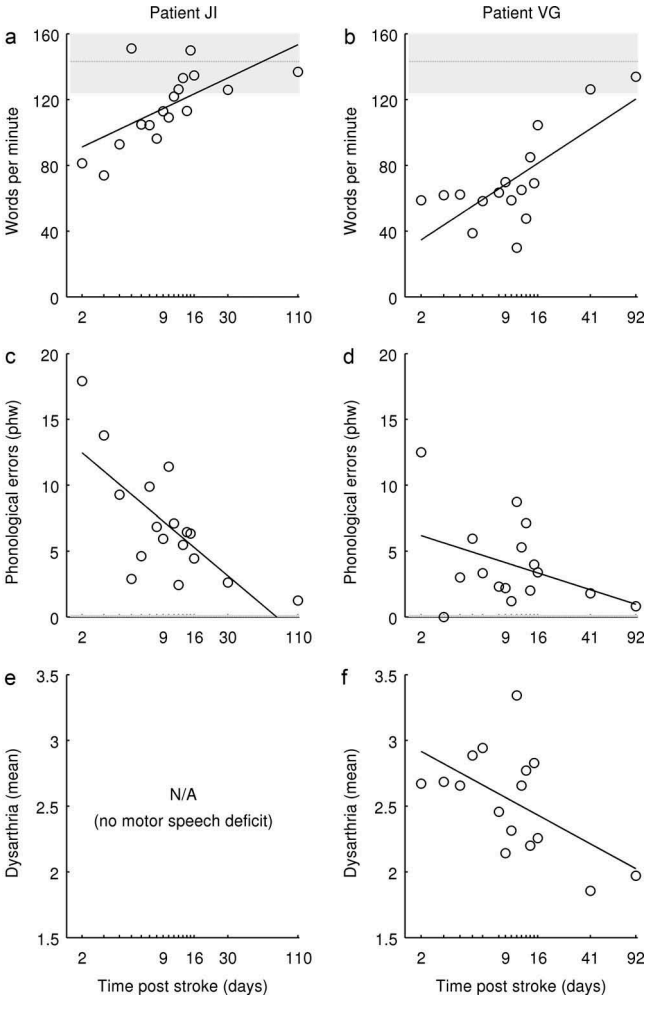

Figure 4. Change over time in speech rate and speech sound errors. (a, b) Words per minute. Data for the six control participants are shown (dotted grey line = mean; light grey-shaded area = one standard deviation). Note that the *x* axis (time post stroke) is shown with a logarithmic scale. (c, d) Phonological errors. phw = per 100 words. Note that controls made negligible numbers of phonological errors. (e) JI had no motor speech deficit, so dysarthria was not rated. (f) Dysarthria in VG.

being described; (ii) phonological errors are transcribed in IPA and contained in brackets; (iii) intended targets are represented in parentheses with an equal sign; (iv) omitted words are indicated by a Ø followed by the target word; (v) retraced sequences are surrounded by angled brackets; (vi) false starts are transcribed in IPA with a hyphen to indicate that the word was incomplete; (vii) pauses greater than 1 second are indicated with (.); (viii) incomplete utterances are indicated with ellipses; (ix) participant initials, days post stroke, and contextual information are included in parentheses following each example.

(1) (a) I fell out went this way and I [slɪmɪrəd] (=slipped) a θ- [θiɔɹəbət] (=unknown). (JI, 2 days post, about his stroke)

(b) <I [dɪdəl] (=did)> I uh uh did a lot of uh oh work with the uh [neɪməˌ] (=unknown). and then we'd [teɪ] (=take) over to here and over here. and we'd take this. and there we got uh all uh [pʊɾeɪ] (=unknown) and [pəɾeɪɹə] (=unknown). (JI, 3 days post, about his first job)

(c) yeah well <[suplub] (=seafood) I just like> I like uh sipl- uh <[sipɹplud] (=seafood)> pɹu- pɹub- [sipɹrub] (=seafood). (JI, 8 days post, about his favourite food)

(d) and everything and my wipe (=wife) s- she locked (=liked) all the people and I locked (=liked) the guy <it just seemed> I think that was probably the best one. (JI, 11 days post, about his favourite job)

(e) <and so> and I wanna go <and uh for more> for eight more [smʌnθs] (=months) for <the> [baɪ] (=my) echo. (JI, 110 days post, about a doctor's appointment)

(2) (a) what do you wanna know? well (.) not sure. well I do get <a> (.) a (.) [kæ̃səm̩] (=unknown) (.) um (.) canes (=unknown). mate (=unknown) uh (.) heart s- (.) um (.) b- b- <[blæs] (=unknown)> [blæsəz] (=unknown) a bunch of stuff. (VG, 2 days post, about her stroke)

(b) <you> you can (.) make a (.) k- <[kæ̃ʔʌ] (=unknown)> (.) [kænləs] (=unknown) (.) əwə- æ̃wɪ- and this (.) outfit (=unknown). (VG, 5 days post, about her activities)

(c) I s- sang. (.) I was a (.) æl- <[ældoʊdɚ] (=alto)> [ælddoʊɚ] (=alto). (VG, 11 days post, about her high school life)

(d) and so I <had a> uh (.) um (.) um (.) had some [taɪn] (=time) away to g- get away from work. (.) so I caught <a a a <[bɜ˞s] (=bus)> a b-> a bus. (VG, 15 days post, about a trip to Canada)

(e) but I'm not a therapist. so how do I know what that's all about? (.) and I just [əksɛkt] (=accept) it and go on. (VG, 41 days post, about her therapy)

Dysarthria was a notable feature of VG's speech. Based on the evaluations of five raters blinded to time point, the severity of her dysarthria decreased over time (strong correlation, $r = -0.55$, $p = 0.013$, Figure 4(f)). This overall improvement was driven by significant improvements in prosody ($r = -0.74$, $p = 0.0005$), rate ($r = -0.59$, $p = 0.0076$), respiration ($r = -0.58$, $p = 0.0092$), and a marginal improvement in intelligibility ($r = -0.34$, $p = 0.087$), whereas articulation ($r = -0.04$, $p = 0.44$) and voice quality ($r = -0.26$, $p = 0.16$) showed no evidence of change over time. JI had no motor speech deficits, so dysarthria was not rated for him.

Lexical content

Lexical access improved over time for both patients, as evidenced by reductions in the composite measure of impaired lexical access (JI: strong correlation, $r = -0.83$, $p < 0.0001$, Figure 5(a); VG: moderate correlation, $r = -0.43$, $p = 0.048$, Figure 5(b)). This measure summed across five phenomena that are indicative of impaired lexical access, all five of which were evident in both patients: word-level semantic paraphasias (e.g., "caught" for "picked" in 3c, "Christmas" for "fourth of July" in 4a), incomplete utterances (e.g., "you can hit the red" in (3d) where the final noun "button" is missing),

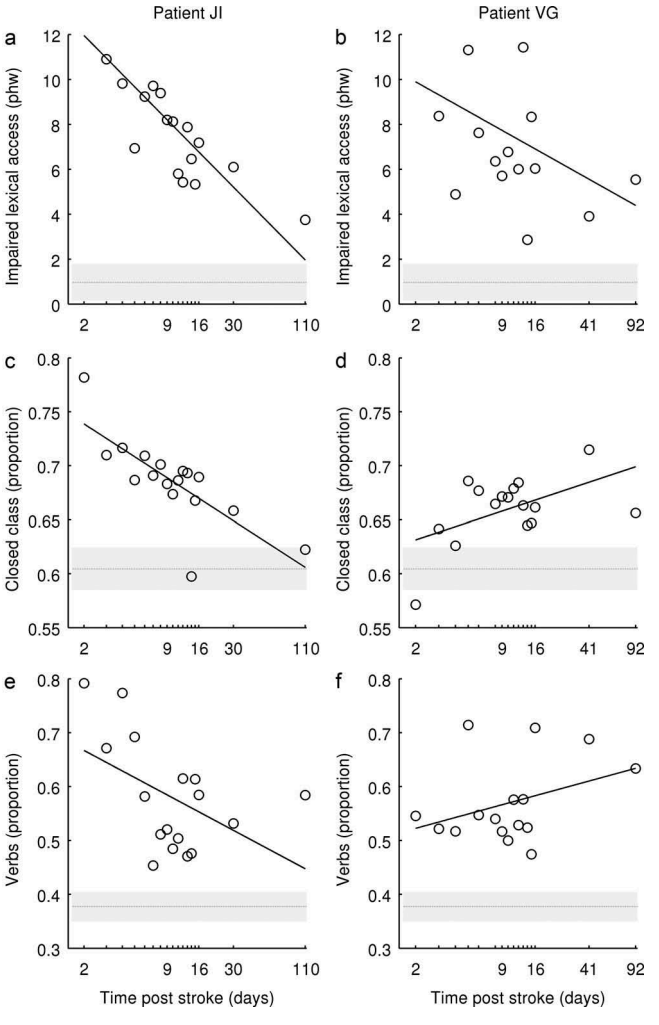

Figure 5. Change over time in measures of lexical content. (a, b) Composite measure of impaired lexical access, reflecting word-level semantic paraphasias, neologisms with unknown targets, incomplete utterances, empty speech, circumlocution, and semantically anomalous utterances. phw = per 100 words. (c, d) Proportion of closed class words (out of all words). (e, f) Proportion of verbs (out of verbs and nouns).

empty speech (3a–c, 4b), circumlocution (4b), and semantically anomalous utterances (4c). Imperfect lexical access is not necessarily abnormal: five of the six healthy control participants had at least one instance of impaired lexical access in their samples. However, both patients were still outside the range of healthy controls by 3 months post stroke.

(3) (a) and then he helped him um (.) <[pi] (=put) the uh the uh uh> put the um uh (.) thing together. and they uh … oh heavens so then … okay so then we got <all k-> all [gæd] (=good). (JI, 4 days post, about getting ready for his wedding)

(b) so that was that. but that said but that was when he was not the first time [ðæs] (=this) one but the second time. (JI, 7 days post, about a scary childhood memory)
(c) okay we got those there. and then <we we d-> we <uh kind uh kit (=picked) uh caught (=picked) uh> picked up some supplies and some food. and just [beʔ] (=unknown). we did uh a lot there. (JI, 14 days post, about running errands)
(d) just give it. you can hit the red … here I'll get it. if you don't know how to work it I'll get it. I just turned it down. I just wasn't even paying attention to it but if you want it there we go. (JI, 110 days post, about turning off the television)

(4) (a) we had to wait for the (.) Christmas (=fourth of July) (.) and then we buried Ø after. (VG, 3 days post, about her mother's funeral)
(b) and the red ones is every now and then. but we have a lot of them. yeah. and I also got a uh <oh the red one (.) bou- [boʊgəngeɪjə] (=bougainvillea)> [geɪjə] (=bougainvillea) whatever Ø <call> call the things. (VG, 14 days post, about flowers in her garden)
(c) uh the Dad uh knew we were having trouble with that house down there. (.) what happened is when they uh came he put the name in <his> d- his outfit (=unknown). (VG, 92 days post, about building a house)

There were only two measures that changed in opposite directions in the two patients: the proportion of closed class words and the proportion of verbs (out of nouns and verbs). In JI, both of these measures decreased over time and approached the healthy control range by 3 months post stroke (closed class words: strong correlation, $r = -0.77$, $p = 0.0003$, Figure 5(c); verbs: moderate correlation, $r = -0.49$, $p = 0.048$, Figure 5(e)), as his speech became progressively less empty (see examples in (3) earlier).

In contrast, in VG, both of these measures increased over time (closed class words: strong correlation, $r = 0.52$, $p = 0.037$, Figure 5(d); proportion of verbs: moderate but non-significant correlation, $r = 0.36$, $p = 0.17$, Figure 5(f)). The significance of the first of these measures was driven by the first time point, during which VG's speech output was minimal and agrammatic, resulting in a below-normal proportion of closed class words. When the first time point was excluded, there was no significant subsequent change in VG's proportion of closed class words ($r = 0.31$, $p = 0.27$). The trends to increased proportions of closed class words and verbs should not be interpreted as indicative of recovery, since these measures were higher than for control participants.

Morphosyntax and complexity

The morphological and syntactic complexity of both patients' utterances increased over time, as shown in the examples in (5) and (6). This was reflected in positive trajectories of three variables. Words per utterance increased in both patients (JI: strong correlation, $r = 0.65$, $p = 0.0024$, Figure 6(a), VG: weak non-significant correlation, $r = 0.28$, $p = 0.15$, Figure 6(b)). The number of bound morphemes per 100 words, a measure of

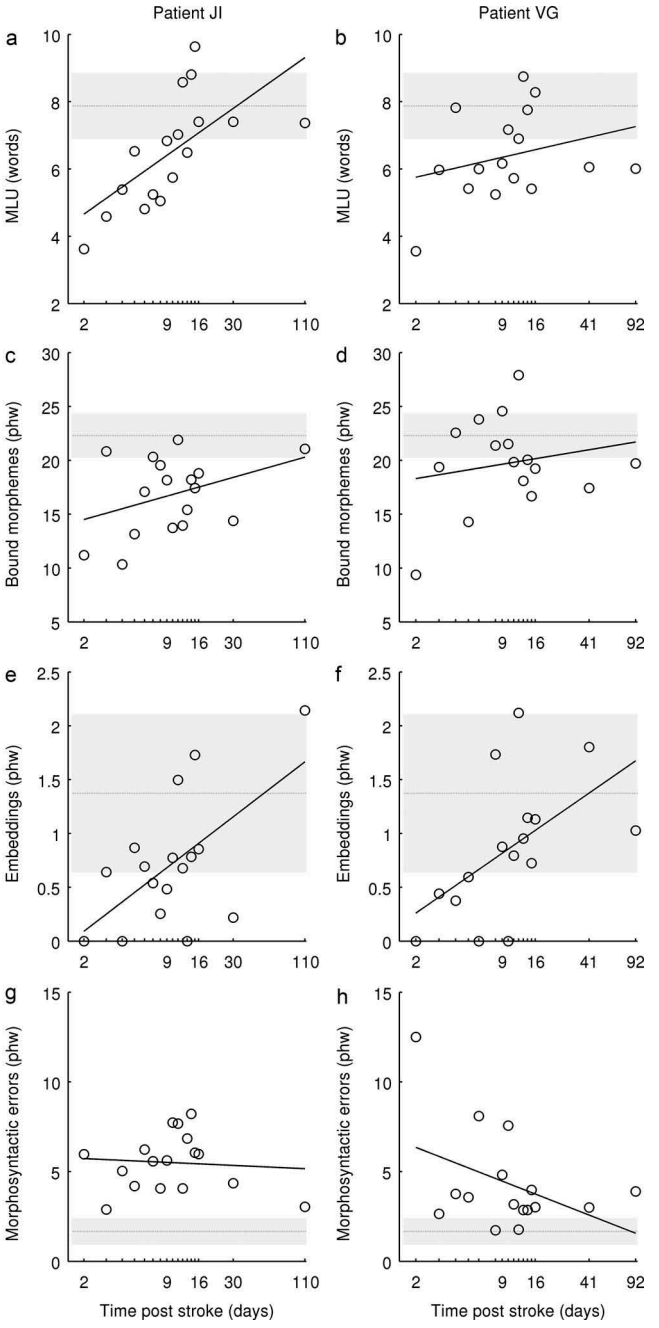

Figure 6. Change over time in measures of morphosyntax and complexity. (a, b) Mean length of utterance, in words. (c, d) Bound morphemes. phw = per 100 words. (e, f) Embeddings. (g, h) Morphosyntactic errors, including word-level morphological errors, errors in formal lexical devices, omitted words, isolated phrases, and agrammatic or paragrammatic utterances.

morphological complexity, increased in both patients (JI: moderate marginally significant correlation, $r = 0.37$, $p = 0.070$, Figure 6(c); VG: weak non-significant correlation, $r = 0.19$, $p = 0.24$, Figure 6(d)). The number of embeddings per 100 words, a measure of

syntactic complexity, increased in both patients (JI: strong correlation, $r = 0.60$, $p = 0.0059$, Figure 6(e); VG: strong correlation, $r = 0.54$, $p = 0.015$, Figure 6(f)). For both patients, all three of these measures approached or were in the control range by 3 months post stroke.

(5) (a) what do they call these thing (=things)? down here. and she got that. and uh Ø (=I) think that's it. (JI, 3 days post, about getting extra blankets)
(b) uh <then> then this little doggy. he's run (=running) around all the park and every (=everything) for four days. and then we f- let him in. (JI, 6 days post, about getting his third dog)
(c) <it's uh> it seem it's been that every (=ever) since you know. except when I first had the stroke you know then I couldn't even say one word. (JI, 16 days post, about word finding)
(d) but I Ø (=am) doing quite well at work because … b- and the people at work think I'm doing okay. (JI, 110 days post, about returning to work)
(e) oh what have I been doing other than uh <doggie> walking the dogs? m- uh we've done that. um uh I think basically we've just been getting all the um um medical stuff taken care of. (JI, 110 days post, about what he has been doing recently)

(6) (a) he was gonna surgery (.) or (.) uh a doctor. but <he> (.) he change it. (VG, 4 days post, about her husband)
(b) I've had a lot of uh (.) s- surgeon (=surgeries). (VG, 6 days post, about being an operating nurse)
(c) The w- birds was (=were) <a color> a color. and we had those whatever it w- … anyway we had a birds coming out of our ear (=ears). But <they are> they're interested (=interesting). <I had> he had build (=built) a h- h- home for the birds. (VG, 10 days post, about past pets)
(d) she said eventually it would come back (.) in time. (.) so I don't know. I'm in hopes it will. (VG, 41 days post, about her language improving; note: "in hopes" appears to be a dialectal feature; VG grew up in Louisiana)
(e) so they lived together for twenty-seven years. everything was calm. then he got married. (VG, 92 days post, about her parents and brother)

While the morphological and syntactic complexity of JI's utterances increased, his prevalence of morphosyntactic errors remained essentially unchanged over time, and was higher than the control range (little to no correlation, $r = -0.08$, $p = 0.38$, Figure 6(g)). Throughout the time period of our study, he omitted some function words (subject in 5a, auxiliary verb in 5d) and bound morphemes (plural in 5a, progressive in 5b, third person singular suffix in 5c), and some of his utterances were paragrammatic (e.g., 5b,c).

In contrast, VG made progressively fewer morphosyntactic errors over the course of her recovery, approaching the control range (moderate marginally significant correlation, $r = -0.42$, $p = 0.054$, Figure 6(h)). In the first few days after her stroke, her speech was sometimes agrammatic (6a), before taking on a paragrammatic character (6b, 6c), which then improved over time (6d, 6e).

Disruptions to fluency

Filled pauses, false starts, and retracings, phenomena that contribute to an overall impression of reduced speech fluency, all decreased in prevalence over time in both patients (filled pauses in JI: moderate marginally significant correlation, $r = -0.34$, $p = 0.091$, Figure 7(a); filled pauses in VG: moderate correlation, $r = -0.47$, $p = 0.032$, Figure 7(b); false starts in JI: weak non-significant correlation, $r = -0.22$, $p = 0.20$, Figure 7(c); false starts in VG: moderate but non-significant correlation, $r = -0.32$, $p = 0.11$, Figure 7(d); retracings in JI: strong correlation, $r = -0.58$, $p = 0.007$, Figure 7(e); retracings in VG: weak non-significant correlation, $r = -0.29$, $p = 0.14$, Figure 7(f)). Examples are shown in (7) and (8). By 3 months post stroke, both patients continued to produce more false starts than controls, but approached controls in numbers of retracings and were in the normal range for filled pauses.

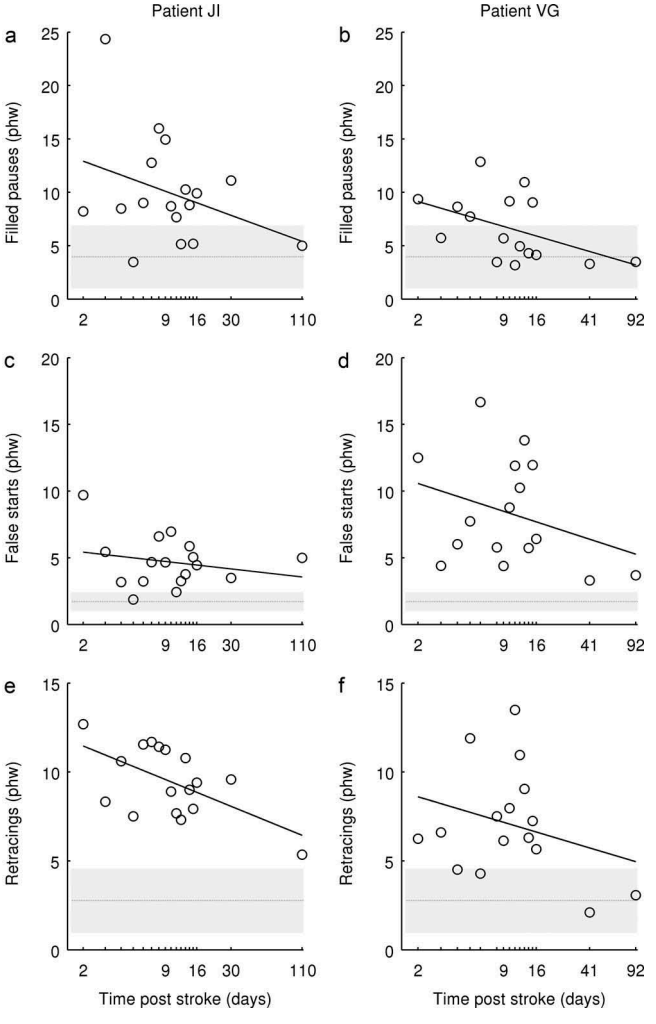

Figure 7. Change over time in disruptions to fluency. (a, b) Filled pauses. phw = per 100 words. (c, d) False starts. (e, f) Retracings.

(7) (a) and <uh they were uh> they're going to uh possibly take me <to to another uh de- uh uh uh go someplace to [sueɪɹouʔ] (=unknown)> somewhere else <to have> to get <a> uh a [biəˀ] (=idea) of what was going in (=on). (JI, 3 days post, about what he did the day prior)

(b) one I did <really bad> really bad. (.) oh this one. but I get it … after a while I got it. but this one I tried. uh but then I got it. it wasn't… <but I θ-> but I <work> work on it. (JI, 5 days post, about practicingpractising a language task)

(c) <and then uh she uh> <so we> and … but <she was> she was good. she was <in> uh I think in the hæspɪ- hospital <for about> for <about seven days I believe> seven or eight days. (JI, 7 days post, about his daughter's childhood)

(d) <and I still did a little> I still Ø (=have) [pɹɑms] (=problems) with some of the words <[tɛmp] (=that)> that won't quite [kʌmp] (=come) out <white (=right)> right. (JI, 15 days post, about his language)

(e) he said "now what you've done for the first f- three and a half wi- months it'll take a good nine or ten or eleven months more." (JI, 110 days post, reporting his doctor's view on his recovery).

(8) (a) a bunch of … (.) uh (.) and <ɪ- it's ɪ- ɪ-> <you can get (.) um (.) the (.) the the> uh you can (.) take the [tʃɜˑmɹɑ] (=chimney) (VG, 5 days post, about building a house)

(b) and it was <on the uh h haɪ highway> (.) on a main n- highway (.) uh (.) uh uh in [tʃænəlʌ] (=Louisiana). (VG, 6 days post, about growing up on her parents' farm)

(c) anyway there's <a kæ-> the capitol <of> of ɜɪɹə- Arizona. (VG, 12 days post, about a travel memory)

(d) they were trying to get me s- to go someplace. I said "I don't wanna go anywhere. I wanna go home." (VG, 16 days post, about people trying to move her).

(e) I've been trying to clean out that room apparently. and I just do that. and they do have people who come in and <help> help me out twice a day. (.) but it gets to the point that there's nothing else to do. (VG, 92 days post, about what she did that day)

Brief aphasia battery

All five measures improved over time in both patients (Figure 8): confrontation naming (JI: strong correlation, $r = 0.96$, $p = 0.0045$; VG: strong correlation, $r = 0.93$, $p = 0.012$), single word comprehension (JI: strong correlation, $r = 0.71$, $p = 0.090$; VG: moderate but non-significant correlation, $r = 0.34$, $p = 0.29$), sentence comprehension (JI: strong but non-significant correlation, $r = 0.51$, $p = 0.19$; VG: strong correlation, $r = 0.85$, $p = 0.034$), repetition (JI: strong correlation, $r = 0.98$, $p = 0.0014$; VG: strong correlation, $r = 0.93$, $p = 0.010$), and reading aloud (JI: strong correlation, $r = 0.94$, $p = 0.0079$; VG: strong but marginally significant correlation, $r = 0.73$, $p = 0.082$).

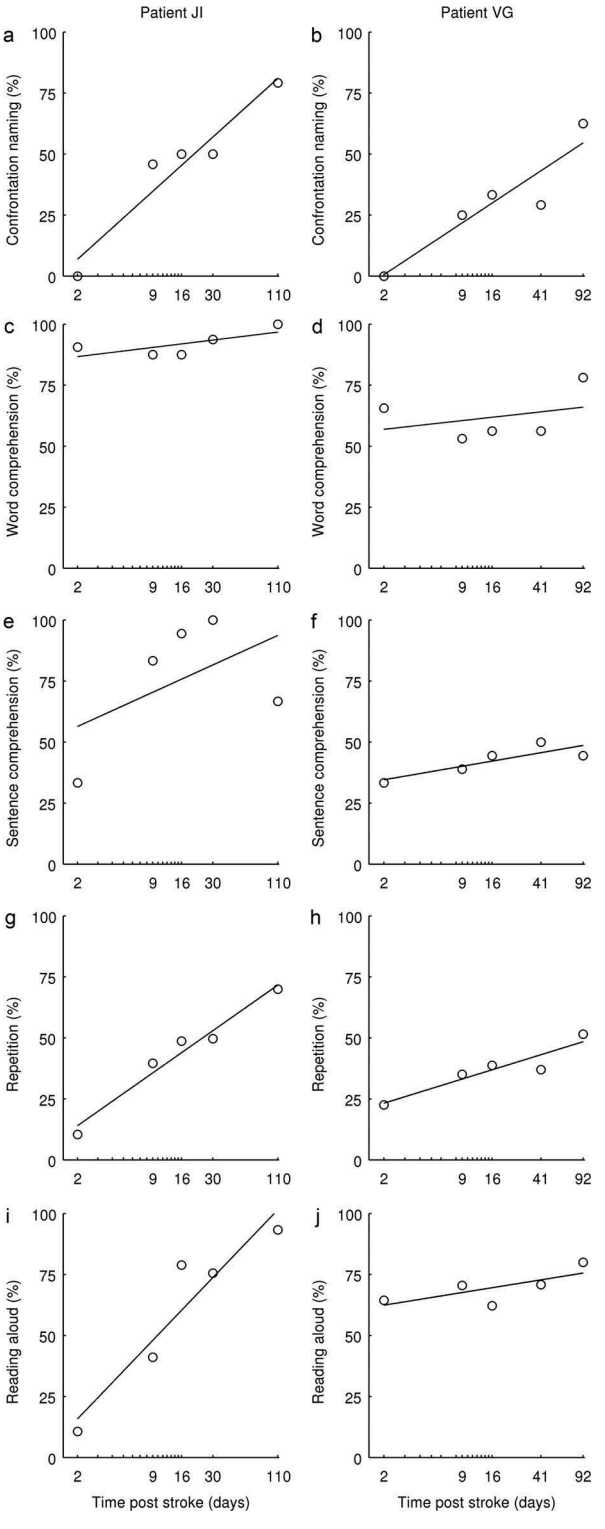

Figure 8. Change over time in scores on the brief aphasia battery. (a, b) Confrontation naming. (c, d) Single word comprehension. (e, f) Sentence comprehension. (g, h) Repetition of words and sentences. (i, j) Reading aloud of words and sentences.

Discussion

In this study, we characterised in detail the recovery of language function over the first 3 months post stroke in two individuals with aphasia following infarction of the left posterior STG: Wernicke's area.

The first aim of our study was to determine which impaired language domains would improve over time. We found that most, but not all, language measures improved over time. Most measures associated with successful language—in particular, words per minute, mean length of utterance, number of bound morphemes, and number of embeddings—increased in both patients; while most measures of erroneous language use—in particular, phonological errors, the composite measure of impaired lexical access, and counts of filled pauses, false starts, and retracings—decreased in both patients. While not all of these changes over time were statistically significant in both patients, the data overall reflected a clear pattern of improvement across multiple language domains, including fluency, lexical access, phonological retrieval and encoding, and syntactic complexity. In VG, amelioration of dysarthria was also observed.

Not every language measure showed improvements. JI produced morphosyntactic errors at a consistent rate over time, even as the syntactic complexity of his utterances increased. Consequently, his speech at the 3-month time point remained somewhat paragrammatic. VG did not show any improvements in proportions of closed class words or verbs, which were greater than the control range, reflecting lexical access difficulties and empty speech (Wilson et al., 2010). On the other hand, VG did show modest improvement on the lexical access composite measure.

Our second aim was to investigate the time course of recovery. We found that the language variables that changed over time were generally well fit by logarithmic models. That is, the most rapid gains were made early, and progress continued but slowed over time. By 3 months, both patients' communication skills were quite effective overall, in the sense that they were usually successful in conveying complex ideas about novel topics, as can be appreciated in some of the examples provided earlier. Yet, both patients still had clearly apparent deficits in connected speech and on the brief aphasia battery, and both met clinical criteria for conduction aphasia, which is a typical chronic outcome for patients with posterior temporal damage (Buchsbaum et al., 2011).

The potential for amelioration of language deficits after damage to Wernicke's area demonstrated in this study is consistent with previous research (Kertesz & Benson, 1970; Kertesz et al., 1993; Kertesz & McCabe, 1977; Laska et al., 2001; Metter et al., 1990; Naeser et al., 1987; Penfield & Roberts, 1959; Selnes et al., 1983, 1984; Weiller et al., 1995), as is the rapid, decelerating time course of recovery (Basso, 1992; Pedersen et al., 1995; Swinburn et al., 2004). The mechanisms behind this rapid recovery are still poorly understood, and may include recruitment and/or reorganisation of ipsilateral perilesional tissue (Heiss, Kessler, Karbe, Fink, & Pawlik, 1993; Heiss, Kessler, Thiel, Ghaemi, & Karbe, 1999; Karbe et al., 1998; Saur et al., 2006) or homotopic regions in the right hemisphere (Crinion & Price, 2005; Weiller et al., 1995). Besides the posterior STG, a broad array of temporo-parietal regions, including the SMG, MTG, angular gyrus, inferior temporal gyrus, and fusiform gyrus have been implicated in language processing in neuropsychological (Kertesz & Benson, 1970; Kertesz et al., 1993; Metter et al., 1990; Naeser et al., 1987; Selnes et al., 1983) and neuroimaging (Binder, Desai, Graves, &

Conant, 2009; Binder et al., 1997) studies. The white matter fibre pathways that underlie and connect these regions, including the arcuate fasciculus, extreme capsule fibre system, inferior longitudinal fasciculus, and middle longitudinal fasciculus are also important (Catani & Mesulam, 2008; Petrides, 2014; Turken & Dronkers, 2011; Wilson et al., 2011). In the two patients we studied, damage was relatively circumscribed to the posterior STG and its underlying white matter, in particular the posterior and direct segments of the arcuate fasciculus (Catani & Mesulam, 2008; Turken & Dronkers, 2011). It appears that in these cases, the surviving temporo-parietal regions and white matter pathways (likely including the anterior indirect segment of the arcuate fasciculus) may provide a substrate for recovery. In contrast, lesions that extend beyond the STG, especially to the MTG, can be expected to impact a much more extensive network of white matter pathways (Turken & Dronkers, 2011) and outcomes are accordingly poorer (Kertesz & Benson, 1970; Kertesz et al., 1993; Metter et al., 1990; Naeser et al., 1987; Selnes et al., 1983).

Despite the remarkable recovery we observed, the persistent conduction aphasia we documented at 3 months post stroke contrasts with outcomes in patients with lesions that are relatively circumscribed to Broca's area, whose language has been reported to be near normal within a month or two (Mohr, 1976). The relative persistence of language deficits after temporal damage as opposed to frontal has been reported by several other authors (Heiss et al., 1999; Penfield & Roberts, 1959; Selnes et al., 1984) and supports the view that the posterior perisylvian region is more indispensable for language than Broca's area.

Quantitative analysis of connected speech samples proved to be an effective means of tracking day-by-day recovery in the acute/subacute post-stroke period (Holland et al., 1985). Changes in language measures across time appear to reflect five main sources of variability: (1) improvement of underlying language function over time; (2) genuine fluctuations in language performance due to varying levels of arousal, motivation, attention, and so on; (3) measurement error due to the relatively brief duration of the speech samples; (4) measurement error due to subjective decisions made in transcription and coding; and (5) measurement error due to quantitative measures that do not perfectly reflect underlying constructs (e.g., the morphosyntactic error composite score includes non-sentence utterances, yet these are not always errors per se). While many measures showed robust logarithmic time courses, no measure was monotone increasing or decreasing, suggesting that one or more of the second, third, fourth, and/or fifth factors contributed significantly to our data. We believe that the third factor—measurement error due to the relatively brief duration of the speech samples—is probably the greatest source of variability unrelated to recovery.

Tracking recovery using quantitative analysis of connected speech minimises the impact of research on participants, which is an important consideration in the acute/subacute post-stroke period. However, while connected speech samples are relatively easy to acquire, they are extremely time consuming to transcribe and code. Transcription and coding of aphasic speech took about 1 hour per minute for this study. Moreover, researchers who carry out transcribing and/or coding require considerable training as well as a strong background in linguistics and/or speech–language pathology (Prins & Bastiaanse, 2004). The fact that we studied only two individuals with aphasia is the most significant limitation of our study and reflects the time-

consuming nature of our approach. To extend our approach to a larger sample of patients, personnel resources would need to be budgeted accordingly. Alternatively, it would be possible to carry out a less nuanced analysis of connected speech samples that would be less time consuming. Two connected speech measures—words per minute (which excludes retraced words) and number of embeddings—hold particular promise as indicators of recovery, since they were strongly and significantly correlated with time post stroke in both patients. These particular measures could be derived from a less comprehensive transcription and coding scheme, and might provide measures of improvement that could have clinical or research utility. Another alternative approach that would also be less time consuming would be to use qualitative instead of quantitative assessments of connected speech samples, as in the Boston Diagnostic Aphasia Examination (Goodglass et al., 2001), and similar to the approach that dominates the field of motor speech disorders (Darley et al., 1969). However, there has been little research to date comparing qualitative and quantitative analyses of connected speech (Grande et al., 2008).

In conclusion, we documented substantial recovery of language function in the first 3 months post stroke in two individuals with aphasia following infarction of the left posterior STG. Recovery was apparent across multiple language domains, including fluency, lexical access, phonological retrieval and encoding, and syntactic complexity. Most language measures improved with a logarithmic time course, with the greatest gains taking place in early in the course of recovery.

Acknowledgements

We thank Audrey Holland for her advice on collecting and analysing speech samples; Pélagie Beeson and Gayle DeDe for providing input on project design and constructive feedback on a draft of this manuscript; Grace Cheifetz, Jane Eustance, Lisa Gordon, Sara McDonald, and Kim Neely for rating motor speech characteristics of speech samples; Kindle Rising and Dana Eriksson for helpful discussions; three anonymous reviewers for their constructive feedback; and the patients for their participation in our research.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was supported in part by the National Institutes of Health (National Institute on Deafness and Other Communication Disorders): grant R01 DC013270.

ORCID

Alexa Bautista http://orcid.org/0000-0002-4487-9593
Stephen M. Wilson http://orcid.org/0000-0001-9884-2852

References

Azuar, C., Leger, A., Arbizu, C., Henry-Amar, F., Chomel-Guillaume, S., & Samson, Y. (2013). The aphasia rapid test: An NIHSS-like aphasia test. *Journal of Neurology*, *260*, 2110–2117. doi:10.1007/s00415-013-6943-x

Basso, A. (1992). Prognostic factors in aphasia. *Aphasiology*, *6*, 337–348. doi:10.1080/02687039208248605

Behrens, T. E. J., Berg, H. J., Jbabdi, S., Rushworth, M. F. S., & Woolrich, M. W. (2007). Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *Neurolmage*, *34*, 144–155. doi:10.1016/j.neuroimage.2006.09.018

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B*, *57*, 289–300.

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, *19*, 2767–2796. doi:10.1093/cercor/bhp055

Binder, J. R., Frost, J. A., Hammeke, T. A., Cox, R. W., Rao, S. M., & Prieto, T. (1997). Human brain language areas identified by functional magnetic resonance imaging. *Journal of Neuroscience*, *17*, 353–362.

Bogen, J. E., & Bogen, G. M. (1976). Wernicke's region—Where is it? *Annals of the New York Academy of Sciences*, *280*, 834–843. doi:10.1111/j.1749-6632.1976.tb25546.x

Buchsbaum, B. R., Baldo, J., Okada, K., Berman, K. F., Dronkers, N., D'Esposito, M., & Hickok, G. (2011). Conduction aphasia, sensory-motor integration, and phonological short-term memory – An aggregate analysis of lesion and fMRI data. *Brain and Language*, *119*, 119–128. doi:10.1016/j.bandl.2010.12.001

Catani, M., & Mesulam, M. (2008). The arcuate fasciculus and the disconnection theme in language and aphasia: History and current state. *Cortex*, *44*, 953–961. doi:10.1016/j.cortex.2008.04.002

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Crinion, J., & Price, C. J. (2005). Right anterior superior temporal activation predicts auditory sentence comprehension following aphasic stroke. *Brain*, *128*, 2858–2871. doi:10.1093/brain/awh659

Darley, F. L., Aronson, A. E., & Brown, J. R. (1969). Differential diagnostic patterns of dysarthria. *Journal of Speech and Hearing Research*, *12*, 246–269. doi:10.1044/jshr.1202.246

Duffy, J. R. (2013). *Motor speech disorders: Substrates, differential diagnosis, and management* (3rd ed.). St. Louis, MO: Elsevier.

Flamand-Roze, C., Falissard, B., Roze, E., Maintigneux, L., Beziz, J., Chacon, A., … Denier, C. (2011). Validation of a new language screening tool for patients with acute stroke: The Language Screening Test (LAST). *Stroke*, *42*, 1224–1229. doi:10.1161/STROKEAHA.110.609503

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*, 189–198. doi:10.1016/0022-3956(75)90026-6

Friston, K. J. (2007). *Statistical parametric mapping: The analysis of funtional brain images*. Amsterdam: Elsevier.

Galantucci, S., Tartaglia, M. C., Wilson, S. M., Henry, M. L., Filippi, M., Agosta, F., … Gorno-Tempini, M. L. (2011). White matter damage in primary progressive aphasias: A diffusion tensor tractography study. *Brain*, *134*, 3011–3029. doi:10.1093/brain/awr099

Gandour, J., Marshall, R., Kim, S., & Neuburger, S. (1991). On the nature of conduction aphasia: A longitudinal case study. *Aphasiology*, *5*, 291–306. doi:10.1080/02687039108248530

Goodglass, H., Kaplan, E., & Barresi, B. (2001). *The Boston Diagnostic Aphasia Examination (BDAE)* (3rd ed.). Baltimore, MD: Lippincott Williamson & Wilkins.

Grande, M., Hussmann, K., Bay, E., Christoph, S., Piefke, M., Willmes, K., & Huber, W. (2008). Basic parameters of spontaneous speech as a sensitive method for measuring change during the course of aphasia. *International Journal of Language & Communication Disorders*, *43*, 408–426. doi:10.1080/13682820701685991

Heiss, W. D., Kessler, J., Karbe, H., Fink, G. R., & Pawlik, G. (1993). Cerebral glucose metabolism as a predictor of recovery from aphasia in ischemic stroke. *Archives of Neurology*, *50*, 958–964. doi:10.1001/archneur.1993.00540090059011

Heiss, W. D., Kessler, J., Thiel, A., Ghaemi, M., & Karbe, H. (1999). Differential capacity of left and right hemispheric areas for compensation of poststroke aphasia. *Annals of Neurology*, *45*, 430–438. doi:10.1002/1531-8249(199904)45:4<430::AID-ANA3>3.0.CO;2-P

Hillis, A. E., Wityk, R. J., Tuffiash, E., Beauchamp, N. J., Jacobs, M. A., Barker, P. B., & Selnes, O. A. (2001). Hypoperfusion of Wernicke's area predicts severity of semantic deficit in acute stroke. *Annals of Neurology*, *50*, 561–566. doi:10.1002/(ISSN)1531-8249

Holland, A. L., Miller, J., Reinmuth, O. M., Bartlett, C., Fromm, D., Pashek, G., … Swindell, C. (1985). Rapid recovery from aphasia: A detailed language analysis. *Brain and Language*, *24*, 156–173. doi:10.1016/0093-934X(85)90101-4

Karbe, H., Thiel, A., Weber-Luxenburger, G., Herholz, K., Kessler, J., & Heiss, W.-D. (1998). Brain plasticity in poststroke aphasia: What is the contribution of the right hemisphere? *Brain and Language*, *64*, 215–230. doi:10.1006/brln.1998.1961

Kertesz, A., & Benson, D. F. (1970). Neologistic jargon: A clinicopathological study. *Cortex*, *6*, 362–386. doi:10.1016/S0010-9452(70)80002-8

Kertesz, A., Lau, W. K., & Polk, M. (1993). The Structural determinants of recovery in Wernicke's aphasia. *Brain and Language*, *44*, 153–164. doi:10.1006/brln.1993.1010

Kertesz, A., & McCabe, P. (1977). Recovery patterns and prognosis in aphasia. *Brain*, *100*, 1–18. doi:10.1093/brain/100.1.1

Kohn, S., & Smith, K. L. (1994). Evolution of impaired access to the phonological lexicon. *Journal of Neurolinguistics*, *8*, 267–288. doi:10.1016/0911-6044(94)90013-2

Laska, A. C., Hellblom, A., Murray, V., Kahan, T., & Von Arbin, M. (2001). Aphasia in acute stroke and relation to outcome. *Journal of Internal Medicine*, *249*, 413–422. doi:10.1046/j.1365-2796.2001.00812.x

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk*. Mahwah, NJ: Lawrence Erlbaum.

MacWhinney, B. (2012). *The CHILDES project: Tools for analyzing talk - -electronic edition*. Retrieved from http://childes.talkbank.org/manuals/chat.pdf.

MacWhinney, B., Fromm, D., Forbes, M., & Holland, A. (2011). AphasiaBank: Methods for studying discourse. *Aphasiology*, *25*, 1286–1307. doi:10.1080/02687038.2011.589893

Marie, P. (1906). The third frontal convolution plays no special role in the function of language. *Semaine Medicale*, *26*, 241–247.

Marshall, R. C. (1982). Language and speech recovery in a case of viral encephalitis. *Brain and Language*, *17*, 316–326. doi:10.1016/0093-934X(82)90024-4

Metter, E. J., Hanson, W. R., Jackson, C. A., Kempler, D., van Lancker, D., Mazziotta, J. C., & Phelps, M. E. (1990). Temporoparietal cortex in aphasia. Evidence from positron emission tomography. *Archives of Neurology*, *47*, 1235–1238. doi:10.1001/archneur.1990.00530110097024

Mohr, J. P. (1976). Broca's area and Broca's aphasia. In H. Whitaker (Ed.), *Studies in Neurolinguistics* (pp. 201–233). New York, NY: Academic Press.

Naeser, M. A., Helm-Estabrooks, N., Haas, G., Auerbach, S., & Srinivasan, M. (1987). Relationship between lesion extent in 'Wernicke's area' on computed tomographic scan and predicting recovery of comprehension in Wernicke's aphasia. *Archives of Neurology*, *44*, 73–82. doi:10.1001/archneur.1987.00520130057018

Pedersen, P. M., Stig Jørgensen, H., Nakayama, H., Raaschou, H. O., & Olsen, T. S. (1995). Aphasia in acute stroke: Incidence, determinants, and recovery. *Annals of Neurology*, *38*, 659–666. doi:10.1002/ana.v38:4

Penfield, W., & Roberts, L. (1959). *Speech and brain mechanisms*. Princeton, NJ: Princeton University Press.

Petrides, M. (2014). *Neuroanatomy of language regions of the human brain*. Amsterdam: Elsevier.

Prins, R., & Bastiaanse, R. (2004). Analysing the spontaneous speech of aphasic speakers. *Aphasiology*, *18*, 1075–1091. doi:10.1080/02687030444000534

Rossion, B., & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's object pictorial set: The role of surface detail in basic-level object recognition. *Perception*, *33*, 217–236. doi:10.1068/p5117

Saffran, E. M., Berndt, R. S., & Schwartz, M. F. (1989). The quantitative analysis of agrammatic production: Procedure and data. *Brain and Language*, *37*, 440–479. doi:10.1016/0093-934X(89)90030-8

Saur, D., Lange, R., Baumgaertner, A., Schraknepper, V., Willmes, K., Rijntjes, M., & Weiller, C. (2006). Dynamics of language reorganization after stroke. *Brain*, *129*, 1371–1384. doi:10.1093/brain/awl090

Selnes, O. A., Knopman, D. S., Niccum, N., Rubens, A. B., & Larson, D. (1983). Computed tomographic scan correlates of auditory comprehension deficits in aphasia: A prospective recovery study. *Annals of Neurology*, *13*, 558–566. doi:10.1002/(ISSN)1531-8249

Selnes, O. A., Niccum, N., Knopman, D. S., & Rubens, A. B. (1984). Recovery of single word comprehension: CT-scan correlates. *Brain and Language*, *21*, 72–84. doi:10.1016/0093-934X(84)90037-3

Swinburn, K., Porter, G., & Howard, D. (2004). *Comprehensive aphasia test*. Hove: Psychology Press.

Turken, A. U., & Dronkers, N. F. (2011). The neural architecture of the language comprehension network: Converging evidence from lesion and connectivity analyses. *Frontiers in Systems Neuroscience*, *5*, 1. doi:10.3389/fnsys.2011.00001

Weiller, C., Isensee, C., Rijntjes, M., Huber, W., Müller, S., Bier, D., … Diener, H. C. (1995). Recovery from Wernicke's aphasia: A positron emission tomographic study. *Annals of Neurology*, *37*, 723–732. doi:10.1002/ana.410370605

Wernicke, K. (1874). *Der Aphasische Symptomencomplex*. Breslau: Cohn and Weigert.

Wilson, S. M., Dronkers, N. F., Ogar, J. M., Jang, J., Growdon, M. E., Agosta, F., … Gorno-Tempini, M. L. (2010). Neural correlates of syntactic processing in the nonfluent variant of primary progressive aphasia. *Journal of Neuroscience*, *30*, 16845–16854.

Wilson, S. M., Galantucci, S., Tartaglia, M. C., Rising, K., Patterson, D. K., Henry, M. L., … Gorno-Tempini, M. L. (2011). Syntactic processing depends on dorsal language tracts. *Neuron*, *72*, 397–403. doi:10.1016/j.neuron.2011.09.014

Wilson, S. M., Henry, M. L., Besbris, M., Ogar, J. M., Dronkers, N. F., Jarrold, W., … Gorno-Tempini, M. L. (2010). Connected speech production in three variants of primary progressive aphasia. *Brain*, *133*, 2069–2088. doi:10.1093/brain/awq129

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). ELAN: A professional framework for multimodality research. In *Proceedings of LREC 2006, fifth International Conference on language resources and evaluation*. Retrieved from http://www.lrec-conf.org/proceedings/lrec2006

Yushkevich, P. A., Piven, J., Hazlett, H. C., Smith, R. G., Ho, S., Gee, J. C., & Gerig, G. (2006). User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *NeuroImage*, *31*, 1116–1128. doi:10.1016/j.neuroimage.2006.01.015