BRAIN
A JOURNAL OF NEUROLOGY

Rhythm in disguise: why singing may not hold the key to recovery from aphasia

Benjamin Stahl, Sonja A. Kotz, Ilona Henseler, Robert Turner and Stefan Geyer

Max Planck Institute for Human Cognitive and Brain Sciences, 04103 Leipzig, Germany

Correspondence to: Benjamin Stahl,
Max Planck Institute for Human Cognitive and Brain Sciences,
Stephanstraße 1A,
04103 Leipzig, Germany
E-mail: stahl@cbs.mpg.de

The question of whether singing may be helpful for stroke patients with non-fluent aphasia has been debated for many years. However, the role of rhythm in speech recovery appears to have been neglected. In the current lesion study, we aimed to assess the relative importance of melody and rhythm for speech production in 17 non-fluent aphasics. Furthermore, we systematically alternated the lyrics to test for the influence of long-term memory and preserved motor automaticity in formulaic expressions. We controlled for vocal frequency variability, pitch accuracy, rhythmicity, syllable duration, phonetic complexity and other relevant factors, such as learning effects or the acoustic setting. Contrary to some opinion, our data suggest that singing may not be decisive for speech production in non-fluent aphasics. Instead, our results indicate that rhythm may be crucial, particularly for patients with lesions including the basal ganglia. Among the patients we studied, basal ganglia lesions accounted for more than 50% of the variance related to rhythmicity. Our findings therefore suggest that benefits typically attributed to melodic intoning in the past could actually have their roots in rhythm. Moreover, our data indicate that lyric production in non-fluent aphasics may be strongly mediated by long-term memory and motor automaticity, irrespective of whether lyrics are sung or spoken.

Keywords: non-fluent aphasia; melodic intonation therapy; basal ganglia; long-term memory; automaticity of formulaic expressions

Introduction

Patients with a left hemisphere stroke frequently suffer from various language-related disorders that restrain or disrupt the spontaneous expression of speech. Such impairments are commonly grouped together under the heading of 'non-fluent aphasia'. A considerable number of these patients never recover completely, despite intensive therapy. For nearly two centuries clinicians have observed that patients with non-fluent aphasia are nevertheless able to sing, with some even being able to sing words (Mills, 1904; Gerstmann, 1964; Yamadori et al., 1977). This astonishing observation has inspired a number of clinical interventions worldwide, among them the highly debated melodic intonation therapy (Albert et al., 1973; Helm-Estabrooks et al., 1989). This therapy consists of a rehabilitation programme with various elements, including three main components: melodic intoning, rhythmic speech and the use of common phrases.

The overall composition of melodic intonation therapy may appear meaningful from a therapeutic point of view. However, when focusing on the different therapeutic elements and their individual contributions to clinical efficacy, some questions arise. To what extent is melody, rhythm or their combination decisive for speech production in aphasics? Does this depend on individual lesion locations within the brain? What role does memory play if

one employs familiar song lyrics? And, since formulaic phrases are used, to what extent may the benefits of melodic intonation therapy be due to preserved motor automaticity? Recent work on these issues has led to a number of ambiguous, sometimes contradictory results.

Melodic intoning

The contribution of singing to melodic intonation therapy has been considered as crucial by the inventors of the treatment. Singing was supposed to stimulate cortical regions in the right hemisphere with homotopic location relative to left language areas. Consequently, this cortical stimulation through singing was assumed to have a positive impact on speech recovery (Albert *et al.*, 1973; Sparks *et al.*, 1974). Indeed, this assumption may appear consistent with right-hemispheric processing of features related to music and prosody (Riecker *et al.*, 2000; Callan *et al.*, 2006; Özdemir *et al.*, 2006). Moreover, some evidence suggests that the right hemisphere may have a compensatory function in speech recovery (Musso *et al.*, 1999; Blasi *et al.*, 2002; Saur *et al.*, 2006).

Several *cross-sectional* studies with non-fluent aphasics, however, failed to support the more effective role of singing, as compared with rhythmic speech (Cohen and Ford, 1995; Boucher *et al.*, 2001) or natural speech (Hébert *et al.*, 2003). Notably, one study found melodic intoning to be advantageous to natural speech when patients were singing along to vocal playback delivered by earphones (Racette *et al.*, 2006). Until now, *longitudinal* evidence for the efficacy of singing in speech recovery is sparse and seems problematic from an experimental point of view. Only two case reports on melodic intonation therapy made use of a control condition, with one study controlling for intoning in an experienced singer (Wilson *et al.*, 2006) and another study controlling for singing but not for rhythmic left-hand tapping in two subjects (Schlaug *et al.*, 2008). Consequently, these results may be confounded with musical training and influences related to rhythm.

Neuroimaging research on the role of singing in speech recovery has given rise to some ambiguous results. Two single case studies suggested right functional differences (Schlaug *et al.*, 2008) and structural changes in the right arcuate fasciculus (Schlaug *et al.*, 2009) after treatment with melodic intonation therapy, while speech production was found to be improved. It could be concluded that singing had a causal, curative effect on speech production in these patients. However, there are different ways to interpret these data. Structural changes in the right arcuate fasciculus, if indeed such findings are validated, may well be the result of intensive singing, whereas the benefits in speech production could be due to massive repetition of the phrases used—recruiting, for instance, perilesional left brain regions. In other words, singing and massive repetition of phrases may be thought of as *two independent mechanisms* that are not causally linked. Conclusions regarding benefits from singing for speech production are therefore questionable in light of these data. In addition, a PET experiment revealed more *left* prefrontal activation in seven non-fluent aphasics when producing phrases that they had trained during melodic intonation therapy (Belin *et al.*, 1996). Indeed, a number of

studies support the crucial role of perilesional left areas in speech recovery (Cao *et al.*, 1999; Rosen *et al.*, 2000; Zahn *et al.*, 2004).

Rhythmic speech

The role of rhythm in speech recovery from aphasia appears to have been neglected for some time. One reason for this may be the experimental problem of how to control for rhythm. Only one study addressed this problem (Cohen and Ford, 1995). Natural speech was chosen as a control for rhythmic speech. Although not mentioned by the authors, the use of natural speech may have resulted in different syllable durations in each condition. Slowing down of syllable duration, however, was suggested to improve articulation, at least in dysarthric patients (Hustad *et al.*, 2003). Moreover, natural speech in stress-timed languages—in this case, English—still implies a distinct metre and may therefore still be considered as rhythmic. This assumption is indirectly supported by evidence from healthy subjects who display increased vocal loudness when producing metrically prominent syllables in English (Kochanski and Orphanidou, 2008). Finally, a metronome accompaniment was chosen for the rhythmic condition only. This may have advantaged the production of natural speech since no additional sound source interfered. Accordingly, the results of this study indicated better performance in the natural speech condition, and may have to be viewed with caution.

However, some indirect evidence points to the contribution of rhythm for speech production. Articulation may be modulated, for example, by visual or auditory rhythmic cues (Pilon *et al.*, 1998; Brendel and Ziegler, 2008) or rhythmic tapping of the left hand, thus engaging sensorimotor networks in the right hemisphere (Gentilucci and Dalla Volta, 2008). It may therefore be noteworthy that melodic intonation therapy includes, among other additional elements, rhythmic left-hand tapping (Albert *et al.*, 1973; Helm-Estabrooks *et al.*, 1989). Rhythmic hand tapping could, at least theoretically, have a profound impact on speech production in aphasics.

Interestingly, research on melodic intonation therapy has focused very much on the dichotomy of left and right cortical functions in speech recovery. However, the contribution from subcortical areas has not drawn much attention. This is all the more surprising as subcortical areas, specifically the basal ganglia, are suggested to mediate rhythmic segmentation in speech perception and production (Kotz *et al.*, 2009). It may therefore be argued that patients with lesions including the basal ganglia should benefit more from rhythmicity than patients without such lesions.

Memory and motor automaticity

Research on the role of memory for speech production in aphasics is based on the observation of a few cases. Two non-fluent aphasics showed improved performance for familiar song lyrics as compared with spontaneous speech (Hébert *et al.*, 2003) or unknown lyrics (Straube *et al.*, 2008). Interestingly, lyric production in these patients was not affected by the circumstance of whether the original melody was used or not. This finding is in accordance with evidence for independent, dual encoding of lyrics and

melody in healthy subjects (Samson and Zatorre, 1991, 1992), although a number of studies suggested perceptual connectedness of melody and lyrics in memory (Crowder et al., 1990; Peretz et al., 2004; Gordon et al., 2010). The case reports presented here indicate that lyric production in aphasics may be largely mediated by verbal long-term memory. However, it remains unclear whether this finding holds true for a larger sample of patients and whether there are factors that determine the contribution of memory for lyric production, such as age. Furthermore, it may be useful to disentangle effects of long-term memory from motor automaticity, since memory and automaticity may affect speech production in different ways. For example, a PET study with healthy subjects revealed diverging activation patterns during recitation of well-known song lyrics as opposed to automatized counting (Blank et al., 2002).

The use of common, formulaic phrases is a substantial component of melodic intonation therapy. It may be presumed that the production of such phrases is, to a considerable extent, automatized at the motor level. Yet, the contribution of preserved automaticity in formulaic phrases for melodic intonation therapy has not been investigated up until now. This is all the more surprising as several lesion studies have suggested that the production of formulaic speech may be processed within the right cortex and the right basal ganglia (Speedie et al., 1993; Sidtis and Postman, 2006; Sidtis et al., 2009). Accordingly, a PET study with healthy subjects revealed similar activation patterns during humming and automatized recitation of weekdays (Ryding et al., 1987). This raises doubt on the consistency of some neuroimaging research on melodic intoning as singing and the production of formulaic speech show a functional overlap in the brain.

In the current study, we aimed to assess the relative importance of melody, rhythm, lyric memory and motor automaticity for speech production in patients with non-fluent aphasia.

Materials and methods

Participants

The present multicentre study was conducted at five rehabilitation centres located in Berlin, Germany. Seventeen stroke patients were included in the study. Table 1 provides an overview of the patients' individual case histories.

Patients were German native-speakers, right-handed and aged 27–80 years [mean (standard deviation) age 56 (14) years]. None of the patients had a pre-morbid history of neurological or psychiatric impairments, nor did any of the patients suffer from dementia. At the time of testing, all patients were at least 3 months post-infarction, except in one case (Patient HS). Eight independent clinical linguists tested the patients within 1 month prior to the study, using a German standard aphasia test battery (Aachen Aphasia Test, Huber et al., 1984). Specified test scores are given in Table 2.

All patients were classified as non-fluent aphasics, diagnosed with Broca's aphasia (n = 10) or global aphasia with prevailing expressive deficits (n = 7). Non-fluent aphasia usually co-occurs with speech disorders that include difficulties to plan and to execute oral, speech-specific movements (apraxia of speech), to coordinate articulatory organs, respiration and larynx (dysarthria) or to swallow

Table 1 Patient history

Patient	Sex	Age (years)	Months since last infarction	Number of infarcts	Pre-morbid handedness	Aetiology	Left basal ganglia lesions	Right hemisphere lesions
AS	F	65	8	1	Right	Ischaemia in left MCA	None	None
BN	F	76	84	1	Right	Ischaemia in left MCA	Putamen, caudate nucleus[a], pallidum	None
CM	M	46	23	1	Right	Ischaemia in left MCA	Putamen, caudate nucleus[a], pallidum[a]	None
DO	M	46	5	1	Right	Ischaemia in left MCA	Putamen[a], caudate nucleus[a], pallidum[a]	None
FF	F	27	12	1	Right	Ischaemia in left MCA, haemorrhage in left putamen	Putamen	None
JD	M	52	4	1	Right	Ischaemia in left MCA	Putamen, caudate nucleus[a]	None
HK	F	52	10	1	Right	Ischaemia in left MCA	Putamen	None
HP	F	68	6	1	Right	Haemorrhage in left basal ganglia	Putamen, caudate nucleus, pallidum	None
HS	F	80	1	1	Right	Ischaemia in left MCA	None	None
IK	M	61	9	1	Right	Ischaemia in left MCA	Putamen, caudate nucleus, pallidum	None
KH	M	39	36	1	Right	Ischaemia in left MCA	None	Right cerebellum
LS	F	53	36	2	Right	Ischaemia in left MCA	Putamen, caudate nucleus, pallidum	None
LT	M	76	5	1	Right	Ischaemia in left MCA	Putamen[a]	Right parietal cortex
PL	M	49	6	1	Right	Ischaemia in left MCA	Putamen, caudate nucleus, pallidum	None
PR	F	58	156	1	Right	Ischaemia in left MCA	Putamen, caudate nucleus, pallidum	None
RK	M	62	12	2	Right	Haemorrhage in left basal ganglia and left pons, left medulla	Putamen, pallidum	Right basal ganglia, right pons
TJ	F	45	7	1	Right	Ischaemia in left MCA	Putamen, caudate nucleus, pallidum	None

[a]Localization with limited certainty; data are therefore excluded from further analysis. F = female; M = male; MCA = middle cerebral artery.

Table 2 Language assessment

Patient	Token test	Comprehension	Naming	Repetition	Diagnosis
AS	2/50	120/120	99/120	122/150	Broca's aphasia, apraxia
BN	16/50	104/120	0/120	91/150	Broca's aphasia, apraxia
CM	21/50	93/120	0/120	43/150	Broca's aphasia, apraxia, dysarthria
DO	37/50	39/120	0/120	32/150	Global aphasia, apraxia
FF	0/50	120/120	88/120	124/150	Broca's aphasia, apraxia
JD	14/50	110/120	57/120	83/150	Broca's aphasia, apraxia
HK	26/50	72/120	0/120	58/150	Global aphasia, apraxia
HP	24/50	76/120	5/120	85/150	Global aphasia, dysarthria, dysphagia
HS	34/50	77/120	0/120	47/150	Global aphasia
IK	16/50	90/120	57/120	100/150	Broca's aphasia, apraxia
KH	0/50	120/120	98/120	144/150	Broca's aphasia, apraxia
LS	31/50	57/120	0/120	24/150	Global aphasia, apraxia
LT	12/50	89/120	82/120	140/150	Broca's aphasia, apraxia
PL	14/50	99/120	60/120	77/150	Broca's aphasia, apraxia, dysarthria
PR	9/50	112/120	75/120	102/150	Broca's aphasia, apraxia
RK	27/50	75/120	21/120	34/150	Global aphasia, apraxia, dysphagia
TJ	19/50	72/120	5/120	11/150	Global aphasia, apraxia

Scores of the Aachen Aphasia Test. Token Test: no/mild disorder (0–6); light (7–21); middle (22–40); severe (>40). Comprehension (including words and sentences in both the visual and auditory modality): no/mild disorder (104–120); light (87–103); middle (58–86); severe (1–57). Naming: no/mild disorder (109–120); light (92–108); middle (41–91); severe (1–40). Repetition: no/mild disorder (144–150); light (123–143); middle (75–122); severe (1–74).

(dysphagia). As non-fluent aphasics commonly show apractic behaviour, we aimed to increase the diagnostic reliability for the assessment of apraxia. Consequently, apraxia of speech had to be diagnosed by at least two experienced clinical linguists on the basis of direct observations, which involved inconsistently occurring phonemic or phonetic errors, word initiation difficulties and visible groping (Brendel and Ziegler, 2008). Correspondingly, dysarthria was diagnosed in case of constantly occurring phonetic errors. As a result, the diagnosed concomitant speech disorders in our subjects involved apraxia of speech ($n = 15$), dysarthria ($n = 2$) and dysphagia ($n = 2$).

Patients were eligible for study inclusion when the aphasia test results indicated a largely preserved simple comprehension, with a comparably limited verbal expression. It should be noted that the patients were considered 'non-fluent' based on the typological classifications as indicated by the aphasia test (global or Broca's aphasia), not based on the concomitant speech disorders. Moreover, aphasia was diagnosed by the clinical linguistics as a prevailing disorder in all of the patients. However, given the large proportion of patients with concomitant apraxia in the current sample, some of the results may be influenced by motor impairments related to apraxia. All patients had undergone speech therapy, which did not comprise singing or explicit rhythmic speech. None of the patients displayed any specific musical training or experience in choral singing. The sample may therefore be considered as exemplary in a clinical context.

CT and MRI scans as well as relevant medical reports were obtained for all patients. A neurologist with special expertise in neuroradiology (I.H.) re-analysed all CT or MRI scans without any knowledge of the corresponding speech output data. All patients showed a left middle cerebral artery infarction, except for three patients with left basal ganglia haemorrhages (Patients FF, HP and RK). To increase the variability in pitch accuracy for subsequent covariation analyses, three aphasic patients (Patients KH, LT and RK) with additional lesions in the right hemisphere were included. All CT or MRI scans were thoroughly analysed for lesions within the left basal ganglia, including the caudate nucleus, the putamen and the pallidum. First, we computed

separate scales for each basal ganglia substructure (1 = lesion; 0 = no lesion). When a lesion could not be identified with satisfying certainty, it was discarded from further analysis (0.5 = lesion identification impossible). Finally, we computed a composite score indicating the number of substructure lesions within the basal ganglia (0–3 = zero to three substructure lesions including the caudate nucleus, the putamen and the pallidum). Figure 1 shows the brain scans of two subjects with lesions either including the basal ganglia or not.

The study was approved by the Ethical Committee at University of Leipzig and by the participating institutions in Berlin, and informed consent was obtained from all patients.

Stimuli

The experimental design focused on melody, rhythm and the selection of the lyrics. A schematic overview of the different conditions is provided in Fig. 2.

Three experimental modalities were applied: melodic intoning, rhythmic speech and a spoken arrhythmic control. In the conditions melodic intoning and rhythmic speech, patients were singing or speaking along to a playback, composed of a pre-recorded voice to mimic and a 4/4 percussion beat according to a chosen song. The pre-recorded voice and the percussion beat were consistently used in every sung and spoken condition, including the spoken arrhythmic control. In the spoken arrhythmic control, however, the percussion beat turned into a 3/4 measure, and was shifted by an eighth note. This arrhythmic interference paradigm was chosen to manipulate the degree of rhythmicity while not confounding the results by different syllable durations. It should be noted that the percussive manipulations did not affect the duration of each syllable throughout the experiment. Rhythmic speech served as the control condition for melodic intoning, whereas the arrhythmic condition provided the control for rhythmic speech. To assess the degree of rhythmicity in each condition, five healthy pilot subjects were asked to perform the different conditions while rating the perceived rhythmicity. All raters independently classified the arrhythmic control as 'highly arrhythmic'.
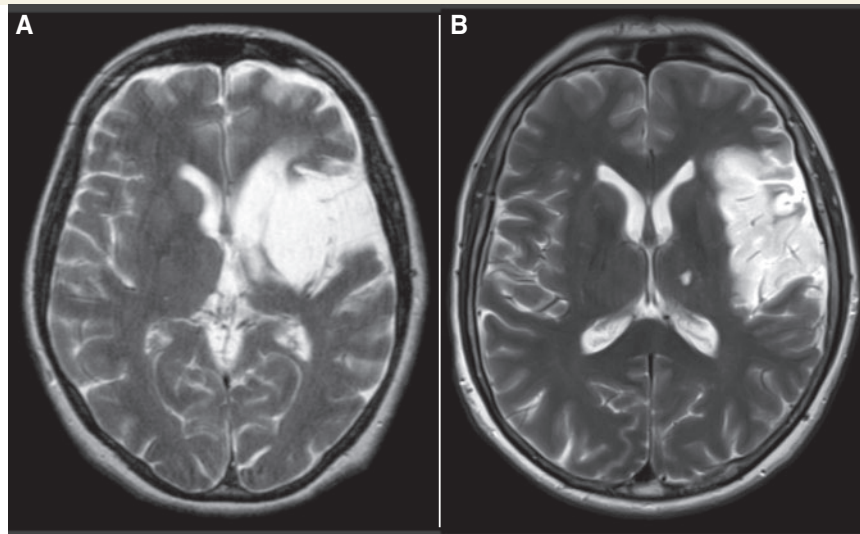
Figure 1 T$_2$-weighted MRI scans (axial view) of Patients PR (**A**) and AS (**B**). Both scans show left middle cerebral artery infarctions, with only Patient PR's lesion including the left basal ganglia.

Figure 2 Schematic overview of the experimental conditions. Three lyric types are employed: original, formulaic and non-formulaic lyrics (from *top* to *bottom*). Each lyric type is produced in three experimental modalities: melodic intoning, rhythmic speech and a spoken arrhythmic control. In the conditions melodic intoning and rhythmic speech, patients sing or speak along with a playback composed of a voice to mimic and a rhythmic percussion beat, which is shown here (rhythmic). The first beat in every 4/4 measure is stressed by lowering the percussion frequency and by accentuating its intensity. In the spoken arrhythmic control, the percussion beat turns into a 3/4 stress pattern, and is shifted by an eighth note (arrhythmic).

Playback voice and percussion beat were mixed in the recording, with both tracks being separately normalized. The sound intensity level of the percussion beat was decreased by 10 dB to make both tracks clearly audible. Vocal playback parts, both sung and spoken, were performed by a male singer. The sung playback parts were recorded in two tonal keys (B and F major) to adopt the patients' individual vocal range, with a piano sound indicating the initial note. Natural prosody was applied for the spoken playback parts. The playback voice was slightly digitally edited to ensure that each syllable was precisely placed within the measure. For the percussion beat, a wooden metronome sound was employed. The first percussion beat in every 4/4 and 3/4 measure was stressed by lowering the percussion frequency and by accentuating its intensity (first beat in every measure: fundamental frequency of 280 Hz, sound intensity level of 80 dB; all remaining beats in every 4/4 or 3/4 measure: fundamental frequency of 420 Hz, sound intensity level of 70 dB; see also Kochanski and Orphanidou, 2008). A tempo of 100 beats per minute was chosen, with a mean duration of 780 ± 25 ms per syllable. The tempo was chosen based on pilot data. With this tempo, patients produced about half of the syllables correctly, thus indicating a medium difficulty level. Every condition was primed by two measures of 4/4 percussion beats. Examples of the playbacks can be downloaded at http://www.cbs.mpg.de/~stahl.

Rhythmic percussive accompaniments are usually not part of spoken utterances in everyday life. To control whether the rhythmic percussion beats in the spoken conditions may have interfered with speech production in the patients, we repeated the experiment with four subjects (Patients JD, KH, LS and LT) while using the same vocal playbacks in the rhythmic speech condition, either with or without percussive accompaniments.

Three types of lyrics were employed in each of the modalities described above: original, formulaic and non-formulaic lyrics. To select a song with very well-known lyrics we explored the familiarity of common German nursery rhymes and folk songs in an age-matched control group of 35 healthy subjects. First, the control subjects were presented with four initial song bars and instructed to complete the melody by humming the remaining notes. Correspondingly, participants were asked in a second step to complete the song lyrics by free recitation. Based on this procedure, a well-known German nursery rhyme was chosen ('Hänschen klein'), with 100% of correctly

produced notes and 87% of correctly produced lyric syllables. It is noteworthy that a correlation between correctly produced syllables and the control participants' age did not reach significance. The melody of the chosen song mainly consists of seconds and thirds, while not exceeding the range of a fifth, and may therefore be considered as very simple.

In a next step, we developed novel lyrics while using the same melody. Formulaic lyrics were composed of stereotyped phrases ('Hello, everything alright? Everything's fine…'), which have been classified as 'largely automatized' by eight clinical linguists. The selected formulaic phrases are highly relevant for communication in everyday life (salutations, farewells, well-being, food) and their order of sequence can be found in a natural conversation. Non-formulaic lyrics comprised very unlikely, but syntactically correct phrases, as they may occur in modern poetry ('Bright forest, there at the boat, thin like oak…').

The probability of the appearance of a word in a language usually depends on the previous word, as denoted by the word transition frequency. From a psycholinguistic point of view, word transition frequencies may serve as a marker for over-learnedness or automaticity in spoken language. In other words, formulaic phrases should be expected to show relatively high word transition frequencies, whereas non-formulaic phrases should show very low word transition frequencies. Non-formulaic lyrics in the present study were therefore conceived in a way so that they showed significantly lower word transition frequencies than formulaic lyrics [$t(33) = 2.3$, $P = 0.029$]. However, formulaic and non-formulaic lyrics did not differ in word frequency [$t(34) < 0.1$, not significant (NS)], word frequency variance [$F(1, 34) = 0.7$, NS], syllable frequency [$t(34) = 0.5$, NS], number of consonants and syntactic phrase structure. All lyrics were consistent with the rhythmically required metre in German. Table 3 provides some characteristics of the lyrics.

Procedure

Testing took place in two sessions during 1 h. Every session was divided in two parts with pauses in between according to the patients' individual needs. To avoid carryover effects, the modalities were presented in blocks and always included three lyric types (original, formulaic, non-formulaic). The stimuli blocks were counterbalanced for each participant: sung, spoken, arrhythmic, pause, arrhythmic, spoken, sung in the first session, with the reversed order in the second session. A correlation between articulatory quality in each condition and the corresponding trial number suggested learning effects in three patients

[Patients JD, FF and AS; $r(34) = 0.67$, 0.57, 0.33; $P < 0.001$, <0.001 and 0.049, respectively]. However, none of these patients exhibited a deviant result pattern of overall means in any of the test conditions.

Participants were seated in front of two loudspeakers in a distance of 75 cm. Patients listened to the vocal playback to sing or speak along with, while being provided with separate sheets of music for each lyric type. It should be noted that lip-reading was not possible. Moreover, rhythmic hand tapping was not allowed as it may interfere with speech production, i.e. by engaging the sensorimotor system. The acoustic setting was conceived to resemble choral singing, with auditory feedback originating from the singer's own voice as well as from surrounding sound sources. In a pilot study with five healthy subjects, the playback intensity was chosen to be approximately balanced with the singer's perceived own vocal loudness. Auditory feedback via earphones was dispensed with to preserve the natural vocal self-monitoring. Utterances were recorded by a head microphone (C520 Vocal Condenser Microphone, AKG Acoustics) and a digital recording device (M-Audio Microtrack II, Avid Technology).

Data analysis

Two speech–language pathology students and the experimenter (B.S.) independently rated the articulatory quality of the produced utterances, based on the digital sound files, with two raters for each patient. Articulatory quality was denoted as the percentage of correct syllables in each condition. Syllables were chosen over words as the critical unit to account for the fact that in apractic patients errors often occur at the syllable level (Aichert and Ziegler, 2004).

A total number of 28 764 syllables were rated. The first two syllables in each condition were discarded from the analyses to control for onset difficulties. Correct syllables were scored with one point (41% of all rated syllables), and half points were given for two conditions: phonemic or phonetic errors occurred in one or more consonants per syllable, but not in the vowel, or vice versa (27% of syllables). No points were allocated when errors occurred in both vowel and in one or more of the consonants within a syllable (24%). Further errors were classified as syllable substitutions as part of a different word (1%) or omissions (7%). The scoring procedure is based on a previous study (Racette *et al.*, 2006), with a more precise definition of the half-point category being applied in the present work.

Pitch accuracy of each sung syllable was assessed separately for each lyric type. It is noteworthy that pitch accuracy did not significantly differ between any of the lyric types employed, irrespective of

Table 3 Characteristics of the lyrics

Feature	Original lyrics	Formulaic lyrics	Non-formulaic lyrics
Mean word frequency (CI)	574 980 ($\pm$400 874)[a]	110 900 ($\pm$58 289)	110 921 ($\pm$67 376)
Mean word transition frequency (right neighbour)	4128	4609	0
Mean syllable frequency (CI)	9510 ($\pm$7893)	10 881 ($\pm$8096)	13 615 ($\pm$11 459)
Number of consonants	93	82	82
Number of syllables	49	49	49
Number of words	38	35	35
Number of ellipsoidal phrases	7	15	14

Syllable frequencies have been computed based on the CELEX database (Baayen *et al.*, 1993). Further values were taken from the online database 'Wortschatz Leipzig' (University of Leipzig, http://wortschatz.uni-leipzig.de/). Values in brackets display the respective confidence interval (CI).
a The average is biased by the use of three articles, which display very high frequencies in German. Formulaic and non-formulaic lyrics, however, do not include articles, since articles are generally not part of formulaic expressions in German.

whether the patients with additional right hemisphere lesions were included or not. As expected, patients with left hemisphere lesions produced more correctly intoned notes (75%; range: 22–96%) than patients with additional right hemisphere lesions (25%; range: 0–47%).

Inter-rater reliabilities for articulatory quality and pitch accuracy in each subject resulted in correlations ranging from 0.93 to 1.00, $P(16) < 0.001$, with an overall inter-rater reliability across participants of 0.98, $P(304) < 0.001$.

Average scores, composed of two raters' judgements for each condition and patient, were computed separately for articulatory quality and pitch accuracy. A repeated-measures ANOVA was performed based on the average scores for articulatory quality in each condition, including the factors modality (sung, spoken, spoken arrhythmic control) and lyrics (original, formulaic, non-formulaic) with patients' age and composite basal ganglia lesion scale as covariates. An alpha level of 0.05 and the Bonferroni correction for multiple comparisons were applied. For frequency analyses we used the software 'Praat' (Boersma and Weenink, 2011).

Results

Part 1: Melodic intoning

This section deals with the issue of whether singing may have proven to be beneficial for speech production in the non-fluent aphasic patients we studied. A repeated-measures ANOVA based on articulatory quality did not indicate an effect of melodic intoning as contrasted with the spoken conditions [$F(1) = 0.55$, NS], nor did a pair-wise comparison of the means reveal a difference between melodic intoning [$M = 53.47$, 95% confidence interval (CI) 41.76–65.18] and rhythmic speech [$M = 56.32$, 95% CI 43.43–69.21, NS]. These results did not change when three patients with additional right hemisphere lesions were excluded. Moreover, we assessed whether the absence of an effect from melodic intoning was found for each lyric type separately. No interaction of modality and lyrics was discovered [$F(2, 766) = 0.51$, NS]. In other words, there was no effect of singing on articulatory quality as compared with rhythmic speech whichever lyric type was used. Means of the results for the conditions melodic intoning and rhythmic speech, separately for each lyric type, are shown in Fig. 3.

To further explore these findings, several *post hoc* analyses were performed. We investigated whether the amount of fundamental frequency variability in the patients' utterances had an effect on articulatory quality. 'Praat' was used to quantify the fundamental frequency variances in the conditions melodic intoning and rhythmic speech separately for each lyric type. In a next step, we computed relative values for fundamental frequency variance and articulatory quality. Each of these variables was expressed as a difference between the conditions melodic intoning and rhythmic speech. Relative values were chosen instead of absolute values to control for interindividual differences. Based on these values, a correlation between fundamental frequency variance and articulatory quality did not yield significant results [$r(16) = -0.19$, NS]. This finding was independent of whether all or specific lyric types were considered.
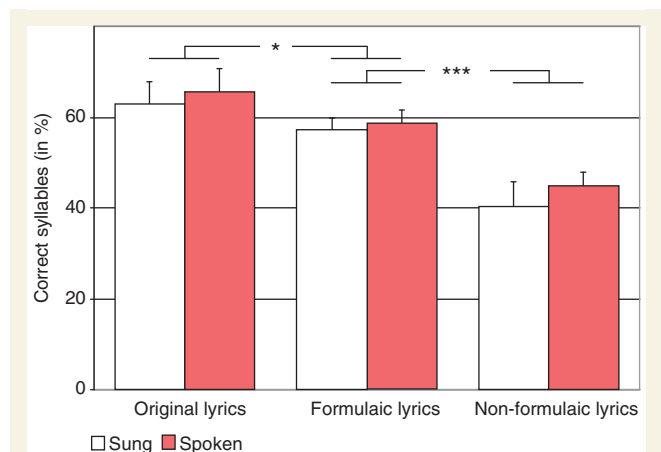

Figure 3 Correctly produced syllables in the conditions melodic intoning (sung) and rhythmic speech (spoken) for three lyric types. Articulatory quality significantly differed for each lyric type, irrespective of whether sung or spoken (*$P < 0.05$; ***$P < 0.001$). Error bars represent confidence intervals corrected for between-subject variance (Loftus and Masson, 1994).

Finally, we addressed the question of whether pitch accuracy in the sung conditions had any impact on articulatory quality. It should be noted that pitch accuracy is conceptually unrelated to frequency variability, as frequency variability reflects the amount of frequency changes over time, irrespective of whether these frequency changes are consistent with the melody or a prosodic pattern. A correlation analysis of pitch accuracy with relative articulatory quality did not yield significant results [$r(16) = 0.29$, NS]. Notably, this finding was independent of whether all or only left hemisphere lesion patients were included.

Interim discussion

Our data do not confirm an effect of singing on speech production in non-fluent aphasic patients. This finding holds true when comparing melodic intoning with natural prosody in rhythmic speech. One may nevertheless claim that frequency variation as such, sung or spoken, could still have positive effects on speech production by engaging the right hemisphere. Yet, no relationship was observed between frequency variability in the patients' utterances and articulatory quality. Our data thus do not support the assumption that frequency variation may facilitate speech production. However, aphasia often co-occurs with amusia, an impairment including the inability to hit the right notes. One may therefore conclude that the patients failed to benefit from singing because they were lacking pitch accuracy. It should therefore be noted that pitch accuracy and articulatory quality were found to be unrelated in our data. This means that patients with good pitch accuracy did not benefit more from singing, whereas patients with poor pitch accuracy did not benefit less from singing. In other words, melodic intoning, frequency variation and pitch accuracy did not affect speech production in the current patient sample.

Whichever lyric type was used, an effect from melodic intoning was consistently absent. Surprisingly, even with original,

well-known song lyrics there was no advantage for singing. High familiarity with the melody therefore failed to help the patients to produce the original lyrics. This finding is in line with earlier case reports (Hébert *et al.*, 2003; Straube *et al.*, 2008). Moreover, if high familiarity with a melody had constrained the patients' sung production of novel lyrics we would have expected worse performance for the sung novel lyrics. However, this was not the case.

Looking closer at one of the few studies that provide evidence for the superiority of singing above natural speech (Racette *et al.*, 2006), one reason for this result may be the use of earphones, which could have altered natural vocal self-monitoring. This view is indirectly supported by research on stuttering patients (Stuart *et al.*, 2008). Moreover, a *post hoc* analysis in that study revealed longer syllable durations for melodic intoning as compared with natural speech. Hence, slowing down of tempo during singing may have caused these patients to commit fewer errors. One further reason may be that the study was conducted in French, a syllable-timed language. English or German, however, is stress-timed language, which predetermines a certain metre in each phrase. Consequently, singing in French could entail a distinct gain in rhythmicity above natural speech, whereas this would not similarly apply in stress-timed languages. Singing in a syllable-timed language such as French may therefore be thought of as 'rhythm in disguise'. It is noteworthy that singing in French was only found to be an efficient tool when using a vocal playback to sing along with. This sung accompaniment may have served as a rhythmic pacemaker.

Part 2: Rhythmic speech

This section is dedicated to the question of whether rhythmicity may have affected speech production in the patients we studied. Based on articulatory quality, a pair-wise comparison of the means revealed a superiority of rhythmic speech [$M = 56.32$, 95% CI 43.43–69.21] as contrasted with the arrhythmic control [$M = 54.60$, 95% CI 42.08–67.12, $P = 0.010$]. To further explore the relationship between basal ganglia lesions and rhythmicity, we included the composite basal ganglia lesion scale as a covariate. A contrast analysis indicated an interaction of basal ganglia lesions with rhythmic speech and the arrhythmic control [$F(1) = 16.90$, $P = 0.001$, partial $\eta^2 = 0.55$]. Such an interaction with basal ganglia lesions was not found for the melodic intoning and rhythmic speech. As indicated in Table 4 and Fig. 4, patients with larger basal ganglia lesions tended to perform worse in the arrhythmic control compared with rhythmic speech. This pattern was not

found in patients with smaller striatal lesions. Moreover, patients with larger basal ganglia lesions showed lower means throughout the experiment. As interindividual differences in lesion size may be responsible for this finding, it should be noted that our design was only sensitive to intraindividual differences.

To control whether rhythmic percussion beats in the spoken conditions may have interfered with speech production in the patients, we repeated the experiment with four control patients. No significant differences were found between the spoken conditions with and without rhythmic percussion beats.

Interim discussion

Our data suggest an effect of rhythm on speech production in non-fluent aphasics. Notably, the benefit from rhythm was found to be strongest in patients with lesions including the basal ganglia. This evidence points to a crucial contribution of the basal ganglia
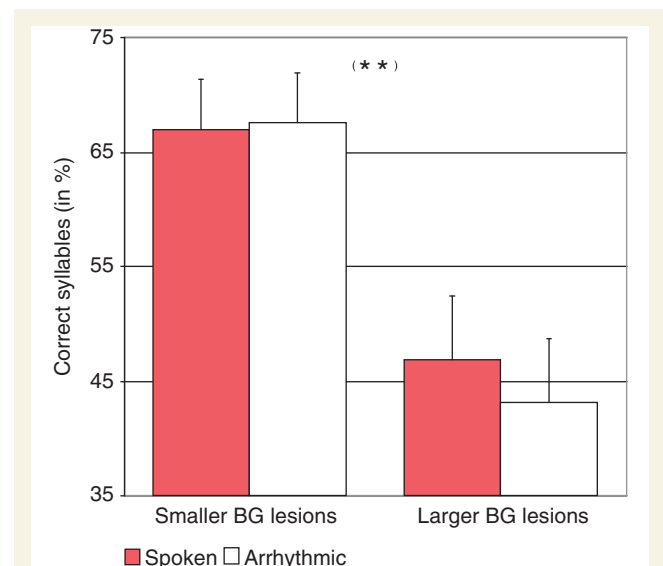

Figure 4 Correctly produced syllables in the conditions rhythmic speech (spoken) and the spoken arrhythmic control (arrhythmic) averaged across lyric types. The results show a significant interaction of basal ganglia (BG) lesions and rhythmicity (**$P < 0.01$). Nine patients with larger basal ganglia lesions (composite basal ganglia lesion score > 1.5) tended to perform worse in the arrhythmic control compared with rhythmic speech. This pattern was not found in eight patients with smaller basal ganglia lesions (composite basal ganglia lesion score $\leqslant 1.5$). Error bars represent confidence intervals corrected for between-subject variance (Loftus and Masson, 1994).

Table 4 Rhythm and basal ganglia lesions

Patient subgroup	Melodic intoning	Rhythmic speech	Arrhythmic control
Composite basal ganglia lesion score > 1.5 ($n = 9$)	42 (± 6.6)	47 (± 3.6)	43 (± 5.5)
Composite basal ganglia lesion score $\leqslant 1.5$ ($n = 8$)	67 (± 6.3)	67 (± 4.5)	68 (± 5.0)

Values represent correct syllables (in percentages), here averaged over lyric types. Values in brackets display confidence intervals corrected for between-subject variance (Loftus and Masson, 1994).

for rhythmic segmentation in speech production. Among the patients we studied, the extent of basal ganglia lesions accounted for ~55% of the variance related to rhythmicity.

One could assume that the use of rhythmic percussive accompaniments may have influenced the patients' utterances in the spoken conditions. However, the presence or absence of rhythmic percussion beats did not affect speech production in the control patients. Hence, it appears rather likely that percussive accompaniments do not interfere with speech production as long as they are rhythmic. This finding should nevertheless be viewed with caution as rhythmic percussion beats are usually not part of spoken utterances with natural prosody in everyday life.

To manipulate rhythmicity, we applied an arrhythmic interference paradigm. This method was chosen in order to keep syllable durations consistent as their impact on articulation is largely unknown. Nevertheless, it should be considered that arrhythmic control is not completely devoid of rhythm but rather provides a gradual decrease in perceived rhythmicity. One could therefore claim that, beyond the constraints of experimental control, the contributions from rhythm to speech production may be even more pronounced. Moreover, rhythm-related interventions in aphasia therapy, such as rhythmic hand tapping, may increase the benefit from rhythm to a considerable extent.

Part 3: Original, formulaic and non-formulaic lyrics

This section addresses the question of whether lyric memory and automaticity in formulaic expressions may have affected speech production in our patients. A repeated-measures ANOVA, based on articulatory quality, indicated a main effect of lyric type [$F(2) = 8.18$, $P = 0.002$], with higher means for original lyrics [$M = 63.53$, 95% CI 50.90–76.17] as opposed to formulaic lyrics [$M = 57.37$, 95% CI 44.84–69.89, $P = 0.027$]. To further explore whether this superiority may be age-dependent, we included the patients' age as a covariate. A contrast analysis revealed an interaction of age with original and formulaic lyrics [$F(1) = 13.18$, $P = 0.003$, partial $\eta^2 = 0.49$]. As can be seen in Table 5 and Fig. 5, the group of elderly patients showed a higher production of original, familiar lyrics as compared with novel lyrics. This difference was not confirmed in the younger group. As to preserved automaticity, higher means were found for formulaic lyrics [$M = 63.53$, 95% CI 50.90–76.17] as compared with non-formulaic lyrics [$M = 43.48$, 95% CI 30.93–56.03, $P < 0.001$]. Figure 3 shows the means for the three lyric types.

Interim discussion

The present data clearly indicate the importance of lyric memory for speech production in aphasics, irrespective of whether the lyrics are sung or spoken. This finding suggests that speech production may be mediated by long-term memory. Our results are therefore consistent with the clinical observation that non-fluent aphasics are sometimes found to be more fluent when producing well-known song lyrics compared with speaking spontaneously.

It should be noted that we controlled for automaticity in formulaic expressions, hence ruling out that memory effects may

Table 5 Memory and age

Patient subgroup	Original lyrics	Formulaic lyrics	Non-formulaic lyrics
Aged >55 years (n = 8)	71 (±7.7)	57 (±2.5)	43 (±7.3)
Aged ≤55 years (n = 9)	55 (±2.6)	57 (±3.3)	45 (±4.1)

Values represent correct syllables (in percentages), here averaged over modalities. Values in brackets display confidence intervals corrected for between-subject variance (Loftus and Masson, 1994).
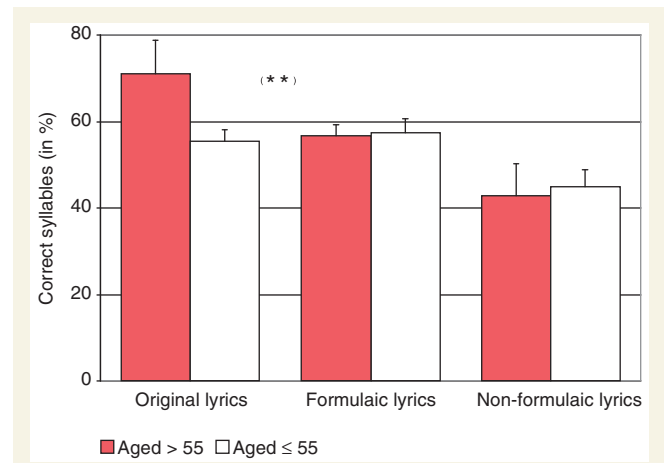

Figure 5 Correctly produced syllables of eight elderly patients (aged >55) and nine younger patients (aged ≤55), averaged over modalities. The results show a significant interaction of age and lyric memory (**$P < 0.01$). Only elderly patients showed an increased performance of original lyrics compared with formulaic lyrics. Error bars represent confidence intervals corrected for between-subject variance (Loftus and Masson, 1994).

actually be driven by overlearned motor sequences carried out in everyday life. In other words, this finding emphasizes that lyric memory and motor automaticity may affect speech production in different ways. Future work on the contribution of lyric memory to speech production in aphasics should therefore take into account to what extent familiar lyrics are automatized at the motor level.

The results also show that age may be crucial for the contribution of memory to speech production as it accounts for ~50% of the variance related to memory. Indeed, the contribution from memory appears comparably large in elderly patients, whereas it is absent in younger patients. One could, for example, claim that the group of elderly patients displayed a lower average in the production of novel lyrics, which could be construed as an advantage for original, familiar lyrics. However, this assumption is not compatible with our data as both younger and elderly patients showed similar results in the production of novel lyrics. Furthermore, age-dependency of the song familiarity is very unlikely to explain this finding as this factor was controlled for in an age-matched group. In other words, age appears be a promising factor in this context. This is all the more important as many studies with aphasics are based on single cases, hence not considering systematic differences related to age.

Finally, the data suggest that preserved motor automaticity in formulaic expressions may be indispensable for speech production in aphasics. The performance of formulaic lyrics showed a considerable superiority over non-formulaic lyrics in every single patient. Speech production in aphasics may therefore be largely mediated by motor automaticity, be it sung or spoken. This finding points to a crucial contribution of preserved automaticity to speech production in aphasics, even beyond questions related to melodic intonation therapy. One may, for instance, consider preserved automaticity as a highly valuable resource for speech therapy.

General discussion

In the current study, we aimed to assess the relative importance of various factors related to singing for speech production in 17 non-fluent aphasics. Contrary to some opinion, our results suggest that singing may *not* be decisive for speech production in non-fluent aphasics. Divergent findings in the past could very likely be a consequence of the acoustic setting, insufficient control of syllable duration or language-specific stress patterns (see Results). However, our results indicate that rhythm may be crucial, particularly for patients with lesions including the basal ganglia. It is noteworthy that lesions within the basal ganglia accounted for >50% of the variance related to rhythmicity. Our findings suggest that benefits typically attributed to melodic intoning in the past may actually have their roots in rhythm.

Moreover, our data demonstrate that *what* patients utter is at least as important as *how* it is uttered, irrespective of whether sung or spoken. Indeed, our data indicate that lyrics play a crucial role in speech production in non-fluent aphasics. Among the patients we studied, long-term memory and preserved motor automaticity appeared to strongly mediate speech production. Memory and automaticity may therefore help to explain effects that have, up until now, been presumed to result from singing. This is all the more critical because automatized, formulaic expressions were suggested to be lateralized in the right hemisphere (see Introduction). In light of this evidence it would seem that some important questions remain unresolved regarding the relationship between right hemisphere correlates in aphasics and melodic intonation therapy.

Acknowledgements

Funding

References

Aichert I, Ziegler W. Syllable frequency and syllable structure in apraxia of speech. Brain Lang 2004; 88: 148–59.

Albert ML, Sparks RW, Helm N. Melodic intonation therapy for aphasia. Arch Neurol 1973; 29: 130–1.

Baayen RH, Piepenbrock R, van Rijn H. The CELEX lexical data base [CD-ROM]. Philadelphia: Linguistic Data Consortium; 1993.

Belin P, Zilbovicius M, Remy P, Francois C, Guillaume S, Chain F, et al. Recovery from nonfluent aphasia after melodic intonation therapy: a PET study. Neurology 1996; 47: 1504–11.

Blank SC, Scott SK, Murphy K, Warburton E, Wise RJS. Speech production: Wernicke, Broca and beyond. Brain 2002; 125: 1829–38.

Blasi V, Young AC, Tansy AP, Petersen SE, Snyder AZ, Corbetta M. Word retrieval learning modulates right frontal cortex in patients with left frontal damage. Neuron 2002; 36: 159–70.

Boersma P, Weenink D. Praat: doing phonetics by computer [computer program]. Version 5.2.21; Amsterdam: Institute of Phonetic Sciences of the University of Amsterdam; 2011.

Boucher V, Garcia LJ, Fleurant J, Paradis J. Variable efficacy of rhythm and tone in melody-based interventions: implications for the assumption of a right-hemisphere facilitation in non-fluent aphasia. Aphasiology 2001; 15: 131–49.

Brendel B, Ziegler W. Effectiveness of metrical pacing in the treatment of apraxia of speech. Aphasiology 2008; 22: 77–102.

Callan DE, Tsytsarev V, Hanakawa T, Callan AM, Katsuhara M, Fukuyama H, et al. Song and speech: brain regions involved with perception and covert production. Neuroimage 2006; 31: 1327–42.

Cao Y, Vikingstad EM, George KP, Johnson AF, Welch KMA. Cortical language activation in stroke patients recovering from aphasia with functional MRI. Stroke 1999; 30: 2331–40.

Cohen NS, Ford J. The effect of musical cues on the nonpurposive speech of persons with aphasia. J Music Ther 1995; 32: 46–57.

Crowder RG, Serafine ML, Repp B. Physical interaction and association by contiguity in memory for the words and melodies of songs. Mem Cognition 1990; 18: 469–76.

Gentilucci M, Dalla Volta R. Spoken language and arm gestures are controlled by the same motor control system. Q J Exp Psychol 2008; 61: 944–57.

Gerstmann HL. A case of aphasia. J Speech Hearing Disord 1964; 29: 89–91.

Gordon RL, Schön D, Magne C, Astésano C, Besson M. Words and melody are intertwined in perception of sung words: EEG and behavioral evidence. PLoS One 2010; 5: 1–12.

Hébert S, Racette A, Gagnon L, Peretz I. Revisiting the dissociation between singing and speaking in expressive aphasia. Brain 2003; 126: 1838–50.

Helm-Estabrooks N, Nicholas M, Morgan A. Melodic intonation therapy. Manual. Austin, TX: Pro-Ed; 1989.

Huber W, Poeck K, Willmes K. The Aachen Aphasia Test. Adv Neurol 1984; 42: 291–303.

Hustad KC, Jones T, Dailey S. Implementing speech supplementation strategies: effects on intelligibility and speech rate of individuals with chronic severe dysarthria. J Speech Lang Hear Res 2003; 46: 462–74.

Kochanski G, Orphanidou C. What marks the beat of speech? J Acoust Soc Am 2008; 123: 2780–91.

Kotz SA, Schwartze M, Schmidt-Kassow M. Non-motor basal ganglia functions: a review and proposal for a model of sensory predictability in auditory language perception. Cortex 2009; 45: 982–90.

Loftus GR, Masson MEJ. Using confidence intervals in within-subject designs. Psychon Bull Rev 1994; 1: 476–90.

Mills CK. Treatment of aphasia by training. JAMA 1904; 43: 1940–9.

Musso M, Weiller C, Kiebel S, Muller SP, Bulau P, Rijntjes M. Training-induced brain plasticity in aphasia. Brain 1999; 122: 1781–90.

Özdemir E, Norton A, Schlaug G. Shared and distinct neural correlates of singing and speaking. Neuroimage 2006; 33: 628–35.

Peretz I, Radeau M, Arguin M. Two-way interactions between music and language: evidence from priming recognition of tune and lyrics in familiar songs. Mem Cognition 2004; 32: 142–52.

Pilon MA, McIntosh KW, Thaut MH. Auditory vs visual speech timing cues as external rate control to enhance verbal intelligibility in mixed spastic ataxic dysarthric speakers: a pilot study. Brain Inj 1998; 12: 793–803.

Racette A, Bard C, Peretz I. Making non-fluent aphasics speak: sing along!. Brain 2006; 129: 2571–84.

Riecker A, Ackermann H, Wildgruber D, Dogil G, Grodd W. Opposite hemispheric lateralization effects during speaking and singing at motor cortex, insula and cerebellum. Neuroreport 2000; 11: 1997–2000.

Rosen HJ, Petersen SE, Linenweber MR, Snyder AZ, White DA, Chapman L, et al. Neural correlates of recovery from aphasia after damage to left inferior frontal cortex. Neurology 2000; 55: 1883–94.

Ryding E, Bradvik B, Ingvar DH. Changes of regional cerebral blood flow measured simultaneously in the right and left hemisphere during automatic speech and humming. Brain 1987; 110: 1345–58.

Samson S, Zatorre RJ. Learning and retention of melodic and verbal information after unilateral temporal lobectomy. Neuropsychologia 1992; 30: 815–26.

Samson S, Zatorre RJ. Recognition memory for text and melody of songs after unilateral temporal lobe lesion: evidence for dual encoding. J Exp Psychol Learn Mem Cogn 1991; 17: 793–804.

Saur D, Lange R, Baumgaertner A, Schraknepper V, Willmes K, Rijntjes M, et al. Dynamics of language reorganization after stroke. Brain 2006; 129: 1371–84.

Schlaug G, Marchina S, Norton A. Evidence for plasticity in white-matter tracts of patients with chronic Broca's aphasia undergoing intense intonation-based speech therapy. Ann N Y Acad Sci 2009; 1169: 385–94.

Schlaug G, Marchina S, Norton A. From singing to speaking: why singing may lead to recovery of expressive language function in patients with Broca's aphasia. Music Percept 2008; 25: 315–23.

Sidtis D, Canterucci G, Katsnelson D. Effects of neurological damage on production of formulaic language. Clin Linguist Phon 2009; 23: 270–84.

Sidtis D, Postman WA. Formulaic expressions in spontaneous speech of left-and right-hemisphere-damaged subjects. Aphasiology 2006; 20: 411–26.

Sparks RW, Helm N, Albert ML. Aphasia rehabilitation resulting from melodic intonation therapy. Cortex 1974; 10: 303–16.

Speedie LJ, Wertman E, Ta'ir J, Heilman KM. Disruption of automatic speech following a right basal ganglia lesion. Neurology 1993; 43: 1768–74.

Straube T, Schulz A, Geipel K, Mentzel HJ, Miltner WHR. Dissociation between singing and speaking in expressive aphasia: the role of song familiarity. Neuropsychologia 2008; 46: 1505–12.

Stuart A, Frazier CL, Kalinowski J, Vos PW. The effect of frequency altered feedback on stuttering duration and type. J Speech Lang Hear Res 2008; 51: 889–97.

Wilson SJ, Parsons K, Reutens DC. Preserved singing in aphasia: a case study of the efficacy of melodic intonation therapy. Music Percept 2006; 24: 23–36.

Yamadori A, Osumi Y, Masuhara S, Okubo M. Preservation of singing in Broca's aphasia. J Neurol Neurosurg Psych 1977; 40: 221–4.

Zahn R, Drews E, Specht K, Kemeny S, Reith W, Willmes K, et al. Recovery of semantic word processing in global aphasia: a functional MRI study. Brain Res 2004; 18: 322–36.