Journal of Fluency Disorders 55 (2018) 120–134

Contents lists available at ScienceDirect

Journal of Fluency Disorders

journal homepage: www.elsevier.com/locate/jfludis

Assisted and unassisted recession of functional anomalies associated with dysprosody in adults who stutter

Katrin Neumann[a,*], Harald A. Euler[a], Malte Kob[b], Alexander Wolff von Gudenberg[c], Anne-Lise Giraud[d], Tobias Weissgerber[e], Christian A. Kell[f]

[a] Department of Phoniatrics and Pediatric Audiology, Clinic of Otorhinolaryngology, Head and Neck Surgery,St. Elisabeth-Hospital, Ruhr University Bochum, Bochum, Germany
[b] Erich-Thienhaus-Institute, University of Music Detmold, Detmold, Germany
[c] Institute of the Kassel Stuttering Therapy, Bad Emstal, Germany
[d] Département des Neuroscience Fondamentales, Université de Genève, Switzerland
[e] Department of Audiological Acoustics, Clinic of Otorhinolaryngology, Goethe University Frankfurt, Frankfurt am Main, Germany
[f] Brain Imaging Center and Department of Neurology, Goethe University Frankfurt, Frankfurt am Main, Germany

ABSTRACT

Purpose: Speech in persons who stutter (PWS) is associated with disturbed prosody (speech melody and intonation), which may impact communication. The neural correlates of PWS' altered prosody during speaking are not known, neither is how a speech-restructuring therapy affects prosody at both a behavioral and a cerebral level.
Methods: In this fMRI study, we explored group differences in brain activation associated with the production of different kinds of prosody in 13 male adults who stutter (AWS) before, directly after, and at least 1 year after an effective intensive fluency-shaping treatment, in 13 typically fluent-speaking control participants (CP), and in 13 males who had spontaneously recovered from stuttering during adulthood (RAWS), while sentences were read aloud with 'neutral', instructed emotional (happy), and linguistically driven (questioning) prosody. These activations were related to speech production acoustics.
Results: During pre-treatment prosody generation, the pars orbitalis of the left inferior frontal gyrus and the left anterior insula were activated less in AWS than in CP. The degree of hypo-activation correlated with acoustic measures of dysprosody. Paralleling the near-normalization of free speech melody following fluency-shaping therapy, AWS normalized the inferior frontal hypo-activation, sooner after treatment for generating emotional than linguistic prosody. Unassisted recovery was associated with an additional recruitment of cerebellar resources.
Conclusions: Fluency shaping therapy may restructure prosody, which approaches that of typically fluent-speaking people. Such a process may benefit from additional training of instructed emotional and linguistic prosody by inducing plasticity in the inferior frontal region which has developed abnormally during childhood in PWS.

1. INTRODUCTION

Prosody – the melody and intonation of speech – is an important carrier of human communication. It is a constituent of both

* Corresponding author at: Department of Phoniatrics and Pediatric Audiology, Clinic of Otorhinolaryngology, Head and Neck Surgery, St. Elisabeth-Hospital, Ruhr University Bochum, Bleichstr. 16, 44787 Bochum, Germany.
E-mail address: Katrin.Neumann@rub.de (K. Neumann).

http://dx.doi.org/10.1016/j.jfludis.2017.09.003
Received 26 October 2016; Received in revised form 4 September 2017; Accepted 5 September 2017
Available online 09 September 2017
0094-730X/ © 2017 Elsevier Inc. All rights reserved.

language perception and production that permits the emission and perception of linguistic and emotional information on contrast, focus, or other elements of language that may not be encoded by grammar or vocabulary. Persons interpret and respond to other persons' speech prosody and produce adequate prosodic cues themselves. Prosody is the most important and ontogenetically earliest processed feature of language development emerging prenatally. It provides access to early phoneme discrimination and phonotactic learning (Friederici, 2005).

Prosody is characterized by several suprasegmental acoustic voice and language parameters, such as the modulation of fundamental frequency; the rhythm, stress, and rate of speech; the sound or syllable duration, intensity, formant frequencies; and pauses between different phrases (intonational phrase boundaries). Beyond isolated acoustic parameters, voice-specific information plays an important role in communication. Prosody details depend on the specific phonological, grammatical, and tonal characteristics of a language but also on social, cultural, status-, gender-, and age-related circumstances.

Speech prosody may be sub-classified into emotional (or affective) and linguistic prosody. Processing of emotional information, for example of happiness, irony, sadness, fear, anger, or sarcasm, may be conveyed by various means of communication, such as propositional content, speech intonation, facial expression, and gestures (Ethofer et al., 2006). Both emitting and perceiving emotional information is central to social communication and evolutionarily relevant for immediate reactions. The processing of linguistic prosodic information enables a listener to decide whether an utterance is a statement, a question, or a command. It requires integrated processing of linguistic and prosodic aspects of speech, such as syntax-prosody mapping and constant adaptation to semantic information, to ensure disambiguity (Anderson & Carlson, 2010).

Dysprosodic or aprosodic speech production occurs in a variety of neurological and psychiatric disorders, such as autism, schizophrenia, cortical and subcortical brain damage, Parkinson's disease, cerebellar ataxia, post-traumatic stress disorder, multiple sclerosis, Alzheimer disease, alcohol abuse, fetal alcohol-exposure (Arnold et al., 2013; Casper et al., 2007; Hesling et al., 2010; Monnot et al., 2002; Ross & Monnot, 2011; Paulmann, Pell, & Kotz, 2009; Paulmann, Seifert, & Kotz, 2010; Paulmann & Pell, 2010; Rusz, Cmejla, Ruzickova, & Ruzicka, 2011; Wymer, Lindman, & Booksh, 2002), and in family members with variants of FOXP2 (Shriberg et al., 2006). Speech in persons who stutter (PWS) is also dysprosodic (Jäncke, Bauer, & Kalveram, 1996; Packman, Onslow, Richard, & Van Doorn, 1996). PWS speak with a flattened speech melody; even during stuttering-free speech PWS produce a significantly smaller range of their fundamental frequency (F0) than persons who do not stutter (PWNS) (Bosshardt, Sappok, Knipschild, & Hölscher, 1997; Healey, 1982). In addition, speech rhythm parameters are altered in the speech of PWS, even in fluent utterances (Maruthy, Venugopal, & Parakh, 2016). PWS have been reported to experience difficulties in placing sentence accent correctly; their stuttering episodes are located mainly on stressed syllables, a fixed timing pattern of speech enhances their fluency, and the intervals between stressed syllables are more variable in their speech, even in symptom-free passages, than in the speech of PWNS (Bergmann, 1986).

Because of the variability of stuttering symptoms with multifaceted combinations of syllable repetitions, prolongations, blocks, circumventions of utterances, or stressed pauses, the disturbance of prosody is highly variable and not easy to parametrize. Nevertheless, an effective stuttering therapy, in particular if it is aimed at speech fluency, should normalize the disturbed speech prosody. Some stuttering treatments, for example speech restructuring approaches such as fluency shaping (Neumann et al., 2016), deal explicitly with prosodic tools, for example, by practicing soft voice onsets or speech bows, thus impacting speech melody. Other treatments, with their focus on speech naturalness, for example stuttering modification (Van Riper, 1973), imply that prosody normalizes with the reduction of the frequency or severity of disfluencies.

Producing emotional and linguistic prosody activates the bilateral inferior frontal gyri, anterior insulae, and large parts of the temporal cortex, together with the striatum and cerebellum (Aziz-Zadeh, Sheng, & Gheytanchi, 2010; Pichon & Kell, 2013). Interestingly, most of these regions are also structurally or functionally altered in PWS compared with non-stuttering control participants (CP) (for example Beal et al., 2010; Beal, Gracco, Lafaille, & De Nil, 2007; Brown, Ingham, Ingham, Laird, & Fox, 2005; Chang, Erickson, Ambrose, Hasegawa-Johnson, & Ludlow, 2008; Cykowski et al., 2007; De Nil & Kroll, 2001; Giraud et al., 2008; Kell et al., 2009; Neumann et al., 2005; Watkins, Smith, Davis, & Howell, 2008). Prosody production in PWS has not yet been addressed with functional neuroimaging. Here our goal was to identify neural correlates of disturbed speech melody, as one aspect of dysprosody, in untreated adults who stutter (AWS) during the generation of instructed emotional and linguistic prosody. Additionally, we investigated whether a fluency-shaping therapy, consisting of an intensive computer-assisted treatment and two to three refresher weekends (for a description see Euler, Wolff von Gudenberg, Jung, & Neumann, 2009), modified intonation. We studied its associated changes in brain activity directly after a three-week in-patient intensive treatment and at least one year later.

The treatment modifies prosody during neutral speech by an initial reduction of speech tempo, by soft voice onsets, continuous phonation, and smooth bows between utterances. This speech-restructuring technique produces an initially unusual and new speech pattern which prevents the occurrence of disfluencies. Soft voice onsets are trained as raising intonation and continuous phonation is exercised by speech bows linking utterances softly to each other. Hence, speech melody is modulated constantly and voluntarily. Patients control these modulations by the computer feedback they get from the envelopes of both the aimed and the realized acoustic speech signals in the time domain by using templates. These changes of speech pattern initially impinge on emotional prosody by a voluntarily changed intonation that is not affect-related. Later on, a more natural sounding speech is targeted by combining both the emotional aspects of speech and the partially automatized new speech pattern. Furthermore, linguistic prosody is also altered by the treatment because the initial artificially-sounding speaking manner is paralleled by unusual phrase boundaries and a voluntarily falling and raising intonation, which in later treatment phases have to be merged with the linguistic speech structure. The treatment thus acts on prosodic speech properties, particularly on speech melody, even if it does not explicitly focus on instructed emotional or linguistic prosody, the facets that are studied here.

Speech rhythm and sound duration are other prosodic aspects that undergo changes during the treatment, in particular by slowing

speech and exercising a regular speech meter, but are not relevant to this study. Neither did we explicitly examine the intensity of speech sound because fundamental frequency and sound pressure level are correlated in voice production (Gramming, Sundberg, Ternström, Leanderson, & Perkins, 1988).

We hypothesized that speech melody would improve after fluency induction and that this improvement should manifest itself in changes in brain activation within and potentially also outside the physiological prosody control network.

Additionally, the effects of such assisted recovery were compared with the activation patterns in males who had recovered spontaneously from their stuttering in adulthood (unassisted recovery in formerly stuttering adults, RAWS). We hypothesized that their brain activation pattern would highlight regions involved in long-term compensation (see also Kell et al., 2009).

2. METHODS

2.1. Participants

The study reported here is an extension of our previous MRI study on therapy effects in AWS compared with recovery effects of former AWS who had recovered in adulthood from stuttering, and with typically fluent participants (Kell et al., 2009), involving the same subjects. In brief, an fMRI speech production experiment was performed with (1) 13 male adults who persisted in stuttering (AWS, in the previous publication named PS; mean age 27 years, range 18-39, mean handedness score 50, measured with the Edinburgh Handedness Inventory, Oldfield, 1971), (2) the same 13 subjects after a three-week intensive treatment phase of the Kassel Stuttering Therapy (Euler & Wolff von Gudenberg, 2000), which is a precision fluency shaping treatment modified after Webster (1974), (3) the same 13 subjects one to three years later (follow-up), after a one-year maintenance phase with regular or occasional self-training, which is an integral part of the Kassel Stuttering Therapy (long-term data not reported in the previous study), (4) 13 males who had recovered from stuttering to 1% SS (stuttered syllables) or less without a preceding therapy (RAWS, previously RS); mean age 40 years, range 16-65, mean handedness score 86), and (5) 13 male typically fluent speaking control participants (CP, in the previous publication named NS; mean age 30 years, range 23-44, mean handedness score 83). All subjects were native Germans and were matched for educational levels (no significant differences in the Kruskal-Wallis test). Because only RAWS were included whose recovery had been maintained for at least five years, there was a significant age difference between the AWS and RAWS groups ($p < .05$). Age was used as a covariate in the group comparisons (see below). Inclusion criteria for RAWS were: (a) age $\geq$16 years, (b) recovery after the 12th birthday, (c) confirmation of former stuttering by a corroborator (in 10 of 13 subjects available), (d) complete or substantial recovery to 1% SS or less, (e) period between last treatment before recovery at least one year (see Supplementary Tables 1 and 2 in Kell et al., 2009). Neurological or other speech-relevant disorders were excluded by a phoniatrician (physician specialized in communication disorders such as speech, language, and voice disorders) and a neurologist, and structural brain images confirmed the absence of focal or diffuse brain lesions. The study was approved by the Ethics Committee of the Medical Faculty of the Frankfurt University (no. 277/04). All subjects gave informed written consent for their participation.

2.2. Procedure

2.2.1. Behavioral assessment

Speech samples were collected during four different speaking conditions (for speech fluency measures, i.e. stuttering severity, defined as percentage of stuttered syllables (%SS), and speech naturalness data, see Supplementary Table 2 in Kell et al., 2009). From these speech situations, audio recordings of only the telephone interviews of the CP and of the AWS before treatment, directly after the intensive course and of seven AWS after the follow-up period were used in this study for the acoustic analysis of prosody during free speech, because they were the only audio samples that were conserved for post-hoc analyses. Nevertheless, the RAWS were kept in the fMRI analysis because their speech fluency measures (%SS and speech naturalness) approached those of the CP (Kell et al., 2009) but differed from them in the functional data (see below).

2.2.2. Experimental paradigm

Participants lay supine with their head immobilized. They viewed visual stimuli presented on a screen through a mirror. The task comprised three seconds of overt reading of declarative sentences under three different prosody conditions: (1) neutral (appropriate to a declarative nonsense sentence), (2) intonated as a question, by raising the voice at the end of the phrase (linguistic prosody), and (3) intonated happily, induced by the imagination of being happy. A covert reading condition served as an additional baseline. The group comparisons of the contrast ‘neutral overt reading > covert reading' are reported in Kell et al. (2009). Here, the group differences of the contrasts of producing ‘linguistic > neutral', ‘emotional > neutral' prosody, ‘linguistic > emotional', and ‘emotional > linguistic' prosody are reported.

The speech material consisted of 90 phonologically balanced, not affectively charged, syntactically identical German sentences (translated example: ‘Blue dwarves like to run through the grass.’) that were presented for three seconds in random order. Each participant had to read aloud 30 sentences in each condition (neutral, emotional, linguistic prosody), i.e. each sentence three times (together 90 sentences per subject). Participants were trained to perform the tasks properly before scanning, because pronouncing neutral sentences happily is difficult; they were instructed to pronounce the sentence as if they were really happy about what they said (for details on training, see Pichon & Kell, 2013).

Two to three seconds before each sentence, participants were instructed verbally about the task they had to perform at the upcoming stimulus. The inter-trial interval varied from 2 to 10 s with a mean of 6 s. Subjects had to stop speaking after 3 s, when the

screen turned black, but all of them completed the sentences adequately within this time limit.

AWS after therapy were instructed to talk 'normally' inside the scanner, without intentionally applying their newly acquired speaking technique. It is important to understand that the fMRI measurements did not focus on the use of the trained speech pattern but on the plastic cerebral changes associated with their acquisition, independently of the actual speech behavior. Before therapy, AWS had already read fluently inside the scanner, owing to a quasi-constant scanner noise and the isolation inside the bore (see also Kell, Neumann, Behrens, von Gudenberg, & Giraud, 2017). After therapy, AWS were asked not to use the newly acquired speaking technique inside the scanner deliberately, but to read freely as they did pre-treatment. Owing to the masking scanner noise and isolation in the scanner, it was unnecessary to apply the newly acquired technique. Because of this same speaking manner before and after therapy, imaging data are well comparable and allow the demonstration of trait- rather than state-related group differences and the robust treatment-related brain plasticity (for meta-analyses on trait- vs. state-related activation differences, see Belyk, Kraft, & Brown, 2015; Budde, Barron, & Fox, 2014). Because the scanner noise was identical during all conditions, direct consequences of scanner noise on brain activation are excluded, while fluency induction may alter activation patterns in general compared with free non-fluent speech.

Participants' speech was recorded with an MRI-compatible microphone (mr confon, Magdeburg, Germany); recordings were analyzed after filtering out the scanner noise (Adobe Audition, San Jose, USA) for task performance. Details on the continuous fMRI data acquisition are reported in Kell et al. (2009).

2.3. Data analysis

2.3.1. Behavioral data

Significant group differences with respect to stuttering severity, defined as percent stuttered syllables (%SS), speech rate, speech naturalness, and several self-report items (negative emotions associated with stuttering, speech aids used to speak fluently, self-perceived stuttering severity) were revealed by ANOVAs and subsequent t-tests, and are reported in Kell et al. (2009).

2.3.2. Acoustic analysis

In order to evaluate the acoustic consequences of the intensive therapy program on speech melody, speech segments of 300 continuous syllables (including stuttered and non-stuttered utterances) sampled from AWS during telephone interviews immediately before and after the intensive treatment phase and at least one year later (follow-up) were analyzed and compared with the values from typically fluent CP using the software Praat 5.3.03 (http://www.fon.hum.uva.nl/praat/). After segmentation into voiced and unvoiced speech components, the following parameters were analyzed for each subject: mean fundamental frequency (meanF0) and standard deviation of F0 (SDF0) as supra-segmental prosodic parameters, jitter (variation of F0 from one glottal cycle to another) and shimmer (variation of waveform amplitude from one glottal cycle to another) as two subsegmental, microprosodic parameters (Farrús, 2007). F0 represents what is perceived as average voice pitch by a listener and reflects vocal fold length and tension. Its modulation, SDF0, characterizes variations in the spectral domain and encodes emotional and linguistic content, even pre-attentively (e.g. Leitman, Sehatpour, Garidis, Gomez-Ramirez, & Javitt, 2011). SDF0 has been shown to be sensitive to instructed changes in produced prosody in a previous neuroimaging study (Pichon & Kell, 2013). Jitter and shimmer characterize microprosodic cycle-to-cycle variations of voice and speaking styles. These four parameters do not cover all aspects of prosody, such as rhythm, accent, dynamics, or sound intensity, but they assumedly represent stuttering-affected speech melody well. Indeed, most of these voice parameters showed a therapy-associated normalization (see Results).

Group values were compared by using independent and pairwise t-tests with the Bonferroni correction in SPSS. Individual values were used for a correlation analysis with the BOLD data (see below). Because not enough speech recordings were available for AWS at least one year after therapy, and for RAWS for these post-hoc analyses, only acoustic values from the CP and AWS before and directly after therapy were considered in the fMRI analyses.

2.3.3. Functional MRI data

The gradient echo planar images (EPI, repetition time 2s, echo time 30 ms, $3 \times 3 \times 3\ mm^3$) were spatially preprocessed (realignment, normalization, and smoothing with an 8 mm full-width at half maximum isotropic Gaussian kernel) using the standard parameters of SPM8. The data were analyzed in the framework of the general linear model; that is, the auditory cue was modelled as an event, and the conditions of interest (3 s of covert or overt reading) in the two sessions were modelled by using a boxcar function convolved with a canonical hemodynamic response function. Data were corrected for serial auto-correlations and globally normalized. Realignment parameters were modelled as covariates of no interest, in order to correct for movement artifacts.

Inferential statistics were performed first on the single subject level by contrasting 'intonation as a question > neutral prosody' which is thought to reveal linguistic prosody, and 'happy > neutral intonation' which indicates emotional prosody. The physiological emotional prosody production network is illustrated in Fig. 1 A, adapted from Pichon and Kell (2013). The prosody production network in the sample of 13 fluent male control participants was illustrated by reporting the results of two one-sample t-tests on the contrasts 'linguistic > neutral' and 'emotional > neutral' prosody in these participants. We additionally tested whether the observed activations in CP were specific for either linguistic or emotional prosody generation by contrasting prosody-related beta maps directly against each other. Results were analyzed with a threshold of $p < .05$, FWE(family-wise error correction)-corrected for multiple comparisons at the cluster level (cluster forming threshold at $p < .001$). In AWS before therapy, all voxels were tested for a correlation between brain activity for the contrast 'linguistic > neutral' and 'emotional > neutral' prosody with % SS, but no significant correlation was found (both $p > .001$, uncorrected).

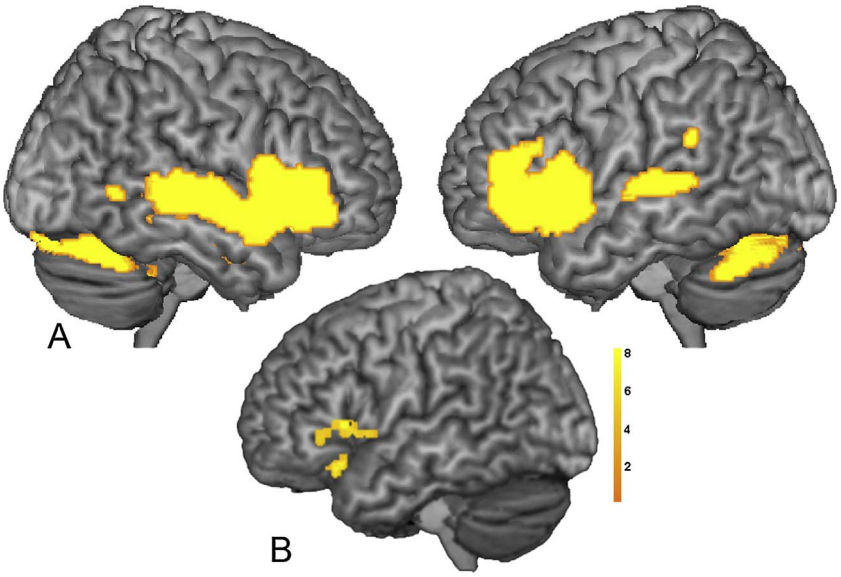

Fig. 1. The emotional prosody generation mask obtained from another study (Pichon & Kell, 2013) is illustrated in A) at $p < .001$, FEW-corrected. This mask was used to investigate brain activity during the prosody production in this experiment for a correlation with SDF0. B) Significant correlation was found between activity in the left inferior frontal gyrus and anterior insula during the generation of emotional > neutral prosody in AWS before therapy and their individual SDF0 during free speech ($p < .05$, FEW-cluster-corrected within the mask).

2.3.4. Inter-subject variability in the modulation of the fundamental frequency

It was additionally tested within groups whether the inter-subject variability in the modulation of the fundamental frequency during free speaking explained any variance in brain activation. This was justified because prosody-related changes in SDF0 modulate brain activity during speech production (Pichon & Kell, 2013). To this end, individual values of SDF0 during free speech were regressed over activation maps for the contrasts 'linguistic > neutral' prosody and 'emotional > neutral' prosody, separately for the groups. Results were considered significant if they occurred within the prosody production network detected in typical fluent CP at $p < .05$, FWE cluster-corrected in the mask volume. The mask, which was thresholded at $p < .001$ and FWE-corrected for multiple comparisons, was obtained from the data published in Pichon and Kell (2013), in order to prevent double-dipping (which would result from defining regions of interest and performing statistical tests in the same dataset) at $p < .05$, FWE cluster-corrected in the mask volume.

2.3.5. Group comparisons

Group comparisons were performed by comparing the beta weights of these two contrasts in either paired t-tests for therapy effects, or two-sample independent t-tests for all other group comparisons. Age and laterality index of the Edinburgh Handedness Inventory served as covariates. To ensure comparability with the report of neutral overt reading in Kell et al. (2009), group (AWS before therapy, AWS after therapy, RAWS, CP) by task (linguistic and emotional prosody vs. neutral prosody) and group by prosody type (linguistic vs. emotional prosody) interactions are reported as group differences at threshold $p < .001$, uncorrected and present only in clusters exceeding the expected size calculated from the smoothness of the data (Friston et al., 1996). The clinical relevance of the observed group effects is demonstrated by correlations with behavioral measures.

Coordinates of activations are given in the Montreal Neurological Institute (MNI) space. Brodmann areas corresponding to the activations were identified from probability maps in the anatomy toolbox for SPM (Eickhoff et al., 2005) or the stereotactic atlas of the human brain (Lancaster et al., 2000).

2.3.6. Post hoc correlations with structural parameters

In order to study the relationship between the observed prosody-related activations and structural brain parameters observed in the same participants, we correlated regional activity with previously published anomalous grey matter volume values in the dorsal inferior frontal gyrus (reduced grey matter volume in AWS) and anomalous values of fractional anisotropy (FA) in white matter pathways below the left anterior insula and orbitofrontal cortex (a measure of local white matter organization with locally increased values in AWS, Kell et al., 2009). To this aim, pre-treatment AWS' beta values for the contrasts emotional > neutral and linguistic > neutral prosody were extracted from the local maxima of observed group differences between typically speaking CP and AWS before therapy (left anterior insula and pars orbitalis of the left inferior frontal gyrus). AWS' individual contrast estimates were Pearson-correlated with grey matter density values in the dorsal inferior frontal gyrus and fractional anisotropy below the left anterior insula and orbitofrontal cortex. Significance was assumed at $p < .05$.

3. Results

3.1. Behavioral data

Detailed therapy outcome data of the involved AWS in comparison with speech data of the CP and RAWS are reported in Kell et al. (2009). The fluency-shaping therapy resulted in an improvement in speech fluency, speech naturalness, and items related to quality of life as reported by the AWS themselves. In brief, the stuttering severity, averaged across four speaking situations, dropped from 7.4%SS before treatment to 0.6%SS immediately after the intensive treatment phase and increased only slightly to 2.2%SS after at least one year of follow-up. The speech rate decreased from 194 to 179 syllables per minute and rose again to 190 syllables per minute, and the mean rating of the speech naturalness improved from 7.1 pre-treatment to 4.1 post-treatment to 3.8 at follow-up (1 = highly natural, 9 = highly unnatural). All these changes are comparable to the treatment-related improvements observed before (Euler et al., 2009; Euler et al., 2016; Euler and Wolff von Gudenberg, 2000).

The speech behavior in the scanner differed from the group-typical behavior assessed before fMRI sessions. Under the masking scanner noise, AWS spoke effortlessly and fluently before and after therapy with less than 1%SS and with comparable speech rates to those of the other participants. Here, this well-known masking effect is rather desired for group comparisons of functional data. This way the latter are not influenced by idiosyncratically varying extents of disfluency, depending on speech situations, and demonstrate the even under best fluent speech conditions for the AWS (the masking effect of the scanner noise) differing group-specific cerebral networks for speaking (Giraud et al., 2008; Neumann et al., 2005; Preibisch et al., 2003). The newly acquired speaking technique after therapy was also unlikely to affect the functional results directly because AWS were instructed not to use it deliberately.

Training AWS before therapy in performing the tasks (reading the neutral sentence happily or pronouncing it as a question) took longer and was more difficult than training RAWS or fluent controls, although this was not quantified.

3.2. Acoustic data

Group differences of acoustic parameters between AWS pre-treatment, post-treatment, at follow-up, and CP, as calculated by t-tests are given in Table 1. Note that the follow-up data set may be biased by incompleteness (data available for only seven of 13 AWS). Three of the four acoustic parameters examined during free speech (meanF0, SDF0, shimmer) in the yet untreated AWS differed significantly from the respective values of the CP. The mean fundamental frequency (meanF0) and its modulation SDF0, the supra-segmental parameters which were our main focus, showed significant group differences between AWS and CP (d = 1.40 and 1.88, respectively). After therapy, meanF0 and SDF0 had increased significantly and with large effects sizes (d = 1.49 and 1.26, respectively) and approached the values of the non-stuttering CP (Fig. 2). In other words, before the fluency-shaping therapy program AWS kept their voice lower, that is more monotonous and “safe”, and modulated their fundamental frequency less than fluent controls. Post-treatment, the differences between AWS and CP were no longer significant for meanF0 and SDF0. Thus, therapy normalized these two parameters. Shimmer values of AWS were not normalized by the therapy, and jitter values did not differ from those of CP before or after treatment. Because absolute jitter and shimmer values as extracted by Praat analysis have been shown only to provide complementary prosodic information, for example in speaker recognition (Farrús, 2007), these parameters were not used for correlation with functional imaging data. Because meanF0 provides only limited information about melody of speech, SDF0, as the best measure for speech melody, was correlated with the functional imaging data (see below).

The correlations between the acoustic parameters and the stuttering-related parameters %SS and speech naturalness were calculated in order to examine the direct relationship between stuttering and prosodic features and treatment-related changes of this relationship. There was neither a significant correlation between the pre-treatment or the post-treatment acoustic parameters with %SS before or after treatment, nor with the speech naturalness before or after treatment. There was also no correlation between the therapy-related normalization of meanF0 or its modulation (SDF0) and the reduction of stuttering severity or the increase in speech naturalness. However, the pre-treatment SDF0 correlated negatively with the post-treatment improvement in SDF0 (r = −.69, p = .009). This is not surprising, because for AWS who master their speech melody well without treatment, a lower improvement after therapy is to be expected, as opposed to those with low pre-treatment variability.

Table 1
Means and confidence intervals of the four acoustic parameters for the four groups (first four data columns) and the p values (t-tests) of the difference between CP and the three AWS groups (last three data columns); p values uncorrected; values that remained significant after Bonferroni correction (p < .0125) marked by asterisk; [1]incomplete follow-up data of only 7 of 13 AWS.

Variable	Means (95% confidence interval)				p Value of difference between CP and		
	CP	AWS pre	AWS post	AWS 1 + yr	AWS pre	AWS post	AWS 1 + yr
MeanF0	136.6 (123.7-149.5)	114.1 (109.5-118.8)	137.5 (127.5-147.5)	120.6 (113.6-127.8)	.003*	.904	.026
SDF0	31.8 (29.0-34.6)	20.7 (16.5-24.9)	28.9 (25.2-32.6)	22.4 (19.6-25.2)	< .001**	.196	< .001**
Jitter	2.65 (2.31-2.98)	2.18 (1.90-2.46)	2.75 (2.38-3.12)	1.96 (1.72-2.21)	.029	.655	< .001**
Shimmer	13.2 (12.1-14.3)	16.4 (15.4-17.4)	17.6 (16.1-19.0)	15.7 (14.2-17.1)	< .001**	< .001**	< .001**

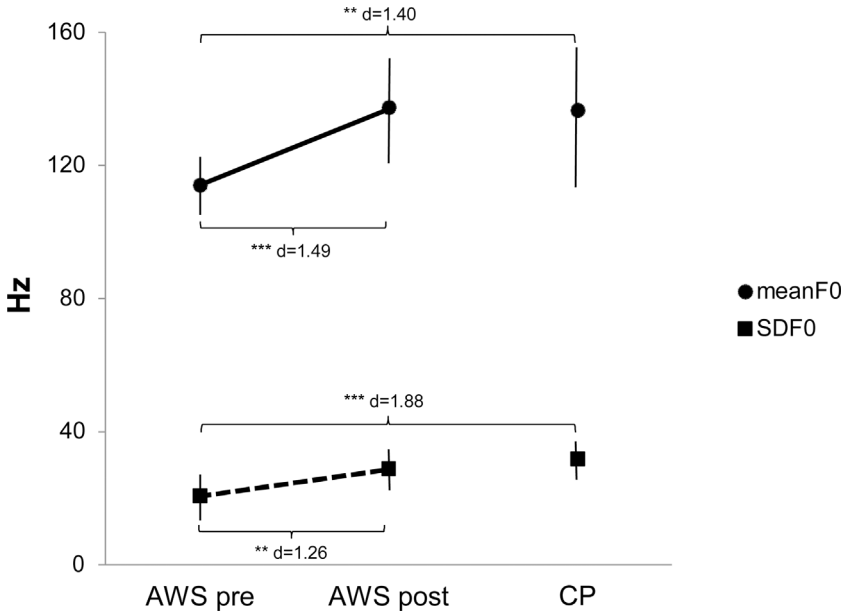

Fig. 2. Mean fundamental frequency and its modulation during free speech. Mean values of F0 and SDF0 and their standard deviations are plotted for CP and for AWS before (pre) and directly after (post) an intensive fluency-shaping treatment. AWS spoke with lower mean F0 and modulated F0 less compared to CP at pre-treatment. Therapy markedly improved F0 ($d = 1.49$) and SDF0 in AWS ($d = 1.26$) approaching CP values; ** = $p < .01$; *** = $p < .001$. Because of incompleteness the follow-up data are not shown here.

3.3. Functional data and their relationship with structural anomalies

The cerebral activation maps for the production of linguistic and emotional prosody were contrasted with those of the generation of neutral prosody. We have done this previously in a larger cohort of fluent CP and revealed a bilateral network including the anterior insula, inferior frontal gyrus, lateral orbitofrontal cortex, superior temporal gyrus and sulcus, together with the bilateral dorsal striatum and superior cerebellum that serves prosody control (Pichon & Kell, 2013, see also Fig. 1A). Given the small number of participants in this study, only clusters in the left hemisphere survived correction for multiple comparisons. For the generation of linguistic prosody, fluent CP activated the left inferior frontal gyrus and the left anterior insula more strongly than for the generation of neutral prosody (blue clusters in Fig. 3, left panel; cluster size 1026, $p < .001$, local maxima at MNI −48, 36, −4 and −38, 22, −6). Two other more strongly activated clusters for pronouncing sentences as a question were the left SMA (cluster size 574, $p < .001$, local maximum at MNI −4, −12, 62) and the left supramarginal gyrus (blue clusters in Fig. 3, left panel; cluster size 116, $p = .04$, local maximum at MNI −58, −46, 30). The latter cluster in the left supra-marginal gyrus was activated more strongly for the

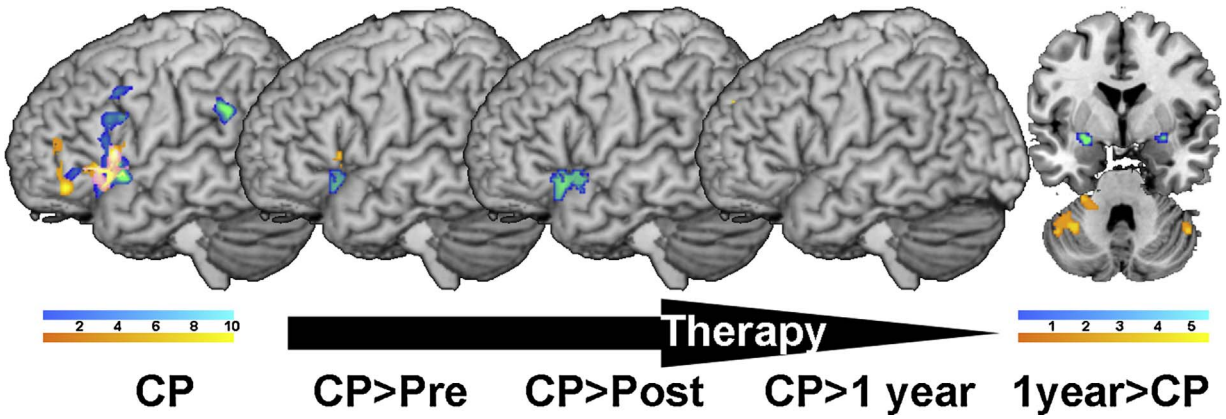

Fig. 3. Linguistic and emotional prosody generation. The left panel illustrates clusters that activate more strongly for the generation of linguistic (blue) and emotional (yellow) prosody than the generation of neutral prosody in typically fluent CP ($p < 0.05$, cluster-corrected, left color bar). All other panels depict group comparisons at $p < 0.001$, uncorrected for multiple comparisons (color bar on the right). The second panel illustrates hypoactivations in AWS before therapy (pre), the middle panel those directly after fluency-inducing therapy (post), and the fourth panel at least one year after therapy (at one year, no significant cortical differences, see Tables 2 and 3). The right panel documents overactivations in AWS at least one year after therapy in both amygdalae for the generation of linguistic (blue) and in both cerebellar hemispheres for the generation of emotional prosody (yellow).

Table 2
Group differences in brain activations between (1) adults who stutter (AWS) before (pre) and directly after (post) a fluency shaping therapy intensive treatment, and at least one year later (1 + yr), (2) recovered formerly stuttering adults (RAWS), and (3) non-stuttering control participants (CP), during the generation of linguistic prosody contrasted against neutral prosody. BA: Brodmann area, MNI: Montreal Neurological Institute, ns: not significant.

Region	BA	AWS pre > CP		AWS pre > AWS post, t values	AWS post > CP		AWS pre & AWS Post > RAWS, t values	RAWS > CP		AWS 1+ yr > AWS pre/AWS 1+ yr > AWS post/AWS 1+ yr > RAWS, t values	AWS 1+ yr > CP	
		MNI	*t* (cluster size)		MNI	*t* (cluster size)		MNI	*t* (cluster size)		MNI	*t* (cluster size)
L frontomesial	9	−4, 56, 22	−4.19 (43)									
R dorsal premotor cortex								22, 10, 56	−4.74 (23)	ns/ns/3.82		
L dorsal premotor cortex								−30, −2, 56	−4.52 (13)			
L inferior frontal gyrus	47/12	−34, 18, −14	−4.44 (76)									
L orbitofrontal region 11		−16, 12, −16	5.41 (78)									
R orbitofrontal region 11		22, 16, −20	4.76 (43)									
L orbitofrontal region 13		−12, 34, −18	4.27 (56)									
R orbitofrontal region 13		44, 42, −18	4.52 (48)									
R rolandic operculum	6	54, −2, 16	4.45 (35)									
L anterior insula	11	−20, 4, −8	−5.59 (47)		38, −10, 8	−5.02 (268)				4.09/4.73/ns		
L superior parietal	7	−16, −62, 68	4.33 (42)									
Table cont'd on next page												
Table cont'd from previous page												
R inferior parietal lobule	40	60, −40, 38	−4.47 (43)	−4.90								
R posterior cingulate	31	0, −14, 44	−3.87 (51)		4, −10, 52	-4.08, (40)		0, −14, 48	−4.59 (84)	ns/ns/3.23		
L dorsal striatum								−16, 8, 10	−4.11 (27)		−8, 0, 12	−4.35 (62)
R dorsal striatum											6, 0, 10	−3.79 (33)
L amygdala										5.06/7.23/ns	−24, −2, −12	4.61 (41)
R amygdala										5.38/5.98/ns	28, −4, −10	5.19 (46)
R ventral striatum											20, 14, −14	4.53 (36)
L cerebellar hemisphere							−4.27/ns	−34, −72, −26	4.07 (28)	ns/ns/−3.52		
R cerebellar hemisphere								40, −74, −26	4.00 (11)			

Table 3
Group differences in brain activations between (1) adults who stutter (AWS) before (pre) and directly after (post) a fluency shaping therapy intensive treatment, and at least one year later (1 + yr), (2) recovered formerly stuttering adults (RAWS), and (3) non-stuttering control participants (CP), during the generation of emotional prosody contrasted against neutral prosody. BA: Brodmann area, MNI: Montreal Neurological Institute, ns: not significant.

Region	BA	AWS pre > CP		AWS pre > AWS post, t values	AWS post > CP		AWS pre & AWS Post > RAWS, t values	RAWS > CP		AWS 1 + yr > AWS pre/AWS 1 + yr > AWS post/AWS 1 + yr > RAWS, t values	AWS 1 + yr > CP	
		MNI	*t* (cluster size)		MNI	*t* (cluster size)		MNI	*t* (cluster size)		MNI	*t* (cluster size)
L frontomesial	10	10, 54, 4	−4.11 (32)									
L middle frontal	10	−22, 56, 22	−4.68 (56)									
L anterior insula	11	−36, 20, 2	−4.17 (33)	−3.00*								
R anterior cingulate	24	0, 26, 16	−5.51 (98)									
L anterior cingulate	24	−2, 54, 44	−4.07 (33)									
R precentral	4							51, 10, 30	−4.67 (19)			
L inferior temporal	20			−4.70	−65, −22, −16	4.03 (20)						
L cerebellar hemisphere		−28, −52, −28	4.02 (23)		−30, −22, −16	3.99 (19)	4.28/ns				−26, −34, −24	5.46 (100)
L cerebellar hemisphere										ns/ns/4.13	−42, −44, −34	4.49 (111)
R cerebellar hemisphere											48, −54, −32	4.26 (24)

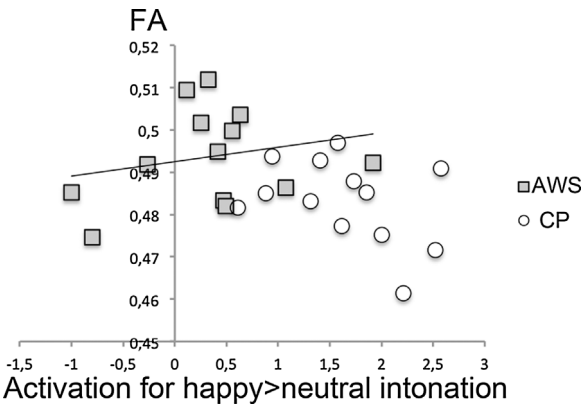

Fig. 4. Structure-function relationship in the left anterior insula. The contrast estimates for happy > neutral intonation are plotted on the x-axis and fractional anisotropy (FA) is plotted on the y-axis. AWS hypo-activate this region during the generation of happy intonation. The positive correlation with white matter integrity suggests that those AWS who activate more strongly, i.e. more physiologically, show increased FA values in white matter pathways directly below.

generation of linguistic than emotional prosody (cluster size 204, $p = .038$). For the production of emotional prosody, only a cluster in the left inferior frontal gyrus and the anterior insula survived the conservative statistical threshold (yellow cluster in Fig. 3, left panel; cluster size 883, $p < .001$, local maxima at MNI −48, 16, 4 and −34, 18, 4).

There was a significant correlation between individual SDF0 values during free speech and brain activity during the generation of emotional prosody only in AWS before therapy. The latter showed the highest variability in the acoustic prosodic marker SDF0 compared with AWS after therapy and CP (see SDF0 standard deviation bar in Fig. 2). Brain activity in the left inferior frontal gyrus (cluster size 137, $p = .014$, MNI −50, 20, 6) and the left anterior insula (cluster size 90, $p = .047$, MNI −34, 22, −16) during the generation of emotional > neutral prosody correlated with individual SDF0 values in the sense that the less AWS activated these regions before therapy, the less they modulated their fundamental frequency during free speech (Fig. 1B). There was no significant correlation of prosody-related activity with stuttering severity (all $p > .05$).

AWS before therapy activated the left anterior insula and the caudal part of the inferior frontal gyrus, pars orbitalis, less than did fluent CP, both for the generation of linguistic and emotional prosody (Fig. 3, second panel, and Tables 2 and 3). AWS showed increased fractional anisotropy in fibers below the left anterior insula (see Kell et al., 2009). The stronger the fractional anisotropy was increased in the white matter of this region, the stronger AWS activated the adjacent cortex (more physiologically) during the generation of happy > neutral intonation ($r = .571$, $p = .041$, see Fig. 4). The AWS overactivated the bilateral medial orbitofrontal cortex, the right frontal operculum, and the left superior parietal cortex for the generation of linguistic prosody (Table 2) and the left cerebellum for the generation of emotional prosody (Table 3) only before therapy. Directly after the fluency-shaping therapy program, the left inferior frontal activity normalized for the generation of emotional prosody but not for the generation of linguistic prosody (Fig. 3, middle panel, and Tables 2 and 3). All the overactivations that were observed before therapy vanished directly after therapy, except for the left cerebellar activation during emotional prosody generation, which persisted at least one year after therapy (Fig. 3, right panel and Table 3). During the generation of emotional prosody, it was this left cerebellar activation that distinguished AWS from RAWS. RAWS differed from CP during emotional prosody generation only in decreased activation of the right primary motor cortex (Table 3).

The activation pattern during the generation of linguistic prosody revealed larger differences between AWS before therapy and CP (see Table 2) and took longer to normalize than emotional prosody. It was only after at least one year of practicing the new speaking technique acquired during intensive fluency-shaping course that the left inferior frontal/anterior insula hypoactivation in AWS normalized during the generation of linguistic prosody (Fig. 3, fourth panel). This normalization of left inferior frontal activation was paralleled by a new hypoactivation in the bilateral dorsal striatum (see Table 2) and an activation of limbic regions at least one year after therapy (amygdala and ventral striatum, see Fig. 3, right panel and Table 2). A dorsal striatal hypoactivation during the generation of linguistic prosody was also observed in spontaneous RAWS. Compared with CP, RAWS suppressed activity in both dorsal premotor cortices during generating linguistic prosody (see Table 2).

4. DISCUSSION

4.1. Findings

This fMRI study compared the brain-functional correlates of the production of two kinds of speech melody carrying either a linguistic or an emotional content, between stuttering and non-stuttering participants. The main finding was a reduced activation of the left inferior frontal gyrus and anterior insula during the generation of both emotional and linguistic prosody in adult males who stuttered before therapy, when compared with fluent control participants. This hypoactivation normalized after a successful fluency-shaping treatment, which led to a modification of the speech melody, nearly normalized its acoustic prosodic marker ‘modulation of fundamental frequency', and successfully reduced stuttering symptoms. It may be noted that the speech melody modification trained

in the fluency-shaping therapy did not target emotional or linguistic prosody explicitly; rather, the treatment changed the overall speech pattern, usually irrespective of local prosodic events, and thus normalized prosody secondarily. The improvement in modulation of fundamental frequency during free speech after therapy could potentially be interpreted as a result of achieved fluency and implicit secondarily normalized prosody. However, there was no direct correlation between disfluency measures and modulation of fundamental frequency, either before or after therapy, or for therapy-related improvements, which speaks against a direct relationship between both domains. This suggests that dysprosody is not a mere consequence of the pathophysiological processes that result in disfluencies, and vice versa: the post-treatment reduction of stuttering severity was not a direct consequence of manipulating prosody. Indeed, imaging results suggest different mechanisms. Our assumption is supported by the fact that AWS spoke fluently inside the scanner, owing to the communicative isolation and scanner noise, which is known to induce fluency. But even under these fluency-inducing conditions, the AWS failed to activate the left inferior frontal region to the same extent as did the fluent CP during prosody generation. Instead, left inferior frontal hypoactivity was directly related to dysprosody at the behavioral level, where the less AWS activated this region, the less they modulated their fundamental frequency during speaking. The observed hypoactivation was not related to differences in the cognitive baseline, which here would refer to the brain activation during overt reading with neutral intonation, because no group differences were observed for the generation of neutral prosody in this region (Kell et al., 2009). Our results suggest that dysprosody constitutes a symptom that is not directly related to stuttering severity. However, training prosody indirectly may contribute to reducing stuttering severity by promoting fluency.

Former breakdown theories on stuttering, which regard stuttering symptoms as a failure of the complicate coordination of the speech-generating systems under some type of pressure, have regarded stuttering a prosodic disorder (Wingate, 1985). Acknowledging the difficulties of PWS in performing sound transitions smoothly it has been recognized that these difficulties concern stressed syllables (Wingate, 1976). Wingate considered the fact that PWS frequently performed poorly in linguistic tests as an indication that it involves processes beyond pure motor execution. He assumed that the planning of an utterance and its generation were poorly synchronized. The representation of the initial phonemes would be depleted and would disturb the phonological encoding process. This would lead, particularly in stressed syllables, to difficulties to generate the syllable nucleus (core of the syllable, most often a vowel or a vowel combination that carries the most sound energy of a syllable) and would result in a prosodic disturbance (Wingate, 1988). Bergmann (1986) has emphasized the association between theories of stuttering as prosodic disturbance and speech motor theories because the production of stressed syllables would require an enhanced effort of planning and execution. Our findings do not support these theories because they indicate a cerebral representation of dysfunctional prosodic networks in the brains of AWS even in the complete absence of stuttering symptoms.

The reduced modulation of the fundamental frequency in voiced speech in untreated AWS indicates a less variable use of the vocal fold muscles resulting in a monotonous F0 contour and therefore a reduced speech melody. The lowered muscle tension also causes a lowering of the absolute F0. Both effects are confirmed by the results for each individual meanF0 value (increases between 3 Hz and 49 Hz) and standard deviation value (only one decrease of 9.5 Hz otherwise increases between 1 Hz and 19 Hz) as well as for the overall meanF0 (increase of 23 Hz) and its standard deviation (increase of 8 Hz). The increase of both meanF0 and SDF0 after treatment, approaching the values of the control group, demonstrates a normalized laryngeal muscle tension and an improved ability to modify this tension.

The observed therapy-associated functional normalization in the ventral inferior frontal gyrus occurred in close vicinity of the known structural anomalies in AWS. Our group has previously reported that the same AWS involved here had anomalous grey matter in more dorsal parts of the inferior frontal gyrus and altered white matter pathways below the left pars orbitalis of the inferior frontal gyrus (Kell et al., 2009). While a positive correlation between the degree of the grey matter anomaly in the left dorsal inferior frontal gyrus with stuttering severity indicates an involvement of this region in the pathophysiology of developmental stuttering, there was no such correlation for white matter integrity and anterior insula (Kell et al., 2009). The current results suggest a compensatory role of left-hemispheric increases in white matter ultrastructure, at least below the anterior insula, because those PWS with stronger increases in fractional anisotropy in this region activated the adjacent cortex more physiologically during prosody generation than did PWS with more normal values of fractional anisotropy in the same region. These results suggest that the therapeutic manipulation of prosody may be efficient in changing the cortical function in the direct vicinity of the observed structural anomalies. This finding is possibly comparable to the functional normalization associated with peri-lesional reorganization in the adjacent cerebral tissue that has been reported in stroke recovery (Seitz & Donnan, 2015). It may be noted that activity in the left anterior insula during the production of neutral prosody was correlated positively with stuttering severity but was not correlated with it after therapy (Kell et al., 2009). Given that group differences in this region were more pronounced and were significant for emotional and linguistic but not for neutral prosody, training prosody may be more efficient in inducing plasticity in this anomalous developed cortical area than training to speak without an emphasis on prosody.

Interestingly, functional normalization occurred earlier for emotional than for linguistic prosody. The normalization of activity associated with the generation of emotional prosody occurred right after the intensive treatment phase whereas the normalization of left inferior frontal activity during the generation of linguistic prosody required another year or more of practice. We assume that for AWS the phylogenetic more recent linguistic prosody is more difficult to manage than the emotional prosody, a statement that needs further confirmation by research. Owing to the necessity for integrated processing of linguistic and prosodic speech aspects, true also for typically fluent persons, the generation of linguistic prosody requires additional steps (e.g. syntax-prosody mapping) compared with emotional prosody (Anderson & Carlson, 2010).

Our finding suggests that future treatment developments could use the easier access to instructed emotional prosody in earlier phases of the treatment programs while training linguistic prosody could be used in later phases. It ought to be stated that emotional prosody in this study did not reflect intonating speech as a function of current affect but rather constituted an artificial condition in

which a neutral sentence was intonated happily. The activation pattern and phenomenology may change considerably when 'real' emotions interfere with speaking.

Remarkably, in patients with Parkinson disease and dysprosodic speech, the left inferior frontal gyrus is functionally less strongly connected with the left striatum than that in matched control participants, which supports our proposal that this region could be central in the control of produced prosody (Arnold, Gehrig, Gispert, Seifried, & Kell, 2013). A physiological role of both cortical and subcortical dopamine in prosody control can be deduced from a modulation of prosody-related left inferior frontal activity during prosody generation by polymorphisms of the dopamine degrading enzyme catechol-O-methyltransferase and the dopamine transporter 1 (Arnold et al., 2016). We take the liberty to speculate here that the relationship between developmental stuttering and dopamine proposed previously (Alm, 2004) could affect communicatively challenging prosody more strongly than neutral prosody.

The functional normalization in the left inferior frontal cortex could be interpreted as a local training effect, but this region was not the only one to show recovery-related changes in prosody-induced activation. Both assisted and unassisted recovery was associated with suppression of activity in the bilateral dorsal striatum during the generation of linguistic prosody compared with fluent controls. In RAWS, this suppression was accompanied by suppressed activity in the bilateral dorsal premotor cortex. Group differences with respect to the generation of emotional prosody were sub-threshold in the dorsal striatum. The dorsal striatum is physiologically engaged in the control of both linguistic and emotional prosody by its connection with the auditory cortex, the motor cortex, and the inferior frontal gyri (Pichon & Kell, 2013). This central position in the control of prosody could potentially result from a specialization of the dorsal striatum in sequencing and rhythm control (Ackermann & Riecker, 2010; Wymbs, Ingham, Ingham, Paolini, & Grafton, 2013). In AWS, activity in the bilateral dorsal striatum during production of neutral prosody is positively correlated with stuttering severity before therapy but not thereafter (Kell et al., 2009). Decreasing this neural activity (as assessed by cerebral blood flow measurements) in the left striatum predicted a successful stuttering therapy, which consisted of modifying short phonation intervals (Ingham, Wang, Ingham, Bother, & Grafton, 2013). Because hypoactivation of this subcortical brain region was observed both for assisted and unassisted recovery and because the observed effects could not be explained by baseline differences, the relative suppression compared with that of fluent controls seems to be meaningful. As this region is physiologically activated and not suppressed, the relative hypoactivation could be interpreted as a consequence of compensatory processes for stuttering elsewhere, rather than constituting an active suppression of physiological processes, although the latter cannot be entirely ruled out. Because RAWS also suppressed activity in both dorsal premotor cortices relative to CP, the former may need to suppress dorsal cortico-striatal loops for maintaining speech fluency.

The observed activation patterns in the cerebellum most likely point to compensatory processes. Unassisted recovery was associated with additional recruitment of cerebellar hemispheres during the generation of linguistic prosody, while the activity pattern of RAWS for the generation of emotional prosody was nearly normal. This suggests again that linguistic prosody constitutes a more difficult problem for persons who stutter and requires additional resources for compensation. During speaking with induced emotional prosody, the only region that RAWS activated less than CP was the right primary motor cortex, suggesting that unassisted recovery largely normalized networks involved in the control of emotional prosody. Assisted recovery instead relied on a cerebellar recruitment for emotional prosody control. Note that cerebellar activation does not automatically imply automatization of articulatory processes, because unassisted recovery disconnects the activated cerebellum from the articulatory neocortical network (Kell et al., 2017). Therapy-induced speech fluency was associated with additional recruitment of limbic regions for the generation of linguistic prosody. Whether this activation profile reflects involvement in compensation or an emotional reaction to this condition is speculative.

In summary, the most convincing recovery mechanism for disturbances in prosody control in male AWS lies in a recruitment of the left pars orbitalis of the inferior frontal gyrus and the anterior insula by means of local plasticity, most likely aided by cerebellar mechanisms.

4.2. Limitations

The difficulty in recruiting German men who had recovered from stuttering spontaneously in adulthood resulted in a group size of only 13 participants. In order to counterbalance groups we decided to include similar numbers of typically fluent control participants and AWS. This is at the lower end of group sizes required for fMRI studies. Nevertheless, this group size was sufficient to generate significant activation clusters that were corrected for multiple comparisons only in the left inferior frontal region of fluent controls. This finding suggests that the right hemispheric contribution is more variable than that one of the left hemisphere even in fluent control participants and that the study of these right-sided activations requires larger group sizes (20 participants in Pichon & Kell, 2013). As a consequence, even the group comparisons could have been biased to the left inferior frontal region because of signal-to-noise ratio issues. Thus, we cannot rule out Type II errors, especially with regard to group differences in the right hemisphere.

Although studying speech samples for prosody evaluation during free speech from telephone interviews might appear as a limitation of the study, we do not assume this to represent an important bias because undetected concomitants are of less importance in studying prosody. Additionally, a previous study revealed that stuttering severity during telephone calls was correlated highly with stuttering severity during face-to-face conversation (Euler & Wolff von Gudenberg, 2000).

We could not analyze sufficient speech recordings of RAWS or of AWS at least one year after therapy in order to quantify their modulation of F0 during free speech. Consequently, we had to rely upon the perceptual judgments of the examiner that the participants modulated their voice in a relatively normal-sounding prosody with sufficient validity to justify our guarded conclusions. Moreover, evaluating the prosody of participants while they read the sentences happily or as a question relied on perceptual judgments, considering that training sessions were not recorded and the in-scanner recordings were too noisy for performing

quantitative acoustic analyses. Finally, it ought to be considered that the extent of imagined happiness during happy intonation assumedly differed between individuals. Although meanF0 and SDF0 are certainly not the only variables that map prosody, we focused on F0 and its modulation as representing speech melody, because they might be target parameters of treatment and can easily be addressed. The lacking correlation between F0 modulation in the overt speech of AWS and their speech naturalness as perceived by others might be considered unexpected; however, speech naturalness, although a sufficient measure of prosody, is a perceptual descriptor and does not reflect prosody alone but also other meta-linguistic aspects of speech (such as pragmatics).

4.3. Conclusion

The data presented here suggest that, in developmental stuttering, dysprosody constitutes a symptom that is not a direct consequence of disfluencies or their underlying pathology. The less AWS modulated their voice during free, fluent speech, the less they activated the left inferior frontal region for the generation of linguistic or instructed emotional prosody. Overall, dysprosody in AWS was associated with hypoactivation of this region during prosody production. Prosody during free speech improved after the fluency-shaping intensive treatment phase. This improvement was paralleled by a normalization of left inferior frontal activation that occurred earlier in time for the generation of emotional than for linguistic prosody. To our knowledge, this study has shown for the first time a normalization of prosody-related activations after an effective stuttering therapy, and this, moreover, with the highest cognitive baseline (speaking overtly with neutral prosody).

For persons who stutter the cerebral processes involved in the production of the evolutionarily older emotional prosody seem to be easier to manage than for linguistic prosody; therefore we venture the suggestion that instructed happy intonation may help future stuttering therapies in inducing plasticity in the anomalous left inferior frontal region. Training linguistic prosody could secondarily normalize inferior frontal activation. Ideally, a randomized controlled study with sufficiently large cohorts should compare the best available treatment against a comparable treatment version that additionally focuses on training happy intonation without tapping into genuinely affective components.

Funding

This work was supported by the German Academic Exchange Service (senior author).

ACKNOWLEDGMENT

We thank Claudia Posenenske for her participation in the acquisition of the behavioral data of this study.

References

Ackermann, H., & Riecker, A. (2010). The contribution(s) of the insula to speech production: a review of the clinical and functional imaging literature. *Brain Structure and Function, 214*, 419–433.

Alm, P. A. (2004). Stuttering and the basal ganglia circuits: a critical review of possible relations. *Journal of Communication Disorders, 37*, 325–369.

Anderson, C., & Carlson, K. (2010). Syntactic structure guides prosody in temporarily ambiguous sentences. *Language and Speech, 53*, 472–493.

Arnold, C., Gehrig, J., Gispert, S., Seifried, C., & Kell, C. A. (2013). Pathomechanisms and compensatory efforts related to Parkinsonian speech. *NeuroImage Clinical, 4*, 82–97.

Arnold, C., Gispert, S., Bonig, H., von Wegner, F., Somasundaram, R., & Kell, C. A. (2016). Dopaminergic modulation of cognitive preparation for overt reading: Evidence from the study of genetic polymorphisms. *Cerebral Cortex, 26*, 1539–1557.

Aziz-Zadeh, L., Sheng, T., & Gheytanchi, A. (2010). Common premotor regions for the perception and production of prosody and correlations with empathy and prosodic ability. *PLoS One, 5*, e8759.

Beal, D. S., Gracco, V. L., Lafaille, S. J., & De Nil, L. F. (2007). Voxel-based morphometry of auditory and speech-related cortex in stutterers. *Neuro-Report, 18*, 1257–1260.

Beal, D. S., Cheyne, D. O., Gracco, V. L., Quraan, M. A., Taylor, M. J., & De Nil, L. F. (2010). Auditory evoked fields to vocalization during passive listening and active generation in adults who stutter. *Neuroimage, 52*, 1645–1653.

Belyk, M., Kraft, S. J., & Brown, S. (2015). Stuttering as a trait or state − an ALE metaanalysis of neuroimaging studies. [Meta-Analysis Research Support, Non-U: SGov't]. *European Journal of Neuroscience, 41*, 275–284.

Bergmann, G. (1986). Studies in stuttering as a prosodic disturbance. *Journal of Speech and Hearing Research, 29*, 290–300.

Bosshardt, H. G., Sappok, C., Knipschild, M., & Hölscher, C. (1997). Spontaneous imitation of fundamental frequency and speech rate by nonstutterers and stutterers. *Journal of Psycholinguistic Research, 26*, 425–448.

Brown, S., Ingham, R. J., Ingham, J. C., Laird, A. R., & Fox, P. T. (2005). Stuttered and fluent speech production: an ALE meta-analysis of functional neuroimaging studies. *Human Brain Mapping, 25*, 105–117.

Budde, K. S., Barron, D. S., & Fox, P. T. (2014). Stuttering, induced fluency: and natural fluency: a hierarchical series of activation likelihood estimation meta-analyses. *Brain & Language, 139*, 99–107.

Casper, M. A., Raphael, L. J., Harris, K. S., & Geibel, J. M. (2007). Speech prosody in cerebellar ataxia. *International Journal of Language and Communication Disorders, 42*, 407–426.

Chang, S. E., Erickson, K. I., Ambrose, N. G., Hasegawa-Johnson, M. A., & Ludlow, C. L. (2008). Brain anatomy differences in childhood stuttering. *Neuroimage, 39*, 1333–1344.

Cykowski, M. D., Kochunov, P. V., Ingham, R. J., Ingham, J. C., Mangin, J. F., Rivière, D., et al. (2007). Perisylvian sulcal morphology and cerebral asymmetry patterns in adults who stutter. *Cerebral Cortex, 18*, 571–583.

De Nil, L. F., & Kroll, R. M. (2001). Searching for the neural basis of stuttering treatment outcome: recent neuroimaging studies. *Clinical Linguistics and Phonetics, 15*, 163–168.

Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., et al. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage, 25*, 1325–1335.

Ethofer, T., Anders, S., Wiethoff, S., Erb, M., Herbert, C., Saur, R., Grodd, W., & Wildgruber, D. (2006). Effects of prosodic emotional intensity on activation of associative auditory cortex. *Neuroreport, 17*, 249–253.

Euler, H. A., Anders, K., Merkel, A., & Wolff von Gudenberg, A. (2016). Mindert eine globale Sprechrestrukturierung wie die Kasseler Stottertherapie (KST) stotterbegleitende negative Emotionen? [Can an integrative speech restructuring treatment reduce stuttering-related negative emotions?]. *Logos, 24*, 84–94.
Euler, H. A., & Wolff von Gudenberg, A. (2000). Die Kasseler Stottertherapie (KST): Ergebnisse einer computergestützten Biofeedbacktherapie für Erwachsene [The Kassel Stuttering Therapy (KST): Results of a computer-based biofeedback therapy for adults]. *Sprache Stimme Gehör, 24*, 71–79.
Euler, H. A., Wolff v. Gudenberg, A., Jung, K., & Neumann, K. (2009). Computergestützte Therapie bei Redeflussstörungen: Die langfristige Wirksamkeit der Kasseler Stottertherapie (KST) [Computer-assisted therapy for speech disfluencies: The long-term effectiveness of the Kassel Stuttering Therapy (KST)]. *Sprache·Stimme·Gehör, 33*, 193–201.
Farrús, M. (2007). Jitter and shimmer measurements for speaker recognition. *8th Annual Conference of the International Speech Communication Association* (pp. 778–781). International Speech Communication Association (ISCA).
Friederici, A. D. (2005). Neurophysiological markers of early language acquisition: from syllables to sentences. *Trends in Cognitive Sciences, 9*, 481–488.
Gramming, P., Sundberg, J., Ternström, S., Leanderson, R., & Perkins, W. H. (1988). Relationship between changes in voice pitch and loudness. *Journal of Voice, 2*, 118–126.
Giraud, A.-L., Neumann, K., Bachoud-Levi, A.-C., Wolff von Gudenberg, A., Euler, H. A., Lanfermann, H., et al. (2008). Severity of dysfluency correlates with basal ganglia activity in persistent developmental stuttering. *Brain and Language, 104*, 190–199.
Healey, E. C. (1982). Speaking fundamental frequency characteristics of stutterers and nonstutterers. *Journal of Communication Disorders, 15*, 21–29.
Hesling, I., Dilharreguy, B., Peppé, S., Amirault, M., Bouvard, M., & Allard, M. (2010). The integration of prosodic speech in high functioning autism: a preliminary FMRI study. *PLoS One, 5*, e11571.
Ingham, R. J., Wang, Y., Ingham, J. C., Bothe, A. K., & Grafton, S. T. (2013). Regional brain activity change predicts responsiveness to treatment for stuttering in adults. *Brain and Language, 127*, 510–519.
Jäncke, L., Bauer, A., & Kalveram, K. T. (1996). Duration of phonation under changing stress conditions in stuttering and nonstuttering adults. *Clinical Linguistics and Phonetics, 10*, 225–234.
Kell, C. A., Neumann, K., Behrens, M., von Gudenberg, A., & Giraud, A. L. (2017). Articulation-related functional connectivity changes associated with assisted and unassisted recovery from developmental stuttering. *Journal of Fluency Disorders*. http://dx.doi.org/10.1016/j.jfludis.2017.02.001 pii: S0094-730X(16)30094-8.
Kell, C. A., Neumann, K., von Kriegstein, K., Posenenske, C., Wolff von Gudenberg, A., Euler, H. A., et al. (2009). How the brain repairs stuttering. *Brain, 132*, 2747–2760.
Lancaster, J. L., Woldorff, M. G., Parsons, L. M., Liotti, M., Freitas, C. S., Rainey, L., et al. (2000). Automated Talairach atlas labels for functional brain mapping. *Human Brain Mapping, 10*, 120–131.
Leitman, D. I., Sehatpour, P., Garidis, C., Gomez-Ramirez, M., & Javitt, D. C. (2011). Preliminary evidence of pre-attentive distinctions of frequency-modulated tones that convey affect. *Frontiers in Human Neuroscience, 5*, 96. http://dx.doi.org/10.3389/fnhum.2011.00096.
Maruthy, S., Venugopal, S., & Parakh, P. (2016). Speech rhythm in Kannada speaking adults who stutter. *International Journal of Speech-Language Pathology,* 1–9.
Monnot, M., Lovallo, W. R., Nixon, S. J., & Ross, E. (2002). Neurological basis of deficits in affective prosody comprehension among alcoholics and fetal alcohol-exposed adults. *Journal of Neuropsychiatry and Clinical Neurosciences, 14*, 321–328.
Neumann, K., Euler, H. A., Bosshardt, H. G., Cook, S., Sandrieser, P., Schneider, P., et al. (2016). *Pathogenese, Diagnostik und Behandlung von Redeflussstörungen. Evidenz- und konsensbasierte S3-Leitlinie, AWMF-Registernummer 049-013, Version 1. [Pathogenesis, diagnostics, and treatment of fluency disorders. Evidence and consensus-based S3-guideline, AWMF-registry number 049-013, version 1. 2016].* Deutsche Gesellschaft für Phoniatrie und Pädaudiologie http://www.awmf.org/leitlinien/detail/ll/049–013. html. Accessed Oct 5, 2016.
Neumann, K., Preibisch, C., Euler, H. A., Wolff von Gudenberg, A., Lanfermann, H., Gall, V., et al. (2005). Cortical plasticity associated with stuttering therapy. *Journal of Fluency Disorders, 30*, 23–39.
Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia, 9*, 97–113.
Packman, A., Onslow, M., Richard, F., & Van Doorn, J. (1996). Syllabic stress and variability: A model of stuttering. *Clinical Linguistics and Phonetics, 10*, 235–263.
Paulmann, S., & Pell, M. D. (2010). Dynamic emotion processing in Parkinson's disease as a function of channel availability. *Journal of Clinical and Experimental Neuropsychology, 32* 822–135.
Paulmann, S., Pell, M. D., & Kotz, S. A. (2009). Comparative processing of emotional prosody and semantics following basal ganglia infarcts: ERP evidence of selective impairments for disgust and fear. *Brain Research, 1295*, 159–169.
Paulmann, S., Seifert, S., & Kotz, S. A. (2010). Orbito-frontal lesions cause impairment during late but not early emotional prosodic processing. *Social Neuroscience, 5*, 59–75.
Pichon, S., & Kell, C. A. (2013). Affective and sensorimotor components of emotional prosody generation. *Journal of Neuroscience, 33*, 1640–1650.
Preibisch, C., Neumann, K., Raab, P., Euler, H. A., von Gudenberg, A. W., Lanfermann, H., et al. (2003). Evidence for compensation for stuttering by the right frontal operculum. *NeuroImage, 20*, 1356–1364.
Ross, E. D., & Monnot, M. (2011). Affective prosody: What do comprehension errors tell us about hemispheric lateralization of emotions, sex and aging effects: and the role of cognitive appraisal. *Neuropsychologia, 49*, 866–877.
Rusz, J., Cmejla, R., Ruzickova, H., & Ruzicka, E. (2011). Quantitative acoustic measurements for characterization of speech and voice disorders in early untreated Parkinson's disease. *Journal of the Acoustical Society of America, 129*, 350–367.
Seitz, R. J., & Donnan, G. A. (2015). Recovery potential after acute stroke. *Frontiers in Neurology, 6*.
Shriberg, L. D., Ballard, K. J., Tomblin, J. B., Duffy, J. R., Odell, K. H., & Williams, C. A. (2006). Speech, prosody, and voice characteristics of a mother and daughter with a 7;13 translocation affecting FOXP2. *Journal of Speech, Language, and Hearing Research, 49*, 500–525.
Van Riper, C. (1973). *The treatment of stuttering.* Englewood Cliffs, NJ: Prentice-Hall.
Watkins, K. E., Smith, S. M., Davis, S., & Howell, P. (2008). Structural and functional abnormalities of the motor system in developmental stuttering. *Brain, 131*, 50–59.
Webster, R. L. (1974). *The Precision Fluency Shaping Program: Speech reconstructions for stutterers.* Roanoke. VA: Communications Development Cooperation.
Wingate, M. E. (1976). *Stuttering. Theory and treatment.* New York, NY: Irvington Publishers.
Wingate, M. E. (1985). Stuttering as a prosodic disorder. In R. F. Curlee, & W. H. Perkins (Eds.). *Nature and treatment of stuttering: New directions* (pp. 215–235). San Diego, CA: College-Hill Press.
Wingate, M. E. (1988). *The structure of stuttering. A psycholinguistic analysis.* New York, NY: Springer.
Wymbs, N. F., Ingham, R. J., Ingham, J. C., Paolini, K. E., & Grafton, S. T. (2013). Individual differences in neural regions functionally related to real and imagined stuttering. *Brain and Language, 124*, 153–164.
Wymer, J. H., Lindman, L. S., & Booksh, R. L. (2002). A neuropsychological perspective of aprosody: features, function, assessment and treatment. *Applied Neuropsychology, 9*, 37–47.

Katrin Neumann, M.D. is professor of phoniatrics and pediatric audiology and otolaryngologist. She is head of the Dept. of Phoniatrics and Pediatric Audiology and the Hearing and Cochlea Implant Center at the ENT clinic of the St. Elisabeth-Hospital, University of Bochum, Germany. She is member of the WHO Expert Advisory Board for the WHO program for prevention of deafness and hearing loss. One of her research focuses is the examination of speech, language, and hearing processes, in particular with neuroimaging techniques. She is associate editor of the Journal of Fluency Disorders.

Harald A. Euler studied psychology at the University of Bonn, Germany, and at Washington State University, USA. From 1974 to 2009 he was Professor of Psychology at the University of Kassel, Germany, where he retired in 2009. His main interest is evolutionary psychology and the treatment of stuttering. He currently serves as a guest scientist at the Department of Phoniatrics and Pediatric Audiology at the University of Bochum, Germany, and as a Professorial Research Fellow for Evolutionary Developmental Psychology at the Department of Developmental Psychology, University of Vienna, Austria.

Malte Kob is professor of theory of music transmission at the Erich-Thienhaus-Institute (ETI) of Detmold University of Music. After his dissertation at the Institute of Technical Acoustics of RWTH Aachen University he worked at the Clinic for Phoniatrics, Paedaudiology and Communication Disorders at University Clinic Aachen. When he became head of ETI he created M. Sc. and Ph.D. programs of Music Acoustics and teaches fundamentals of engineering such as acoustics and electronics, as well as music acoustics and measurement technique. His research interest spans from room acoustics and signal processing to voice physiology and musicians' health. He is member of the executive councils of the German (DEGA) and European Acoustic Associations (EAA).

Alexander Wolff von Gudenberg, M.D., born 1957, studied medicine at Berlin and Hannover University. He specialized as a general practitioner, and in speech and language. In his dissertation he compared stuttering therapies in Germany and the US. He is the medical director of the Institute of the Kassel Stuttering Therapy.

Anne-Lise Giraud, PhD. is a professor of Neuroscience. She heads the Auditory Language Laboratory, in the Basic Neuroscience Department, at the Geneva Biotech Campus. She is a specialist in auditory and speech sciences. Her main research interest is in the understanding of the basic neurocomputational operations that govern speech processing in the human brain, and the diruptions of these operations in language development disorders.

Tobias Weissgerber, Dr Eng., is deputy head of the Audiological Acoustics Department at the ENT Clinic of University Hospital Frankfurt. His responsibilities are research, teaching, and clinical audio-technology in the field of hearing impairment. His main research interests are psychoacoustics, room acoustics, and virtual acoustics and its application in audiology.

Christian Kell, M.D. is a neurologist and group leader of a Cognitive Neuroscience group at the Brain Imaging Center of Goethe University Frankfurt. Besides other research interests (see www.brainclocks.com), he investigates neural plasticity in speech and language disorders, including developmental stuttering, Parkinson's disease, and tumor aphasia.