

Available online at www.sciencedirect.com

SciVerse ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Research report

Cerebellar contribution to the prediction of self-initiated sounds

Franziska Knolle^{a,*}, Erich Schröger^b and Sonja A. Kotz^{a,*}^a Research Group “Subcortical Contributions to Comprehension”, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany^b Institute of Psychology, Leipzig University, Leipzig, Germany

ARTICLE INFO

Article history:

Received 22 July 2012

Reviewed 4 September 2012

Revised 1 November 2012

Accepted 11 December 2012

Action editor Henry Augustus

Buchtel

Published online xxx

Keywords:

Cerebellum

Auditory prediction

Selectivity

Forward-model

Event-related potentials (ERP)

ABSTRACT

In everyday life we frequently make the fundamental distinction between sensory input resulting from our own actions and sensory input that is externally-produced. It has been speculated that making this distinction involves the use of an internal forward-model, which enables the brain to adjust its response to self-produced sensory input. In the auditory domain, this idea has been supported by event-related potential and evoked-magnetic field studies revealing that self-initiated sounds elicit a suppressed N100/M100 brain response compared to externally-produced sounds. Moreover, a recent study reveals that patients with cerebellar lesions do not show a significant N100-suppression effect. This result supports the theory that the cerebellum is essential for generating internal forward predictions. However, all except one study compared self-initiated and externally-produced auditory stimuli in separate conditions. Such a setup prevents an unambiguous interpretation of the N100-suppression effect when distinguishing self- and externally-produced sensory stimuli: the N100-suppression can also be explained by differences in the allocation of attention in different conditions. In the current electroencephalography (EEG)-study we investigated the N100-suppression effect in an altered design comparing (i) self-initiated sounds to externally-produced sounds that occurred intermixed with these self-initiated sounds (i.e., both sound types occurred in the same condition) or (ii) self-initiated sounds to externally-produced sounds that occurred in separate conditions. Results reveal that the cerebellum generates selective predictions in response to self-initiated sounds independent of condition type: cerebellar patients, in contrast to healthy controls, do not display an N100-suppression effect in response to self-initiated sounds when intermixed with externally-produced sounds. Furthermore, the effect is not influenced by the temporal proximity of externally-produced sounds to self-produced sounds. Controls and patients showed a P200-reduction in response to self-initiated sounds. This suggests the existence of an additional and probably more conscious mechanism for identifying self-generated sounds that does not functionally depend on the cerebellum.

© 2012 Elsevier Ltd. All rights reserved.

* Corresponding authors. Research Group “Subcortical Contributions to Comprehension”, Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstraße 1a, 04103 Leipzig, Germany.

E-mail addresses: fknolle@cbs.mpg.de, franzigrimm@hotmail.com (F. Knolle), kotz@cbs.mpg.de (S.A. Kotz).

0010-9452/\$ – see front matter © 2012 Elsevier Ltd. All rights reserved.

<http://dx.doi.org/10.1016/j.cortex.2012.12.012>

1. Introduction

Living in a highly complex world we constantly receive input to our sensory system, some of which results from our own actions (e.g., applauding hands or rubbing one's own arm) and some of which is produced externally. Thus, one fundamental distinction in sensory processing we frequently make is between externally-produced sensory input and input from our own actions. Self-produced sensations are highly predictable: if a sensation matches the prediction of an expected sensation (e.g., a sound resulting from one's knocking on a door) then brain activity directed to the actual input is reduced. On the other hand, externally-produced sensations (e.g., someone else knocks on a door) lead to increased processing activity as external sensations may provide new and important information (Creutzfeldt et al., 1989). To account for the distinction between self-initiated and externally-produced sensations, Wolpert et al. (1998) developed the concept of an internal forward-model for motor planning and motor control. The forward-model postulates that the primary motor cortex sends an efference copy (von Holst and Mittelstädt, 1950) of a motor command to the cerebellum. The cerebellum uses this input to generate a motor-to-somatosensory prediction that mimics the sensory consequences of an action and prepares the musculoskeletal system to successfully execute a movement. A corollary discharge mechanism (Sperry, 1950) is applied to compare the predicted sensations to the actual incoming sensation. If a sensation matches, the same pattern is applied in a subsequent movement. If no match occurs, the cerebellum receives feedback information from respective cortical and subcortical areas to reset its prediction and to apply an online correction of the movement (Kotz and Schwartze, 2010; for a review, see Miall, 1998). Investigations of the internal forward-model have provided ample evidence that the cerebellum generates motor-to-somatosensory predictions (e.g., Blakemore et al., 2001; Imamizu and Kawato, 2008; Tseng et al., 2007; Wolpert et al., 1998).

Based on this notion, we hypothesize that the cerebellum is not only involved in generating motor-to-somatosensory, but also motor-to-auditory predictions. In a previous study, we tested patients with focal cerebellar lesions utilizing an auditory N100-suppression paradigm (EEG: Baess et al., 2008, 2011; Hazemann et al., 1975; McCarthy and Donchin, 1976; Schäfer and Marcus, 1973; magnetoencephalography (MEG): Martikainen et al., 2005). In this paradigm, self-initiated sounds are compared to externally-produced sounds. If a sound is self-initiated, a motor-to-auditory prediction is generated, which leads to a strong N100-suppression effect. In this case, knowledge of agency allows the generation of a motor-to-auditory prediction of the consequences of an action (Frith, 2005). The prediction, in turn, leads to the suppression of an incoming sensation. In contrast to self-initiated sensations, the sensory consequences of external events lead to an accentuated sensation as sensations based on external events cannot be predicted (Blakemore et al., 2000; Wolpert et al., 1995). In fact, our previous results have clearly shown that patients with focal cerebellar lesions do not show a significant N100 amplitude suppression in response to self-initiated sounds (Knolle et al., 2012). These data provide

novel evidence that the cerebellum not only generates somatosensory, but also auditory predictions.

However, studies investigating the N100-suppression effect generally compare self-initiated to external auditory sounds in separate conditions (e.g., Baess et al., 2008; Knolle et al., 2012; Martikainen et al., 2005; Schäfer and Marcus, 1973). In such a design the allocation of attention may differ between conditions: when eliciting self-initiated sounds, attention may be drawn to performing a finger tap, whereas during the presentation of externally-generated sounds attention may be fully directed to perceiving sounds. As the N100 amplitude is sensitive to attentional manipulations (Näätänen, 1990), the N100 reduction in response to self-initiated sounds could merely result from different condition requirements (Baess et al., 2011; Curio et al., 2000; Ford et al., 2007). Thus, the possible condition effects prevent an unambiguous interpretation with respect to a forward-model that distinguishes self-produced from externally-produced sensory events.

Therefore, we extended the classical N100-suppression paradigm by adding externally-produced sounds that occurred randomly after 40% of the self-initiated sounds. Mixing self-initiated and externally-produced sounds in one condition allowed investigating whether a prediction is generated selectively in response to self-initiated sounds or not. Moreover, we aimed to test the influence of the temporal interval of external sounds following the self-initiated sounds: if an additional external sound occurs within a short interval after or before a self-initiated sound, it may erroneously be perceived as a self-initiated sound. Furthermore, we also tested externally-produced sounds in a separate condition to replicate previous results.

In the current experiment additional external sounds occurred randomly in one of four possible intervals following a self-initiated sound: two intervals in the sub-second range (i.e., 400 msec, 800 msec) and two intervals in the supra-second range (i.e., 1200 msec, 1800 msec) as the cerebellum is one of the key structures involved in millisecond timing (e.g., Ivry et al., 2002; Koch et al., 2009; Lee et al., 2007; for review see Buhusi and Meck, 2006) with an upper limit of approximately 1000 msec (Del Olmo et al., 2007; Fierro et al., 2007; Koch et al., 2007). Having external sounds occur within very short intervals either before (i.e., a sound that occurs with a delay of 1800 msec after the last self-initiated sound may occur only 200 msec before the next self-initiated sound) or after (i.e., a sound follows the last self-initiated sound with a delay of 400 msec) a self-initiated sound could lead to external sounds being perceived as self-initiated sounds. In this case, the added external sounds may elicit an N100-suppression effect similar to the one elicited by self-initiated sounds. Thus, the closer an added external sound occurs before or after a self-initiated sound, the more likely it is that this added external sound is processed in the same fashion as a self-initiated sound. Such a result may indicate that the cerebellum does not generate predictions selectively in response to a sound that is self-produced, but rather generates unspecific predictions based on temporal information.

However, if the cerebellum is involved in generating motor-to-auditory predictions, we hypothesize that healthy controls should also maintain an N100-suppression effect in

response to self-initiated sounds when self-initiated and externally-produced sounds are intermixed, whereas we do not expect to find such N100-suppression effect in the cerebellar patients. Consequently, we expect similar results, independent of whether self-initiated and externally-produced sounds occur in the same or in separate condition(s). Based on our previous results (Knolle et al., 2012), we also expect a P200-reduction in response to self-initiated sounds in healthy controls as well as cerebellar patients. We suggest that the P200 effect represents the effect of a mechanism, which more consciously distinguishes self-produced from externally-produced sounds as opposed to the N100-suppression effect, which reflects the effect of an automatic formation of a prediction. More specifically, the P200 effect may reveal the conscious detection of a self-initiated sound. We assume cerebellar patients and healthy controls are able to consciously identify sounds that are self-produced. Therefore, both groups should show comparable P200 responses.

2. Methods

2.1. Participants

Ten patients with focal cerebellar lesions (five females; mean age, 45.82 years; range 25–61 years; all right handed according to the Edinburgh Handedness Inventory; Oldfield, 1971; see Table 1) and ten healthy controls, matched in age, gender, handedness, and educational background, participated in the experiment. All gave informed consent and were paid for their participation. All participants reported normal or corrected-to-normal visual acuity, and normal hearing. Nevertheless, central and peripheral hearing levels were assessed in each participant (i.e., 500–4000 Hz to both ears covering the

standard speech input). In Table 2 for patients and Table 3 for controls results are presented by number of decibel (dB) below the standard hearing curve. A loss of up to 50 dB is considered in the range of normal hearing. Patients and healthy controls were also tested in a number of neuropsychological tests, including the Digit-Span of the Wechsler-Intelligence-Test (Wechsler, 1997), the Listening-Span, which is a standardized and auditory version of the German Reading Span (Daneman and Carpenter, 1980), the Token-Test (De Renzi and Vignolo, 1962), and the Phoneme-Discrimination-Test (in-house used, audio recorded syllables consisting of a consonant and the vowel “a”, e.g., Garrido-Vásquez et al., in press) to assess neurocognitive functions and the National Institutes of Health (NIH) Stroke Scale (Brott et al., 1989; see Table 2 for results).

In the current study we tested patients with focal vascular cerebellar lesions (see Fig. 1a for individual lesion sections and Fig. 1b for lesion overlay; see also Table 1 for lesion size estimated based on affected number of voxels in the magnetic resonance imaging – MRI scan) that resulted from an ischemic or a hemorrhagic stroke (apart from one patient with a cerebellar tumor resection). Patients with posttraumatic lesions were excluded. MRI (3 T Siemens Vario; sequences: T1 3D, axial T2, sagittal Flair, axial T2*) was used to diagnose and locate the lesions. The individual 3 T lesion scans provided in Fig. 1a present three axial sections for each patient (ID P01–P10). To generate the lesion scans the Unified Segmentation Approach was applied. Scans were normalized using an MNI-template (ICBM 152 fitted Brain). Three patients also had extra-cerebellar lesions (see Table 1 for description). The neurological examination was carried out by medical experts at the Day Clinic of Cognitive Neurology at the University Medical Complex, Leipzig, Germany; the procedure involves a neuropsychological, an orthoptic, a psychotherapeutic, and a sociotherapeutic diagnosis at the onset of a rehabilitation

Table 1 – Basic demographics and clinical information.

Patient ID	Gender	Handedness	Age	Lesion site and etiology	Additional lesions	Lesion size (cm ³)
01	m	R	60	R medial PCL; PICA infarction		40.87
02	m	R	31	L ACL and superior PCL, upper pons; SCA infarction following colloid cyst resection		29.88
03	f	R	46	L medial PCL, tonsil, L lateral vermis and R superior PCL; PICA infarction		3.38
04	m	R	59	L medial PCL; PICA infarction	L anterior thalamus	.86
05	f	R	39	R lateral PCL; PICA infarction		.31
06	f	R	33	Bilateral vermis; R deep cerebellar nuclei, superior cerebellar peduncle; tumor postoperatively		3.1
07	m	R	26	PCL bilaterally; ICH and AVM	L superior frontal gyrus and anterior corpus callosum	13.53
08	f	R	50	MCL tonsil; PICA infarction		23.66
09	m	R	46	R deep cerebellar nuclei, superior and middle cerebellar peduncle, pontine tegmentum; SCA infarction	L middle insula, R paramedian thalamus	.65
10	f	R	59	R lateral ACL; SCA infarction		14.27

Abbreviations: m = male, f = female, L = Left, R = Right, ACL = anterior cerebellar lobule, PCL = posterior cerebellar lobule, MCL = medial cerebellar lobule, SCA = superior cerebellar artery, PICA = posterior cerebellar artery, ICH = intracerebral/cerebellar hemorrhage, AVM = arteriovenous malformation.

Table 2 – Neuropsychological and neuropathological test results.

Patient ID	Digit-Span	Listening-Span	Token-Test	Phoneme-Discrimination-Test	NIH-Scale	Ataxia (AFM-Test)	L 500–4000 Hz (db)	R 500–4000 Hz (db)
01	21	4.0	50	21	2	– (0)	48	43
02	17	3.0	50	24	8	L (1)	25	31
03	16	4.0	50	24	0	L (1)	22	25
04	18	5.0	50	24	1	– (0)	18	18
05	16	4.0	50	24	4	– (0)	31	35
06	22	4.5	50	25	3	R (1)	13	13
07	16	4.0	50	22	6	– (0)	23	23
08	17	3.0	50	24	0	– (0)	19	20
09	17	3.0	49	24	9	R (2)	17	14
10	5	2.0	49	22	1	– (0)	14	13

Abbreviations: AFM-Test (test for fine motor skills) = Allensbacher Feinmotoriktest (degree of disturbance: 0 = no disturbance, 1 = mild disturbance, 2 = medial disturbance); L = left, R = right.

program, which permits to assess the patients' cognitive, sensory and motor functioning.

As the current experiment involved a motor task, in which participants have to tap on a tapping device, we collected motor scores reporting the degree of fine motor deficits. The Allensbacher Feinmotoriktest (AFM-Test, Minkwitz, 2000) is a frequently used motor skill test, which measures fine motor skills of patients with a neuropathological diagnosis. The AFM-Test contains 31 tasks drawn from 15 functional and object-related areas (e.g., usage of cutlery, writing with a pen, imitating a rhythm). The performance in each area is measured on a scale from 0 to 4 (i.e., 0 = no performance; 4 = perfect performance, see Table 2 for final scores and information on the affected side). The test measures the level of disturbance in daily life due to fine motor skill impairments. Overall the level of disturbance is also measured on a scale from 0 to 4. In this context 4 indicates that the patient was strongly impaired in the execution of several performances and still requires support in daily situations.

Healthy controls had no history of neurological, psychiatric disorders, or other medical problems. They were matched in age, gender, educational and socioeconomic background. The experiment was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of the University of Leipzig.

2.2. Experimental conditions

The study contained two experimental conditions and one control condition (Fig. 2). In the auditory-motor condition (AMC) participants induced finger taps about every 2.4 sec.¹ Each tap

elicited an immediate presentation of a sinusoidal tone (delay of 2–4 msec due to the uploading of the sound file) via headphones. After 40% of the self-initiated sounds, an additional externally-produced sound was elicited. These added external sounds occurred randomly in one of four intervals in the AMC: 400 msec (AC4), 800 msec (AC8), 1200 msec (AC12), or 1800 msec (AC18). The acoustic stimulation, including self-initiated and externally initiated sounds, was recorded online and used as an “external sound sequence” in an auditory-only condition (AOC). In the AOC, participants did not produce finger taps, but were simply asked to listen and attend to auditory stimuli. Lastly, participants carried out a control condition – the motor-only condition (MOC) –, in which they also performed self-paced finger taps every 2.4 sec. However, in contrast to the AMC, no tone was induced via the finger tap, thus controlling for motor activity. As explained in further detail in the Section 2.4.2, the MOC was subtracted from the AMC to compare the purely sensory activity elicited in the two experimental conditions (AMC – MOC).

The experimental run was preceded by two training blocks. In the first block, participants practiced tapping every 2.4 sec by imitating the interval played back from a metronome. In the second training block, the participants had to tap the sound sequence themselves. They received visual feedback to indicate whether a tapping-trial was too slow (tapping interval longer than 3 sec) or too fast (tapping interval shorter than 1.8 sec).² This training block allowed participants to learn to estimate the time between two successive finger taps without counting. The training block contained 200 trials. The participants had to perform correct taps in 75%. All participants achieved this result in the first run. On average the controls performed correct taps in 85.2% (SD 7.5%) and the patients

¹ The 2.4 sec interval was chosen based on the tapping and timing literature; 600 msec are considered to be a reliable measure for the spontaneous motor tempo (Fraisse, 1982; Drake et al., 2000). The literature provides evidence that the processing of sounds presented in an interval of two or three multiples of 600 msec is facilitated (Essens, 1986; Povel, 1984). Furthermore, sounds occurring within an interval of more than 1.8 sec are no longer perceived as a chain of events, but rather as discrete events (Fraisse, 1982). As we were aiming for a discrete event structure, we chose the first multiple of 600 msec following 1.8 sec which is 2.4 sec.

² We choose to set a minimum and a maximum length of a tapping interval, for two reasons: first, the correct interval was motivated by the timing literature which states that 600 msec is considered a reliable measure of spontaneous motor tempo (Fraisse, 1982; Drake et al., 2000). We therefore set the boundaries of the correct tapping rate at ± 600 msec. Second, we wanted to ensure that the two experimental conditions were as similar as possible. Although the sounds in the AOC also had a regular structure as they appear approximately every 2.4 sec, they were not predictable in terms of precise timing. Hence, the possibility to generate a precise prediction is the only difference when comparing the AOC and the ACC.

Table 3 – Neuropsychological and neuropathological test results.

Control ID	Age	Handedness	Gender	Digit-Span	Listening-Span	Token-Test	Phoneme-Discrimination-Test	NIH-Scale	Ataxia (AFM-Test)	L 500–4000 Hz (db)	R 500–4000 Hz (db)
01	61	R	m	19	4.5	50	24	0	– (0)	29	27
02	31	R	m	17	3.5	50	22	0	– (0)	14	23
03	46	R	f	16	4.5	50	22	0	– (0)	29	26
04	60	R	m	18	4.0	50	24	0	– (0)	21	19
05	39	R	f	21	4.5	50	22	0	– (0)	20	17
06	33	R	f	20	4.5	50	25	0	– (0)	17	16
07	25	R	m	22	5.0	50	22	0	– (0)	18	16
08	50	R	f	18	3.5	50	24	0	– (0)	23	27
09	47	R	m	25	5.5	50	23	0	– (0)	17	16
10	60	R	f	16	3.5	50	24	0	– (0)	19	18

Abbreviations: AFM-Test (test for fine motor skills) = Allensbacher Feinmotoriktest (degree of disturbance: 0 = no disturbance, 1 = mild disturbance, 2 = medial disturbance); L = left, R = right.

produced correct taps in 83.2% (SD 6.4%). Thus, the training block never had to be repeated. During the experimental run, no feedback was given.

2.3. Experimental procedure and recordings

Participants were comfortably seated in an electrically shielded and sound-attenuated experimental chamber. A fixation cross was displayed in the middle of a computer screen. To ensure that the motor activation pattern was similar across participants during AMC and MOC, subjects were instructed to change the index finger after 150 trials. The point of change and the performing index finger were indicated on the screen. Thus, the participants did not have to count the trials or memorize which finger they had to use. The alteration of index fingers was counterbalanced across participants. As cerebellar patients suffered from different intensities and sidedness of motor deficits, they tapped in equal parts with their affected and unaffected hand. Each tap triggered the instantaneous presentation of a tone (680 Hz, 50 msec in duration, 70 dB sound pressure level (SPL)) via headphones (Sennheiser HD 202) to both ears in the AMC and AOC but not in the MOC. An in-house-built, highly sensitive tapping device was used to record the finger taps. No measurable sound was emitted by the taps. In each of AMC and AOC 300 trials were recorded. Altogether 120 added external sounds were recorded, 30 trials per interval type (i.e., 400 msec, 800 msec, 1200 msec, 1800 msec). MOC consisted of 150 trials. All conditions together resulted in a total of 750 trials. Experimental conditions were presented in blocks of 100 trials each. Block order was restricted: the AMC always preceded AOC, but MOC was randomized across participants.

The EEG was recorded continuously from 29 Ag–AgCl electrodes according to the International 10–20 system (FP1, FP2, F7, F3, FZ, F4, F8, FC3, FC4, FT7, FT8, C3, T7, CZ, C4, T8, CP5, CP6, P7, P3, PZ, P4, P8, O1, O2). In addition, we recorded from the left and right mastoids, and the sternum as the ground electrode. The EEG was sampled at a rate of 500 Hz (Refa amplifiers system, TMS International, Enschede, NL) and an anti-aliasing filter of 135 Hz was applied. To control for eye movements, vertical and horizontal electro-oculograms (EOG) were recorded bipolarly. The impedance of all electrodes was

kept below 5 k Ω . The recordings were online referenced to the left mastoid.

2.4. Data analysis

2.4.1. Behavioral data

We acquired tapping intervals for AMC and MOC using the Presentation software (Neurobehavioral Systems, Inc., Albany, CA, USA). For each participant we generated the overall performance accuracy (percent correct; ACCURACY) separately for AMC and MOC. Additionally, we computed separate accuracy values for the left and right index finger; as well as mean values (MEAN) for correct (1.8 sec–3.0 sec) tapping intervals separately for AMC and MOC and the performing index finger (i.e., intervals tapped with the left index finger vs the right index finger). Furthermore, we acquired taps that were correct and too fast or too slow (CORRECTNESS) after each of the four different external sounds which occurred in four different intervals (i.e., 400 msec, 800 msec, 1200 msec, 1800 msec).

2.4.2. EEG data

The EEG data were filtered with a .3–15 Hz bandpass filter (1601 Hamming windowed filter). Epochs were rejected when they exceeded 100 μ V. The EEG data were re-referenced to linked mastoids. ERPs were time-locked to the stimulus onset of the critical trials. Each analyzed epoch lasted 600 msec, including a 100 msec pre-stimulus baseline. Tapping intervals shorter than 1.8 sec or longer than 3.0 sec were treated as errors, and excluded from further EEG analysis. We controlled for motor activity by computing a difference-wave (auditory-motor-corrected condition – ACC) between AMC and MOC. In ACC we investigate self-initiated sounds (AC1) and additional external sounds (AC4, AC8, AC12, AC18). In AOC we examined formerly self-initiated sounds, now externally presented (AO1). Furthermore, we combined all added external sounds in AMC as this allowed to not only comparing self-initiated and added externally-generated sounds in the same sound stream (AC1 and ACex), but also to compare externally-produced sounds and self-initiated once in different sound streams (AC1 and AO1).

The group-average ERPs were generated for each sound type. Difference waves revealed two effects: one negative

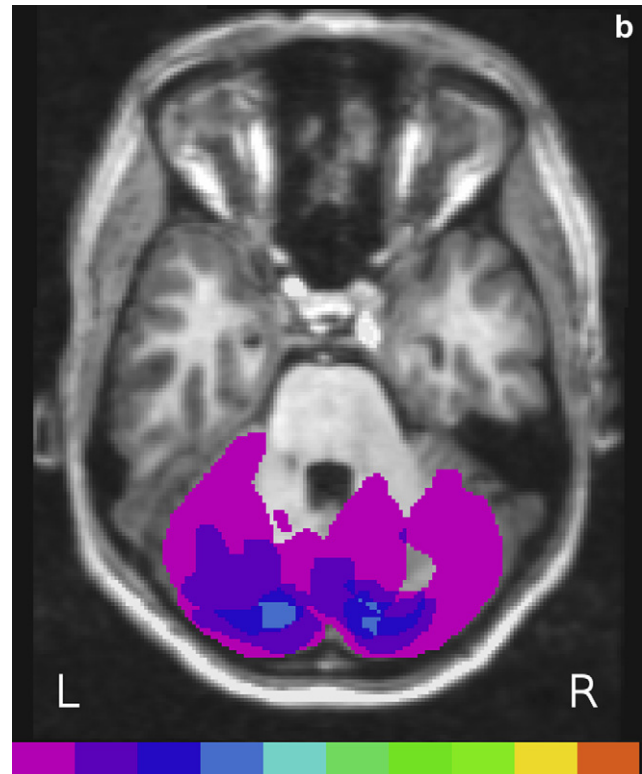
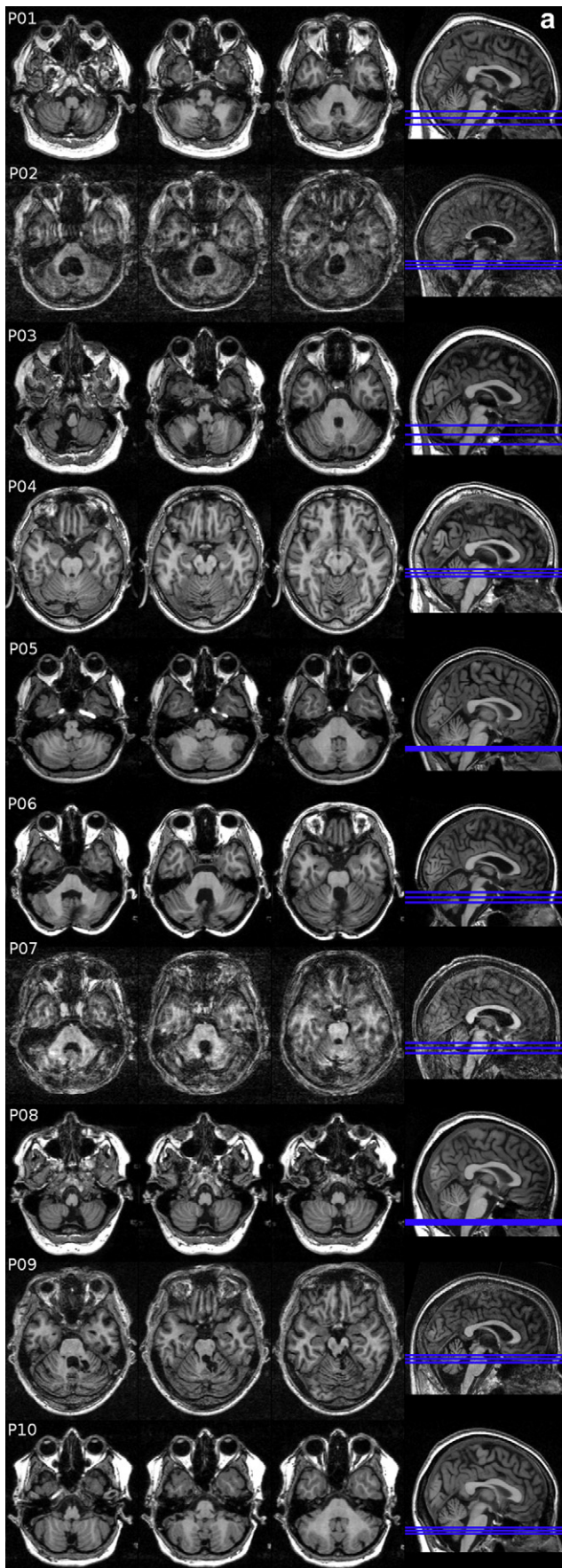


Fig. 1 – (continued).

component in the N100 time window and one positive component, peaking at approximately 190 msec. Statistical analyses were conducted, based on individual amplitudes, in the time windows of 80–120 msec for the N100 and 175–215 msec for the P200. Furthermore we applied a regions of interest (ROIs) analysis using five ROIs [central (ZZ): FZ, CZ, PZ; left lateral (LL): F7, T7, P7; left medial (LM): F3, C3, P3; right lateral (RL): F8, T8, P8; right medial (RM): F4, C4, P4]. Although it is an interesting and important issue to investigate which regions in the cerebellum may be engaged in forming predictions, our sample is too small to allow for a meaningful statistical analysis.

2.4.3. Statistical analyses

For the statistical analysis, the SAS 8.20.20 (Statistical Analysis System, SAS Institute Inc., Cary, North Carolina, USA) software package was used. Only significant results are presented and where required, the Greenhouse–Geisser correction was applied.

For the statistical analysis of the behavioral data we applied a two-way repeated measures analysis of variance

Fig. 1 – a. Individual sections of cerebellar lesions. For each patient (ID P01–P10), three axial sections are presented. The Unified Segmentation Approach was applied. Scans were normalized using an MNI-template (ICBM 152 fitted Brain). b. Cerebellar lesion overlay. Each color represents one patient (i.e., overlapping lesions of ten patient in one area). Purple indicates minimal overlay, orange indicates maximal overlay.

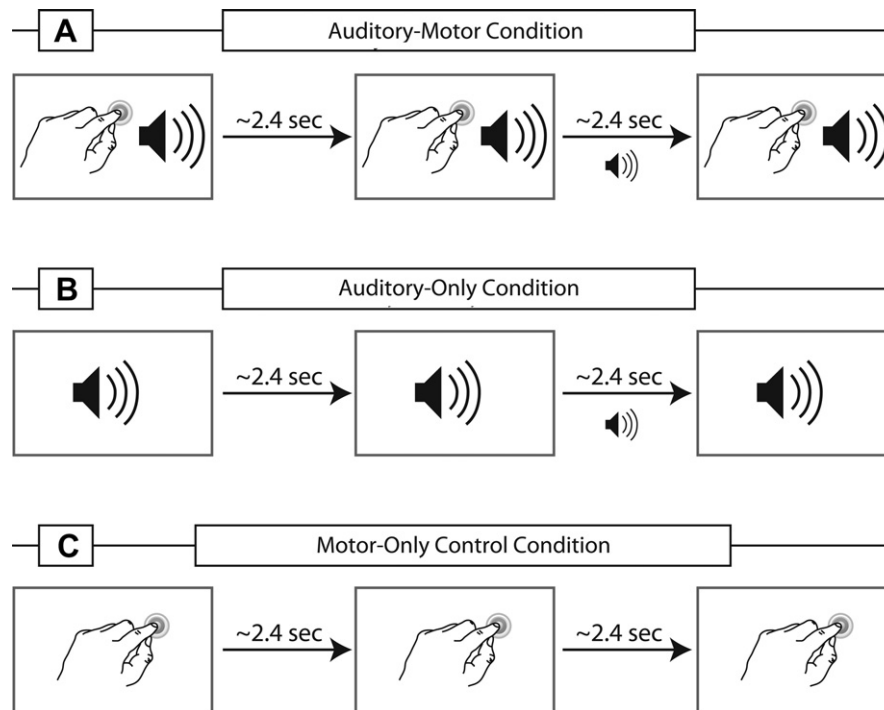


Fig. 2 – Schematic illustration of the three conditions. A) AMC: a sound is self-initiated via a finger tap. In 40% after the self-initiated sounds additional external sounds are introduced (small sound symbol). B) AOC: the tonal sequence is presented externally, containing all sounds from the AMC. C) MOC: taps are required, but no sound is elicited.

(ANOVA). We tested for possible differences in tapping interval means and accuracy rates generated with the right or the left index finger with a between-subjects factor GROUP (patients, controls). We also applied an ANOVA testing the accuracy differences with CORRECTNESS (correct, incorrect) as the within-subject factor and GROUP (patients, controls) as the between-subject factor.

For the statistical analyses of the ERP components we ran an overall ANOVA using the between-subject factor GROUP, to distinguish between patients and controls, and all possible within-subject factors, which were CONDITION (containing self and external sounds: AC1, AO1; as well as all added external sounds: AC4, AC8, AC12, AC18) and ROI (5).

Furthermore, we compared self- and externally-produced sounds that occurred in separate conditions. In this separate setting, we ran an ANOVA including the within-subjects factors GENERATOR (AC1, AO1) and ROI, and the between-subjects factor GROUP. Furthermore, we compared self-initiated and externally-generated sounds that are intermixed in the same condition (i.e., mixed condition). In this mixed setting, we ran one ANOVA including the within-subjects factors GENERATOR (AC1, ACex) and ROI and the between-subjects factor GROUP. Furthermore, we ran a second ANOVA (INTERVAL \times ROI \times GROUP) comparing self-initiated sounds (AC1) to the added external sounds separately according to their interval (AC4, AC8, AC12, AC18).

In addition, the patient sample included three patients with additional lesions outside the cerebellum. In order to assure that these additional lesions did not impact the results, we conducted a Jack-knifing measure (Shao and Tu, 1995) to look at the impact of each patient to the overall ERP result.

Excluding patients one at a time did not significantly alter the result.

3. Results

3.1. Behavioral data

In AMC, the control subjects correctly tapped in 83.5% (SD 10.4%) of the trials and in MOC in 81.0% (SD 24.6%) of the trials. The patients produced correct intervals in 62.7% (SD 22.4%) of the trials in AMC, and in 62.6% (SD 26.7%) in MOC. Statistical analysis of participants' tapping performance revealed a significant group difference in AMC, with controls producing fewer errors than patients [GROUP $F(1,19) = 9.757, p < .05$]. No significant group difference was found in MOC.

Looking at the performance as a function of the possible intervals between self-initiated and added external sounds, we found that patients only performed significantly fewer correct taps after each external sound type than controls {GROUP \times INTERVAL $F(1,4) = 3.56, p < .05$ }; after 400 msec: GROUP $F(1,19) = 4.65, p < .05$ }; after 800 msec: GROUP $F(1,19) = 3.89, p < .05$ }; after 1200 msec: GROUP $F(1,19) = 6.73, p < .05$ }; after 1800 msec: GROUP $F(1,19) = 8.19, p < .05$ }. However, controls and patients performed similarly well after trials that were not interrupted by added external sounds.

3.2. ERP-results – N100 component

The OMNIBUS ANOVA the between-factor (GROUP: patients, controls) and all within-factors (CONDITION: AC1, AO1, AC4,

AC8, AC12, AC18) revealed a nicely pronounced N100, which differed across groups and conditions {CONDITION \times GROUP [F(5,1) = 3.34, $p < .05$]. As expected, we only found a significant group difference in response to self-initiated sounds {GROUP [F(1,19) = 4.42, $p < .05$]. All added external sounds were processed similarly across and within the groups (statistical comparison was not significant (n.s.)). In the control group, the results revealed an N100-suppression in response to self-initiated sounds compared to all externally-generated sounds {CONDITION \times ROI [F(5,4) = 2.82, $p < .05$]. The effect was strongest in the middle and central regions [LM: CONDITION [F(5,9) = 2.57, $p < .05$]; ZZ: CONDITION [F(5,9) = 2.87, $p < .05$]; RM: CONDITION [F(5,9) = 2.73, $p < .05$]. The mean N100 amplitude in response to self-initiated sounds (AC1 mean: $-2.26 \mu\text{V}$) was significantly suppressed compared to the externally-produced sounds (AO1 mean: $-3.19 \mu\text{V}$; AC4 mean: $-3.53 \mu\text{V}$; AC8 mean: $-3.51 \mu\text{V}$; AC12 mean: $-3.30 \mu\text{V}$; AC18 mean: $-3.56 \mu\text{V}$). In the patient group no N100-suppression effect was elicited in response to self-initiated sounds (AC1 mean: $-4.17 \mu\text{V}$; AO1 mean: $-3.66 \mu\text{V}$; AC4 mean: $-4.68 \mu\text{V}$; AC8 mean: $-3.21 \mu\text{V}$; AC12 mean: $-3.90 \mu\text{V}$; AC18 mean: $-3.98 \mu\text{V}$).

As all added external sounds that occur in AMC were processed similarly across and within the groups (i.e., no significant main effects or interactions) we combined the individual 'added external sounds' into one added external sound (ACex), which was used in further analyses.

In the separate conditions the statistical analyses of the ERPs in response to the self-initiated and the externally-produced sounds revealed a nicely pronounced N100, which differed across groups and conditions {GENERATOR \times GROUP [F(1,18) = 6.04, $p < .05$]; Fig. 3A}. As expected, we only found a significant group difference in response to self-initiated sounds {GROUP [F(1,19) = 4.42, $p < .05$], the effect is the same as in the overall and mixed presentation, as in all three cases the same responses to self-initiated sounds are compared. In the control group, the results revealed an N100-suppression in response to self-initiated sounds compared to all externally-generated sounds {GENERATOR [F(1,9) = 5.42, $p < .05$]; GENERATOR \times ROI [F(3,4) = 8.35, $p < .01$]. As can be seen in Fig. 4A, the mean N100 amplitude in response to self-initiated sounds (AC1 mean: $-2.26 \mu\text{V}$) was significantly suppressed compared to the externally-produced sounds (AO1 mean: $-3.20 \mu\text{V}$). The suppression effect was most pronounced in the medial and central ROIs [LM: GENERATOR [F(3,9) = 7.12, $p < .05$]; ZZ: GENERATOR [F(3,9) = 9.77, $p < .05$]; RM: GENERATOR [F(3,9) = 8.46, $p < .05$]; Fig. 5A}.

In the patient group, we did not find an N100-suppression effect in response to self-initiated sounds (Fig. 3B). The result is supported by the mean amplitudes of the ERPs in response to self-initiated (AC1 mean: $-4.17 \mu\text{V}$) and externally-produced sounds (AO1 mean: $-3.66 \mu\text{V}$; Fig. 5A).

In the mixed condition, we obtained very similar results. Both types of sounds elicited a nicely pronounced N100. However, the pattern differed across groups and the generator, which distinguishes between self and externally induced sounds {GENERATOR \times GROUP [F(1,1) = 5.05, $p < .05$]; Fig. 3B}. The results confirm that a significant group difference is only significant in response to self-initiated sounds {GROUP [F(1,19) = 4.42, $p < .05$]. In the control group self-initiated

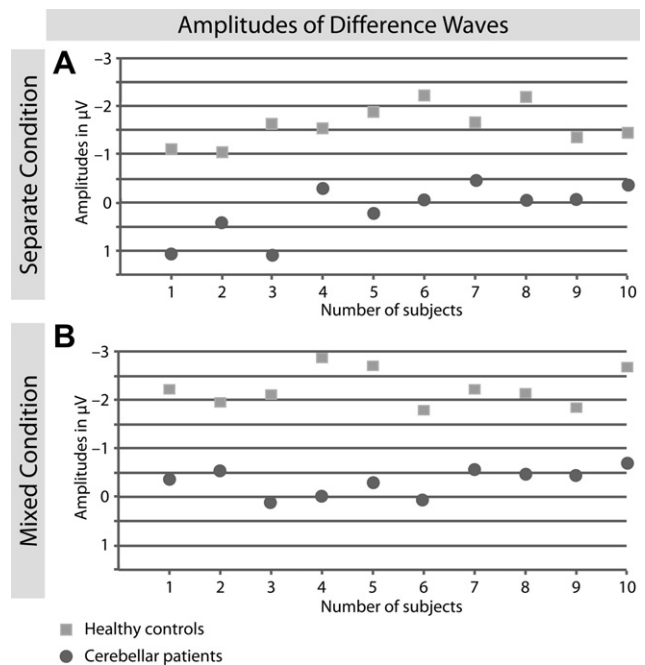


Fig. 3 – Scatter plots of the N100-effect in the central ROI for controls and cerebellar patients. A) Separate conditions: scatter plot shows individual amplitudes of difference waves (AO1–AC1). B) Mixed condition: scatter plot shows individual amplitudes of difference waves (ACex–ACG).

sounds elicit a significantly reduced N100 amplitude compared to the added external sounds in the same sound stream {GENERATOR [F(1,9) = 6.09, $p < .05$]; GENERATOR \times ROI [F(1,3) = 8.54, $p < .01$]; AC1 mean: $-2.26 \mu\text{V}$; ACex: $-3.56 \mu\text{V}$). The suppression effect was most prominent in the medial and central regions [LM: CONDITION [F(1,9) = 7.12, $p < .05$]; ZZ: CONDITION [F(1,9) = 8.46, $p < .05$]; RM: CONDITION [F(1,9) = 9.77, $p < .05$]; Fig. 5B), showing exactly the same pattern when comparing self-initiated to the matching externally-produced sounds as in the mixed condition (Fig. 4B). In contrast to controls, patients did not show a suppression effect to self-initiated sounds compared to added external sounds in the same sound stream (AC1 mean: $-4.17 \mu\text{V}$; ACex mean: $-3.98 \mu\text{V}$; Fig. 4B). The analyses revealed identical results across the separate and the mixed setup (Fig. 4A & B).

Looking at the intervals separately (Fig. 4D), we found results similar to these when looking at the combined ERP in response to added external sounds {INTERVAL \times GROUP [F(4,1) = 3.40, $p < .05$]. The results confirmed a significant group difference only found in response to self-initiated sounds {GROUP [F(1,19) = 4.42, $p < .05$], all added external sounds were processed similarly in the two groups. In the control group, self-initiated sounds elicited a significantly reduced N100 amplitude compared to the added external sounds in the same sound stream {INTERVAL \times ROI [F(4,4) = 2.97, $p < .05$]; AC1 mean: $-2.26 \mu\text{V}$; AC4 mean: $-3.53 \mu\text{V}$; AC8 mean: $-3.51 \mu\text{V}$; AC12 mean: $-3.30 \mu\text{V}$; AC18 mean: $-3.56 \mu\text{V}$ }. We also replicated previous results in cerebellar patients. They did not show an N100-suppression effect to self-initiated sounds compared to single added external sounds in the same sound stream (AC1

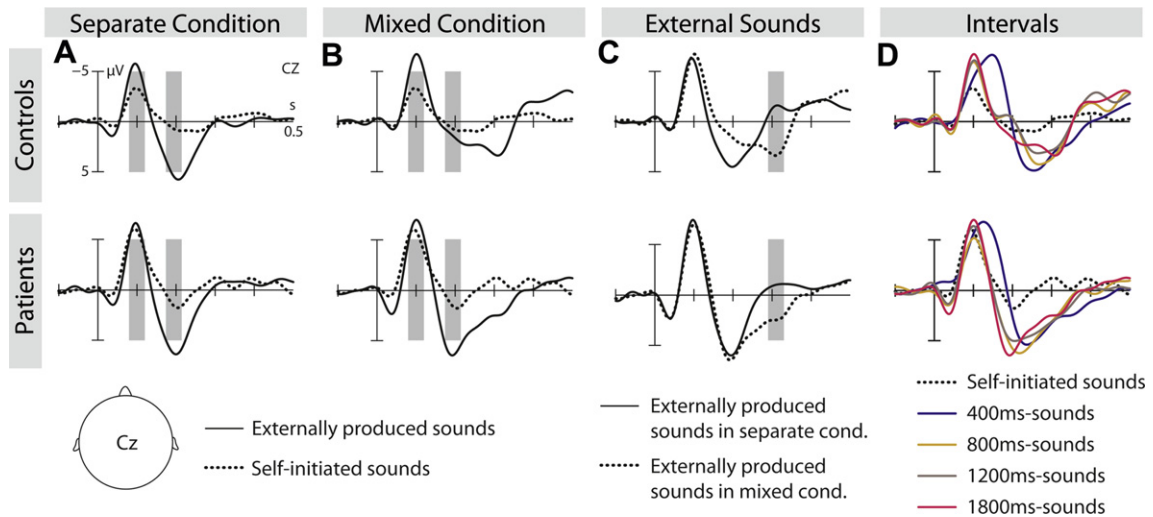


Fig. 4 – ERP responses shown for controls and cerebellar patients. **A) Separate conditions:** ERPs elicited by externally-produced sounds (AO1; black solid line) in AOC and by self-initiated sounds (AC1; black dotted line) in ACC. **B) Mixed conditions:** ERPs elicited by added external sounds (ACex; black solid line) and by self-initiated sounds (AC1; black dotted line) in ACC. **C) External sounds:** ERPs elicited by externally-produced sounds (AO1; black solid line) in AOC and by added external sounds (ACex; black solid line) in AMC. **D) Intervals:** ERPs elicited by added external sounds presented according to their temporal occurrence (AC4, AC8, AC12, AC18) and the self-initiated sounds (AC1; black dotted line), all sounds are elicited in AMC.

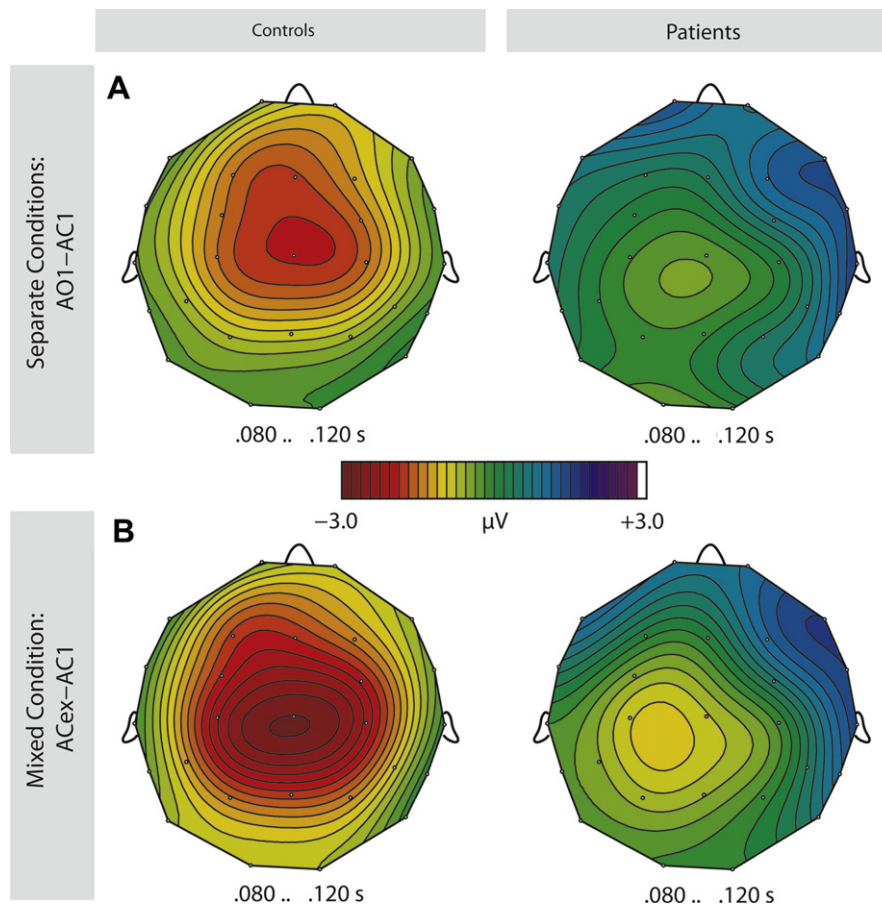


Fig. 5 – Brain maps for both groups. **A) Separate conditions:** grand average scalp maps show the spatial distribution of the difference waves (AO1–AC1) in the analyzed N100 time window. **B) Mixed condition:** grand average scalp maps show the spatial distribution of the difference waves (ACex–AC1) in the analyzed N100 time window.

mean: $-4.17 \mu\text{V}$; AC4 mean: $-4.68 \mu\text{V}$; AC8 mean: $-3.21 \mu\text{V}$; AC12 mean: $-3.90 \mu\text{V}$; AC18 mean: $-3.98 \mu\text{V}$).

3.3. ERP-results – P200 and P3a

In the separate condition (Fig. 4A), the statistical analyses revealed no group difference but a highly significant difference between self-initiated and external sounds {GENERATOR [F(1,19) = 49.43, $p < .0001$]}. Both groups show a reduced P200 in response to self-initiated sounds (controls: ACC mean: $-.04 \mu\text{V}$, AOC mean: $2.92 \mu\text{V}$; patients: ACC mean: $.17 \mu\text{V}$, AOC mean: $3.24 \mu\text{V}$). In the mixed condition (Fig. 4B), however, we found a strong effect of self-initiated versus added external sounds {GENERATOR [F(1,19) = 15.06, $p < .01$]}. The groups also behaved slightly differently with respect to the two different sound types {GENERATOR \times GROUP [F(1,1) = 3.79, $p = .068$]}. Whereas there was no group difference for the self-initiated sounds, we did find a group difference for the added external sounds {GROUP [F(1,19) = 3.10, $p < .05$]}. The mean amplitudes (controls: ACC mean: $-.04 \mu\text{V}$, ACex mean: $1.09 \mu\text{V}$; patients: ACC mean: $.17 \mu\text{V}$, ACex mean: $3.54 \mu\text{V}$) revealed that controls showed a suppressed P200 in response to the added external sounds, but patients did not.

Interestingly, in the control group we found a pronounced P3a in response to externally-produced sounds occurring in AMC {CONDITION [F(1,9) = 11.27, $p < .01$]; ACex mean: $3.13 \mu\text{V}$; AO1 mean: $-.29 \mu\text{V}$ }. Cerebellar patients also displayed a P3a-like component {CONDITION [F(1,9) = 7.05, $p < .05$] in response to externally-generated sounds in AMC. However, the morphology of the ERP component varied enormously. In the controls, the P3a developed as a discrete component following an N100 and a reduced P200, while the P3a-like component in the patients unfolded after a fully pronounced P200 (Fig. 4C).

4. Discussion

In the current study we examined the role of the cerebellum in generating motor-to-auditory predictions by presenting self-initiated and externally-produced sounds either in a mixed or two separate condition(s). In the separate conditions we replicated the results from a standard N100-suppression paradigm used in a previous study (Knolle et al., 2012). Results revealed that the cerebellum generates not only somatosensory, but also auditory predictions, as evidenced by the absence of an N100-suppression effect in response to self-initiated sounds in patients with focal cerebellar lesions. However, when external and self-initiated sounds are elicited in separate conditions, as is the case in most N100-suppression studies, the suppression effect could be explained by different condition demands (Baess et al., 2011; Curio et al., 2000; Ford et al., 2007), rather than as a forward-model, which distinguishes self-produced from externally-produced sounds. Therefore, in the current study we also utilized a mixed condition to investigate whether the prediction is generated selectively in response to self-initiated sounds. In the mixed condition we introduced additional external sounds that randomly followed the self-initiated sounds (in 40% of the cases) in one of four possible intervals (i.e., sub-second:

400 msec and 800 msec; supra-second: 1200 msec and 1800 msec). If these added external sounds occur in close proximity to self-initiated sounds, they could be mistaken as self-initiated sounds and could therefore elicit a comparable N100-suppression. The current results confirm that independent of the condition type (separate or mixed), the cerebellum generates precise motor-to-auditory predictions only for self-initiated sounds, as shown by the lack of a suppression effect in the patient group. In the two separate, as well as the mixed condition(s) only the control group showed a significant N100-suppression and P200-reduction in response to self-initiated sounds, whereas the cerebellar patients only showed a P200-reduction. In both conditions, controls showed a strong N100-suppression effect, as seen in scatter plots (Fig. 3). While it seems that the suppression effect is even stronger in the mixed condition than the separate condition, this difference does not reach significance. Furthermore, only in the control group all external sounds were clearly distinguished from the self-initiated sounds, which revealed a suppression effect in the N100 and P200 response. We, therefore, suggest that the N100-suppression and the P200-reduction in response to self-initiated sounds may reflect two ways of attributing a sensory event to one's own action: one operating in a prospective way and the other in a retrospective way. The N100-suppression may reflect an automatic prediction, generated to prepare the auditory cortex to receive an expected sensory input (Creutzfeldt et al., 1989). The reduced P200, on the other hand, may respond to a conscious post-hoc reflection that a sound closely following a finger tap must have been self-produced (i.e., a conscious detection of a self-initiated sound).

In contrast to controls, cerebellar patients do not show an N100-suppression effect to self-initiated sounds in either of the separate or the mixed condition(s). In fact, the ERP amplitude in response to self-initiated and externally-produced sounds is nearly identical. These results indicate that cerebellar patients are not able to generate a prediction leading to an N100-suppression. Hence, the current finding supports the notion that the cerebellum is critically involved in generating motor-to-auditory predictions. The scatter plot (Fig. 3) shows that in all patients the suppression effect is greatly reduced or not present at all, whereas all matching controls show a significant suppression effect.

As the patients still show a P200-reduction in response to self-initiated sounds, we believe that the self-initiated sounds are detected as such, due to later, more conscious processes. This notion is supported by the patients' reports after the experiment, stating that they always recognized the sound immediately following the tap as being self-produced. We hypothesize that the intact P200-reduction in the patients reflects the conscious detection of a self-initiated sound or, similarly, that a sound following a tap must have been produced by oneself. Unfortunately, the literature on the auditory P200 is sparse (for review see Crowley and Colrain, 2004). However, it has been proposed that a P200-reduction reflects the processing of attended auditory stimuli (De Chicchis et al., 2002; Crowley and Colrain, 2004). This relates nicely to our finding that self-initiated sounds are predictable and expected, which is why they are processed under attention compared to unpredictable externally-produced sounds, which are unattended. In other words, self-generated sounds

do not bear new information for the listeners, as they already know what will come next. Thus processing of this sound is attenuated. Externally-generated sounds, however, do contain new information – the listeners do not know about them in advance. Hence, the processing is not suppressed. If the externally-generated sound is unexpected and infrequent, it may elicit a call for attention, which may then be indexed by the P3a. Moreover, [Minati et al. \(2010\)](#) propose that a P200 amplitude modulation may be related to differences in the integrative processing of successive auditory events, dependent on whether or not a sound is presented within a melodic structure. They found that sounds occurring within a melody elicited a reduced P200, compared to sounds occurring in a randomly-generated sound sequence. This interpretation fits our data as self-initiated sounds and sounds occurring in a melody are highly linked to each other due to their context, and both sound types elicit a reduced P200. However, both interpretations are speculative, and open for future investigation.

In the current study we also tested the selectivity of the prediction leading to a suppression effect. By adding externally-produced sounds that occur in near proximity to self-initiated sounds, we tested the possibility that externally initiated sounds may be processed as self-initiated sounds when presented in the same sound stream. This could lead to a similar N100-suppression effect for externally presented sounds. However, we find a significant suppression effect only in response to self-initiated sounds in the control group. This suggests that a selective prediction is generated for self-initiated sounds only. Furthermore, neither controls nor patients show significant differences between the four different external sound types, indicating that the temporal intervals of external sounds did not influence their processing. This may be due to the setup of the experiment: on the one hand, all physical parameters of the self-initiated and all external sounds were identical, while on the other hand, a sound stream was created that established a relative temporal relation between the self-initiated and the added external sounds, occurring in one of four intervals. A theoretical explanation for this phenomenon is provided by [Buhusi and Meck \(2006\)](#), who suggested that the cerebellum is involved in automatic temporal processing of discrete events, but not in the processing of continuous events in sound streams. Therefore, we believe that mere sound processing in cerebellar patients is not affected by the varying temporal presentation of added external sounds.

When comparing the added external sounds occurring in AMC with externally-produced sounds in AOC, we report an unexpected result. The control group shows a much reduced P200 and a clear P3a effect in response to externally-generated sounds in AMC, but only a P200 in AOC. We suggest that more attention is drawn to added external sounds as they are unexpected and infrequent. This may lead to a reduced P200 response ([Minati et al., 2010](#)). The attentional effect is furthermore reflected in the enhanced P3a. We consider that the P3a response to infrequent and unexpected added external sounds may reflect attentional reorienting, that is, next to the self-initiated sounds the added external sounds take the function of distracters or deviants ([Linden, 2005](#); [Polich, 2007](#); [Snyder and Hillyard, 1976](#)). In contrast to the

controls, the patients show a fully pronounced P200 response (see [Fig. 4C](#)), which is elicited in AOC and in AMC. Furthermore, we also find a P3a-like effect in response to added external sounds in AMC. We therefore speculate that the cerebellar patients may only detect the added external sounds, as being infrequent and unexpected, after having processed the incoming sound completely (i.e., fully pronounced P200 in both conditions). As cerebellar patients fail to generate a prediction based on self-initiated sounds, a differentiation of self-initiated and externally-produced sounds fails in the automatic time range (i.e., no N100-suppression effect), but they may rely on conscious processes to detect infrequent and unexpected sounds (i.e., a P3a-like component after a P200). The current finding therefore suggests that a deficit in generating auditory predictions may also influence attention-dependent processing stages such as the detection of deviants via the reorienting of attention.

Contrary to our previous results ([Knolle et al., 2012](#)), behavioral results reveal significant group differences in AMC, in which cerebellar patients produced fewer correct taps. However, this finding is not surprising, because the added external sounds may have disrupted the tapping flow and enhanced attention to task requirements. As lesion patients are affected in their motor performance, it may be harder for patients than for controls to keep up the tapping rhythm when distracters (i.e., added external sounds) are introduced, leading to an increase in tapping errors. This is supported by the finding that patients only produce more incorrect tapping trials after an added external sound, but not after trials that are not interrupted by an added external sound. Additionally, no differences were found between patients and controls in the performance during MOC. Though, this is a rather surprising finding, it may be explained with the standard deviation measures, which are twice as high in the controls when comparing MOC and AMC. This increase in standard deviation may be due to the fact that MOC is the most difficult of the three conditions, as the participants do not receive any auditory feedback on their performance; therefore they have to fully rely on the internal timing which may be equally hard for controls and patients.

In conclusion, the current study set out to investigate whether the cerebellum is involved in generating predictions selectively for self-initiated sounds when intermixed with externally-produced sounds. The results clearly suggest that the cerebellum generates selective predictions in the auditory domain. Cerebellar patients, in contrast to healthy controls, did not display a significant N100-suppression in response to self-initiated sounds. As we presented self-initiated and externally-produced sounds in a mixed condition, as well as in separate conditions, we have ruled out the possibility of condition effects that could have influenced previous results. More importantly, the mixed design tested whether the prediction is selectively generated in response to self-initiated sounds when occurring together with externally-produced sounds. The current findings suggest that the prediction leading to a suppression effect is generated automatically and selectively, as it is not influenced by added external sounds occurring in close proximity to self-initiated sounds. Therefore, our results extend previous functional interpretations of

the cerebellum as being engaged in forward (somatosensory) predictions (Jordan and Rumelhart, 1992; Wolpert et al., 1998; for reviews see Blakemore and Sirigu, 2003; Ito, 1970) by motor-to-auditory predictions. This also suggests that the cerebellum generates predictions across various domains (Knolle et al., 2012; Ramnani, 2006; Ito, 2008).

Acknowledgments

This work was funded by a DFG-Reinhart-Koselleck grant to ES and DFG KO 2268/6-1 to SAK. FK received financial support from the IMPRS NeuroCom. We thank Christian Obermeier for his support in programming and Ulrike Barth for support in data collection. Also, we would like to thank Anika Stockert and Ingmar Brilmayer for support in conducting the patient imaging data. Lastly, thanks to Kerstin Flake for graphics support and Helga Smallwood for proofreading.

REFERENCES

- Baess P, Jacobsen T, and Schröger E. Suppression of the auditory N1 event-related potential component with unpredictable self-initiated tones: Evidence for internal forward models with dynamic stimulation. *International Journal of Psychophysiology*, 70: 137–143, 2008.
- Baess P, Horvath J, Jacobsen T, and Schröger E. Selective suppression of self-initiated sounds in an auditory stream: An ERP study. *Psychophysiology*, 48: 1276–1283, 2011.
- Blakemore SJ, Wolpert DM, and Frith CD. Why can't you tickle yourself? *NeuroReport*, 11: 11–16, 2000.
- Blakemore SJ, Frith CD, and Wolpert DM. The cerebellum is involved in predicting the sensory consequences of action. *NeuroReport*, 12: 1879–1884, 2001.
- Blakemore SJ and Sirigu A. Action prediction in the cerebellum and in the parietal lobe. *Experimental Brain Research*, 153: 239–245, 2003.
- Buhusi CV and Meck WH. What makes us tick? Functional and neural mechanisms of interval timing. *Nature Review Neuroscience*, 6: 755–765, 2006.
- Brott T, Adams Jr HP, Olinger CP, Marler JR, Barsan WG, Biller J, et al. Measurements of acute cerebral infarction: A clinical examination scale. *Stroke*, 20: 864–870, 1989.
- Creutzfeldt O, Ojemann G, and Lettich E. Neuronal activity in the human lateral temporal lobe. II. Responses to the subjects own voice. *Experimental Brain Research*, 77: 476–489, 1989.
- Crowley KE and Colrain IM. A review of the evidence for P2 being an independent component process: Age, sleep and modality. *Clinical Neurophysiology*, 115: 732–744, 2004.
- Curio G, Neuloh G, Numminen J, Jousmaki V, and Hari R. Speaking modifies voice-evoked activity in the human auditory cortex. *Human Brain Mapping*, 9: 183–191, 2000.
- Daneman M and Carpenter P. Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19: 450–466, 1980.
- De Chicchis AR, Carpenter M, Cranford JL, and Hymel MR. Electrophysiologic correlates of attention versus distraction in young and elderly listeners. *Journal of the American Academy of Audiology*, 13: 383–391. quiz 400–401, 2002.
- Del Olmo MF, Cheeran B, Koch G, and Rothwell JC. Role of the cerebellum in externally paced rhythmic finger movements. *Journal of Neurophysiology*, 98: 145–152, 2007.
- De Renzi E and Vignolo LA. The Token Test: A sensitive test to detect receptive disturbances in aphasics. *Brain*, 85: 665–678, 1962.
- Drake C, Jones MR, and Baruch C. The development of rhythmic attending in auditory sequences: Attunement, referent period, focal attending. *Cognition*, 77: 251–288, 2000.
- Essens PJ. Hierarchical organization of temporal patterns. *Perception and Psychophysics*, 40: 69–73, 1986.
- Fierro B, Palermo A, Puma A, Francolini M, Panetta ML, Daniele O, et al. Role of the cerebellum in time perception: A TMS study in normal subjects. *Journal of the Neurological Science*, 263: 107–112, 2007.
- Ford JM, Gray M, Faustman WO, Roach BJ, and Mathalon DH. Dissecting corollary discharge dysfunction in schizophrenia. *Psychophysiology*, 44: 522–529, 2007.
- Fraisse P. Rhythm tempo. In Deutsch D (Ed), *The Psychology of Music*. New York: Academic Press, 1982: 149–180.
- Frith CD. The self in action: Lessons from delusions of control. *Consciousness and Cognition*, 14: 752–770, 2005.
- Garrido-Vásquez P, Pell MD, Paulmann S, Strecker K, Schwarz J, and Kotz SA. An ERP study of vocal emotion processing in asymmetric Parkinson's disease. *Social Cognitive and Affective Neuroscience*, in press.
- Hazemann P, Audin G, and Lille F. Effect of voluntary self-paced movements upon auditory and somatosensory evoked potentials in man. *Electroencephalography and Clinical Neurophysiology*, 39: 247–254, 1975.
- Imamizu H and Kawato M. Neural correlates of predictive and postdictive switching mechanisms for internal models. *Journal of Neuroscience*, 28: 10751–10765, 2008.
- Ito M. Neurophysiological aspects of the cerebellar motor control system. *International Journal of Neurology*, 7: 162–176, 1970.
- Ito M. Control of mental activities by internal models in the cerebellum. *Nature Review Neuroscience*, 9: 304–313, 2008.
- Ivry RB, Spencer RM, Zelaznik HN, and Diedrichsen J. The cerebellum and event timing. *Annals of the New York Academy of Science*, 978: 302–317, 2002.
- Jordan MI and Rumelhart DE. Forward models: Supervised learning with a distal teacher. *Cognitive Science*, 16: 307–354, 1992.
- Knolle F, Schröger E, Baess P, and Kotz SA. The cerebellum generates motor-to-auditory predictions: ERP lesion evidence. *Journal of Cognitive Neuroscience*, 24: 698–706, 2012.
- Koch G, Oliveri M, Torriero S, Salerno S, Gerfo EL, and Caltagirone C. Repetitive TMS of cerebellum interferes with millisecond time processing. *Experimental Brain Research*, 179: 291–299, 2007.
- Koch G, Oliveri M, and Caltagirone C. Neural networks engage milliseconds and seconds time processing: Evidence from transcranial magnetic stimulation and patients with cortical or subcortical dysfunction. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 264: 1907–1918, 2009.
- Kotz SA and Schwartze M. Cortical speech processing unplugged: A timely subcortico-cortical framework. *Trends in Cognitive Sciences*, 14: 392–399, 2010.
- Lee KH, Egleston PN, Brown WH, Gregory AN, Barker AT, and Woodruff PW. The role of the cerebellum in subsecond time perception: Evidence from repetitive transcranial magnetic stimulation. *Journal of Cognitive Neuroscience*, 19: 147–157, 2007.
- Linden D. The P300: Where in the brain is it produced and what does it tell us? *Neuroscientist*, 6: 563–576, 2005.
- Martikainen MH, Kaneko K, and Hari R. Suppressed responses to self-triggered sounds in the human auditory cortex. *Cerebral Cortex*, 15: 299–302, 2005.
- McCarthy G and Donchin E. The effects of temporal and event uncertainty in determining the waveforms of the auditory

- event related potential (ERP). *Psychophysiology*, 13: 581–590, 1976.
- Miall RC. The cerebellum, predictive control and motor coordination. *Novartis Foundation Symposium*, 218: 272–284. discussion 284–290, 1998.
- Minati L, Salvatoni L, Rosazza C, Pietrocini E, Visani E, Panzica F, et al. Event-related potential (ERP) markers of melodic processing: The N2 component is modulated by structural complexity, not by melodic ‘meaningfulness’. *Brain Research Bulletin*, 83: 23–28, 2010.
- Minkwitz K. Allensbacher Feinmotoriktest (AFM-Test). *Ergotherapie & Rehabilitation*, 1: 19–22, 2000.
- Näätänen R. The role of attention in auditory information-processing as revealed by event-related potentials and other brain measures of cognitive function. *Behavioral and Brain Sciences*, 13: 201–232, 1990.
- Oldfield R. The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9: 97–113, 1971.
- Polich J. Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, 118(10): 2128–2148, 2007.
- Povel DJ. A theoretical framework for rhythm perception. *Psychological Research*, 45(4): 315–337, 1984.
- Ramnani N. The primate cortico-cerebellar system: anatomy and function. *Nature Review Neuroscience*, 7: 511–522, 2006.
- Schäfer EW and Marcus MM. Self-stimulation alters human sensory brain responses. *Science*, 181: 175–177, 1973.
- Shao J and Tu D. *The Jackknife and Bootstrap*. Springer-Verlag, Inc., 1995.
- Snyder E and Hillyard S. Long-latency evoked potentials to irrelevant, deviant stimuli. *Behavioral Biology*, 16: 319–331, 1976.
- Sperry RW. Neural basis of the spontaneous optokinetic response produced by visual inversion. *Journal of Comparative and Physiological Psychology*, 43: 482–489, 1950.
- Tseng YW, Diedrichsen J, Krakauer JW, Shadmehr R, and Bastian AJ. Sensory prediction errors drive cerebellum-dependent adaptation of reaching. *Journal of Neurophysiology*, 98: 54–62, 2007.
- von Holst E and Mittelstädt H. Das Reafferenzprinzip. *Naturwissenschaften*, 37: 464–476, 1950.
- Wechsler DA. *Wechsler Adult Intelligence Scale-III*. New York: Psychological Corporation, 1997.
- Wolpert DM, Ghahramani Z, and Jordan MI. An internal model for sensorimotor integration. *Science*, 269: 1880–1882, 1995.
- Wolpert DM, Miall RC, and Kawato M. Internal models in the cerebellum. *Trends in Cognitive Sciences*, 2: 338–347, 1998.