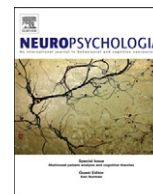




ELSEVIER

Contents lists available at SciVerse ScienceDirect

Neuropsychologia

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

Note

Auditory skills and brain morphology predict individual differences in adaptation to degraded speech

Julia Erb^{a,*}, Molly J. Henry^a, Frank Eisner^b, Jonas Obleser^a

^a Max Planck Research Group "Auditory Cognition", Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstraße 1A, 04103 Leipzig, Germany

^b Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6525 XD Nijmegen, the Netherlands

ARTICLE INFO

Article history:

Received 16 January 2012

Received in revised form

9 May 2012

Accepted 10 May 2012

Available online 18 May 2012

Keywords:

Language

Noise-vocoded speech

Cochlear implant simulation

Perceptual learning

Amplitude modulation rate

Voxel-based morphometry

ABSTRACT

Noise-vocoded speech is a spectrally highly degraded signal, but it preserves the temporal envelope of speech. Listeners vary considerably in their ability to adapt to this degraded speech signal. Here, we hypothesised that individual differences in adaptation to vocoded speech should be predictable by non-speech auditory, cognitive, and neuroanatomical factors. We tested 18 normal-hearing participants in a short-term vocoded speech-learning paradigm (listening to 100 4-band-vocoded sentences). Non-speech auditory skills were assessed using amplitude modulation (AM) rate discrimination, where modulation rates were centred on the speech-relevant rate of 4 Hz. Working memory capacities were evaluated (digit span and nonword repetition), and structural MRI scans were examined for anatomical predictors of vocoded speech learning using voxel-based morphometry. Listeners who learned faster to understand degraded speech also showed smaller thresholds in the AM discrimination task. This ability to adjust to degraded speech is furthermore reflected anatomically in increased grey matter volume in an area of the left thalamus (pulvinar) that is strongly connected to the auditory and prefrontal cortices. Thus, individual non-speech auditory skills and left thalamus grey matter volume can predict how quickly a listener adapts to degraded speech.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

Noise-vocoding has been used to simulate cochlear-implant (CI) transduced speech in normal-hearing listeners (Shannon, Zeng, Kamath, Wygonski, & Ekelid, 1995). Vocoding removes most of the spectral properties of the auditory signal, while its temporal envelope cues remain largely preserved. Normal-hearing listeners learn to understand this type of degraded speech surprisingly quickly on the first exposure (Davis, Johnsruide, Hervais-Adelman, Taylor, & McGettigan, 2005; Eisner, McGettigan, Faulkner, Rosen, & Scott, 2010). However, there is considerable variability across listeners in their ability to perceptually adapt to this degraded auditory input. Although there are larger individual differences in speech perception performance in CI users than in normal-hearing subjects listening to a CI simulation, normal-hearing listeners still differ considerably in learning to understand vocoded speech (Shannon, Galvin, & Baskent, 2002). Currently it is unclear what drives the adaptation to this degraded speech signal. In cochlear-implanted children, demographic factors such as age at implantation, duration of deafness, residual hearing before implantation or

the number of electrodes of the CI can only explain about half of the variance in outcome (Blamey et al., 2001; Sarant, Blamey, Dowell, Clark, & Gibson, 2001).

1.1. Behavioural correlates of perceptual adaptation to degraded speech

Phonological working memory is one cognitive factor that has been linked to the development of speech recognition abilities (Gathercole, Willis, Baddeley, & Emslie, 1994). According to the model proposed by Baddeley and colleagues, working memory relies on two processes: the "buffer", which holds memory traces for a few moments, and the "phonological loop", which can be imagined as a subvocal rehearsal process (Baddeley, 2012). Pisoni and colleagues have investigated the relationship between phonological working memory and individual speech recognition performance in CI users, and found that working memory scores, as measured by digit span, significantly correlated with spoken word recognition scores in paediatric CI users, even after statistical partialling out of demographic factors (Pisoni & Cleary, 2003).

In the present study, we wanted to examine whether we could replicate these results in normal-hearing adults listening to vocoded speech. Digit span requires the participant to hold online representations of perceived auditory information, taxing the

* Corresponding author. Tel.: +49 341 9940 2641; fax: +49 341 9940 2499.
E-mail addresses: jerb@cbs.mpg.de (J. Erb), henry@cbs.mpg.de (M.J. Henry), f.eisner@mpi.nl (F. Eisner), obleser@cbs.mpg.de (J. Obleser).

“buffer” in Baddeley’s working memory model. However, as digit span demands memorizing of highly familiar items that have a long-term semantic representation, it may not be a reliable test of phonological working memory (Jacquemot & Scott, 2006), especially if semantic and phonological representations have separable stores as some patient studies suggest (Romani & Martin, 1999). In order to test more specifically phonological working memory, we therefore additionally employed a nonword repetition test in the current study. Degraded speech recognition challenges phonological analysis and requires the maintenance of auditory input in the working memory for a short period of time. Thus, we expected that both tests of working memory, digit span and nonword repetition, should independently explain some of the variability associated with vocoded speech perception.

From an acoustic perspective, noise-vocoded speech has very poor spectral resolution. This forces listeners to rely more on temporal envelope cues for speech recognition. The utilisation of these cues might be easier for listeners with a higher sensitivity to envelope fluctuations in an auditory signal. We therefore hypothesised that the individual capability to adapt to vocoded speech is related to amplitude modulation (AM) rate discrimination thresholds. AM discrimination performance provides a measure of the auditory system’s ability to encode temporal information in a waveform’s envelope (Wakefield & Viemeister, 1990). The most prominent low-frequency amplitude modulations in the temporal envelope of speech are near the syllabic rate of approximately 4 Hz (Giraud et al., 2000; Houtgast & Steeneken, 1985). As these low AM frequencies are essential for speech perception, we decided to test listeners in their sensitivity to AM rates centred on 4 Hz. In the present study, we deliberately chose to assess sensitivity to AM rate rather than AM depth, which is traditionally used to determine temporal modulation transfer functions (Viemeister, 1979). We assessed thresholds to modulation rate instead, as it captures the temporal aspect of envelope encoding which is emphasised when listening to vocoded speech.

Hence we anticipated that, complementarily to phonological working memory scores, auditory processing skills that are not speech-specific (i.e., AM rate sensitivity) could account for part of the individual differences in perceptual learning of vocoded speech.

1.2. Brain correlates of perceptual adaptation to degraded speech

A small number of studies have investigated the relationship between brain activation patterns and perceptual adaptation to degraded speech. Notably, Adank and Devlin (2010) conducted a functional MRI study of adaptation to time-compressed speech. They observed adaptation-related activity bilaterally in the auditory cortex (posterior superior temporal sulcus) and in the left ventral premotor cortex. The authors concluded that perceptual learning of degraded speech involves mapping novel acoustic patterns onto articulatory motor plans (Adank & Devlin, 2010).

Using a short-term vocoded speech-learning paradigm, Eisner et al. (2010) found that individual performance improvement in vocoded speech perception was correlated with the functional MRI signal change in the inferior frontal gyrus (IFG). Additionally they observed a positive correlation between activity in the angular gyrus (AG) and individual learning curves over the course of 100 noise-vocoded sentences (but did not have continuous measures of performance throughout the scan period).

Golestani and colleagues explored the structural brain correlates of novel speech sound learning. They found that faster phonetic learners had increased white matter density in the parieto-occipital sulcus (Golestani, Paus, & Zatorre, 2002). Though, to our knowledge, structural correlates of vocoded

speech learning have not been investigated so far. Here, we examine whether it is possible to predict perceptual learning of a vocoded speech from the brain structure of a listener.

The present study tested normal-hearing adults in a short-term (ca. 20 min long) vocoded speech-learning paradigm. A set of tests including digit span, nonword repetition and AM rate discrimination were administered. Additionally, we analysed structural MRI scans using voxel-based morphometry (VBM). The aim of the study was to elucidate whether and to what extent these cognitive, non-speech auditory, and neuroanatomic measures could predict how well a listener perceptually adapts to vocoded speech.

2. Materials and methods

2.1. Participants

Eighteen participants (nine females, age range 22–30 years, mean 26.2 years) took part in this study. Participants were recruited from the participant database of the Max Planck Institute for Human Cognitive and Brain Sciences. All were right-handed, monolingual speakers of German with no known hearing or language impairments or neurological problems. They were naïve to noise-vocoded speech and had not participated in a vocoded speech perception experiment before. Participants received financial compensation of € 7, and informed consent was obtained from all subjects. Procedures were in accordance with the guidelines of the local ethics committee (University of Leipzig).

2.2. Design and materials

2.2.1. Vocoded speech perception

Stimuli. Sentence material was drawn from a German version of the English speech in noise (SPIN) sentences (Kalikow, Stevens, & Elliott, 1977). We developed a German version of the SPIN sentences based on the same criteria on which the English version were based; as in the English version, the German speech in noise (GSPIN) sentences were controlled for predictability of the final word (high vs. low predictability) which was always a mono- or bisyllabic noun. Homogeneity in sentence length was achieved by restricting all sentences to five to eight words and nine to eleven syllables. According to the “Leipziger Wortschatz” (<http://wortschatz.uni-leipzig.de/>) all final words occurred with relatively high frequency (range 6–16, where frequency f means that the specified word occurs 2^f times more rarely than the reference word “der”) in the German language.

For the present study, only low-predictability sentences were chosen, in order to control and keep low the influence of semantic cues. Thus, the listener had to rely primarily on acoustic properties of the sentence to understand it as opposed to high-predictability items where more semantic information would be available. Typical sentences were “Morgen nehmen wir mal wieder Obst” [Tomorrow we will be taking again some fruit; literal translation] or “Nina hatte keine Ahnung von der Post” [Nina did not know about the mail].

One hundred and thirty low-predictable sentences were recorded by a female speaker of standard German in a sound-damped chamber. The length of the recorded sentences varied between 1623 and 2759 ms. For the short-term adaptation experiment, 100 sentences were used that had proven unambiguously low in predictability in a written cloze test (i.e., sentence completion) with 10 (different) subjects. A list of these 100 low predictable GSPIN sentences and their high-predictable counterparts are provided in the Appendix.

The sentences were noise-vocoded, which involves dividing the raw speech signal into frequency bands, extracting the amplitude envelope from each band and reapplying the amplitude envelope to bandpass-filtered noise carriers with matched cut-off frequencies. The fewer frequency bands used, the poorer the spectral resolution of the resulting speech signal will be, meaning that the speech signal is more distorted and therefore more difficult to recognise. In the present experiment, GSPIN sentences were noise-vocoded using four frequency bands that were spaced according to Greenwood's cochlear frequency-position function (Greenwood, 1990). Noise-vocoding was carried out in MATLAB 7.11 as described in Rosen, Faulkner, and Wilkinson (1999). The envelopes were extracted using a zero-phase, 4th-order Butterworth lowpass filter; the lowpass filter cutoff was set at 256 Hz. Speech signals were half-wave rectified. The filters for band extraction are specified in Table 1; they were extracted using a 6-th order Butterworth bandpass filter. Scaling for equal root mean square (RMS) energy was performed channel-wise for each channel envelope.

Task. One hundred low-predictable 4-band noise-vocoded GSPIN sentences were presented by a personal computer through Sennheiser HD201 headphones using Presentation software. On each trial, participants heard a sentence which was followed by visual presentation of a green light ("go" signal for response) on the screen lasting for 3 s. During this time, participants were to respond by repeating the sentence, which was recorded for later off-line scoring. As the experiment was designed to be compatible with a future fMRI sparse sampling experiment, participants heard a recording of scanner noise at the end of each trial. For the same reason, we interspersed the experiment with twenty off-trials lacking any auditory presentation. In sum, the experiment comprised 120 trials lasting 9 s each.

Vocoded speech perception performance was evaluated from recordings of participants' responses. Responses were scored as proportions of correctly repeated words per sentence. Scoring comprised all words of a sentence including function words. The marking scheme was liberal such that errors in declension or conjugation were accepted as correct.

2.2.2. Measure of individual adaptation to vocoded speech

To establish that participants learned to understand vocoded speech, i.e., that their performance improved over the course of the experiment, the experiment was divided post-hoc into two blocks of 50 trials each, which were compared using a paired-sample *t*-test.

In order to find a measure for individual perceptual learning of vocoded speech over time, we modelled each individual's performance improvement in two different ways: as a power law or as a linear performance increase. It has frequently been claimed that learning follows a power law (Anderson & Schooler, 1991), and power law functions have been used previously to describe short-term perceptual adaptation to time-compressed speech (Peelle, Olafsen, Davis, & Wingfield, 2010). However, experiments examining perceptual adaptation to vocoded speech have commonly assumed a linear increase in performance over time (Davis et al., 2005).

Table 1
Cut-off and centre frequencies (in Hz) of the four-band noise-vocoded stimuli.

Band	1	2	3	4
Lower cut-off	70	423	1304	3504
Centre frequency	207	764	2156	5634
Upper cut-off	423	1304	3504	9000

Here, to test whether a linear or a power law function would better describe the data, both functions were fitted to the individual performance scores over time using a least-squares estimation procedure in MATLAB 7.11 (for an example linear fit to the scores averaged over participants, see Fig. 1B). The linear trend was described by the function $y = a*t + b$ where y is performance (proportion words correct), t is time (trial), a is the slope and b the intercept of the fitted function. The power law curve was defined by a function of the shape $y = a*t^m + b$, where y is performance, t is time (trial), m is the learning rate parameter, a the slope and b the intercept of the power law curve. Thus, the power law function has one more parameter than the linear function.

We compared goodness of fit by determining the Bayesian information criterion (BIC; Schwarz & Gideon, 1978) of the linear and the power law fits within each subject. Note that BIC takes into account also the number of fitted parameters, thereby allowing for a fair comparison between the linear and the power law fit.

2.2.3. Working memory tests

All participants underwent two short-term memory tests: digit span (forward and backward) and nonword repetition. Digit span is supposed to measure working memory capacity whereas nonword repetition is thought to test more specifically phonological working memory capacity (Baddeley, 2012).

2.2.3.1. Auditory forward and backward digit span. Lists of single digits between one and nine were read out and recorded by a female speaker at a rate of approximately one digit per second. Participants listened to these recordings and had to repeat immediately the list of digits in the same order (forward digit span) or in the reverse order (backward digit span), i.e., starting with the last digit and finishing with the first digit they heard. The test had seven levels featuring list lengths starting from three digits increasing by one digit up to nine digits for forward digit span, and lists lengths from two to eight digits for backward digit span. Each level comprised two items. The participants' response was marked as correct only if all digits were correctly repeated in the required order. The testing stopped when the participant could not report correctly any item of one level or when all 14 items had been presented. The individual forward and backward digit span scores were calculated as the sum of items the participant had repeated correctly.

2.2.3.2. Auditory nonword repetition test. Stimuli. An extended version of Mottier's test (Mottier, 1951) was used to test phonological working memory. This test comprises 90 nonwords; each nonword consists of two up to 15 syllables with a consonant-vowel structure, e.g., "kapeto". None of these nonwords belonged to the German vocabulary. The test had 14 levels, one for each nonword length (counted as the number of syllables) and each level featured six items. The first level contained nonwords with a length of two syllables; word length increased by one syllable for each subsequent level. Nonwords were recorded by a female speaker.

Task. Participants listened to recordings over headphones and had to repeat back immediately what they had heard. Responses were judged online by an experimenter but also recorded for later verification. Only exact repetitions of the nonwords were scored correct. Participants had to report correctly at least four nonwords per level in order to proceed to the next level, otherwise the test was terminated. Nonword repetition test scores were calculated as the total number of correctly repeated nonwords.

2.2.4. Amplitude modulation (AM) rate discrimination

Stimuli. Participants were also tested in an AM rate discrimination task. Stimuli were one-second long sinusoidally amplitude-modulated white noises. The standard stimulus was modulated at 4 Hz. Deviant stimuli were modulated at eight different rates that were linearly spaced between 2 and 6 Hz in steps of 0.57 Hz (hence deviant stimuli were centred around 4 Hz) with a constant modulation depth of $m=60\%$. The onset phase of the sinusoidal modulation was randomly varied for all stimuli. Stimuli were presented with an inter-stimulus interval of 500 ms.

Task. Participants listened to stimuli on a PC in MATLAB 7.11 via Sennheiser HD-201 headphones at approximately 60 dB SPL. The paradigm was a 3-alternative forced choice (3-AFC) task: In each trial, participants heard two standard stimuli and one deviant sound; the position of the deviant within the trial was randomly varied from trial to trial. Participants' task was to indicate at which position they heard the sound with the different modulation rate (faster or slower). Participants responded via keyboard by pressing the keys "1", "2" or "3" and immediately received feedback after each trial with "right" or "wrong" appearing on the screen. On the whole, every participant listened to 160 trials, such that each deviant modulation rate level was presented 20 times in a random order.

2.2.5. Calculation of just noticeable difference (JND) thresholds in AM discrimination

As a measure of individual sensitivity to AM rate, we determined just noticeable difference (JND) thresholds. Proportions of correct responses were calculated for each participant for each of the eight comparison levels and averaged over the twenty trials. JND thresholds for the resulting psychometric curves were then estimated using the z-transform method described by Macmillan and Creelman (2005). Based on this analysis, JND is a measure of half the distance between the 25th and 75th percentiles (i.e., the semi-interquartile range; IQR) of the cumulative probability density function. It provides a measure of sensitivity to changes in AM rate, where smaller JND thresholds correspond to a higher sensitivity. JND thresholds are reported as absolute deviation from the standard (4 Hz) in Hz. Due to experimenter error, JNDs for two participants were calculated based on only 60 trials (7 or 8 trials per deviant AM rate).

As for the vocoded speech learning task, in order to test for performance increases over time, data were divided post-hoc into two blocks of 80 trials each and compared using a paired-sample *t*-test. The two previously mentioned participants were excluded from this analysis, because the time scale of their performance was not comparable to the one of the remaining participants (60 vs. 160 trials).

2.3. Procedure

In order to maximise the comparability of performance between individuals, all participants were tested in the same order as follows: Digit span, nonword repetition, vocoded speech perception, AM discrimination.

2.4. MRI data acquisition and analysis

Structural brain scans of all eighteen participants were available through the Institute's brain data base. Scanning was carried out on a 3T Siemens Trio TIM scanner. A T1-weighted MPRAGE sequence was used to acquire 176 sagittal slices, with an acquisition matrix of 256×240 , yielding a resolution of 1 mm^3 .

Repetition time was 1300 ms and echo time was 3.46 ms ($n=15$, n being the number of participants) or 3.97 ms ($n=3$). For acquisition of the images either a birdcage head coil ($n=3$), a 12-channel head coil ($n=7$), or a 32-channel head coil ($n=8$) was used. The scans were between 1 and 62 months old ($M \pm SEM = 20.6 \pm 4$).

We used voxel-based morphometry (VBM) to examine the relationship between brain structure and adaptation to vocoded speech. VBM is an exploratory whole-brain technique running in SPM8 (Wellcome Department of Imaging Neuroscience, London, UK). In accordance with the protocol developed by Ashburner and colleagues (Ashburner & Friston, 2000), we first segmented images into grey matter, white matter and cerebrospinal fluid. Grey and white matter images were imported into DARTEL format. In order to create a group-specific grey and white matter template we used the DARTEL algorithm (Ashburner, 2007), which estimates the deformations that best align the individual images together by iteratively registering the images to their average. These deformation estimations were then used to spatially normalise the grey and white matter images to MNI space. The images were smoothed with an isotropic Gaussian kernel of 10 mm FWHM. The final pre-processed images were modulated, i.e., adjusted for the total volume of grey matter identified in the normalisation procedure. Therefore, the voxels of these images represent a measure of regional volume (Ashburner, 2009).

Modulated images were statistically analysed in a multiple regression analysis in the context of the general linear model (GLM). We set up separate models where either individual adaptation slopes for learning of vocoded speech or AM discrimination thresholds were entered as covariates of interest. As these covariates of interest were highly correlated, we preferred not to enter them into the same model to avoid that shared variance would lead to uninterpretable results. In order to control for the variance in sex, age, coil of acquisition and time since acquisition (in months) which might confound our results, we added to each model four nuisance variables coding for these factors. In both analyses, we corrected for total grey matter volume (GMV) using global normalisation, because we were not interested in the absolute amount of grey matter, but in how regional GMV compares to elsewhere in the brain. That is, we tested the hypothesis that amount of local GMV relative to total GMV was predictive of the behavioural effect of interest (i.e., adaptation slope or AM discrimination threshold). We used this approach to assure not to confound our results with existing age or sex differences in total GMV (Peelle, Cusack, & Henson, 2012).

The results were corrected across the whole brain based on cluster extent. Corrections based on cluster extent can be problematic in VBM studies due to the potential nonstationarity of the data, but it is possible to use a non-stationary cluster extent correction (Hayasaka, Phan, Liberzon, Worsley, & Nichols, 2004). Therefore we corrected the cluster extent threshold for non-isotropic smoothness using the vbm8-toolbox (developed by Christian Gaser; available at <http://dbm.neuro.uni-jena.de/vbm/download/>).

In addition to the whole-brain analysis, we also conducted a region of interest (ROI) analysis based on the results by Eisner et al. (2010). We used the peak coordinate of their correlation between individual learning scores and fMRI signal in the left IFG (MNI-coordinates: $-46 \ 26 \ 20$) to define an ROI (radius: 3 mm; ROI was built using MarsBar toolbox (Brett, Anton, Valabregue, & Poline, 2002) running in SPM8).

Results were displayed in MRICRON on the ch2 template in MNI space (Rorden & Brett, 2000; available at <http://www.mccauslandcenter.sc.edu/mricron/mricron/>).

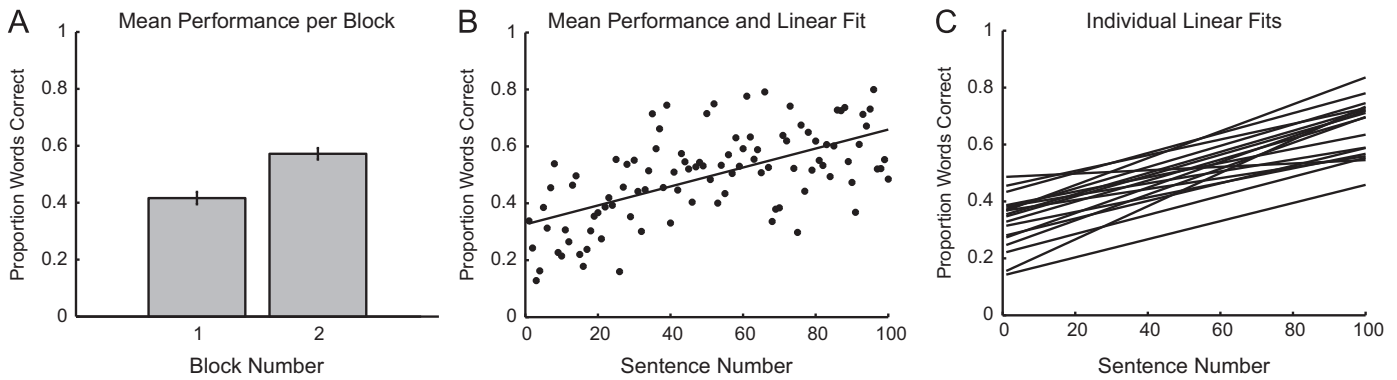


Fig. 1. Perceptual adaptation to vocoded speech. (A) Mean word recognition performance increases from the first to the second half of the experiment (indicating that learning took place). (B) Linear fit to mean word recognition scores for all trials (averaged over participants) and (C) linear fits to individual speech perception scores, showing individual differences in adaptation to vocoded speech.

3. Results

3.1. Adaptation to vocoded speech

First, we established that participants exhibited performance increases in vocoded speech recognition over the course of the experiment. A paired-sample *t*-test showed that participants' mean report scores in the first half of the experiment were significantly lower than the ones in the second half ($t(17) = -7.64$, $p < 0.001$; see Fig. 1A). Another index for participants' overall performance improvement over time is the correlation between mean report scores with sentence position within the experiment ("sentence number"), which was also significant (Pearson's $r = 0.61$, $p < 0.001$, Fig. 1B).

To measure individual learning rate, we fitted a linear and a power law curve to each subject's performance over time (see Section 2.2.2). To determine which model better described the data, we calculated BICs for the two fits within each subject. A Wilcoxon signed-rank test showed that the BICs of the linear fit (median 239.99, range 223.13–252.73) were significantly smaller than those of the power law fit (median 243.57, range 227.14–261.18; $p < 0.001$). Smaller BIC scores indicate a better fit. Thus, the slope of the linear fit was chosen as an estimate for all individual adaptation rates; in the following analysis we will refer to this simply as "adaptation slope". Table 2 shows the descriptive statistics associated with this perceptual learning measure.

As expected, there was considerable variability in adaptation slopes across listeners; this is easily visible from Fig. 1C where the linear fits for all participants are shown.

Furthermore, the slopes of the linear fits correlated negatively with their intercepts (Pearson's $r = -0.69$, $p < 0.01$), indicating that listeners who started out with a good performance improved less over the course of the experiment.

3.2. Working memory test scores

Forward digit span and backward digit span scores were calculated as the sum of all correctly-repeated items of each test. Similarly, scores for the nonword repetition test were evaluated as the sum of all correctly-repeated items. Descriptive statistics of forward, backward digit span and nonword repetition scores are shown in Table 2.

3.3. AM rate discrimination

JND thresholds for AM discrimination performance were determined as described in Section 2.2.5. Table 2 shows

Table 2

Descriptive statistics for test scores across participants.

	Mean	SD	Min	Max	IQR
Adaptation slope	0.0033	0.0012	0.0006	0.0058	0.0012
AM discrimination (JND in Hz)	1.3	0.43	0.66	2.03	0.71
Forward digit span (score)	9	2.4	4	14	4
Backward digit span (score)	8.39	1.97	5	11	3
Nonword repetition test	25.61	3.68	17	30	6

Table 3

Pearson's correlation coefficients for two-tailed correlations between behavioural test scores. no indication = non significant.

	AM discrimination	Forward digit span	Backward digit span	Nonword repetition
Adaptation slope	-0.515*	0.404	0.234	0.145
AM discrimination	-	-0.727**	-0.526*	-0.262
Forward digit span	-	-	0.304	0.417
Backward digit span	-	-	-	-0.213

* = $p < 0.05$.

** = $p < 0.01$.

descriptive statistics for JND thresholds. There was no difference in mean performance between the first ($M \pm SEM = 0.768 \pm 0.018$) and the second half (0.767 ± 0.016) of the experiment as shown by a paired sample *t*-test ($t(15) = 0.08$, $p = 0.95$), indicating that participants' performance did not improve over time in the AM discrimination task.

3.4. Pearson's correlations between test scores

In order to examine whether other test scores could predict adaptation slope, a two-tailed Pearson's correlation was calculated for all pairs of test scores. There was a significant negative correlation between adaptation slope and AM discrimination thresholds, confirming that participants who adapted faster to vocoded speech also performed better in the AM discrimination task. AM discrimination thresholds also correlated negatively with forward as well backward digit span, revealing that better AM discrimination performers also had a higher digit span. No significant correlation was found between other pairs of test scores (all correlation coefficients can be found in Table 3; a scatterplot of the correlation between AM discrimination thresholds and adaptation slope is shown in Fig. 2A).

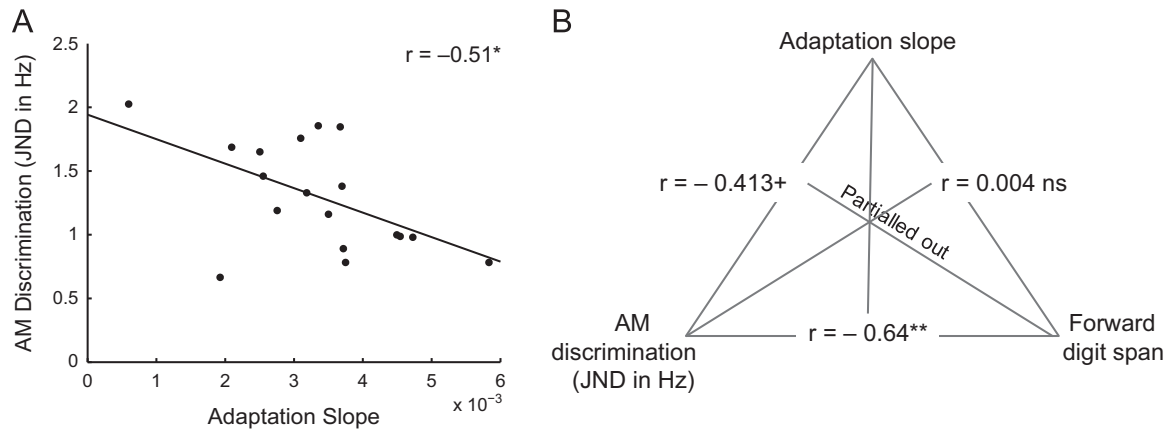


Fig. 2. (A) Adaptation slope and AM discrimination threshold correlate significantly. (B) Partial correlation coefficients between adaptation slope, AM discrimination threshold and forward digit span (** $p < 0.01$, * $p < 0.05$, + $p < 0.1$, ns=non significant).

3.5. *Partial correlation coefficients*

To analyse more closely the relationship between adaptation slope, AM discrimination and forward digit span, we calculated partial correlation coefficients between these three variables of interest (Fig. 2B). When partialling out forward digit span, the correlation between adaptation slope and AM discrimination threshold dropped to a non-significant trend; AM discrimination and forward digit span remained highly significantly correlated when partialling out adaptation slope.

3.6. *Voxel-based morphometry results*

To identify brain areas predictive of individual adaptation to vocoded speech, a multiple regression analysis of the VBM pre-processed modulated grey matter images with adaptation slope as covariate of interest and four nuisance variables coding for sex, age, coil of acquisition and time since acquisition was carried out. Regional volume correlated positively with adaptation slope in the left thalamus (pulvinar), indicating that better learners had increased GMV in the pulvinar. On the other hand, poorer learners had larger GMV in right-lateralised middle frontal gyrus

and inferior temporal gyrus, as shown by a negative correlation between regional volume and adaptation slope (see Fig. 3 and Table 4).

As AM discrimination thresholds were predictive of individual adaptation to vocoded speech at the behavioural level, we wanted to test whether there was an overlap in brain structures predictive of both measures. The whole-brain multiple regression analysis was repeated with AM discrimination thresholds as covariate of interest. We found a positive correlation between AM discrimination thresholds and GMV in the left superior parietal gyrus, meaning that poorer performers had increased GMV. A negative correlation was found in the occipital pole and posterior cingulate gyrus, signifying that better AM discrimination abilities were associated with higher GMV in these regions (Table 4). There were no overlaps between areas predictive of adaptation slope and AM discrimination rate.

In addition to the whole-brain analysis, we performed an ROI-analysis in the left IFG, where Eisner et al. (2010) had found a higher fMRI signal during vocoded speech perception in better learners. There was a non-significant trend for a positive correlation between mean GMV in the left IFG (at $-46\ 26\ 20$) and adaptation slope ($z = 1.47, p = 0.075$).

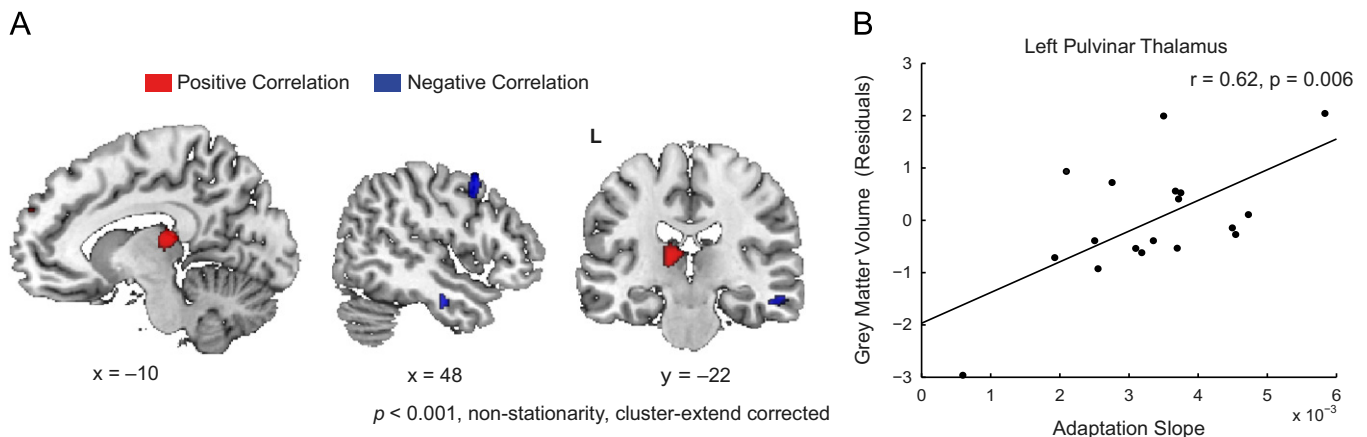


Fig. 3. Results of voxel-based morphometry analysis. (A) Better learners had increased grey matter volume in the left thalamus whereas poorer learners had increased GMV in right-lateralised middle frontal and inferior temporal gyrus. (B) Correlation between GMV residuals (having removed the effects of all covariates of no interest) and adaptation slope in the left thalamus.

Table 4
Summary of VBM results.

Location	Coordinates (MNI space)		Z- score	Extent (mm ³)	
Adaptation to vocoded speech (slope)					
<i>Positive correlation</i>					
L thalamus (pulvinar)	−12	−24	12	3.56	280
<i>Negative correlation</i>					
R middle frontal gyrus	48	9	51	3.92	152
R inferior temporal gyrus	48	−11	−25	3.68	66
AM discrimination threshold (JND)					
<i>Positive correlation</i>					
L superior parietal gyrus	−15	−47	63	3.71	141
<i>Negative correlation</i>					
L occipital pole	−26	−102	−3	4.07	137
L posterior cingulate gyrus	−12	−44	5	3.47	85

4. Discussion

In the present study, we tested individual short-term adaptation to degraded (i.e., noise-vocoded) speech. We hypothesised that individual differences in adaptation to vocoded speech should be predictable by non-speech auditory, cognitive, and neuroanatomical factors. The main finding of the present study is the predictive potential of AM rate discrimination ability and morphology of the left thalamus in individual perceptual learning of vocoded speech.

4.1. Perceptual adaptation

Perceptual adaptation to degraded speech affords a robust understanding of speech in adverse listening situations. Our results corroborate that perceptual learning of vocoded speech is possible in a short period of time. It has been shown that various other forms of degraded speech or speech that deviates from the norm such as unfamiliar accents (Bradlow & Bent, 2008), speech in noise (Song, Skoe, Banai, & Kraus, 2011), and even more artificial forms of degradation such as time-compression (Dupoux & Green, 1997; Pallier, Sebastian-Galles, Dupoux, Christophe, & Mehler, 1998) lead to fast perceptual adaptation.

As expected, we found considerable individual variability in perceptual learning of vocoded speech. Nonetheless, all participants were far from ceiling performance after listening to 100 sentences of 4-band vocoded speech. The finding that the learning curves in this experiment were better described by a linear than a power law fit reflects the fact that performance changes at a constant rate rather than changing at a different rates throughout the learning process (the latter being reflected by a power law curve).

The observed negative correlation between slope and intercept of the linear fits suggests that participants starting better had a more limited scope to improve. This corroborates previous findings that poorer baseline performance is correlated with greater learning in auditory tasks (Amitay, Hawkey, & Moore, 2005) and perceptual learning of speech (Stacey & Summerfield, 2007).

Despite demonstrations that feedback can greatly enhance the speed of vocoded speech learning (Davis et al., 2005; Hervais-Adelman, Davis, Johnsrude, & Carlyon 2008), we showed in the present study that fast perceptual learning of vocoded speech is still achieved in the absence of feedback.

In contrast to perceptual learning of vocoded speech, no performance improvement over time was observed in the AM discrimination task. This suggests that sensitivity to AM rate is rather a pre-existing, non-improvable feature. Alternatively, exposure to the paradigm during 20 min only might not be an appropriately long period to achieve performance improvement.

The latter view is supported by a study on perceptual learning of AM rate discrimination where participants' performance improved after training with 720 trials per day for 6–7 days (Fitzgerald & Wright, 2011).

4.2. The role of working memory in perceptual learning

Contrary to our expectations, nonword repetition test scores did not correlate with any of the other measures. This result could reflect a reality, indicating that adaptation to vocoded speech is independent of phonological working memory, or it may simply reflect the inability of the test to capture phonological working memory. The nonword repetition test had initially been designed for diagnostics of dyslexic children (Mottier, 1951); it was adapted in-house for phonological working memory testing in adults by extending the upper limit to a list length of 15 syllables. Yet, this extended version remains to be validated carefully in the normal-hearing adult population.

In contrast to Pisoni and Cleary (2003), we did not find a correlation between forward digit span and vocoded speech learning. Nevertheless, digit span did explain part of the common variance between adaptation slope and AM discrimination threshold as shown by the decrease in correlation strength when partialling out forward digit span (Fig. 2B).

The role of working memory in degraded speech perception has been debated extensively. Rabbitt (1968) suggests with his "effortful hypothesis" that working memory becomes a limiting factor when perceptual effort is required for degraded speech recognition. He observed that masking a word (with noise) makes not only the word itself more difficult to recall, but also the clearly-presented word prior to it. This is proposed to be due to the fact that resources are drawn on to accomplish "perceptual effort" that would otherwise be available for storage in working memory.

Based on this hypothesis, Piquado and colleagues recently proposed a model by which masked words disrupt the short-term memory buffer (Piquado, Cousins, Wingfield, & Miller, 2010). Previous work of the same group offered a biologically motivated computational model where, following the idea of Hebbian learning, synaptic strengthening is necessary for the recall of a word. They assume that a degraded stimulus causes weaker neuronal activity (synaptic firing) and therefore leads to fewer synaptic strengthening than a clear stimulus. Thus, recall probability is reduced when the stimulus is degraded (Miller & Wingfield, 2010). The proposed mechanism implies that working memory is a limiting factor in degraded speech perception.

When summing across the results cited here, the following picture emerges: An increased working memory buffer should facilitate adaptation to degraded speech, and it is a reasonable speculation that the relationship between adaptation slope and digit span (here a Pearson's r of 0.4) would reach significance in a larger sample. But it is also becoming clear that the story of working memory and speech perception is not fully told yet (Jacquemot & Scott, 2006). Moreover, working memory performance itself has been shown to be at least partly predictable by perceptual uncertainty of the sensory input (e.g., Burkholder, Pisoni, & Svirskey, 2005). Thus, in line with our main finding (see next paragraph), specifically developed auditory short-term memory tests might ultimately prove most predictable of individuals' adaptation to sensory challenges.

4.3. AM rate discrimination

One of the most intriguing findings of the present study is that individual AM rate discrimination thresholds predicted adaptation to vocoded speech (Fig. 2A). This supports our hypothesis

that non-speech auditory skills contribute to perceptual learning of degraded speech.

Clearly, different sources of information can contribute to perceptual learning of degraded speech. Some studies have found evidence for “top-down” influences (i.e., higher-level cognitive processes) on vocoded speech perception, pointing out the importance of lexical information for perceptual learning (Davis & Johnsrude, 2007; Davis et al., 2005) and the facilitatory effect of semantic predictability on speech recognition (Obleser, Wise, Alex Dresner, & Scott, 2007). Other studies have suggested a sublexical mechanism for perceptual learning of vocoded speech. For example, Hervais-Adelman et al. (2008) showed that training with noise-vocoded nonwords is as effective as with real words for listeners learning to understand vocoded speech. Similarly, Sebastian-Galles, Dupoux, Costa, and Mehler (2000) proposed that adaptation to time-compressed speech operates at a sublexical, phonological level. They found that native speakers of Spanish benefit from training with time-compressed Greek sentences (which is phonologically similar but lexically distant from Spanish; the participants did not understand Greek) as much as from training with Spanish sentences. The authors concluded that adaptation does not depend on comprehension of the adapting sentences, but rather on the phonological properties of the adapting language (Sebastian-Galles et al., 2000). Another study showed that learning can be driven by low-level acoustic cues when listeners adapt to an artificial accent (Idemaru & Holt, 2011). Our findings provide further evidence for the importance of a “bottom-up” acoustic analysis that the auditory system has to perform when confronted with degraded speech (see also Mattys, Brooks, & Cooke, 2009). Initial degraded speech perception studies have underlined that speech waveform envelopes are almost sufficient for speech recognition (Shannon et al., 1995; Van Tasell, Soli, Kirby, & Widin, 1987). Here we show that individual sensitivity to AM is even predictive of degraded speech recognition performance.

Non-speech auditory skills were dependent on short-term working memory, as shown by the correlation with forward digit span. This could partly be due to the design of the AM discrimination task, a 3-AFC task which makes considerable demands on working memory. Thus, listeners who have a poorer auditory short-term memory might perform worse on the 3-AFC task, because they may have more difficulties keeping the three stimuli available for comparison.

We cannot conclude from our results whether the observed relationship between perceptual adaptation and AM discrimination abilities is specific to a modulation rate around 4 Hz, since we did not test other modulation frequencies. There is only cursory evidence that such a relationship is not present for higher modulation rates (i.e., vocoded speech recognition performance and sensitivity to AM depth at 8 Hz; personal communication, Carolyn McGettigan). We thus speculate that especially slow amplitude modulations in the syllabic range are predictive of adaptation to degraded speech as they are necessary for accurate speech recognition (Drullman, Festen, & Plomp, 1994a, 1994b).

4.4. Anatomical predictors of adaptation to vocoded speech

We found that individual differences in grey matter volume in the left pulvinar thalamus predicted degraded speech learning (Fig. 3B). There is evidence from electrophysiological studies that the left pulvinar is involved in processing speech and short term memory. Ojemann and colleagues performed electrical stimulations of the thalamus prior to therapeutic thalamectomy in patients with chronic pain, Parkinson's disease and other disorders to verify localisation of ablation. During these procedures, the patient was awake and had to perform language or memory tasks. They observed that electrical stimulation of the dominant pulvinar (i.e., left pulvinar in right-handers) interfered with object

naming and increased recall errors during a verbal short-term memory task. The authors attributed a gating function to the pulvinar where the thalamus controls retrieval from memory (Johnson & Ojemann, 2000).

In deaf subjects, the pulvinar appears to play a role in auditory imagery. Lazard, Giraud, Truy, and Lee (2011) conducted an fMRI study, where post-lingually deaf candidates for cochlear implantation imagined non-speech, non-human sounds and colours (the latter was used as baseline condition). They found increased activity in the right pulvinar in deaf patients compared to controls in the sound imagery versus colour imagery task. Additionally, within the CI candidate group, activity in the right thalamus was positively correlated with duration of hearing loss (Lazard et al., 2011).

In the first-order auditory thalamus, the medial geniculate body (MGB), von Kriegstein, Patterson, and Griffiths (2008) found activity to be modulated dependent on specific features of speech processing. In this fMRI study, listeners performed either syllable recognition or a speaker/loudness recognition task on the same stimuli. Activity in the MGB was modulated dependent on the task performed.

Furthermore, Peelle, Troiani, Grossman, and Wingfield (2011) observed hearing-related changes in neural activity in the thalamus. This fMRI study compared older adults with different hearing abilities (pure tone audiometric thresholds). Individuals with worse hearing thresholds had less activity in the thalamus when listening to grammatically more complex sentences. The authors inferred that the thalamus plays an important role in transferring auditory information to prefrontal and premotor areas and thus is crucial for speech comprehension.

The high interconnectivity of the thalamus with cortical areas could also be an explanation for a role of the pulvinar in adaptation to degraded speech. In mouse studies it has been shown that connectivity between the sensory thalamus (ventral MGB) and auditory cortex is critical for auditory processing (Barkat, Polley, & Hensch, 2011). A recent animal study on visual perceptual learning stated that learning can be explained by changes in the feed-forward connections between the sensory thalamus and the primary sensory cortex, improving the inference performed by neural circuits (Bejjanki, Beck, Lu, & Pouget, 2011). In non-human primates it has been shown that auditory belt areas receive input from the thalamus, including the pulvinar (Hackett et al., 2007). Our knowledge is of course more limited in humans and is mainly based on diffusion weighted brain imaging data. According to the Oxford Thalamic Connectivity Atlas (<http://www.fmrib.ox.ac.uk/connect/>; Behrens et al., 2003), which provides probabilistic connectivity values based on diffusion tractography, the peak voxel of our VBM results (at the MNI coordinates $-12 -24 12$) exhibited a high probability ($p=0.82$) of being connected to the temporal cortex.

In addition to this positive correlation with perceptual learning of vocoded speech, we found a negative correlation between GMV and adaptation slope in right-lateralised middle frontal gyrus and inferior temporal gyrus (Fig. 3A). In the brain morphometry literature, a common explanation of how increased regional brain volume/thickness can reflect impaired processing capabilities is a lack of normal neuronal pruning during brain development. For example, Hyde et al. (2007) found increased cortical thickness in the right auditory cortex and inferior frontal gyrus in subjects suffering from amusia (i.e., showing impairments in music production and perception). Another recent study found that paediatric patients of pseudoheya, i.e., children with an abnormal audiogram despite no known organic causes (normal auditory brain stem responses etc.), have increased GMV in the superior temporal gyrus as well as the medial frontal gyrus. The authors suggested that insufficient neuronal pruning during development in these areas might cause higher GMV and the associated symptoms (Tomoda, Kinoshita, Korenaga, & Mabe, 2012). Both studies, although drawn

from patient populations, indicate that increased GMV in the temporal cortex (i.e., in a region close to where we found higher GMV in slower vocoded speech learners) is associated with difficulties in auditory processing.

The fact that we did not identify any other regions associated with vocoded-speech learning might be due to the rather small number of subjects. Although we did not find increased GMV in the left IFG corresponding to the functional MRI results of Eisner et al. (2010) at the whole-brain corrected level of significance, we did observe a trend-level correlation between IFG volume and learning scores in an ROI-based analysis. Though we do not want to overinterpret this trend, higher brain activity in the IFG in better learners might be paralleled by increased grey matter, corroborating the left IFG's role in perceptual learning.

4.5. Clinical implications

The present study adds two important findings to the vocoded speech perception literature: In normal-hearing adults listening to a CI simulation, AM sensitivity as well as brain morphology in the left thalamus are predictive of perceptual adaptation.

These experiments in normal-hearing listeners could be usefully transferred to CI patients in two different ways: First, we suggest that pre-implantation brain scans can be used to assess anatomical predictors of post-implantation speech comprehension. So far, there have been attempts to prospectively predict CI patients' speech

recognition performance from pre-implantation functional MRI scans (Lazard et al., 2010), but to our knowledge no studies of brain structural predictors of CI success have been conducted.

Second, after cochlear implantation, CI patients could be tested with an AM discrimination paradigm to verify whether their AM sensitivity is also correlated with speech comprehension. This could help to clarify how much of the individual variability in CI patients' speech recognition can be explained by basic auditory processing capabilities.

Acknowledgements

Research was supported by the Max Planck Society (Max Planck Research Group fund to J.O.) and F.E. was supported by NWO grant 275-75-009. Zoe Schlüter helped acquire and analyse the data; Lars Meyer provided scripts and was very helpful with the VBM analysis; Antje Strauß helped develop the sentence material; Claudia Männel provided an adapted version of the nonword repetition test. Two anonymous reviewers contributed constructive comments and helped improve this manuscript.

Appendix. The German SPIN sentences

See Table A1.

Table A1
The German SPIN sentences.

	Low-predictable sentences	High-predictable sentences
1	Wir freuen uns sehr über die Laken	Sie bezieht das Bett mit frischen Laken
2	Paul erzählte von dem neuen Netz	Er fing diesen Fisch in seinem Netz
3	Frau Meier spricht über die Küste	Das Boot segelte entlang der Küste
4	Ich möchte das mit der kleinen Blüte	Alle Blumen standen in voller Blüte
5	Peter ärgert sich über die Gläser	Die Biertrinker hoben ihre Gläser
6	Er kümmert sich jetzt nicht um den Pfeil	Ihn traf dort ein vergifteter Pfeil
7	Regelmäßig holt er sich eine Wunde	Am Knie bildet sich Schorf über der Wunde
8	Maria hielt das sehr lange für Weizen	In der Auswahl trennt sich die Spreu vom Weizen
9	Daniel dachte überhaupt nicht an Sport	Boxen ist ein gefährlicher Sport
10	Frau Maier sorgte sich um den Arm	Nehmen Sie das Baby auf den Arm
11	Johannes benötigte schnell den Block	Reiß doch ein paar Zettel aus dem Block
12	Wir betrachteten aufmerksam die Truhe	Der Schatz befindet sich in einer Truhe
13	Kathrin entdeckte vorhin ein Nest	Die Wespen schwärmten aus ihrem Nest
14	Eva und Anton mustern die Scheiben	Schneide den Schinken in dünne Scheiben
15	David hoffte vergeblich auf einen Kuss	Sie gab ihm einen innigen Kuss
16	Ich glaube dir die Sache mit dem Finger	Ich schnitt mir aus Versehen in den Finger
17	Max verschwindet eilig mit dem Wachs	Die Flamme der Kerze schmolz das Wachs
18	Er bittet ihn dauernd um das Schloss	Dieser Schlüssel passt nicht in das Schloss
19	Hans ärgerte sich über den Kratzer	Woher hat dein Auto diesen Kratzer
20	Er erkundigte sich nach einer Wiege	Das Baby schlief in seiner Wiege
21	Herr Koch rätselt über die Dornen	Ein Rosenstrauch hat spitze Dornen
22	Andreas weiß mehr über das Floß	Die schiffbrüchigen Segler bauten ein Floß
23	Der alte Mann registrierte das Tor	Der kleine Junge schießt den Ball ins Tor
24	Er überprüft sehr genau den Unfall	Es gab keine Verletzten bei dem Unfall
25	Unmöglich findet sie auf Anhieb das Deck	Der Matrose schrubbte sorgfältig das Deck
26	Der junge Mann bemerkte nicht die Maus	Sie spielten mit der Polizei Katz' und Maus
27	Frau Krause widmet sich mit Anna dem Spiel	Sie machte gute Miene zum bösen Spiel
28	Sie wussten schon lange von dem Fell	Der große Bär hat ein dickes Fell
29	Ich verkaufe Marie niemals das Sofa	Jens schlief auf dem ausgezogenen Sofa
30	Leider fehlt uns eine passende Dose	Die Plätzchen sind in der silbernen Dose
31	Thomas war gestern wieder nicht im Stall	Klaus geht zum Kühmelken in den Stall
32	Frau Kaiser reagierte nicht auf den Schock	Seit dem Zusammenstoß steht er unter Schock
33	Hanna untersucht penibel die Hüften	Lars stemmte die Hände in die Hüften
34	Anja und ich kümmern uns um den Saft	Zum Frühstück trank er frisch gepressten Saft
35	Er achtete nicht auf seine Faust	Er schlug mich mit der geballten Faust
36	Er fragte mich nach der neuen Mannschaft	Der Trainer qualifiziert seine Mannschaft
37	Julia zeigte Felix einen Schwan	Das Entlein schwamm mit dem weißen Schwan
38	Anton lügt bezüglich seiner Münze	Zur Entscheidung werfen wir eine Münze
39	Jeder braucht unbedingt einen Besen	Das Dienstmädchen fegt das Haus mit dem Besen
40	Sie fragten nach dem hinteren Schuppen	Als Holzlager baute er einen Schuppen

Table A1 (continued)

	Low-predictable sentences	High-predictable sentences
41	Julius fragt nach der neuen Straße	Er parkte den Wagen am Rand der Straße
42	Philipp ignoriert die Fragen der Leute	Sie hörte nicht auf das Geschwätz der Leute
43	Wir hörten schon von der Beute	Der Einbrecher entkam mit der Beute
44	Wieder befassten sie sich mit der Hose	Er machte sich aus Angst in die Hose
45	Zerstreut erwähnte sie nicht mal die Rippe	Gott schuf Eva aus Adams Rippe
46	Herr Wegener observiert die Uhr	Wir hörten das Ticken der alten Uhr
47	Sie waren überrascht vom Applaus	Nach dem Stück ernteten sie viel Applaus
48	Monika überprüfte den Film	Der Fotograf wechselte den Film
49	Stephan lernte alles über den Kiefer	Vom Kaugummikauen schmerzt mein Kiefer
50	Das ist für heute meine erste Suppe	Sie kochte ihm eine herzhafte Suppe
51	Wir unterhalten uns noch über den Preis	Mit dem Beitrag gewann sie den ersten Preis
52	Frau Lehmann ergatterte das letzte Brot	Der Asket lebt nur von Wasser und Brot
53	Wir erinnern uns alle an den Schrei	Vor Schreck entfuhr ihr ein gellender Schrei
54	Niemand besprach genau die Handlung	Er beging eine strafbare Handlung
55	Lena redet nicht gern über ihr Bein	Der Schwimmer bekam einen Krampf im Bein
56	Frau Keller zeigte mir meine Reihe	Unsre Plätze sind in der zweiten Reihe
57	Sie dachte gar nicht mehr an die Höhle	Der Löwe schlief in seiner Höhle
58	Alexandra wirft es in das Feuer	Das gebrannte Kind meidet Feuer
59	Der Junge holt draußen die Falle	Die kleine Maus tappte in die Falle
60	Frau Jung erinnerte an den Tee	Nach dem Aufgießen zieht der schwarze Tee
61	Mario interessiert sich für das Recht	Der Richter gab dem Ankläger Recht
62	Ein Mann untersuchte gründlich den Ochsen	Vor den Pflug spannte man früher den Ochsen
63	Helena unterschätzte den Tribut	Die Katastrophe fordert ihren Tribut
64	Sie musste nach Hause wegen ihres Rocks	Sie kürzte den Saum ihres neuen Rocks
65	Morgen entfernen sie schließlich den Stamm	Ein alter Baum hat einen dicken Stamm
66	Kathrin versucht es mit den Karten	Laut Spielregeln mischt man nun die Karten
67	Tom freut sich über diese schöne Seide	Das Halstuch besteht aus glänzender Seide
68	Er unterhielt sich lange mit dem Gast	Herr Fischer ist ein gern gesehener Gast
69	Er sieht nicht gut aus ohne Gürtel	In Karate hat er den schwarzen Gürtel
70	Sie stellte sich neben die Ecke	Folgen Sie der Straße um die Ecke
71	Am liebsten fotografiert er die Taube	Als Friedensvogel gilt die weiße Taube
72	Täglich produziert sie ein neues Kleid	Die Braut trägt ein langes weißes Kleid
73	Er erspät in der Ferne die Flut	Die Gezeiten sind Ebbe und Flut
74	Gewöhnlich nutzen wir die Räder	Ein Auto hat immer vier Räder
75	Die Leute begutachteten die Perlen	Sie fädelt auf die Kette teure Perlen
76	Die Gruppe orientiert sich in der Bucht	Der Hafen liegt in einer stillen Bucht
77	Ich zeige dir noch einmal das Kalb	Eine Kuh säugt nur ihr eigenes Kalb
78	Sven gab ihr letzte Woche den Leim	Zusammengeklebt wird das Holz mit Leim
79	Isabel erwähnte auch die Klinge	Sein Klappmesser hat eine scharfe Klinge
80	Nina hatte keine Ahnung von der Post	Briefmarken bekommst du bei der Post
81	Sein Blick schweifte kurz über das Meer	Der Mann ertrank im stürmischen Meer
82	Ich untersuche den grünen Lack	Vom rostigen Auto blätterte der Lack
83	Heute geht Katharina nicht zum Bach	Sie fischten in einem plätschernden Bach
84	Harald verfolgte den Anfang der Spuren	Der Täter hinterließ sichtbare Spuren
85	Brigitte vertraut nur ihrem Arzt	Fragen Sie Ihren behandelnden Arzt
86	Du verwechselst immer die Löffel	Für Suppen nehmen wir silberne Löffel
87	Mir wird immer schlecht von der Rinde	Wir ritzen ein Herz in die Rinde
88	Gewöhnlich verarbeitet sie kein Mehl	Nun mischt man das Ei mit gesiebt Mehl
89	Am meisten mag er den roten Ball	Der Mittelfeldspieler köpfte den Ball
90	Den ganzen Tag schon will er keinen Spaß	Austricksen macht ihm diebischen Spaß
91	Zeige mir mal genauer die Bahn	Ich fahre dieses mal mit der Bahn
92	Morgen nehmen wir mal wieder Obst	Sie kaufte Gemüse und frisches Obst
93	Hoffentlich fragt Sascha nach dem Riegel	Der Verbrecher sitzt hinter Schloss und Riegel
94	Frau Schulze wusste von den Flaschen	Bier kauft man in Dosen oder Flaschen
95	Das Mädchen beachtet nicht die Flammen	Aus dem Haus züngelten lodernde Flammen
96	Peter berichtete von dem Regen	Nimm den Schirm mit bei Schnee und Regen
97	Martin erwähnte nicht einmal die Wanzen	In ihrem Bett krabbelten die Wanzen
98	Karin fühlte sich sehr wohl im Wald	Die Arbeiter rodeten den Wald
99	Sie beneidet mich um meine Sauna	Im Bad ist auch eine finnische Sauna
100	Offensichtlich fehlt heute Morgen das Schiff	Der Kapitän lenkte geschickt sein Schiff

References

- Adank, P., & Devlin, J. T. (2010). On-line plasticity in spoken sentence comprehension: Adapting to time-compressed speech. *Neuroimage*, 49, 1124–1132.
- Amitay, S., Hawkey, D. J., & Moore, D. R. (2005). Auditory frequency discrimination learning is affected by stimulus variability. *Perception and Psychophysics*, 67, 691–698.
- Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science*, 2, 396–408.
- Ashburner, J. (2007). A fast diffeomorphic image registration algorithm. *Neuroimage*, 38, 95–113.
- Ashburner, J. (2009). Computational anatomy with the SPM software. *Magnetic Resonance Imaging*, 27, 1163–1174.
- Ashburner, J., & Friston, K. J. (2000). Voxel-based morphometry—the methods. *Neuroimage*, 11, 805–821.
- Baddeley, A. (2012). Working memory: theories, models, and controversies. *Annual Review of Psychology*, 63, 1–29.
- Barkat, T. R., Polley, D. B., & Hensch, T. K. (2011). A critical period for auditory thalamocortical connectivity. *Nature Neuroscience*, 14, 1189–1194.
- Behrens, T. E., Johansen-Berg, H., Woolrich, M. W., Smith, S. M., Wheeler-Kingshott, C. A., Boulby, P. A., et al. (2003). Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. *Nature Neuroscience*, 6, 750–757.
- Bejjanki, V. R., Beck, J. M., Lu, Z. L., & Pouget, A. (2011). Perceptual learning as improved probabilistic inference in early sensory areas. *Nature Neuroscience*, 14, 642–648.

- Blamey, P. J., Sarant, J. Z., Paatsch, L. E., Barry, J. G., Bow, C. P., Wales, R. J., et al. (2001). Relationships among speech perception, production, language, hearing loss, and age in children with impaired hearing. *Journal of Speech, Language, and Hearing Research*, 44, 264–285.
- Bradlow, A. R., & Bent, T. (2008). Perceptual adaptation to non-native speech. *Cognition*, 106, 707–729.
- Brett, M., Anton, J., Valabregue, R., & Poline, J. (2002). Region of interest analysis using an SPM toolbox. *Presentation at the 8th International Conference on Functional Mapping of the Human Brain, Sendai, Japan, June 2–6, 2002*.
- Burkholder, R. A., Pisoni, D. B., & Svirsky, M. A. (2005). Effects of a cochlear implant simulation on immediate memory in normal-hearing adults. *International Journal of Audiology*, 44, 551–558.
- Davis, M. H., & Johnsrude, I. S. (2007). Hearing speech sounds: top-down influences on the interface between audition and speech perception. *Hearing Research*, 229, 132–147.
- Davis, M. H., Johnsrude, I. S., Hervais-Adelman, A., Taylor, K., & McGettigan, C. (2005). Lexical information drives perceptual learning of distorted speech: evidence from the comprehension of noise-vocoded sentences. *Journal of Experimental Psychology: General*, 134, 222–241.
- Drullman, R., Festen, J. M., & Plomp, R. (1994a). Effect of reducing slow temporal modulations on speech reception. *Journal of the Acoustical Society of America*, 95, 2670–2680.
- Drullman, R., Festen, J. M., & Plomp, R. (1994b). Effect of temporal envelope smearing on speech reception. *Journal of the Acoustical Society of America*, 95, 1053–1064.
- Dupoux, E., & Green, K. (1997). Perceptual adjustment to highly compressed speech: effects of talker and rate changes. *Journal of Experimental Psychology: Human Perception and Performance*, 23, 914–927.
- Eisner, F., McGettigan, C., Faulkner, A., Rosen, S., & Scott, S. K. (2010). Inferior frontal gyrus activation predicts individual differences in perceptual learning of cochlear-implant simulations. *Journal of Neuroscience*, 30, 7179–7186.
- Fitzgerald, M. B., & Wright, B. A. (2011). Perceptual learning and generalization resulting from training on an auditory amplitude-modulation detection task. *Journal of the Acoustical Society of America*, 129, 898–906.
- Gathercole, S. E., Willis, C. S., Baddeley, A. D., & Emslie, H. (1994). The children's test of nonword repetition: a test of phonological working memory. *Memory*, 2, 103–127.
- Giraud, A. L., Lorenzi, C., Ashburner, J., Wable, J., Johnsrude, I., Frackowiak, R., et al. (2000). Representation of the temporal envelope of sounds in the human brain. *Journal of Neurophysiology*, 84, 1588–1598.
- Golestani, N., Paus, T., & Zatorre, R. J. (2002). Anatomical correlates of learning novel speech sounds. *Neuron*, 35, 997–1010.
- Greenwood, D. D. (1990). A cochlear frequency-position function for several species—29 years later. *Journal of the Acoustical Society of America*, 87, 2592–2605.
- Hackett, T. A., Smiley, J. F., Ulbert, I., Karmos, G., Lakatos, P., de la Mothe, L. A., et al. (2007). Sources of somatosensory input to the caudal belt areas of auditory cortex. *Perception*, 36, 1419–1430.
- Hayasaka, S., Phan, K. L., Liberzon, I., Worsley, K. J., & Nichols, T. E. (2004). Nonstationary cluster-size inference with random field and permutation methods. *Neuroimage*, 22, 676–687.
- Hervais-Adelman, A., Davis, M. H., Johnsrude, I. S., & Carlyon, R. P. (2008). Perceptual learning of noise vocoded words: effects of feedback and lexicality. *Journal of Experimental Psychology: Human Perception and Performance*, 34, 460–474.
- Houtgast, T., & Steeneken, H. J. (1985). A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria. *Journal of the Acoustical Society of America*, 77, 1069–1077.
- Hyde, K. L., Lerch, J. P., Zatorre, R. J., Griffiths, T. D., Evans, A. C., & Peretz, I. (2007). Cortical thickness in congenital amusia: when less is better than more. *Journal of Neuroscience*, 27, 13028–13032.
- Idemaru, K., & Holt, L. L. (2011). Word recognition reflects dimension-based statistical learning. *Journal of Experimental Psychology: Human Perception and Performance*, 37, 1939–1956.
- Jacquemot, C., & Scott, S. K. (2006). What is the relationship between phonological short-term memory and speech processing?. *Trends in Cognitive Sciences*, 10, 480–486.
- Johnson, M. D., & Ojemann, G. A. (2000). The role of the human thalamus in language and memory: evidence from electrophysiological studies. *Brain and Cognition*, 42, 218–230.
- Kalikow, D. N., Stevens, K. N., & Elliott, L. L. (1977). Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *Journal of the Acoustical Society of America*, 61, 1337–1351.
- Lazard, D. S., Giraud, A. L., Truy, E., & Lee, H. J. (2011). Evolution of non-speech sound memory in postlingual deafness: implications for cochlear implant rehabilitation. *Neuropsychologia*, 49, 2475–2482.
- Lazard, D. S., Lee, H. J., Gaebler, M., Kell, C. A., Truy, E., & Giraud, A. L. (2010). Phonological processing in post-lingual deafness and cochlear implant outcome. *Neuroimage*, 49, 3443–3451.
- Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: a user's guide* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates pp 219–220.
- Mattys, S. L., Brooks, J., & Cooke, M. (2009). Recognizing speech under a processing load: dissociating energetic from informational factors. *Cognitive Psychology*, 59, 203–243.
- Miller, P., & Wingfield, A. (2010). Distinct effects of perceptual quality on auditory word recognition, memory formation and recall in a neural model of sequential memory. *Frontiers in Systems Neuroscience*, 4, 14.
- Mottier, G. (1951). Examination of the speech of children with reading difficulties. *Folia Phoniatrica (Basel)*, 3, 170–177.
- Obleser, J., Wise, R. J., Alex Dresner, M., & Scott, S. K. (2007). Functional integration across brain regions improves speech perception under adverse listening conditions. *Journal of Neuroscience*, 27, 2283–2289.
- Pallier, C., Sebastian-Galles, N., Dupoux, E., Christophe, A., & Mehler, J. (1998). Perceptual adjustment to time-compressed speech: a cross-linguistic study. *Memory and Cognition*, 26, 844–851.
- Peelle, J. E., Cusack, R., & Henson, R. N. (2012). Adjusting for global effects in voxel-based morphometry: gray matter decline in normal aging. *Neuroimage*, 60, 1503–1516.
- Peelle, J. E., Olafsen, T., Davis, M. H., & Wingfield, A. (2010). Perceptual learning in speech comprehension governed by power law dynamics. *Poster presentation at the 40th SfN annual meeting, San Diego, US, November 13–17, 2010*.
- Peelle, J. E., Troiani, V., Grossman, M., & Wingfield, A. (2011). Hearing loss in older adults affects neural systems supporting speech comprehension. *Journal of Neuroscience*, 31, 12638–12643.
- Piquado, T., Cousins, K. A., Wingfield, A., & Miller, P. (2010). Effects of degraded sensory input on memory for speech: behavioral data and a test of biologically constrained computational models. *Brain Research*, 1365, 48–65.
- Pisoni, D. B., & Cleary, M. (2003). Measures of working memory span and verbal rehearsal speed in deaf children after cochlear implantation. *Ear and Hearing*, 24, 1065–1205.
- Rabbitt, P. M. (1968). Channel-capacity, intelligibility and immediate memory. *Quarterly Journal of Experimental Psychology*, 20, 241–248.
- Romani, C., & Martin, R. (1999). A deficit in the short-term retention of lexical-semantic information: forgetting words but remembering a story. *Journal of Experimental Psychology: General*, 128, 56–77.
- Rorden, C., & Brett, M. (2000). Stereotaxic display of brain lesions. *Behavioural Neurology*, 12, 191–200.
- Rosen, S., Faulkner, A., & Wilkinson, L. (1999). Adaptation by normal listeners to upward spectral shifts of speech: implications for cochlear implants. *Journal of the Acoustical Society of America*, 106, 3629–3636.
- Sarant, J. Z., Blamey, P. J., Dowell, R. C., Clark, G. M., & Gibson, W. P. (2001). Variation in speech perception scores among children with cochlear implants. *Ear and Hearing*, 22, 18–28.
- Schwarz, & Gideon, E. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6(2), 461–464.
- Sebastian-Galles, N., Dupoux, E., Costa, A., & Mehler, J. (2000). Adaptation to time-compressed speech: phonological determinants. *Perception and Psychophysics*, 62, 834–842.
- Shannon, R. V., Galvin, J. J., 3rd, & Baskent, D. (2002). Holes in hearing. *Journal of the Association for Research in Otolaryngology*, 3, 185–199.
- Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science*, 270, 303–304.
- Song, J. H., Skoe, E., Banai, K., & Kraus, N. (2011). Training to improve hearing speech in noise: biological mechanisms. *Cerebral Cortex*.
- Stacey, P. C., & Summerfield, A. Q. (2007). Effectiveness of computer-based auditory training in improving the perception of noise-vocoded speech. *Journal of the Acoustical Society of America*, 121, 2923–2935.
- Tomoda, A., Kinoshita, S., Korenaga, Y., & Mabe, H. (2012). Pseudohypacusis in childhood and adolescence is associated with increased gray matter volume in the medial frontal gyrus and superior temporal gyrus. *Cortex*, 48, 492–503.
- Van Tasell, D. J., Soli, S. D., Kirby, V. M., & Widin, G. P. (1987). Speech waveform envelope cues for consonant recognition. *Journal of the Acoustical Society of America*, 82, 1152–1161.
- Viemeister, N. F. (1979). Temporal modulation transfer functions based upon modulation thresholds. *Journal of the Acoustical Society of America*, 66, 1364–1380.
- von Kriegstein, K., Patterson, R. D., & Griffiths, T. D. (2008). Task-dependent modulation of medial geniculate body is behaviorally relevant for speech recognition. *Current Biology*, 18, 1855–1859.
- Wakefield, G. H., & Viemeister, N. F. (1990). Discrimination of modulation depth of sinusoidal amplitude modulation (SAM) noise. *Journal of the Acoustical Society of America*, 88, 1367–1373.