

Facilitation of Inferior Frontal Cortex by Transcranial Direct Current Stimulation Induces Perceptual Learning of Severely Degraded Speech

Bernhard Sehm,^{1,3} Tim Schnitzler,¹ Jonas Obleser,² Agnes Groba,^{1,4} Patrick Ragert,¹ Arno Villringer,^{1,3} and Hellmuth Obrig^{1,3}

¹Department of Neurology, and ²Max Planck Research Group “Auditory Cognition,” Max-Planck-Institute for Human Cognitive and Brain Sciences, 04103 Leipzig, Germany, ³Clinic of Cognitive Neurology, University of Leipzig, 04103 Leipzig, Germany, and ⁴Department of Linguistics, University of Erfurt, 99089 Erfurt, Germany

Perceptual learning requires the generalization of categorical perceptual sensitivity from trained to untrained items. For degraded speech, perceptual learning modulates activation in a left-lateralized network, including inferior frontal gyrus (IFG) and inferior parietal cortex (IPC). Here we demonstrate that facilitatory anodal transcranial direct current stimulation (tDCS_{anodal}) can induce perceptual learning in healthy humans. In a sham-controlled, parallel design study, 36 volunteers were allocated to the three following intervention groups: tDCS_{anodal} over left IFG, IPC, or sham. Participants decided on the match between an acoustically degraded and an undegraded written word by forced same-different choice. Acoustic degradation varied in four noise-vocoding levels (2, 3, 4, and 6 bands). Participants were trained to discriminate between minimal (/Tisch/-FISCH) and identical word pairs (/Tisch/-TISCH) over a period of 3 d, and tDCS_{anodal} was applied during the first 20 min of training. Perceptual sensitivity (d') for trained word pairs, and an equal number of untrained word pairs, was tested before and after training. Increases in d' indicate perceptual learning for untrained word pairs, and a combination of item-specific and perceptual learning for trained word pairs. Most notably for the lowest intelligibility level, perceptual learning occurred only when tDCS_{anodal} was applied over left IFG. For trained pairs, improved d' was seen on all intelligibility levels regardless of tDCS intervention. Over left IPC, tDCS_{anodal} did not modulate learning but instead introduced a response bias during training. Volunteers were more likely to respond “same,” potentially indicating enhanced perceptual fusion of degraded auditory with undegraded written input. Our results supply first evidence that neural facilitation of higher-order language areas can induce perceptual learning of severely degraded speech.

Introduction

Speech comprehension requires the rapid mapping of complex sounds to meaning. Remarkably, the human brain affords this task almost without effort, although natural listening environments regularly provide highly variable, often degraded and distorted input. Moreover, listeners readily adapt even to highly adverse listening conditions, suggesting rapid plasticity in the speech comprehension network. Behavioral evidence indicates that linguistic competence at various levels supports adaptation to degraded speech: lexical information augments learning (Davis et al., 2005), but is not mandatory because training in an unknown language generalizes to the (untrained) native language depending on their phonological similarity (Bent et al., 2011). Adaptation profits from sentential semantics (Obleser et al.,

2007) and phonological information supplied by nonwords (Hervais-Adelman et al., 2008), and even basic durational and spectral cues may support learning (Winn et al., 2012). Similarly, generalization to untrained material, a crucial trait of perceptual learning (Wohllwill, 1958), may apply to domains beyond speech, such as identification of environmental sounds (Loebach et al., 2009).

Supporting the notion of a tight interplay between linguistic top-down influences and auditory bottom-up processing, neuroimaging studies have delineated a left-lateralized network including superior temporal [superior temporal gyrus (STG)/superior temporal sulcus (STS)], inferior frontal [inferior frontal gyrus (IFG)], and inferior parietal [inferior parietal cortex (IPC)] cortices. It is relatively uncontroversial that STG/STS activation correlates with intelligibility (Scott et al., 2000). Activation may extend more anteriorly as a function of lexical predictability (Obleser and Kotz, 2010) and more posteriorly depending on the syntactic information supplied (Friederici et al., 2010). The IFG shows more complex behavior: activation increases with semantic integration demands, but this effect is “gated” by intelligibility (Obleser and Kotz, 2010). As shown in an electrophysiological study on single words, the frontotemporal loop impacts on de-

Received Nov. 27, 2012; revised July 26, 2013; accepted Aug. 22, 2013.

Author contributions: B.S., A.G., P.R., A.V., and H.O. designed research; B.S. and T.S. performed research; B.S., T.S., J.O., P.R., and H.O. analyzed data; B.S., J.O., A.G., A.V., and H.O. wrote the paper.

The authors declare no competing financial interests.

Corresponding author: Dr. Bernhard Sehm, Max Planck Institute for Human Cognitive and Brain Sciences, Department of Neurology, Stephanstrasse 1a, D-04103 Leipzig, Germany. E-mail: sehm@cbs.mpg.de.

DOI:10.1523/JNEUROSCI.5466-12.2013

Copyright © 2013 the authors 0270-6474/13/3315868-11\$15.00/0

graded speech perception at a very early stage (Sohoglu et al., 2012). Most relevant here, left IFG activation has been shown to correlate with interindividual learning success in a short degraded speech learning paradigm performed in the scanner (Eisner et al., 2010).

In the same study, individual learning progress correlated with left IPC activation [angular gyrus (AG)], identifying IPC as another key area for top-down control during perceptual learning. More generally, IPC may afford meta-integration (Sharp et al., 2010a), converging pre-established semantic and contextual information. In this vein, a recent imaging study elegantly showed its role in “repair processes” during degraded speech comprehension, when auditory information is elusive (Shahin et al., 2009). On a supramodal level, the IPC is involved in decision-making processes particularly under uncertainty (Vickery and Jiang, 2009) and may modulate response bias (Eickhoff et al., 2011).

Proceeding from both behavioral and neuroimaging evidence, here we investigate whether facilitation of two higher-order nodes in the speech comprehension network modulates perceptual learning. Given its role during degraded speech comprehension, we hypothesize that facilitating left IFG during training enhances perceptual learning. This effect should be most prominent when severe degradation requires strong “top-down” predictions. For left IPC, the existing literature provides less concise predictions. Facilitation can be expected to modulate integration and decision processes during training, most likely increasing the tendency to converge auditory and written input. Such enhanced integration of multimodal percepts will be reflected in a shift in the criterion (C). Since learning requires optimization of C (i.e., reduction in potentially pre-existing response bias), the effect of IPC facilitation can be expected to depend on several factors. In case of a response bias before training, minimization of the latter should support learning, while induction or augmentation of a response bias should not support learning.

Materials and Methods

Participants

Thirty-six healthy volunteers, who were native speakers of German, without a history of hearing impairment or any other neurological or medical condition, participated in the study (mean age, 26.6 years; age range, 21–31 years; 18 females). None of the participants took any CNS-active medication during the experiments. Before participation, all subjects underwent a comprehensive neurological examination to screen for potential exclusion criteria. Participants who did not meet the protocol criteria and/or had contraindications for transcranial direct current stimulation (tDCS) were not included (Nitsche et al., 2008). According to the Oldfield questionnaire (Oldfield, 1971), all participants were strongly right handed. They gave written informed consent according to the Declaration of Helsinki and were financially compensated for participation according to the standard practice at the Institute. The Ethics Committee of the University of Leipzig approved of the study.

Stimuli and experimental conditions

We studied 200 minimal pairs (MPs) from a large corpus of monosyllabic and bisyllabic monomorphemic German nouns, which differed in the initial consonant and were phonological and graphematic neighbors without elision or addition of any phoneme or grapheme [e.g., Fisch–Tisch (English: fish–table)]. An equal number of corresponding identical pairs (IPs) of words were also constructed (e.g., Fisch–Fisch). These target stimuli were presented intermixed with 64 distractor stimuli, which were MPs that differed in the final consonant [e.g., Maus–Maul (English: mouse–mouth)] and their corresponding identical word pairs (e.g., Maus–Maus). These distractor stimuli did not enter the analysis, but

were introduced to prevent participants from selectively attending to the initial phoneme during the task. All stimuli were controlled for word frequency (according to the Wortschatz Lexikon of the University of Leipzig; <http://wortschatz.uni-leipzig.de>) and for the total number of competing MP neighbors differing in the initial phoneme [e.g., Mutter (English: mother) has three minimal competitors, which are Butter, Kutter, and Futter (English: butter, boat, fodder)]. Target stimuli were also controlled for phonetic distance of the initial phonemes of both words of an MP regarding the features voicing coding, place coding, and manner coding.

Stimuli spoken by a female trained speaker were recorded in a sound-attenuated room at a sampling rate of 44.1 kHz. Postediting included downsampling to 22.5 kHz; cutting at zero crossings before and after each spoken word, including a fade-in and fade-out of 2 ms; and root mean square normalization. To vary intelligibility, each word was noise vocoded at 2, 3, 4, and 6 bands. Bands were chosen according to a pilot study ($n = 16$, different but demographically similar volunteers), which indicated error scores to provide a sufficient dynamic range for learning. Noise vocoding preserves the temporal detail but can parametrically vary the spectral detail of the auditory signal (Shannon et al., 1995). We followed a procedure that has been previously outlined (Rosen et al., 1999) and often applied in neuroimaging and behavioral studies (Obleser et al., 2008; Erb et al., 2012); all spectral information between 0.07 and 9 kHz entered the vocoding routine and was divided into filter bands according to the Greenwood formula (Greenwood, 1990), yielding approximately logarithmic spacing. A 400 Hz envelope low-pass filter was applied to each band. Since two-band noise vocoding conveys the spectral information only in two bands, this is the least intelligible version of the stimuli, while six-band vocoded speech can usually be decoded rather well after a short training session (Dahan and Mead, 2010).

All target stimuli (200 MPs and 200 IPs) and distractor stimuli (64 MPs and 64 IPs) were allocated to different conditions: training to discriminate between word pairs of half of the target stimuli (100 MPs and 100 IPs) and half of the distractor stimuli (32 MPs and 32 IPs) took place over 3 consecutive days, whereas participants were not trained on the other half. Instead, these were only presented at pretesting and post-testing (henceforth, these stimuli are referred to as “trained/untrained”; see training procedure below). Both trained and untrained stimuli as well as the distractors were allocated to the four different noise-vocoding levels (2, 3, 4, and 6 bands), yielding a total amount of 25 MPs, 25 IPs, and 8 distractor MPs and IPs per experimental condition. Across participants the allocation of the specific stimuli to the eight experimental conditions (2, 3, 4, and 6 bands; trained and untrained stimuli) was changed, thus further attenuating potential item-specific differences (see Fig. 2 for a graphical display of the stimulus categories).

Experimental procedure and trial design

Our behavioral measures were changes in perceptual sensitivity (d') and C in a discrimination task between MPs and IPs. Each stimulus consisted of the auditory presentation of a word parametrically degraded by noise vocoding followed by the same (IP) or a different (MP) word presented in an undegraded written form. To compare item-specific learning (trained stimuli) and generalization (untrained stimuli), participants were trained on half of the stimuli over 3 days. Generalization (untrained stimuli) can be considered an operationalization of perceptual learning in the present task. In this article, “perceptual learning” is operationalized in that an increase in perceptual sensitivity and the resulting increase in “same-different” discrimination for untrained items is considered to signal generalization. Notably, this transfer is bound to impact at a sub-lexical level, while training intentionally relied on lexical items. We consider this transfer effect to allow one to probe perceptual learning for the case of degraded speech on the single-word level [for a more general discussion of perceptual learning, see Goldstone (1998)].

To investigate whether IFG or IPC facilitation modulates learning compared with sham stimulation, a between-subject (3 groups of 12 participants) was required, because our primary interest in perceptual learning excludes a within-subject design. To sum up, we used a mixed factorial design including the within-subject factors of stimulus type (IP/MP), auditory degradation (2, 3, 4, and 6 bands), training status (trained/

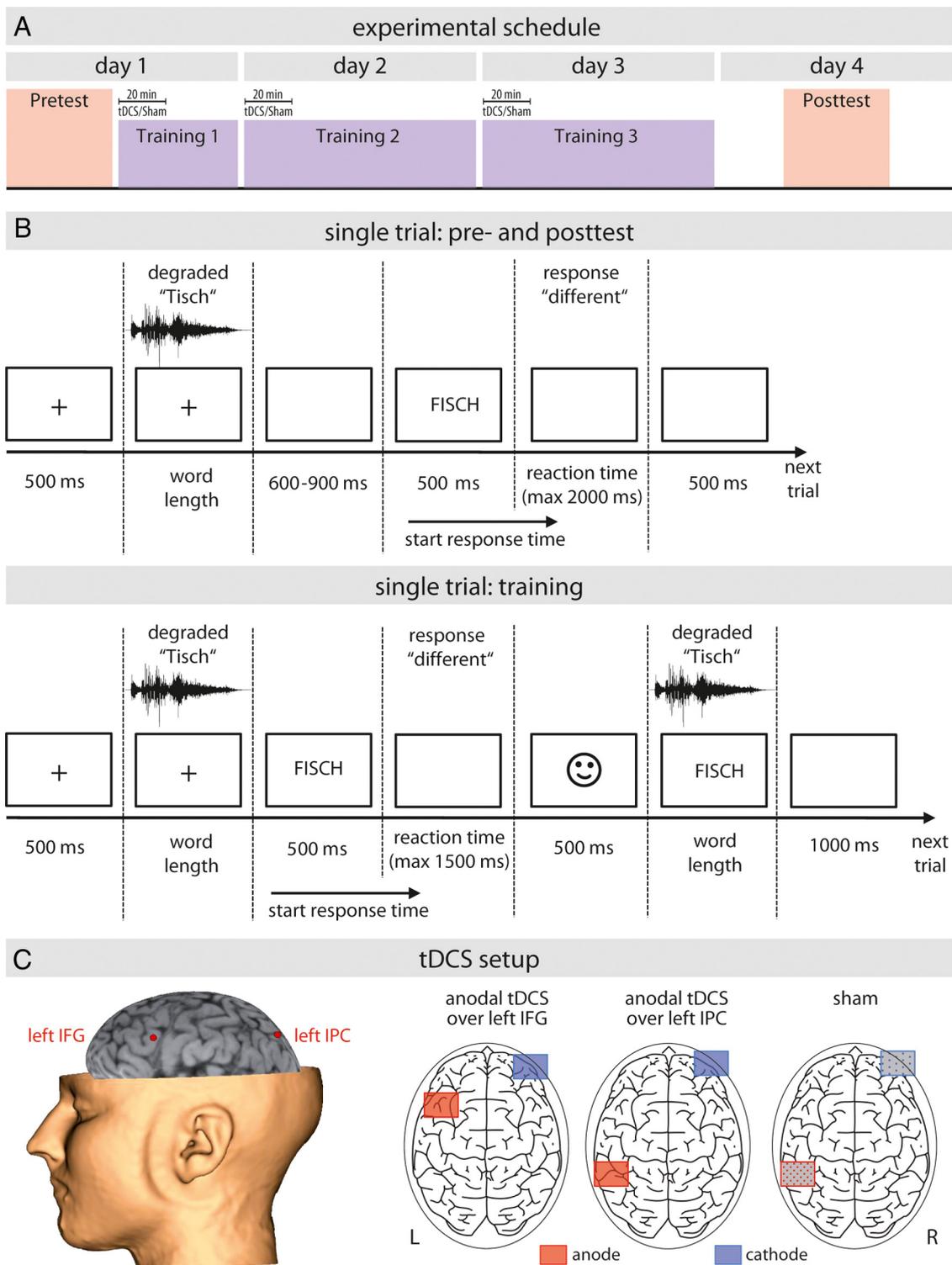


Figure 1. Experimental setup and design. **A**, Baseline perceptual sensitivity to discriminate between MPs and IPs was assessed on day 1, followed by a training over 3 days where participants were trained intensively on half of the stimuli ($n = 100$). During the first 20 min of each training session, either tDCS over IFG, IPC, or sham stimulation was applied. After training (day 4), perceptual sensitivity of all stimuli (trained and untrained) were tested again. **B**, Single trial design of the same-different forced-choice paradigm during both pretest and post-test (top row) and training (bottom row). **C**, For tDCS, neuronavigation was used to target brain areas (left IFG and IPC) based on the subjects' individual structural MRI scans.

untrained), time (pretraining/post-training), and the across-subjects factor stimulation type (IFG/IPC/sham stimulation). For the latter, participants were randomly allocated to three intervention groups that differed in the mode of brain stimulation during learning, as follows: anodal tDCS (tDCS_{anodal}) over stimulated left IFG (IFG^{STIM}); tDCS_{anodal} over stimulated left IPC (IPC^{STIM}), or sham stimulation (SHAM^{STIM}; see section on tDCS below).

The experiment proceeded over 4 consecutive days with a pretest on the first day, a post-test on the fourth day, and a total of three training sessions, one each on days 1, 2, and 3 (Fig. 1A). During both the training and pretest/post-test periods, a same-different forced-choice paradigm was used to train participants and to test their perceptual sensitivity of word discrimination. Stimulus presentation and behavioral response recording were controlled by Presentation software (version 14.7,

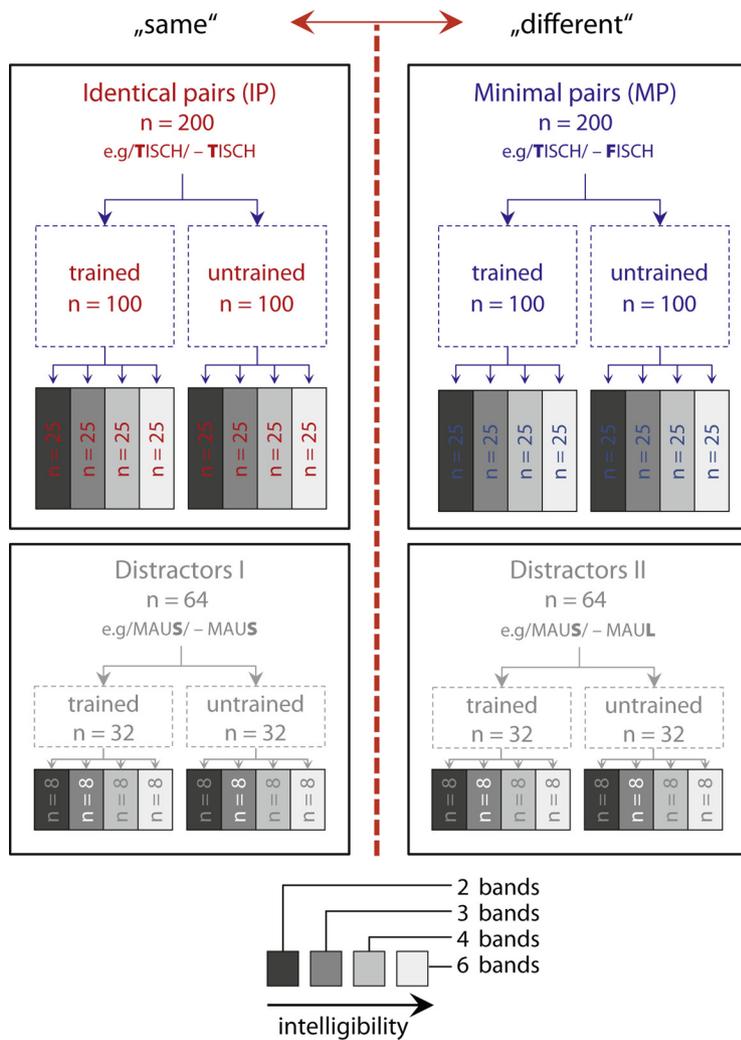


Figure 2. Illustration of the material used in the present study. This altogether included 200 MPs, 200 IPs, and 128 distractors. The latter were either of type I (identical) or type II (minimal pairs) and were not analyzed. They solely served as a goal to enforce judgments on the complete lexical item and prevent judgments based selectively on the first phoneme.

Neurobehavioral Systems). Participants were seated in front of the computer screen with the right index and middle finger placed on response buttons.

Pretest and post-test—trial design and stimuli. During the pretest, each participant’s baseline perceptual sensitivity to discriminate between MPs and IPs was assessed (e.g., MP /Fisch/ – TISCH = “different” vs IP /Fisch/ – FISCH = same; please note that throughout the manuscript “/Xxx/” indicates the acoustic, degraded presentation, while “XXX” indicates the written, undegraded presentation of the respective word). The first word in each word pair was presented acoustically, degraded by noise vocoding in one of the four different intelligibility levels (2, 3, 4, and 6 bands; see above). After a variable interstimulus interval of 600–900 ms, the second word was visually presented for 500 ms (capital letters, font size 72, black on white background). After the presentation of the second word, participants had to decide by button press whether both words were identical or different. They were encouraged to do so as fast as possible. After the button press or after time out (2000 ms, if the participant failed to respond; a missed trial), there was a break of 500 ms before the next trial started (Fig. 1B). Across participants, the allocation of the index or middle finger for the response of same or different was counterbalanced.

All target stimuli (200 MPs and 200 IPs) and distractor stimuli (64 MPs and 64 IPs) were tested during the pretest. To prevent order recognition effects, each auditorily presented word was presented a third time followed by the visual presentation of either the identical or the minimal

pair word (hence, a total of 300 MPs, 300 IPs, and 192 distractors in pseudorandomized order). This third presentation was not included in the data analysis. On day 4, a post-test assessed perceptual sensitivity for both trained and untrained stimuli. The procedure of the post-test was identical to the pretest, except that stimuli were presented in a different randomization.

Training—trial design and stimuli. The training proceeded over 3 days and included five training blocks (one block on day 1; two blocks each on days 2 and 3). As sketched above (Fig. 2), participants were trained to discriminate only half of the stimuli tested in the pretest, yielding a total amount of 25 MPs, 25 IPs, 8 distractor MPs, and IPs per degradation level. Each block comprised the threefold auditory presentation of each stimulus, once followed by the visual presentation of the identical word, once followed by the visual presentation of the minimal word pair, and once randomly followed by either of both to prevent order recognition effects. In analogy, 32 distractor MPs and 32 IPs (8 per degradation level) were also presented auditorily three times, resulting in a total of 300 training and 96 distractor trials per training block. To allow perceptual learning, two additional steps were added to the procedure otherwise identical to the pretesting: (1) after the response, feedback (correct/incorrect) was given by icons; and (2) thereafter the first word was simultaneously presented in its acoustically degraded and in the (undegraded) written forms (Fig. 1B). The latter procedure was chosen because perceptual learning of degraded speech has been shown to be most robust when acoustically degraded verbal material and its written transcription are presented simultaneously (Loebach et al., 2010). One training block (~30 min) followed the pretest on day 1, while two training blocks (~60 min) were performed on day 2 and day 3. On day 4 (post-test), there was no further training. To investigate its potential to induce facilitation of perceptual learning, tDCS or sham stimulation was applied during the first 20 min of each training session (for tDCS methods, see below).

Neuro-navigated transcranial direct current stimulation
T1-weighted high-resolution structural images of each participant were available to individually identify the target areas for tDCS. Target coordinates for left IFG and left IPC/AG were chosen according to a recent fMRI study on perceptual learning of degraded speech (Eisner et al., 2010): *x*, *y*, *z* MNI space, –46, 26, 20 for left IFG; –36, –58, 50 for left IPC (Fig. 1C). For the sham group (SHAM^{STIM}), one of each of the target areas was randomly chosen. A battery-driven DC stimulator delivered tDCS (Neuroconn GmbH) using a pair of electrodes in a 5 × 5 cm saline-soaked sponge. Before training, the electrodes were attached to the participant’s head using elastic bands. The anodal electrode was centered over the respective target coordinate, while the cathode was attached to the contralateral supraorbital region (Flöel et al., 2008). The application of tDCS was single blinded. For all experimental conditions (anodal tDCS over IFG, anodal tDCS over IPC, and SHAM), the current was increased in a ramp-like fashion over 30 s to a maximum of 1 mA eliciting a transient tingling sensation on the scalp. In the verum groups, tDCS was delivered for 20 min (IFG^{STIM}, IPC^{STIM}), but in the sham group stimulation was faded out after 30 s (SHAM^{STIM}). The current density at the stimulation electrodes amounts to 0.04 mA/cm² for our 1 mA anodal

tDCS, and the total charge (current density \times total stimulation duration, in seconds) was 0.048 C/cm^2 on each day for the verum conditions. Currents were turned off slowly after a few seconds precluding sensory differences between conditions (Nitsche et al., 2003), which has been shown to be efficient in the blinding of the procedure (Gandiga et al., 2006; Ragert et al., 2008).

Data analysis

The analysis of the same-different forced-choice discrimination task followed signal detection methodology rendering measures of d' and the underlying C . The former measure is independent of response bias, and the latter renders a numerical value that indicates response bias when different from zero. The analytic framework takes into account that individuals are not merely passive receivers but use an internal decision criterion for response selection. Stimulus degradation induces an increase in perceptual uncertainty rendering response bias more likely to occur. In the following, an increase of d' indicates an increase in perceptual sensitivity for the discrimination of MPs versus IPs. Successful learning thus increases d' . This measure is calculated by $d' = z(\text{hits}) - z(\text{false alarms})$.

Since d' is a measure independent of response bias, C renders additional information on the individual's response behavior. C was calculated using the following formula: $C = -0.5 * (z(\text{hits}) + z(\text{false alarms}))$ (McMillan, 2005). When $C = 0$, there is no response bias. In the present study $C > 0$ indicates a more "conservative" response behavior, meaning that the participant showed a tendency to respond differently. Conversely, $C < 0$ (more liberal) indicates the tendency to respond the same.

Four different statistical models were calculated separately for both d' and C . These models combined the following: one across-subject factor, STIM (3 groups, i.e., IFG^{STIM}, IPC^{STIM}, SHAM^{STIM}); four within-subject factors, DEGRAD (four levels, i.e., 2, 3, 4, 6-band vocoding); BLOCK (five levels, i.e., training blocks only available for trained items); TIME (two levels, i.e., pretest, post-test); and TRAIN (two levels, i.e., trained, untrained).

Model 1. At baseline, we assessed potential influences of degradation and/or tDCS group on d' and C in two separate 4×3 repeated-measures mixed-factorial ANOVAs (RM-ANOVAs ; DEGRAD \times STIM). These were performed on the pretest data.

Model 2. Changes in d' and C during training, were assessed by two separate $5 \times 4 \times 3$ RM-ANOVAs (BLOCK \times DEGRAD \times STIM). These obviously only included the trained items.

Model 3. The influence of the full training program on d' and C was assessed by two $2 \times 2 \times 4 \times 3$ RM-ANOVAs (TIME \times TRAIN \times DEGRAD \times STIM).

Model 4. The influence of tDCS on perceptual learning (i.e., the change in perceptual sensitivity for untrained stimuli) was assessed by its influence on changes in d' ($\Delta d' = d'_{\text{post-test}} - d'_{\text{pretest}}$; Zaehle et al., 2011). Likewise, changes in C were analyzed to assess potential changes in response behavior associated with perceptual learning ($\Delta C = C_{\text{post-test}} - C_{\text{pretest}}$). Both parameters were tested by using two separate 4×3 RM-ANOVAs (DEGRAD \times STIM). In a second step, univariate between-subject ANOVAs for the factor STIM were calculated separately for each degradation level.

The level of statistical significance was set at a threshold of $p < 0.05$. When ANOVAs yielded significant main effects, *post hoc* tests were calculated according to Fisher's least significant difference.

Results

Our study targets the improvement in perceptual sensitivity to discriminate between MPs and IPs under acoustically degraded conditions. The measure to assess changes in perceptual sensitivity was d' in a forced-choice paradigm. Here, our main interest was the effect of tDCS on changes in d' (i.e., pretest vs post-test) for the untrained items, because this transfer of perceptual sensitivity from trained to untrained stimuli can be considered an operationalization of perceptual learning in the present study. Additionally, we looked into the changes in C , potentially disclos-

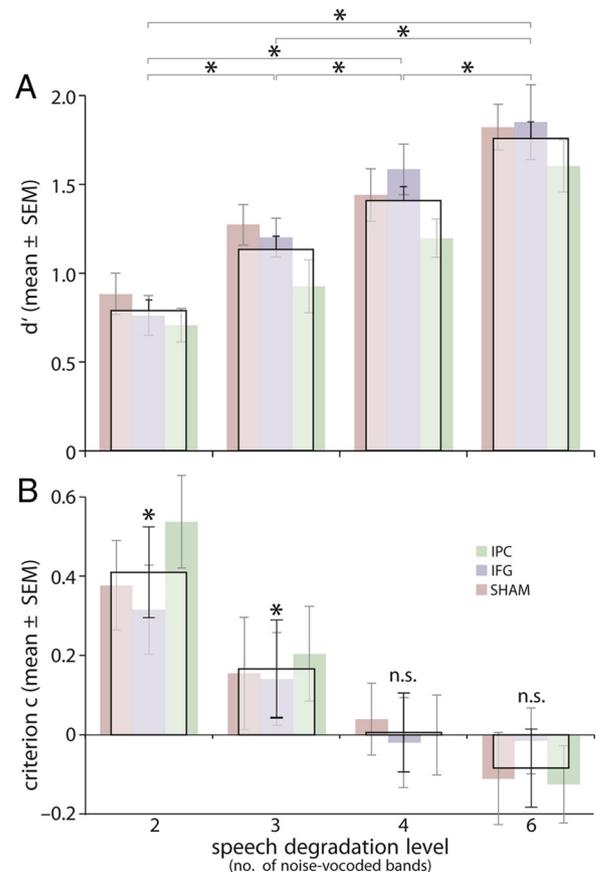


Figure 3. Behavioral data at baseline (pretest). **A**, d' decreased as a function of signal degradation (number of bands) in a roughly linear fashion. **B**, The amount of degradation influenced the participant's criterion, C : severely degraded speech (2 and 3 bands) led to a positive criterion; for example, subjects showed a response bias to judge pairs differently for severely degraded speech, but not for lesser degraded speech (4 and 6 bands). Wide columns, Mean of all participants ($n = 36$); narrow columns, mean per stimulation group ($n = 12$ each). Error bars indicate \pm SE.

ing influences on learning and influences of tDCS on response bias during stimulation.

d' and C

As described in the Materials and Methods section, four models were tested with regard to d' and C : (1) the influence of degradation level and stimulation group at baseline; (2) the effect of training for IP, and MP respecting the stimulation group (essentially this only refers to the trained items); (3) changes in both parameters from baseline to post-test after the full training program regarding differential effects on trained and untrained items; and (4) the effect of the three different tDCS stimulation conditions on the improvement of perceptual sensitivity after training (determined by $\Delta d' = d'_{\text{post-test}} - d'_{\text{pretest}}$), and changes in response criterion ($\Delta C = C_{\text{post-test}} - C_{\text{pretest}}$).

Influence of degradation on d' and C at baseline

d'
The parametric degradation of the acoustically presented words significantly reduced perceptual sensitivity (Fig. 3A). This steady increase in d' with intelligibility was statistically confirmed by a main effect of DEGRAD ($F_{(3,99)} = 56.6, p < 0.001$; *post hoc* tests between all vocoding levels at $p < 0.015$). The factor STIM did not influence perceptual sensitivity at baseline, precluding a pre-training bias due to uncontrolled group differences (main effect

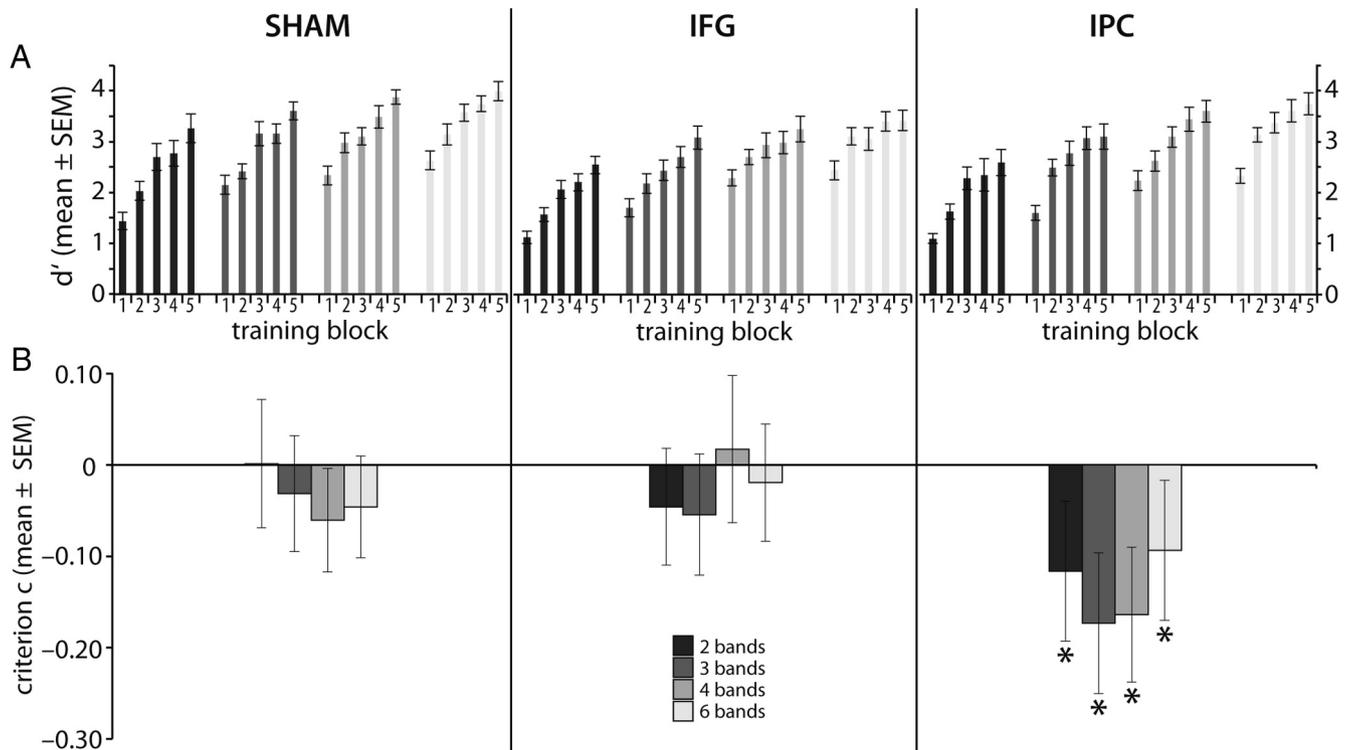


Figure 4. Behavioral data of all subjects during the training. **A**, Graph shows improvements in d' during five blocks of training (days 1–3) in the three different tDCS groups. More strongly degraded stimuli showed a steeper learning curve compared with the less degraded stimuli, as supported by a significant interaction of signal degradation and training block. **B**, Application of tDCS during training led to a modulation of C toward a more liberal response tendency in the IPC group only. No response bias (e.g., significant difference of C from zero) was found in the IFG or sham group. Error bars indicate \pm SE.

for STIM: $F_{(2,33)} = 0.6, p = 0.9$; interaction of DEGRAD \times STIM: $F_{(6,99)} = 0.71, p = 0.64$). In detail, baseline of d' for all stimulation groups and degradation levels were as follows (mean \pm SEM): SHAM^{STIM}: 2 bands, $d' = 0.89 \pm 0.12$; 3 bands, $d' = 1.27 \pm 0.11$; 4 bands, $d' = 1.44 \pm 0.15$; 6 bands, $d' = 1.82 \pm 0.13$; IFG^{STIM}: 2 bands, $d' = 0.77 \pm 0.11$; 3 bands, $d' = 1.2 \pm 0.11$; 4 bands, $d' = 1.58 \pm 0.14$; 6 bands, $d' = 1.85 \pm 0.21$; IPC^{STIM}: 2 bands, $d' = 0.71 \pm 0.1$; 3 bands, $d' = 0.93 \pm 0.15$; 4 bands, $d' = 1.2 \pm 0.11$; 6 bands, $d' = 1.6 \pm 0.15$.

C

Higher degradation levels led to more conservative values for C (Fig. 3B), indicating that participants showed a bias to judge word pairs differently for severely degraded speech, not present for lesser degradation levels as follows (mean \pm SEM): SHAM^{STIM}: 2 bands, $C = 0.38 \pm 0.11$; 3 bands, $C = 0.15 \pm 0.14$; 4 bands, $C = 0.04 \pm 0.09$; 6 bands, $C = -0.11 \pm 0.12$; IFG^{STIM}: 2 bands, $C = 0.32 \pm 0.11$; 3 bands, $C = 0.14 \pm 0.11$; 4 bands, $C = -0.02 \pm 0.11$; 6 bands, $C = -0.02 \pm 0.08$; IPC^{STIM}: 2 bands, $C = 0.54 \pm 0.12$; 3 bands, $C = 0.20 \pm 0.12$; 4 bands, $C = -0.00 \pm 0.10$; 6 bands, $C = -0.12 \pm 0.1$. Statistics confirmed a main effect of DEGRAD ($F_{(3,99)} = 55.75, p < 0.001$) and confirmed a significant response bias only for the two severe degradation levels (2 bands, $p < 0.001$; 3 bands, $p = 0.03$, 4 bands, $p = 0.92$; 6 bands, $p = 0.15$). At baseline no influence of tDCS group on C was found (factor STIM: $F_{(2,33)} = 0.6, p = 0.9$; interaction STIM \times DEGRAD: $F_{(6,99)} = 2.13, p = 0.06$).

To sum up, before training, the three stimulation groups showed no differences with regard to d' and C . The within-subject factor speech degradation level (DEGRAD) showed the

expected increase in d' with intelligibility and a bias to judge pairs as different for severe degradation.

Influence of training and tDCS on d' and C

d' In the course of the training, perceptual sensitivity as indexed by d' improved for the trained items (Fig. 4A). Visual inspection suggests steeper learning curves for the more strongly degraded items when compared with the less degraded items. The increase in d' levels off toward the later training blocks only for the least degraded stimuli (6 bands). These findings were confirmed by the multifactorial RM-ANOVA, showing main effects for BLOCK ($F_{(132,4)} = 132.08, p < 0.001$) and DEGRAD ($F_{(99,3)} = 89.11, p < 0.001$), and for the interaction BLOCK \times DEGRAD ($F_{(396,12)} = 1.99, p < 0.024$). *Post hoc* testing of consecutive blocks revealed the following relations in magnitude: 2 band, block1 $< 2 < 3$, and block 4 < 5 (all $p < 0.045$); 3 band, block 1 $< 2 < 3$ (all $p < 0.007$); 4 band, block1 $< 2 < 3$ (all $p < 0.042$); and 6 band, block 1 $< 2 < 0.001$). All other comparisons were not significant ($p > 0.05$). Importantly, we did not find a significant main effect of STIM ($F_{(33,2)} = 2.093, p < 0.14$) or significant interactions with STIM (BLOCK \times STIM, $F_{(132,8)} = 1.47, p = 1.74$; BLOCK \times STIM \times DEGRAD, $F_{(396,24)} = 0.92, p < 0.574$) during training.

C

During training, tDCS over the IPC yielded a response bias (Fig. 4B), while C was not significantly different from 0 (i.e., no response bias) for IFG or sham stimulation. IPC stimulation led to a decrease in C for all degradation levels, with the response bias shifting toward the judgment same as indicated by the negative,

that is, more liberal, C . A multifactorial RM -ANOVA revealed a strong trend for a significant main effect of STIM (IFG^{S-TIM} , IPC^{S-TIM} , $SHAM^{S-TIM}$; $F_{(33,1)} = 3.26$, $p = 0.051$). *Post hoc* testing revealed differences in criterion between IPC^{S-TIM} versus $SHAM^{S-TIM}$ ($p = 0.042$) and IPC^{S-TIM} versus IFG^{S-TIM} ($p = 0.028$), respectively. No difference in criterion was found between IFG^{S-TIM} and $SHAM^{S-TIM}$ ($p = 0.862$). Furthermore, a significant interaction of the factors BLOCK and DEGRAD was found ($F_{(280,56,12)} = 2.53$, $p = 0.01$; Greenhouse–Geisser corrected). This interaction was driven by the influence of the factor DEGRAD. No significant main effect was found for DEGRAD ($F_{(99,3)} = 0.7$; $p = 0.55$) and BLOCK ($F_{(132,4)} = 0.7$, $p = 0.59$), or the interactions DEGRAD \times STIM ($F_{(99,6)} = 1.16$; $p = 0.34$), BLOCK \times STIM ($F_{(132,8)} = 1.11$; $p = 0.36$), or DEGRAD \times BLOCK \times STIM ($F_{(396,24)} = 0.91$; $p = 0.6$).

In summary, perceptual sensitivity for the trained items increased during the training. For less degraded stimuli, this increase leveled off for the later training blocks. Regarding the criterion, a response bias was induced only by IPC stimulation during training, which showed no changes over the course of the training.

Item-specific learning versus generalization

Changes in d' from pretest to post-test

As expected, and according to the results in Model 2, the comparison between pretest and post-test confirms the improvement for the trained items (Fig. 5A). More relevant to our research question, item-specific learning elicited generalization because d' also increased for the untrained items (Fig. 5B). Comparison between the different degradation levels indicates better performance for lesser degradation (as shown above) but also suggests an interaction with training: there is a parallel increase on all degradation levels for the trained items, whereas learning curves for the untrained items fan out as a function of degradation level. Stronger degradation leads to a lesser generalization. For the 2-band degradation, no increase in d' is seen for the untrained items, suggesting no perceptual learning for this degradation level. These findings were confirmed by the results of the RM -ANOVA yielding significant main effects for TIME ($F_{(33,1)} = 202.9$, $p < 0.001$), TRAIN ($F_{(33,1)} = 119.47$, $p < 0.001$), and DEGRAD ($F_{(99,3)} = 129.0$, $p < 0.001$). The main effect STIM did not reach significance ($F_{(33,2)} = 3.23$; $p = 0.052$). Regarding interactions involving the factor of interest (TIME), we observed significant interactions of TIME \times DEGRAD ($F_{(99,3)} = 4.56$, $p < 0.005$) and TIME \times TRAIN \times DEGRAD ($F_{(99,3)} = 3.46$, $p < 0.02$), which confirmed differential learning-related changes for trained versus untrained items at different degradation levels. *Post hoc* testing revealed significant increases in d' from pretest to post-test on all degradation levels (2, 3, 4, and 6 bands) for the trained stimuli ($p < 0.001$), and for 3, 4, and 6-bands degradation condition for the untrained stimuli ($p < 0.01$). No difference from pretest to post-test was found for the 2-band condition in the untrained stimuli (not significant at $p > 0.05$).

The other significant interactions did not involve the factor TIME (TRAIN \times DEGRAD, $F_{(99,3)} = 5.4$, $p < 0.002$; and

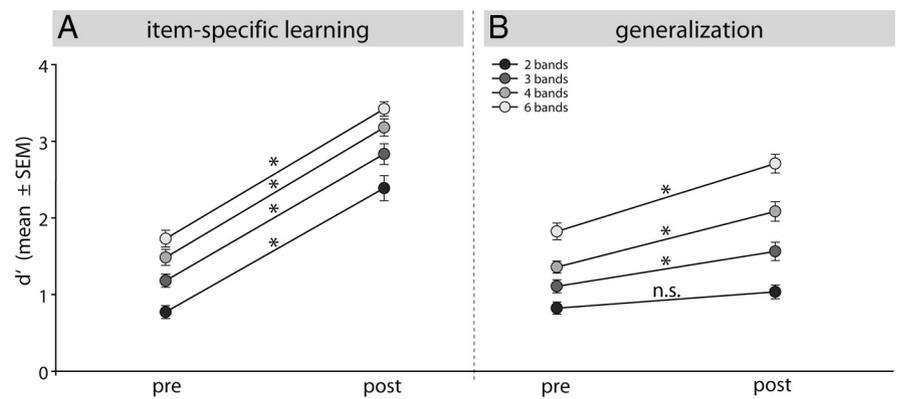


Figure 5. Perceptual sensitivity of all subjects ($n = 36$) before and after training (pretest and post-test). **A**, Graph shows an improvement in perceptual sensitivity of trained stimuli in all degradation levels from pretest to post-test, indicating item-specific learning. **B**, Improvement in untrained stimuli, i.e., generalization, revealed perceptual learning for the lesser degraded conditions (3, 4, and 6 bands). No perceptual learning was observed for the most severely degraded stimuli (2 bands). Perceptual and item-specific learning show differential learning patterns as a function of degradation: while equal discrimination improvements were observed across degradation levels for item-specific learning, during perceptual learning the slopes fan out as a function of degradation level, i.e. stronger degradation leads to a lesser generalization. Error bars indicate \pm SE. Asterisks indicate significant differences ($p < 0.05$).

TRAIN \times DEGRAD \times STIM, $F_{(99,6)} = 2.54$, $p < 0.03$). The interactions involving the factors TIME and STIM were not significant, as follows: TIME \times STIM, $F_{(29,2)} = 1.05$, $p = 0.362$; TIME \times STIM \times TRAIN, $F_{(33,2)} = 1.56$, $p = 0.226$; TIME \times STIM \times DEGRAD, $F_{(99,6)} = 0.79$, $p = 0.583$.

Changes in C from pretest to post-test

The analysis of C confirms the above significant influence of degradation on C as indexed by the main effect DEGRAD ($F_{(99,3)} = 11.34$, $p < 0.001$). Furthermore, we observed significant interactions of TIME \times DEGRAD ($F_{(99,3)} = 8.3$, $p < 0.001$) and TIME \times DEGRAD \times TRAIN ($F_{(99,3)} = 3.21$, $p < 0.026$). *Post hoc* tests comparing C from pretest to post-test revealed significant differences in 2-band condition for both trained ($p < 0.002$) and untrained stimuli ($p < 0.008$) and in the 6-band condition for the untrained stimuli only ($p < 0.023$). In detail, at baseline for 2-band vocoding, trained and untrained stimuli showed a response bias significantly different from zero as follows (mean \pm SEM): trained, 0.43 ± 0.08 , $p < 0.001$; untrained, 0.39 ± 0.06 , $p < 0.001$. At post-testing, this response bias decreased for the untrained stimuli (0.19 ± 0.04 , $p < 0.001$) and was abolished for the trained stimuli (trained -0.01 ± 0.11 , n.s.). For 6-band vocoding, no significant response bias, that is, C different from zero, was detected at pretesting (trained, -0.07 ± 0.07 , n.s.; untrained, -0.1 ± 0.05 , n.s.) and post-testing (trained, -0.19 ± 0.17 , n.s.; untrained, -0.06 ± 0.04 , n.s.). No influence was found for between-subject factor STIM ($F_{(33,2)} = 0.35$, $p = 0.71$) or the interactions involving the factors STIM and TIME: TIME \times STIM, $F_{(33,2)} = 0.59$, $p = 0.556$; TIME \times STIM \times TRAIN, $F_{(33,2)} = 0.32$, $p = 0.727$; TIME \times STIM \times DEGRAD, $F_{(99,6)} = 0.77$, $p = 0.592$.

To sum up, training reduced (untrained stimuli) or abolished (trained stimuli) a pre-existing response bias to judge stimuli as different for the most severely degraded condition. No change in response bias was seen for 3- and 4-band vocoding. The increase in C for the 6-band condition was purely numerical, since neither pretesting nor post-testing showed a significant response bias ($C \approx 0$ at all times).

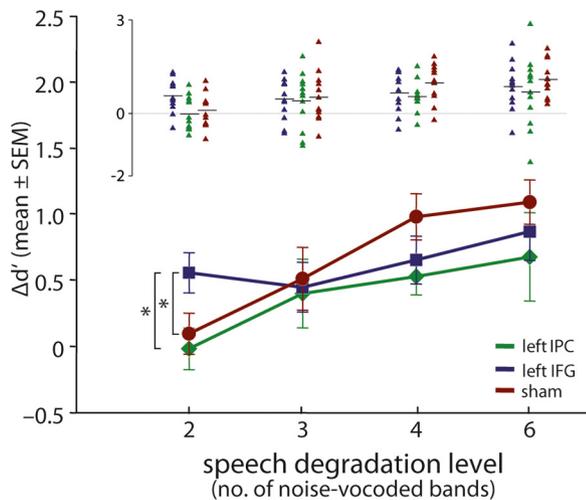


Figure 6. Effect of tDCS on perceptual learning ($\Delta d'$ of untrained items; lower graph displaying group averages, upper graph displaying single subject data). tDCS over IFG ($n = 12$) induced perceptual learning in the condition with the most severely degraded items (2 bands), while here no perceptual learning was observed for tDCS over IPC or SHAM ($n = 12$ each). No significant differences between groups were observed in the other degradation levels (3, 4, and 6 bands). Error bars indicate \pm SE. Asterisks indicate significant differences ($p < 0.05$).

Effect of tDCS on perceptual learning and related changes in response behavior

$\Delta d'$

The omnibus ANOVA (DEGRAD \times STIM) of $\Delta d'$ revealed a significant influence of degradation on perceptual learning as indexed by the main effect DEGRAD ($F_{(99,3)} = 8.73$, $p < 0.001$). In this statistical model, the factor STIM or the interaction STIM \times DEGRAD did not reach significance ($p > 0.05$). Since we expected the effect of top-down influences to be more relevant in severely degraded speech, we additionally tested the four degradation levels separately by using univariate between-subject ANOVAs (factor STIM) for each degradation level. tDCS over IFG led to an increase in perceptual learning of degraded speech only in the condition with severely degraded speech (2 bands, $\Delta d' = 0.6 \pm 0.52$; Fig. 6). At this degradation level, no perceptual learning was seen for IPC^{STIM} or SHAM^{STIM} (mean \pm SD: IPC-^{STIM}, 0.0 ± 0.54 ; SHAM^{STIM}, 0.1 ± 0.53 ; $F_{(33,2)} = 3.88$; $p < 0.031$; $\eta^2 = 0.19$; *post hoc t*-tests: IFG vs sham, $p < 0.043$; IFG vs IPC, $p < 0.013$; sham vs IPC, n.s., $p = 0.605$), while the differences for the other degradation levels (3, 4, and 6 bands) did not reach statistical significance (3 bands, $F_{(33,2)} = 0.06$, $p = 0.94$; 4 bands, $F_{(33,2)} = 1.974$, $p = 0.155$; 6 bands, $F_{(33,2)} = 0.69$; $p = 0.51$). No stimulation group effects were seen for the trained stimuli (2 bands, $F_{(33,2)} = 1.82$, $p = 0.177$; 3 bands, ($F_{(33,2)} = 0.26$; $p = 0.78$; 4 bands, $F_{(33,2)} = 1.484$, $p = 0.242$; 6 bands, $F_{(33,2)} = 0.64$; $p = 0.541$).

ΔC

The omnibus ANOVA (DEGRAD \times STIM) of ΔC revealed a significant influence of degradation on changes in response behavior for perceptual learning as indexed by the main effect DEGRAD ($F_{(99,3)} = 10.28$, $p < 0.001$). The factor STIM or the interaction STIM \times DEGRAD did not reach significance ($p > 0.05$). Likewise, the analyses using univariate ANOVAs performed separately for each degradation level did not show an effect of STIM on ΔC ($p \geq 0.19$ for all comparisons).

Discussion

Using a sham-controlled design, our results demonstrate that facilitatory tDCS_{anodal} impacts on perceptual learning when applied during training over two key areas involved in the top-down processing of noise-vocoded speech. Facilitation of the left IFG induced a transfer from trained to untrained stimuli for the increase in d' . This applied only for the most severely degraded speech. Such transfer can be considered evidence for perceptual learning. Notably, no perceptual learning occurred when IPC was facilitated by tDCS_{anodal}. On the contrary, left IPC facilitation induced a change in the decision criterion (C), in that participants showed a robust bias to judge stimuli as same during training. Additionally, the response bias ($C > 0$) before training for the most severely degraded stimuli was attenuated for the untrained, and abolished for the trained stimuli after the training.

The major findings can be conceived in the framework of signal detection theory. As sketched in Figure 7, two effects must be considered. (1) Forced choice between two categories (i.e., same vs different) will improve when stimuli (i.e., IP vs MP) elicit less overlapping neuronal response patterns (Fig. 7A). Our data support the hypothesis that for severely degraded speech IFG supports such a formation of more distinct response patterns during training, as evidenced by an increase in d' for untrained stimuli. We suggest that in our paradigm a linguistic feature (here phonetic categories) is “sharpened” by training. (2) More general cognitive control processes such as C impact on forced choice in a more complex way. If no response bias is present at baseline (i.e., $C \approx 0$), the induction of a response bias will not enhance performance. Only if an intervention reduces a pre-existing response bias ($C \neq 0$) will this increase performance in a balanced forced-choice task. The latter effect is seen for more severely degraded speech in the comparison between pretesting and post-testing. On the contrary, the induction of a response bias during training by IPC facilitation does not increase perceptual sensitivity for the untrained items (Fig. 7B).

Clearly, for perceptual learning during natural connected speech perception, these processes interact. However, our design, selectively providing lexical information without any syntactic, pragmatic, or contextual cues, allows us to discuss more linguistic and more general cognitive aspects during perceptual learning of vocoded speech.

Linguistic and general cognitive top-down processes

During speech perception, linguistic competence is incrementally recruited when the auditory input is degraded. At the same time, decision bias is likely to occur under challenging perceptual conditions, leading to the ambiguity of a stimulus (Vickery and Jiang, 2009). Using facilitatory anodal tDCS, we here supply experimental evidence that left IFG and IPC are important hubs differentially impacting on these two processes pertaining to perceptual learning of degraded speech. This noninvasive brain stimulation technique has been previously shown to modulate language functions and language learning processes (Flöel et al., 2008; Sparing et al., 2008; de Vries et al., 2010; Cattaneo et al., 2011). In line with our hypothesis anodal tDCS over left IFG supports perceptual learning of degraded speech in a same-different forced-choice paradigm. Consistent with the recruitment of linguistic knowledge for adverse listening environments, left IFG indeed showed a facilitatory effect for perceptual learning of highly degraded speech: Only when IFG was facilitated during training did perceptual learning occur for the most severely degraded stimuli. Note that these stimuli provide almost no spectral information to allow for the phonemic discrimination required

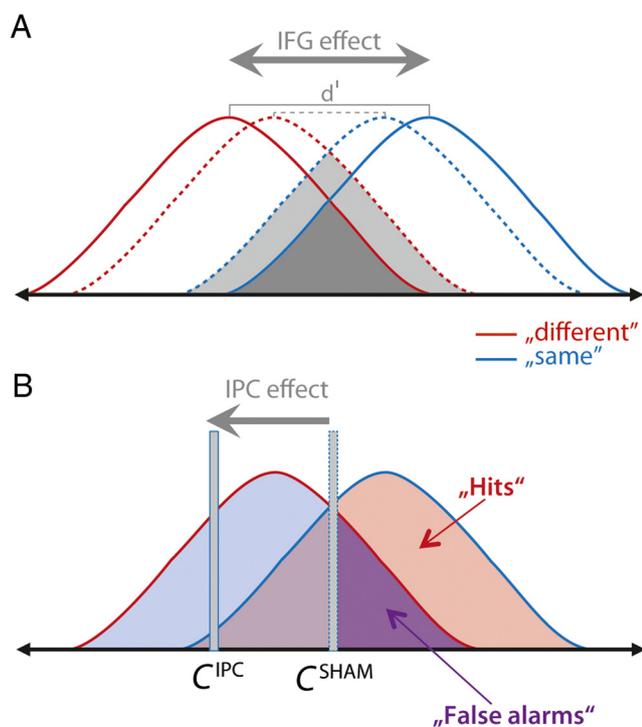


Figure 7. Sketch illustrating roles of IFG and IPC during perceptual learning of degraded speech. **A**, Conceptualization of the effect of tDCS over IFG for the most severely degraded items. Before training (dashed curves), the neuronal response to the auditory stimuli is noisy, as illustrated by broad overlapping response distributions and a comparatively small d' (dashed gray line). Induced by training under IFG stimulation, the response distributions drift apart (solid lines), which signals enhanced categorical phoneme discrimination. Note that for the 2-band-vocoded words this effect relies on IFG facilitation. **B** sketches the effect of a change in the internal C as seen only during training with IPC facilitation for all vocoding levels. IPC facilitation induced liberalization of the criterion, yielding a near-optimal hit rate for identical pairs, at the expense of a grossly enhanced false alarm rate. The bias to integrate the acoustically degraded and the undegraded written word is a bias toward a unified percept. Such a bias may enhance perceptual learning when a larger-scale linguistic context (semantic, syntactic, or pragmatic) is supplied, but may not enhance the phonetic nucleus of perceptual learning of degraded speech.

by the task, and, accordingly, participants showed no perceptual learning for these items under natural, that is, non-tDCS, conditions. On the contrary, IPC facilitation did not affect perceptual learning but instead modulated participants' decision criterion during training, eliciting a robust liberalization of the response bias toward judging pairs as identical.

On a theoretical account, our data support the notion that integration of linguistic knowledge becomes essential especially when the auditory input is severely degraded (Obleser et al., 2007). The data provide the first evidence that top-down influences can be efficiently augmented by facilitatory tDCS to allow for learning even when learning is absent under "natural," that is, non-tDCS learning conditions. Since noise vocoding simulates the degraded auditory input supplied by cochlear implants (Shannon et al., 1995), our findings are encouraging to further explore the potential clinical benefits of noninvasive brain stimulation in hearing impairment and restorative hearing therapy.

Operationalization of perceptual learning

Our results are based on an operationalization because transfer of the increase in d' from trained to untrained items is thought to signal perceptual learning. Principally, perceptual adaptation and learning processes ensure the relatively stable processing of sensory information despite a highly variable and/or distorted signal.

Depending on the experimental operationalization, the generalization of a previously established differential response to a new stimulus can be a mandatory feature for perceptual learning (Wohlwill, 1958; Goldstone, 1998; Loebach et al., 2009). In our study, participants were intensively trained on a same-different forced-choice paradigm to discriminate between MPs and IPs. Importantly, after the training on day 4, the generalization to untrained items was tested. Because improvements for the untrained stimuli indicate perceptual learning, our study design allows disentangling of perceptual learning from item-specific stimulus–response associations.

It should be noted that our design targets perceptual learning, which has been differentiated from within-session adaptation, in that multiple sessions and overnight consolidation may be required to stabilize the augmented perceptual skill. Both perceptual adaptation and learning have been suggested to require an interaction of primary sensory with higher-order, more cognitive processes in different brain areas (Censor et al., 2012).

Implications for the neuronal network affording perceptual learning of degraded speech

With regard to the adaptation to degraded speech, a recent fMRI study provides support for the differential role of the two higher-order hubs of the network (IFG and IPC) to interact with bottom-up processes (mostly taxing STG/STS). Participants were trained on moderately degraded 8-band-vocoded (and frequency-shifted) sentences in a comparatively short single training session in the scanner (Eisner et al., 2010). BOLD contrast changes over the course of training disclosed that vocoded speech, when compared with nonlearnable material, elicited IFG and STS activations. However, only left IFG activation correlated with the learning success, suggesting a pivotal role of the IFG for perceptual learning. Interestingly, inferior parietal regions [AG and supramarginal gyrus (SMG)] correlated with the individual progress over the course of the training. Based on a functional connectivity analysis, the authors proposed that IFG is the key structure linking bottom-up (STS/STG) and multimodal inferior parietal integration areas (AG and SMG) during perceptual learning.

Advancing from such correlational imaging findings, our study provides (1) novel evidence for a causal relationship between IFG facilitation and perceptual learning of severely degraded speech, and (2) highlights novel aspects with regard to the interplay between IFG and IPC, affording the modulation of phonemic categorical decisions during the learning process, as follows.

First, we show that learning is not only enhanced, but also enabled by tDCS in the case of severely degraded, highly unintelligible speech. No learning took place for 2-band degraded speech in the absence of IFG tDCS. To our knowledge, this is the first demonstration of a cognitive process that depends on external stimulation. How might tDCS enable perceptual learning? Anodal stimulation causes a tonic depolarization of the neuronal membrane potential, increasing the spontaneous firing rate and the excitability of cortical neurons in the stimulated area. Importantly, anodal stimulation acts upon a modulation of the synaptic strength and thus resembles long-term potentiation-like mechanisms (Stagg and Nitsche, 2011). We suggest that tDCS-enhanced plasticity in the left IFG allowed for the modulation and sharpening of phonemic categories, even when the auditory input supplied minimally discernible features (Fig. 7A). This effect is constitutive for perceptual learning under severely de-

graded conditions where stronger contributions of linguistic knowledge are required.

Second, previous studies on the neural correlates of perceptual learning of degraded speech used training paradigms providing a high level of semantic, syntactic, and even pragmatic context. When subjects are trained on sentences (Eisner et al., 2010; Erb et al., 2012) or in a natural listening environment using a portable real-time cochlear-implant simulator (Smalt et al., 2011), IFG involvement cannot be attributed to a certain level of linguistic processing. In contrast, our task targets phoneme discrimination and minimizes further contextual (e.g., semantic, syntactic, or pragmatic) information. Since our data show that IFG stimulation enables perceptual learning of highly degraded single words, we propose that in this context left IFG affords the critical formation and adaptation of phonological categories. Our findings support a more linguistic role of the IFG, while the more general cognitive processes of perceptual integration and decision bias may rely on the highly multimodal, parietal parts of the network. Such a view converges with previous studies showing an involvement of the left IFG in different aspects of phonological processing (Poldrack et al., 1999; Nixon et al., 2004; Tremblay et al., 2004; Hartwigsen et al., 2010). Future studies should target the question of the exact locus of facilitation from a neurolinguistic perspective, since the importance of lexical information of the training items has been controversially discussed in the light of behavioral findings (Davis et al., 2005; Hervais-Adelman et al., 2008). Based on our present findings, we propose that IFG is the key area to readjust sublexical, more specifically phonemic, categories during learning (Fig. 7A). In turn, this categorical readjustment modulates intelligibility (Obleser and Eisner, 2009).

It is important to note that tDCS does not selectively increase local cortical activity but rather modulates neuronal networks to change functional connectivity between both local and interconnected brain areas (Sehm et al., 2012). We assume that the behavioral effect does not exclusively rely on a modulation of the left IFG, but is supported by the modulation of a fronto-temporo-parietal network, with the left IFG as the key node. This assumption needs to be put under scrutiny by investigating approaches that combine facilitatory brain stimulation and functional connectivity imaging.

Beyond IFG, left IPC has been considered a key node within the network affording perception and/or perceptual learning of degraded speech (Shahin et al., 2009; Eisner et al., 2010; Obleser and Kotz, 2010; Sharp et al., 2010a; Clos et al., 2012). We do not find an IPC facilitation in perceptual learning, which is most likely due to the use of single words and a phonological decision task, which neither requires nor allows higher-order semantic or sentential integration. We suggest that the IPC comes into play when a higher-order semantic context is provided, or when the temporo-frontal loop is compromised by disease (Sharp et al., 2010b). In the present data, IPC stimulation had a more subtle, indirect effect on performance and modulated the decision criterion during learning. Only when IPC was facilitated by anodal tDCS, did participants show a bias to identify identical pairs (Fig. 7B). The results raise the possibility that facilitation of the IPC leads to an increased integration of the degraded acoustic word and the undegraded written word, resulting in a response bias toward identical pairs. On a more general level, our data support the putatively supramodal role of the IPC during discrimination tasks, where this area is a hub of a network involved in perceptual decision making (Kühn et al., 2011).

Potential limitations

A principal limitation of studies targeting perceptual learning is the necessity to test interventions between groups. Since perceptual sensitivity, training efficiency, and tDCS effects will vary between individuals, the resulting variance may elicit substantial type II errors. In that vein, interactions involving the factor STIM in the critical comparisons failed to reach significance (e.g., STIM \times TIME). With regard to the network affording processing of degraded speech, one key area (STS/STG) was not targeted in the present study. Here, the major limitation is that tDCS cannot selectively target these areas without affecting primary auditory processing. Hence, a different training schedule including tests for primary auditory processing would be required, which we considered to overly complicate the already complex design of the study. Another technical issue regarding tDCS is bipolar electrode arrangement. Based on our experimental design, we cannot entirely rule out an effect of the cathodal “reference” electrode over the right supraorbital region. We consider it unlikely, however, that prefrontal inhibitory tDCS can explain our findings because (1) we find differential effects for anodal stimulation of the IFG and IPC, though the reference electrode position was kept the same in all tDCS conditions; (2) inhibitory effects by cathodal stimulation in cognitive and language tasks must be considered small (Jacobson et al., 2012), especially at the intensity of 1 mA used here; and (3) a modulatory effect on perceptual learning by potential inhibition of right supraorbital region would be a very unexpected finding given the putative key areas in this task, as discussed in previous work (Eisner et al., 2010).

Summary

To sum up, our results allow us to disentangle the contributions of two key structures of the speech perception network during perceptual learning of degraded speech: IFG is the crucial brain area to adjust and sharpen the mapping of degraded input to phonemic categories. Upregulation in IFG allows perceptual learning even under most severely reduced auditory bottom-up information. On the contrary, IPC plays a modulatory role in the same-different discrimination, as indicated by a clear influence on the decision criterion during learning. Such an influence on the criterion may be relevant in perceptual learning, but may only influence learning success when the categories to be differentiated have grossly different occurrence probabilities.

References

- Bent T, Loeback JL, Phillips L, Pisoni DB (2011) Perceptual adaptation to sinewave-vocoded speech across languages. *J Exp Psychol Hum Percept Perform* 37:1607–1616. [CrossRef Medline](#)
- Cattaneo Z, Pisoni A, Papagno C (2011) Transcranial direct current stimulation over Broca's region improves phonemic and semantic fluency in healthy individuals. *Neuroscience* 183:64–70. [CrossRef Medline](#)
- Censor N, Sagi D, Cohen LG (2012) Common mechanisms of human perceptual and motor learning. *Nat Rev Neurosci* 13:658–664. [CrossRef Medline](#)
- Clos M, Langner R, Meyer M, Oechslin MS, Zilles K, Eickhoff SB (2012) Effects of prior information on decoding degraded speech: an fMRI study. *Hum Brain Mapp*. Advance online publication. Retrieved August 29, 2013. doi:10.1002/hbm.22151. [CrossRef Medline](#)
- Dahan D, Mead RL (2010) Context-conditioned generalization in adaptation to distorted speech. *J Exp Psychol Hum Percept Perform* 36:704–728. [CrossRef Medline](#)
- Davis MH, Johnsrude IS, Hervais-Adelman A, Taylor K, McGettigan C (2005) Lexical information drives perceptual learning of distorted speech: evidence from the comprehension of noise-vocoded sentences. *J Exp Psychol Gen* 134:222–241. [CrossRef Medline](#)
- de Vries MH, Barth AC, Maiworm S, Knecht S, Zwitserlood P, Flöel A (2010)

- Electrical stimulation of Broca's area enhances implicit learning of an artificial grammar. *J Cogn Neurosci* 22:2427–2436. [CrossRef Medline](#)
- Eickhoff SB, Pomjanski W, Jakobs O, Zilles K, Langner R (2011) Neural correlates of developing and adapting behavioral biases in speeded choice reactions—an fMRI study on predictive motor coding. *Cereb Cortex* 21:1178–1191. [CrossRef Medline](#)
- Eisner F, McGettigan C, Faulkner A, Rosen S, Scott SK (2010) Inferior frontal gyrus activation predicts individual differences in perceptual learning of cochlear-implant simulations. *J Neurosci* 30:7179–7186. [CrossRef Medline](#)
- Erb J, Henry MJ, Eisner F, Obleser J (2012) Auditory skills and brain morphology predict individual differences in adaptation to degraded speech. *Neuropsychologia* 50:2154–2164. [CrossRef Medline](#)
- Flöel A, Rösser N, Michka O, Knecht S, Breitenstein C (2008) Noninvasive brain stimulation improves language learning. *J Cogn Neurosci* 20:1415–1422. [CrossRef Medline](#)
- Friederici AD, Kotz SA, Scott SK, Obleser J (2010) Disentangling syntax and intelligibility in auditory language comprehension. *Hum Brain Mapp* 31:448–457. [CrossRef Medline](#)
- Gandiga PC, Hummel FC, Cohen LG (2006) Transcranial DC stimulation (tDCS): a tool for double-blind sham-controlled clinical studies in brain stimulation. *Clin Neurophysiol* 117:845–850. [CrossRef Medline](#)
- Goldstone RL (1998) Perceptual learning. *Annu Rev Psychol* 49:585–612. [CrossRef Medline](#)
- Greenwood DD (1990) A cochlear frequency-position function for several species—29 years later. *J Acoust Soc Am* 87:2592–2605. [CrossRef Medline](#)
- Hartwigsen G, Price CJ, Baumgaertner A, Geiss G, Koehnke M, Ulmer S, Siebner HR (2010) The right posterior inferior frontal gyrus contributes to phonological word decisions in the healthy brain: evidence from dual-site TMS. *Neuropsychologia* 48:3155–3163. [CrossRef Medline](#)
- Hervais-Adelman A, Davis MH, Johnsrude IS, Carlyon RP (2008) Perceptual learning of noise vocoded words: effects of feedback and lexicality. *J Exp Psychol Hum Percept Perform* 34:460–474. [CrossRef Medline](#)
- Jacobson L, Koslowsky M, Lavidor M (2012) tDCS polarity effects in motor and cognitive domains: a meta-analytical review. *Exp Brain Res* 216:1–10. [CrossRef Medline](#)
- Kühn S, Schmiedek F, Schott B, Ratcliff R, Heinze HJ, Düzel E, Lindenberger U, Lövdén M (2011) Brain areas consistently linked to individual differences in perceptual decision-making in younger as well as older adults before and after training. *J Cogn Neurosci* 23:2147–2158. [CrossRef Medline](#)
- Loebach JL, Pisoni DB, Svirsky MA (2009) Transfer of auditory perceptual learning with spectrally reduced speech to speech and nonspeech tasks: implications for cochlear implants. *Ear Hear* 30:662–674. [CrossRef Medline](#)
- Loebach JL, Pisoni DB, Svirsky MA (2010) Effects of semantic context and feedback on perceptual learning of speech processed through an acoustic simulation of a cochlear implant. *J Exp Psychol Hum Percept Perform* 36:224–234. [CrossRef Medline](#)
- McMillan NA (2005) *Detection theory*. Mahwah, NJ: Erlbaum.
- Nitsche MA, Liebetanz D, Lang N, Antal A, Tergau F, Paulus W (2003) Safety criteria for transcranial direct current stimulation (tDCS) in humans. *Clin Neurophysiol* 114:2220–2222. [CrossRef Medline](#)
- Nitsche MA, Cohen LG, Wassermann EM, Priori A, Lang N, Antal A, Paulus W, Hummel F, Boggio PS, Fregni F, Pascual-Leone A (2008) Transcranial direct current stimulation: state of the art 2008. *Brain Stimul* 1:206–223. [CrossRef Medline](#)
- Nixon P, Lazarova J, Hodinott-Hill I, Gough P, Passingham R (2004) The inferior frontal gyrus and phonological processing: an investigation using rTMS. *J Cogn Neurosci* 16:289–300. [CrossRef Medline](#)
- Obleser J, Eisner F (2009) Pre-lexical abstraction of speech in the auditory cortex. *Trends Cogn Sci* 13:14–19. [CrossRef Medline](#)
- Obleser J, Kotz SA (2010) Expectancy constraints in degraded speech modulate the language comprehension network. *Cereb Cortex* 20:633–640. [CrossRef Medline](#)
- Obleser J, Wise RJ, Alex Dresner M, Scott SK (2007) Functional integration across brain regions improves speech perception under adverse listening conditions. *J Neurosci* 27:2283–2289. [CrossRef Medline](#)
- Obleser J, Eisner F, Kotz SA (2008) Bilateral speech comprehension reflects differential sensitivity to spectral and temporal features. *J Neurosci* 28:8116–8123. [CrossRef Medline](#)
- Oldfield RC (1971) The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9:97–113. [CrossRef Medline](#)
- Poldrack RA, Wagner AD, Prull MW, Desmond JE, Glover GH, Gabrieli JD (1999) Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *Neuroimage* 10:15–35. [CrossRef Medline](#)
- Ragert P, Vandermeeren Y, Camus M, Cohen LG (2008) Improvement of spatial tactile acuity by transcranial direct current stimulation. *Clin Neurophysiol* 119:805–811. [CrossRef Medline](#)
- Rosen S, Faulkner A, Wilkinson L (1999) Adaptation by normal listeners to upward spectral shifts of speech: implications for cochlear implants. *J Acoust Soc Am* 106:3629–3636. [CrossRef Medline](#)
- Scott SK, Blank CC, Rosen S, Wise RJ (2000) Identification of a pathway for intelligible speech in the left temporal lobe. *Brain* 123:2400–2406. [CrossRef Medline](#)
- Sehm B, Schäfer A, Kipping J, Margulies D, Conde V, Taubert M, Villringer A, Ragert P (2012) Dynamic modulation of intrinsic functional connectivity by transcranial direct current stimulation. *J Neurophysiol* 108:3253–3263. [CrossRef Medline](#)
- Shahin AJ, Bishop CW, Miller LM (2009) Neural mechanisms for illusory filling-in of degraded speech. *Neuroimage* 44:1133–1143. [CrossRef Medline](#)
- Shannon RV, Zeng FG, Kamath V, Wygonski J, Ekelid M (1995) Speech recognition with primarily temporal cues. *Science* 270:303–304. [CrossRef Medline](#)
- Sharp DJ, Awad M, Warren JE, Wise RJ, Vigliocco G, Scott SK (2010a) The neural response to changing semantic and perceptual complexity during language processing. *Hum Brain Mapp* 31:365–377. [CrossRef Medline](#)
- Sharp DJ, Turkheimer FE, Bose SK, Scott SK, Wise RJ (2010b) Increased frontoparietal integration after stroke and cognitive recovery. *Ann Neurol* 68:753–756. [CrossRef Medline](#)
- Smalt CJ, Talavage TM, Pisoni DB, Svirsky MA (2011) Neural adaptation and perceptual learning using a portable real-time cochlear implant simulator in natural environments. *Conf Proc IEEE Eng Med Biol Soc* 2011:1145–1148. [CrossRef Medline](#)
- Sohoglu E, Peelle JE, Carlyon RP, Davis MH (2012) Predictive top-down integration of prior knowledge during speech perception. *J Neurosci* 32:8443–8453. [CrossRef Medline](#)
- Sparing R, Dafotakis M, Meister IG, Thirugnanasambandam N, Fink GR (2008) Enhancing language performance with non-invasive brain stimulation—a transcranial direct current stimulation study in healthy humans. *Neuropsychologia* 46:261–268. [CrossRef Medline](#)
- Stagg CJ, Nitsche MA (2011) Physiological basis of transcranial direct current stimulation. *Neuroscientist* 17:37–53. [CrossRef Medline](#)
- Tremblay T, Monetta L, Joannette Y (2004) Phonological processing of words in right- and left-handers. *Brain Cogn* 55:427–432. [CrossRef Medline](#)
- Vickery TJ, Jiang YV (2009) Inferior parietal lobule supports decision making under uncertainty in humans. *Cereb Cortex* 19:916–925. [CrossRef Medline](#)
- Winn MB, Chatterjee M, Idsardi WJ (2012) The use of acoustic cues for phonetic identification: effects of spectral degradation and electric hearing. *J Acoust Soc Am* 131:1465–1479. [CrossRef Medline](#)
- Wohllwill JF (1958) The definition and analysis of perceptual learning. *Psychol Rev* 65:283–295. [CrossRef Medline](#)
- Zaehle T, Sandmann P, Thorne JD, Jäncke L, Herrmann CS (2011) Transcranial direct current stimulation of the prefrontal cortex modulates working memory performance: combined behavioural and electrophysiological evidence. *BMC Neurosci* 12:2. [CrossRef Medline](#)