PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/116736

Please be advised that this information was generated on 2016-05-01 and may be subject to change.
The role of accent imitation in sensorimotor integration during processing of intelligible speech

Patti Adank1,2*, Shirley-Ann Rueschemeyer2,3 and Harold Bekkering2

1 Department of Speech, Hearing and Phonetic Sciences, Division of Psychology and Language Sciences, University College London, London, UK
2 Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Nijmegen, Netherlands
3 Department of Psychology, University of York, York, UK

INTRODUCTION

In everyday communication, humans have to deal with a variety of challenging listening situations, such as the presence of background noise, poor telephone connections, or unfamiliar accents. Recent studies have investigated cognitive and neural mechanisms underlying the ability to process speech effectively in such challenging listening situations (Davis and Johnsrude, 2003; Rodd et al., 2005, 2010; Obleser et al., 2007, 2011; Adank, 2012). The aforementioned studies have established that effective speech comprehension recruits areas involved in speech perception including left Superior Temporal Sulcus (STS), areas involved in linguistic and articulatory processes including left Inferior Frontal Gyrus (IFG), areas involved in speech production including Precentral Gyrus and Supplementary Motor Area (SMA).

There is growing consensus that cortical regions associated with speech production, such as ventral premotor cortex, IFG, SMA/pre-SMA, and primary motor cortex play an active and essential role in effective speech comprehension (Skipper et al., 2006; D’Ausilio et al., 2009; Londei et al., 2009; Callan et al., 2010; Tremblay and Small, 2011; Adank, 2012). Despite this emerging agreement (but see Venezia et al., 2012), much is still unclear about precisely how speech motor and premotor regions contribute to the comprehension process.

Recent theories suggest that mental simulation of perceived actions may aid listeners when predicting upcoming speech signals (Wilson and Knoblich, 2005; Pickering and Garrod, 2007). Simulation theories of action perception argue that observing actions results in the automatic generation of motor plans required to perform these actions. The motor plans are used to produce a forward model of the incoming action. These forward models use information about the movement properties of muscles to generate a simulated course of movement in parallel with, or even in anticipation of, the perceived movement (Grush, 2004). Any discrepancy between the simulated movement from the forward model and the real-world movement results in prediction errors and leads to corrective commands. This type of forward model serves to anticipate others’ (speech) actions as if they were produced by the observer (Locatelli et al., 2012). Forward models thus, generate a series of predictions that may be used as disambiguating information in situations of how action perception is made more difficult due to noisy or ambiguous observing conditions (Wilson and Knoblich, 2005; Pickering and Garrod, 2007).

For speech, these type of conditions involve listening to speech in the presence of background noise, or listening to someone speak with an unfamiliar regional accent (Adank et al., 2009).

Some propose that the prediction signal generated by the forward models is optimized for the type of actions the observers have had experience executing themselves (Knoblich and Sebanz, 2006; Grafon, 2009). Behavioral evidence for this prediction comes from a study on basketball players by Aglioti et al. (2008). Aglioti et al. compared elite basketball players (experts at observing the action and also at performing the action) with a control group (basketball referees; experts at observing the action but not playing themselves) on how effectively each group could judge
whether a basketball would be thrown in the basket. They showed that being experienced in performing an action such as throwing a basketball allows one to efficiently predict the outcome of other players throwing a basketball.

For speech, we recently found that short-term experience with speaking in an unfamiliar regional accent helps comprehension of sentences spoken in that accent (Adank et al., 2010). Adank et al.’s participants listened to sentences spoken in an unfamiliar accent in background noise in a pre-test phase and verbally repeated key words. Next, participants were divided into six groups and either received no training, listened to sentences in the unfamiliar accent without speaking, repeated the accented sentences in their own accent, listened to and transcribed the accented sentences, listened to and imitated the accented sentences, or listened to and imitated the accented sentences without being able to hear their own vocalizations. Post-training measures showed that participants who imitated the speaker’s accent could repeat key words in poorer signal-to-noise ratios (i.e., under more challenging listening conditions) than participants who had not imitated the accent. Adank et al.’s results demonstrate that having experience with speaking in a specific way (i.e., in an unfamiliar accent) can positively affect speech processing by optimizing comprehension of sentences spoken in a similar way.

The neural underpinnings of the optimizing effect of experience with performing an action on action perception have been investigated a several fMRI experiments (Calvo-Merino et al., 2005; Wilson and Iacoboni, 2006; Lahav et al., 2007). Calvo-Merino et al. (2005) studied professional dancers and compared neural activity when male and female dancers viewed dance moves that were performed by their own gender (i.e., moves they would perform when dancing themselves), compared to other moves performed by the other gender (i.e., moves they would always observe, but not perform, when dancing themselves). Calvo-Merino’s design allowed them to separate brain responses related to dance motor representation from those related to visual knowledge about the observed moves. Left dorsal premotor cortex, left inferior parietal sulcus, and right cerebellum showed a higher Blood Oxygenated-Level Dependent (BOLD-) response for dance moves they would perform themselves than for observed moves.

Lahav et al. (2007) trained non-musicians in playing several simple melodies on the piano. Participants were scanned while listening to pieces they had learnt to play, plus pieces they had not learnt. An increased BOLD-response was found in bilateral IFG, the posterior middle premotor region and the inferior parietal lobule (IPL) (Supramarginal gyrus and Angular gyrus) bilaterally and left cerebellum.

Finally, Wilson and Iacoboni tested neural responses for 5 native and 25 non-native speech sounds varying in how easy they were to produce for their native American-English speaking participants. They report increased activation in bilateral superior temporal areas for foreign speech sounds that are more difficult to produce. No correlations were found between the ease with which the foreign sounds could be produced in speech motor areas. Nevertheless, in a region-of-interest analysis, they report that the BOLD-response in premotor cortex, a speech motor area, was elevated for listening to foreign speech sounds compared to native speech sounds. The results from Wilson and Iacoboni thus, contradict the expectation that having had more experience with a motor action (which is the case for native speech sounds) activates motor areas to a higher degree. However, Wilson did not explicitly evaluate how experience with producing the foreign speech sounds affected brain activation when listening to those speech sounds. Thus, it remains unclear how speech motor experience affects neural activation during speech comprehension.

The present study will examine how motor experience with imitating a novel accent affects the activation of cortical areas associated with speech perception and production when subsequently listening to the accented speech in background noise. Based on Adank et al.’s behavioral study, we hypothesized that speech imitation experience supports speech comprehension through the integrating of information from specific speech perception and production areas.

We mapped out the neural bases associated with increased robustness of speech comprehension after imitating an unfamiliar accent and after repeating the accented speech in one’s own accent. Two groups of listeners were scanned using an adapted version of the staircase procedure described in Adank et al. (2010). In a pre-test, participants were examined on their comprehension of sentences spoken in an unfamiliar accent in background noise. Next, one group of participants repeated a series of accented sentences in their own accent, while the second group vocally imitated the sentences in the unfamiliar accent in a training session. Finally, both groups were tested again on their comprehension of accented speech in background noise. If imitation supports sensorimotor integration, we expected a different pattern of activation of cortical areas associated with either speech comprehension or speech production for listeners who have had experience with imitating the unfamiliar accent during the training session. We focused on three left-lateralized regions, namely posterior STS, IFG, and SMA, all of which have been associated with speech perception and speech production tasks (for STS: Scott et al., 2000; Blank et al., 2002; Narain et al., 2003; Crinion and Price, 2005; Tremblay and Gracco, 2009, for IFG: Davis and Johnsrud, 2003; Tremblay and Gracco, 2009; Adank and Devlin, 2010, for SMA: Alario et al., 2006; Wong et al., 2008; Fridriksson et al., 2009).

MATERIALS AND METHODS

PARTICIPANTS

We tested 36 participants (23F and 13M, mean 21.8 years, SD 2.9 years, range 18–29 years). All participants were right-handed, native speakers of Dutch, with no history of oral or written language impairment, or neurological or psychiatric disease. All gave written informed consent and were paid for their participation. The study was approved by the local ethics committee of the Radboud University Nijmegen. Four participants were excluded from further analysis: one participant (F) originally allocated to the imitation group was excluded as the second run was not collected due to technical difficulties, another participant allocated to the imitation group (M) was excluded due to an abnormality in his structural scan, a repeat group participant (M) was excluded as his scan was interrupted due to him feeling unwell, and a final participant (F) in the repeat group was excluded due to poor
performance on the behavioral task; her results showed more than 50% missed responses. All analyses were run on the results from the remaining 32 participants. The repeat group consisted of 15 participants (9F and 6M, mean 21.6 years, SD 3.2 years, range 18–29 years) and the imitation group consisted of 17 participants (10F and 6M, mean 22.2 years, SD 2.8 years, range 18–28 years).

**STIMULUS MATERIAL**

The test stimulus set consisted of 96 sentences (see Table A1 in the Appendix) that were taken from the speech reception threshold (SRT) corpus (Plomp and Mippen, 1979). This corpus has been widely used for assessing intelligibility of different types of stimuli, for example, for foreign-accented speech (Van Wijngaarden et al., 2002). Sentences in the SRT corpus are designed to resemble short samples of conversational speech. All consist of maximally eight or nine syllables and do not include words longer than three syllables. The 48 sentences used in the training (listed in Table A2) were taken from an earlier study on comprehension of accented speech (Adank et al., 2012) and produced by the same speaker as the 96 test sentences. Finally, 30 sentences were recorded in Standard Dutch and used in a hearing pre-test (cf. Table A3).

All sentences were spoken by a single female speaker of Standard Dutch. The test sentences were recorded in both Standard Dutch and an unfamiliar, novel accent. The training sentences were recorded in the novel accent only. The novel accent was created by instructing the speaker to read sentences with an adapted orthography (see also Adank et al., 2010). This accent was designed to sound different from Standard Dutch and not intended to replicate an existing Dutch accent. The orthography was systematically altered to achieve changes in all 15 Dutch vowels as listed in Table 1. Only vowels bearing primary or secondary stress were included in the orthography conversion. An example of a sentence in Standard Dutch and a converted version is given below, including a broad phonetic transcription using the International Phonetic Alphabet (IPA, 1999):

**Table 1 | Intended vowel conversions for obtaining the novel accent.**

<table>
<thead>
<tr>
<th>Orthography</th>
<th>Phonetic (IPA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a → aa</td>
<td>/æ/ → /a:/</td>
</tr>
<tr>
<td>aa → a</td>
<td>/ɑː/ → /ə/</td>
</tr>
<tr>
<td>e → ee</td>
<td>/ɛ/ → /e/</td>
</tr>
<tr>
<td>ee → e</td>
<td>/ɛ/ → /i/</td>
</tr>
<tr>
<td>i → ie</td>
<td>/ɪ/ → /i/</td>
</tr>
<tr>
<td>ie → i</td>
<td>/ɪ/ → /ɪ/</td>
</tr>
<tr>
<td>o → oo</td>
<td>/ɔ/ → /ɔ/</td>
</tr>
<tr>
<td>oo → o</td>
<td>/ɔ/ → /ɔ/</td>
</tr>
<tr>
<td>uu → u</td>
<td>/u/ → /u/</td>
</tr>
<tr>
<td>u → uu</td>
<td>/u/ → /u/</td>
</tr>
<tr>
<td>oe → u</td>
<td>/o/ → /u/</td>
</tr>
<tr>
<td>eu → u</td>
<td>/œ/ → /u/</td>
</tr>
<tr>
<td>au → oe</td>
<td>/ɔu/ → /u/</td>
</tr>
<tr>
<td>ei → ee</td>
<td>/ɛi/ → /ɛi/</td>
</tr>
<tr>
<td>ui → uu</td>
<td>/ɛi/ → /ɪi/</td>
</tr>
</tbody>
</table>

The left column shows the altered orthography in Standard Dutch, and the right column shows the intended change in pronunciation of the vowel in broad phonetic transcription, using the International Phonetic Alphabet (IPA, 1999).

The stimulus materials used in the scanner was created as follows. The sentences in background noise were created by adding continuous speech-shaped noise to the accented sentences in quiet at so that the final signal-to-noise ratio (SNR) was set to −2 dB, 0 dB, +2 dB, +4 dB, +6 dB, +8 dB, +10 dB, +12 dB, +14 dB, or +16 dB, thus, resulting in 10 versions for each sentence. Speech-shaped noise was added using Matlab (Mathworks Inc.). An acoustically matched—yet unintelligible—baseline version of each sentence was created by spectrally rotating (Blesser, 1972) and time-reversing the sentence, using Praat (Boersma and Weenink, 2003). No noise was added to the baseline sentences.

**MRI DATA ACQUISITION**

Whole-brain imaging was performed at the Donders Centre for Brain, Cognition and Behavior, Centre for Cognitive Neuroimaging, at a 3T MR scanner (Magnetom Trio, Siemens Medical Systems, Erlangen, Germany), in Nijmegen, The Netherlands. The sentences were presented over sound attenuating (~30 dB SPL) electrostatic headphones (MRConFon, Magdeburg, Germany) during continuous scanner acquisition (GE-EPI, echo time 35 ms; 32 axial slices; voxel size 3.5 × 3.5 × 3.5 mm, slice thickness 3 mm, inter-slice gap 0.5 mm; field of view 224 mm; flip angle 70). All functional images were acquired in two runs. Functional scans were obtained every 2 s in a continuous scanning paradigm (TR 2 s). Listeners watched a fixation cross that was presented on a screen and viewed through a mirror attached to the head coil. After the acquisition of functional images, a high-resolution structural scan was acquired (T1-weighted MP-RAGE, 192 slices, echo time 3.93 ms; field of view 256 mm, slice thickness 1 mm). Finally, a diffusion-weighted scan was acquired with 68 directions (not included in the present analysis). Total scanning time was around 60 min.

**PROCEDURE**

The experiment had a mixed-subject design with two groups (Figure 1): Repeat or Imitate. Participants were randomly...
allocated to the Repeat or Imitation group. Imaging data was collected in two runs and participants performed a training task in between these runs. Before the start of the scanning session, participants completed a hearing test in a quiet room. This test established the SRT as described in Plomp and Mimpen (1979) for sentences spoken in Standard Dutch by the same female speakers as used for the test in the scanner. Using a similar procedure as in (Adank et al., 2010), participants responded by repeating back what they had heard and the experimenter scored key words (see Table A1 in the Appendix). In this “standard” version, of the SRT, participants were presented the first sentence at +10 dB SNR. If they correctly repeated more than two out of four keywords, the next sentence was played at +2 dB SNR, if they got more than two key words in this sentence correct, the next sentence was played at -6 dB. This procedure continued until the participant either got two keywords right, or if he or she got fewer than two keywords right. If they repeated two keywords correctly, the next sentence was played at the same SNR as the previous sentence, while the SNR was increased with +2 dB if fewer than two keywords were repeated. If the participant than got more than two keywords right, the SNR decreased in steps of -2 dB. This staircase procedure was repeated for 30 sentences (Table A3). The final SNR, or SRT, was calculated as the average across all instances for which a reversal occurred, following Adank et al. (2010) and Plomp and Mimpen (1979). A reversal occurs whenever the SNR changed direction (e.g., from 0 dB SNR to +2 dB SNR, or to -2 dB SNR). The participants showed an average SRT of -3.69 dB (0.95 dB SD), indicating that their hearing was in the normal range for participants with normal hearing (Plomp and Mimpen, 1979).

While in the scanner, participants in both groups performed an intelligibility judgment in the first run. Following the procedure in Adank et al. (2010), participants in the repeat group were instructed to listen to each accented sentence and then to repeat it in their own accent, without imitating the accent. Participants were explicitly instructed not to imitate the speaker’s accent. If participants imitated the accent, they were reminded by the experimenter to repeat the sentence in Standard Dutch. Participants repeated 48 sentences. In the imitation group, participants were instructed to imitate vocally the precise pronunciation of the accented sentence. If participants repeated the sentence in their own accent, the experimenter instructed to imitate the accent as they heard it spoken. Participants imitated 48 sentences during the training and training was not attested using a formal scoring device. The scanner was turned off during the training phase. Next, the second run commenced and participants performed the intelligibility judgment task from run 1 for another series of accented sentences. A total of 96 test sentences was presented across the two runs, plus 48 training sentences in between runs. Per run, 24 acoustically-matched baseline stimuli and 48 test sentences were presented. The distribution of the test sentences and baseline stimuli (which were created based on the 96 test sentences) were counterbalanced across runs and participants.

A single trial in the intelligibility judgment task began with a tone signal of 250 Hz with a duration of 200 ms, followed by a pause of 200 ms, followed by the presentation of a sentence. After the sentence was presented, participants judged its intelligibility, using a button-box with four buttons (one per finger) that they were holding in their right hand. If they found the sentence completely unintelligible, they pressed with their index finger (score 1), if they found the sentence slightly more intelligible they pressed the button under their middle finger (score 2), if they found the sentence intelligible they pressed with their ring finger (score 3) and if they found the sentence very intelligible they pressed the button under their little finger (score 4). This procedure was the same across participants. The SNR of the following sentence depended on the score of the previous sentence. If the participant had rated a specific sentence as 1 or 2, the next sentence was played at an easier (higher) SNR, and if the participant had rated a specific sentence as 3 or 4, the next sentence was played at a less favorable SNR. For instance, if a participant heard a sentence with a SNR of +4 dB and rated the sentence as unintelligible (score 2), then SNR for the next sentence was increased with 2 dB and the next sentence was presented at +6 dB. Alternatively, if a participant heard a sentence with a SNR of +4 dB and rated the sentence as very intelligible (score 4), then SNR for the next sentence was decreased with 2 dB and the next sentence was presented at +2 dB. Each run started with a sentence presented at +10 dB. The SNR decreased in steps of 2 dB until reaching -2 dB SNR and increased in steps of 2 dB up to +16 dB SNR. If these limits were reached, the SNR stayed the same until the participant pressed 3 or 4 (for +16 dB) or 1 or 2 (for -2 dB). As in the behavioral test outside the scanner, SRT was calculated as the average across all instances for which a reversal occurred. This procedure was identical across both runs. Intensity of stimulus presentation was set at a comfortable level for each participant in a familiarization session in which six sentences in Standard Dutch in quiet spoken by the same speaker (not included in the both runs experiment) were presented while the scanner was running. Stimulus presentation was performed using Presentation (Neurobehavioral Systems, Albany, CA), running on a Pentium 4 with 2 GB RAM, and a 2.8 GHz processor. The two runs plus training lasted ~40 min.

**DATA PROCESSING**

The neuroimaging data were pre-processed and analyzed using SPM8 (Wellcome Imaging Department, University College London, London, UK). The first four volumes of every functional run from each participant were excluded from the analysis to minimize T1-saturation effects. Next, the image time series were spatially realigned using a least-squares approach that estimates six rigid-body transformation parameters (Friston et al., 1995) by minimizing head movements between each image and the reference image, that is, the first image in the time series. Next, the time series for each voxel was temporally realigned to acquisition of the middle slice. Subsequently, images were normalized onto a custom Montreal Neurological Institute (MNI)-aligned EPI template (standard in SPM8) using both linear and non-linear transformations and resampled at an isotropic voxel size of 2 mm. All participants’ functional images were smoothed using an 8 mm FWHM Gaussian filter. Each participants’ functional image was processed using a unified segmentation procedure as implemented in SPM8. After segmentation of the T1 structural image (using the unified segmentation model)
and co-registration to the mean functional image (Ashburner and Friston, 1997), functional images were spatially normalized into standard stereotaxic space of the MNI using segmentation parameters of the T1 structural image. A high-pass filter was applied with a 0.0078 Hz (128 s) cut-off to remove low-frequency components from the data, such as scanner drifts. The fMRI time series were analyzed within the context of the General Linear Model using an event-related approach, including an autoregressive AR (1) model during Classical (ReML) parameter estimation.

Three events of interest were identified and entered into a subject-specific General Linear Model: acoustically matched baseline sentences (Baseline), Intelligible Speech (i.e., all sentences rated 1 or 2 by the participant), Intelligible Speech (i.e., all sentences rated 3 or 4 by the participant) for each run. Parameter estimates were calculated for each voxel, and contrast maps were constructed for each participant. The statistical model also considered six separate covariates describing the head-related movements (as estimated by the spatial realignment procedure). Linear-weighted contrasts were used to specify the main contrasts.

Intelligibility (having two levels: Intelligible Speech and Baseline), Run (Run 1 or Run 2), and Group (Repeat and Imitate) were analyzed in a $2 \times 2 \times 2$ mixed ANOVA design, with Run as a within-subject and Group as a between-subject factor. We first identified areas showing a higher BOLD response for Intelligible Speech than for Baseline. Second, we identified areas that showed a higher BOLD-response as a result of having imitated during the training. We compared the activation across runs for both groups. We used the following contrast to identify areas that showed a higher BOLD-response for the Imitate group than the Repeat group by directly comparing the second to the first run for each group: [Intelligible speech Imitation Group Run 2 > Intelligible speech Imitation Group Run 1] > [Intelligible speech Repeat Group Run 2 > Intelligible speech Repeat Group Run 1] (in short: [Imitation Run 2 > 1] > [Repeat Run 2 > 1]). Second, we identified areas showing a higher BOLD-response as a result of having repeated vs. imitated during the training using the contrast: [Intelligible speech Repeat Group Run 2 > Intelligible speech Repeat Group Run 1] > [Intelligible speech Imitation Group Run 2 > Intelligible speech Imitation Group Run 1] (in short: [Repeat Run 2 > 1] > [Imitation Run 2 > 1]). We also tested the contrasts [Imitate Run 2 > Run 1] and [Repeat Run 2 > Run 1] to get a general idea of areas that showed increases in BOLD-response across runs for both groups.

Significant activations in the whole-brain analysis were assessed with a Family Wise Error (FWE) correction for multiple comparisons ($p < 0.05$), in combination with a minimal cluster extent of 20 voxels. Finally, anatomical localization was performed using the Anatomy toolbox in SPM8 (Eickhoff et al., 2005) and verified with MRICRON using the ch2 template in MNI space (Rorden and Brett, 2000).

RESULTS

BEHAVIORAL RESULTS

A repeated-measures ANOVA with Group (Repeat or Imitate) and Run (1 or 2) on the SRT value in dB per run (Figure 2) revealed no main effects, but did show an interaction between Run and Group [$F_{(1, 30)} = 6.21, p = 0.018, \eta^2_p = 0.17$]. The absence of a main effect of Run indicated that both groups of listeners judged the sentence materials to be equally intelligible when played at similar SNRs. The presence of an interaction between Run and Group indicates that listeners in both groups were affected differently by the training. Two post-hoc analyses showed a marginally significant trend for the repeat group to require more favorable SNRs in the second run [$t_{(14)} = -2.033, p = 0.061$], while no such trend was found for the imitate group [$t_{(16)} = 0.946, p = 0.358$].

NEUROIMAGING RESULTS

Table 2 and Figure 2 show the network of areas involved in processing intelligible speech (the contrast Intelligibility included all stimuli that had been rated as 3 or 4 in the intelligibility judgment test) vs. the acoustically matched baseline. Intelligible speech was associated with higher BOLD values with peak coordinates in left and right anterior and posterior STS, left IFG, left insula, and SMA.

We tested the contrasts [Imitation Run 2 > 1] > [Repeat Run 2 > 1] and [Repeat Run 2 > 1] > [Imitation Run 2 > 1] to probe for areas showing an effect of training in the second run for both groups (see Table 2 and Figure 3) in a whole-brain analysis. The areas showing an increased BOLD-response for the contrast [Imitation Run 2 > 1] > [Repeat Run 2 > 1] effect included left STS, an area in left IFG, and a cluster in left SMA. For the reverse contrast [Repeat Run 2 > 1] > [Imitation Run 2 > 1] no significant clusters were found. However, at a less stringent significance level of $p < 0.001$, we found two clusters in left and right Supramarginal Gyrus (SMG) at $[-62, -38, 36]$ (157 voxels) and $[-60, -42, 22]$ (50 voxels). Finally we found no supra-threshold voxels for the contrast [Repeat Run 2 > Run 1], while for [Imitate Run 2 > Run 1] significant clusters were found in MTG bilaterally, left STG, SMA, bilateral Anterior Insula, Right Temporal Pole, and left Fusiform Gyrus.
DISCUSSION

MAIN FINDINGS

We investigated the neural bases associated with increased robustness of speech comprehension after imitating an unfamiliar accent and after repeating the accent in one's own accent. The aim of the study was to establish the effect of short-term experience with imitating accent on speech production. We performed an fMRI experiment with a two-level 2×2 factorial design. The participants included 30 right-handed native English speakers from the University of California, Los Angeles. The design included two runs of vocal imitation, one with an unfamiliar accent and one with an accentation group. Adank and Devlin report that areas that showed increased BOLD-response when participants were first presented with the time-compressed sentences, performance on the sentence-verification task initially deteriorated sharply, but it quickly improved, signaling successful perceptual adaptation. Four cortical areas showed a pattern in the BOLD-activations in line with the behavioral results: one area in left ventral premotor cortex, two left lateralized areas in STG/STS, and one right-lateralized area in STG/STS. Eisner et al. studied the cortical changes in activity related to learning to perceptually adapt to noise-vocoded speech (Shannon et al., 1995). They included two types of distortion: one that was learnable and another was not. Eisner et al. found that activity in left IFG correlated positively with individual variation in working memory scores, and that left IFG and left STS were sensitive to the extent to which the stimulus displayed learnable variation.

Nevertheless, a difference between the present study and Adank and Devlin's is that Adank and Devlin used an online learning design. They monitored neural responses as participants perceptually adapted. In the present study, we used an offline design: we did not monitor participants' neural responses during the training, to avoid problems with head motion during speaking. Consequently, it is possible that left IFG, STS, and SMA were even more active during the training phase for the imitation group. Adank and Devlin report that areas that showed an increase in their BOLD-response when participants were first
confronted with the distortion (time-compressed sentences) in the speech signal also remained more active after participants had adapted. Thus, speech perception and speech production regions continued to be active when the listening conditions remained challenging.

However, others have suggested that other areas other than IFG, SMA, and STS are the focus of adaptation-related changes in speech perception. For instance, Rauschecker and Scott (2009) suggest that the IPL provides an interface “where feed-forward signals from motor preparatory networks in the inferior frontal cortex and premotor cortex can be matched with feedback signals from sensory areas” (p. 722). Increased BOLD-activation was found—albeit at a lowered significance level—for the repeat group in the second run for the contrast [Repeat Run 2 > 1] > [Imitation Run 2 > 1]. However, no activation was found in IPL for the imitation group in the second run compared to the repeat group. It is possible that IPL was more active during the training phase for the imitation group and that the current design has not been able to record this activity as the scanner was turned off during training.

The question of which areas are active while participants are acquiring motor experience may be resolved in an online imitation design, for instance in an experiment in which participants are scanned while vocally imitating or repeating a novel speech stream. Such a design may allow for scrutiny of imitation-related activity in IPL, IFG, STS, and SMA.

**INTERACTION BETWEEN GROUP AND RUN**

The absolute pattern of results for the interaction effect observed in the behavioral data as well as in the neuroimaging data was not completely in line with our predictions. Namely, while we hypothesized an interaction in participants’ performance between Run × Group, we expected that interaction to be driven by a significant increase in performance in Run 2 for participants who had imitated in contrast to participants who repeated in their own accent. Instead, the interaction in the current data set is driven by a relative decrease in performance for participants in the Repeat as compared to the Imitate Group. Thus, while the overall and relative pattern of results is in line with our predictions, the pattern within the second alone warrants further discussion.
The first explanation of these results is in terms of the simulation theories discussed in the Introduction. Simulation models presume that imitative motor involvement helps perceivers anticipate other people's actions better by generating forward models (Grush, 2004). These models serve to anticipate others' actions as if they were our own, something which is proposed to be beneficial for ambiguous or noisy signals or actions (Wilson and Knoblich, 2005; Pickering and Garrod, 2007). Simulation models predict that imitation improves perception, and several studies have recently confirmed this prediction for manual actions (Locatelli et al., 2012) and for speech (Adank et al., 2010). Simulation models generally presume that imitation is predominantly covert (and that any overt imitative action is actively suppressed, cf. Baldissera et al., 2001), but it is unclear how the act of overt motor imitation can support perception.

Moreover, recent behavioral research on perceptual learning in speech indicates that listeners adapt to accented speech by updating their internal phoneme representations for the speech sounds in question (e.g., Norris et al., 2003; Evans and Iverson, 2004). Information from speech articulators may be used to inform feed forward models used in the simulation process. Two recent studies support the possibility of such a supporting mechanism. D'Ausilio et al. (2009) demonstrated that repetitive Transcranial Magnetic Stimulation (rTMS) to lip and tongue motor cortex shifted perception of speech toward speech sounds involving these articulators, while Ito et al. (2009) demonstrated that perception of vowel sounds is affected by feedback from facial muscles while speaking these vowels with an altered pronunciation. Information from muscles can therefore, be used to inform perception. This information could be used by the forward models to improve prediction of utterances in the unfamiliar accent. In contrast, in the repeat group no updated information was available from motor neurons in articulation muscles. In turn, this could have resulted in the trend toward poorer comprehension after training in the current experiment.

However, an alternative explanation of the interaction between group and run and the trend toward lower intelligibility judgments in repeat group could be that participants in the repeat group performed a less engaging task. It seems plausible that having to replicate the way in which a sentence is being said is a task that requires more attention and effort than a task in which participants had to pronounce what was being said in their own accent. This possible increased lower engagement in the Repeat group may have had the following two consequences.

First, participants in the Repeat group could have gotten more fatigued and bored than those in the Imitation group. Consequently, participants in the Repeat group may not have been able (or willing) to pay as much attention in the first run. There exists anecdotal and experimental support for the notion that it is harder to maintain attention in unchallenging, monotonous tasks than for cognitively demanding but interesting tasks (see Robertson and O'connell, 2010, for an overview). Generally, behavioral performance deteriorates for longer sustained tasks, and this deterioration is faster for unchallenging tasks (such as repeating compared to imitating). A recent meta-analysis of neuroimaging studies on vigilant attention (i.e., the cognitive function that enables prolonged engagement in intellectually unchallenging, uninteresting activities) identified a network of—mostly right-lateralized—brain regions subserving vigilant attention, including dorsomedial, mid- and ventrolateral prefrontal cortex, anterior insula, parietal areas (intraparietal sulcus, temporoparietal junction), and finally subcortical structures (cerebellar vermis, thalamus, putamen, and midbrain) (Langner and Eickhoff, 2013). Note that our network of areas displaying the interaction effect ([Imitation Run 2 > 1] > [Repeat Run 2 > 1]) was exclusively left-lateralized and does not appear to overlap with the network identified by Langner and Eickhoff. Furthermore, inspecting of the reverse contrast ([Repeat Run 2 > 1] > [Imitation Run 2 > 1]) at a less stringent significance level indicated involvement of left IPL, whereas Langner and Eickhoff’s analysis showed no clusters in the left parietal lobe.

Second, participants in the Repeat group performed a task that may have not allowed them to engage with the speaker of the accent as much as participants in the Imitation group. This decreased engagement could then have made them less willing to put as much effort into understanding the accented sentences in the second run as in the first run. This notion is speculative, but is supported by an experiment on the effect of imitating a target’s (such as a speaker) behavior on the perception of the social characteristics of that target. In a within-subject design, Adank et al. (2013) asked a group of participants to listen to sentences spoken by two speakers of a regional accent (Glaswegian) of British-English. Next, they repeated the sentences spoken by one of the Glaswegian speakers in their own accent, and subsequently participants imitated sentences spoken by the second Glaswegian speaker. The order of repeating and imitating was counterbalanced across participants. After each repeating or imitation session, participants completed a questionnaire probing the speakers’ perceived power, competence, and social attractiveness. Imitating had a positive effect on the perceived social attractiveness of the speaker compared to repeating. The authors explained the positive results of imitating by stating that the act of imitating another’s accent makes the speaker part of participants’ social in-group in a way that repeating does not. Since people are more positively biased toward people in their in-group than those outside (Brewer, 1979), such ingroup favoritism then made the speaker seem more subjectively pleasant. However, it is unclear whether a comparable effect of imitation on perception of the speaker would also occur for the constructed accent used in our study. Hearing accented speech automatically invokes attitudes associated with speakers of the accent (Giles, 1970; Bishop et al., 2005; Coupland and Bishop, 2007) and it seems unlikely that constructed accents are associated with specific language attitudes as is the case for existing accents.

The issue of differences in engagement with the task across groups issue may be addressed in future studies by explicitly matching training tasks for “level of interest.” One way to achieve this would be to create two different constructed accents that are matched for intelligibility and have participants repeat sentences for one of the accents and imitate for the other accent in a within-subject design. Subsequently, their intelligibility of both accents
could be tested to verify that imitation had a positive effect on the intelligibility of the imitated constructed accent. However, we chose not to take this route in the present study, as we aimed to replicate the design in Adank et al. (2010) as closely as possible using an fMRI design. Yet, future studies could, for instance, pre-test a variety of tasks on how much they are perceived to be equally interesting or engaging by participants or use multiple constructed accents.

A final explanation for the pattern in the behavioral results could be that the behavioral task used in the experiment, perhaps in combination with the presence of continuous scanner noise, did not accurately reflect the extent to which participants’ performances changed over the two runs. Our earlier behavioral study described in Adank et al. (2010) used the standard SRT procedure (which is the same as described for the hearing test in the Materials and Methods section), and it seems plausible that this test provides a more fine-grained measure of intelligibility processing than the test used in the scanner. Adank et al.’s SRT procedure involved overt vocal responses, which was not optimal given our fMRI design. Future studies should consider using a more fine-grained behavioral task, possibly in combination with a sparse sampling design (Hall et al., 1999).

REFERENCES


CONCLUSION

Previous studies (e.g., Davis and Johnsrude, 2003; Adank, 2012) have shown that processing of intelligible speech involves neural substrates associated with speech production and speech perception. This study adds to a growing body of literature showing that processing of intelligible speech involves sensorimotor processes. However, further research is required to establish in what manner overt imitation of speech contributes to these sensorimotor processes in challenging listening situations. Ideally, future studies involving imitation related-training should make a point of explicitly disentangling the possible contribution of factors such as participant (dis)engagement and listening environment. Only by addressing these issues can the contribution of imitative actions to speech perception be elucidated.

ACKNOWLEDGMENTS

We wish to thank Pascal de Water and Paul Gaalman for technical assistance and Ivan Toni for useful comments on the design of the study. The study was supported by the Dutch Organization for Scientific Research (NWO) through a VENI grant 016904–053 to Shirley-Ann Rueschemeyer and a VICI grant 453–05–001 to Harold Bekkering.

Frontiers in Human Neuroscience

October 2013 | Volume 7 | Article 634 | 9

www.frontiersin.org
Vocal imitation and sensorimotor processing


## APPENDIX

### Table A1: Test sentences from the Speech Reception Threshold corpus (Plomp and Mimpen, 1979).

<table>
<thead>
<tr>
<th>Nr.</th>
<th>Standard Dutch</th>
<th>Accented Dutch</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>De bal vloog over de schutting</td>
<td>De baal vloog offer de schuutting</td>
</tr>
<tr>
<td>2</td>
<td>Morgen wil ik maar één liter melk meelk</td>
<td>Moorgen wiel iek mar én litter meelk</td>
</tr>
<tr>
<td>3</td>
<td>Deze kerk moet gesloopt worden</td>
<td>Desse keerk mut geslopt woorden</td>
</tr>
<tr>
<td>4</td>
<td>De spoorrein was al gauw kapot</td>
<td>De sportrein waas aal goew kaapoot</td>
</tr>
<tr>
<td>5</td>
<td>De nieuwe fiets is gestolen</td>
<td>De niewe fiets ies gestollen</td>
</tr>
<tr>
<td>6</td>
<td>Zijn manier van werken ligt mij niet</td>
<td>Zeen manir vaan werken liegt mee nit</td>
</tr>
<tr>
<td>7</td>
<td>Het slot van de voordeur is kapot</td>
<td>Het slot vaan de vordur ies kaapoot</td>
</tr>
<tr>
<td>8</td>
<td>Dat hotel heeft een slechte naam</td>
<td>Daat hotteel heft ’n sleechte nam</td>
</tr>
<tr>
<td>9</td>
<td>De jongen werd stevig aangepakt</td>
<td>De joongen weerd steffig angepaakt</td>
</tr>
<tr>
<td>10</td>
<td>Het natte hout siet in het vuur</td>
<td>Het naatte hoet siest ien het vur</td>
</tr>
<tr>
<td>11</td>
<td>Zijn fantasie kent geen grenzen</td>
<td>Zeen faantassih keent gèn grenzen</td>
</tr>
<tr>
<td>12</td>
<td>De aardappels liggen in de schuur</td>
<td>De ardaappels liegen ien de schur</td>
</tr>
<tr>
<td>13</td>
<td>Alle prijzen waren verhoogd</td>
<td>Aalle preezen warren verhogt</td>
</tr>
<tr>
<td>14</td>
<td>Zijn fantasie kent geen grenzen</td>
<td>Zeen faantassih keent gèn</td>
</tr>
<tr>
<td>15</td>
<td>Het dak moet nodig hersteld worden</td>
<td>Het daak mut noddied heersteeld woorden</td>
</tr>
<tr>
<td>16</td>
<td>De kachel is nog steeds niet aan</td>
<td>De kaachel ies noog stèds nit an</td>
</tr>
<tr>
<td>17</td>
<td>Van de viool is een snaar kapot</td>
<td>Vaan de vij-jol ies ’n snar kaapoot</td>
</tr>
<tr>
<td>18</td>
<td>De tuinman heeft het gras gemaiad</td>
<td>De tuinmaan heft het graas gemait</td>
</tr>
<tr>
<td>19</td>
<td>De appels aan de boom zijn rijp</td>
<td>De ardaappels liegen ien de schur</td>
</tr>
<tr>
<td>20</td>
<td>Voor het eerst was er nieuwe haring</td>
<td>Vor het erst waas eer niwwe harrieng</td>
</tr>
<tr>
<td>21</td>
<td>Het loket bleef lang gesloten</td>
<td>Het lokket bleef laang geslotten</td>
</tr>
<tr>
<td>22</td>
<td>Er werd een diepe kuil gegraven</td>
<td>Eer weerd ’n dippe koel gegraffen</td>
</tr>
<tr>
<td>23</td>
<td>Zijn gezicht heeft een rode kleur</td>
<td>Zeen lèfteed liegt boffen de deertieg</td>
</tr>
<tr>
<td>24</td>
<td>Het begon vroeg donker te worden</td>
<td>Het beggoon bruik doonker te woorden</td>
</tr>
<tr>
<td>25</td>
<td>Het gras was helemaal verdroogd</td>
<td>Het graas waas hélemaal verdrogt</td>
</tr>
<tr>
<td>26</td>
<td>Spoedig kwam er een einde aan</td>
<td>Spuddieg kwaam eer ’n eende an</td>
</tr>
<tr>
<td>27</td>
<td>Ieder half uur komt hier een bus langs</td>
<td>Idder haalf ur koomt hir ’n buus langs</td>
</tr>
<tr>
<td>28</td>
<td>De bel van de voordeur is kapot</td>
<td>De beel vaan de vordur ies kaapoot</td>
</tr>
<tr>
<td>29</td>
<td>De wind waait vandaag uit het westen</td>
<td>De wiend wajt vaandag ute het westen</td>
</tr>
<tr>
<td>30</td>
<td>De slang bewoog zich door het gras</td>
<td>De slang bewog zich dor het graas</td>
</tr>
<tr>
<td>31</td>
<td>De kamer rook naar sigaren</td>
<td>De kammer rok nar siggarren</td>
</tr>
<tr>
<td>32</td>
<td>De appel had een zure smaak</td>
<td>De aappel haad ’n zurre smak</td>
</tr>
<tr>
<td>33</td>
<td>De trein kwam met een schok tot stilstand</td>
<td>De treen kwaam meet ’n schook toot stielstaand</td>
</tr>
</tbody>
</table>

(Continued)

<table>
<thead>
<tr>
<th>Nr.</th>
<th>Standard Dutch</th>
<th>Accented Dutch</th>
</tr>
</thead>
<tbody>
<tr>
<td>34</td>
<td>De koeien werden juist gemolken</td>
<td>De kujjen weerd juust gemoolken</td>
</tr>
<tr>
<td>35</td>
<td>Het duurt niet langer dan een minuut</td>
<td>Het durt nit laanger daan ’n minuut</td>
</tr>
<tr>
<td>36</td>
<td>De grijze lucht voorspeelt regen</td>
<td>De greeze luucht vorspeelt règen</td>
</tr>
<tr>
<td>37</td>
<td>Hij kon de hamer neergens vinden</td>
<td>Hee koorn de hammer neergens vienden</td>
</tr>
<tr>
<td>38</td>
<td>Deze berg is nog niet bekloommen</td>
<td>Desse beel ias noog nit bekloommen</td>
</tr>
<tr>
<td>39</td>
<td>De bel van mijn fiets is kapot</td>
<td>De beel vaan meen fits ies kaapoot</td>
</tr>
<tr>
<td>40</td>
<td>De auto heeft een lekke band</td>
<td>De oetoh héft ’n leekke baand</td>
</tr>
<tr>
<td>41</td>
<td>Het moeilijke werk bleef liggen</td>
<td>De muj-leeke werk blef liegen</td>
</tr>
<tr>
<td>42</td>
<td>Het vliegtuig vertrekt over een uur</td>
<td>Het vlügtuig vertrekt ofter ’n ur</td>
</tr>
<tr>
<td>43</td>
<td>De jongens vechten de hele dag</td>
<td>De joongens vechten de hèlle daag</td>
</tr>
<tr>
<td>44</td>
<td>De schoenen moeten verzold worden</td>
<td>De schunnen mutten verzold woorden</td>
</tr>
<tr>
<td>45</td>
<td>In de krant staat vandaag niet veel nieuws</td>
<td>Ien de kraant staat vaandag nit vèl niws</td>
</tr>
<tr>
<td>46</td>
<td>Door de neus ademen is beter</td>
<td>Dor de nus addemmen ies better</td>
</tr>
<tr>
<td>47</td>
<td>Het kind was niet in staat te spreken</td>
<td>Het kiend waas nit ien stat te spreken</td>
</tr>
<tr>
<td>48</td>
<td>De witte zwaan dook onder water</td>
<td>De wiette zwaan dook onderer water</td>
</tr>
<tr>
<td>49</td>
<td>Hij nam het pak onder zijn arm</td>
<td>Hee naam het paak oonder zeen aarm</td>
</tr>
<tr>
<td>50</td>
<td>Gelukkig sloeg de motor niet af</td>
<td>Geluukkieg slug de mottor nit aaf</td>
</tr>
<tr>
<td>51</td>
<td>De leeraar gaf hem een laag cijfer</td>
<td>De lèrrar gaaf heem ’n lag seeffer</td>
</tr>
<tr>
<td>52</td>
<td>Een kopje koffie zal goed smaken</td>
<td>Een koopje kooffih zaal gud smaken</td>
</tr>
<tr>
<td>53</td>
<td>De foto is mooi ingelijst</td>
<td>De fotto ies moi iengeleest</td>
</tr>
<tr>
<td>54</td>
<td>Mijn broer gaat elke dag fietsen</td>
<td>Meen brur gat eelke daag fitsen</td>
</tr>
<tr>
<td>55</td>
<td>De foto is mooi ingelijst</td>
<td>De fotto ies moi iengeleest</td>
</tr>
<tr>
<td>56</td>
<td>De schreffer van dit boek is dood</td>
<td>De schreffer van dit buk ies dot</td>
</tr>
<tr>
<td>57</td>
<td>Zij heeft haar proefwerk slecht gemaakt</td>
<td>Zee heft har prufwerk slecht gemakt</td>
</tr>
<tr>
<td>58</td>
<td>De sigaar ligt in de asbak</td>
<td>De siggar ies in de aasbaak</td>
</tr>
<tr>
<td>59</td>
<td>De appelboom stond in volle bloei</td>
<td>De appelboom stond in volle bloei</td>
</tr>
<tr>
<td>60</td>
<td>Er wordt in dit land geen rijst verbouwd</td>
<td>Eer woordt ien diet laand gèn reest verbuwd</td>
</tr>
<tr>
<td>61</td>
<td>Hij kon er nu eenmaal niets aan doen</td>
<td>Hee kaan er nu eenmaal nits an doen</td>
</tr>
<tr>
<td>62</td>
<td>De kleren waren niet gewassen</td>
<td>De kleren waren nit gewaassen</td>
</tr>
<tr>
<td>63</td>
<td>Het gedicht werd voorgelezen</td>
<td>Het gediecht weerd vorgelèssen</td>
</tr>
<tr>
<td>64</td>
<td>Haar gezicht was zwart van het vuil</td>
<td>Haar geziecht waas zwaart vaan het vuil</td>
</tr>
<tr>
<td>65</td>
<td>Het slot van de voordeur is kapot</td>
<td>Het sloot vaan de vordur ies kaapoot</td>
</tr>
</tbody>
</table>

(Continued)
Table A1 | Continued

<table>
<thead>
<tr>
<th>Nr.</th>
<th>Standard Dutch</th>
<th>Accented Dutch</th>
</tr>
</thead>
<tbody>
<tr>
<td>66</td>
<td>De groene appels waren erg zuur</td>
<td>De grune appels waren er eeg zuur</td>
</tr>
<tr>
<td>67</td>
<td>In het gebouw waren vier liften</td>
<td>Ien het gebouw waren vier liften</td>
</tr>
<tr>
<td>68</td>
<td>Lopen is gezonder dan fietsen</td>
<td>Lopen ies gezoonder dan fietsen</td>
</tr>
<tr>
<td>69</td>
<td>Het lawaai maakte hem wakker</td>
<td>Het lawaai maakte hem wakker</td>
</tr>
<tr>
<td>70</td>
<td>Mijn buurman heeft een auto gekocht</td>
<td>Mijn buurman heeft een auto gekocht</td>
</tr>
<tr>
<td>71</td>
<td>Als het flink vriest kunnen we schaatsen</td>
<td>Als het flink vriest kunnen we schaatsen</td>
</tr>
<tr>
<td>72</td>
<td>De kast was een meter verschoven</td>
<td>De kaast was een meter verschoven</td>
</tr>
<tr>
<td>73</td>
<td>Oude meubels zijn zeer in trek</td>
<td>Oude meubels zijn zeer in trek</td>
</tr>
<tr>
<td>74</td>
<td>De portier ging met vakantie</td>
<td>De portier ging met vakantie</td>
</tr>
<tr>
<td>75</td>
<td>De lantaarn gaf niet veel licht</td>
<td>De lantaarn gaf niet veel licht</td>
</tr>
<tr>
<td>76</td>
<td>Door zijn snelheid vloog hij uit de bocht</td>
<td>Door zijn snelheid vloog hij uit de bocht</td>
</tr>
<tr>
<td>77</td>
<td>Het is hier nog steeds veel te koud</td>
<td>Het is hier nog steeds veel te koud</td>
</tr>
<tr>
<td>78</td>
<td>De oude man was kaal geworden</td>
<td>De oude man was kaal geworden</td>
</tr>
<tr>
<td>79</td>
<td>De bomen waren helemaal kaal</td>
<td>De bomen waren helemaal kaal</td>
</tr>
<tr>
<td>80</td>
<td>Een flinke borrel zal mij goed doen</td>
<td>Een flinke borrel zal mij goed doen</td>
</tr>
<tr>
<td>81</td>
<td>Onze bank geeft vijf procent rente</td>
<td>Onze bank geeft vijf procent rente</td>
</tr>
<tr>
<td>82</td>
<td>Het verslag in de krant is kort</td>
<td>Het verslag in de krant is kort</td>
</tr>
<tr>
<td>83</td>
<td>In de vijver zwemmen veel vissen</td>
<td>In de vijver zwemmen veel vissen</td>
</tr>
<tr>
<td>84</td>
<td>Honden mogen niet in het gebouw</td>
<td>Honden mogen niet in het gebouw</td>
</tr>
<tr>
<td>85</td>
<td>Een flinke borrel zal mij goed doen</td>
<td>Een flinke borrel zal mij goed doen</td>
</tr>
<tr>
<td>86</td>
<td>Gisteren waaide het nog harder</td>
<td>Gisteren waaide het nog harder</td>
</tr>
<tr>
<td>87</td>
<td>Het meisje stond lang te wachten</td>
<td>Het meisje stond lang te wachten</td>
</tr>
<tr>
<td>88</td>
<td>De volgende dag kwam hij ook niet</td>
<td>De volgende dag kwam hij ook niet</td>
</tr>
<tr>
<td>89</td>
<td>Het gescheurde is duidelijk hoorbaar</td>
<td>Het gescheurde is duidelijk hoorbaar</td>
</tr>
<tr>
<td>90</td>
<td>Eindelijk kwam de trein op gang</td>
<td>Eindelijk kwam de trein op gang</td>
</tr>
<tr>
<td>91</td>
<td>De grote stad trok hem wel aan</td>
<td>De grote stad trok hem wel aan</td>
</tr>
<tr>
<td>92</td>
<td>De bus is vandaag niet op tijd</td>
<td>De bus is vandaag niet op tijd</td>
</tr>
<tr>
<td>93</td>
<td>Onze dochter speelt goed blokfluit</td>
<td>Oonze dochter speelt goed blokfluit</td>
</tr>
<tr>
<td>94</td>
<td>Ook in de zomer is het hier koel</td>
<td>Ook in de zomer is het hier koel</td>
</tr>
<tr>
<td>95</td>
<td>Zij moesten vier uur hard werken</td>
<td>Zij moesten vier uur hard werken</td>
</tr>
<tr>
<td>96</td>
<td>Niemand kan de Fransman verstaan</td>
<td>Niemand kan de Fransman verstaan</td>
</tr>
</tbody>
</table>

Only vowels bearing primary or secondary lexical stress were included in the conversion of the orthography.

---

Table A2 | Training sentences.

<table>
<thead>
<tr>
<th>Nr.</th>
<th>Standard Dutch</th>
<th>Accented Dutch</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Architekten hebben een beroep</td>
<td>Aarchitekten hebben een beroep</td>
</tr>
<tr>
<td>2</td>
<td>Asperges zijn groenten</td>
<td>Aaspeerges zijn groenten</td>
</tr>
<tr>
<td>3</td>
<td>Aardappels worden geschild</td>
<td>Aardappels worden geschild</td>
</tr>
<tr>
<td>4</td>
<td>Bananen zijn fruit</td>
<td>Bananen zijn fruit</td>
</tr>
<tr>
<td>5</td>
<td>Bijen vliegen rond op zoek naar voedsel</td>
<td>Bijen vliegen rond op zoek naar voedsel</td>
</tr>
<tr>
<td>6</td>
<td>Bevers bouwen dammen in de rivier</td>
<td>Bevers bouwen dammen in de rivier</td>
</tr>
<tr>
<td>7</td>
<td>Bieren hebben vier poten</td>
<td>Bieren hebben vier poten</td>
</tr>
<tr>
<td>8</td>
<td>Bisschoppen dragen kleren</td>
<td>Bisschoppen dragen kleren</td>
</tr>
<tr>
<td>9</td>
<td>Biefstukken worden verkocht door slagers</td>
<td>Biefstukken worden verkocht door slagers</td>
</tr>
<tr>
<td>10</td>
<td>Blikopeners kunnen in winkels gekocht worden</td>
<td>Blikopeners kunnen in winkels gekocht worden</td>
</tr>
<tr>
<td>11</td>
<td>Brokmijten rijden op de weg</td>
<td>Brokmijten rijden op de weg</td>
</tr>
<tr>
<td>12</td>
<td>Chirurgen moeten lang studeren</td>
<td>Chirurgen moeten lang studeren</td>
</tr>
<tr>
<td>13</td>
<td>Druiven zijn eetbaar</td>
<td>Druiven zijn eetbaar</td>
</tr>
<tr>
<td>14</td>
<td>Ezels dragen zware lasten</td>
<td>Ezels dragen zware lasten</td>
</tr>
<tr>
<td>15</td>
<td>Ezels kunnen kopig zijn</td>
<td>Ezels kunnen kopig zijn</td>
</tr>
<tr>
<td>16</td>
<td>Forellen hebben sterke tanden</td>
<td>Forellen hebben sterke tanden</td>
</tr>
<tr>
<td>17</td>
<td>Ganzen kunnen zwemmen en lopen</td>
<td>Ganzen kunnen zwemmen en lopen</td>
</tr>
<tr>
<td>18</td>
<td>Heien hebben sterke tanden</td>
<td>Heien hebben sterke tanden</td>
</tr>
<tr>
<td>19</td>
<td>Heggenscharen worden in de tuin gekocht</td>
<td>Heggenscharen worden in de tuin gekocht</td>
</tr>
<tr>
<td>20</td>
<td>Kapiteins voeren het bevel op schepen</td>
<td>Kapiteins voeren het bevel op schepen</td>
</tr>
<tr>
<td>21</td>
<td>Kasten zijn van hout</td>
<td>Kasten zijn van hout</td>
</tr>
<tr>
<td>22</td>
<td>Kroketten zijn gefrituurd</td>
<td>Kroketten zijn gefrituurd</td>
</tr>
<tr>
<td>23</td>
<td>Lammetjes komen van schapen</td>
<td>Lammetjes komen van schapen</td>
</tr>
<tr>
<td>24</td>
<td>Makrelen ademen door kieuwen</td>
<td>Makrelen ademen door kieuwen</td>
</tr>
<tr>
<td>25</td>
<td>Leeuwen hebben manen</td>
<td>Leeuwen hebben manen</td>
</tr>
<tr>
<td>26</td>
<td>Luipaarden hebben een vacht</td>
<td>Luipaarden hebben een vacht</td>
</tr>
<tr>
<td>27</td>
<td>Meuses ademen door kieuwen</td>
<td>Meuses ademen door kieuwen</td>
</tr>
<tr>
<td>28</td>
<td>Messen worden gebouwd als keukengerei</td>
<td>Messen worden gebouwd als keukengerei</td>
</tr>
<tr>
<td>29</td>
<td>Monniken wonen in een klooster</td>
<td>Monniken wonen in een klooster</td>
</tr>
<tr>
<td>30</td>
<td>Nachtegalen hebben een bijenkorf</td>
<td>Nachtegalen hebben een bijenkorf</td>
</tr>
<tr>
<td>31</td>
<td>Ooms zijn deel van de familie</td>
<td>Ooms zijn deel van de familie</td>
</tr>
<tr>
<td>32</td>
<td>Otters kunnen goed zwemmen</td>
<td>Otters kunnen goed zwemmen</td>
</tr>
<tr>
<td>33</td>
<td>Pinguïns eten veel vis</td>
<td>Pinguïns eten veel vis</td>
</tr>
<tr>
<td>34</td>
<td>Presidents wonen in de politiek</td>
<td>Presidents wonen in de politiek</td>
</tr>
<tr>
<td>35</td>
<td>Ratelslangen spelen op hun buik</td>
<td>Ratelslangen spelen op hun buik</td>
</tr>
<tr>
<td>36</td>
<td>Rodborstjes hebben een snaald</td>
<td>Rodborstjes hebben een snaald</td>
</tr>
<tr>
<td>37</td>
<td>Schuurtjes worden gebruikt voor opslag</td>
<td>Schuurtjes worden gebruikt voor opslag</td>
</tr>
<tr>
<td>38</td>
<td>Slagers hebben winkels</td>
<td>Slagers hebben winkels</td>
</tr>
<tr>
<td>39</td>
<td>Sloffen worden gemaakt in een fabriek</td>
<td>Sloffen worden gemaakt in een fabriek</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Nr.</th>
<th>Standard Dutch</th>
<th>Accented Dutch</th>
</tr>
</thead>
<tbody>
<tr>
<td>40</td>
<td>Tantes zijn altijd vrouwen</td>
<td>Taantes zeen aalteed vroewen</td>
</tr>
<tr>
<td>41</td>
<td>Tigers hebben een staart</td>
<td>Teegers heeben ’n start</td>
</tr>
<tr>
<td>42</td>
<td>Tomaten groeien aan planten</td>
<td>Tommatten grujjen an plaanten</td>
</tr>
<tr>
<td>43</td>
<td>Vaders zijn ouders</td>
<td>Vadders zeen oeders</td>
</tr>
<tr>
<td>44</td>
<td>Vlinders hebben voelsprieten</td>
<td>Vlienders heeben vulspritten</td>
</tr>
<tr>
<td>45</td>
<td>Wandelschoenen zijn gefabriceerde goederen</td>
<td>Waandelschunnen zeen gefaabriserde gudderen</td>
</tr>
<tr>
<td>46</td>
<td>Wijflessen hebben een kuur</td>
<td>Weenfleessen heeben ’n kuurk</td>
</tr>
<tr>
<td>47</td>
<td>Wetenschappers moeten lang studeren</td>
<td>Wetenschaapers mutten laang studderren</td>
</tr>
<tr>
<td>48</td>
<td>Wortels groeien in een moestuin</td>
<td>Woortels grujjen ien ’n mustuun</td>
</tr>
</tbody>
</table>

*Only vowels bearing primary or secondary lexical stress were included in the conversion of the orthography.*

<table>
<thead>
<tr>
<th>Nr.</th>
<th>Standard Dutch</th>
<th>Key words</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Het regent al de hele dag</td>
<td>regent al hele dag</td>
</tr>
<tr>
<td>2</td>
<td>De schaatsen zijn in het vet gezet</td>
<td>schaatsen zijn vet gezet</td>
</tr>
<tr>
<td>3</td>
<td>In juni zijn de dagen het langst</td>
<td>juni zijn dagen langst</td>
</tr>
<tr>
<td>4</td>
<td>De bakkers bezorgen vandaag niet</td>
<td>bakkers bezorgen vandaag niet</td>
</tr>
<tr>
<td>5</td>
<td>Hij rookte zijn sigaret op</td>
<td>rookte zijn sigaret op</td>
</tr>
<tr>
<td>6</td>
<td>Het was heel stil in de duinen</td>
<td>was heel stil duinen</td>
</tr>
<tr>
<td>7</td>
<td>Door zijn haast maakte hij veel fouten</td>
<td>Door haast veel fouten</td>
</tr>
<tr>
<td>8</td>
<td>De kat likt het schoteltje leeg</td>
<td>kat likt schoteltje leeg</td>
</tr>
<tr>
<td>9</td>
<td>De hond blafte de hele nacht</td>
<td>hond blafte hele nacht</td>
</tr>
<tr>
<td>10</td>
<td>Het tempo lag voor hem veel te hoog</td>
<td>tempo hem veel hoog</td>
</tr>
<tr>
<td>11</td>
<td>Lawaai maakt je op den duur doof</td>
<td>Lawaai maakt duur doof</td>
</tr>
<tr>
<td>12</td>
<td>Moeizaam klom de man naar boven</td>
<td>Moeizaam klom man boven</td>
</tr>
<tr>
<td>13</td>
<td>Hij probeerde het nog een keer</td>
<td>Hij probeerde nog keer</td>
</tr>
<tr>
<td>14</td>
<td>Toch lijkt me dat een goed voorstel</td>
<td>Toch lijkt goed voorstel</td>
</tr>
<tr>
<td>15</td>
<td>Dat was voor hem een bittere pil</td>
<td>Dat hem bittere pil</td>
</tr>
<tr>
<td>16</td>
<td>De nieuwe zaak was pas geopend</td>
<td>nieuwe zaak pas geopend</td>
</tr>
<tr>
<td>17</td>
<td>De rivier trad buiten haar oevers</td>
<td>rivier trad buiten oevers</td>
</tr>
<tr>
<td>18</td>
<td>De biefstuk was vandaag erg mals</td>
<td>biefstuk vandaag erg mals</td>
</tr>
<tr>
<td>19</td>
<td>Op het gras mag men niet lopen</td>
<td>gras mag niet lopen</td>
</tr>
<tr>
<td>20</td>
<td>De trein vertrekt over twee uur</td>
<td>trein vertrekt twee uur</td>
</tr>
<tr>
<td>21</td>
<td>De kat van de buren is weg</td>
<td>kat van buren weg</td>
</tr>
<tr>
<td>22</td>
<td>De wagen reed snel de berg af</td>
<td>wagen reed snel berg</td>
</tr>
<tr>
<td>23</td>
<td>Iedereen genoot van het uitzicht</td>
<td>iedereen genoot van uitzicht</td>
</tr>
<tr>
<td>24</td>
<td>Steile trappen zijn gevaarlijk</td>
<td>Steile trappen zijn gevaarlijk</td>
</tr>
<tr>
<td>25</td>
<td>De zon gaat in het westen onder</td>
<td>zon gaat westen onder</td>
</tr>
<tr>
<td>26</td>
<td>De zak zat vol oude rommel</td>
<td>zak vol oude rommel</td>
</tr>
<tr>
<td>27</td>
<td>Zij werd misselijk van het rijden</td>
<td>Zij werd misselijk rijden</td>
</tr>
<tr>
<td>28</td>
<td>Het licht in de gang brandt nog steeds</td>
<td>licht gang brandt steeds</td>
</tr>
<tr>
<td>29</td>
<td>In de kerk werd mooi orgel gespeeld</td>
<td>kerk mooi orgel gespeeld</td>
</tr>
<tr>
<td>30</td>
<td>De jongens gingen er gauw vandoor</td>
<td>jongens gingen gauw vandoor</td>
</tr>
</tbody>
</table>