


Comparing Clinically Applicable Behavioral and Electrophysiological Measures of Speech Detection, Discrimination, and Comprehension

Trends in Hearing
Volume 26: 1–13
© The Author(s) 2022
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/23312165221139733
journals.sagepub.com/home/tia


Pushkar Deshpande^{1,2} , Christian Brandt^{1,2} , Stefan Debener³
and Tobias Neher^{1,2} 

Abstract

Effective communication requires good speech perception abilities. Speech perception can be assessed with behavioral and electrophysiological methods. Relating these two types of measures to each other can provide a basis for new clinical tests. In audiological practice, speech detection and discrimination are routinely assessed, whereas comprehension-related aspects are ignored. The current study compared behavioral and electrophysiological measures of speech detection, discrimination, and comprehension. Thirty young normal-hearing native Danish speakers participated. All measurements were carried out with digits and stationary speech-shaped noise as the stimuli. The behavioral measures included speech detection thresholds (SDTs), speech recognition thresholds (SRTs), and speech comprehension scores (i.e., response times). For the electrophysiological measures, multichannel electroencephalography (EEG) recordings were performed. N100 and P300 responses were evoked using an active auditory oddball paradigm. N400 and Late Positive Complex (LPC) responses were evoked using a paradigm based on congruent and incongruent digit triplets, with the digits presented either all acoustically or first visually (digits 1–2) and then acoustically (digit 3). While no correlations between the SDTs and SRTs and the N100 and P300 responses were found, the response times were correlated with the EEG responses to the congruent and incongruent triplets. Furthermore, significant differences between the response times (but not EEG responses) obtained with auditory and visual-then-auditory stimulus presentation were observed. This pattern of results could reflect a faster recall mechanism when the first two digits are presented visually rather than acoustically. The visual-then-auditory condition may facilitate the assessment of comprehension-related processes in hard-of-hearing individuals.

Keywords

speech perception, speech audiometry, electroencephalography, cortical evoked potentials

Received 15 January 2022; Revised received 14 October 2022; accepted 1 November 2022

Introduction

Spoken language comprehension and daily-life communication require intact speech perception. Hearing loss affects approximately 466 million people worldwide, and this number is projected to rise to 900 million by 2050 (Lawrence et al., 2021). Hearing loss negatively affects speech perception (e.g., Blamey et al., 2001). For optimal hearing rehabilitation, it is thus essential to assess speech perception abilities reliably and to understand the underlying processes better.

According to Kiessling et al. (2003), speech perception encompasses hearing, listening, comprehension, and communication. Hearing is a passive function that provides access to audition via the perception of sound. It is concerned with sensing the presence of sound. That is, it reflects

detection. Listening, which is the next step, is the process of hearing with intention and attention. This step is purposeful and requires effort, as it involves the discrimination of

¹Institute of Clinical Research, University of Southern Denmark, Odense, Denmark

²Research Unit for ORL – Head & Neck Surgery and Audiology, Odense University Hospital & University of Southern Denmark, Odense, Denmark

³Department of Psychology, University of Oldenburg, Oldenburg, Germany

Corresponding Author:

Pushkar Deshpande, Institute of Clinical Research, University of Southern Denmark, Odense, Denmark; Research Unit for ORL – Head & Neck Surgery and Audiology, Odense University Hospital & University of Southern Denmark, Odense, Denmark.
Email: pdeshpande@health.sdu.dk



auditory stimuli. Comprehension is the final step before communication can take place. It is the reception of information, meaning, and intent.

Impaired hearing may affect speech perception at all levels. Speech perception abilities can be assessed using behavioral and electrophysiological measures. Behavioral measures include speech detection thresholds (SDTs) and speech recognition thresholds (SRTs), both of which are widely used in hearing clinics (Carhart, 1952). Electrophysiological measures such as event-related potentials (ERPs) measure electrical brain activity in response to a given stimulus. Commonly used ERPs for studying speech perception are the N100, P300, N400, and late positive complex (LPC) components (Sur & Sinha, 2009). The N100 is a negative deflection with a fronto-central topography that peaks around 100 ms following stimulus onset, reflecting stimulus detection (Sur & Sinha, 2009). It is an obligatory response that – unlike the mismatch negativity (Sussman, 2007) – is not influenced by attention. The P300 is a positive deflection occurring around 300–400 ms post-stimulus onset with a centro-parietal topography reflecting attention to the presented stimuli (Sur & Sinha, 2009). It can be elicited using an oddball paradigm based on a standard and a deviant stimulus (Duarte et al., 2009). The N400 is a negative deflection that occurs around 300–400 ms following stimulus onset. It is thought to occur when the brain encounters an incongruity in terms of the meaningfulness of the presented stimulus (Niedeggen & Rösler, 1999). For example, a sentence like “The dog was flying” or an arithmetic expression like “ $3 + 4 = 9$ ” will elicit an N400 response (Kutas & Federmeier, 2011). The N400 is usually followed by an ongoing positivity which is referred to as the LPC. The LPC is believed to occur when the brain tries to repair the previously detected incongruity and then concludes about the correctness or meaningfulness of the presented stimulus (Yang et al., 2019). The N400 and LPC are reflected in the difference wave derived from the ERPs for the congruent and incongruent stimuli (Niedeggen & Rösler, 1999).

In audiological practice, SDT and SRT measurements are typically performed using monosyllabic words or digits as speech signals. Under quiet conditions, normal-hearing listeners obtain SDTs and SRTs that lie approximately 10–15 dB above their pure-tone audiometric thresholds averaged across 500, 1000, and 2000 Hz. At 10 dB above the SRT, listeners can generally recognize speech effortlessly (Jerger et al., 1968). On the electrophysiological side, the N100 and P300 are well-established responses (e.g., Stapells, 2002). Usually, the N100 can be reliably evoked at a presentation level that is a few decibels above the corresponding audiometric threshold (e.g., Lütkenhöner & Klein, 2007). The P300, which is typically evoked at higher levels, has been employed for studying discrimination abilities in clinical populations (e.g., Cone-Wesson & Wunderlich, 2003). In contrast, comprehension-related aspects (e.g., semantic or syntactic processing) are neglected in the clinic so far.

While behavioral and electrophysiological measurements can provide much information about speech perception on their own, more knowledge about the neural correlates of speech detection, discrimination, and comprehension could inspire new clinical test procedures. Several studies have looked for such correlates, with most of them focusing on relatively early cortical processes. For example, Billings et al. (2013) found strong positive correlations between N100 amplitude and latency values and speech recognition at various signal-to-noise ratios (SNRs). Another study found that both N100 and P300 latencies and behavioral reaction times to speech stimuli increased when the SNR decreased (Kaplan-Neeman et al., 2006). Overall, these findings support the idea that electrophysiological measures can be used to predict speech-in-noise difficulties. Regarding later cortical speech processes, there is a scarcity of audiological motivated research studies. In the field of linguistics, researchers have relied on the N400 component to study aspects related to language, cognition, vocabulary, and emotion. For instance, Henderson et al. (2011) tried to correlate the N400 response with behavioral measures of expressive or receptive vocabulary knowledge in children. Other researchers used the LPC component to investigate foreign language competencies (Jiao et al. (2022) or the effects of aging on semantic processing (Xu et al., 2017), for example.

While detection is a prerequisite for intact speech perception, discrimination and comprehension are needed for effective communication to occur. As mentioned above, detection and discrimination are routinely assessed in audiological practice using speech audiometry, whereas comprehension-related aspects are ignored. Clinical experience shows that although many hearing-impaired individuals perform rather well in terms of speech audiometry, they report significant hearing problems in daily life. A possible explanation for this could be that they struggle with speech understanding. In this case, it could be beneficial to have effective tests for assessing detection-, discrimination-, and comprehension-related speech processing abilities.

The purpose of the current study was to address this issue. For the assessment of speech detection and discrimination, we relied on established audiological measures (i.e., SDTs, SRTs, N100 responses, and P300 responses). For the assessment of comprehension-related abilities, we devised a simple test paradigm based on congruent and incongruent digit triplets. This made it possible for us to evoke N400 and LPC responses and to measure behavioral response times to the digit triplets. Furthermore, we performed these measurements in two conditions: auditory and visual-then-auditory stimulus presentation. In the first condition, all digits were presented acoustically. In the second condition, the first two digits were presented visually and only the third one acoustically. Our motivation for including the visual-then-auditory condition was to facilitate the assessment of comprehension abilities in follow-up studies with hard-of-hearing individuals. If a hearing loss interferes with speech detection and/or discrimination

comprehension will also be affected. In the case of intact visual processing, the information needed for comprehension to occur can be provided via that channel. In that manner, issues related to inadequate stimulus audibility or supra-threshold auditory deficits can be circumvented. All the behavioral and electrophysiological measurements were performed in the presence of speech-shaped noise because speech perception is adversely affected by background noise, especially in individuals with a hearing loss (Anderson & Kraus, 2010).

Overall, the main aims of our study were as follows:

- To relate clinically applicable behavioral and electrophysiological measures of speech detection, discrimination, and comprehension to each other;
- To compare auditory and visual-then-auditory versions of the measures of speech comprehension with each other.

These aims gave rise to the following research questions:

1. Are the behavioral and electrophysiological measures correlated when studying speech detection, discrimination, and comprehension?
2. Do the auditory and visual-then-auditory versions of the measures of speech comprehension provide similar results?

In view of the general lack of audiological motivated research into brain-behavior correlations based on more complex tasks than speech detection, the current study was rather exploratory in nature.

Materials and Methods

Ethical approval for the current study was obtained from the Regional Committee on Health Research Ethics for Southern Denmark (case no. S-20190117).

Participants

Thirty participants were recruited from the University of Southern Denmark student population. The participants were native Danish speakers aged 18–30 years with no reported history of any neurological disorders and with normal or corrected-to-normal vision. They all had pure-tone hearing thresholds of maximally 25 dB HL at the standard audiometric test frequencies from 250 to 8000 Hz in both ears. The participants provided written informed consent and received financial compensation for their efforts.

Test Setup

The tests were performed in an electrically shielded sound booth. The participants were seated in a comfortable chair during the measurements. Inside the sound booth, a pair of headphones for audiometry measurements (RadioEar DD45), a pair of insert phones for electrophysiological

testing (Etymotic ER3A), a monitor, an amplifier (NeurOne Tesla, Bittium), a battery for powering the amplifier for the electrophysiological measurements, and a response pad (Cedrus, USA) were located. Outside the sound booth, the EEG main unit (NeurOne, Bittium) was placed and connected to both the stimulus presentation computer and the main computer for recording the electrophysiological data.

Behavioral Setup. The behavioral tests were performed using standard clinical equipment (Interacoustics Affinity 2.0, RadioEar DD45 headphones) and procedures (ISO, 2010).

Electrophysiological Setup. The NeurOne main unit was connected to the main computer via an ethernet cable and the presentation computer via an 8-bit trigger cable. The stimulus computer was used to present the auditory stimuli via an RME Fireface UC soundcard to the insert phones and the visual stimuli on the monitor inside the sound booth. The response pad was used for logging the participants' behavioral responses to the congruent and incongruent digit triplets. The behavioral responses were delivered to the presentation computer, where event codes were routed to the EEG computer and saved along with the EEG measurements.

The stimuli were presented using Presentation software version 21.0 (Neurobehavioral Systems, USA). The auditory stimuli were presented diotically through the insert phones. The visual stimuli were presented on the monitor inside the sound booth, which was placed 2 m in front of the participant (at 0° azimuth).

Continuous EEG signals were recorded using a custom 39-channel electrode cap. The electrode montage followed the extended 10–20 layout and included one ocular channel under the left eye to record horizontal and vertical eye movements and a vertex reference channel (Cz). Electrode impedances were kept below 5000 Ohm. The EEG signals were band-pass filtered from 0.1 to 100 Hz with 12 dB/octave roll-off and digitized at a sampling rate of 1000 Hz.

Stimulus Materials

The Dantale-I corpus (Elberling et al., 1989) was used for the speech stimuli for both the behavioral and electrophysiological measurements. More specifically, the digit materials from lists 9, 10, and 11 were used. The digits from these lists are organized in sets of triplets. The original .wav files from the Dantale-I lists were edited to prepare the stimulus material for the EEG measurements. The files were first cropped to obtain all monosyllabic digits available from the Dantale-I test lists, that is, 0, 1, 2, 3, 5, 6, 7, and 12 (see the appendix for phonetic transcriptions). The resultant files were all 400 ms in length.

The digit materials were then manipulated for the N400 and LPC measurements to yield different triplets. The triplets constituted either ascending (e.g., 1-2-3) or descending

(e.g., 3-2-1) congruent sequences or ascending or descending incongruent sequences. The incongruent sequences were divided into ‘error-by-1’ and ‘error-by-4’ conditions. For the error-by-1 triplets, the first two digits were in sequence, whereas the last digit deviated by ± 1 from the expected digit (e.g., 1-2-4 or 1-2-2 instead of 1-2-3). Similarly, the error-by-4 condition included incongruent triplets whose first two digits were in sequence, whereas the last digit deviated by ± 4 from the expected digit (e.g., 3-4-9 or 3-4-1 instead of 3-4-5). In this manner, a total of 30 distinct digit triplets were constructed. During the measurements, a total of 240 triplets were presented, that is, 120 congruent and 120 incongruent ones. Among the 120 incongruent triplets, 60 corresponded to the error-by-1 condition, while the other 60 corresponded to the error-by-4 condition.

All behavioral and electrophysiological measurements were performed in continuous speech-shaped noise with a fixed level of 50 dB HL or 67 dB(C) SPL.

Behavioral Measurements

Pure-tone audiometry was carried out using the modified Hughson-Westlake procedure (Carhart & Jerger, 1959). Left and right ears were tested individually for all the octave frequencies from 250 to 8000 Hz.

The SDT and SRT measurements were performed diotically and at least twice. A third measurement was made if the absolute difference between the first and second measurements was ≥ 5 dB. The final threshold values were calculated using the median of all measurements per participant. For the SDT measurements, the participants were instructed to press a button every time they detected a speech sound. They had to press the button for every digit they heard. The measurement started with a speech level of 50 dB HL. For every correct response, the presentation level of the stimulus was decreased by 5 dB until an incorrect response was registered. A correct response corresponded to the detection of all three digits of a given triplet, while an incorrect response corresponded to the detection of < 3 digits. At the presentation level corresponding to the first incorrect response, the participant was presented with 10 triplets, and the percentage of correctly responded digits was calculated. Next, the stimulus intensity was reduced by 5 dB. Another set of 10 triplets was presented, and the percentage of correctly identified digits was calculated. The two percentage points obtained in this manner were then connected by a line. If that line passed through the 50%-correct point on the Affinity user interface, the corresponding level was taken as the SDT. If the line did not pass through the 50%-correct point, a third measurement was made using a presentation level that was 5 dB lower. The three percentages were then connected to determine the level corresponding to 50%-correct detection performance.

The participants were asked to repeat the digits they heard for the SRT measurements. The same steps as used for the SDT measurements were then followed to determine the

presentation level at which the participant could recognize 50% of the presented digits.

Speech Comprehension Scores (Response Times). The time the participants took to log a response (via a button press) to the congruent or incongruent digit triplets from the corresponding EEG paradigm was taken as an indirect estimate of speech comprehension. Only responses to digit triplets that were correctly classified as congruent or incongruent were included in these calculations. The scores were calculated by averaging the response times to the individual digit triplets per condition and participant.

Speech-Evoked Cortical Potentials

For all EEG measurements, the speech presentation level was set to 10 dB above a given individual’s SRT to ensure good detectability and discriminability. The order of the different electrophysiological measurements was randomized across participants.

N100 and P300 Responses. The N100 and P300 responses were evoked using an auditory oddball paradigm. The digit ‘3’ was used as the standard and the digit ‘5’ as the deviant stimulus. These two digits were chosen for their clear onset characteristics. The deviants were presented with a probability of 15% (60 times) and the standards with a probability of 85% (340 times). The inter-stimulus interval was randomized in the range of 1000–1500 ms. During the measurements, the participants were instructed to attend to the deviants and to count them silently in their heads. At the end of a given measurement, they were asked to report how many deviants they had heard. The participants’ data for the P300 were only included in the analyses if they could detect the deviants with at least 80% accuracy (i.e., 48/60). The responses to the standards were used for the N100 analysis and the responses to the deviants for the P300 analysis.

N400 and LPC Responses. As mentioned above, the N400 and LPC responses were evoked using auditory and visual-then-auditory stimulus presentation (see Introduction). In the auditory condition, all digits were presented acoustically via insert phones (see Figure 1A). In the visual-then-auditory condition, the first two digits were presented visually and the third digit acoustically (see Figure 1B). Otherwise, the two conditions were identical in terms of their designs. The last digit had an onset jitter of ± 100 ms. The participants were asked to indicate the ‘correctness’ of the third digit via a button press as soon as they had reached a conclusion.

The two paradigms elicited condition-specific responses for the congruent, error-by-1, and error-by-4 stimulus conditions. To obtain the N400 and LPC waveforms, the responses to the two incongruent stimulus conditions were subtracted from the corresponding responses to the congruent stimuli. A given participant’s data were only included in the analyses

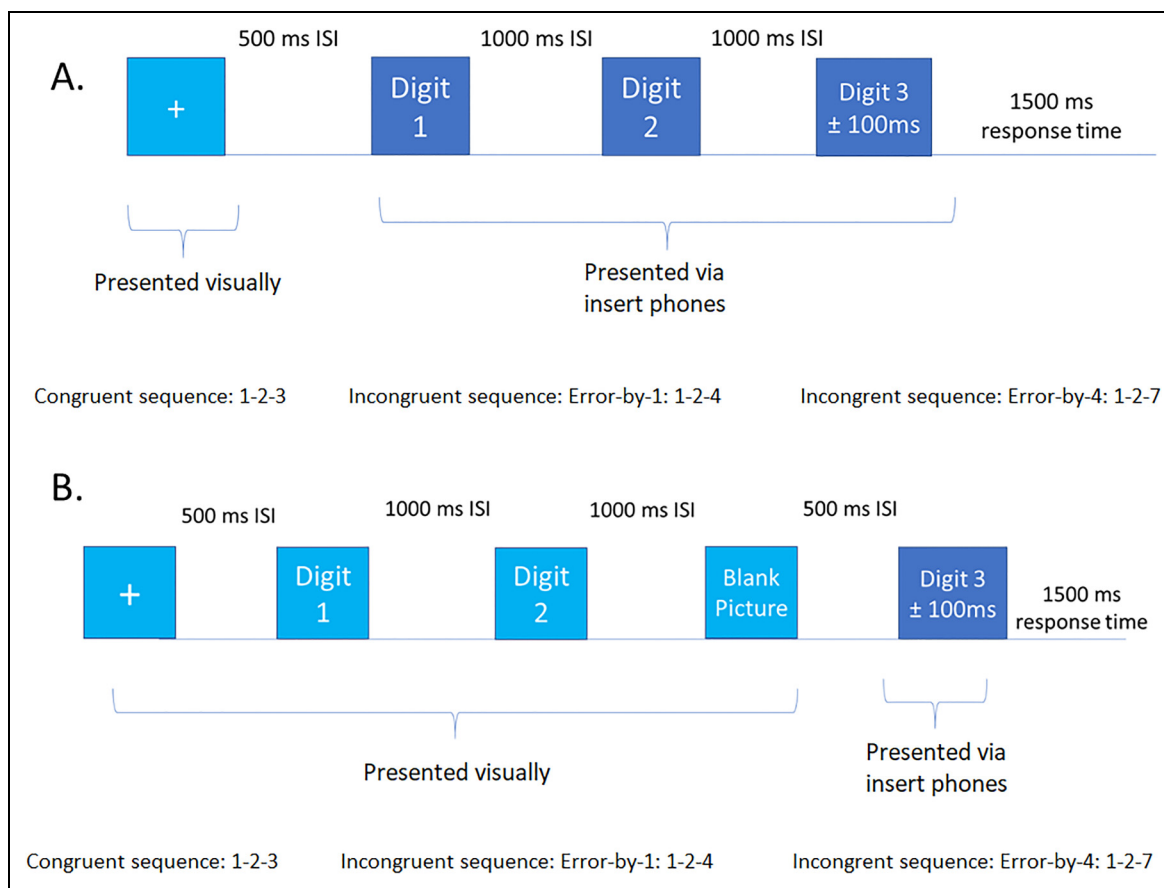


Figure 1. Illustration of the stimulus design for the N400 and LPC measurements. A: Auditory stimulus presentation. B: Visual-then-auditory stimulus presentation.

if, for both conditions (congruent and incongruent), more than 40 out of 60 responses (>68%) were correct.

EEG Data Preprocessing

The preprocessing was performed offline after the acquisition of the EEG data. EEGLAB (Delorme & Makeig, 2004) was used for this purpose. The raw data were imported into EEGLAB, and the channel locations were added. The data were then downsampled to 250 Hz. Next, the data were highpass-filtered at 1 Hz, partitioned into 1-s epochs, and subjected to typical artifact rejection by probability and kurtosis matching. An Independent Component Analysis (ICA) was then carried out. After the ICA was completed, the ICA weights were back-projected onto the raw dataset. The ICA components representing common artifacts were identified and rejected. Subsequently, the data were band-pass filtered from 1 to 10 Hz for the N100 analysis and from 0.3 to 8 Hz for the P300, N400, and LPC analysis. Different filter settings were carefully chosen based on pilot testing to isolate the corresponding EEG components as much as possible from the background noise, as recommended in the literature (e.g., Duncan et al., 2009). The filtered data were epoched to

the events of interest and baseline-corrected to -100 to 0 ms. Residual artifacts not identified by the ICA were rejected using probability-matching artifact rejection. The responses of interest were then visualized for further inspection and illustration of the data. The recording electrodes were either Cz for the N100 or Pz for the P300, N400 and LPC.

Amplitude and Latency Extraction. Based on the research literature (Bassok et al., 2009; Dickson et al., 2018; Niedeggen & Rösler, 1999; Pawlowski et al., 2018; Sur & Sinha, 2009), the following time windows were chosen for the ERP amplitude and latency extraction: 60–160 ms for the N100 component, 250–450 ms for the P300 component, 250–450 ms for the N400 component, and 450–800 ms for the LPC component. For the condition-specific responses, 300–800 ms was used. For a given EEG component, the amplitude value was calculated by taking the average of all sample values within a time window of ± 32 ms surrounding the maximum peak amplitude. The ± 32 ms time window was chosen based on an intermediate analysis. To that end, the ERP responses for the four EEG components were divided into odd and even trials. Based on the two resultant data subsets, the component amplitudes were calculated by taking the average of all

sample values in ± 0 ms, ± 16 ms, ± 32 ms, and ± 64 ms wide time windows. Following this, a correlation analysis was carried out based on the amplitudes calculated for the odd and even trials. The ± 32 -ms time window resulted in the strongest correlation between these two datasets ($r=0.68$, $p<0.05$). The ± 64 -ms window gave a slightly weaker correlation. Also, the risk of including other EEG components in the averaging increased with a greater window length. Hence, ± 32 ms was chosen for the analyses reported below.

Statistical Analyses

The statistical analyses were carried out using IBM SPSS version 26. The variables of interest were inspected for normality using Kolmogorov-Smirnov's test and quantile-quantile plots. Variables that did not fulfill the normality criterion were adjusted by removing extreme values or by log-transforming the raw data. In the case of the SRTs, normal distributions were not achievable, which is why non-parametric statistical tests were used for the analysis of these data.

To address the first research aim (see Introduction), a correlation analysis was carried out on the corresponding behavioral and electrophysiological measurements, that is, the SDT vs. N100 data (amplitude and latency), the SRT vs. P300 data (amplitude and latency), and the response times vs. the condition-specific ERPs from the digit-triplet paradigm (amplitude and latency). The latter was done separately for the data collected with auditory and visual-then-auditory stimulus presentation.

To address the second research aim (see Introduction), the results from the two stimulus presentation modes (auditory, visual-then-auditory) were compared. For the condition-specific waveforms, the amplitude and latency values were analyzed using a 2-way repeated-measures analysis of variance (ANOVA) with presentation mode (auditory, visual-then-auditory) and congruency (congruent, error-by-1, error-by-4) as factors. For the N400 and LPC responses (or difference waveforms), the amplitude and latency values were analyzed using a 2-way repeated-measures ANOVA with presentation mode (auditory, visual-then-auditory) and congruency (error-by-1, error-by-4) as factors. To analyze the response times, a 2-way repeated-measures ANOVA with presentation mode (auditory, visual-then-auditory) and congruency (congruent, error-by-1, error-by-4) as factors was performed.

To summarize, the EEG components reflecting comprehension abilities were analyzed at two levels: the condition-specific responses and the difference waveforms. For the first research aim, the condition-specific responses were compared to the response times. This was because the N400 and LPC components were derived from the difference waveforms and thus could not be directly correlated with the response times. For the second research aim, the condition-specific waveforms were analyzed in addition to the N400 and LPC components. This was because the condition-specific

waveforms allowed for a sanity check to be made, that is, if the responses to incongruent stimuli would be delayed relative to those to congruent stimuli.

Results

Behavioral Measurements

SDT and SRT Data. Figure 2 shows the SNRs corresponding to the SDTs and SRTs (measured in 50-dB-HL speech-shaped noise) for each of the 30 participants. In general, the threshold values lie within -24 to -12 dB SNR and can therefore be considered representative of normal-hearing listeners. There was a positive correlation between the SDT and SRT values ($r=0.52$, $p<0.01$).

Speech Comprehension Scores (Response Times). Table 1 shows the means and standard deviations of the response times of the participants to congruent and incongruent digit triplets. The response times obtained with auditory presentation were longer than those obtained with visual-then-auditory presentation. For both presentation modes, the response times to congruent stimuli were shorter than those to incongruent stimuli.

EEG Measurements

N100 Responses. Figure 3 shows the grand average response from the N100 analysis. As can be seen, a negative deflection with a peak latency of 111 ms, an amplitude of $1.3 \mu\text{V}$, and a fronto-central topography resembling literature N100 data (e.g., Sur & Sinha, 2009) was observed.

P300 Responses. Figure 4 shows the grand average response from the P300 analysis. As can be seen, a positive deflection with a peak latency of 341 ms, an amplitude of $10.4 \mu\text{V}$, and a fronto-central topography resembling literature P300 data (e.g., Sur & Sinha, 2009) was observed.

N400 and LPC Responses. Figures 5 and 6 show the condition-specific grand average responses and difference waveforms (N400, LPC) obtained with auditory and visual-then-auditory presentation together with the corresponding scalp topographies. As can be seen, the responses to congruent stimuli have shorter latencies than those to incongruent stimuli. The responses for the two presentation modes look similar and will be described further below.

Table 2 summarizes the extracted amplitude and latency values for the different EEG components.

Statistical Analyses

Aim 1. Regarding research aim 1 (see Introduction), no correlations were found between the SDTs and the N100 amplitude or latency values (both $p>0.05$). Neither were there any

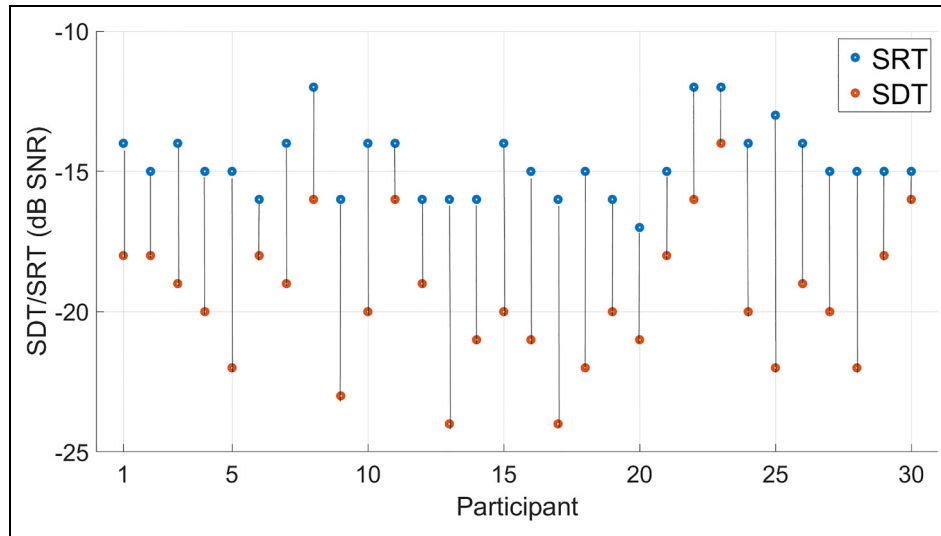


Figure 2. SDTs (red circles) and SRTs (blue circles) for each of the 30 participants. All measurements were performed in 50-dB-HL stationary speech-shaped noise.

Table 1. Response Times for the Three Congruency Conditions and the two Stimulus Presentation Modes. Std. dev. = Standard Deviation. ms = Milliseconds.

Congruency	Auditory presentation		Visual-then-auditory presentation	
	Mean (ms)	Std. dev. (ms)	Mean (ms)	Std. dev. (ms)
Congruent	717	188	521	123
Error-by-1	769	171	649	151
Error-by-4	830	178	616	127

correlations between the SRTs and the P300 amplitude or latency values (both $p > 0.05$). The response times, however, were correlated with some of the condition-specific amplitude values obtained with auditory and visual-then-auditory stimulus presentation (see Table 3). For illustrative purposes, Figure 7 shows the two clearest correlations for the error-by-1 condition and auditory presentation in the form of scatter plots.

Aim 2. Regarding research aim 2, the ANOVA performed on the N400 and LPC amplitude values revealed no effects of stimulus presentation mode or congruency (both $p > 0.05$). However, a significant interaction between stimulus presentation and congruency was found [$F(2,60) = 5.9$, $p < 0.005$]. Post-hoc testing with Bonferroni correction showed a significant difference between the congruent and error-by-4 conditions ($p < 0.005$) for auditory presentation, with larger amplitudes for the error-by-4 condition-specific responses.

For the N400 and LPC latency values, no effect of stimulus presentation mode was found ($p > 0.05$), but the effect of congruency was significant [$F(2,60) = 30.3$, $p < 0.005$].

Post-hoc testing with Bonferroni correction revealed a significant difference between the congruent and error-by-1 condition-specific responses and between the congruent and error-by-4 condition-specific responses (both $p < 0.005$). Longer latencies were observed for the error-by-1 and error-by-4 condition-specific responses. No interaction between condition and congruency was found ($p > 0.05$).

The two ANOVAs performed on the N400 amplitude and latency values revealed no effects of stimulus presentation mode or congruency (both $p > 0.05$). The same was true for the LPC amplitude and latency values. There was, however, a significant interaction between stimulus presentation mode and congruency [$F(1,29) = 9.0$, $p < 0.005$]. Post-hoc testing with Bonferroni correction revealed significant differences between the error-by-1 and error-by-4 condition-specific responses for auditory presentation ($p < 0.005$), with larger amplitudes for the error-by-1 condition-specific responses.

The analysis of the response times revealed effects of stimulus presentation mode [$F(1,30) = 47.5$, $p < 0.005$] and congruency [$F(2,60) = 65.4$, $p < 0.005$]. Post-hoc testing revealed significant differences between the congruent and error-by-1 condition-specific responses and between the congruent and error-by-4 condition-specific responses (both $p < 0.005$). The interaction between stimulus presentation mode and congruency was also significant [$F(2,60) = 23.8$, $p < 0.005$]. Post-hoc testing revealed significant differences between the congruent and error-by-1 condition-specific responses and between the congruent and error-by-4 condition-specific responses (both $p < 0.005$) but not between the error-by-1 and error-by-4 condition-specific responses ($p > 0.05$). The interaction occurred due to the error-by-1 condition leading to shorter response times than the error-by-4 condition with the visual-then-auditory presentation, while the opposite was true for auditory presentation.

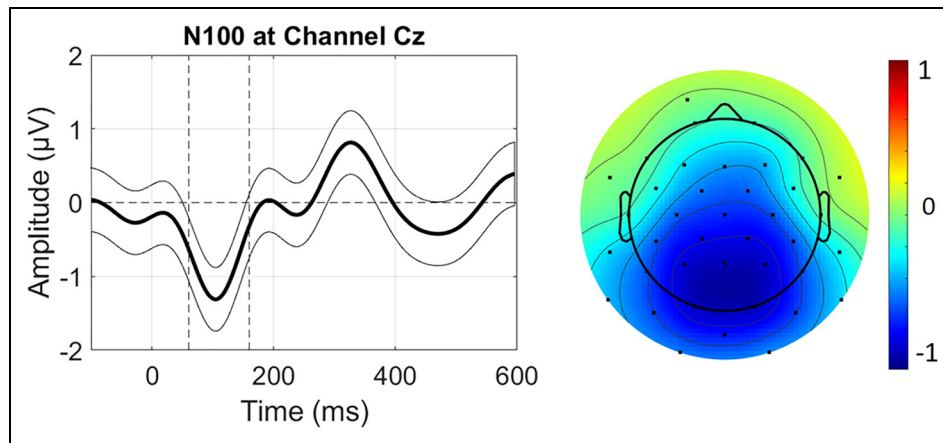


Figure 3. N100 results showing the grand average response (thick black line) with ± 1 standard deviation (thin lines) at channel Cz and the corresponding topography for the 60–160 ms time window. The black dots indicate the electrode positions. The color legend shows the amplitude in microvolts. The vertical dashed lines in the figure indicate the time window used for the analysis.

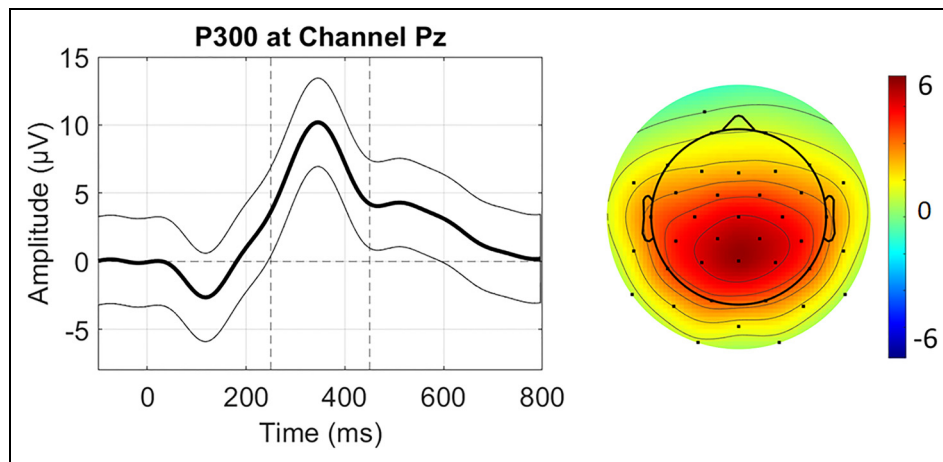


Figure 4. P300 results showing the grand average response (thick black line) with ± 1 standard deviation (thin lines) at channel Pz and the corresponding topography for the 250–450 ms time window. The black dots indicate the electrode positions. The color legend shows the amplitude in microvolts. The vertical dashed lines in the figure indicate the time window used for the analysis.

Discussion

In the current study, digit-based behavioral and electrophysiological measures of speech detection, discrimination, and comprehension were compared. Data were collected from 30 young normal-hearing adults and analyzed in terms of correlations between the behavioral and electrophysiological outcomes. Furthermore, differences between the results obtained with auditory and visual-then-auditory stimulus presentation were investigated. The analyses did not reveal any correlations between the SDTs and N100 (detection) or the SRTs and P300 (discrimination) responses. Neither were there any differences between the EEG responses obtained with auditory and visual-then-auditory presentation. However, the response times were significantly longer with auditory presentation than with visual-then-auditory presentation. In terms

of correlations, the response times were correlated with the amplitude values of the condition-specific waveforms for both presentation modes in most cases.

Measures of Speech Detection, Discrimination, and Comprehension

Digits are frequently used for audiological assessments of speech perception. In the current study, the SDT and SRT measurements ranged from -24 to -12 dB SNR (Figure 2). Overall, these thresholds cover a narrow range, which was a consequence of young normal-hearing listeners being tested. In terms of the electrophysiological responses, the N100 and P300 responses were very similar to literature data (Bidelman et al., 2013; Jasinski & Coch, 2012). The

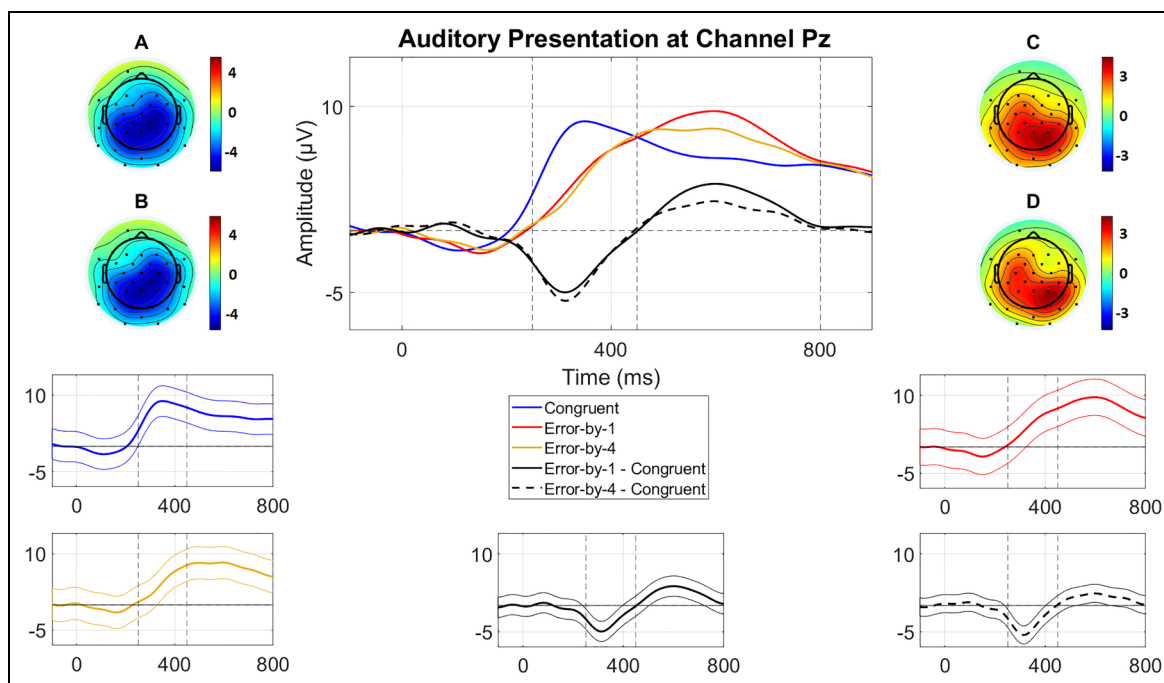


Figure 5. Condition-specific grand average responses and corresponding N400 and LPC components obtained with auditory stimulus presentation. The central panel at the top shows all grand average responses in one figure. In the five panels below, the individual grand average responses are shown together with ± 1 standard deviation. On the top-lefthand side, the topographies for the N400 component for the error-by-1 (A) and error-by-4 (B) conditions are shown. On the top-righthand side, the topographies for the LPC component for the error-by-1 (C) and error-by-4 (D) conditions are shown. The vertical dashed lines in the figures indicate the time windows used for the analyses. The black dots in the topography plots indicate the electrode positions, while the color legend shows the amplitude in microvolts.

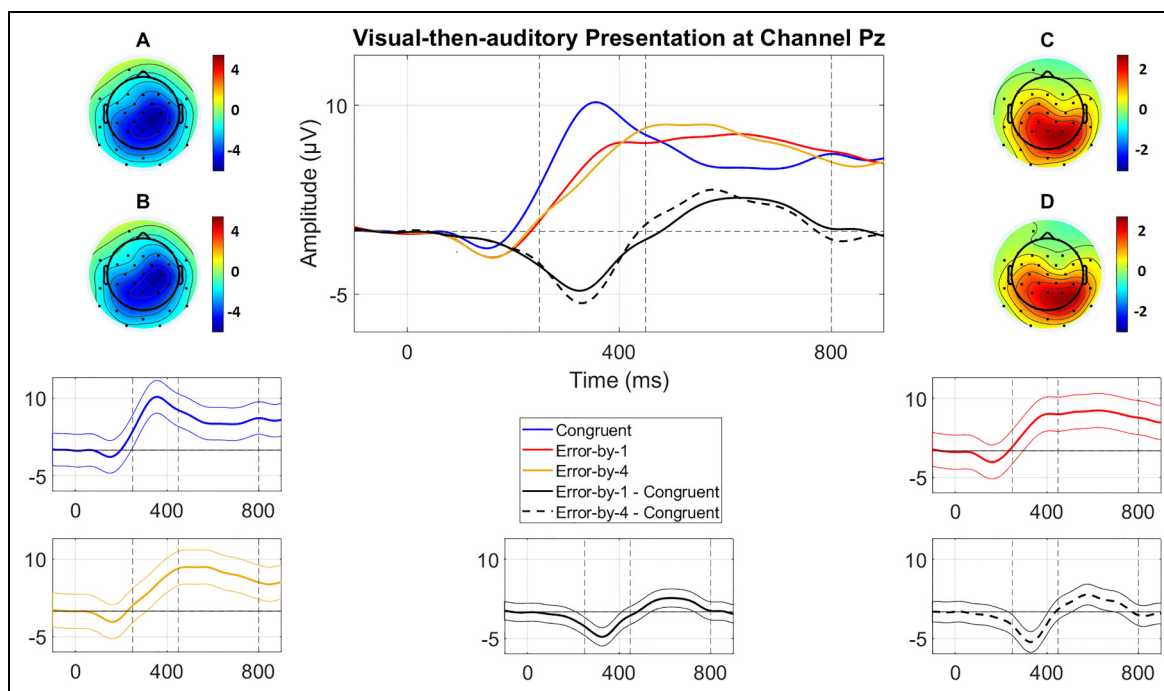


Figure 6. Condition-specific grand average responses and corresponding N400 and LPC components obtained with visual-then-auditory stimulus presentation. The central panel at the top shows all grand average responses in one figure. In the panels below, the individual grand average responses are shown together with ± 1 standard deviation. On the top-lefthand side, the topographies for the N400 component for the error-by-1 (A) and error-by-4 (B) conditions are shown. On the top-righthand side, the topographies for the LPC component for the error-by-1 (C) and error-by-4 (D) conditions are shown. The vertical dotted lines in the figures indicate the time windows used for the analyses. The black dots in the topography plots indicate the electrode positions, while the color legend shows the amplitude in microvolts.

Table 2. Mean Amplitude and Latency Values with Standard Deviations for the Condition-Specific Responses as Well as the N400 and LPC Components Obtained with Auditory and Visual-Then-Auditory Stimulus Presentation.

Component	Auditory presentation		Visual-then-auditory presentation	
	Mean amplitude (μ V) (SD)	Mean latency (ms) (SD)	Mean amplitude (μ V) (SD)	Mean latency (ms) (SD)
Congruent	10.3 (4.4)	450 (173)	11.5 (4.2)	421 (143)
Error-by-1	11.7 (4.9)	584 (78)	10.2 (4.0)	605 (222)
Error-by-4	10.3 (4.7)	569 (133)	10.5 (4.4)	547 (99)
N400, error-by-1	-5.9 (3.9)	321 (42)	-5.5 (3.3)	333 (65)
N400, error-by-4	-6.2 (3.8)	325 (48)	-6.2 (3.4)	349 (55)
LPC, error-by-1	5.0 (2.5)	652 (93)	4.2 (3.4)	637 (87)
LPC, error-by-4	3.9 (2.4)	659 (106)	4.6 (3.1)	632 (93)

Table 3. Results from the Correlation Analysis Performed on the Response Times and Condition-Specific Amplitude Values Obtained with Auditory and Visual-Then-Auditory Stimulus Presentation.

Congruency	Auditory presentation	Visual-then-auditory presentation
Congruent	$\rho = -0.63, p < 0.005$	$\rho = -0.42, p < 0.05$
Error-by-1	$\rho = -0.43, p < 0.005$	Not significant
Error-by-4	$\rho = -0.42, p < 0.05$	$\rho = -0.39, p < 0.05$

N400, however, occurred 100 ms earlier than a typical 'linguistic' N400. This is in accordance with other studies, which traced earlier responses to mathematical tasks back to a different cortical pathway being responsible for their processing (Bassok et al., 2009). Niedeggen and Rösler (1999) pointed out that simple arithmetic tasks are stored as facts rather than rule-based semantic entities. Hence, it is likely that the digit triplet task used here can be carried out faster than tasks involving linguistic processes. The N400 response was followed by the LPC, which peaked around 600 ms after stimulus onset. In general, the LPC is believed to reflect the repair of an erroneous stimulus (Niedeggen & Rösler, 1999). With regards to the current study, it appears indicative of the process that occurs when an incongruent digit sequence is presented and the brain decides about its degree of congruency.

The effects of congruency can be studied by examining the condition-specific waveforms and the response times. The congruent condition-specific waveform had a peak amplitude around 300-400 ms, whereas the incongruent condition-specific waveform had a larger amplitude around 600 ms. For the congruent condition-specific waveform, we observed a positive-going deflection resembling a P300 response, similar to what Dickson et al. (2018) described in their study. For the incongruent conditions, the positive-going deflection was delayed by approximately 200-300 ms compared to a typical P300. This could be indicative of increased processing demands incurred by the incongruent stimuli used here.

The behavioral response times were generally longer than the mean latencies of the positive-going deflections observed

in the EEG responses for the three congruency conditions. In fact, there was a rather constant delay of around 200-400 ms for auditory presentation and around 70-100 ms for visual-then-auditory presentation between the EEG latencies and the point in time when the button presses occurred (see Table 1 for mean response times and Table 2 for peak latencies). Among the condition-specific responses, mean response times were significantly shorter for the congruent conditions than for the incongruent conditions, although no difference between the error-by-1 and error-by-4 conditions was found. Overall, this indicates that the processing of incongruent stimuli needs more cognitive resources, probably because the underlying neural processes for identifying congruent and incongruent sequences are slightly different (Dickson et al., 2018).

Aim 1: Behavioral vs. EEG-Based Measures

In the current study, no correlations between the behavioral and EEG-based measures of speech detection and discrimination were observed. This could perhaps be attributed to the SDT involving active participation, whereas the N100 is an obligatory neural response that does not require attention to the stimulus, even though its amplitude can be affected by it (Gutschalk et al., 2008). Regarding the SRT and P300 measurements, a clear difference between them was the stimuli that were used. The P300 response involves paying attention to a given deviant, that is, a single digit. The SRT, on the other hand, requires more complex processing in the sense that multiple digits need to be discriminated by the participant. Another consideration is the presentation level. As the N100 and P300 responses were evoked using the same paradigm, the presentation levels used for them were identical (i.e., 10 dB above the individual SRT). Given that detection is generally possible at lower presentation levels (see Introduction), the N100 and P300 responses, as evoked here, may not be the most appropriate measures to correlate with SDTs or SRTs. In future studies, a different EEG paradigm should ideally be developed to ensure more similarity with the stimuli used for the behavioral measurements.

The negative correlations observed between the response times and EEG amplitudes for the condition-specific responses

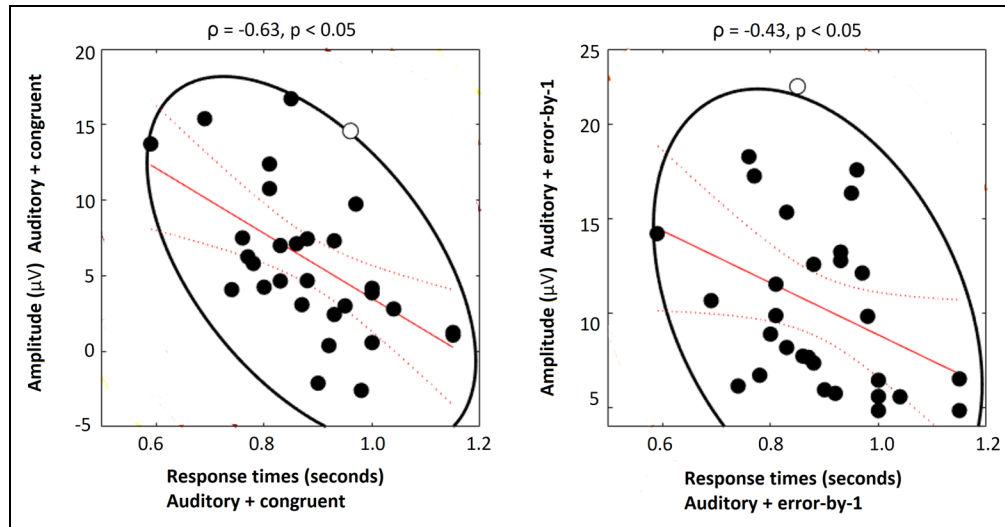


Figure 7. Scatter plots illustrating correlations between the response times and EEG amplitude values. Filled circles denote data included in the analysis, while open circles denote outliers. The solid red lines correspond to least-squares regression lines and the red dotted lines to 95% confidence intervals after outlier removal. Spearman's ρ values and corresponding p-values are also shown.

imply that longer behavioral response times are associated with reduced cortical responses. Smaller cortical potentials have been interpreted as a sign of less cognitive effort (e.g., Ghani et al., 2020). According to this view, the negative correlation observed here would suggest that longer response times are indicative of less effort, which is at odds with the findings of many other studies (e.g., Gatehouse & Gordon, 1990; Neher et al., 2014). Further research is therefore needed to resolve this issue.

Aim 2: Auditory vs. Visual-Then-Auditory Presentation

Regarding the comparison of auditory and visual-then-auditory presentation for the comprehension-related measures, the data analysis revealed a difference in the response times but not the EEG responses. It is well known that the visual and auditory pathways are different and that visual detection can be up to 100 ms faster (Pawlowski et al., 2018). Consistent with this, the mean response times obtained with visual-then-auditory presentation were significantly shorter than with auditory presentation (Table 1). For the triplet sequences to be processed correctly, all three digits need to be understood. In terms of timing, the auditory and visual-then-auditory stimulus sequences were very comparable. In both cases, the task required the recall of the first two digits to be able to assess congruency based on the third digit. When the first two digits were presented visually, a different pathway was activated than when they were presented acoustically. Because the visual pathway is generally faster (as discussed above), this could have speeded up the recall process. This may be why the response times were significantly shorter with visual-then-auditory presentation.

Broadly speaking, this finding can be compared to the benefit of visual aids for communication purposes. Some

researchers have proposed that the use of the visual pathway for processing speech cues is more efficient (Noda et al., 2014). It is also well known that individuals with a hearing loss rely on compensatory mechanisms to improve speech perception (e.g., Başkent et al., 2016). For the N400 and LPC responses, however, no influence of presentation mode was found. This would seem to suggest that auditory and visual-then-auditory presentation can be used interchangeably for investigating the cortical processes captured by these two responses. In the future, it will be interesting to apply these conditions to the study of the effects of hearing impairment and deafness. Visual processing often remains intact or is sometimes even improved in hard-of-hearing individuals (Mitchell & Maslin, 2007). With visual-then-auditory stimulus presentation, it will be possible to minimize any issues related to auditory deprivation on speech detection and discrimination. In this manner, it should be possible to assess comprehension-related abilities more clearly.

Influence of Individualized Test SNRs

In the current study, test SNRs were chosen based on individual speech audiometry results (see Methods and Materials). Early cortical potentials such as the N100 can be evoked at an intensity close to the perceptual hearing threshold (Lütkenhöner & Klein, 2007). In contrast, later cortical potentials can only be successfully evoked if a given stimulus is detected, discriminated, and adequately comprehended. Hence, by presenting the digits 10 dB above the individual SRT we ensured good speech detectability in the current study. It is well known, however, that the test SNR influences ERP measurements (e.g., Billings et al., 2009). Nevertheless,

the focus of the current study was on later cortical processes, particularly those reflecting comprehension abilities. It is meant to serve as the basis for follow-up studies with hearing aid and cochlear implant users, for whom large inter-individual differences in terms of speech perception are very common. In such cases, individualizing the test SNR based on speech audiometry is expected to be even more important.

Limitations

For the N400 and LPC paradigms, the third digit was jittered by ± 100 ms (see Materials and Methods) to avoid habituation effects. For the behavioral measurements, the Interacoustics Affinity system was used, so the digits were not jittered. However, there is some inherent variability in the length of the digit triplets used for the speech audiometry. The order of magnitude of these differences is, in fact, comparable to the jitter introduced during the N400 and LPC measurements. While a signal detection analysis would have been optimal to analyze the speech audiometry data, it was unfortunately not possible to separate the participants' responses into hits, misses, and false alarms. We do not think that different response strategies played a role in our results, although we acknowledge this as a potential limitation.

Acknowledgments

The authors thank Vivi Tran and Louise Plougheld for their help with the data collection.




Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was funded by a Ph.D. stipend from the Institute of Clinical Research, SDU.

ORCID iDs

Pushkar Deshpande  <https://orcid.org/0000-0002-7239-282X>
 Christian Brandt  <https://orcid.org/0000-0003-0672-9487>
 Tobias Neher  <https://orcid.org/0000-0002-1107-9274>

References

- Anderson, S., & Kraus, N. (2010). Sensory-cognitive interaction in the neural encoding of speech in noise: A review. *Journal of the American Academy of Audiology*, *21*(9), 575–585. <https://doi.org/10.3766/jaaa.21.9.3>
- Başkent, D., Clarke, J., Pals, C., Benard, M. R., Bhargava, P., Saija, J., Sarampalis, A., Wagner, A., & Gaudrain, E. (2016). Cognitive compensation of speech perception with hearing impairment, cochlear implants, and aging: How and to what degree can it be achieved? *Trends in Hearing*, *20*, 233121651667027. <https://doi.org/10.1177/2331216516670279>
- Bassok, M., Fisher, K., & Osterhout, L. (2009). *Conceptual integration in arithmetic is the same for digits and words: It's the meaning, stupid!* Paper presented at the Proceedings of the Annual Meeting of the Cognitive Science Society.
- Bidelman, G. M., Moreno, S., & Alain, C. (2013). Tracing the emergence of categorical speech perception in the human auditory system. *Neuroimage*, *79*, 201–212. <https://doi.org/10.1016/j.neuroimage.2013.04.093>
- Billings, C. J., McMillan, G. P., Penman, T. M., & Gille, S. M. (2013). Predicting perception in noise using cortical auditory evoked potentials. *Journal of the Association for Research in Otolaryngology*, *14*(6), 891–903. <https://doi.org/10.1007/s10162-013-0415-y>
- Billings, C. J., Tremblay, K. L., Stecker, G. C., & Tolin, W. M. (2009). Human evoked cortical activity to signal-to-noise ratio and absolute signal level. *Hearing Research*, *254*(1–2), 15–24. <https://doi.org/10.1016/j.heares.2009.04.002>
- Blamey, P. J., Sarant, J. Z., Paatsch, L. E., Barry, J. G., Bow, C. P., Wales, R. J., Wright, M., Psarros, C., Rattigan, K., & Tooher, R. (2001). Relationships among speech perception, production, language, hearing loss, and age in children with impaired hearing. *Journal of Speech, Language, and Hearing Research*, *44*(2), 264–285. [https://doi.org/10.1044/1092-4388\(2001\)022](https://doi.org/10.1044/1092-4388(2001)022)
- Carhart, R. (1952). Speech audiometry in clinical evaluation. *Acta Oto-Laryngologica*, *41*(1–2), 18–42. <https://doi.org/10.3109/00016485209124357>
- Carhart, R., & Jerger, J. F. (1959). Preferred method for clinical determination of pure-tone thresholds. *Journal of Speech and Hearing Disorders*, *24*(4), 330–345. <https://doi.org/10.1044/jshd.2404.330>
- Cone-Wesson, B., & Wunderlich, J. (2003). Auditory evoked potentials from the cortex: Audiology applications. *Current Opinion in Otolaryngology & Head and Neck Surgery*, *11*(5), 372–377. <https://doi.org/10.1097/00020840-200310000-00011>
- Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. <https://doi.org/10.1016/j.jneumeth.2003.10.009>
- Dickson, D. S., Cerda, V. R., Beavers, R. N., Ruiz, A., Castañeda, R., & Wicha, N. Y. Y. (2018). When 2×4 is meaningful: The N400 and P300 reveal operand format effects in multiplication verification. *Psychophysiology*, *55*(11), e13212. <https://doi.org/10.1111/psyp.13212>
- Duarte, J. L., Alvarenga Kde, F., Banhara, M. R., Melo, A. D., Sás, R. M., & Costa Filho, O. A. (2009). P300-long-latency auditory evoked potential in normal hearing subjects: Simultaneous recording value in Fz and Cz. *Brazilian Journal of Otorhinolaryngology*, *75*(2), 231–236. [https://doi.org/10.1016/s1808-8694\(15\)30783-7](https://doi.org/10.1016/s1808-8694(15)30783-7)
- Duncan, C. C., Barry, R. J., Connolly, J. F., Fischer, C., Michie, P. T., & Näätänen, R., ... Van Petten, C. (2009). Event-related potentials in clinical research: Guidelines for eliciting, recording, and quantifying mismatch negativity, P300, and N400. *Clinical Neurophysiology*, *120*(11), 1883–1908.
- Elberling, C., Ludvigsen, C., & Lyregaard, P. E. (1989). DANTALE: A new Danish speech material. *Scandinavian Audiology*, *18*(3), 169–175. <https://doi.org/10.3109/01050398909070742>
- Gatehouse, S., & Gordon, J. (1990). Response times to speech stimuli as measures of benefit from amplification. *British Journal of*

- Audiology*, 24(1), 63–68. <https://doi.org/10.3109/03005369009077843>
- Ghani, U., Signal, N., Niazi, I. K., & Taylor, D. (2020). ERP Based measures of cognitive workload: A review. *Neuroscience & Biobehavioral Reviews*, 118, 18–26. <https://doi.org/10.1016/j.neubiorev.2020.07.020>
- Gutschalk, A., Micheyl, C., & Oxenham, A. J. (2008). Neural correlates of auditory perceptual awareness under informational masking. *PLoS Biology*, 6(6), e138. <https://doi.org/10.1371/journal.pbio.0060138>
- Henderson, L. M., Baseler, H. A., Clarke, P. J., Watson, S., & Snowling, M. J. (2011). The N400 effect in children: Relationships with comprehension, vocabulary and decoding. *Brain and Language*, 117(2), 88–99. <https://doi.org/10.1016/j.bandl.2010.12.003>
- ISO (2010). *ISO 8253-1:2010. Acoustics - Audiometric test methods - Part 1: Pure-tone air and bone conduction audiometry*.
- Jasinski, E. C., & Coch, D. (2012). ERPs across arithmetic operations in a delayed answer verification task. *Psychophysiology*, 49(7), 943–958. <https://doi.org/10.1111/j.1469-8986.2012.01378.x>
- Jerger, J., Speaks, C., & Trammell, J. L. (1968). A new approach to speech audiometry. *Journal of Speech and Hearing Disorders*, 33(4), 318–328. <https://doi.org/10.1044/jshd.3304.318>
- Jiao, L., Duan, X., Liu, C., & Chen, B. (2022). Comprehension-based language switching between newly learned languages: The role of individual differences. *Journal of Neurolinguistics*, 61, 101036. <https://doi.org/10.1016/j.jneuroling.2021.101036>
- Kaplan-Neeman, R., Kishon-Rabin, L., Henkin, Y., & Muchnik, C. (2006). Identification of syllables in noise: Electrophysiological and behavioral correlates. *The Journal of the Acoustical Society of America*, 120(2), 926–933. <https://doi.org/10.1121/1.2217567>
- Kiessling, J., Pichora-Fuller, M. K., Gatehouse, S., Stephens, D., Arlinger, S., Chisolm, T., Davis, A. C., Erber, N. P., Hickson, L., Holmes, A., Rosenhall, U., & von Wedel, H. (2003). Candidature for and delivery of audiological services: Special needs of older people. *International Journal of Audiology*, 42(sup2), 92–101. <https://doi.org/10.3109/14992020309074650>
- Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, 62, 621–647. <https://doi.org/10.1146/annurev.psych.093008.131123>
- Lawrence, R. J., Wiggins, I. M., Hodgson, J. C., & Hartley, D. E. H. (2021). Evaluating cortical responses to speech in children: A functional near-infrared spectroscopy (fNIRS) study. *Hearing Research*, 401, 108155. <https://doi.org/10.1016/j.heares.2020.108155>
- Lütkenhöner, B., & Klein, J.-S. (2007). Auditory evoked field at threshold. *Hearing Research*, 228(1-2), 188–200. <https://doi.org/10.1016/j.heares.2007.02.011>
- Mitchell, T. V., & Maslin, M. T. (2007). How vision matters for individuals with hearing loss. *International Journal of Audiology*, 46(9), 500–511. <https://doi.org/10.1080/14992020701383050>
- Neher, T., Grimm, G., & Hohmann, V. (2014). Perceptual consequences of different signal changes due to binaural noise reduction: Do hearing loss and working memory capacity play a role? *Ear and Hearing*, 35(5), e213–e227. <https://doi.org/10.1097/AUD.0000000000000054>
- Niedeggen, M., & Rösler, F. (1999). N400 effects reflect activation spread during retrieval of arithmetic facts. *Psychological Science*, 10(3), 271–276. <https://doi.org/10.1111/1467-9280.00149>
- Noda, K., Yamaguchi, Y., Nakadai, K., Okuno, H. G., & Ogata, T. (2014). *Lipreading using convolutional neural network*. Paper presented at the fifteenth annual conference of the international speech communication association.
- Pawłowski, G. M., Ghosh-Hajra, S., Fickling, S. D., Liu, C. C., Song, X., Robinovitch, S., Doesburg, S. M., & D'Arcy, R. C. N. (2018). Brain vital signs: Expanding from the auditory to visual modality. *Frontiers in Neuroscience*, 12, 968. <https://doi.org/10.3389/fnins.2018.00968>
- Stapells, D. R. (2002). Cortical event-related potentials to auditory stimuli. *Handbook of Clinical Audiology*, 5, 378–406.
- Sur, S., & Sinha, V. K. (2009). Event-related potential: An overview. *Industrial Psychiatry Journal*, 18(1), 70. <https://doi.org/10.4103/0972-6748.57865>
- Sussman, E. S. (2007). A new view on the MMN and attention debate. *Journal of Psychophysiology*, 21(3-4), 164–175. <https://doi.org/10.1027/0269-8803.21.34.164>
- Xu, N., Hou, X., Zhao, B., Zhu, Z., & Yang, Y. (2017). Age-related temporal-spatial dynamic ERP changes during sentence comprehension. *Neuroscience Letters*, 645, 74–79. <https://doi.org/10.1016/j.neulet.2017.02.074>
- Yang, H., Laforge, G., Stojanoski, B., Nichols, E. S., McRae, K., & Köhler, S. (2019). Late positive complex in event-related potentials tracks memory signals when they are decision relevant. *Scientific Reports*, 9(1), 9469. <https://doi.org/10.1038/s41598-019-45880-y>

Appendix

The table below shows phonetic transcriptions of the digits used for the behavioral and EEG measurements.

Digit	Spelling	Pronunciation (IPA)
0	nul	[ˈnɔl]
1	en	[ˈeːn]
2	to	[ˈtʰoːʔ]
3	tre	[ˈtʰɹæːʔ]
5	fem	[ˈfɛmʔ]
6	seks	[ˈsɛgˈs]
7	syv	[ˈsyvʔ]
12	tolv	[ˈtʰɔlʔ]