Objective Listening Effort Estimation by Electroencephalographic Methods
• **Listening**
  – is the process of hearing with intention and attention for purposeful activities demanding the expenditure of mental effort (Kiessling et al., 2003)
  – is effortless for normal hearing persons in ideal listening situations

• **Hearing Impaired Persons**
  – speech comprehension in noisy situations is usually an effortful process
  – cognitive processing demands are needed to obtain an acceptable level of audibility
Listening Effort and Fatigue in School-Age Children With and Without Hearing Loss

The Influence of Sentence Intelligibility and Pupil Response as an Indicator of Effortful Listening

Auditory Scrambling, Sridhar Kalluri, Brent Edwards, Erin R. Baker, and Joost M. Fensen

Listening Effort

Objective measures of listening effort: Effects of background noise and noise reduction

Testing listening effort for speech comprehension using the individuals' background noise and noise reduction

Adriana A. Zacvak, Sophia E. Kramer, and Joost M. Fessen
Main objective:

*Development of a neurodiagnostic system, which reduces listening effort in hearing aid wearers.*

- Development of a quantitative neurophysiological model of brain structure interaction
  - Simulation of attention effects on auditory late responses (ALRs)
- Analysis of a new approach to estimate listening effort
  - Extraction of the instantaneous phase of ALR sequences
- Validation of the new measure by experimental data (gained from different studies)
Listening - processing of auditory information

1. analytical stage: decomposition into discrete sensory elements
2. synthetical stage: recombination into a perceptual stream

bidirectional, complementary mechanism

bottom-up  
exogenous,  
purely data-driven,  
**effortless**

top-down  
endogenous,  
subconsciously/  
consciously driven,  
**effortful**

endogenous, effortful modulation:
perceptual filling
phonemic restoration
serial stream scanning

degraded automatic stream formation:
missing fine structure
suppressed frequencies
missing modulations


Attention Correlates in ALRs

Attention Effects on ALRs

• N1 wave
  – influenced by exogenous and endogenous factors
  – reflects selective attention
  – amplitude is enhanced by increased attention to the stimulus


• Model of corticothalamic feedback dynamics
  – predicts endogenous and effortful corticofugal modulations of ALR single sweeps using large-scale ALR simulations
  – predicts larger (instantaneous) phase synchronization stability for an increased endogenous modulation of the bottom-up data in the range of the N1 wave


Calculation of the Synchronization Stability

- set of $M$ ALR sweeps

$$\mathcal{X} = \{ x_m \in L^2(\mathbb{R}) : m = 1, \ldots, M \}$$

- complex wavelet transform

$$\mathcal{W}_\psi : L^2(\mathbb{R}) \longrightarrow L^2(\mathbb{R}^2, \frac{da db}{a^2})$$

$$(\mathcal{W}_\psi x)(a, b) = \langle x, \psi_{a,b} \rangle_{L^2}$$

- we define the wavelet phase synchronization stability (WPSS) by

$$\Gamma_{a,b}(\mathcal{X}) := \frac{1}{M} \left| \sum_{m=1}^{M} e^{i \arg((\mathcal{W}_\psi x_m)(a,b))} \right|$$

- the WPSS was calculated for $M = 100$ target stimuli using a scale of $a = 40$
Synchronization Stability and Listening Effort

• we define the wavelet phase synchronization stability (WPSS) by

\[
\Gamma_{a,b}(\mathcal{X}) := \frac{1}{M} \left| \sum_{m=1}^{M} e^{i \arg((\mathcal{W}_\psi x_m)(a,b))} \right|
\]

• we suggest for a fixed \( a \) and \( b \) and a suitable experimental paradigm

\[
\text{Listening Effort} \propto \Gamma_{a,b}(\mathcal{X})
\]

• we define the measure Listening Effort (LE) as the mean of the WPSS in the range of the N1-wave

• the larger the WPSS, the larger the cognitive effort to solve an auditory paradigm


Comparison: Experiment vs. Model (Feasibility Study)*


Stimulus effects (different noise types, noise onsets)


Investigation of later ALR components (P300 component) and extrauditive factors


Extraction of the WPSS in a more realistic listening environment*


Extraction of the WPSS in different age groups and hearing impaired persons*

Comparison: Experiment vs. Model (Feasibility Study)
Experimental Paradigm

- auditory stimuli: pure tones (duration: 40ms)
- set of paradigms
  - "Difficult Paradigm (DP)“: three tones (1kHz, target:1.3kHz, 1.6kHz)
  - "Easy Paradigm (EP)“: three tones (0.5kHz, target:1.3kHz, 2.1kHz)
- enhancement of the entropy of the paradigms
  - tones had randomized order and randomized interstimulus intervall of 1-2 s
  - maximum effort is required to detect the target tone (response button)
- presentation at 70 dB SPL to the right ear
  (calibration, EN 60645-3: 2007, Test signals of short duration)

Subjects

- 20 subjects (27 ± 4.1 years, 11 male, 9 female)
- student volunteers, no history of hearing problems
- normal hearing thresholds (< 15 dB HL)
Data Acquisition and Inclusion Criteria

Data Acquisition

- biosignal amplifier
- sampling frequency: 512 Hz
- artifacts: rejected by an amplitude threshold of 50 µV
- bandpass filter: 1 to 30 Hz
- Ag/AgCl-electrodes
  - right mastoid (ipsilateral to the stimulus), vertex (common reference), upper forehead (ground)
  - electrodes impedances < 5kΩ

Inclusion Criteria

- identifiable waveform of the N1-P2-complex

20 included subjects
Results (I): Comparison Experiment vs. Model

Single-Sweep-Matrices

Grand Average of the WPSS

Results (II): Topological Mapping of the WPSS

Individual Result

- 64 channel recording

- WPSS for the target tone is much larger in temporal and parietal areas
  - for the difficult compared to the easy condition
  - for solving the paradigm compared to the relaxation phase

Extraction of the WPSS in a more realistic listening environment
Experimental Paradigm

• auditory stimuli: consonant-vowel syllables (female voice, adapted, calibrated (EN 60645-3: 2007))

• set of paradigms
  – „Difficult Syllable Paradigm (DSP)“:
    different plosives, same vowel
  – „Easy Syllable Paradigm (ESP)“:
    different plosives and different vowels

• maximum entropy paradigms
  – randomized order of the syllables and the interstimulus interval (1-2s)

• the paradigms were embedded in female multi-talker babble noise (SNR +5dB)

• intensity level: 65 dB SPL
Subjects and Inclusion Criteria

Subjects

- 21 subjects (25 ± 3.52 years, 12 male, 9 female)
- student volunteers
- no history of hearing problems
- normal hearing thresholds (< 15 dB HL)

Inclusion Criteria

- identifiable waveform of the N1-P2-complex
- 80% correctly detected target syllables

18 included subjects

Subject’s instruction

- detection of the target stimuli (response button)
- ignoring the background noise
Results (I): Grand Average of the WPSS

Grand Average of the WPSS (different scales)

<table>
<thead>
<tr>
<th>scales</th>
<th>frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>60</td>
<td>4.27 Hz</td>
</tr>
<tr>
<td>50</td>
<td>5.12 Hz</td>
</tr>
<tr>
<td>40</td>
<td>6.40 Hz</td>
</tr>
<tr>
<td>30</td>
<td>8.53 Hz</td>
</tr>
<tr>
<td>20</td>
<td>12.80 Hz</td>
</tr>
</tbody>
</table>

black areas: significant difference (p<0.05)

Results (II): Grand Average of the WPSS

Grand Average of the WPSS
(over all the 18 subjects, M=70 sweeps, scale a=40)

Bernarding C, Corona-Strauss Fl, Latzel M, Strauss DJ.
Results (III): Individual Results

Auditory Late Responses

- "good" ALR example

Synchronization Stability

- "bad" ALR example

Extraction of the WPSS in different age groups and hearing impaired persons
Experimental Paradigm

- auditory stimuli: consonant-vowel syllables (female voice, adapted, calibrated (EN 60645-3: 2007))

- set of paradigms
  - „Difficult Syllable Paradigm (DSP)“:
    different plosives, same vowel
  - „Easy Syllable Paradigm (ESP)“:
    different plosives and different vowels

- maximum entropy paradigms
  - randomized order of the syllables and the interstimulus interval (1-2s)

- intensity level
  - 65 dB SPL (normal hearing, mild hearing impaired subjects)
  - adjusted intensity (moderate hearing impaired subjects)
A total of 94 subjects participated in the study

- 24 young subjects (13 m/11 f), normal hearing levels
  (ynh; aged 21 to 35 years, mean age: 25.25 ± 4.01 years)
- 21 middle-aged subjects (9 m/13 f) normal hearing levels
  (manh; 40 to 60 years, mean age: 51.15 ± 5.64 years)
- 25 middle-aged subjects (15 m/10 f) with mild hearing loss
  (mild; 45 to 63 years, mean age: 51.87 ± 5.82 years)
- 24 middle-aged subjects (9 m/15 f) moderate hearing loss
  (mod; 43 to 57 years, mean age: 51.12 ± 5.53 years)
Individual Results

Normal hearing subject
(PTA: 18.75±24.28, 57 years)

Mild hearing impaired subject
(PTA: 33.75±9.31, 55 years)

Moderate hearing impaired subject
(PTA: 45±21.98, 50 years)

WPSS

ALRs
Summary and Conclusions

- Listening effort correlates can be assessed by the instantaneous phase synchronization stability of auditory late responses in:
  - different listening conditions
  - different age groups
  - normal hearing persons
  - hearing impaired persons

WPSS was extracted in a total of 160 subjects.

- Strong theoretical basis for the experimental results:
  - Development and validation of a new neuroscientific/corticothalamic model of selective attention/listening effort.

- WPSS is a solid measure (compared to the fragile amplitude information).
Main objective:

*Development of a neurodiagnostic system, which reduces listening effort in hearing aid wearers.*

- analysis of a new approach to estimate listening effort
  - extraction of the instantaneous phase of the oscillatory EEG
  - calculation of the angular entropy
Let \( \psi_{a,b}(\cdot) = |a|^{-1/2} \psi((\cdot-b)/a) \) where \( \psi \in L^2(\mathbb{R}) \) is the wavelet with \( 0 < \int_{\mathbb{R}} |\Psi(\omega)|^2 |\omega|^{-1} \mathrm{d}\omega < \infty \) (\( \Psi(\omega) \) is the Fourier transform of the wavelet), and \( a, b \in \mathbb{R}, a \neq 0 \). The wavelet transform \( \mathcal{W}_\psi : L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R}^2, \frac{dadb}{a^2}) \) of a signal \( x \in L^2(\mathbb{R}) \) with respect to the wavelet \( \psi \) is given by the inner \( L^2 \)-product \( (\mathcal{W}_\psi x)(a, b) = \langle x, \psi_{a,b} \rangle_{L^2} \). The instantaneous phase of a signal \( x \in L^2(\mathbb{R}) \) can be achieved by taking the complex argument from the complex wavelet transform with the signal: \( \phi_{a,b} = \text{arg}(\mathcal{W}_\psi x)(a, b) \). We divided the phase values into \( N \) bins and each bin has the probability \( p_i, I = \{-\pi, -\pi + \frac{\pi}{10}, \ldots, +\pi\} \), with \( \sum_i^N p_i = 1 \). Then, the normalized angular entropy can be defined by

\[
H = -\sum_{i \in I} \frac{p_i \cdot \ln p_i}{\ln N}.
\]
Focus: Angular Entropy as a possible new measure for the quantification of large-scale listening effort correlates.

- Assumption: Angular entropy reflects phase synchronization effects of the ongoing activities due to an increased attention on the relevant (speech) signal.

- We expect that smaller values of the angular entropy reflect a more "ordered" process of the phase distribution.
**Data acquisition**

- EEG, 32 channels, fs=512Hz, Cz reference, forehead ground, impedances <5kOhm

**Data analysis**

- Signals were filtered (0.5-40Hz)
- Extraction of the EEG data during the presentation of the sentence (trigger signal)
- Artifacts were rejected if either the maximum amplitude threshold exceeded ±70μV or the standard deviation exceeded ±40μV within a moving time window (window size: 50ms)
Study I

Paradigm 1 (PD1):
This paradigm consisted of the original sentences, presented without any background noise.

Paradigm 2 (PD2):
The second paradigm was built by the original sentences, embedded in the speech simulating background noise (SNR of 0dB, cocktail-party environment).

Paradigm 3 (PD3):
25% of the information of each sentence was removed. The sentences were also embedded in the speech simulating background noise.
Study I

Stimulus presentation
- monaural (right side), intensity 65dB SPL
- duration of the complete experiment: approx. 40min

Subjects
- 13 normal hearing (<15dB HL) subjects (mean age 24.28±3.12 yrs, 7 F/ 6 M).
- 12 included subjects (too many EEG artifacts)

Instruction of the subjects
- First part of the experiment (more active condition)
  - pay attention to the sentence, not to the distracting background noise,
  - to follow each sentence/try to understand each sentence,
  - to repeat the last word of each sentence after the signal (response was noted by the experimenter).
  - After the presentation of each paradigm: subjective rating of the required listening effort.

- Second part of the experiment (more passive condition)
  - The subjects were instructed to relax, not to listen actively to the sentences, to repeat if the signal was presented (signals were randomly presented).

Listening Effort Scale

<table>
<thead>
<tr>
<th>mühelig</th>
<th>sehr wenig anstrengend</th>
<th>wenig anstrengend</th>
<th>mittelgradig anstrengend</th>
<th>deutlich anstrengend</th>
<th>sehr anstrengend</th>
<th>extrem anstrengend</th>
</tr>
</thead>
</table>


Power Spectrum Analysis
In order to compare the new proposed measure with traditional analysis methods, we analyzed also the power spectrum of the EEG data. For the calculation of the power spectrum of each band, the Fourier transform was applied. The following frequency bands were analyzed:
- Theta (4-8Hz)
- alpha (8-12Hz)
- beta (12-30Hz)

Study I: Results

Listening Effort Scale
• a number was added to each level of the scale (ranging from 1 (very little effort) to 7 (extreme effort)).
• For each paradigm, the following ratings were obtained (mean ± s.d.):
  • PD1: 1.16±0.38 (no effort),
  • PD2: 2.75±0.96 (very little effort - little effort),
  • PD3: 3.33±1.15 (little effort – moderate effort).

Speech Intelligibility
• All subjects could repeat correctly 100% of the last words of the first paradigm (PD1; original sentences).
• For the other paradigms, the performance was only slightly reduced:
  • PD2 a mean of 97.91±1.62%
  • PD3 a mean of 98.41±3.17%.
Study I: Results

**EEG Power Spectrum**

- The power for each frequency band and condition (A and B) is illustrated as a bar graph (from left to right (light grey to black): theta-, alpha- and beta band).

- None of the power spectra showed a statistical significance ((one-way) ANOVA, p>0.05) between the two conditions.

Results of the EEG power spectrum analysis for PD1 (representative for all paradigms; p>0.05). The three main blocks of the power spectrum correspond to one frequency band (left to right (light grey to black): theta-, alpha- and beta band). Each bar of one block corresponds to one condition (left: condition A, right: condition B).
Angular Entropy

- Results of the ANOVA for the analysis of the angular entropy (cond. A vs. B) for each paradigm
- Only electrode positions are depicted, where the difference of the angular entropy between the conditions was significantly different (p<0.05).
- The angular entropy was always significantly enhanced for condition B (relaxing part) compared to condition A (solving the paradigm) for the shown electrode positions.
- Most of the involved electrodes are located in the frontal areas (e.g. F3, F4, FC4) within the theta range.

<table>
<thead>
<tr>
<th>frequency band</th>
<th>scale</th>
<th>PD1</th>
<th>PD2</th>
<th>PD3</th>
</tr>
</thead>
<tbody>
<tr>
<td>β</td>
<td>10</td>
<td>FC4</td>
<td>F3,FP2</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>O3, F4</td>
<td>FC4</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>F4, FC4</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>F4</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>22</td>
<td>FP2</td>
<td>-</td>
<td>FC5</td>
</tr>
<tr>
<td>α</td>
<td>24</td>
<td>-</td>
<td>PO4</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>26</td>
<td>-</td>
<td>-</td>
<td>FC3</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>-</td>
<td>T7</td>
<td>F8, F06</td>
</tr>
<tr>
<td></td>
<td>32</td>
<td>-</td>
<td>F3</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>34</td>
<td>T7</td>
<td>FP1</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>36</td>
<td>-</td>
<td>P4, F3</td>
<td>CP2</td>
</tr>
<tr>
<td></td>
<td>38</td>
<td>-</td>
<td>P8</td>
<td>-</td>
</tr>
<tr>
<td>θ</td>
<td>40</td>
<td>-</td>
<td>-</td>
<td>T7, P8</td>
</tr>
<tr>
<td></td>
<td>44</td>
<td>F3</td>
<td>F3,FP1,F2</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>46</td>
<td>-</td>
<td>FP2</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>48</td>
<td>F7, FC4</td>
<td>F4</td>
<td>T7</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>F7</td>
<td>FC2</td>
<td>FC4, P8</td>
</tr>
<tr>
<td></td>
<td>52</td>
<td>F8</td>
<td>CP2</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>54</td>
<td>P7</td>
<td>OZ</td>
<td>P4</td>
</tr>
<tr>
<td></td>
<td>56</td>
<td>-</td>
<td>-</td>
<td>CP6</td>
</tr>
<tr>
<td></td>
<td>58</td>
<td>-</td>
<td>CP2</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>PO3, T7</td>
<td>PO4</td>
<td>-</td>
</tr>
</tbody>
</table>
Angular Entropy

- The grand averages of the angular entropy (over all the 12 included subjects) for two scales
  - $a=40$ (corresponds to the $\alpha/\theta$-border)
  - $a=48$ (corresponds to the center of the theta band)
- The angular entropy is enhanced for condition B (relaxing part) compared to condition A (solving the paradigm).
- This means, that the angular phase for the effortful condition is not uniformly distributed, i.e. the phase is more "ordered" and synchronized in these conditions.
Study II

Objective

Estimating neural correlates of listening effort in realistic hearing aid settings by means of oscillatory EEG activity
• Speech material:
  – German sentence test (Oldenburger Sentence Test (OLSA)) Each sentence has a five-words structure: subject - verb - numeral - adjective - object
  – filtered (1/3 octave band, 18 channels, freq. range 250-6kHz (DIN EN 61260:1995)) and attenuated speech material to simulate a moderate hearing loss (N3, IEC 60118-15:2008).

• Background noise:
  – „multitalker babble noise“
  – consists of 5 speech sequences (composed by 150 Sentences of the OLSA Test, male speaker). Each sequence was shifted in time by one sentence and finally, the sequences were added together.
  – the babble noise starts 5s before the presentation of the sentences (Luts et al., 2010)
Design: Multitalker Babble noise

sequence 1  sentence 1  sentence 2  sentence 3  sentence 4  sentence 5  sentence 6

sequence 2

sequence 3

sequence 4

sequence 5
Experimental Paradigms

Auditory Paradigms (duration: 5min each)

• Test Ia: Presentation with hearing aid (Control Measurement, „First Fit“) for acclimatization
  – UCL 0.5, 1, 2, 4kHz = 80dB HL
  – Mastergain: 18dB
  – Speech enhancement 8-15dB
• Test II: Gain reduction
  – Mastergain: 9dB
• Test III: Gain reduction in the speech area,
  – resulting in 11dB Mastergain
• Test IV: Reduction: Speech enhancement
  – to 4-5dB
• Test V: Increase: Speech enhancement
  – to maximum level 12-24dB
• Test VI: Omnidirectionale microphone
• Test VII: Automatic/ TruEar
• Test Ib: Presentation with hearing aid (Control Measurement, „First Fit“, c.f.Test Ia)
Study Design & Setup

- Stimuli presentation:
  - speech material: 60 dB SPL
  - background noise: 60 dB SPL (SNR: 0 dB)

- Loudspeakers arrangement (S=Signal, N=Noise):
  - Test Ia,b-IV.: S 0° N 0°
  - Test VI-VII.: S 0° N 180°

- Auditory task
  - Repetition of the last word (after tone- signalization (1kHz))
  - Evaluation of listening effort by a subjective scale after each test

<table>
<thead>
<tr>
<th></th>
<th>1 mühelos</th>
<th>2 sehr wenig anstrengend</th>
<th>3 wenig anstrengend</th>
<th>4 mittelgradig anstrengend</th>
<th>5 deutlich anstrengend</th>
<th>6 sehr anstrengend</th>
<th>7 extrem anstrengend</th>
</tr>
</thead>
</table>
Data acquisition and Analysis

• Continuous EEG recording (Electrode positions according to 10/20 system, 16 channels)

• Extraction and analysis of the instantaneous angular phase entropy as possible objective measure of listening effort.
Normal hearing subjects

- 15 subjects (7m/8f)
- Mean age: 24.8±2.59 years

Inclusion criteria:
- 50 % of correctly recognized-repeated words in test 1b (control measurement)
- 85 % artefactfree EEG data

Included subjects:
- 14 subjects (7m/7f)
- Mean age: 24.78±2.69 years
Linear Fit: Different Approaches

Approach 1
- „pseudo“ frequency $f_a$ (wavelet transform) to fit the line
- Res. structure = unwrapped_phase - 2$\pi$t$f_a$

Approach 2
- Linear fit
- Res. structure = unwrapped_phase – linear fit
- c.f. phase modulated signals

Approach 3
- Calculation of the instantaneous frequency
- Fit a line using the mean of the instantaneous frequency
- Res. structure = unwrapped_phase - 2$\pi$mean(instfreq)$t$
Examples Linear Fit: Different Approaches
Instantaneous phase

unwrapped phase-linear fit=residual structure

Results: Subjective LE scaling

- Subjective LE-Scale (median)
- Repeated words (mean)

Charts showing subjective LE scaling and repeated words across different test conditions.
Future Work

Multidimensional data analysis (e.g. Parafac)

– decomposing EEG-data

(Channel x Frequency x Subjects x Paradigm)

– extracting significant activities from EEG
  • region of interest (electrode channels)
  • scales/frequencies

Morten Mørup, Lars Kai Hansen, Christoph S. Herrmann, Josef Parnas, and Sidse M. Arnfred, Parallel Factor Analysis as an exploratory tool for wavelet transformed event-related EEG
2005
Thank you for your attention!