Title: Interpersonal synchronization as an objective measure of listening engagement

Authors: Lotte Lambrechts, Bernd Accou, Jonas Vanthornhout, Bart Boets, Tom Francart

Abstract

Listening engagement has received increasing attention for its role in effective communication and knowledge transfer. It encompasses a state of deep listening in which a person not only pays attention to a stimulus but is fully absorbed in it. Behavioral measures have been used to assess listening engagement; however, they have limitations and do not fully capture this multidimensional state. Measuring listening engagement objectively using brain and body signals is a compelling alternative. In particular, interpersonal synchronization—observed through synchronized physiological responses among listeners—offers a promising objective measure of this phenomenon. We measured brain, heart rate, and electrodermal activity synchronization between people who listened to engaging and non-engaging stories. As a ground truth, we also measured subjective listening engagement through a questionnaire. Results showed higher synchronization when participants listened to an engaging stimulus compared to a non-engaging one across different modalities of objective response. Furthermore, we found a significant correlation between synchronization measures and subjective engagement ratings. These results confirm that synchronization may serve as a reliable measure of listening engagement.

INTRODUCTION

When listening to an auditory stimulus, sound is processed and interpreted by the brain. It has already been extensively demonstrated that the brain is capable of tracking the complex dynamics inherent in speech (Brodbeck & Simon, 2020; Crosse et al., 2021; Vanthornhout et al., 2018). Neural tracking of speech has been proven to be a reliable and valid measure of natural speech processing and perception, allowing us to objectively measure whether someone can understand speech or not (e.g. Iotzov & Parra, 2019; Vanthornhout et al., 2018). Additionally, it has been shown that neural tracking is modulated by attention to speech (e.g., Ding & Simon, 2012; Vanthornhout et al., 2019). With this paper, we will go even a step further by exploring listening engagement and building further on the currently existing objective measures for listening engagement. Listening engagement involves a state of deep listening where a person not only pays attention to a stimulus but is completely absorbed by what they are listening to (Busselle & Bilandzic, 2009).

Busselle & Bilandzic (2009) were able to evaluate how listening engagement is intertwined with comprehension, attentional focus, emotional engagement and, if available in the narrative, narrative presence (Busselle & Bilandzic, 2009; Kuijpers et al., 2014). This involves the experience of losing awareness of the present world and feeling as if you are part of the narrative's world (transportation). In the model of listening engagement, the concept is approached from a cognitive perspective, where higher engagement corresponds to greater recruitment of cognitive resources (Herrmann & Johnsrude, 2020). Additionally, listening engagement is hypothesized to be modulated by one's intrinsic motivation (Pichora-Fuller et al., 2016), the amount of available auditory attention (Massar et al., 2016; Seli et al., 2015), someone's specific interest in a topic (Tobin & Guadagno, 2022) and someone's hearing abilities and fatigue (Herrmann & Johnsrude, 2020; Irsik et al., 2022).

Taken together, this suggests that listening engagement involves a crucial balance between conscious and unconscious processing of sound. On the one hand, engaged listening consists of sensory processing of the sound stimulus influenced by auditory attention and listening motivation. On the other hand, listening engagement is an involuntary process in which individuals may lose track of time and experience reduced awareness of their surroundings. Listening engagement is important because

it relates to high-quality speech processing and communication. High auditory engagement stimulus can yield several beneficial outcomes, including enhanced memory retention (Cohen & Parra, 2016; Song et al., 2021), increased listening enjoyment (Bevilacqua et al., 2018; Herrmann & Johnsrude, 2020) and a state of flow (Busselle & Bilandzic, 2009; Cziksentmihalyi, 1990).

So far, mostly behavioral measures have been used to evaluate listening engagement (Busselle & Bilandzic, 2009; Hannum & Simons, 2020; Kuijpers et al., 2014). These are well-validated in large cohorts of people and are indicative of how one experiences listening engagement. Furthermore, they have been used previously in studies and have proved to be a reliable measure of listening engagement (Irsik et al., 2022). Nevertheless, relying on behavioral, self-reported measures of listening engagement holds uncertainty as this can result in a biased outcome and involve only limited sampling. An objective measure of listening engagement could significantly contribute to these behavioral measures, as it is less prone to bias. Previous attempts have been made to objectively measure (correlates of) engagement. Previous research has shown that when people consciously process engaging stories, their bodies tend to synchronize (Madsen & Parra, 2024; P. Pérez et al., 2021). We refer to this phenomenon as interpersonal synchronization. It's been shown that interpersonal synchronization can arise at different levels of the body (Madsen & Parra, 2022; Palumbo et al., 2017), e.g., brain synchronization, heart rate synchronization, and electrodermal activity synchronization. Interpersonal synchronization of the body in response to listening has already been found to arise in both conversational interactions (back and forth) (Kinreich et al., 2017) as well as in situations where people simply listened to a speaker or an auditory stimulus (Madsen & Parra, 2022, 2024; A. Pérez et al., 2017; P. Pérez et al., 2021). Following this, there is still speculation on where this occurrence of synchronization comes from. Bevilacqua et al. (2018) and Madsen & Parra (2022, 2024) claim that synchronization is a consequence of conscious and attentive processing of an engaging stimulus. Both argue that it is driven by similar cognitive processing and neural oscillations in response to engaging content. For example, attentive listening to an engaging stimulus modulates brain responses, which are shared among people. This shared brain activity is reflected in synchronization. Perez et al. (2021) support this view, stating that heart rate synchronizes in people listening to the same engaging stimulus, as it is influenced by various cognitive and conscious processes. In contrast, other studies suggest that heart rate may also be modulated by the specific emotions elicited by the stimulus, moving beyond a purely cognitive perspective (Deng et al., 2024). Given that body responses, like changes in heart rate and electrodermal activity, are often related to emotions and arousal (Anthiyur Aravindan et al., 2023; Kołodziej et al., 2019; Sánchez-Reolid et al., 2022), this hypothesis seems plausible.

Synchronization is not solely driven by physical interaction with others (Palumbo et al., 2017; A. Pérez et al., 2017), though such interactions may enhance the observed effects. Parra et al. (2021, 2022, 2024) have found significant interpersonal synchronization even when participants did not listen to the stimulus at the same time, suggesting that synchronization is not uniquely driven by interpersonal social interaction but also simply by being similarly engaged or excited by an auditory stimulus. Interestingly, they also found a decrease in synchronization when they distracted their participants from listening by giving them a cognitively engaging task (distracted condition). This suggests that synchronization may be influenced by attention. However, to link synchronization to listening engagement, it is necessary to compare synchronization across conditions where engagement levels, rather than attention, are explicitly varied—such as by having participants listen to stimuli that yield significantly different levels of engagement. In addition, previous researchers have, or not included a behavioral ground truth (Irsik et al., 2022; Madsen & Parra, 2022, 2024), or they included one that was not directly relevant for engagement (Chatterjee et al., 2021). This makes it challenging to ascertain whether these stimuli fulfil their goal of being engaging or not.

This study aims to employ an objective synchronization measure to distinguish between engaging and non-engaging conditions and explore the relationship between synchronization and listening engagement. We tackle some limitations found in earlier studies by meticulously selecting engaging and dull stimuli appropriate for the participant group. Additionally, we incorporated a subjective measure of engagement. We hypothesize that interpersonal synchronization increases when individuals are engaged with the stimulus and that the level of synchronization can predict self-rated listening engagement.

Methods

Participants

Participants applied to participate in the experiment together with someone they knew well (e.g., friends, romantic partners, family…). We included 40 participants. Participants had a mean age of 21 years, ranging from 18 to 25 years old. Of the 40 participants, 4 were male. All participants had normal hearing, as confirmed by a digit triplet screening test (SRT = -11.7 dB SNR) (Denys et al., 2019) and/ or a tonal audiogram (thresholds at all audiometric frequencies below 25 dB HL), with a mean Fletcher index of 13 ± 6 dB HL. Other exclusion criteria were concentration or attention-related problems, neuropathological disorders, communication disorders and language disorders. This study was approved by the Medical Ethics Committee UZ Leuven/Research (KU Leuven) with reference S68492.

Setup

Participants were tested in a soundproof room (Faraday cage). They sat next to each other, with a distance of approximately 1.35 meters between them. As you can see in figure 1, participants' chairs were turned towards each other (and towards the loudspeaker) by 45°, so they sat at a 90° angle to each other. A loudspeaker and monitor were positioned 1.5 meters in front of the participants. The setup was calibrated, and stimuli were presented at 60 dB A, i.e. at everyday conversation level.

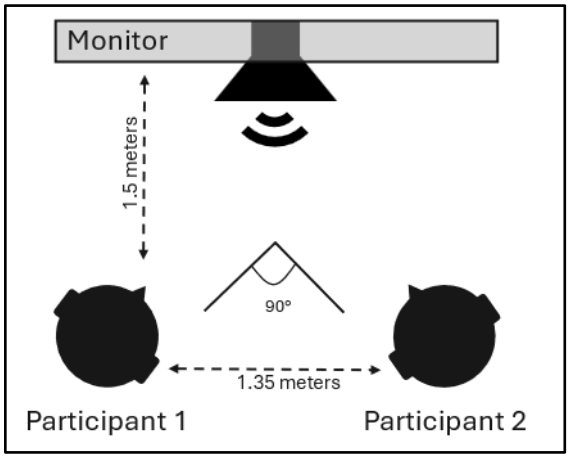

Figure 1: Graphical representation of our experimental setup

Procedure

In this experiment, we will measure brain/ body measures as well as subjective listening engagement in participants who listened to various stimuli. First, we evaluated interpersonal closeness using the 'Unidimensional relation closeness questionnaire' (URCQ) by Dibble et al. (2012), which contains 13 questions in which participants had to evaluate the closeness they felt towards the other participant. Prior to the experiment, we used a shortened version of the self-disclosure paradigm (Aron et al., 1997). This is a paradigm developed and validated experimentally to generate closeness between

people. Additionally, we included a baseline condition in which we measured all the brain/body responses at rest.

The actual experiment consisted **of three different conditions.** In all of the conditions, an auditory stimulus was presented:

1. Engaging condition
2. Nonengaging condition
3. Distracted condition

In these first two conditions, participants were instructed to listen attentively to the speech stimulus while stimulus engagement was manipulated. We did this by presenting stimuli with different levels of listening engagement, i.e., engaging and nonengaging stimuli. In the distracted condition, the participants were encouraged not to pay attention anymore to the auditory stimulation. They were given riddles to solve during the stimulus presentation. The order in which the riddles were given was different for each participant. In the engaging condition, participants listened to an engaging stimulus, i.e., a stimulus they would find interesting, that is easy to follow and that captured people's focus naturally. More specifically, as an engaging stimulus, we selected an episode, namely episode 19 'Maddie McCann' of a popular Flemish podcast 'De Volksjury' (second place on "Spotify Wrapped" of 2024), allowing us to assume that it would resonate with a large group of people. The engaging stimulus was 46:29 minutes in total. As a nonengaging stimulus, we self-recorded excerpts from Dutch debates in the European parliament. This stimulus was 25:19 minutes in duration. Both the engaging stimulus as well as the nonengaging one were split into two, so we had two engaging conditions and two nonengaging.

We also used a nonengaging stimulus. The specific stimulus was also a self-recorded excerpt of the European Parliament of 12:42 minutes. The goal of this condition was to bring people in a state, where not only no attention but also no engagement was experienced, serving as an excellent control. This would help us to better understand synchronization as a measure of engagement, disentangling it from a measure of attention.

The full experimental procedure and order of the protocol can be found in **Table 1**. Each dyad participated in the same experimental procedure, with the order of conditions (engaging, nonengaging and distracted) randomized across dyads to mitigate potential training or fatigue effects.

Table 1: Overview of the entire protocol. Participants always started with the baseline condition. The other conditions were presented in random order but were always administered in their entirety.

ENGAGING CONDITION	NONENGAGING CONDITION	DISTRACTED CONDITION
Engaging podcast part 1 (23:27 minutes)	Nonengaging recording 1 (12:37 minutes)	Nonengaging stimulus and riddles (12:42 minutes)
Engaging podcast part 2 (23:02 minutes)	Nonengaging recording 2 (12:42 minutes)	
Behavioral questionnaires	Behavioral questionnaires	

Brain/ Body measures

We included the following objective measures: brain response as measured through electroencephalography (EEG), electrodermal activity (EDA) and heart rate (HR). We used a BioSemi ActiveTwo system (BioSemi - The Netherlands) to record the EEG, HR, and EDA. We fit each person with a headcap and applied 64 electrodes in the conventional 10-20 placement. For HR, we measure electrocardiography using two flat electrodes applied at the left lumbar region and below the left

clavicle, which are application locations that have been used before (Madsen & Parra, 2022, 2024; Metshein et al., 2022). EDA was measured using 2 passive Nihon Kohden electrodes, applied on the pointer and middle finger of the left hand. All signals were recorded at a sampling rate of 2048 Hz, and the recording systems of the two participants were synchronized.

Signal processing

All signal processing (both preprocessing and signal processing) was done in Python (version 3.10), using Numpy (Harris et al., 2020), Scipy (Virtanen et al., 2020) and Pandas (DOI 10.5281/zenodo.3509134).

Interpersonal synchronization

Preprocessing was different for each included body sensor. For fair comparison, all body responses were cut to the same length as the shortest stimulus: 12:37 minutes.

EEG

EEG was first high-pass filtered using a 4th-order Butterworth filter with a 0.5 Hz cutoff to remove sensor drift. Artifacts exceeding 500 µV were removed, and linear interpolation was performed across the affected intervals. Eyeblink artifacts were removed using a multichannel Wiener Filter (Somers et al., 2018). All channels were re-referenced to a common average reference. Lastly, the signal was downsampled to 64 Hz. All filtering steps were done using zero-phase filtering. Additionally, to reduce the dimensionality (for the synchronization analysis), we applied correlated component analysis (corrCA) (Parra et al., 2018) to calculate components from the EEG that are maximally correlated between participants. Input for corrCA is all of the EEG data between which the intersubject correlations have to be calculated.

ECG

From the raw ECG signal, we extracted instantaneous heart rate based on the ECG processing code made available by Perez et al. (2021). The two ECG- leads were re-referenced to each other (ECG1-ECG2) to suppress common-mode noise and enhance the signal-to-noise ratio. We applied the algorithm by Pan & Tompkins et al. (1985) to find the R-peaks in the ECG signal. Instantaneous heart rate was determined by calculating the time intervals between the peaks across time using the *ECG-detectors* package in Python.

Electrodermal activity

For EDA, we applied a 1.5-second Bartlett window to the raw signal to smooth it. Next, we computed the difference between consecutive samples to remove low-frequency trends and emphasize rapid changes in the signal (Marci et al., 2007). We applied a sigmoid transformation (i.e., $y = \frac{1}{1+e^{-x}}$, with x = EDA amplitude) to compress large peaks. This scaled the data between 0 and 1, where 0 indicates no change or a decrease, and values above 0 indicate an increase.

To assess interpersonal synchronization, we computed intersubject correlations, which measure the correlations of body responses (EEG, HR, EDA) among multiple participants. This involved calculating the average correlations of one participant's responses with those of all other participants. This resulted in a single average correlation value for each participant for each condition, which we could then interpret as the mean listening engagement performance of that specific participant in relation to the whole group. A high synchronization value would indicate a high degree of listening engagement and vice versa. The significance of each data point was calculated using circular shifting,

where one of the signals was randomly shifted in time. Then the correlation was recalculated, after which we fisher transformed each value. This was done over 1000 times to create a null distribution, from which we derived a p-value for each synchronization value.

Neural tracking of the auditory stimulus

No CorrCA was applied for neural tracking analysis. Prior to the following neural tracking analysis, both EEG data and the speech envelope were bandpass filtered between 0.5 and 4 Hz with a 4th-order Butterworth filter. To measure neural tracking, we reconstructed the envelope from the EEG by training a linear decoder with an integration window of 250 ms (or 16 samples). The decoder was trained using ridge regression with the regularization parameter (lambda) set to max(abs(EEG covariance matrix)) to prevent overfitting. Linear weights for each channel at different time delays were calculated so that the mean squared error between the reconstructed and actual speech envelope was minimized. We applied leave-one-fold-out cross-validation, where each recording was split into *N* non-overlapping folds of 120 seconds. The decoder was trained on all but one fold and then tested on the held-out fold.

After the decoder was trained for a specific stimulus, we reconstructed the envelope from the EEG. The reconstructed envelope was then correlated to the actual envelope to obtain a correlation reconstruction accuracy value. This value would be interpreted as a degree of neural tracking or how well the stimulus was processed in the brain over time.

Statistical significance was calculated for each stimulus separately using circ-shifting (Crosse et al., 2021), which is similar to how we calculated significance for the synchronization results. In this case, the pre-processed speech envelope was shifted in time by a random amount of time samples to calculate new noisy correlations. We applied the same Fisher transformation before saving all the new correlation values in the null distribution.

All stimuli were sampled at 48 kHz and normalized for RMS. We extracted the envelope using a gammatone filterbank (Brian2 and Brian2hears toolkits; https://github.com/brian-team/brian2hears), where the soundwave was filtered in 28 subbands between 500 and 5000Hz. Next, the absolute value of each signal was raised by a power of 0.6 to extract the slow modulations (Biesmans et al., 2017). The envelope was downsampled to 64 Hz. This envelope is used as the sound feature in neural tracking. For further fair analysis, all stimuli were cut to the same length as the shortest one (Nonengaging 1), namely 12:37 minutes 37.

Behavioral measure

As a behavioral measure, we evaluated participants' experienced listening engagement and enjoyment after each active listening condition (engaging - nonengaging). We combined and adapted two validated scales that evaluated listening engagement or general engagement (Hannum & Simons, 2020; Kuijpers et al., 2014). Specifically, we combined the 'attention' subscale from the narrative absorption scale and adapted the engagement questionnaire by Hannum & Simons to focus more on listening experiences. The scale was administered after every active listening condition, giving us a good idea of how each stimulus was experienced. Participants rated 15 statements related to their listening experience on a 1-7 Likert scale. For further analysis, we calculated the mean of these 15 statements for each condition separately per participant. We also administered a short multiple-choice content questionnaire, which was mainly intended to motivate participants.

Data analysis

All data analysis and visualization were conducted in R Statistical Software (version 4.3.1, R Core Team, 2023). Subjective results and synchronization measures across different conditions were evaluated for statistical significance using a linear mixed effects model, where synchronization was treated as the dependent variable, with the condition as a fixed effect and participants included as a random effect. The main effect of the condition was assessed by comparing two linear mixed-effects models—one including the condition as a predictor and one excluding it—using ANOVA to determine significant differences. We used the following models:

Full model: *Synchronization value ~ condition + (1|participants)*

Reduced model: *Synchronization value ~ (1|participants)*

Next, we calculated post-hoc comparisons with a holm-correction for multiple comparisons to unravel the different comparisons between groups.

In addition, we calculated correlations between the synchronization values and the subjective measure. We calculated one Spearman correlation per modality for all participants and conditions. Each correlation contains 4 data points per participant (two engaging and two nonengaging conditions). Consequently, the assumption of independence for correlations was violated. To solve this problem, we calculated a new null distribution to compare our results to. More specifically, we shuffled the synchronization data and then recalculated the correlation with a subjective measure. Shuffling was done in such a way that we did not violate any within-participant effects. We shuffled conditions while matching the values of each participant to the values of another participant. We reshuffled the data 1000 times and then recalculated a new correlation each time. All of these analyses are done exactly the same for each body signal.

RESULTS

<u>Participants rate engaging stimuli higher than nonengaging stimuli on the Listening engagement questionnaire</u>

Figure 2 shows the averaged (across all questions) subjective engagement ratings for the different conditions. The mean listening engagement score for the engaging condition was 4.8. Mean scores for the nonengaging stimuli were respectively 2.1. The linear mixed effects model showed significant differences in subjective engagement between conditions ($\beta = -2.70$, $SE = 0.098$, $t = -27.73$, $p < .001$).

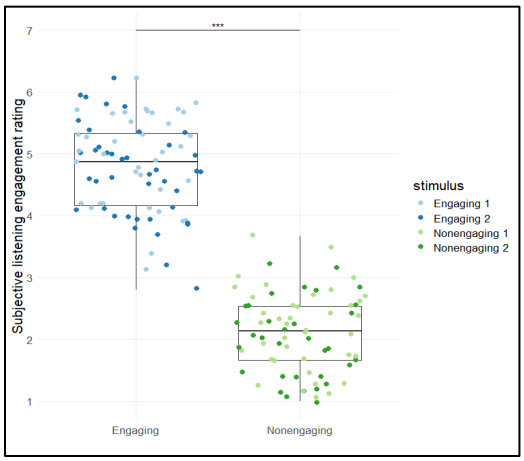

Figure 2: Scores on the subjective listening engagement 7-point-Likert scale. Different shades of blue and green represent different parts of the engaging and nonengaging stimuli. The engaging conditions were rated significantly higher than the nonengaging stimuli (p<0.0001).

Due to system failure, we lost the subjective data for the first engaging stimulus for one participant, for the first part of the nonengaging stimulus for three participants and data for the second part of the nonengaging stimulus for 4 participants.

Participants synchronize the most when listening to an engaging stimulus

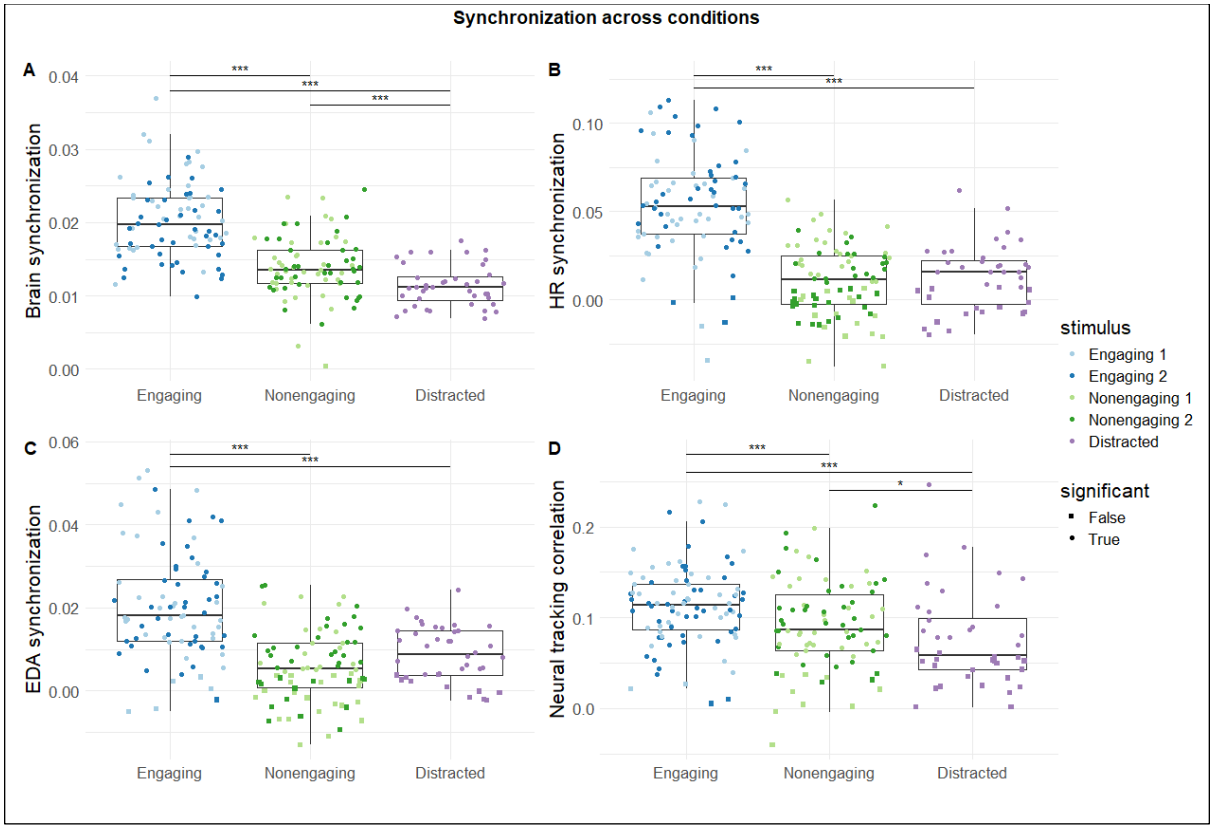

Figure 3: In all cases more synchronization and neural tracking could be identified in the engaging condition in comparison to the nonengaging condition. For brain synchronization and neural tracking also a significant difference could be found between the nonengaging and distracted condition. Each data point represents the mean of correlations calculated between the responses of one participant with all the other participants for that specific combination of stimulus and body response.

Panels A, B and C in **Figure 3** show synchronization between the body responses of different people in response to listening to different auditory stimuli (different degrees of engagement). The significance of each data point was calculated using circ-shifting.

A significant effect of condition could be identified for all the body responses, by comparing two linear mixed effect models with and without condition using ANOVA: brain synchronization ($\chi^2(2) = 158.24$, $p < .0001$), HR synchronization ($\chi^2(2) = 119.44$, $p < .0001$), EDA synchronization ($\chi^2(2) = 73.392$, $p < .0001$). Consecutive post-hoc comparisons with holm correction were performed and the results can be found in **table 2.** All significant contrasts are indicated in bold. For all body responses synchronization was significantly higher in the engaging condition (E) than in the other two conditions (N, D). For brain synchronization we also found higher results in the nonengaging condition in comparison to the distracted condition. For HR and EDA this effect did not hold.

We also looked at neural tracking results. Significant neural tracking could be found for each condition, but not for every participant, especially in the distracted condition. We performed a similar analysis here as we did for the synchronization measures. Condition could be identified as a significant neural

tracking ($\chi^2(2) = 45.033$, $p < .0001$). Post-hoc comparisons showed significant differences for neural tracking between the three conditions (see **table 2**).

Table 2: Post hoc comparisons; bold indicates significant contrasts

	Contrast	β (estimate)	SE	t	Adjusted p-value
BRAIN	**E - N**	0.00644	0.0005	12.878	**<.0001**
	E - D	-0.00883	0.000608	-14.529	**<.0001**
	N - D	-0.00239	0.000611	-3.913	**0.0001**
Heart rate	**E - N**	0.04332	0.00356	12.174	**<.0001**
	E - D	-0.04136	0.00433	-9.557	**<.0001**
	N - D	0.00196	0.00435	0.451	0.6525
Electrodermal activity	**E - N**	0.0145	0.00156	9.261	**<.0001**
	E - D	-0.0115	0.00190	-6.036	**<.0001**
	N - D	0.0030	0.00191	1.571	0.1181
Neural tracking	**E - N**	0.0228	0.00471	4.839	**<.0001**
	E - D	-0.0388	0.00574	-6.766	**<.0001**
	N - D	0.0228	0.00471	4.839	**<.0001**

Relation between objective and behavioral data

We analyzed whether any relations could be found between the objective and behavioral results. **Figure 4** contains scatter plots showing the relation between the different interpersonal synchronization values and the results of the engagement questionnaires. Data for all participants, for all conditions that involved attentive listening can be found in each correlation plot (not the distracted condition as we did not have any subjective results). Panel A contains a scatter plot for brain synchronization, panel B for HR synchronization and panel C for EDA synchronization. All correlations were found to be significant, indicating that relationships between both the objective and the subjective measure for engagement could be found. This was the case for each modality. For the different body responses the scatter plots and the correlations suggest that if participants rated the stimuli as more engaging, there is also higher synchronization between them.

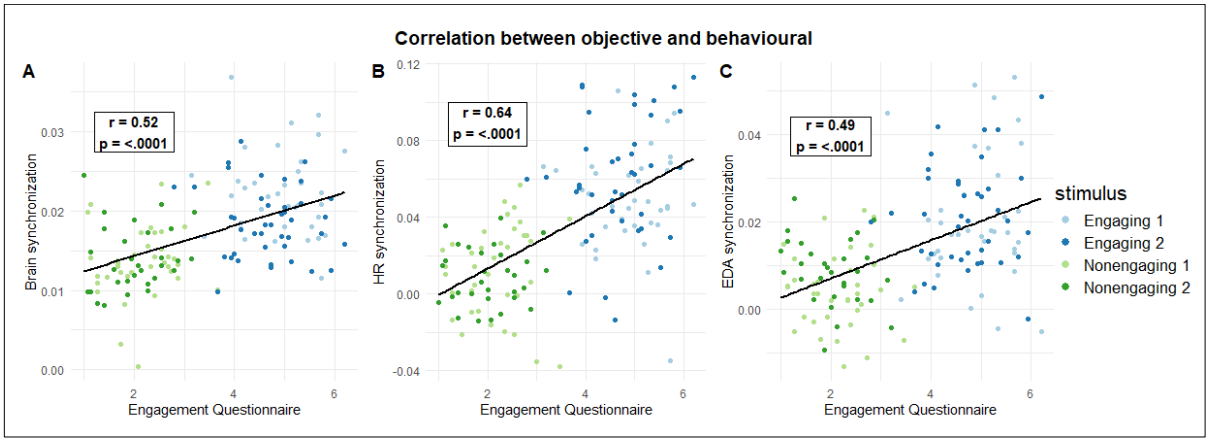

Figure 4: Correlation between synchronization results for the different body responses and the behavioral measure. Significant correlations could be found for all the different body responses

DISCUSSION

This paper aims to evaluate interpersonal synchronization as a measure of listening engagement. Twenty duos were invited to our lab and listened to different auditory stimuli, ranging from engaging and interesting to nonengaging and uninteresting. Meanwhile, we measured several physiological brain and body responses. As an extra control condition, we also gave participants a distracting task while listening to the stimuli. Results suggested that synchronization is promising as a measure of listening engagement.

Objective data

For the different body responses we found significant synchronization when people were presented with an auditory stimulus for all conditions. This finding supports past results where it was shown that interpersonal synchronization was related to conscious processing of spoken stories (Irsik et al., 2022; Madsen & Parra, 2022, 2024). Our results show that synchronization is the highest when people listened in an engaged manner (i.e. engaging condition) and that it significantly declined when they listened to a boring stimulus. This was true for the different brain and body signals (EDA and HR). This finding suggests that some additional cognitive processing happens when you are listening to something that you are engaged in. It is plausible that when you listen to something engaging, more brain regions and cognitive processes get activated, which is then reflected in the brain responses (Marois et al., 2024; Thayer et al., 2009). These cognitive effects might drive the same synchronization in the body. Determining the exact factors driving this is difficult at this stage and needs further research. Nonetheless, these results suggest that, at least to a certain extent, the synchronization-engagement relationship is driven by an involuntary component as the participants were encouraged to listen attentively in both conditions. To verify this, adding a nonengaging stimulus as a control was essential.

Additionally, we showed that synchronization declines when people listen passively (are distracted from listening by a cognitive task) in comparison to when they listen in an engaged manner. This supports previous findings (Madsen & Parra, 2022; P. Pérez et al., 2021). From the brain response, reflected in both neural tracking and brain-to-brain-synchronization, we can identify a difference between listening to something nonengaging and being distracted from listening. This all together suggests that there is a difference between attentive and engaged listening, which can be reflected in measures of synchronization. In addition to attention and conscious speech processing (Madsen & Parra, 2022; P. Pérez et al., 2021), listening engagement can also influence synchronization so that

synchronization can be used as a proxy for listening engagement. For HR and EDA, we did not find the same difference in synchronization for attended-nonengaged listening and distracted listening.

The objective measure as a predictor for subjective listening engagement

We also examined the relationship between self-rated listening engagement scores and objective synchronization data. The results revealed clear and robust relationships between the subjective ratings and objective synchronization measures across the three body responses analyzed. These findings validate the hypothesis that synchronization measures can serve as predictors of listening engagement and/or an individual's listening experience.

Limitations and future directions

We aimed to find the relationship between synchronization and listening engagement, so, based on content interest, we made the difference between the engaging and nonengaging stimulus as large as possible (very interesting vs. very boring). Of course, in daily life we might be more confronted with sounds that are more in the middle. In follow-up studies, it will be important to also include stimuli that are more in between and see whether a continuous gradient in our objective measure could be found.

Listening engagement is a multidimensional concept, and someone's listening experience is influenced by many different factors. Identifying and controlling these factors is challenging. For example, the speaker was not the same for all stimuli. It's possible that the engageability of a voice timbre is different for each participant. To partly control for this, we ensured that all speakers were female, so no potential speaker's sex effects could be of influence. Other factors to keep in mind are things like fatigue, anxiety, interest (Busselle & Bilandzic, 2009; Herrmann & Johnsrude, 2020; Pichora-Fuller et al., 2016). In the future, engageability of the speaker, e.g. speech rate, prosody, voice… might also be of influence for a listening engagement.

Another limitation of this study is the challenge of differentiating between listening engagement and auditory attention. In all active listening conditions, participants were instructed to listen attentively to the presented stimulus and to focus on it actively. To ensure compliance, participants were informed beforehand that they would be asked multiple-choice questions about the stimulus after listening. With this design, it can be assumed, theoretically and at the group level, that participants maintained similar levels of attention across both listening activities. However, in practice, engaged listening and attentive listening are likely to be interrelated in a more complex manner. Consequently, it is difficult to ascertain whether the observed differences between conditions are solely driven by listening engagement or potentially influenced by other factors, such as attention. Future studies will need to explore this relationship further to disentangle these variables.

In future analyses, we also aim to explore the dyadic effects in these data further.

Conclusion

In this study we let people listen to auditory stimuli while we measure their brain response, heart rate and electrodermal activity. We included stimuli that were either engaging or nonengaging. We evaluated whether synchronization between participants' bodies was different in the engaging condition than in the nonengaging condition. This was clearly the case for all the body responses. This suggests that synchronization can be used to objectively evaluate listening engagement.

References

Anthiyur Aravindan, A., Kalyan Chappidi, S., Thumma, A., & Palanisamy, R. (2023). Prediction of Arousal and Valence State from Electrodermal Activity using Wavelet based ResNet50 Model. *Current Directions in Biomedical Engineering*, *9*(1), 555–558. https://doi.org/10.1515/cdbme-2023-1139

Aron, A., Melinat, E., Aron, E. N., Vallone, R. D., & Bator, R. J. (1997). The Experimental Generation of Interpersonal Closeness: A procedure and Some Preliminary Findings. *Personality and Social Psychology Bulletin* , *23*(4).

Bevilacqua, D., Davidesco, I., Wan, L., Chaloner, K., Rowland, J., Ding, M., Poeppel, D., & Dikker, S. (2018). Brain-to-brain synchrony and learning outcomes vary by student–teacher dynamics: Evidence from a real-world classroom electroencephalography study. *Journal of Cognitive Neuroscience*, *31*(3), 401–411. https://doi.org/10.1162/jocn_a_01274

Biesmans, W., Das, N., Francart, T., & Bertrand, A. (2017). Auditory-inspired speech envelope extraction methods for improved EEG-based auditory attention detection in a cocktail party scenario. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *25*(5), 402–412. https://doi.org/10.1109/TNSRE.2016.2571900

Brodbeck, C., & Simon, J. Z. (2020). Continuous speech processing. In *Current Opinion in Physiology* (Vol. 18, pp. 25–31). Elsevier Ltd. https://doi.org/10.1016/j.cophys.2020.07.014

Busselle, R., & Bilandzic, H. (2009). Measuring Narrative Engagement. *Media Psychology*, *12*(4), 321–347. https://doi.org/10.1080/15213260903287259

Chatterjee, I., Goršič, M., Clapp, J. D., & Novak, D. (2021). Automatic Estimation of Interpersonal Engagement During Naturalistic Conversation Using Dyadic Physiological Measurements. *Frontiers in Neuroscience*, *15*. https://doi.org/10.3389/fnins.2021.757381

Cohen, S. S., & Parra, L. C. (2016). Memorable audiovisual narratives synchronize sensory and supramodal neural responses. *ENeuro*, *3*(6). https://doi.org/10.1523/ENEURO.0203-16.2016

Crosse, M. J., Zuk, N. J., Di Liberto, G. M., Nidiffer, A. R., Molholm, S., & Lalor, E. C. (2021). Linear Modeling of Neurophysiological Responses to Speech and Other Continuous Stimuli: Methodological Considerations for Applied Research. In *Frontiers in Neuroscience* (Vol. 15). Frontiers Media S.A. https://doi.org/10.3389/fnins.2021.705621

Cziksentmihalyi, M. (1990). Flow: The Psychology of Optimal Experience. *Harper & Row*.

Deng, X., Chen, K., Chen, X., Zhang, L., Lin, M., Li, X., & Gao, Q. (2024). Parental involvement affects parent-adolescents brain-to-brain synchrony when experiencing different emotions together: An EEG-based hyperscanning study. *Behavioural Brain Research*, *458*. https://doi.org/10.1016/j.bbr.2023.114734

Denys, S., Hofmann, M., van Wieringen, A., & Wouters, J. (2019). Improving the efficiency of the digit triplet test using digit scoring with variable adaptive step sizes. *International Journal of Audiology*, *58*(10), 670–677. https://doi.org/10.1080/14992027.2019.1622042

Dibble, J. L., Levine, T. R., & Park, H. S. (2012). The unidimensional relationship closeness scale (URCS): Reliability and validity evidence for a new measure of relationship closeness. *Psychological Assessment*, *24*(3), 565–572. https://doi.org/10.1037/a0026265

Ding, N., & Simon, J. Z. (2012). Emergence of neural encoding of auditory objects while listening to competing speakers. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(29), 11854–11859. https://doi.org/10.1073/pnas.1205381109

Hannum, M. E., & Simons, C. T. (2020). Development of the engagement questionnaire (EQ): A tool to measure panelist engagement during sensory and consumer evaluations. *Food Quality and Preference*, *81*. https://doi.org/10.1016/j.foodqual.2019.103840

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., … Oliphant, T. E. (2020). Array programming with NumPy. In *Nature* (Vol. 585, Issue 7825, pp. 357–362). Nature Research. https://doi.org/10.1038/s41586-020-2649-2

Herrmann, B., & Johnsrude, I. S. (2020). A model of listening engagement (MoLE). *Hearing Research*, *397*. https://doi.org/10.1016/j.heares.2020.108016

Iotzov, I., & Parra, L. C. (2019). EEG can predict speech intelligibility. *Journal of Neural Engineering*, *16*(3). https://doi.org/10.1088/1741-2552/ab07fe

Irsik, V. C., Johnsrude, I. S., & Herrmann, B. (2022). Neural Activity during Story Listening Is Synchronized across Individuals Despite Acoustic Masking. *Journal of Cognitive Neuroscience*, *34*(6), 933–950. https://doi.org/10.1162/jocn_a_01842

Kinreich, S., Djalovski, A., Kraus, L., Louzoun, Y., & Feldman, R. (2017). Brain-to-Brain Synchrony during Naturalistic Social Interactions. *Scientific Reports*, *7*(1). https://doi.org/10.1038/s41598-017-17339-5

Kołodziej, M., Tarnowski, P., Majkowski, A., & Rak, R. J. (2019). Electrodermal activity measurements for detection of emotional arousal. *Bulletin of the Polish Academy of Sciences: Technical Sciences*, *67*(4), 813–826. https://doi.org/10.24425/bpasts.2019.130190

Kuijpers, M. M., Hakemulder, F., Tan, E. S., & Doicaru, M. M. (2014). Exploring absorbing reading experiences. *Scientific Study of Literature*, *4*(1), 89–122. https://doi.org/10.1075/ssol.4.1.05kui

Madsen, J., & Parra, L. C. (2022). Cognitive processing of a common stimulus synchronizes brains, hearts, and eyes. *PNAS Nexus*, *1*(1). https://doi.org/10.1093/pnasnexus/pgac020

Madsen, J., & Parra, L. C. (2024). Bidirectional brain-body interactions during natural story listening. *Cell Reports*, *43*(4). https://doi.org/10.1016/j.celrep.2024.114081

Marci, C. D., Ham, J., Moran, E., & Orr, S. P. (2007). Physiologic correlates of perceived therapist empathy and social-emotional process during psychotherapy. *Journal of Nervous and Mental Disease*, *195*(2), 103–111. https://doi.org/10.1097/01.nmd.0000253731.71025.fc

Marois, A., Mouratille, D., Pratviel, Y., Chamberland, C., & Tremblay, S. (2024). Using Cardiac and Electrodermal Activity as Cognitive Markers for Interruptions and Distraction in a Surveillance Simulation. *Neuroergonomics and Cognitive Engineering*, *126*. https://doi.org/10.54941/ahfe1004733

Massar, S. A. A., Lim, J., Sasmita, K., & Chee, M. W. L. (2016). Rewards boost sustained attention through higher effort: A value-based decision making approach. *Biological Psychology*, *120*, 21–27. https://doi.org/10.1016/j.biopsycho.2016.07.019

Metshein, M., Krivošei, A., Abdullayev, A., Annus, P., & Märtens, O. (2022). Non-Standard Electrode Placement Strategies for ECG Signal Acquisition. *Sensors*, *22*(23). https://doi.org/10.3390/s22239351

Palumbo, R. V., Marraccini, M. E., Weyandt, L. L., Wilder-Smith, O., McGee, H. A., Liu, S., & Goodwin, M. S. (2017). Interpersonal Autonomic Physiology: A Systematic Review of the Literature. *Personality and Social Psychology Review*, *21*(2), 99–141. https://doi.org/10.1177/1088868316628405

Pan, J., & Tompkins, W. J. (1985). A Real-Time QRS Detection Algorithm. *IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING*, *32*(3).

Parra, L. C., Haufe, S., & Dmochowski, J. P. (2018). *Correlated Components Analysis - Extracting Reliable Dimensions in Multivariate Data*. https://doi.org/10.51628/001c.7125

Pérez, A., Carreiras, M., & Duñabeitia, J. A. (2017). Brain-To-brain entrainment: EEG interbrain synchronization while speaking and listening. *Scientific Reports*, *7*(1). https://doi.org/10.1038/s41598-017-04464-4

Pérez, P., Madsen, J., Banellis, L., Türker, B., Raimondo, F., Perlbarg, V., Valente, M., Niérat, M. C., Puybasset, L., Naccache, L., Similowski, T., Cruse, D., Parra, L. C., & Sitt, J. D. (2021). Conscious processing of narrative stimuli synchronizes heart rate between individuals. *Cell Reports*, *36*(11). https://doi.org/10.1016/j.celrep.2021.109692

Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W. Y., Humes, L. E., Lemke, U., Lunner, T., Matthen, M., Mackersie, C. L., Naylor, G., Phillips, N. A., Richter, M., Rudner, M., Sommers, M. S., Tremblay, K. L., & Wingfield, A. (2016). Hearing impairment and cognitive energy: The framework for understanding effortful listening (FUEL). *Ear and Hearing*, *37*, 5S-27S. https://doi.org/10.1097/AUD.0000000000000312

Sánchez-Reolid, R., López de la Rosa, F., Sánchez-Reolid, D., López, M. T., & Fernández-Caballero, A. (2022). Machine Learning Techniques for Arousal Classification from Electrodermal Activity: A Systematic Review. In *Sensors* (Vol. 22, Issue 22). MDPI. https://doi.org/10.3390/s22228886

Seli, P., Cheyne, J. A., Xu, M., Purdon, C., & Smilek, D. (2015). Motivation, intentionality, and mind wandering: Implications for assessments of task-unrelated thought. *Journal of Experimental Psychology: Learning Memory and Cognition*, *41*(5), 1417–1425. https://doi.org/10.1037/xlm0000116

Somers, B., Francart, T., & Bertrand, A. (2018). A generic EEG artifact removal algorithm based on the multi-channel Wiener filter. *Journal of Neural Engineering*, *15*(3). https://doi.org/10.1088/1741-2552/aaac92

Song, H., Finn, E. S., & Rosenberg, M. D. (2021). Neural signatures of attentional engagement during narratives and its consequences for event memory. *PSYCHOLOGICAL AND COGNITIVE SCIENCES*, *118*(33).

Thayer, J. F., Hansen, A. L., Saus-Rose, E., & Johnsen, B. H. (2009). Heart rate variability, prefrontal neural function, and cognitive performance: The neurovisceral integration perspective on self-regulation, adaptation, and health. In *Annals of Behavioral Medicine* (Vol. 37, Issue 2, pp. 141–153). https://doi.org/10.1007/s12160-009-9101-z

Tobin, S. J., & Guadagno, R. E. (2022). Why people listen: Motivations and outcomes of podcast listening. *PLoS ONE*, *17*(4 April 2022). https://doi.org/10.1371/journal.pone.0265806

Vanthornhout, J., Decruy, L., & Francart, T. (2019). Effect of Task and Attention on Neural Tracking of Speech. *Frontiers in Neuroscience*, *13*. https://doi.org/10.3389/fnins.2019.00977

Vanthornhout, J., Decruy, L., Wouters, J., Simon, J. Z., & Francart, T. (2018). Speech Intelligibility Predicted from Neural Entrainment of the Speech Envelope. *JARO - Journal of the Association for Research in Otolaryngology*, *19*(2), 181–191. https://doi.org/10.1007/s10162-018-0654-z

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., … Vázquez-Baeza, Y. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods*, *17*(3), 261–272. https://doi.org/10.1038/s41592-019-0686-2