Predictive encoding of a speech: a neural tracking analysis in ‘cocktail party’ situations

Thaiz Priscilla Sánchez Costa

Maestría en Ciencias Cognitivas
Facultad de Ciencias, Facultad de Psicología

Universidad de la República

Montevideo, Uruguay
Diciembre, 2024

Predictive encoding of a speech: a neural tracking analysis in ‘cocktail party’ situations

Thaiz Priscilla Sánchez Costa

Tesis de Maestría presentada al Programa de Posgrado en Ciencias Cognitivas, Facultad de Ciencias de la Universidad de la República, como parte de los requisitos necesarios para la obtención del título de Magíster en Ciencias Cognitivas.

Director:
Prof. Adj. Dr. Francisco Cervantes Constantino

Director académico:
Prof. Agda. Dra. Alejandra Carboni

Montevideo, Uruguay
Diciembre, 2024

INTEGRANTES DEL TRIBUNAL DE DEFENSA DE TESIS

Ph.D. Prof. Ignacio Rebollo

Ph.D. Prof. Gabriela Fernández

Ph.D. Prof. Gabriel Barg

Montevideo, Uruguay
Diciembre, 2024

Abstract

The brain's ability to focus on one conversation in a noisy environment, known as the cocktail party effect, remains a significant challenge in neuroscience. We still lack a clear understanding of how the brain uses experience to prioritize elements of auditory information from speech successfully. Research has shown that neural processing is predominantly dedicated to speech from an attended speaker compared to ignored speech. However, familiar and previously experienced stimuli may trigger relatively weaker neural responses than novel ones do, a phenomenon known as priming effect. This study investigates the interplay between auditory processing in the cock- tail party effect and the influence of prior knowledge through different levels of familiarity with speech. Our research specifically investigates how prior experience with speech, including familiarity with the voice or the content of the voiced message, plays a role in this information segregation process.

In a series of single-trial electroencephalography (EEG) studies, we measured the temporal attention indices allocated to neural responses to continuous speech using the temporal response function (TRF) method. The experiments were conducted to investigate the interaction between attention and experience at three different levels.

In Experiment 1, we primed listeners with single-speaker speech prior to hearing the exact same sequence in a two-speaker cocktail-party scenario, which involved full prior knowledge of one of the two speech objects. In Experiment 2, listeners were similarly primed with single-speaker speech, and in the subsequent cocktail-party setting were then exposed to one of the two speakers having the same voice but a different message. In Experiment 3, the speech on which listeners were primed via a single-speaker was then restated in the same paraphrasal content from one of two new speakers. Importantly, we addressed the separate impact that prior knowledge may have on processing of the attended versus unattended speech, i.e., whether prior experience relates to target or masker speech.

We identify and analyse the effects of attentional and experience factors on components P1, N1, and P2 from the resulting EEG TRFs. The results show that attention and experience factors interact during the late P2 processing stage (220-320 ms), as long as priming involves full prior knowledge (Experiment 1), not being observed on partially primed auditory information (Experiments 2a and 2b. The interaction consists of a reduced neural response to the primed target, compared to the case when the target is unprimed. The reduction occurs while maintaining an attentional enhancement in the P2 response, relative to the concurrent masker's. The finding of reductions in P2 but not earlier stages is consistent with this later cortical stage operating a single auditory object representation of neurally segregated speech.

Keywords:
cocktail party effect, selective attention, speech encoding, temporal response function.

Resumen

La capacidad del cerebro para enfocarse en una conversación en un ambiente ruidoso, conocido como efecto cóctel, sigue siendo un desafío significativo en la neurociencia. Todavía carecemos de una comprensión clara de cómo el cerebro utiliza la experiencia para priorizar elementos de la información auditiva del habla con éxito.

La investigación ha demostrado que el procesamiento neuronal se dedica predominantemente al habla de un interlocutor atendido en comparación con el habla ignorada. Sin embargo, los estímulos familiares y previamente experimentados pueden desencadenar respuestas neuronales relativamente más débiles que los novedosos, un fenómeno conocido como efecto de priming. Este estudio investiga la interacción entre el procesamiento auditivo en el efecto de fiesta de cóctel y la influencia del conocimiento previo a través de diferentes niveles de familiaridad con el habla. Nuestra investigación investiga específicamente cómo la experiencia previa con el habla, incluida la familiaridad con la voz o el contenido del mensaje vocalizado, desempeña un papel en este proceso de segregación de información.

En una serie de estudios de electroencefalografía (EEG) de ensayo único, medimos los índices de atención temporal asignados a las respuestas neuronales al habla contínua utilizando el método de función de respuesta temporal (FRT). Los experimentos se llevaron a cabo para investigar la interacción entre la atención y la experiencia en tres niveles diferentes.

En el Experimento 1, familiarizamos a los oyentes con el discurso de un solo hablante antes de escuchar la misma secuencia en un escenario de fiesta de cóctel con dos hablantes, que implicaba un conocimiento previo completo de uno de los dos objetos de habla. En el Experimento 2, los oyentes fueron familiarizados de manera similar con discurso de un solo hablante y, en el posterior escenario de cóctel, fueron expuestos a uno de los dos interlocutores con la misma voz pero un mensaje diferente. En el Experimento 3, el habla sobre la que se familiarizó a los oyentes a través de un solo interlocutor se repitió luego en el mismo contenido parafraseado de dos nuevos hablantes en la posterior escucha del cóctel. Es impor-

tante destacar que abordamos el impacto separado que el conocimiento previo puede tener en el procesamiento del habla atendida versus no atendida, es decir, si la experiencia previa se relaciona con el habla objetivo o el habla desatendida.

Identificamos y analizamos los efectos de los factores atencionales y de experiencia en los componentes P1, N1 y P2 de las FRT de EEG resultantes. Los resultados muestran que los factores de atención y experiencia interactúan durante la etapa de procesamiento P2 tardío (230-320 ms), siempre que el priming implique un conocimiento previo completo (Experimento 1), y no se observa en la información auditiva parcialmente familiarizada (Experimentos 2a y 2b). La interacción consiste en una respuesta neuronal reducida al objetivo conocido, en comparación con el caso en que el objetivo no conocido. La reducción ocurre mientras se mantiene un aumento atencional en la respuesta P2, en relación con el enmascarador concurrente. El hallazgo de reducciones en P2 pero no en etapas anteriores es consistente con esta etapa cortical posterior que opera una representación de objeto auditivo único de habla neuronalmente segregada.

Palabras claves:

efecto fiesta de cóctel, atención selectiva, codificación neural, función de respuesta temporal.

Agradecimientos

Quisiera expresar mi agradecimiento a mis mentores académicos, Francisco Cervantes y Alejandra Carboni, por su constante acompañamiento y supervisión durante mi desarrollo académico como investigadora en formación.

A la Agencia Nacional de Investigación e Innovación, por financiar este proyecto. A Jonathan Z. Simon por las colaboraciones en el proyecto Tom Francart.

A Ignacio Rebollo, Gabriela Fernández y Gabriel Barg, por la revisión de este trabajo y por sus sugerencias constructivas destinadas a su mejora.

Al CIBPsi, por brindar un entorno de trabajo interdisciplinario y enriquecedor, así como por promover mi desarrollo académico en el campo de las neurociencias.

A mis colegas de laboratorio, de maestría y amigos Soledad Assis, Maria Noel Macedo, Anaclara Gerosa, Jenny Nogueira, Germán Cipriani, María Clara Ferreira y Richard Rodriguez quienes han realizado valiosos aportes y han sido compañeros indispensables en este desafiante trayecto.

A los Licenciados Rodrigo Caramés y Franco Caballero por colaborar en el proyecto.

Agradezco a mi familia por su incondicional apoyo y cariño durante esta etapa a María Costa, Sandra Sánchez, Aline Sánchez y Rhitter Sánchez.

A mis amigos por acompañarme en este proceso. A los voluntarios que participaron en este estudio.

List of tables

List of figures

Table of Contents

Chapter 1

Introduction

Following intense social lives, humans have developed a remarkable ability to communicate and understand each other. Many important soundscapes involve multiple conversations, such as parties where people talk simultaneously amid ever-changing background sounds. Yet, one may still discriminate the speech of a friend from a stranger's. Despite such acoustically challenging environments, listeners retain the ability to comprehend a message of their interest from a target speaker while disregarding others. This phenomenon is known as the cocktail party effect (Cherry, 1953) and is based on selective attention, a cognitive process that enables the selection of different sound sources by shaping the representation of sounds from their initial mixed entry into the ear, and follows on different stages up to the parsing of the message (Har-shai Yahav and Zion Golumbic, 2021; Wood and Cowan, 1995).

The formation of auditory objects involves the grouping of sounds in the form of temporally connected objects (Bronkhorst, 2000; Moore and Gockel, 2012) with selection leading to their segregation from a scene (Middlebrooks and Simon, 2017; Shinn-Cunningham et al., 2017). Selection is motivated by listener's demands in a multi-talker environment and is constrained by limited resources and neural processing costs, shaping the distinction between relevant and irrelevant sources. Auditory selection may be based on extraction of physical features of sound including timbre, tone, intensity, or location (Brodbeck et al., 2020; Bronkhorst, 2000; Giraud and Poeppel, 2012; Luo and Poeppel, 2007; Obleser and Kayser, 2019; Ozmeral et al., 2021). In speech, these features represent low-hierarchy 'primitive' grouping processing (Bronkhorst, 2000) that later serve as a basis to determine the binding of sensory information under common causes as auditory object units (Poeppel and Assaneo, 2020; Shinn-Cunningham et al., 2017). Such processing entails top-down higher-level cognitive functions, with speech comprehension depending on both formation and selection processes feedback to each other heterarchically (Shinn-Cunningham et al., 2017). This is opposed to a process structured into successively higher-level control mechanisms. For instance, determining a target in an auditory scene involves estimating individual sound sources, such as a particular voice, on the basis of relevance guided by top-down action mechanisms. However, strongly salient features of a scene can instantaneously capture and redirect relevance away from the elements of the target. In another example, strong prior expectations about speech may guide its representation when it is replaced by noise, overcoming missing sensory input (Cervantes Constantino and Simon, 2018, Leonard et al., 2016).

Examples in which forming an individual auditory object involves interacting top-down and bottom-up processes typically relate to the balance between attentional control (which promotes exploitation of prior knowledge) versus attentional capture (which promotes feature exploration). Critically, the balance depends on experience and observer expectations (Shinn-Cunningham, 2008). In an effort to gain a deeper understanding of how auditory objects are encoded in the brain, numerous studies have addressed the representational differences between attended and unattended auditory objects at the cocktail party problem (Ding and Simon, 2012; Haykin and Chen, 2005; O'Sullivan et al., 2015; Power et al., 2012). The question is motivated in particular because in complex cocktail party situations, individuals can adapt and follow a specific stream despite considerable disadvantages (Fiedler et al., 2019), including when the target and maskers have a voice of the same gender, or when the intensity of the target is lower than that of the background (Ding and Simon, 2012; Kidd et al., 2016) for instance.

Important insight has been gained with high temporal precision neuroimaging methods such as electroencephalography (EEG), in combination with temporal response function analyses (TRF). The TRF is particularly useful to describe the neural processing during speech perception, as it details the stages of auditory processing given specific acoustical characteristics of speech (Bednar and Lalor, 2020; Broderick et al., 2019; Di Liberto et al., 2015; Har-shai Yahav and Zion Golumbic, 2021; Wöstmann et al., 2019; Zion Golumbic et al., 2013). These studies have clarified which auditory features are relevant for speech encoding, and how does selective attention modulate their representation in relation to a target versus a masker.

It remains unclear however how selective attention interacts with prior knowledge in the cocktail party. For this, we set out an experimental design to investigate the nature of the statistical interaction between selective attention and prior experience in speech processing when individuals are exposed to multiple simultaneous speakers. Using the TRF method, we investigated this interaction at each of the relevant stages of speech processing, using non-invasive EEG data. In particular, we addressed the problem of prior experience of speech influences neural patterns under a speech comprehension task, representing the typical challenges facing the ‘cocktail party’ problem.

Chapter 2

Theoretical background

Attentional filter models

Different models have approached the attentional mechanisms involved in sensory information processing. The work of Cherry (1953) led to the advent of the Early Selection model of attention (Broadbent, 1957), which assumes a rigid filter that allows only relevant information to pass for processing at later stages, similar to a bottleneck. In this model, the processing of irrelevant acoustic information is limited, and extraction of information from ignored auditory objects is not fully accounted for (Cherry, 1953). It does not explain however how attentional mechanisms can be directed based on individual needs and stimulus properties, for example, when student wishes to listen to a lecture despite off-topic discussions from his classmates. Even if thorough comprehension of the lecturer's speech can be achieved despite the noise, should the student be named by peers this would likely engage his attention. Thus even as off-target input can be mostly ignored some may leak in affecting performance.

By contrast, from the visual literature the Attenuation Model proposes that object-dependent feature integration during serial search as the process driving selection (Treisman, 1960). Unlike exclusive processing of the target versus background input (all or nothing), the latter is only subject to an attenuation mechanism that modifies background perception (Driver, 2001). This introduces a preattentive stage where objects may be formed and integrated through their various features, setting the conditions for later selection of the target.

Understanding how the brain processes acoustic information in challenging multi-talker scenarios during the early stages of processing may help disentangle the neural coding and attentional mechanisms that manage basic features of sound before we focus on specific auditory inputs. This process may act like a spotlight, enabling us to concentrate on one sound at a time.

Three key time windows for speech processing

Evidence suggests that the auditory brain aligns its activity to temporal regularities found in speech. This neural entrainment phenomenon is known as 'phase coding' or 'speech tracking' (Obleser and Kayser, 2019; Zion Golumbic et al., 2013). Speech tracking and related event-related potential (ERP) reports from EEG and MEG studies typically indicate a pattern of neural activation with latencies structured in a triphasic series of auditory processing stages, namely the P1, N1, and P2 components (Aiken and Picton, 2008; Fiedler et al., 2019; Martin et al., 2008; Steinschneider et al., 2011).

The first processing peak, P1, is recognized with a positive polarity, and also called the P50 component, or P1m in the MEG literature (Aiken and Picton, 2008) and occurs in the pre-attentive stage, approximately 50 ms after stimulation. P1 is considered to emerge in the Heschl gyrus and is related to sensory perception and the encoding of acoustic characteristics of the incoming auditory stream. P1 has been associated with sensory gating (Jones et al., 2016), to the ability to reduce representation of repetitive sounds, similar to filtering out redundant sensory information (Haykin and Chen, 2005). It is a component widely linked to bottom-up processes that has however rarely featured in cocktail party studies given its early nature and relatively low amplitude. Some evidence links the P1 to goal-directed attentional tasks, suggesting an interactive flow of information between participating top-down and bottom-up processes (Golubic et al., 2019; Jones et al., 2016).

The next stage of processing, the N1 component, is indicated by a negative peak starting 100 ms after stimulus onset in the EEG. It has been the most studied in the cocktail party as it is in this time window when dominance of the representation of the attended stream response is typically observed. Based on ERP data, Hillyard et al. (1973) first observed attentional effects in the auditory N1 response to attended tones, which elicited larger responses than N1 responses from unattended tones. The N1 modulation appears consistent with the gain model for the representation of attended stimuli and suppression of unattended stimuli, according to Treisman's filter-attenuation hypothesis. In particular, N1 changes are critical to account for in step-wise frameworks to disentangle aspects of early versus late selection (Fiedler et al., 2019; Giard et al., 2000). This is particularly relevant when processing salient stimuli within complex acoustic scenes, where attention plays a crucial role in segregating important auditory information. In natural acoustic scenes, other factors can also contribute to the processing within this temporal window, including familiarity with the attended or ignored stimulus. This phenomenon, known as the priming effect, can influence the processing of a stimulus (Wang et al., 2019).

The emerging phase-locking of auditory objects at the N1 may reflect synchronization of endogenous neural activity with the target's own exogenous temporal structure (Kachlicka et al., 2022). A hypothesis is that this stage enables masker-attenuated object-based processing of the selected target, in a mechanism that is considered to be influenced and guided by top-down processes (Ding and Simon, 2012; Kaya and Elhilali, 2017).

Brodbeck et al. (2020) investigated how speech segregation is processed cortically during the cocktail party. Specifically, they examined MEG responses to attended and ignored speech, as well as responses to single-speaker (solo) listening. The researchers presented pairs of audiobook segments in a continuous mixture of two speakers, and the listeners were instructed to focus on only one speaker. They investigated how the auditory cortex tracks acoustic onsets (Hertrich et al., 2012), which are necessary cues for speech segmentation. The activity revealed, as expected, attentional modulation of the N1 in the form of enhanced responses to the target acoustic onsets, but also an earlier response compared to responses to the ignored speaker. The responses to acoustic onsets were obtained with spectro-temporal response functions (STRFs), which generally followed a similar pattern for both solo speech and cocktail listening. The general pattern consisted of at least two peaks, of which the earliest, P1, had a latency of approximately 70 ms, followed by an N1 peak of approximately 150 ms. Their N1 results were in line with previous studies that have indicated enhanced target processing at this stage. With regards to the background, the attenuated and delayed responses may be interpreted as the auditory system maintaining a limited ability to retrieve acoustic onsets from the background response. The available representations of the background sensory noise then become further reduced as different stages of information processing advance from the periphery through the auditory cortex (Rabinowitz et al., 2013).

Finally, the P2 component is characterized by a positive peak in the EEG waveform, which typically occurs around 150-260 ms after stimulus onset (Pinto et al., 2019; Lewald and Getzmann, 2015; Getzmann et al., 2016). The stage has been is associated with top-down and higher-order perceptual processing that is also modulated by attention, and appears delayed in older populations (Getzmann et al., 2016). In the cocktail party situations, the P2 component of the target response typically increases compared with the background (Charest et al., 2009). In particular, the auditory P2 may be behaviorally relevant, as the magnitude of its attentional modulation may indicate task performance (O'Sullivan et al., 2015). One hypothesis is that during this later stage of processing, also known as P300, representations of the background are further prevented from entering working memory (Power et al., 2012). Evidence for this higher-order component in particular to be linked to working memory remains unclear, however (Hjortkjær et al., 2020). Nevertheless, as memory load increases, encoding accuracy of the speech envelope was shown to generally decreases in TRF models, suggesting a potential interplay between attentional mechanisms and working memory (Gazzaley and Nobre, 2012) across encoding stages.

Does prior knowledge play a significant role in understanding speech in cocktail party scenarios?

In everyday scenarios, prior knowledge or familiarity may offer advantages in managing complex and challenging situations. Investigating the significance of the specific circumstances set by priming speech input may provide new insight into ecological contexts. Examples include the easier distinguishing of a friend's voice in a crowd, or understanding a noisy message when one is familiar with the topic. Processing voices that are familiar improves target intelligibility in natural communication settings with multiple speakers, relative to unfamiliar voices (Domingo et al., 2020). However, the behavioral benefits on comprehension may not be extended to cases where the unattended speaker alone is familiar (Newman and Evers, 2007). Nevertheless, there is evidence that in a low cognitive demand tasks and with extensive experience, masker's voice familiarity may facilitate cocktail party performance (Johnsrude et al., 2013).

At the level of prior knowledge about speech content, behavioral evidence shows detection benefits for target words that have been semantically primed shortly before and which correspond to the same semantic category during masking (Dekerle et al., 2014). Park et al. (2023) studied the influence of topic familiarity on selective attention; using a dichotic listening paradigm, selective comprehension of speech was investigated across simple versus complex message narratives (including challenging philosophical discourse), and, separately, on fixed versus variable speaker targets. The findings showed, in the case of complex narratives, steep comprehension error increases that were not in place for listeners that were presented with simple narrations but had no prior knowledge of the speaker. The complex narrations entail contexts of lower sentential predictability, and this is a factor that may interact with speech intelligibility especially in noise-degraded conditions (Bhandari et al., 2021).

In the case of complete prior knowledge about speech, full auditory priming of an identical stimulus is particularly helpful, in terms of performance when target and masker speech presentations are co-localized (Wang et al., 2021). Evidence suggests that prior knowledge of speech facilitates the segregation of targets in multi-speaker scenes through top-down processing. Using the TRF method, Wang et al. (2019) investigated the effect of prior knowledge of speech stimuli by analysing magnetoencephalography responses to a mixture of stories narrated by the same speaker. In half of trials, listeners were be primed with the target speech sequences of the mixture, by listening to them in advance as a solo (single-speaker) speech presentation. Neural encoding was compared under primed and unprimed conditions for attended and unattended speech. Attentional modulation of the TRF emerged from around 100 ms and was influenced by priming, with the authors reporting significantly enhanced attentional effects for the primed compared to the unprimed condition. Source analyses revealed significant activity changes related to differential attentional enhancement at the level of bilateral superior temporal gyrus (STG) and the superior temporal sulcus (STS). However, the nature of the statistical interaction between attentional enhancement and the priming effect at early, middle and late stages of processing, remained unclear. Moreover, the experimental design of the task did not directly involve speech comprehension, as participants were instructed to detect brief silences (acoustic gaps) of the target speaker ignoring those from the masker's. In addition, the cocktail stimuli were presented using the same voice for target and background, in all trials. If neural processing of the scene were involved for a speech comprehension task, we would expect performance differences to be dependent on the combination of participating voices (Brungart, 2001), as target/masker voice characteristics may affect the separability of auditory objects.

Top-down expectation allows for a complex interplay between perception and understanding, using knowledge to drive perceptual inference (Clark, 2013; Ten Oeve & Martin, 2024). For human hearing, predictive mechanisms are considered to interact at the bottom-up and top-down levels playing an important role in speech encoding. In auditory cortex, top-down input is hypothesized to be compared against incoming sensory inputs, according to existing knowledge and expectations of the observer, in order to adjust prediction errors and interpret the auditory scene accordingly more precisely. In a study that evaluated the influence of visual predictions on auditory speech processing (Pinto et al., 2019), it was found that the observer's reliance on the gestural information of a speaker during speech listening resulted in a reduction in the amplitude of the components N1 and P2. The effects of decreasing activity under priming of certain features may be interpreted as a learning outcome that reflects a reduced neural processing cost regarding those features. This is similar to the 'repetition suppression effect' phenomenon that may indicate the efficient encoding of redundant information (Grill-Spector et al., 2006; Summerfield et al., 2008). It is unclear how does this reduction interact with the key suppression of distractor-related activity under voluntary attention, an effect attributed to inhibitory process (Wöstmann et al., 2022).

In this thesis, we address this interaction in complex cocktail-party scenarios. We ask at what level does prior experience benefit the auditory segregation that is reflected in activity at the auditory cortex. One level of experience involves the prior knowledge about acoustic characteristics of a speaker, such as their distinctive voice spectral cues. At a higher level, we consider the case where prior knowledge arises from the intended message, leading to facilitated semantic and phonemic content inference. The third, most fundamental scenario, involves full knowledge of target input, with prior experience therefore resolving the 'cocktail-party' selection problem beforehand.

Chapter 3

Research problem

The neural mechanisms that underlie the brain's segregation of relevant speech from competing noise, the cocktail party effect, remain poorly understood. Furthermore, the nature of the auditory cortical responses of primed targets versus maskers has not yet been described at the principal stages of auditory cortical processing.

3.1 Objectives

3.1.1 General Objective

- To investigate the interaction between selective attention and prior experience during speech processing when listeners face simultaneous speakers, using non-invasive electrophysiological estimates of speech encoding (P1-N1-P2 components) of the TRF.

3.1.2 Specific Objectives

- To investigate the role of full prior knowledge in selective attention in a speech comprehension task under the 'cocktail party' problem (Experiment 1).

- To investigate the role of prior knowledge of speaker's voice in selective attention in the same task (Experiment 2a).

- To investigate the role of prior knowledge of speech content message in selective attention in the same task (Experiment 2B).

3.2 Hypothesis

3.2.1 Behaviour

- We hypothesize a higher level of comprehension when listeners are primed with information about the target.

3.2.2 Neural response

- We hypothesize the interaction between selective attention and prior experience in the neural processing associated with resolving the 'cocktail-party'.
- Regarding selective attention, we anticipate a systematic enhancement in neural tracking response for the attended speaker compared to the unattended speaker at N1 and P2 stages of cortical processing.
- Regarding prior experience, we anticipate that priming effects observed, if any, will be consistent with suppression of the neural tracking response to the primed stream.

Chapter 4

Methodology

4.1 Participants

Seventy-four human subjects participated across three experiments (48 female; mean age 25.6 ± 4.9 SD; 11 left-handed). Subjects were native Rioplatense Spanish speakers who reported no hearing impairment or neurological disorder and normal or corrected vision. The participants completed the Edinburgh Handedness Inventory.

The Ethics in Research Committee at the Faculty of Psychology, Universidad de la República Uruguay, approved the study. All participants gave their written consent in accordance with the guidelines of the Declaration of Helsinki (World Medical Association, 2009). As a token of appreciation for their time, the participants received a chocolate bar at the conclusion of the experimental session.

Thirty-six subjects participated in Experiment 1 (27 female; mean age 26.5 ± 4.7 SD; 7 left-handed). One participant was excluded due to a negative score on the task, while the behavioural response file of the other participant was not saved correctly.

Seventeen subjects participated in Experiment 2a (7 female; mean age 24.1 ±4.8 SD; 3 left-handed). One participant was excluded due to the behavioural response file was not saved correctly. Twenty-one subjects participated in Experiment 2b (16 female; mean age 25,7 ± 6 SD; 1 left-handed).

4.2 Setup

Participants were seated 50 cm in front of a 40 cm size monitor (E. Systems, Inc., CA). For audio presentation, we used a Sound Blaster Z sound card (Creative Labs, Singapore) in combination with a Scarlett 414 sound interface (Focusrite Plc, UK) and high-quality Sennheiser HD 25 headphones (Sennheiser, Germany). Neural recordings were made with BioSemi ActiveTwo 64-channel system with an Active two ERGO Opticlink (BioSemi, The Netherlands) to connect synchronous auditory signals to the ActiveTwo output system. The participants adjusted the presentation of the sound volume to a comfortable listening level before starting the computerized tasks.

4.3 Stimuli

Auditory stimuli were obtained from two databases constructed explicitly for the study by our team (see Database 1, Database 2). Database audios consisted of short phrases narrated by a single speaker, where keywords and topics have been annotated per audio. Additionally, each audio was classified according to the perceived age and sex group of the speaker.

MATLAB® software (MathWorks, Natick, United States) was used to convert audio recordings into diotic formats (simultaneous presentation to both ears) at a sampling rate of 44100 Hz. The low- frequency amplitude envelopes in the speech signals were obtained using Hilbert's method.

Long silences and repeated filler words were manually removed based on the stimulus envelope and content features on a case-by-case basis. To prevent the audible perception of clicks, 5 ms long cosine ramps were applied at the beginning or end of the stimulus and any excisions. All individual stimuli were normalized.

4.3.1 Database 1

Database 1 (D1) consisted of 259 single-speaker audios (129 female voices, mean duration 8.6 s, ±0.8 s). D1 was based on various source formats, such as news broadcasts, audiobooks, and interviews with Uruguayan Spanish speakers. In addition, audio clips were extracted from radio and TV channels, podcasts, and YouTube videos. All audio sentences were selected for explicit thematic content issued by a single speaker without background noise. In D1, four keywords were chosen from each audio sentence, including adjectives, nouns, verbs, and adverbs, but excluding those near the end of the sentence. Additionally, one or two topical categories were assigned to each sentence, from a fixed pool of topics that included politics, government, sports, religion, education, health, leisure, personal, culture, nature, labour, commerce, and money. Furthermore, apparent voice gender (male or female) was designated according to the name, face, and voice quality of the source material. The audios were also identified according to the perceived age of the speakers, i.e., less than or older than 35 years, resulting in 59 younger and 200 older speakers.

Audio	Politics	Gov't	Sports	Religion	Educ.	Key1	Key2	Key3	Key4	M/F	Age
1	X	X		X		Biases	Brain	Financial	Invest	F	Younger
2	X					Retire	Write	Journal	Stuff	M	Older
3		X		X		Tourism	Illusio n	Transfor m	Experienc e	M	Younger
4	X					Advance s	Rights	Money	Depressio n	F	Older

Table 1: Example of the Database registration

4.3.2 Database 2

Database 2 (D2) consisted of 348 single-speaker audios containing 184 female voices. The mean duration of the audios was 8.08 s, SD ± 1.97 s. Individual sentences were created and recorded using the Microsoft Research Paraphrase Corpus (MRPC) (Quirk et al., 2004), a database that records pairs of headlines or sentences containing similar information on the world news from different newspapers and sources. We selected 169 paraphrase pairs from the MRPC rated as semantically equivalent, and translated them into Spanish. Non-Spanish names (e.g., people, companies, or localities) from the local context and substituted them with familiar and local/regional names using an online random name generator (https://generadordenombres.online/). An example of a resulting paraphrase pair used for D2 audios shown in Table 2.

MRPC sentence 202294	MRPC sentence 202143 (paraphrase)
El procedimiento comenzó con la presentación por parte de los fiscales del caso contra Adolfo, leyendo 33 páginas de documentos que resumían las acusaciones contra él.	El procedimiento comenzó con la expresión por parte de los fiscales de su caso contra Adolfo, leyendo una carta de acusación de 33 páginas al tribunal.
The proceedings were taken up with prosecutors outlining their case against Amrozi, reading 33 pages of documents outlining allegations against him.	Proceedings were taken up with prosecutors outlining their case against Amrozi, reading a 33-page accusation letter to the court.

Table 2: Example of a paraphrase pair used in D2

Audio recordings were obtained from locally surveyed volunteers, who read aloud one or two pairs of paraphrases simulating a spontaneous conversation. An

Olympus recorder model VN-541OC (OM Digital Solutions Corporation, Japan) was used with settings suitable for recording in the open air (44.1 kHz sampling rate, 40 Hz - 13 kHz frequency response, 5 - 320 kbps bit rate range). Selected recordings were made in noise-free open or closed spaces, and contained clear message narrations as per the indicated sentence(s). Resulting audios were annotated for their speaker age, gender and keywords following similar criteria as in D1, with the exception that keywords were limited to two per phrase. The selected keywords always differed across paraphrasing pairs.

4.3.3 Cocktail party stimuli

A unique set of cocktail party (CP) stimuli was created per study participant. Each CP stimulus was generated from a selection of two solo speech stimuli from the same database by additive combination of the two streams using a 750 ms inter-stream onset asynchrony. This design was used for target instruction purposes, to ensure that listeners attend either the one single speech stimulus that was presented initially or the other speech stimulus interjected later. Each CP stimulus end time was equal to the resulting mixture overlap time. The root mean squared (RMS) intensity was equalized between the overlapping epochs of both streams, and the amplitude of the entire cocktail stimulus was normalized.

4.4 Experimental task

Participants were instructed to listen to every auditory stimulus presentation with their eyes closed, focused solely on speech without visual distractions while minimizing all other motor activities. A two-stage design was used for each trial in the task, to investigate the interaction between previous listening experiences and attention, specifically how previous listening experiences may affect listeners' capacity to concentrate on a particular speaker in a multi-talker scenario. In the first part, 'pre-listening', participants were presented with a single-speaker speech stimulus. In the second part, 'CP listening', participants were tasked with selectively attending to either the first (leading) or second (interjecting) speech stream. Leading or interjecting targets were cued immediately beforehand, and were instructed with equal probability. Upon listening, the trial continued with a questionnaire designed to evaluate participants' comprehension of the target speaker, and involved selecting three options from a multiple-choice format containing nine choices. Subsequent to participant's response validation, the score feedback of the trial was provided along with an accumulated score for the experiment trial. See Figure 1.

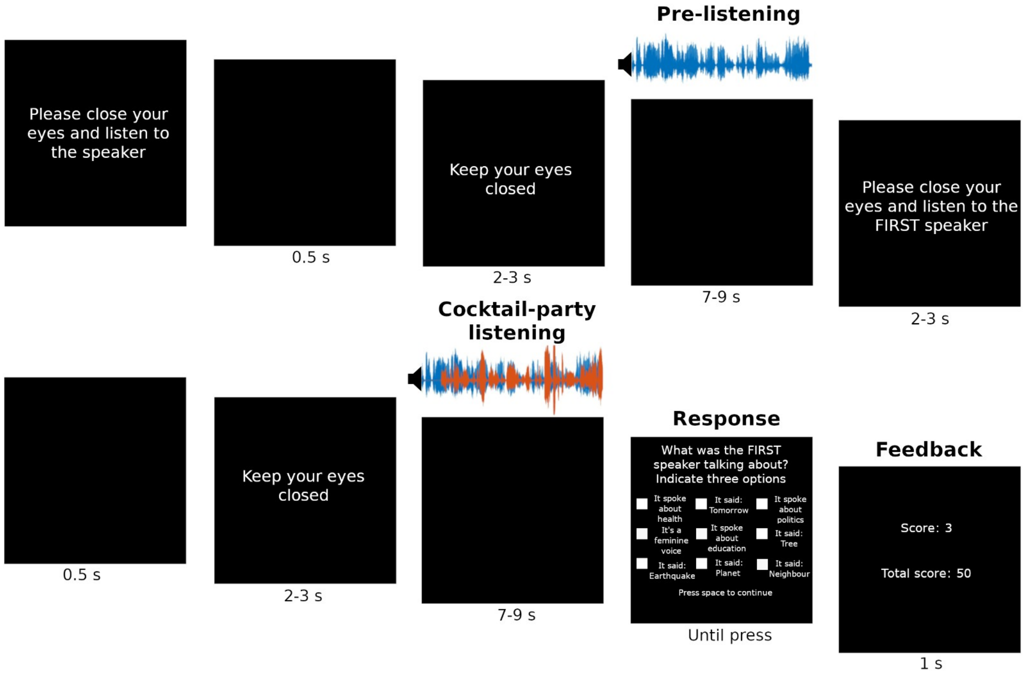

Figure 1: Experimental design

Each trial consisted of two parts, the first of which was a pre-listening stage where participants listened to a single speaker speech stimulus. In the second part, CP listening, participants were given instructions to attend to one speaker from a 'cocktail-party' stimulus, i.e., a speech mixture of two different speakers. After each trial, participants were required to answer multiple-choice questions about the attended speech. This example illustrates the 'attended known' condition (AK), where the pre-listening stage provides full prior knowledge about the target element of the CP mix.

4.4.1 Experimental conditions

Three conditions, namely, attended known (AK), unattended known (UK), and no prior knowledge (NP), were employed in all studies. Within each experiment, these conditions provided the listener with different degrees of knowledge about individual streams of the CP listening stage. Generally, for the AK trials, the pre-listening presentation matched with information contained in the target stream. In UK trials, the pre-listening presentation matched with information contained in the background stream. NP trials by contrast used pre-listening presentations without any connection or prior exposure to either the attended or unattended stream of the CP listening stage; these trials served as a control condition to assess baseline performance across all experiments. Attentional target cues were instructed at the start of the CP-listening stage, and were issued on the basis of leading or interjecting order ("first/second") of the asynchronous order mixture (Figure 1).

Each pre- and CP-listening stimulus was presented only once. Prior to the main experimental session, participants performed a 3-trial practice session with fixed, non-database stimuli that were specifically generated for this purpose. In all experiments, all speakers varied from trial to trial, as random and unique stimuli combinations were created per participant without stimulus replacement. All three conditions were equally represented and randomly ordered during the experimental trials.

4.4.1.1 Experiment 1

Experiment 1 (full priming) was designed for participants to gain prior knowledge with a clean auditory stimulus that will later act as foreground or background during masking at the selective attention stage. Therefore, for AK or UK trials the pre-listening stimulus was identical to one of the two elements of the mixture. In the AK condition, participants were instructed so that they effectively focus, during CP-listening, on a target stream that is identical to the one presented in the pre-listening period (Figure 1 shows this case). Conversely, in UK trials, participants were instructed so that they effectively disregard, during CP-listening, a stimulus that is identical with the one presented during the pre-listening stage as it now acted as the background stream. All stimuli were constructed using samples from D1. The task comprised 108 trials, and the total experimental session lasted approximately 2.5 hours.

4.4.1.2 Experiment 2a

Experiment 2a (voice priming) was designed for participants to gain prior knowledge with a speaker's voice that will later act as foreground or background during masking at the selective attention stage. NP trials were generated as in Experiment 1, with stimuli from this condition also sampled or constructed from D1. AK/UK trials were generated so that each new speaker presenting at pre-listening corresponded with one of the two speakers of the mixture during cocktail presentation. In the AK condition, participants were instructed so that they effectively focus during CP-listening on a target stream whose speaker's voice is the same as presented in the pre-listening period, thus with voice-priming of the target. Conversely, in UK trials participants were instructed so that they effectively disregard, during CP-listening, a speaker whose voice coincided with that during the pre-listening stage as it now acted as masker. This created voice-priming of the background stream. Because the same speaker was involved at both stages in the same trial, we ensured that the repeating speaker stimuli pairs always involved two different sentences and contents. AK/UK stimuli were exclusively sampled or constructed from D2. We also discarded the possibility that repeated content sentences (e.g., paraphrases) of the same or potentially two different speakers were re-used at any point for a participant. The experiment consisted of 78 trials and the total duration of the experimental session was approximately 1.5 hours.

4.4.1.3 Experiment 2b

Experiment 2b (content priming) was designed for participants to gain prior knowledge with a spoken message that will later act as foreground or background during masking at the selective attention stage. NP trials were generated as in the previous experiments, with stimuli also sampled or constructed from D1. AK/UK trials were generated so that each spoken message at pre-listening corresponded with one of the two spoken messages of the mixture during cocktail presentation. In the AK condition, participants were instructed so that they effectively focus, during CP-listening, on a target stream whose message (i.e., paraphrase) matches the one presented during the pre-listening period, thus with content-priming of the target. Conversely, in UK trials participants were instructed so that they effectively disregard, during CP-listening, speech whose content coincided with the one presented during the pre-listening stage as this stream now acted as a masker. In this experiment, all speaker voices always differed from each other within and across trials. AK/UK stimuli were exclusively sampled or constructed from D2. The presentation order of the paraphrase sentence pairs was equally counterbalanced across participants. The task consisted of 78 trials, and the experimental session lasted approximately 1 hour and 30 minutes.

4.5 Task questionnaire

For all experiments, participant behavior was assessed via a multiple-choice questionnaire that was presented at the response stage screen. Based on stimulus-specific information collected from the corresponding databases, 9 options were generated for each trial. One of the options presented was related to vocal characteristics, with possibilities including determining the voice's gender (male or female) or age (young or advanced age) attributes, with equal probability. Four of the additional options involved sentence keywords, randomly drawn from a list of all 8 keywords from the attended and unattended stream (experiment 1) or 4 keywords (experiments 2a and 2b). The remaining four options listed possible topics of the target stream, which were presented based on a random selection of four out of the 13 (experiment 1) or 14 (experiments 2a and 2b) categories. It was possible that none of the presented topic options corresponded to either stream. The placement of the nine different options was randomized for every trial.

4.5.1 Scoring system

In all three experiments, identifying an option that was correctly related to a speaker's voice (such as sex or age) earned one point. Selecting an option that identified voice characteristics of the unattended speaker and not the attended one always resulted in a point deduction (e.g., if participants had to attend to a younger age and selected an older one). Similarly, each correctly chosen keyword always awarded one point, and each selected keyword related to the unattended speaker always deducted one point. Each correct topic selection earned one point, but topic selections unrelated to the attended speech did not result in any point deductions. In Experiment 1, not selecting any option resulted in a deduction of -2 points (equivalent to missing two correct keywords on average). In Experiments 2a and 2b, not selecting any option resulted in a score of 0 for the trial. Participants received feedback on the trial score and on the accumulated score (see Figure 1).

4.6 Data Acquisition and Preprocessing

Presentations were made with PsychoPy3 software (Peirce, 2007). Electroencephalography (EEG) data were recorded for all subjects using the BioSemi ActiveTwo 64 scalp channels system (BioSemi, The Netherlands), with 10/20 layout, at a digitization rate of 2048 Hz with CMS/DRL (ground) and the tip of the nose as reference electrodes. EEG subject data resampled at 1024 Hz were common average-referenced to the 64 scalp channels, after which DC offset was removed. A 5th order cascaded integrator-comb low-pass filter with -3 dB at 410 Hz was applied online, after which signals were decimated to 1024 Hz. The online high-pass response was fully DC coupled. Electrooculographic data were recorded, supra- and infra-orbitally as well as from the left versus right orbital rim. Complete experimental sessions lasted ~ 2.5 h.

Single channel data were subsequently (offline) and channels were subject to an automated blind rejection procedure based ona variance-based criterion (Junghöfer et al., 2000). In order to account for individual differences that could impact spatial filter estimation, a blind rejection procedure was applied to the channel and trial time series, incorporating data from all participants simultaneously, with a confidence coefficient of $\lambda_P = 4$. The process was repeated exclusively for external reference channels with a more restrictive confidence coefficient of $\lambda_P = 2$. Sensor and reference data sets were separately re-referenced based on the channel sets median values, after which the channels rejection procedure was again conducted. FastICA (Hyvarinen, 1999) was applied to minimize the impact of general movement artifacts on the EEG data. Two independent components were automatically selected for their maximal proportion of broadband power and projected out of the raw data. To further reduce ocular artefacts, a time-shifted principal component analysis (de Cheveigné and Simon, 2008) was applied to discard environmental signals recorded on the oculogram reference sensors (shift: ±4 ms). A sensor noise suppression algorithm (de Cheveigné and Simon, 2008) was applied to attenuate artifact components specific to any single channel (up to 63 neighbors). The blind variance-based rejection procedure was repeated, resulting in fewer than 1% rejected single channel trial time series on average (subject range 0.09 - 1.52%).

To extract the reproducible denoising by spatial filtering (DSS) component, first, EEG data were bandpass-filtered with a second-order Butterworth filter in the 1 – 8 Hz region. To address variability across participants, which may affect spatial filter estimation, the blind rejection procedure described above was performed over channel and trial time series, including all participants' data simultaneously and with a more restrictive confidence.

4.6.0.1 Spatial filtering

To emphasize signal reproducibility and generalization across subjects, joint decorrelation (JD) (de Cheveigné and Parra, 2014; de Cheveigné and Simon, 2008) methods were implemented. For this, a generic spatial filter obtained after solo speech listening data from a separate study (Vanthornhout et al., 2018) based on the same EEG system was used with the aim to improve the SNR of speech responses in the present study and reduces dimensionality of single-trial, single speaker speech listening EEG datasets. Joint decorrelation was trained on this full independent dataset involving 28 listeners' recordings during passive listening to a 14.5 minute long story and presented without noise (Vanthornhout et al., 2018). The study files were pre-processed as described above, and each dataset was epoched, downsampled to 32 Hz, and collated across all participants. The spatial profile associated from this filter is represented in Figure 5, as the component with maximal evoked/induced activity ratio (de Cheveigné and Simon, 2008). This filter component represents the most reproducible aspect of evoked responses and was selected as a single virtual sensor in all analyses, i.e., was retained and applied to all subjects in the present studies. The procedure is effectively analogous to sensor selection in standard ERP studies, solved here via a data-driven method.

4.7 Temporal response function estimation

The Temporal Response Function (TRF) is a method to analyze the encoding dynamics in neural recordings of neural time series data. The TRF is a linear systems tool that can be used to represent the neural response to changing auditory input over time (Crosse et al., 2016). The stimulus S(t) was here represented by the time series of the speech envelope onsets of pre-listening stimuli, $S_{prelistening}(t)$, and of CP stimuli, which was decomposed as $S_{cocktail}(t) = S_{target}(t) + S_{masker}(t)$, where the two-time series separately index the envelope onsets of the target and masker stimuli, respectively. To correctly reflect the required task demands, we excluded from neural tracking analyses all trials where participants earned non-positive scores (10-20% of trials, see Behavioral results). For each stimulus category and condition, all resulting single trial presentations were concatenated, resulting in data equivalent to up to 15 min. presentation time (pre-listening) and 2.5 min. (cocktail listening, each condition), per participant. The linear model is formulated as:

$$r'(t) = \sum_{\tau} TRF(\tau)S(t - \tau) + \epsilon(t)$$

where r'(t) is the neural response predicted by the model $TRF(\tau)$ under the conditions set by the chosen stimulus representation $S(t)$, and $\epsilon(t)$ is the residual contribution to the evoked response not explained by the linear model and cross-validation was performed to prevent overfitting. TRFs were estimated by boosting (David et al., 2007) between concatenated stimulus and EEG time series, scaled to z-units, with 20 times cross-validation, and assessed at the -100 to 400 ms window. Each grand average TRF was obtained by averaging the individual TRFs across participants.

4.8 Statistical analysis

We first determined the presence of general attentional effects (AEs) at the cocktail party, to determine the time windows relevant for selection unbiased by any priming effects. For this, non-parametric tests (Maris, 2011; Martin et al., 2008) were applied using cluster-based non-parametric testing, corrected for multiple comparisons. We used subject TRFs from the no prior knowledge (NP) condition across all three experiments (71 subjects, see Participants section). This also served to compare our data with previous literature on the classic cocktail party effect. For each contrast between NP-attended and NP-unattended TRFs, its t-value time series was estimated and thresholded at the t distribution's 0.05th percentile for randomisation-based testing. Any cluster exceeding the threshold was deemed significant if its associated t-statistic (sum of t-values within the cluster) exceeded the t-statistics estimated from the randomization distribution (Maris and Oostenveld, 2007). Nonparametric statistical testing of EEG-and MEG-data. Journal of neuroscience methods, 164(1), 177-190.. The distribution was estimated after $N = 2^{17}$ resamplings, each generated by random reshuffling between attended/unattended conditions, per participant.

Once the relevant attentional windows were determined (P1, N1 and P2 components, see Results), we examined the statistical interactions between prior experience (AK, UK, and NP conditions) and attention (attended and unattended condition) by means of two-way repeated measures ANOVAs of time-averaged TRF data as dependent variable, quantified per component and listener. To interpret valid F-ratios, data sphericity violations were determined based on Mauchly testing; for cases where sphericity could not be assumed, Greenhouse-Geisser ($\varepsilon < 0.75$) or Huynh-Feldt ($\varepsilon > 0.75$) corrections were applied (Field, 2017).

Chapter 5

Results

5.1 Behavioural results

5.1.1 Experiment 1- Full priming

Participants completed the task with accuracy, i.e., scored 1 point or more on target-specific speech information, for greater than 50% of trials in all conditions, indicating that the instructions were understood and generally answered correctly. In the no prior knowledge conditions (NP), participants successfully performed for most of the trials (mean success rate 80.9 ± 1.6% SD). Given full prior knowledge of the foreground (AK condition), the participants presented a higher percentage of correct answers on average (84.3 ± 1.3%). Given full priming of the background stimulus (UK condition), average performance decreased (79 ± 1.7%).

To compare the accuracy of CP-listening performance of the correct rates in the three conditions, a one-way repeated measures ANOVA was performed on participant data with prior knowledge condition as independent variable. There was a significant main effect of prior knowledge ($F(2, 66) = 5.093$, $p = 0.009$). Post-hoc pairwise comparisons of correct rates revealed a significant difference between AK and UK conditions (corrected $p = 0.001$). As a result of target priming, improved performance was observed for known targets in comparison to known maskers. However, neither the NP vs. UK (corr. $p = 0.799$) nor the AK vs. NP (corr. $p = 0.267$) comparisons reached statistical significance. See Figure 2. The behavioral data indicate that prior knowledge facilitated participants' solving of the cocktail party effect task when it carried information about the attended stream compared to the unattended stream.

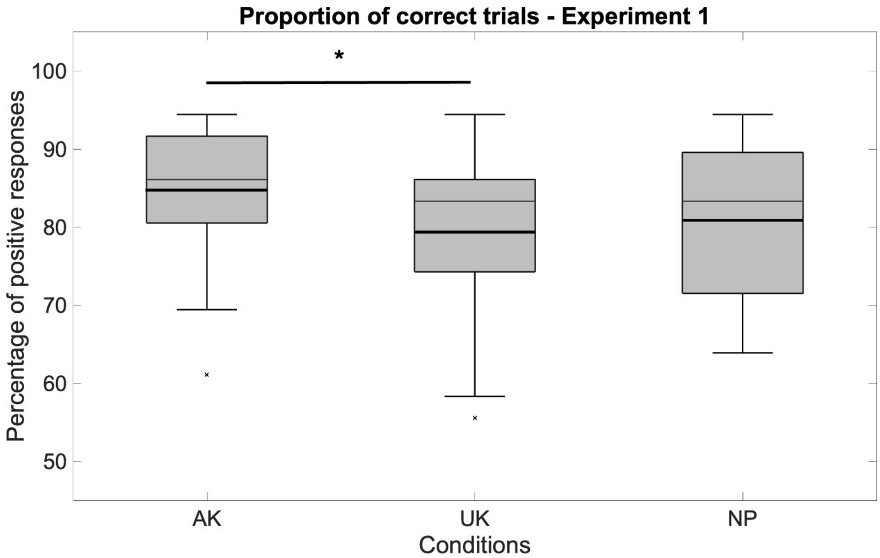

Figure 2: Behavioural results in Experiment 1 - Full priming

Percentage of positive trial responses by prior knowledge condition. A significant effect of prior knowledge about the full auditory stimulus was found, where target priming improves participants' correct trials relative to masker priming ($p < 0.001$). Bars represent standard error of the mean. The black divider line represents the distribution mean and the grey divider the median.

5.1.2 Experiment 2a - Voice priming

The participants generally performed as well as in Experiment 1 on the comprehension task, with greater than 80% of correct trials. The percentage of positive responses for trials without any prior knowledge (NP) was on average 89.4 ± 4.7%. Given prior knowledge of the target speaker's voice (AK), participants achieved on average 81.7 ± 14.2% correct trials. Given prior knowledge of the masker's voice (UK), participants achieved on average 89.9 ± 5.5% correct trials.

The accuracy of CP-listening performance of the correct rates in the three conditions was analyzed using a one-way repeated measures ANOVA. There was no significant main effect of prior knowledge ($F(1.146, 17.190) = 3.941$, $p = 0.059$). See Figure 3.

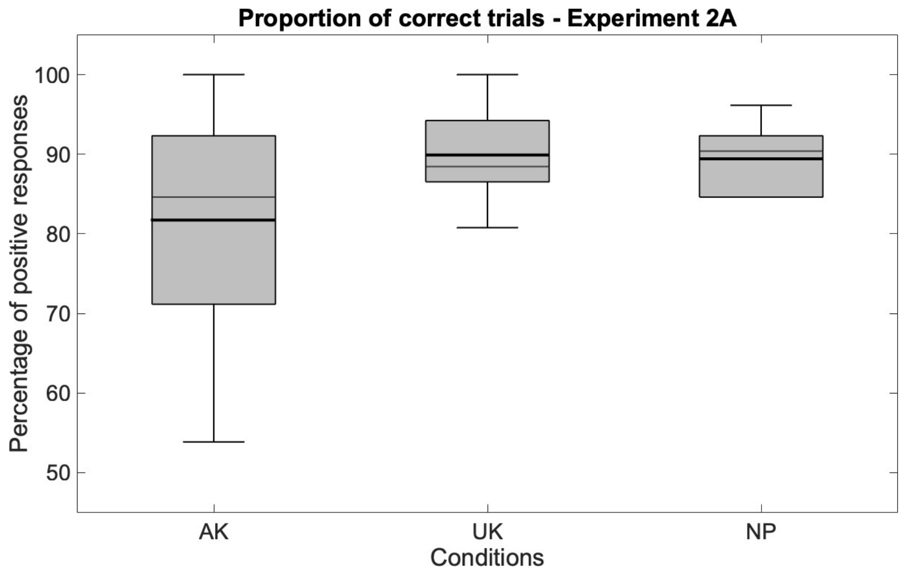

Figure 3: Behavioural results in Experiment 2A - Voice priming

Percentage of positive trial responses by prior knowledge condition. No significant effects of prior knowledge about speaker's voice were observed on CP task performance.

5.1.3 Experiment 2b - Content priming

The participants generally performed as well as in the previous experiments on the comprehension task, with greater than 80% of correct trials. The percentage of positive trial responses for trials without any prior knowledge (NP) was on average 90.1 ± 8.3%. Given prior knowledge of the target speaker's voice (AK), participants achieved on average 84.4 ± 12.7% correct trials. Given prior knowledge of the masker's voice (UK), participants achieved on average 86.2 ± 7.7% correct trials.

The accuracy of CP-listening performance of the correct rates in the three conditions was analyzed using a one-way repeated measures ANOVA. There was no significant main effect of prior knowledge ($F(2, 40) = 1.921$, $p = 0.160$). See Figure 4.

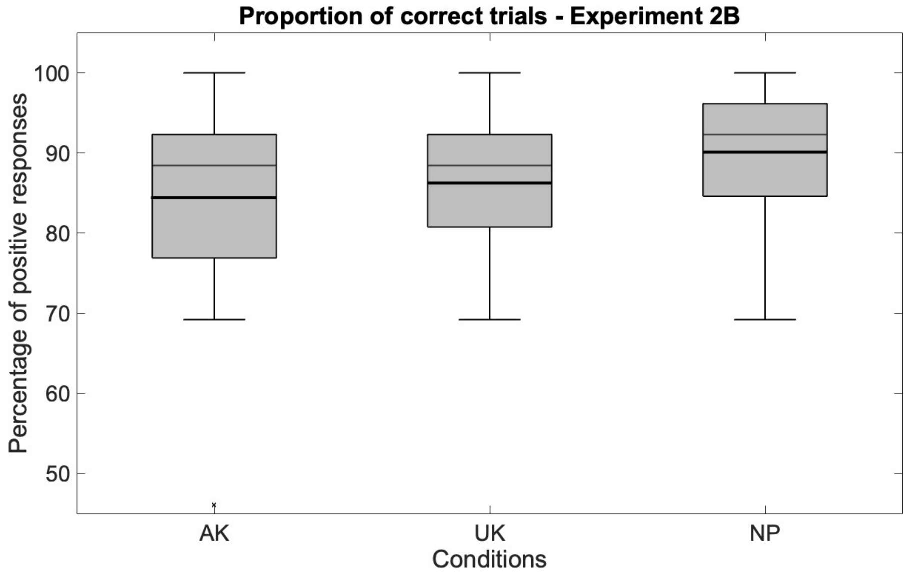

Figure 4: Behavioural results in Experiment 2b - Content priming

Percentage of positive trial responses by prior knowledge condition. No significant effects of prior knowledge about speech content were observed on CP task performance.

5.2 Neural results

The relevant windows of auditory processing were first addressed for single speaker (solo) speech using data from pre-listening stages, collated across all studies. See Figure 5. This clean speech grand-average TRF presented a characteristic triphasic response, where a smaller positive deflection is observed within the first 50-78 ms, corresponding to the P1 component. A negative peak, N1, is observed at approximately 125 ms. Furthermore, a greater positive peak, P2, emerges later at approximately 200 ms. The relevant windows of auditory processing during the CP stage were then defined from a non-parametric analysis of subject TRFs from attended and unattended speech at the no prior knowledge (NP) condition across all three experiments (71 subjects, see Statistical analysis). This initial test revealed significant modulations of selective attention on the attended versus unattended speech TRFs, i) between 31-78 ms (corresponding to the P1 window), ii) between 100-180 ms (N1 window), and iii) between 230-320 ms (P2 window), p

= 1.5259×10^{-5}. For all three windows, the differences related a higher amplitude for window-averaged TRF estimates to the target than the background speech. The three component windows were then fixed for statistical analyses of TRF data across all experiments.

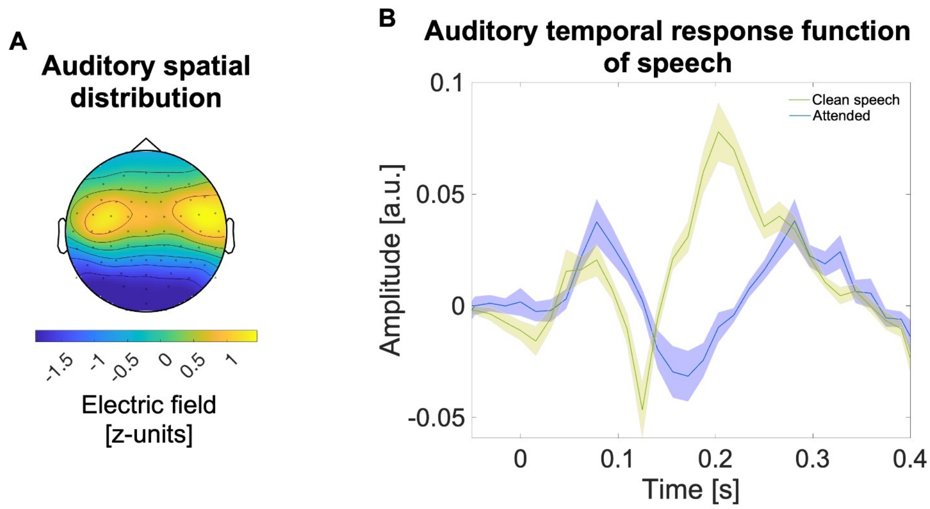

Figure 5: Auditory tracking of the envelope onset signals of solo and CP speech

A. Auditory EEG distribution from the top spatial component resulting from the data-driven spatial filtering procedure that was applied to extract an auditory recording from EEG data (see Spatial filtering). B. Auditory temporal response function (TRFs) of clean and cocktail-party speech. Grand average of TRFs from 71 participants in pre-listening (clean speech condition, yellow) and CP-listening (NP-attended, blue). In the pre-listening condition, participants listen to a single speaker, which mimics optimal conditions for neural speech encoding. By contrast the NP-attended TRF represents processing of a target under the relatively less intelligible cocktail party paradigm. Despite differences in latency, three distinct responses are observed in auditory processing across both paradigms, the positive P1 component, the negative N1, and the P2. Across all studies, TRFs model the neural tracking of the envelope onsets of speech.

5.2.1 Experiment 1. Full priming

To investigate the statistical interaction between selective attention and prior auditory experience/knowledge of a speech stimulus at each of the relevant speech processing stages (P1, N1, P2), a two-way repeated measures ANOVA was conducted per TRF component estimate of CP listening data.

5.2.1.1 P1

A two-way repeated measures ANOVA (Attention × Experience) conducted on TRF estimates of the P1 component did not reveal a statistically significant interaction between selective attention and prior knowledge (F(2, 66) = 0.105, p = 0.901). At this early stage, there was a significant main effect of selective attention (F(1, 33) = 7.015, p = 0.012), and there was no significant main effect of prior knowledge (F(1.662, 54.833) = 0.987, p = 0.366). The data suggest that, while selective attention may modulate the neural tracking of onsets in the envelope of speech at this early stage (Figure 6), this modulation may not be influenced by prior experience of the target or masker auditory stimuli.

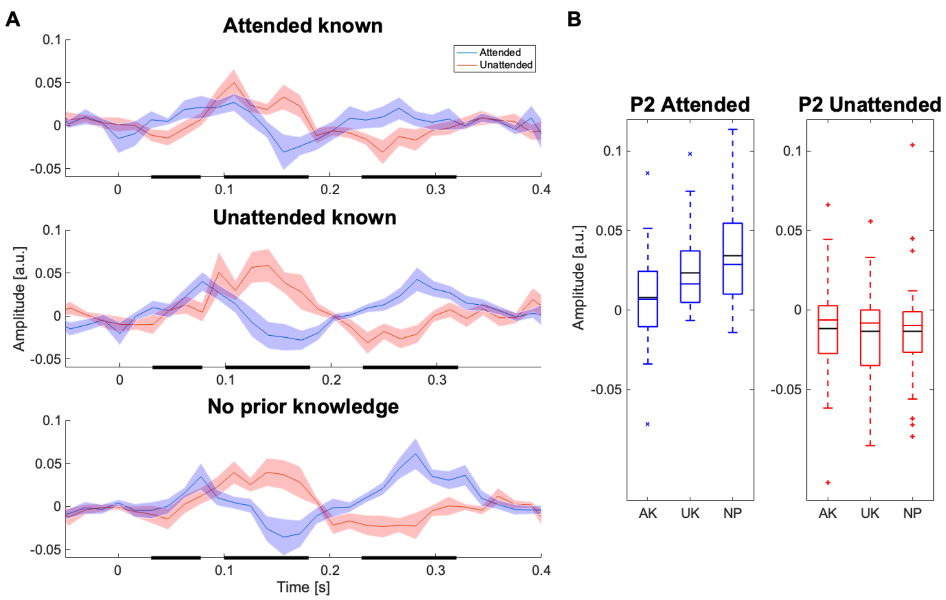

Figure 6: Results from Experiment 1 - Full priming

A. TRF models of neural tracking of the envelope onsets from attended (blue) and unattended speech (red), across prior knowledge conditions. The shaded areas represent the standard error of the mean. The thick bars on the time axis represent analysis windows corresponding to the P1, N1, and P2. The shaded areas represent the standard error of the mean. B. Interaction between attention and prior knowledge at the P2 stage. TRF estimates of the P2 (230-320 ms) from attended and unattended streams under AK, UK and NP conditions, where a significant decrease of the P2 corresponding to the target stream is observed under priming.

5.2.1.2 N1

A two-way repeated measures ANOVA (Attention × Experience) conducted on TRF estimates of the N1 component did not reveal a statistically significant interaction between selective attention and prior knowledge (F(2, 66) = 1.488, p = 0.233). At this relevant stage, there was a significant main effect of selective attention (F(1, 33) = 26.568, p < 0.001), and there was no significant main effect of prior knowledge (F(2, 66) = 0.221, p = 0.802). The data are in agreement with previous findings that selective attention may modulate the neural tracking of onsets in the envelope of speech at this stage, with enhanced responses to the target versus masker speech (Figure 6). The results however do not support the hypothesis that such modulation is influenced by prior experience of the target or masker auditory stimuli.

5.2.1.3 P2

A two-way repeated measures ANOVA (Attention × Experience) conducted on TRF estimates of the P2 component revealed a statistically significant interaction between selective attention and prior knowledge ($F(2, 66) = 4.745$, $p = 0.012$). At this late stage, there was a significant main effect of selective attention ($F(1, 33) = 49.128$, $p < 0.001$), and there was no significant main effect of prior knowledge ($F(2, 66) = 2.927$, $p = 0.061$). The data are also agreement with previous findings that selective attention may modulate the neural tracking of onsets in the envelope of speech at this stage, with enhanced responses to the target versus masker speech (Figure 5). The results however indicate that such modulation is influenced by prior experience of the target or masker auditory stimuli.

To address the significant interaction, the effect of prior knowledge was evaluated on post-hoc analyses of attended and, separately, unattended speech . A one-way repeated measures ANOVA (Experience) conducted on TRF estimates of the P2 component related to the attended stream revealed a significant effect of prior knowledge ($F(2, 66) = 7.186$, corr. $p = 0.004$). A similar analysis conducted on estimates related to the unattended stream did not reveal a significant effect of prior knowledge ($F(2, 66) = 0.044$, corr. $p > 0.957$). Hence, priming modulated the neural coding of envelope onsets from the attended stream, while for unattended streams these responses remained similar across priming conditions.

Subsequent post-hoc pairwise comparisons of TRF estimates of the attended stream at the P2 stage revealed a significant difference between AK and NP conditions (corr. $p = 0.006$). As a result of target priming, a considerably reduced P2 was observed for known targets in comparison to control conditions. However, neither the NP vs. UK (corr. $p = 0.289$) nor the AK vs. UK (corr. $p = 0.082$) comparisons reached statistical significance. See Figure 6.

In summary, the results confirm that, in the cocktail party paradigm, selective attention improves the neural tracking of envelope onsets from attended speech compared to unattended speech at the P2 stage. This enhancement can be modulated by prior experience, indicated by reduced tracking of envelope onsets from attended speech if the same single auditory stimulus has been fully primed. This prior experience had an impact on the late neural tracking of attended speech.

5.2.2 Experiment 2a. Voice priming

To investigate the statistical interaction between selective attention and prior auditory experience/knowledge of a speaker's voice at each of the relevant speech processing stages (P1, N1, P2), a two-way repeated measures ANOVA was conducted per TRF component estimate of CP listening data.

5.2.2.1 P1

A two-way repeated measures ANOVA (Attention × Experience) conducted on TRF estimates of the P1 component did not reveal a statistically significant interaction between selective attention and prior knowledge ($F(2, 30) = 0.087$, $p = 0.917$). At this early stage, there was no significant main effect of selective attention ($F(1, 15) = 3.532$, $p = 0.080$), and there was no significant main effect of prior knowledge ($F(2, 30) = 1.293$, $p = 0.289$). The data from Experiment 2a do not confirm the finding from Experiment 1 that selective attention may modulate the neural tracking of onsets in the envelope of speech at this early stage. Similar to Experiment 1, this component was not found to be influenced by prior experience of the target or masker speaker's voice.

5.2.2.2 N1

A two-way repeated measures ANOVA (Attention × Experience) conducted on TRF estimates of the N1 component did not reveal a statistically significant interaction between selective attention and prior knowledge ($F(1.416, 21.235) = 0.014$, $p = 0.958$). At this relevant stage, there was a significant main effect of selective attention ($F(1, 15) = 17.179$, $p = 0.001$), and there was no significant main effect of prior knowledge ($F(1.453, 21.792) = 2.496$, $p = 0.118$). As in Experiment 1, the data are in agreement with previous findings of auditory object segregation via enhanced responses to foreground versus background speech at this stage (Figure 7). Similarly, the results do not support the hypothesis that such modulation is influenced by prior experience of the target or masker speaker's voice.

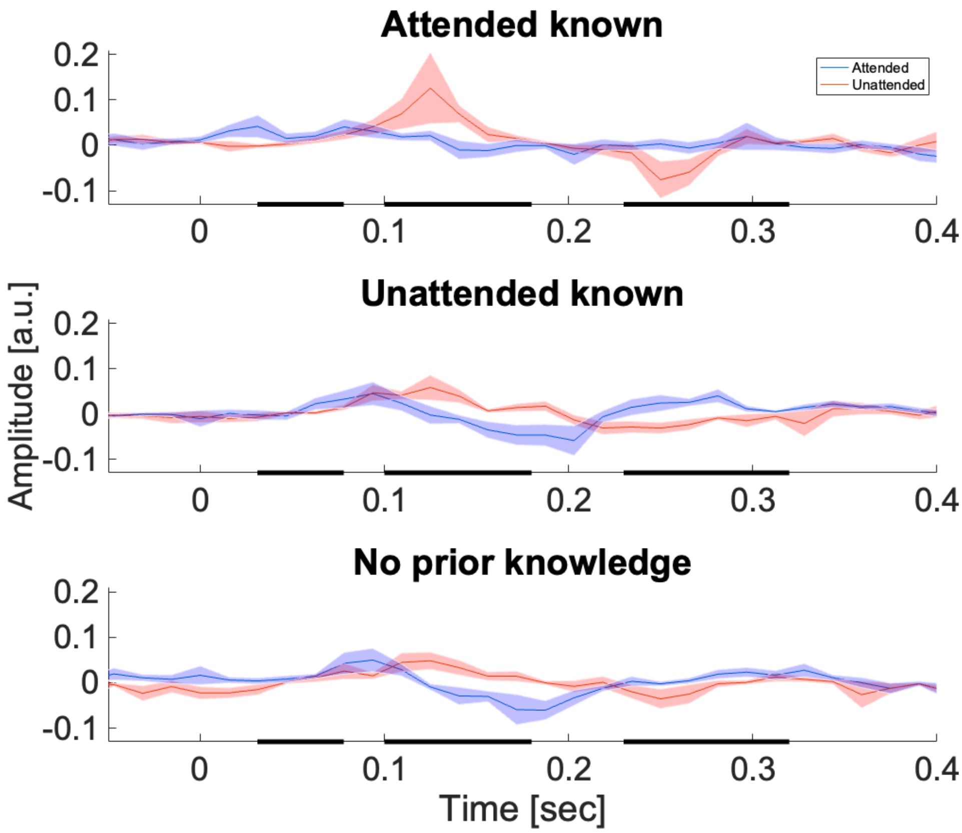

Figure 7: TRF results from Experiment 2A - Voice priming

TRF models of neural tracking of the envelope onsets from attended (blue) and unattended speech (red), across speaker's voice prior knowledge conditions. The shaded areas represent the standard error of the mean. The thick bars on the time axis represent analysis windows corresponding to the P1, N1, and P2.

5.2.2.3 P2

A two-way repeated measures ANOVA (Attention × Experience) conducted on TRF estimates of the P2 component did not reveal a statistically significant interaction between selective attention and prior knowledge ($F(2, 30) = 1.019$, $p = 0.373$). At this late stage, there was a significant main effect of selective attention ($F(1, 15) = 16.108$, $p = 0.001$), and there was no significant main effect of prior knowledge ($F(1.318, 19.771) = 1.009$, $p = 0.351$). The data are also in agreement with previous findings that selective attention may modulate the neural tracking of onsets in the envelope of speech at this stage, with enhanced responses to the target versus masker speech (Figure 7). The results however do not support the hypothesis that such modulation is influenced by prior experience of the target or masker speaker's voice.

5.2.3 Experiment 2b - Content priming

To investigate the statistical interaction between selective attention and prior auditory experience/knowledge of voiced content at each of the relevant speech processing stages (P1, N1, P2), a two-way repeated measures ANOVA was conducted per TRF component estimate of CP listening data.

5.2.3.1 P1

A two-way repeated measures ANOVA (Attention × Experience) conducted on TRF estimates of the P1 component did not reveal a statistically significant interaction between selective attention and prior knowledge ($F(2, 40) = 0.201$, $p = 0.819$). At this early stage, there was a significant main effect of selective attention ($F(1, 20) = 8.982$, $p = 0.007$), and there was no significant main effect of prior knowledge ($F(2, 40) = 0.904$, $p = 0.413$). The data from Experiment 2b replicate the finding from Experiment 1 that selective attention may modulate the neural tracking of onsets in the envelope of speech at this early stage. Similar to Experiments 1 and 2a, this component was not found to be influenced by prior experience of the target or masker speech content. See Figure 8.

5.2.3.2 N1

A two-way repeated measures ANOVA (Attention × Experience) conducted on TRF estimates of the N1 component did not reveal a statistically significant interaction between selective attention and prior knowledge ($F(2, 40) = 1.122$, $p = 0.336$). At this relevant stage, there was a significant main effect of selective attention ($F(1, 20) = 8.025$, $p = 0.010$), and there was no significant main effect of prior knowledge ($F(2, 40) = 0.663$, $p = 0.521$). As in Experiments 1 and 2a, the data are in agreement with previous findings of auditory object segregation via enhanced responses to foreground versus background speech at this stage (Figure 8). Similarly, the results do not support the hypothesis that such modulation is influenced by prior experience of the target or masker speech content.

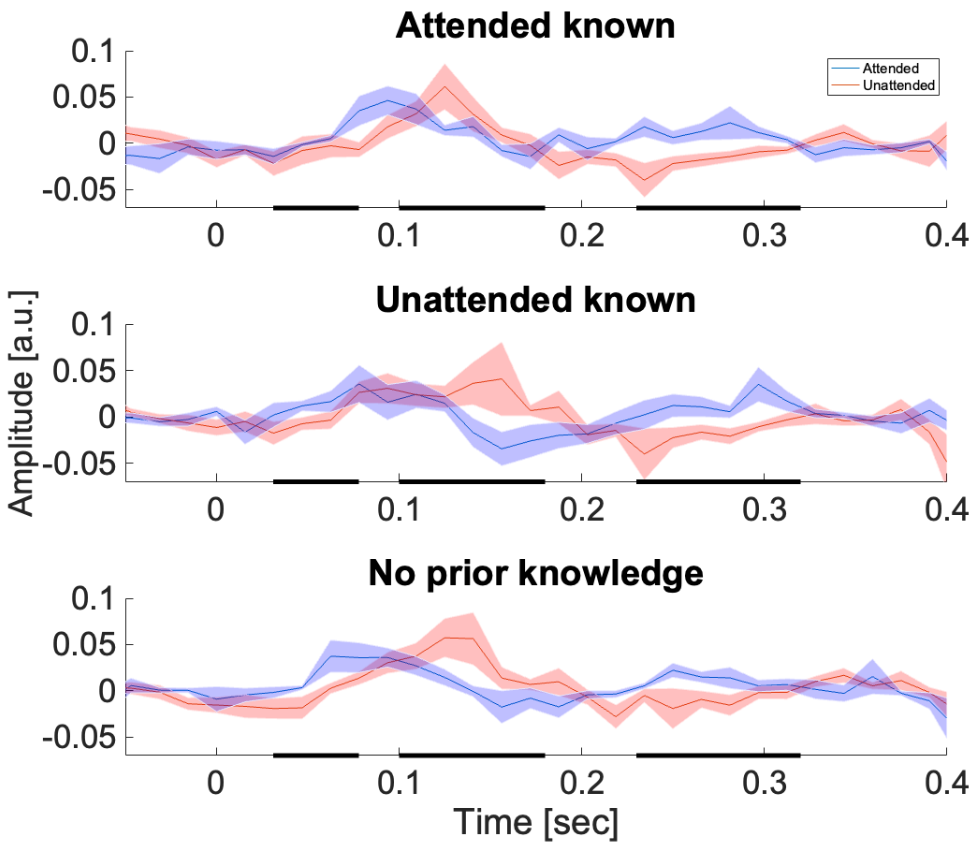

Figure 8: TRF results from Experiment 2b - Content priming

TRF models of neural tracking of the envelope onsets from attended (blue) and unattended speech (red), across speech content prior knowledge conditions. The shaded areas represent the standard error of the mean. The thick bars on the time axis represent analysis windows corresponding to the P1, N1, and P2.

5.2.3.3 P2

A two-way repeated measures ANOVA (Attention × Experience) conducted on TRF estimates of the P2 component did not reveal a statistically significant interaction between selective attention and prior knowledge ($F(2, 40) = 0.664$, $p = 0.520$). At this late stage, there was a significant main effect of selective attention ($F(1, 20) = 21.299$, $p < 0.001$), and there was no significant main effect of prior knowledge ($F(2, 40) = 0.352$, $p = 0.706$). See Figure 8. The data also confirm Experiments 1 and 2a findings that selective attention may modulate the neural tracking of onsets in the envelope of speech at this stage, with enhanced responses to the target versus masker speech. Similarly, to Experiment 2a, the results however do not support the hypothesis that such modulation is influenced by prior experience of the target or masker speech content.

Experiment	**Effect**	**P1 (31-78 ms)**	**N1 (100-180 ms)**	**P2 (230-320 ms)**
1 – Full priming	Attention	**0.012**	**<0.001**	**<0.001**
	Experience	0.366	0.802	0.061
	Attention x Experience	0.901	0.233	**0.012**
2a – Voice priming	Attention	0.080	**0.001**	**0.001**
	Experience	0.289	0.118	0.351
	Attention x Experience	0.917	0.958	0.373
2b – Content priming	Attention	**0.007**	**0.010**	**<0.001**
	Experience	0.413	0.521	0.706
	Attention x Experience	0.819	0.336	0.520

Table 3: Summary of neural results.
Summary of statistical significance levels for interactions and main effects across components and experiments, during speech envelope onset EEG tracking. Bold entries indicate significant effects (a<0.05).

Chapter 6

Discussion

6.1 General discussion

In demanding auditory situations, like a party, the human brain must allocate necessary resources to process relevant information. Prior experience with voiced speech input may enable the brain to form predictions that influence how speech is processed at subsequent times (Daly and Pitt, 2023; Pinto et al., 2019). Given the simultaneous operation of two complex cognitive mechanisms, attention and experience, this study investigated the impact of prior knowledge on the selective processing of speech during a 'cocktail party' task. Auditory processing in the human cortex was addressed using temporal response function (TRF) models from single-trial EEG recordings. To evaluate the effects of priming elements of speech in multi-talker environments, we focused on the three stages that were identified as modulated by attention, and which are typically associated with speech processing of envelope onset cues. Three experiments were designed to address whether and when does prior knowledge impact processing – from priming based on the lower-level analyses of acoustic features of vocal cues, to priming based on the higher-level analyses of lexical and semantic content.

In line with previous results, our data confirm findings of significantly enhanced N1 and P2 components of the TRF for attended compared with unattended speech, two processing stages typically modulated by selective attentional mechanisms. The main result of this work is however the observation of a significant interaction between attention and experience, observed during the late stage of envelope onset processing. Specifically, a reduction in the P2 component encoding of the attended stream was observed when the target was fully primed, i.e., when the identical auditory stimulus was presented masker-free beforehand. As for performance, target priming facilitated the comprehension compared to masker priming. We further provide some evidence of a modulatory effect of selective attention on the P1 stage.

6.2 Cortical processing under full auditory priming of a cocktail party speech stimulus

Behavioural responses showed that clean primed auditory objects promoted accurate comprehension when they subsequently acting as targets amid masking, then as the competing maskers of an unprimed target. These behavioural results are consistent with enhanced object detection by prior experience, resulting in a benefit to selective attention when the same object is a target, but interfering with by attentional capture it if presented as a distractor. Secondly, in the masker priming condition, the performance was poorer, suggesting an interference effect due to previous experience.

In parallel, neural responses of listeners to each speech object were analysed using the TRF technique (Crosse et al., 2016). Mapping from acoustic envelope onset features of each speech stimulus to their corresponding cortical tracking dynamics, listeners' neural responses were compared according to whether target or masker objects were primed or not. Mirroring cortical tracking results from clean speech "pre-listening" conditions, the neural representation of both target and masker auditory objects was found to display a similar triphasic structure, albeit with amplitude and delay variations in the P1, N1, and P2 components of attended and unattended streams. At the early stage of processing (P1), the temporal response functions (TRFs) of both streams were similar in terms of prior experience. We obtained evidence of an attentional modulation in the P1 stage, which was however not modulated by prior experience. While the P1 component is associated with the facilitation of sensory processing (Haykin and Chen, 2005), task-relevant visual spatial attention has been shown to enhance P1 (Luck, et al., 1990). Although this is typically observed in visual processing, similar attentional mechanisms may be implicated in auditory processing in this context (Brodbeck et al., 2020).

Our results also replicated the characteristic modulation of the N1 component by attention, in the form of a significant enhancement of the attended stream compared to the unattended speech, across condition. This represents a stage of processing proposed to reflect auditory object formation and segregated (Ding and Simon, 2012).

Importantly, the full priming study results indicated a modulation of attention and experience, with a significant interaction between attention and experience in the P2 processing stage. The result suggests that knowledge of a clean speech stimulus prior to its repetition in masked 'cocktail-party' form leads to reduced neural processing, as participants may anticipate with clarity the full auditory object, in particular, when to expect its temporal edges. These results are consistent with findings that precise temporal expectations about the occurrence and type of speech sounds lead to reduced P2 (and possibly even also N1) components (Pinto et al., 2019). In their study, the researchers addressed speech perception using syllables and audiovisual stimuli presentations. The conditions of their study included complementing the auditory stream with visual cues of the presentation timing and/or of lexical identity, both of which were effective to reduce the components' level of activation.

The present findings demonstrate that the neural mechanisms involved in selective attention may be subject to modulatory changes by repetition suppression. In the case of the P2 stage, commonly associated with semantic processing (Broderick et al., 2020), cortical activity still prioritized the parsing and segmentation of auditory events associated with the attended speaker suppressing; yet, prior exposure to its clean version led to activity reduction. In interpretations from predictive coding frameworks, coincidence with prior expectations produces a reduction in prediction error responsible for reduced neural activity (Broderick et al., 2019; Park et al., 2023). The decrease may also be due to differing parsing and segmentation processes normally cued by onsets of the speech envelope, for example, through diminished populations of neurons or narrower timeframes involved (Grill-Spector et al., 2006). The effective attenuation through gradual learning suggests an adaptive response that is based on task demands, introduced by the neural savings made when of re-processing, e.g., re-segmenting, the target stream.

6.3 Cortical processing under voice priming of a cocktail party speech stimulus

In Experiment 2a, participants underwent a pre-listening phase where they may be exposed to the speaker's voice. We had hypothesized that this brief experience may influence their perception of the same speaker presented within a competing masker, and the neural processing associated with segregating new sentences.

The performance data did not show a modulation by prior experience of speaker voice on the task. Previous research suggests the influence of voice familiarity on speech processing in scenarios with competing stimuli, when such experience has been gained over the long-term (Johnsrude et al., 2013). However, whether performance improvements by experience of voice stands under masking as it does in clean condition remains unclear (Daly and Pitt, 2023). Whether participants achieve greater accuracy under masker priming with long-term familiar voices may be worth further investigation in future studies, as familiarity with vocal musical distractors has been shown to assist enhance task comprehension (Brown and Bidelman, 2022) for instance.

In terms of neural encoding, attentional modulations were first observed at the N1 component, with the attended response eliciting a larger amplitude compared to the unattended response. As in the previous experiment, a similar modulation was found at the P2 component but, by contrast, this time we did not find any significant interaction effects with prior knowledge, such as activity reductions by target priming. It is important to note that, in this experiment's comprehension task, participants solved the cocktail party for an unknown message. In this case, the envelope onsets from the target speech (or the masker) yield a different temporal pattern as the sentence content has changed, despite the voice being the same.

6.4 Cortical processing under content priming of a cocktail party speech stimulus

In Experiment 2b, we examined the impact of prior knowledge about high-level lexical and sentential semantic information on neural segregation of speech. Participants first underwent a brief pre-listening phase where they could be exposed to a clean sentence that paraphrased one of the two speech stimuli, from two new, independent speakers, at the upcoming cocktail party mixture. We did not find that such prior exposure influenced their comprehension in the task however, as behavioural results revealed similar performance levels across conditions. As in Experiment 2a the target stimulus was not identically distributed over time as the primed subset of information about the target, which may matter for segmentation purposes. The null behavioral results of prior experience may suggest that partial foreknowledge of the message may not be sufficient to predict speech in the current behavioral task, which demands detailed speech information on short (e.g., ~10 s) timeframes.

In terms of neural processing of speech envelope onsets, the P1 responses showed a significant attentional effect, as in Experiment 1, suggesting early attentional segregation around 50 ms post-stimulus. Auditory responses at this staga may be dominated by excitatory activity triggered by the auditory scene's acoustic characteristics and physical features (Grunwald et al., 2003; Lewald and Getzmann, 2015), which have been linked to bottom-up mechanisms and potentially subject to interaction with top-down processes (Shinn-Cunningham, 2008). In the case of the N1 component, we again found an enhanced response for the attended stimulus compared to the unattended stream across priming conditions. The modulation was not influenced by prior semantic knowledge, which was a consistent null result across studies. Moreover, in the present study, as in Experiment 2a, we did not find any significant interaction by experience with selective attention at the level of the P2 component. In both cases, the temporal patterns of onsets in the envelope of individual speech streams did not coincide with the clean speech presented at pre-listening stage. We propose that he lack of repetition suggestion here suggests that this stage may be directly concerned with top-down processing leading to segmentation of an ‘unmasked’ representation of the stimulus generated post the N1 stage.

6.5 Auditory processing of clean and masked speech

Our results are generally consistent with observations that the auditory processing of clean and masked speech are processes that occur in stages. Our findings include evidence of cortical encoding of transient increases in the speech envelope that is organized in a triphasic profile, as in previous studies (Fiedler et al., 2019; Gansonre et al., 2018, Kraus et al., 2021; Wikman, et al., 2024, Peterson et al., 2017, Han, 2010). It begins with a small positive deflection that reaches its peak within 50 ms of the encoding response, corresponding to the P1 component that is also observed in the ERP literature. The P1 component has been associated with acoustic feature analysis of speech, involving Heschl's gyrus in primary auditory cortex (Brodbeck et al., 2020). Neural generators of the P1 may also include the hippocampus, the planum temporale posterior to the auditory cortex, and the lateral temporal cortex (Han, 2010). This stage is often related with bottom-up processing, and may be relevant for processing target features, such as timbre, intensity, or location (Nakagawa et al., 2014). After the P1, the N1 emerges as a negative peak around 125 ms. Evidence suggests that this component plays a central role in auditory segregation, since it is modulated by top-down mechanisms that influence auditory selectivity according to task demands (Hillyard 1973; Ding & Simon 2012; Zion Golumbic et al., 2013; Brodbeck et al., 2020; Fiedler et al., 2019). At a later stage of processing (approximately 200 ms latency), the P2 appears as a prominent positive peak (Alain and Tremblay, 2007, Wang et al., 2020, Verschueren et al., 2018). The P2 component could be considered in relation to the auditory encoding of the effectively tracked target, including access to its higher-order lexical and contextual information. Yet, the extent to which ignored stream information remains processed at this stage is unclear, as there are magnetoencephalography reports of residual representations in place (Brodbeck et al., 2020). Using EEG, Fiedler et al. (2019) for instance did not find neural tracking of the ignored speaker at the P2 stage.

Chapter 7

Conclusions

The cocktail party scenario presents a challenge for current studies of auditory processing. This thesis set out to survey the involvement of priming effects on neural responses to speech temporal cues that are modulated by selective attention in the cocktail party. The study found evidence of such interaction as a result of full prior knowledge of the target, which led to reduced P2 activation. While substantial, the reduction nevertheless did not override the attentional enhancement effect that is usually observed at this late stage of processing. The neural changes were in this case accompanied by behavioral benefits of increased accuracy in the comprehension task, underscoring an improved target detection under a now facilitated selective attention process. Critically, the reduction was exclusive to priming of the identical target sequence at the clean speech. The reduction, and generally, interaction, disappeared in cases of partial experience of target information, including prior knowledge of the speakers' voice or of the speech content. Because of this, we consider that P2 envelope onset coding may involve the operation of segmenting the neurally unmasked version of target speech, a process that would have only been repeated under full priming conditions here. As for the N1 component, our results systematically confirmed known attentional modulation effects, for a stage that is considered to reflect the formation and segregation of auditory objects. Contrary to a previous report (Wang et al., 2019), we did not observe evidence of priming at this stage. In additionally, new findings are reported for the P1 stage of processing (P1) of a weak selective attentional effect, which suggest the early presence of biased representations of a speech mixture with partial filtering of irrelevant information. Overall, the thesis contributes valuable insights into the implications of priming effects on neural mechanisms underlying selective attention in complex listening environments. This provides avenues for further research in auditory processing and attention.

Chapter 8

Limitations

Moving forward, incorporating experimental designs that encompass a holistic view of the complex phenomenon of learning in the cocktail party problem is crucial. This necessitates venturing beyond unimodal auditory interventions and incorporating natural contexts that integrate multimodal input, reflecting the real-world scenarios where listeners actively engage with speech. Our initial aim was to investigate auditory processing under at least two levels of priming, voice and content, yet our general approach was limited to the coding of the speech envelope onsets, a low-level auditory feature. Behavior was further assessed against anovel scoring system that may not have accounted relatively more for higher-level semantic processing such as words and topic than low-level auditory processing, such as gender and age recognition. Future studies may use different tasks and scoring systems to specifically evaluate outcomes of specific low-level auditory or high-level semantic processing.

Furthermore, as the study set out to address the interaction between attention and experience, the sample size differed across the three experiments which may have resulted ininconsistent observations of the weak attentional effect at the P1 stage. This effect was present across the larger collection of control data (N=72, NP trials), and was replicated in the Full priming (N=35, all conditions) and Content priming (N=21, all conditions) experiments. The early attentional modulation only appeared to approach marginal statistical significance in the Voice priming (N=16) experiment. For this, a power analysis is recommended in future research aiming to confirm this early effect of attentional selection to guarantee an adequate sample size in its detection.

Chapter 9

Scientific communications

The results or partial results of this work have been presented in regional and international conferences:

- Sánchez-Costa, T, Cervantes Constantino, F.; Carboni, A. (2023). Selective attention and the priming effect in the cocktail party problem: a psychophys- iological study. Presented at Jornada de Investigación de Ciencias Cognitivas y del Comportamiento, Uruguay, Montevideo.
- Sánchez-Costa, T, Cervantes Constantino, F.; Carboni, A. (2023). How do attention and experience interact in the cocktail party problem?. Presented at Sociedad de neurociencias del Uruguay, Montevideo.
- Sánchez-Costa, T, Cervantes Constantino, F.; Carboni, A. (2022). “Communication in humans: might familiarity affect how we attend to the environment?”. Presented at Humboldt Kolleg Montevideo.
- Sánchez-Costa, T, Cervantes Constantino, F.; Carboni, A. (2022). Cortical dynamics of speech processing in the cocktail party: Enhancement effects of masker priming. Presented at Falan congress, online.
- Cervantes Constantino, F.; Carboni, A.; Sánchez-Costa, T. (2022). Attention and expectation interacting modulations in the auditory cocktail party. Presented at International Conference in Cognitive Neuroscience, Helsinki, Finland
- Sánchez-Costa, T.; Carboni, A.; Cervantes Constantino, F. (2021). Enhancement dynamics of cortical representations by masker priming during attentional selection in the cocktail party problem. Presented at Advances and Perspectives in Auditory Neuroscience, online. Selected as a poster teaser.
- Sánchez-Costa, T.; Carboni, A.; Cervantes Constantino, F. (2021). The en-

hancement effects of background knowledge on selective auditory attention: an analysis of cortical responses in the cocktail party problem. Presented at Society for Neuroscience, online.

- Sánchez-Costa, T.; Carboni, A.; Cervantes Constantino, F. (2021). Dinámica de la corteza cerebral en el procesamiento del habla para el "efecto de fiesta de cóctel": interacción entre atención y experiencia. Presented at Jornada Interdisciplinaria de Ciencias Cognitivas, Montevideo.

Bibliography

Aiken, S. J. and Picton, T. W. (2008). Human Cortical Responses to the Speech Envelope. *Ear and Hearing*, 29(2):139–157.

Alain, C. and Tremblay, K. (2007). The Role of Event-Related Brain Potentials in Assessing Central Auditory Processing. *Journal of the American Academy of Audiology*, 18(07):573–589.

Bednar, A. and Lalor, E. C. (2020). Where is the cocktail party? Decoding loca- tions of attended and unattended moving sound sources using EEG. *NeuroImage*, 205:116283.

Bhandari, P., Demberg, V., and Kray, J. (2021). Semantic Predictability Facilitates Comprehension of Degraded Speech in a Graded Manner. *Frontiers in Psychol- ogy*, 12.

Broadbent, D. E. (1957). A mechanical model for human attention and immediate memory. *Psychological Review*, 64(3):205–215.

Brodbeck, C., Jiao, A., Hong, L. E., and Simon, J. Z. (2020). Neural speech restora- tion at the cocktail party: Auditory cortex recovers masked speech of both at- tended and ignored speakers. *PLOS Biology*, 18(10):e3000883.

Broderick, M. P., Anderson, A. J., and Lalor, E. C. (2019). Semantic Context En- hances the Early Auditory Encoding of Natural Speech. *The Journal of Neuro- science*, 39(38):7564–7575.

Broderick, M. P., Zuk, N. J., Anderson, A. J., & Lalor, E. C. (2020). More than words: Neurophysiological correlates of semantic dissimilarity depend on comprehension of the speech narrative. *In bioRxiv*. https://doi.org/10.1101/2020.12.14.422789

Bronkhorst, A. (2000). The Cocktail Party Phenomenon: A Review of Research on Speech Intelligibility in Multiple-Talker Conditions. *Acta Acustica united with Acustica*, 86:117–128.

Brown, J. A. and Bidelman, G. M. (2022). Familiarity of background music modulates the cortical tracking of target speech at the "cocktail party". *Brain Sciences*, 12(10):1320.

Brungart, D. S. (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. *The Journal of the Acoustical Society of America*, 109(3):1101–1109.

Cervantes Constantino, F. and Simon, J. Z. (2018). Restoration and Efficiency of the Neural Processing of Continuous Speech Are Promoted by Prior Knowledge. *Frontiers in Systems Neuroscience*, 12.

Charest, I., Pernet, C. R., Rousselet, G. A., Quiñones, I., Latinus, M., Fillion-Bilodeau, S., Chartrand, J.-P., & Belin, P. (2009). Electrophysiological evidence for an early processing of human voices. *BMC Neuroscience*, 10(1), 127. https://doi.org/10.1186/1471-2202-10-127

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *The Journal of the acoustical society of America*, 25(5):975–979.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(3):181–204.

Crosse, M. J., Di Liberto, G. M., Bednar, A., and Lalor, E. C. (2016). The multi- variate temporal response function (mtrf) toolbox: a matlab toolbox for relating neural signals to continuous stimuli. *Frontiers in human neuroscience*, 10:604.

Daly, H. R. and Pitt, M. A. (2023). Target voice probability influences enhancement in auditory selective attention. *Attention, Perception, & Psychophysics*.

David, S. V., Mesgarani, N., and Shamma, S. A. (2007). Estimating sparse spectro- temporal receptive fields with natural stimuli. *Network: Computation in Neural Systems*, 18(3):191–212.

de Cheveigné, A. and Parra, L. C. (2014). Joint decorrelation, a versatile tool for multichannel data analysis. *Neuroimage*, 98:487–505.

de Cheveigné, A. and Simon, J. Z. (2008). Denoising based on spatial filtering. *Journal of Neuroscience Methods*, 171(2):331–339.

Dekerle, M., Boulenger, V., Hoen, M., and Meunier, F. (2014). Multi-talker back- ground and semantic priming effect. *Frontiers in human neuroscience*, 8:878.

Di Liberto, G. M., O'Sullivan, J. A., and Lalor, E. C. (2015). Low-Frequency Cortical Entrainment to Speech Reflects Phoneme-Level Processing. *Current Biology*, 25(19):2457–2465.

Ding, N. and Simon, J. Z. (2012). Emergence of neural encoding of auditory objects while listening to competing speakers. *Proceedings of the National Academy of Sciences*, 109(29):11854–11859. Publisher: National Academy of Sciences Section: Biological Sciences.

Domingo, Y., Holmes, E., and Johnsrude, I. S. (2020). The benefit to speech in- telligibility of hearing a familiar voice. *Journal of Experimental Psychology: Applied*, 26(2):236–247. Place: US Publisher: American Psychological Association.

Driver, J. (2001). A selective review of selective attention research from the past century. *British Journal of Psychology*, 92(1):53–78.

Evans, S., McGettigan, C., Agnew, Z. K., Rosen, S., and Scott, S. K. (2016). Get- ting the cocktail party started: masking effects in speech perception. *Journal of cognitive neuroscience*, 28(3):483–500.

Fiedler, L., Wöstmann, M., Herbst, S. K., and Obleser, J. (2019). Late cortical tracking of ignored speech facilitates neural selectivity in acoustically challenging conditions. *NeuroImage*, 186:33–42.

Field, A. (2017). GLM: Repeated Measures ANOVA. In *Discovering statistics using IBM SPSS Statistics: And sex and drugs and rock ’n’ roll*. Sage, London, 4th edition.

Gansonre, C., Højlund, A., Leminen, A., Bailey, C., and Shtyrov, Y. (2018). Task- free auditory EEG paradigm for probing multiple levels of speech process- ing in the brain. *Psychophysiology*, 55(11).

Gazzaley, A. and Nobre, A. C. (2012). Top-down modulation: Bridging selective attention and working memory. *Trends in cognitive sciences*, 16(2):129–135.

Getzmann, S., Golob, E. J., and Wascher, E. (2016). Focused and divided attention in a simulated cocktail-party situation: ERP evidence from younger and older adults. *Neurobiology of Aging*, 41:138–149.

Giard, M. H., Fort, A., Mouchetant-Rostaing, Y., & Pernier, J. (2000). Neurophysiological mechanisms of auditory selective attention in humans. *Frontiers in bioscience*, *5*(1), 84-94.

Giraud, A.-L. and Poeppel, D. (2012). Cortical oscillations and speech process- ing: emerging computational principles and operations. *Nature neuroscience*, 15(4):511–517.

Golubic, S. J., Jurasic, M. J., Susac, A., Huonker, R., Gotz, T., and Haueisen, J. (2019). Attention modulates topology and dynamics of auditory sensory gating. *Human Brain Mapping*, 40(10):2981–2994.

Grunwald, T., Boutros, N. N., Pezer, N., von Oertzen, J., Fernández, G., Schaller, C., and Elger, C. E. (2003). Neuronal substrates of sensory gating within the human brain. *Biological psychiatry*, 53(6):511–519.

Han, W. (2010). P1-n1-p2 complex and acoustic change complex elicited by speech sounds: current research and applications. *Audiology*, 6(2):121–127.

Har-shai Yahav, P. and Zion Golumbic, E. (2021). Linguistic processing of task- irrelevant speech at a cocktail party. *eLife*, 10:e65096.

Haykin, S. and Chen, Z. (2005). The Cocktail Party Problem. *Neural Computation*, 17(9):1875–1902.

Hertrich, I., Dietrich, S., Trouvain, J., Moos, A., and Ackermann, H. (2012). Mag- netic brain activity phase-locked to the envelope, the syllable onsets, and the fun- damental frequency of a perceived speech signal. *Psychophysiology*, 49(3):322– 334.

Hillyard, S. A., Hink, R. F., Schwent, V. L., and Picton, T. W. (1973). Electrical Signs of Selective Attention in the Human Brain. *Science*, 182(4108):177–180. Publisher: American Association for the Advancement of Science.

Hjortkjær, J., Märcher-Rørsted, J., Fuglsang, S. A., and Dau, T. (2020). Cor- tical oscillations and entrainment in speech processing during working memory load. *European Journal of Neuroscience*, 51(5):1279–1289. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/ejn.13855.

Hyvarinen, A. (1999). Fast and robust fixed-point algorithms for independent component analysis. *IEEE Transactions on Neural Networks*, 10(3):626–634.

Johnsrude, I. S., Mackey, A., Hakyemez, H., Alexander, E., Trang, H. P., and Carlyon, R. P. (2013). Swinging at a cocktail party: Voice familiarity aids speech perception in the presence of a competing voice. Psychological science, 24(10):1995–2004.

Jones, L. A., Hills, P. J., Dick, K. M., Jones, S. P., and Bright, P. (2016). Cognitive mechanisms associated with auditory sensory gating. Brain and Cognition, 102:33–45.

Junghöfer, M., Elbert, T., Tucker, D. M., & Rockstroh, B. (2000). Statistical control of artifacts in dense array EEG/MEG studies. Psychophysiology, 37(4), 523–532. https://doi.org/10.1111/1469-8986.3740523

Kachlicka, M., Laffere, A., Dick, F., and Tierney, A. (2022). Slow phase-locked modulations support selective attention to sound. NeuroImage, 252:119024.

Kaya, E. M. and Elhilali, M. (2017). Modelling auditory attention. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1714):20160101. Publisher: Royal Society.

Kidd, Jr., G., Best, V., and Mason, C. R. (2008). Listening to every other word: Examining the strength of linkage variables in forming streams of speech. *The Journal of the Acoustical Society of America*, 124(6):3793–3802.

Kidd, Jr., G., Mason, C. R., Swaminathan, J., Roverud, E., Clayton, K. K., and Best, V. (2016). Determining the energetic and informational components of speech-on-speech masking. *The Journal of the Acoustical Society of America*, 140(1):132–144.

Kraus, F., Tune, S., Ruhe, A., Obleser, J., & Wöstmann, M. (2021). Unilateral acoustic degradation delays attentional separation of competing speech. *Trends in Hearing*, *25*, 233121652110132. https://doi.org/10.1177/23312165211013242

Leonard, M. K., Baud, M. O., Sjerps, M. J., & Chang, E. F. (2016). Perceptual restoration of masked speech in human cortex. *Nature communications*, *7*(1), 13619.

Lewald, J. and Getzmann, S. (2015). Electrophysiological correlates of cock-tail- party listening. *Behavioural Brain Research*, 292:157–166.

Luck, S. J., Heinze, H. J., Mangun, G. R., & Hillyard, S. A. (1990). Visual event-related potentials index focused attention within bilateral stimulus arrays. II. Functional dissociation of P1 and N1 components. *Electroencephalography and clinical neurophysiology*, *75*(6), 528-542.

Luo, H. and Poeppel, D. (2007). Phase Patterns of Neuronal Responses Reliably Discriminate Speech in Human Auditory Cortex. *Neuron*, 54(6):1001–1010.

Maris, E. (2011). Statistical testing in electrophysiological studies. *Psychophysiology*, 49:549–65.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. Journal of neuroscience methods, 164(1), 177-190.

Martin, B. A., Tremblay, K. L., and Korczak, P. (2008). Speech Evoked Potentials: From the Laboratory to the Clinic. *Ear & Hearing*, 29(3):285–313.

Middlebrooks, J. C. and Simon, J. Z. (2017). Ear and Brain Mechanisms for Parsing the Auditory Scene. In Middlebrooks, J. C., Simon, J. Z., Popper, A. N., and Fay, R. R., editors, *The Auditory System at the Cocktail Party*, volume 60, pages 1–6. Springer International Publishing, Cham. Series Title: Springer Handbook of Auditory Research.

Moore, B. C. J. and Gockel, H. E. (2012). Properties of auditory stream forma- tion. *Philosophical Transactions: Biological Sciences*, 367(1591):919–931. Pub- lisher: Royal Society.

Newman, R. S. and Evers, S. (2007). The effect of talker familiarity on stream segregation. *Journal of Phonetics*, 35(1):85–103.

Nakagawa, K., Otsuru, N., Inui, K., & Kakigi, R. (2014). Change-related auditory P50: A MEG study. NeuroImage, 86, 131-137.

Obleser, J. and Kayser, C. (2019). Neural Entrainment and Attentional Selection in the Listening Brain. *Trends in Cognitive Sciences*, 23(11):913–926. Publisher: Elsevier.

O'Sullivan, J. A., Power, A. J., Mesgarani, N., Rajaram, S., Foxe, J. J., Shinn-Cunningham, B. G., Slaney, M., Shamma, S. A., and Lalor, E. C. (2015). Attentional Selection in a Cocktail Party Environment Can Be Decoded from Single-Trial EEG. *Cerebral Cortex*, 25(7):1697–1706.

Ozmeral, E. J., Eddins, D. A., and Eddins, A. C. (2021). Selective auditory attention modulates cortical responses to sound location change in younger and older adults. *Journal of Neurophysiology*, 126(3):803–815. Publisher: American Physiological Society.

Park, J. J., Baek, S.-C., Suh, M.-W., Choi, J., Kim, S. J., and Lim, Y. (2023). The effect of topic familiarity and volatility of auditory scene on selective auditory attention. *Hearing Research*, 433:108770.

Petersen, E. B., Wöstmann, M., Obleser, J., & Lunner, T. (2017). Neural tracking of attended versus ignored speech is differentially affected by hearing loss. Journal of neurophysiology, 117(1), 18-27.

Pinto, S., Tremblay, P., Basirat, A., and Sato, M. (2019). The impact of when, what and how predictions on auditory speech perception. *Experimental Brain Research*, 237(12):3143–3153.

Poeppel, D. and Assaneo, M. F. (2020). Speech rhythms and their neural foun- dations. *Nature Reviews Neuroscience*, 21(6):322–334. Number: 6 Publisher: Nature Publishing Group.

Power, A. J., Foxe, J. J., Forde, E.-J., Reilly, R. B., and Lalor, E. C. (2012). At what time is the cocktail party? A late locus of selective attention to natural speech: A late locus of attention to natural speech. *European Journal of Neuroscience*, 35(9):1497–1503.

Quirk, C., C. Brockett, and W. B. Dolan. (2004). Monolingual Machine Translation for Paraphrase Generation. *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing*, Barcelona Spain. Association for Computational Linguistics.

Rabinowitz, N. C., Willmore, B. D. B., King, A. J., and Schnupp, J. W. H. (2013). Constructing Noise-Invariant Representations of Sound in the Auditory Pathway. *PLOS Biology*, 11(11):e1001710. Publisher: Public Library of Science.

Scanlon, J. E. M., Cormier, D. L., Townsend, K. A., Kuziek, J.

W. P., and Mathewson, K. E. (2019a). The ecological cocktail party: Measuring brain activity during an auditory oddball task with background noise. *Psychophysiology*, 56(11):e13435. _eprint: https://onlinelibrary.wiley.-

com/doi/pdf/10.1111/psyp.13435.

Scanlon, J. E. M., Townsend, K. A., Cormier, D. L., Kuziek, J. W. P., and Mathew- son, K. E. (2019b). Taking off the training wheels: Measuring auditory P3 during outdoor cycling using an active wet EEG system. *Brain Research*, 1716:50–61.

Shinn-Cunningham, B., Best, V., and Lee, A. K. C. (2017). Auditory Object Forma- tion and Selection. In Middlebrooks, J. C., Simon, J. Z., Popper, A. N., and Fay, R. R., editors, *The Auditory System at the Cocktail Party*, Springer Handbook of Auditory Research, pages 7–40. Springer International Publishing, Cham.

Shinn-Cunningham, B. G. (2008). Object-based auditory and visual attention. *Trends in Cognitive Sciences*, 12(5):182–186.

Steinschneider, M., Liégeois-Chauvel, C., and Brugge, J. (2011). Auditory Evoked Potentials and Their Utility in the Assessment of Complex Sound Processing. *The Auditory Cortex*, pages 535–559.

Summerfield, C., Trittschuh, E. H., Monti, J. M., Mesulam, M.-M., & Egner, T. (2008). Neural repetition suppression reflects fulfilled perceptual expectations. Nature Neuroscience, 11(9), 1004–1006. https://doi.org/10.1038/nn.2163

Ten Oever, S., & Martin, A. E. (2024). Interdependence of “what” and “when” in the brain. *Journal of Cognitive Neuroscience*, *36*(1), 167–186. https://doi.org/10.1162/jocn_a_02067

Treisman, A. M. (1960). Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology*, 12(4):242–248. Publisher: Routledge _eprint: https://doi.org/10.1080/17470216008416732.

Vanthornhout, J., Decruy, L., Wouters, J., Simon, J. Z., and Francart, T. (2018). Speech intelligibility predicted from neural entrainment of the speech envelope. *Journal of the Association for Research in Otolaryngology*, 19:181–191.

Verschueren, E., Vanthornhout, J., & Francart, T. (2018). The effect of stimulus choice on an EEG-based objective measure of speech intelligibility. In *bioRxiv*.

https://doi.org/10.1101/421727.

Wang, J., Chen, J., Yang, X., Liu, L., Wu, C., Lu, L., Li, L., and Wu, Y. (2021). Common brain substrates underlying auditory speech priming and perceived spa- tial separation. *Frontiers in Neuroscience*, 15:664985.

Wang, L., Wu, E. X., & Chen, F. (2020). Robust EEG-based decoding of auditory attention with high-RMS-level speech segments in noisy conditions. *Frontiers in Human Neuroscience*, 14. https://doi.org/10.3389/fnhum.2020.557534.

Wang, Y., Zhang, J., Zou, J., Luo, H., and Ding, N. (2019). Prior Knowledge Guides Speech Segregation in Human Auditory Cortex. Cerebral Cortex, 29(4):1561– 1571.

Wikman, P., Salmela, V., Sjöblom, E., Leminen, M., Laine, M., & Alho, K. (2024). Attention to audiovisual speech shapes neural processing through feedback-feedforward loops between different nodes of the speech network. PLoS Biology, 22(3), e3002534.

Wood, N. and Cowan, N. (1995). The cocktail party phenomenon revisited: How frequent are attention shifts to one's name in an irrelevant auditory channel? *Jour- nal of Experimental Psychology: Learning, Memory, and Cognition*, 21:255–260. Place: US Publisher: American Psychological Association.

World Medical Association (WMA) (2009). Declaration of Helsinki. Ethical principles for medical research involving human subjects. Jahrbuch Fur Wissenschaft Und Ethik, 14(1), 233–238. https://doi.org/10.1515/9783110208856.233

Wöstmann, M., Alavash, M., and Obleser, J. (2019). Alpha Oscillations in the Human Brain Implement Distractor Suppression Independent of Target Selection. *Journal of Neuroscience*, 39(49):9797–9805. Publisher: Society for Neuro- science Section: Research Articles.

Wöstmann, M., Störmer, V. S., Obleser, J., Addleman, D. A., Andersen, S. K., Gaspelin, N., Geng, J. J., Luck, S. J., Noonan, M. P., Slagter, H. A., and Theeuwes, J. (2022). Ten simple rules to study distractor suppression. *Progress* *in* *Neurobiology*, 213:102269.

Zion Golumbic, E., Ding, N., Bickel, S., Lakatos, P., Schevon, C., McKhann, G., Goodman, R., Emerson, R., Mehta, A., Simon, J., Poeppel, D., and Schroeder,

C. (2013). Mechanisms Underlying Selective Neuronal Tracking of Attended Speech at a “Cocktail Party”. *Neuron*, 77(5):980–991.

ANNEXES

Annex 1

Information sheet

Figure 9. Information sheet presented to participants

Hoja de información - PCCP

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

Facultad de Psicología
UNIVERSIDAD DE LA REPÚBLICA

Hoja de información para participantes

Se le entregará una copia de esta hoja de información para que la conserve como referencia.

Título del proyecto: Codificación predictiva en el efecto de fiesta de coctel

Este estudio cuenta con aval para ejecución por el Comité de Ética en Investigación de la Facultad de Psicología de la UdelaR.

Nombre: Dr. Francisco Cervantes

Datos de contacto: Centro de Investigación Básica en Psicología, Tristán Narvaja 1674, Montevideo 11.200.
Correo electrónico: fcervantes@psico.edu.uy

¿A quién estamos reclutando?
Nos gustaría invitar a personas de 18 a 35 años de edad, para participar de forma voluntaria en una investigación psicológica. Estamos buscando personas que completen una tarea conductual y un registro de datos fisiológicos mediante electroencefalografía (EEG).

Detalles del estudio

Nos gustaría invitarle a participar en este proyecto de investigación. Solo debe participar si quiere hacerlo; elegir no participar no le perjudicará de ninguna manera. Si elige participar, tiene la oportunidad de retirarse en cualquier momento, sin dar ninguna razón.

¿Por qué estamos haciendo este estudio?
Deseamos entender cómo algunas personas orientan su atención a hablantes del entorno.

Para probar esto, usaremos algunas tareas sencillas por computadora y cuestionarios. Se le pedirá que escuche ciertos elementos hablados en donde atenderá selectivamente a algunos mensajes e ignorará otros simultáneamente.

¿Qué pasará durante la prueba?

Necesitará venir al Centro de Investigación Básica en Psicología de la UdelaR, Dr. Tristán Narvaja 1674, Montevideo. Luego de verificar si es elegible para participar, y que no hay factores que puedan impedirle participar, y de leer y refrendar el formulario de consentimiento, llenará cuestionarios que miden rasgos psicológicos relevantes a la atención.

Le colocaremos un gorro de EEG (electroencefalografía), una técnica no invasiva e indolora. El procedimiento implica frotar suavemente el cuero cabelludo para quitar el cabello de los electrodos. Se le dejará un gel conductor en el cuero cabelludo, que podrá lavar en la sala de pruebas al final.

No debe participar si no se siente cómodo o cómoda con el gorro ajustado y el gel aplicado en el cuero cabelludo. El gorro del EEG podría dejar algunas marcas rojas leves en la frente donde los electrodos presionan contra la piel. Estos no son dolorosos y se desvanecerán a los pocos minutos de quitar la gorra.

Cuando la gorra se coloque, completará las tareas de computadora antes indicadas. La sesión experimental (menos preparación) durará aproximadamente 1 hora.

¿Cómo se almacenarán mis datos?
Toda la información que se recopila sobre usted durante el curso de la investigación se mantendrá estrictamente confidencial y se almacenará de forma segura, de manera electrónica o en repositorios bloqueados, con su nombre reemplazado por un código numerado para que no pueda ser identificado. (Ref: Ley N° 18.331, 11-08-2008).

Nota: Si tiene más preguntas sobre este estudio, no dude en ponerse en contacto con los investigadores enviando un correo electrónico a fcervantes@psico.edu.uy

Si decide participar, todavía puede retirarse en cualquier momento y sin dar una razón. Comparta esta información anterior con otros si desea o pregúntenos si hay algo que no esté claro o si desea obtener más información.

Por favor tenga en cuenta que **NO PUEDE** participar si:
1. Ha participado en este estudio con anterioridad.
2. Sufre de cualquier problema de salud física o mental importante.
3. Presenta visión y/o audición no corregidas.
4. Ha consumido ciertas sustancias psicoactivas recientemente, por lo menos 24 horas antes de la sesión de prueba.
5. Tiene un peinado permanente como rastas, trenzas, etc., lo que significa que no podremos colocarle un gorro de EEG.

Este estudio cuenta con el aval para ejecución por el Comité de Ética en la Investigación de la Facultad de Psicología de la Universidad de la República (ID de proyecto)

Annex 2

Informed consent

Figure 10. Informed consent

Facultad de Psicología
UNIVERSIDAD DE LA REPÚBLICA

CIBPsi
CENTRO DE INVESTIGACIÓN BÁSICA EN PSICOLOGÍA

cibpsi@psico.edu.uy
www.cibpsi.psico.edu.uy
Tel.: (598 2) 400 85 55 / Int. 285-286

Hoja de consentimiento informado

Nombre del proyecto: **Codificación predictiva en el efecto de fiesta de coctel.**

Yo, (nombre y apellidos):

DECLARO que:

- He leído la hoja de información que se me ha entregado con respecto al presente proyecto (resumen anexo)
- He podido hacer preguntas sobre el estudio y me las han respondido satisfactoriamente.
- He recibido suficiente información sobre el estudio, conozco sus objetivos, el tiempo que durará mi participación en él y que tipo de actividad que deberé realizar.
- Reconozco que la información que yo provea en el curso de esta investigación es estrictamente confidencial y no será usada para ningún otro propósito fuera de los de este estudio sin mi consentimiento.
- Comprendo que mi participación es voluntaria y libre y que puedo retirarme del estudio:
  · Cuando quiera.
  · Sin tener que dar explicaciones.
  · Sin que esto represente perjuicio alguno sobre mi persona
- Entiendo que una copia de esta ficha de consentimiento me será entregada, y que puedo pedir información sobre los resultados de esta investigación cuando ésta haya concluido.

Por lo anteriormente dicho, presto libremente mi conformidad para participar en el estudio
Montevideo,
Firma del participante
Aclaración del participante: ___________________

Firma del investigador
Aclaración del investigador: ___________________

Por cualquier consulta puede dirigirse a:
Dr. Francisco Cervantes Constantino
Centro de Investigación Básica en Psicología
Facultad de Psicología, Universidad de la República
Tristán Narvaja 1674, Montevideo
fcervantes@psico.edu.uy

Tristán Narvaja 1674 CP 11200 Tel.: (598) 2 400 85 55 Fax: (598) 2 400 86 40
Montevideo, Uruguay
www.psico.edu.uy

Facultad de Psicología
UNIVERSIDAD DE LA REPÚBLICA
CIBPsi
CENTRO DE INVESTIGACIÓN BÁSICA EN PSICOLOGÍA
cibpsi@psico.edu.uy
www.cibpsi.psico.edu.uy
Tel.: (598 2) 400 85 55 / Int. 285-286

Resumen de la Hoja de información

Nos gustaría invitarle a participar en este proyecto de investigación. Solo debe participar si quiere hacerlo; elegir no participar no le perjudicará de ninguna manera. Si elige participar, tiene la oportunidad de retirarse en cualquier momento, sin dar ninguna razón.

Si usted elige participar, usaremos algunas tareas sencillas por computadora y cuestionarios. Se le pedirá que escuche ciertos elementos hablados en donde atenderá selectivamente a algunos mensajes e ignorará otros simultáneamente.

Le colocaremos un gorro de EEG (electroencefalografía), una técnica no invasiva e indolora. Se le dejará un gel que podrá lavar en la sala de pruebas al final.

No debe participar si no se siente cómodo o cómoda. El gorro del EEG podría dejar algunas marcas rojas leves en la frente donde se presiona contra la piel. Estos no son dolorosos y se desvanecerán a los pocos minutos de quitar la gorra.

Cuando la gorra se coloque, completará las tareas de computadora antes indicadas. La sesión experimental (menos preparación) durará aproximadamente 1 hora.

Toda la información que se recopila sobre usted durante el curso de la investigación se mantendrá estrictamente confidencial y se almacenará de forma segura.

<*En caso de ser aprobado:* Este estudio cuenta con el aval para ejecución por el Comité de Ética en la Investigación de la Facultad de Psicología de la Universidad de la República (ID de proyecto)>

Tristán Narvaja 1674 CP 11200 Tel.: (598) 2 400 85 55 Fax: (598) 2 400 86 40
Montevideo, Uruguay
www.psico.edu.uy

Hoja de consentimiento informado

Nombre del proyecto: **Codificación predictiva en el efecto de fiesta de coctel.**

Yo, (nombre y apellidos):

DECLARO que:

- He leído la hoja de información que se me ha entregado con respecto al presente proyecto (resumen anexo)
- He podido hacer preguntas sobre el estudio y me las han respondido satisfactoriamente.
- He recibido suficiente información sobre el estudio, conozco sus objetivos, el tiempo que durará mi participación en él y que tipo de actividad que deberé realizar.
- Reconozco que la información que yo provea en el curso de esta investigación es estrictamente confidencial y no será usada para ningún otro propósito fuera de los de este estudio sin mi consentimiento.
- Comprendo que mi participación es voluntaria y libre y que puedo retirarme del estudio:

· Cuando quiera.
· Sin tener que dar explicaciones.
· Sin que esto represente perjuicio alguno sobre mi persona

- Entiendo que una copia de esta ficha de consentimiento me será entregada, y que puedo pedir información sobre los resultados de esta investigación cuando ésta haya concluido.

Por lo anteriormente dicho, presto libremente mi conformidad para participar en el estudio
Montevideo,
Firma del participante
Aclaración del participante: ___________________

Firma del investigador
Aclaración del investigador: ___________________

Por cualquier consulta puede dirigirse a:
Dr. Francisco Cervantes Constantino
Centro de Investigación Básica en Psicología
Facultad de Psicología, Universidad de la República
Tristán Narvaja 1674, Montevideo
fcervantes@psico.edu.uy

Annex 3

Edinburgh Handedness Inventory

Figure 11. Figure A.3: Edinburgh Handedness Inventory

CUESTIONARIO DE DOMINANCIA MANUAL DE EDINBURGH

¿QUÉ MANO UTILIZA PARA?	DERECHA	IZQUIERDA	PUNTOS
escribir			1-2-3-4-5
dibujar			1-2-3-4-5
lanzar un objeto			1-2-3-4-5
lavarse los dientes			1-2-3-4-5
utilizar un cuchillo (sin tenedor)			1-2-3-4-5
cortar con tijeras			1-2-3-4-5
comer con la cuchara			1-2-3-4-5
barrer (la mano que coloca arriba)			1-2-3-4-5
rascar una cerilla			1-2-3-4-5
levantar la tapa de una caja			1-2-3-4-5

neuropediatra.org

Instrucciones: marque la casilla de la mano que prefiere con

+ una cruz, si es la mano que utiliza de modo preferente
++ dos cruces, si es la mano que utiliza de modo muy preferente y además le resultaría imposible o muy difícil hacerlo con la otra mano
+ una cruz, en las dos casillas cuando pueda hacerlo tan bien con una mano como con la otra

Puntos:

5 si ++ solo en la mano izquierda
4 si + solo en la mano izquierda
3 si + en manos izquierda y derecha
2 si + solo en la mano derecha
1 si ++ solo en la mano derecha

Resultado:

El resultado máximo son 50 puntos, el mínimo 10.
Cuanto más cerca a 50 puntos más zurdo, cuanto más cerca a 10 más diestro.

Nota: los resultados de este cuestionario no tienen valor clínico si no se contrastan con su médico. Si por razones médicas necesita examinar su preferencia manual debe consultar con su médico.

Referencia: EDINBURGH HANDEDNESS INVENTORY (Oldfield, 1971; Bryden, 1977).

MJ Mas - Neuropediatra