



# From psychophysiology to brain imaging: forty-five years MMN history of investigating acoustic change sensitivity

Valéria Csépe<sup>1,2</sup> · Ferenc Honbolygó<sup>1,3</sup>

Received: 4 June 2023 / Accepted: 18 March 2024 / Published online: 12 April 2024  
© The Author(s) 2024

## Abstract

Forty-five years have passed since the first publication of the mismatch negativity (MMN) event-related brain potential (ERP) component. The first 10 years of research hardly gained any particular attention of the scientific community interested in acoustic perception. Debates on the nature of sensation versus perception were going on, and the technical possibilities to record ERPs, called in general evoked potentials, were very limited. Subtle changes in pure tone frequency or intensity giving rise to the MMN component were first investigated in humans. The background of the theoretical model developed by Risto Näätänen was the orientation reaction model of E.N. Sokolov published in 1963 so that the MMN was seen first as an electrophysiological correlate of auditory change detection. This fundamental ability of the auditory system seen as crucial for survival led to the development of the first animal model of the MMN (Csépe et al. in *Clin Neurophysiol* 66: 571–578, 1987). Indeed, it was confirmed that the MMN was the brain correlate of subtle changes detected that might alert to potential threats in the environment and direct the behavioral orientation. The investigations performed after 2000 introduced complex models and more sophisticated methods, both in animal and human studies, so that the MMN method was on the way to become a tool on the first place and not the main goal of research. This approach was further strengthened by the increasing number of studies on different clinical populations aiming at future applications. The aim of our review is to describe and redefine what the MMN may reflect in auditory perception and to show why and how this brain correlate of changes in the auditory scene can be used as a valuable tool in cognitive neuroscience research. We refer to publications selected to underly the argument the MMN cannot be classified anymore as a sign of simple change detection and not all the indicators used to confirm how genuine the MMN elicited by variations of tones are valid for those to speech contrasts. We provide a fresh view on the broadly used MMN models, provided by some influential publications as well as on the unwritten history of MMN research aiming to give revised picture on what the MMN may truly reflect. We show how the focus and terminology of the MMN research have changed and what kind of misunderstandings and seemingly contradictory results prevent the MMN community to accept a generally usable cognitive model.

**Keywords** Mismatch negativity · Auditory events · Context dependence · Tones versus speech · Repetition suppression · Stimulus-specific adaptation · Predictive coding · Word stress processing · Development

## Introduction

Psychophysiology. A multidisciplinary field of research that has always provided different approaches, fresh views, and new methods biological in nature for those psychologists who were interested in perception, learning, memory, attention, and motivation beyond behavior. In Hungary, psychophysiology was not just a simple choice of interest in the 1960s and 1970s as physiological laboratories, departments, institutions served as scientific shelters for several researchers in psychology. These laboratories were the places of incubation for new ideas and gave chances

---

✉ Valéria Csépe  
csepe.valeria@ttk.hu

<sup>1</sup> Brain Imaging Centre, HUN-REN Research Centre of Natural Sciences, Magyar Tudósok Krt. 2, Budapest 1117, Hungary

<sup>2</sup> University of Pannonia, Veszprém, Hungary

<sup>3</sup> Eötvös Loránd University, Budapest, Hungary

to the reemergence of psychology surviving the nearly two decades long political suppression between 1945 and 1963. The book on Psychophysiology (1972) edited by György Ádám was a breakthrough, a collection of translated publications on learning, memory, motivation, and consciousness of influential authors, from Pavlov through Skinner, Pribram and Lurija to Freud. The thought-provoking introduction was written by the editor, whose introduction on learning, motivation, and consciousness, ‘instead of preface’ as mentioned, full of brilliant ideas on the multi-level construct of psychology and the role of physiology in getting closer to the real nature of mental operations. Ádám’s main question was “how the polymorphic bricks of physiological psychology are built into the high-rise buildings of psychological sciences” (1972, p 16). The emphasis on the role of physiology in understanding psychology was rather new that time and is still prevalent in the twenty-first century’s third decade.

“In the sense of what has been said, it is not in dispute that they (*authors’ comment: the polymorphic bricks*) participate in creating the foundation. The more solid this physiological base becomes, the more stable the fundament of the discovered, though still fragile psychological functions will be. Both higher importance of the insemination and increase of the essential and foundational role of this frontier research will contribute to gain recognition by psychologists of different branches. (Ádám, 1972, p 16.)”

The increasing use of physiological methods in psychology, especially those used to study the brain correlates of different psychological processes provided new insight and contributed to specialization of the field including a rising multidisciplinary area of research called cognitive neuroscience in the last three decades. Research on auditory perception started in psychophysiology and the field is classified by the time of our publication as cognitive neuroscience. Naming is the consequence of the attitude change not foreseen by Ádám (1972, p5): “Brain physiologists and experimental psychologists in psychophysiology research have never claimed that the research field they cultivate should be recognized as an independent discipline, physiological psychology lacks ‘professional self-awareness’ in the good sense.”

Nowadays cognitive neuroscientists show more self-awareness than ever before, probably due to the vast repertoire of methods they use and the increasing multidisciplinary of the field required by the complexity of the topics investigated. This transformation is well seen in the development of the methodological repertoire and the growing number of influential theories on auditory perception. It was psychophysiology when the first paper (Näätänen et al. 1978) on the event-related brain potential (ERP) correlate of automatic detection of acoustic changes called mismatch negativity (MMN) was published. The first 10 years of

research hardly gained any particular attention of the scientific community interested in acoustic perception, debates on the nature of sensation versus perception were going on and the technical possibilities to record ERPs, called in general as evoked potentials, were very limited. Moreover, the first MMN model followed the classical psychophysiological approach and was based on an existing physiological theory of orientation reaction. How the MMN theories moved from detection to coding theories? What kind of changes might contribute to the recent view on the psychological importance and assumed physiological mechanisms? Our paper is a subjective summary of those important steps psychophysiology helped us to make for understanding the processing of subtle deviations in rules, roles and regularities occurring in the acoustic environment.

### The first decades of MMN research—from orientation to memory

The auditory MMN appears in electrophysiological recordings as a small negative deflection in response to sounds deviating from established repetition or consistency in the recent past. In a laboratory setting, MMN is typically studied in passive oddball paradigms and observed as responses to low probability “deviant” sounds irregularly interspersing the highly repetitive “standards” they differ from in some dimension. The easiest way to observe the MMN is to produce a difference waveform via subtracting the response to the standard from that elicited by the deviant. The early studies interpreted the MMN as a difference-detection response representing an automatic stimulus discrimination. This interpretation formulated as a model has since been referred to as deviance detection (DD) theory, considered in our today’s terminology as neurophysiological prediction error signal elicited by an error in perceptual inference. This, however, is just rewording the initially postulated mechanisms of auditory object abstraction of physical properties of discrete sounds represented together. According to the first animal model of the MMN (Csépe, Karmos, Molnár 1987, 1989), this processing is presumably occurring in the ascending pathway as the sensory signals travel through, from the brainstem up to the cortex. Moreover, the different nature of the MMN as compared to the obligatory components, called exogenous that time, supported the authors’ view on a genuine process often criticized and seen by leading physiologists as artifact or dishabituation. Therefore, the main objective of the first studies was to characterize the MMN. Systematics studies, though not standardized at all, were performed to show how the deviance magnitude, physical properties, stimulus frequency and probability influence the ERP’s component structure. Although debates on the genuine MMN are still going on, the focus of the MMN studies

and models has changed. The concept of object abstraction gained attention in the late 1990s and seen as representation formulated through the auditory neurons tuned to different physical properties. However, how these cellular activities are integrated is still uncertain. It is, however, clear that the brain does not simply respond to changes in the acoustic environment as it also interprets them based on experiences. With this approach we are not far from the view of Helmholtz (1867 cited in 1948), who argued that perception operated in an inferential manner, and not simply emerging from a purely sensory operation.

According to the first models, the MMN essentially reflects changes detected via an automatic and active process that is scanning the acoustic environment. Therefore, it cannot be classified as a simple sensory act. The processing of subtle changes means more than the neural dishabituation do, so that the sensitivity of the auditory system is crucial for survival, such as alerting one to potential threats in the environment and directing orienting behavior if this change signals something important (Sokolov 1960a, 1960b, 1963). Moreover, the MMN was seen as the sign of an ‘oddness’ detection, a neural indicator that does not require attention or a certain level of saliency. Unfortunately, the influential DD theory led to several misunderstandings about the MMN stating that it merely reflects tone feature discrimination (e.g., frequency, intensity, duration, spatial location). Indeed, the most widely used paradigm to elicit MMN has been the auditory oddball paradigm using two tones and interpreted as the best index of sound feature discrimination. Unfortunately, the parallel use of the concepts detection and discrimination did not support a stable distinction between two important abilities of the neural system serving sensation and perception. It is valid to interpret the MMN when elicited as the sign of feature discrimination contributing to change detection. However, the lack of MMN does not necessarily mean that the two features are not discriminable on their own, so that the MMN is not necessarily an objective index of the auditory feature discrimination. The main reason is that the MMN is context-dependent (see Sussman 2007) so that the basis for MMN elicitation is regularity extraction from the ongoing input structured in short- (trace) or long-term (representation) memory. It is important to mention here that the change detection model thought to be valid for tones only led to further misunderstandings when the first manuscripts on phoneme MMN submitted by some laboratories led to skeptical comments and rejections.

The leading model of the 1990s postulated that auditory objects were stored in the sensory memory, and they were represented by a perceptual prediction (PP) model. It was proposed that memory representations were automatically compared with the auditory input, and MMN generation could be expected when they mismatched and acted as a prediction error signal (Giard et al. 1990). This approach also

meant that an insufficient object-input difference could not elicit an MMN, and the neural mechanism supposed to act was called repetition suppression (RS) or stimulus-specific adaptation (SSA) referred to as model adjustment or adaptation (for review see Song et al. 2023). While construction of the perceptual models predictive in nature and comparative processes based on detected changes in context-dependent discrimination can be classified as top-down, mechanisms associated with auditory object formation are seen as bottom-up processes. However, one of the many questions is how different the processes in the main domains, here tones versus speech, are. Are there general mechanisms in processing probabilities and regularities? Is the perception of significant features of speech the same as that of tone patterns often defined as auditory scene, or they are different in a sense that both rules anchored in representations and roles assigned to feature complexes contribute to detection giving rise to MMN. The immense proliferation of paradigms and models in the first two decades on searching for common mechanisms of the various types of MMN gave rise to many terms, including statistical learning and inference, sensory learning, auditory perceptual learning, change detection, sensory memory, predictive coding, auditory pattern learning, prediction error signaling, novelty processing, hierarchical rule learning, and automatic auditory discrimination. This richness of the terminology supposedly explanatory in nature shows well that the DD theory widely supported over four decades by thousands of publications is not universally accepted. The ‘why’ question can be better answered when we move to a special topic of research that is language. Here the paradigms used and the range and nature of interpretations are different from those used for the tone elicited MMN. This means the explanatory strength of short-term, trace-like object formations is limited, their role is not exactly the same in all domains.

## From detection to prediction

It seems that at least a partial consensus about the MMN exists stating this response may reflect longer-term stimulus characteristics too including temporal and spectral dynamics of the signal extracted from the stimulus history and maintained in memory. However, how and to what extent our MMN models rely on the existing theories of memory, especially those on linguistic processing, show variations between the different schools of cognitive neuroscience and psychology. This, however, is not new at all. Already at rebirth of the Hungarian psychology György Ádám wrote in the editorial chapter of Psychophysiology (1972): “The truth is that there are currently almost as many memory hypotheses as there are brain researchers or psychologists dealing with this complex issue, and many more theoretical

and methodological foundations must precede the birth of a realistic and verifiable unified theory of memory.” (p12) It seems that we have even more memory theories than in the year of publication of the book, so that we try here to come up with a construct that may explain the role of long-term memory processes in deviance discrimination present on behavioral and neural level associated with MMN emerging in different context and age.

MMN experiments need stimulus repetitions, and this is valid for speech processing studies as well, although this technique is necessary only for refreshing the canonical form of items to investigate (isolated speech sounds, syllables, words, pseudo-words). It seems that the DD model, although detection is an important part of the processing, is not the best candidate for interpreting the variations found. Therefore, we should investigate the interpretational strength of the predictive coding (PC) theory when rely on as a possible theoretical frame for understanding deviance processing in speech perception, especially because the PC theories are still underspecified when it comes to language. The PC theory (Friston 2005, 2009, 2010) consolidated with Bayes theorem (Kersten et al. 2004; Knill and Pouget 2004) has been momentarily influential as it aimed at integrating action, perception, attention, and learning (see Winkler and Schröger 2015). Moreover, it was and is seen as a unified theory of cortical function (Heilbron and Chait 2018) that postulates that the brain relies on a generative model that combines top-down predictions with bottom-up sensory input (de Lange et al. 2018), and pre-activates the cortical representation of a predicted stimulus. The pre-activated representation is compared with the sensory input; (1) a match between the sensory input and prediction induces a suppression in the neural response, a mechanism called expectation suppression (see Garrido et al. 2017; Summerfield et al. 2008; Todorovic and de Lange 2012). (2) A mismatch results in a prediction error signal (Friston 2005; Summerfield and de Lange 2014). Despite the success of using the model for explaining changes observed in basic visual and auditory domains, how the PC principles apply to higher level such as speech remains a matter of debate.

In general, the results of human MMN studies fit well into many models if non-speech sequences are manipulated. The traditional and novel models take different aspects into account and one of the leading ones suggests the application of the Bayesian learning not fully applicable yet for the MMN to segmental and suprasegmental variations and violations of speech. It seems, we better consider all the models developed when try to link the complex stimulus-, context- and representation-specific processes, as they all deal with different aspects and spatial and temporal variations, including asynchrony, of the underlying mechanisms. The main MMN models based on different assumptions are as follows.

- *The Deviance Detection (DD) model* sees the MMN as reflection of the detection of local physical changes in the sensory input.
- *The Stimulus-Specific Adaptation (SSA) theory* assumes the activity difference of adapted and non-adapted sensory neurons is indexed by the MMN.
- *The Model-Adjustment (MA) model* focuses on the pivotal role of the auditory cortex in the MMN generation by maintaining a model of the acoustic environment updated by the incoming stimulus.
- *The Novelty Detection (ND) model* suggests the MMN reflects the degree of novelty or surprise induced by the actual event in contrast with the pre-set context. In this model not the change per se but the violated prediction gives rise to the mismatch response. Moreover, it gives a possible explanation why the MMN is elicited absence of a predictable change, e.g., a surprise different in nature from the prediction error.
- *The Prediction Coding (PC) or Predictive Error (PE) model* provides one of the leading explanations and states the human brain implements approximate Bayesian inferences based on predictive coding. This means the MMN reflects the difference between the actual and predicted inputs and the result, called prediction error, indicates the direction in which the event deviates from the prediction. The PC model is different in this sense from the ND model.

Although there are several models explaining the generation of MMN, the researcher community still waits for a universal, possibly unified theory. It is possible the animal models help us further in getting closer to this goal. Indeed, the animal studies using these models as theoretical frames may provide a deeper insight into the neural processes and contribute to understand the major processes the MMN correlates with.

### **Animal models on stimulus-specific adaptation and prediction error**

The first animal model of MMN (Csépe et al. 1987) revealed that this response is not exclusively human and associated with processes of broader biological significance. The MMN elicited by subtle acoustic changes of the deviant stimulus could be recorded in freely moving cats with chronically implanted electrodes on the primary and secondary auditory cortex and from subcortical structures such as the important relay nuclei of the auditory pathway in the inferior colliculus and medial geniculate body (Csépe et al. 1988, 1989, for review see Csépe 1995). Responses of single neurons in the inferior colliculus, medial geniculate body, and auditory cortex of other mammals, such as rats and primates,

showed responses that shared several properties with the MMN. Based on the results of these studies, a novel proposal stated that the SSA of single neurons in the auditory cortex is the cellular substrate of MMN (Ulanovsky et al. 2003, Nelken and Ulanovsky 2007). The SSA proposal gave impetus to further studies addressing the neuronal basis of MMN recorded in anesthetized animals. SSA was identified at different stages of the auditory pathway: in the cat auditory cortex (Ulanovsky et al. 2003), the mouse auditory thalamus (Anderson et al. 2009), and rat inferior colliculus (Malmierca et al. 2009). However, we must bear in mind that the many attempts to demonstrate MMN with recording event-related brain potentials in rodents resulted in weak or ambiguous patterns (Lazar and Metherate 2003; Sambeth et al. 2003; Eriksson and Villa 2005; Umbricht et al. 2005; Astikainen et al. 2006; Tikhonravov et al. 2008). A well-designed study (Von der Behrens 2009) investigating single and multiunit neuronal activity recorded parallel with local field potentials in awake rat shed light on the contradicting results found in the various studies. The authors' conclusion was that single neurons in the rat auditory cortex adapt in a stimulus-specific manner and contribute to corresponding changes in the field potential. However, the SSA did not contribute to the late deviant response directly equivalent to the human MMN, so that it might reflect a certain part of the processes underlying sound discrimination.

Moreover, the strength of the animal models is based on the possibility to record on the neural level (single- and multiunit activity and field potential recordings, measure of receptor functions, etc.) and through all these to shed light on the neural mechanisms contributing to the MMN generation. The animal models crucial in the sense that all the MMN models based on human studies are qualitative in nature and do not make quantitative predictions. Even the very original proposal based on trial-by-trial modeling of the MMN (Lieder et al. 2013) has limitations. Here, according to the authors' suggestion the MMN reflects Bayesian learning of sensory regularities. This also means the main process generating the MMN is to adjust a probabilistic model of the environment according to the predictive error.

The animal models have more possibilities to contribute to our view on the possible role of predictive error in the MMN generation. As it seems there is a consensus on the contribution of two important brain areas playing a pivotal role in the MMN generation, the auditory cortex (AC) and the prefrontal cortex (PFC). Two major MMN generators, the frontal and the auditory cortical, have always been assumed, although their precise assignment to the processing types was not possible. The reason for investigating these areas in animals is to search for the precise location of the acoustic regularity encoding as well as to know more about the real nature of stimulus- and representation-specific processes assumed to have a temporal asynchrony. As it was

shown in rats by recording neuronal spiking activity and local field potentials (Casado-Román et al. 2020), the AC-recorded responses were driven by stimulus-dependent changes, and the PFC activity was sensitive to unpredictability, context-dependent, delayed, robust, and longer-lasting and signaling the prediction error. According to the authors' conclusion the time course of the mismatch responses as followed by parallel recording of the spiking activity and the local field potentials of the AC and the PFC corresponded the different MMN-like signals reported in the rat brain. These findings contributed a lot to resolving one of the main concerns about the animal studies giving contradicting results probably due to the levels used for recording the brain correlates of automatic deviance detection.

As mentioned above the animal models had and have an immense contribution to understanding the real nature of the biologically important predictions in the auditory scene. Beyond this the MMN seen as biological marker can be further used in animal only in studies designed for genetic manipulations. One of the first steps in this direction is the whole-cortical recording with multi-channel electrocorticograms (Komatsu et al. 2015) performed by investigating common marmosets. The authors found an exquisite sensitivity of the temporal area of the deviant stimuli and planned further studies aiming at developing a non-human model of schizophrenia (see Featherstone et al. 2018 for review).

## Role and rule

One of the surviving myths in the MMN research is that this response occurs only when the stimuli are repeated. As Fitzgerald and Todd (2020) wrote in their review paper; Functionally MMN is defined by two key characteristics: that it is context-dependent and does not rely on conscious attention to the stimulus. "Whereas both the N1 and N2b ERP components can be elicited by a deviant stimulus alone, the MMN response occurs only when the sound is interspersed among a series of repetitive standards." The problem is not only with the statement but with the generalized view that N1 and N2b (obligatory ERP components) are elicited by the deviant per se and MMN occurs only when it appears as interspersed within standards. This is basically valid for acoustic features when they used to serve the actual object formation. However, existing representations (speech sounds, spoken syllables, words, even pseudo-words, and familiar melodies, etc.) require repetitions only for technical reasons as object formation and rules applied have long-term representations. Complex units like, for example, word level prosody are tightened to representations and get activated in a unified process already by a single presentation that activates the representation via prediction. Deviating feature(s), complex(es) or rule(s) get detected when matching fails.

Therefore, we may assume that expectation-matching suppression and prediction-related error detection—based on still debated neural mechanisms—may occur in a single process and result in a neural response. This may also mean one presentation of the deviating stimulus or scene is sufficient to evoke the neural response, e.g., the MMN on its own. Moreover, prediction must work even without repetitions whenever the feature(s) and the rule(s) (realization, assignment, etc.) have a long-term representation. Representation-based predictions rely on feature and rule-extractions and not on short-lived traces, so that ‘repeated stimuli interspersed by deviants’ cannot be a crucial part of MMN definition, at least not for linguistic processes. The MMN studies seem to be unable to get rid of unconditional requirement on repetition which should only be technical and not definitional. The technique we use for computing the event-related responses embedded in the EEG (electroencephalography) recorded over the human scalp with surface electrodes has several limitations. One of these is the relatively high number of stimuli we need for averaging to get a response of relatively good signal-to-noise ratio in different ways. It is worth to note that we could investigate the unified suppression-prediction process assumed if we had reliable methods for single epoch analysis. Instead, we use several types of processes including averaging, all linear in nature. The general procedure used by many laboratories is averaging multiple single-trial epochs to create an averaged ERP waveform, and we also average the values at multiple time points within a time window when aim at quantifying the amplitude of a component with a mean amplitude measurement. *t* means that repetition is the prerequisite of the techniques used, so that we should not think, that the neural processing of linguistic features such as phonemic deviances, foreign accent, syllabic stress violations need repetitions for an adult brain.

Traces or representations? Is that a crucial question? The answer is yes and no. The expression trace is broadly used in the tonal MMN literature and less and less accepted in the speech MMN studies. The comprehensive MMN review published by Risto Näätänen states (2001, p 1) that “each sound, both speech and nonspeech, develops its neural representation corresponding to the percept of this sound in the neurophysiological substrate of auditory sensory memory” and the perception of phonemes, even of syllables and words is “based on language-specific phonetic traces developed in the posterior part of the left-hemisphere auditory cortex.” Are they traces or representations? Psychologist and psycholinguists consistently use the term representation that refers to categorically organized and flexibly usable neural imprints of the canonical realization of speech units prone to variations to a certain extent. Although the view and the model applied were modern 20 years ago, the rigid frame of traces led sometimes to dogmas in speech perception

research using MMN. Especially the general approach relying on the assumption that accuracy of the representation can be probed by this method for all auditory features. Unfortunately, speech units (sounds, syllables, words) do not differ in single features only as even the smallest units are determined by feature-complexes, we may call them compounds; different ones can be assigned to the same role and same or very similar ones to a different role. While the rules on how to compose the features is language-related, the role of units is mostly universal where just the feature complex is different. For example, while a salient feature-complex is used for word-level stress, the accentuation relies on qualitatively different characteristics with rules for assignment and role-associated acoustic features. The model we use deals with the rules, e.g., how the representation of acoustic feature-complexes emerges, and with the role of these representations, e.g., how they may contribute to speech perception and learning.

### Beyond trace—the role of mental lexicon

Several models of spoken word recognition have been developed, and the most current models explain it in terms of activation of processing units within a mental lexicon. These models consider spoken words’ processing as matching speech signals to representations stored in the mental lexicon (Moss and Gaskell 1999:59). Although there is no clear consensus about the exact structure and organization of the mental lexicon among the existing models, they all concur on the fact that the mental lexicon is certainly not a list of word forms. A fundamental but unsettled issue for spoken word processing is the type of information that constrains the activation in each entry. Several studies have demonstrated since the influential paper of Näätänen et al. (1997) published in *Nature* that the deviation detection of native speech sounds is based on long-term representations of any relevant language-specific information. The question is whether the processing of suprasegmental features follows the same matching process as it does for the phonemic ones.

In the early 2000s two laboratories were so far in modeling this representation-based matching that they started the first studies on suprasegmental feature-matching by using the MMN paradigm. The first publication from the Max Planck Institute (Weber et al. 2004) and from ours at the Research Institute of Psychology (Honbolygó et al. 2004) did not gain immediate attention, and any further study has only started a decade later in other laboratories.

While our study of 2004 (Honbolygó et al. 2004) used words to investigate the matching process for stress pattern, the one from 2013 applied pseudo-words (Honbolygó and Csépe 2013). The idea emerged from the critics on our paper arguing that the phonemic deviation used as reference is

lexical in nature, and this might influence the processing of words deviating for stress. Our hypotheses on the processing of unfamiliar stress patterns were based on the idea that accentuation in a fixed-stress language like Hungarian relies on long-term representation of a general rule on stress assignment. This rule to apply for any spoken word-like unit, here pseudo-words, was referred to as stress template activated as a general object formation valid for processing word-level suprasegmental features. This template is used when Hungarian adults pronounce any word-like unit be Hungarian words, pseudowords or foreign words not familiar for the speaker. We also supposed these templates to be language-specific and hypothesized as pre-lexical. To reveal whether matching the stress representation with the input relies on short-term traces or on long-term representations, two experimental conditions using pseudoword stress variations for the deviant's legality (the legal term here refers to the canonical native stress pattern) were introduced. In the illegal deviant condition, the pseudoword with legal stress pattern, e.g., stressed first syllable (the canonical stress-assignment in Hungarian) served as standard, and the pseudoword with illegal stress pattern served as deviant. In the legal deviant condition, stress pattern had a reversed role, e.g., a stress assignment non-existent in Hungarian was used for as standard. According to our hypothesis, in adults (1) the illegal stress pattern used as deviant would elicit two MMN components; but (2) the legal stimulus as deviant would not. We assumed that a legally stressed deviant does not elicit an MMN like the illegal one and will occur only if the formation of a short-term trace is successful by repeating the standard of illegal pattern and it overcomes the impact of the long-term representation valid only for the legal template. In line with our expectations based on our previous studies, two consecutive MMNs were elicited by the illegal deviant, and no MMN by the legal one. We proposed that the stress template might play a prominent role in speech perception and can be crucial in the development of and access to the mental lexicon. Although our results confirmed that word-level stress processing is based on a strong representation, we assumed that this is not fully valid for languages using several rules and variations so that learning should have a delicate role in long-term object formation.

### Expectations, predictions, and word level stress processing

Our adult studies on stress pattern violation gave rise to a new question emerged from our experience on reversed experimental blocks (stimuli in one block as standard and in another one as deviant) used for confirming on the genuine MMN in tonal and phonemic paradigms. Words and pseudowords of legal stress delivered as deviants within repeated

standards of illegal pattern did not elicit reliable MMNs and the changes found were not typical. We assumed the processing of word stress associated with rigid rule application in Hungarian differed in a sense from any other processes of linguistic units eliciting MMN. We also assumed the word stress representation might undergo a one rule applies for all developmental change during infancy. This would explain why the legal pattern as deviant did not work and why the illegal pattern could not be repeated in sufficient number for building up a trace serving reference of the neural matching/mismatching process. Our assumption agreed with the stress deafness theory (for details see Dupoux et al. 2008) only so far that we hypothesized the impact of learning on decreasing sensitivity to foreign language feature complexes as well as to non-native stress templates. However, it is important to bear in mind that the syllabic stress is realized by using feature complexes in different languages and their assignment varies as one or more templates are used according to the syllabic structure or lexical status. This means that the existing variations of syllabic stress as feature complex as well as of the template (s) or rule(s) are to learn from birth on.

Unfortunately, only a few MMN studies have provided reliable evidence for the emerging stress templates in infancy. It seems the strong, nearly rigid representations in fixed-stressed languages are the best candidates for testing our hypotheses on the existence of templates used as rules. The study of Weber et al. (2004), mentioned above investigated 4- and 5-month-old infants' growing up in monolingual German families. In German, 90% of the bisyllabic content words have a strong–weak trochaic (emphasis is on the first syllable) stress pattern. The pseudoword *baba* (this as word in Hungarian was used in our studies performed later) with two different stress patterns (stress on the first or on the second syllable) were presented in two conditions. The deviant stress could elicit a significant mismatch response of positive polarity (MMR) and was present in the 5-month-olds. We may interpret these data nowadays as a clear sign of maturation. The next study, a German–French crosslinguistic comparison of word-level stress processing was published by Friederici et al. (2007) implied the emergence of language-specific stress representation at 4.5 months of age. Native French and German monolingual infants were tested in the study using a mismatch paradigm very similar to the one used by Weber et al. (2004). In German, the stress is predominantly on the first syllable, while in French the dominant pattern is stress on the second. The word stress in this study was realized by changes in the vowel length and formant structure. While in German babies pseudowords stressed on the second syllable elicited an MMR, MMR to the first syllable was elicited in the French learners. The MMR occurred only when the dominant stress pattern of the native language was violated. However, we may ask why lexicality

has a weak impact on word stress processing in infants. We can give a reliable answer if the emergence of long-term stress representation in infancy is systematically investigated. Despite the emergence of a few studies related to stress perception in infants, the testing of Hungarian infants is relevant in revealing how cross-linguistic differences in word stress organization result in word stress processing variations. Around 6 months of life infants appear to shift their attention from prosody to phonemic structure of their native language. Moreover, lexicality does not play a role for several months, so that a question to answer is the interaction of lexical status and word stress in the early stages of language acquisition.

As we mentioned above, the MMN is context-dependent (see Sussman 2007), and as it is shown by most studies on non-speech paradigms, the expectations are related to short-term traces (tones and tone patterns) or in some cases (music) to representations. However, speech is a different story in many aspects as native language-related predictions are based on representations. A recent version of the dual-stream model of language processing proposed that the predictive sequential processing of linguistic information is performed by hierarchically organized internal models where the posterodorsal stream has a pivotal role. Our study (Honbolygó et al. 2020) using functional magnetic resonance imaging (fMRI) aimed to shed light on an unexplored area of predictive processes in speech, e.g., the role of expectation of prosodic segmentation of linguistic information.

The main hypothesis is that predictive inferences are processed by a dual auditory stream network of the brain and involve both the ventral and dorsal streams of slightly different functions (Bornkessel-Schlesewsky et al. 2015). The proposal is of special interest as the dorsal stream involves the superior longitudinal fascicles for parieto-frontal and the arcuate fascicle for temporo-frontal connections. Word stress, a significant component of speech segmentation, was investigated in an event-related acoustic fMRI repetition suppression (RS) paradigm. The RS modulation as the marker of predictive processing. Here, the main idea was derived from the free energy principle (Friston 2010) where the brain is seen as “prediction machine” that is optimized for diminishing free energy. Pairs of pseudowords of same or different stress patterns were delivered in blocks, and the BOLD (blood-oxygen-level-dependent) signal was significantly lower for the same than for the different trials. These results speak for the important role of the superior temporal gyrus in predictive processing of word stress as well as for the dorsal auditory stream’s involvement in activating the representation-based expectation on language-typical word stress templates used as rules in a fixed-stressed languages, probably not only in Hungarian.

## The development of prosody-lexicality integration

The first MMN study on newborns was published in 1990 by the research group of Näätänen (Alho et al. 1990) followed years later by several studies using different paradigms, including those designed for investigating the processing of phonemic contrasts (for review see Csépe 1995). Results of the further studies shed light not only on the mismatch responses’ nature as one of the main aims was to know more about the maturation of possible neural networks responsible for generation of the MMN as well as the development of acoustic sensitivity to changes formed via experimental variations (tones) or exposition to one or more languages. Moreover, a decade of debates led to a consensus on naming of the component associated with mismatching. The name MMR was broadly used by many developmental studies. It seemed the MMR was stable over the developmental timeline (Kushnerenko et al. 2002) and showed a significant latency decrease during the first two years of life and reached its typical onset and peak latency around the 3rd year of age (Morr et al. 2002). In the next decade, the component has been widely used in clinical samples of infants to investigate the assumed anomalies in auditory processing (Fellman and Huotilainen 2006; Jansson-Verkasalo et al. 2010; Leipälä et al. 2011; Ragó et al. 2014). The amplitude of the MMR was found to associate with gestational age (Leppänen et al. 2004), and several perinatal factors, like the intrauterine growth restriction (Fellman et al. 2004), or perinatal asphyxia (Leipälä et al. 2011). The MMR was also seen as a possible tool to predict the language development of full-term (Friedrich et al. 2009; Weber et al. 2005) and pre-term infants (Jansson-Verkasalo et al. 2010).

The focus of the MMR studies performed before 2004 was on the maturation and development of phoneme processing shown as basically prosody-free before the 6th month of age. Saffran and Thiessen (2003) provided empirical evidence that 6-month-old infants used statistical regularity of the different phoneme sequences when recognized repeatedly presented nonsense words, not relying on the predominant stress pattern of their native language. It makes sense if we agree on the model of Becker et al. (2018) who suggested different developmental trajectories for reliable representations of the phonemic and prosodic features of spoken utterances. As the authors assumed German infants showed a reliable integration of prosodic and phoneme-relevant information only by the 9th months of age, in line with other studies (Johnson and Jusczyk 2001; Thiessen and Saffran 2003).

However, the influential study of Skoruppa et al. (2009) demonstrated that the assumed interplay depends on the



role of stress in the given language with or with no use of lexical stress. They found that 9-month-old Spanish infants discriminated the various stress patterns of segmental variability, unlike their French peers. This means that several features of the segmental and suprasegmental structure contribute to an accurate processing, and the prerequisites of a successful integration—stress assignment rules (e.g., variable stress, fixed stress) and roles (e.g., lexical, non-lexical)—is not clear.

A series of our MMR studies aimed at investigating the impact of lexical status we assumed to contribute to our intriguing results found of the adult study (Honbolygó and Csépe 2013), e.g., the emergence of modulation effect on stress processing. We succeeded to find the expected suppression and facilitation effects of lexical status (Varga et al. 2019) with respect to word stress processing at 6 and 10 months of age. Moreover, we found an age effect in the legal deviant (first syllabic stress) condition of the word paradigm, where the lexical and stress cues conflicted. Our results demonstrated that only the 10-month-old infants, in contrast with the 6-month-olds, were able to integrate the lexical and stress cues. A further study (Varga et al. 2021) performed in a group of pre-term infants demonstrated that the shortened intrauterine language experience is one of the explanatory factors of atypical prosodic development. The results of all studies performed by our research group suggest that the MMR is a reliable tool for investigating the development of word level stress processing. It seems that even the integration of prosodic and lexical cues starting around the 6th months of age, completed primarily around the 10th month of age, can be followed by using this method. Moreover, the start of integration was found to be at an earlier time point than the date suggested by Becker et al. (2018) in their investigation of German infants (9 months). We attributed this difference to the independent phoneme-relevant and word level stress-relevant rules applied in the Hungarian language as compared to lexical and variable stress languages.

## Lessons learnt

We learnt a lot during the 45 years history of MMN concerning its generation, maturation, development, variations in clinical cases as well the changing interpretations and models aiming at explaining how this event-related response might contribute to our understanding of the complexity of auditory scene analysis. The MMN is one of the biological correlates psychologists may use to go beyond the behavioral correlates of these processes. As in the editorial chapter of Psychophysiology (Ádám 1972, p 6) stated: “The physiological “invasion” of psychology promises to be productive just as the “intrusion” of sociological or pedagogical

knowledge and methods into the field of psychology proves to be fruitful. Well, today many classical disciplinary ideas and methods struggle with the aged boundaries of their own traditionally developed system of knowledge. The opening of the borders provides an opportunity for an ideological and methodological renewal!” Indeed, the psychophysiological approach both in classical and modern terms largely contributed to understand how our acoustic environment was continuously scanned, how we predicted the incoming events based on rules, roles, regularities, and short- and long-term representations.

The concept on perception and memory classified by György Ádám as “reflections of a biologist” (Ádám 1980) found a fertile soil in the psychophysiological research in Hungary and contributed to the rise of a new research field within the cognitive neuroscience in Hungary. The smooth transition from psychophysiology to cognitive neuroscience was helped by the pioneer work of Ádám published in English in 1980. Neither György Ádám nor the psychologists in Hungary used this name although it has been coined by Michael Gazzaniga around this time using it as the title of a new institute developed according to the cognitive science initiative in the USA (for reference see Posner and DiGirolamo 2000). Results of the MMN studies performed first in Finnish and Hungarian laboratories have gained recognition by researchers of several disciplines. Today a richer than ever methodology is available, and a significant development of the scan-adapt-predict models support the complex nature of the acoustic change-detection. However, we should not forget the first models on the underlying mechanisms of MMN elicited by non-speech stimuli assumed memory traces to have a pivotal role in the comparison processes. The mechanisms seen by György Ádám were “functional or structural changes of some sort in the central neurons, producing what is called a memory trace or engram for the lack of a better term” (Ádám 1980, p 193). Ádám following the contemporary theories of memory stated that none of them “lived as long as the currently favored one of memory traces” although no one have “discovered a more plastic and illustrative metaphor to describe the essence of memory that than of the waxen tablet used by Plato” (Ádám 1980, p 194).”

The MMN research went through different phases ranging from misunderstandings, skeptics to broad acceptance, though its use and interpretation is still not free from myths and dogmas. The future of the curiosity-driven and applied research using MMN research depends on whether we successfully overcome the reproducibility crises, develop generally accepted protocols, use standards, and design our paradigms in the view of ecological validity as well. The most important benchmark studies, systematic and comprehensive reviews may help us to find the right compass to see the wood from the tree. A reliable compass is well needed

to find the most significant contributions in form of publications on original research and reviews found in Scopus (more than 19 thousand) cited in high number (more than 17 thousand) as given in the Web of Science Core Collection as well as in different forms of personal communications and discussions. Therefore, we need a particular awareness in several topics:

- First, designing and applying an MMN paradigm may have several aims, including crucial questions raised by psychologists about the nature of auditory perception of various stimuli delivered in different modalities, here acoustic classified as speech and non-speech.
- Second, we must be aware that all measured MMN parameters including polarity, amplitude, latency, distribution, and estimated sources show a large variance associated with the processes contributing to.
- Third, infant and child studies should pay attention to the fact that maturation and development contribute to variations in the recorded activity, like, for example, the MMR a term broadly used in the recent developmental studies related to its polarity reversal, not fully known for the contributing factors.
- Fourth, the ideal or optimal MMN paradigms and models of high explanatory strength are still under development, and many questions be psychological and/or neuroscientific in nature wait for being better addressed and answered.
- Fifth, the MMN paradigms developed for studying the processing of linguistic features require a different focus not biased by the results gained from non-speech studies and models used as leading hypotheses.
- Sixth, a proper account on timescales of learning, especially that of language under a heavy impact of maturation and development, should count to a high extent, so that informative considerations may add to understanding the differences of paradigms used and the groups investigated who differ across paradigms.
- Seventh, in linguistic paradigms, the MMN may reflect both lower level auditory and higher level limited (pseudo-words) or full (word) lexical processes and predictions complex in nature.
- Eighth, it is expected that theories focusing on the modulating effects of long-term experience in sensory and linguistic processing will gain a broader acceptance. However, for this a better understanding of the different models—the predictive coding framework, the model-adjustment hypothesis, the neuronal memory circuits, or the dual auditory stream model—is needed.

**Acknowledgements** This research was supported by the Research Programme for Public Education Development, Hungarian Academy of

Sciences (SZKF-20/2021; PI: FH), the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant agreement No. 813546 (co-authors and PIs VCs, FH). Special thanks of the first author go to her late mentors, György Ádám, György Karmos, Evgeny N. Sokolov and Risto Näätänen.

**Funding** Open access funding provided by HUN-REN Research Centre for Natural Sciences.

## Declarations

**Conflict of interest** This work is directly or indirectly related to financial or non-financial interests of the authors.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Ádám György (1972) Tanulás, motiváció, tudat—Előszó helyett. In: Ádám György (szerk.) *Pszichofiziológia*, Gondolat Könyvkiadó, Budapest, p 5–16
- Ádám György (1980) Perception, consciousness, memory: reflections of a biologist, Akadémiai Kiadó. Budapest and Plenum Press, New York; London
- Alho K, Sainio K, Sajaniem N, Reinikainen K, Näätänen R (1990) Event-related brain potential of human newborns to pitch change of an acoustic stimulus. *Electroenceph Clin Neurophysiol* 77:151–155. [https://doi.org/10.1016/0168-5597\(90\)90031-8](https://doi.org/10.1016/0168-5597(90)90031-8)
- Anderson LA, Christianson GB, Linden JF (2009) Stimulus-specific adaptation occurs in the auditory thalamus. *J Neurosci* 29:7359–7363. <https://doi.org/10.1523/JNEUROSCI.0793-09.2009>
- Astikainen P, Ruusuvirta T, Wikgren J, Penttonen M (2006) Memory-based detection of rare sound feature combinations in anesthetized rats. *NeuroReport* 17:1561–1564. <https://doi.org/10.1097/01.wnr.0000233097.13032.7d>
- Becker A, Schild U, Friedrich CK (2018) Tracking independence and merging of prosodic and phonemic processing across infancy. *Dev Sci* 21:1–11
- Bornkessel-Schlesewsky I, Schlesewsky M, Small SL, Rauschecker JP (2015) Neurobiological roots of language in primate audition: common computational properties. *Trends Cognit Sci* 19:142–150. <https://doi.org/10.1016/j.tics.2014.12.008>
- Casado-Román L, Carbajal GV, Pérez-González D, Malmierca MS (2020) Prediction error signaling explains neuronal mismatch responses in the medial prefrontal cortex. *PLoS Biol* 18(12):e3001019. <https://doi.org/10.1371/journal.pbio.3001019>
- Csépe V, Karmos G, Molnár M (1987) Evoked potential correlates of stimulus deviance during wakefulness and sleep in cat—animal model of mismatch negativity. *Electroenceph Clin Neurophysiol* 66:571–578. [https://doi.org/10.1016/0013-4694\(87\)90103-9](https://doi.org/10.1016/0013-4694(87)90103-9)

- Csépe V, Karmos G, Molnár M (1988) Evoked potential correlates of sensory mismatch process during sleep in cats. In: Koella, WP, Obál F, Schulz H, Visser P (eds) *Sleep '86*. Gustav Fischer Verlag, Stuttgart, p 281–283
- Csépe V, Karmos G, Molnár M (1989) Subcortical evoked potential correlates of early information processing: mismatch negativity in cats. In: Başar E, Bullock TH (eds) *Brain Dynamics*. Springer, Berlin, Heidelberg, p 279–289. [https://doi.org/10.1007/978-3-642-74557-7\\_23](https://doi.org/10.1007/978-3-642-74557-7_23)
- de Lange FP, Heilbron M, Kok P (2018) How do expectations shape perception? *Trends Cogn Sci* 22:764–779. <https://doi.org/10.1016/j.tics.2018.06.002>
- Dupoux E, Sebastián-Gallés N, Navarrete E, Sharon Peperkamp S (2008) Persistent stress ‘deafness’: the case of French learners of Spanish. *Cognition* 106:682–706. <https://doi.org/10.1016/j.cognition.2007.04.001>
- Eriksson J, Villa AEP (2005) Event-related potentials in an auditory oddball situation in the rat. *Biosystems* 79:207–212. <https://doi.org/10.1016/j.biosystems.2004.09.017>
- Featherstone RE, Melnychenko O, Siegel SJ (2018) Mismatch negativity in preclinical models of schizophrenia. *Schizophr Res* 191:35–42. <https://doi.org/10.1016/j.schres.2017.07.039>
- Fellman V, Huotilainen M (2006) Cortical auditory event-related potentials in newborn infants. *Semin Fetal Neonatal Med* 11:452–458. <https://doi.org/10.1016/j.siny.2006.07.004>
- Fellman V, Kushnerenko E, Mikkola K, Ceponiene R, Leipälä J, Näätänen R (2004) Atypical auditory event-related potentials in preterm infants—a possible sign of cognitive dysfunction? *Pediatr Res* 56:291–297. <https://doi.org/10.1203/01.PDR.0000132750.97066.B9>
- Fitzgerald K, Todd J (2020) Making sense of mismatch negativity. *Front Psych* 11:468. <https://doi.org/10.3389/fpsy.2020.00468>
- Friederici AD, Friedrich M, Christophe A (2007) Brain Responses in 4-month-old infants are already language specific. *Curr Biol* 17:1208–1211. <https://doi.org/10.1016/j.cub.2007.06.011>
- Friedrich M, Herold B, Friederici AD (2009) ERP correlates of processing native and non-native language word stress in infants with different language outcomes. *Cortex* 45:662–676. <https://doi.org/10.1016/j.cortex.2008.06.014>
- Friston KJ (2005) A theory of cortical responses. *Philos Trans R Soc B Biol Sci* 36:815–836. <https://doi.org/10.1098/rstb.2005.1622>
- Friston KJ (2009) The free-energy principle: a rough guide to the brain? *Trends Cogn Sci* 13:293–301. <https://doi.org/10.1016/j.tics.2009.04.005>
- Friston KJ (2010) The free-energy principle: a unified brain theory? *Nat Rev Neurosci* 11:127–138. <https://doi.org/10.1038/nrn2787>
- Garrido MI et al (2017) Bayesian mapping reveals that attention boosts neural responses to predicted and unpredicted stimuli. *Cereb Cortex* 28:1771–1782. <https://doi.org/10.1093/cercor/bhx087>
- Giard M-H, Perrin F, Pernier J, Bouchet P (1990) Brain generators implicated in the processing of auditory stimulus deviance: a topographic event-related potential study. *Psychophysiology* 27:627–640. <https://doi.org/10.1111/j.1469-8986.1990.tb03184.x>
- Heilbron M, Chait M (2018) Great expectations: is there evidence for predictive coding in auditory cortex? *Neuroscience* 389:54–73. <https://doi.org/10.1016/j.neuroscience.2017.07.061>
- Honbolygó F, Csépe V (2013) Saliency or template? ERP evidence for long-term representation of word stress. *Int J Psychophysiol* 87:165–172. <https://doi.org/10.1016/j.ijpsycho.2012.12.005>
- Honbolygó F, Csépe V, Ragó A (2004) Suprasegmental speech cues are automatically processed by the human brain: a mismatch negativity study. *Neurosci Lett* 363:84–88. <https://doi.org/10.1016/j.neulet.2004.03.057>
- Honbolygó F, Kóbor A, Hermann P, Kettinger ÁO, Vidnyánszky Z, Kovács G, Csépe V (2020) Expectations about word stress modulate neural activity in speech-sensitive cortical areas. *Neuropsychologia* 143:107467. <https://doi.org/10.1016/j.neuropsychologia.2020.107467>
- Jansson-Verkasalo E, Ruusuvirta T, Huotilainen M, Alku P, Kushnerenko E, Suominen K et al (2010) Atypical perceptual narrowing in prematurely born infants is associated with compromised language acquisition at 2 years of age. *BMC Neurosci* 11:88. <https://doi.org/10.1186/1471-2202-11-88>
- Johnson EK, Jusczyk PW (2001) Word segmentation by 8-month-olds: when speech cues count more than statistics. *J Mem Lang* 44:548–567. <https://doi.org/10.1006/jmla.2000.2755>
- Kersten D, Mamassian P, Yuille A (2004) Object perception as Bayesian inference. *Annu Rev Psychol* 55:271–304. <https://doi.org/10.1146/annurev.psych.55.090902.142005>
- Knill DC, Pouget A (2004) The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends Neurosci* 27:712–719. <https://doi.org/10.1016/j.tins.2004.10.007>
- Komatsu M, Takaura K, Fujii N (2015) Mismatch negativity in common marmosets: Whole-cortical recordings with multi-channel electrocorticograms. *Sci Rep* 5:15006. <https://doi.org/10.1038/srep15006>
- Kushnerenko E, Ceponiene R, Balan P, Fellman V, Näätänen R (2002) Maturation of the auditory change detection response in infants: longitudinal ERP study. *NeuroReport* 13:1843–1848. <https://doi.org/10.1097/00001756-200210280-00002>
- Lazar R, Metherate R (2003) Spectral interactions, but no mismatch negativity, in auditory cortex of anesthetized rat. *Hear Res* 181:51–56. [https://doi.org/10.1016/s0378-5955\(03\)00166-7](https://doi.org/10.1016/s0378-5955(03)00166-7)
- Leipälä JA, Partanen E, Kushnerenko E, Huotilainen M, Fellman V (2011) Perinatal cerebral insults alter auditory event-related potentials. *Early Human Dev* 87:89–95. <https://doi.org/10.1016/j.earlhumdev.2010.11.009>
- Leppänen PHT, Guttorm TK, Pihko E, Takkinen S, Eklund KM, H. Lyytinen H. (2004) Maturation effects on newborn ERPs measured in the mismatch negativity paradigm. *Exp Neurol* 190:91–101. <https://doi.org/10.1016/j.expneurol.2004.06.002>
- Lieder F, Daunizeau J, Garrido MI, Friston KJ, Stephan KE (2013) Modelling trial-by-trial changes in the mismatch negativity. *PLoS Comput Biol* 9:e1002911. <https://doi.org/10.1371/journal.pcbi.1002911>
- Malmierca MS, Cristaudo S, Pérez-González D, Covey E (2009) Stimulus-specific adaptation in the inferior colliculus of the anesthetized rat. *J Neurosci* 29:5483–5493. <https://doi.org/10.1523/JNEUROSCI.4153-08.2009>
- Morr ML, Shafer VL, Kreuzer JA, Kurtzberg D (2002) Maturation of mismatch negativity in typically developing infants and preschool children. *Ear Hear* 23:118–136. <https://doi.org/10.1097/00003446-200204000-00005>
- Moss HE, Gaskell MG (1999) Lexical semantic processing during speech. In: Garrod S, Pickering M (eds) *Language processing*. Psychology Press, Hove, UK. <https://doi.org/10.4324/9781315784724>
- Näätänen R (2001) The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). *Psychophysiology* 38:1–21. <https://doi.org/10.1017/s0048577201000208>
- Näätänen R, Gaillard AWK, Mantysalo S (1978) Early selective attention effect on evoked potential reinterpreted. *Acta Physiol (oxf)* 42:313–329. [https://doi.org/10.1016/0001-6918\(78\)90006-9](https://doi.org/10.1016/0001-6918(78)90006-9)
- Näätänen R, Lehtokoski A, Lennes M, Cheour M, Huotilainen M, Iivonen A et al (1997) Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature* 385:432–434. <https://doi.org/10.1038/385432a0>
- Nelken I, Ulanovsky N (2007) Mismatch negativity and stimulus-specific adaptation in animal models. *J Psychophysiol* 21:214–223. <https://doi.org/10.1027/0269-8803.21.34.214>

- Posner M, DiGirolamo GJ (2000) Cognitive neuroscience: origins and promise. *Psychol Bull* 126:873–889. <https://doi.org/10.1037/003-2909.126.6.873>
- Ragó A, Honbolygó A, Zs R, Beke A, Csépe V (2014) Effect of maturation on suprasegmental speech processing in full- and preterm infants: a mismatch negativity study. *Res Dev Disabil* 35:192–202. <https://doi.org/10.1016/j.ridd.2013.10.006>
- Saffran JR, Thiessen ED (2003) Pattern induction by infant language learners. *Dev Psychol* 39:484–494. <https://doi.org/10.1037/0012-1649.39.3.484>
- Sambeth A, Maes JH, Van Luitelaar G, Molenkamp IB, Jongsma ML, Van Rijn CM (2003) Auditory event-related potentials in humans and rats: effects of task manipulation. *Psychophysiology* 40:60–68. <https://doi.org/10.1111/1469-8986.00007>
- Skoruppa K, Pons F, Christophe A, Bosch L, Dupoux E, Sebastián-Gallés N, Limissuri RA, Peperkamp S (2009) Language-specific stress perception by 9-month-old French and Spanish infants. *Dev Sci* 12:914–919. <https://doi.org/10.1111/j.1467-7687.2009.00835>
- Sokolov EN (1960b) Neuronal models in the orienting reflex. In: Brazier MA (ed) *The Central Nervous System and Behavior*. Macy Foundation, New York, pp 187–271
- Sokolov EN (1963) Higher nervous functions: the orienting reflex. *Annu Rev Physiol* 25:545–580. <https://doi.org/10.1146/annurev.ph.25.030163.002553>
- Sokolov EN (1960a) The neural model of the stimulus and the orienting reflex (**In Russian**). *Voprosy Psichologii*, 4, 61–72
- Song P, Zhai Y, Yu X (2023) Stimulus-specific adaptation (SSA) in the auditory system: functional relevance and underlying mechanisms. *Neurosci Biobehav Res* 149:105190
- Summerfield C, de Lange FP (2014) Expectation in perceptual decision making: neural and computational mechanisms. *Nat Rev Neurosci* 15:745–756. <https://doi.org/10.1038/nrn3838>
- Summerfield C et al (2008) Neural repetition suppression reflects fulfilled perceptual expectations. *Nat Neurosci* 11:1004–1006. <https://doi.org/10.1038/nn.2163>
- Sussman ES (2007) A new view on the MMN and attention debate. *J Psychophysiol* 21(3–4):164–175. <https://doi.org/10.1027/0269-8803.21.34.164>
- Thiessen ED, Saffran JR (2003) When cues collide: use of stress and statistical cues to word boundaries by 7- to 9-month-old infants. *Dev Psychol* 39(4):706–716. <https://doi.org/10.1037/0012-16>
- Tikhonravov D, Neuvonen T, Pertovaara A, Savioja K, Ruusuvirta T, Näätänen R, Carlson S (2008) Effects of an NMDA-receptor antagonist MK-801 on an MMN-like response recorded in anesthetized rats. *Brain Res* 1203:97–102. <https://doi.org/10.1016/j.brainres.2008.02.006>
- Todorovic A, De Lange FP (2012) Repetition suppression and expectation suppression are dissociable in time in early auditory evoked fields. *J Neurosci* 32:13389–13395. <https://doi.org/10.1523/JNEUROSCI.2227-12.2012>
- Ulanovsky N, Las L, Nelken I (2003) Processing of low-probability sounds by cortical neurons. *Nat Neurosci* 6:391–398. <https://doi.org/10.1038/nn1032>
- Umbricht D, Vysotki D, Latanov A, Nitsch R, Lipp HP (2005) Deviance-related electrophysiological activity in mice: is there mismatch negativity in mice? *Clin Neurophysiol* 116:353–363. <https://doi.org/10.1016/j.clinph.2004.08.015>
- Valéria C (1995) On the origin and development of the mismatch negativity. *Ear Hear* 16:91–104
- Varga Z, Garami L, Ragó A, Honbolygó F, Csépe V (2019) Does intra-uterine language experience modulate word stress processing? An ERP study. *Res Dev Disabil* 90:59–71. <https://doi.org/10.1016/j.ridd.2019.04.01>
- Varga Zs, Ragó A, Honbolygó F, Csépe V (2021) Disrupted or delayed? Stress discrimination among preterm as compared to full-term infants during the first year of life. *Infant Behav Dev* 62:101520 <https://doi.org/10.1016/j.infbeh.2020.101520>
- Von der Behrens W, Bäuerle P, Kössl M, Gaese BH (2009) Correlating stimulus-specific adaptation of cortical neurons and local field potentials in the awake rat. *J Neurosci* 29(44):13837–13849. <https://doi.org/10.1523/JNEUROSCI.3475-09.2009>
- Von Helmholtz H (1867) *Handbuch der physiologischen Optik*. Voss, Leipzig
- Weber C, Hahne A, Friedrich M, Friederici AD (2004) Discrimination of word stress in early infant perception: electrophysiological evidence. *Cogn Brain Res* 18:149–161. <https://doi.org/10.1016/j.cogbrainres.2003.10.001>
- Weber C, Hahne A, Friedrich M, Friederici AD (2005) Reduced stress pattern discrimination in 5-month-olds as a marker of risk for later language impairment: neurophysiological evidence. *Cogn Brain Res* 25:180–187. <https://doi.org/10.1016/j.cogbrainres.2005.05.007>
- Winkler I, Schröger E (2015) Auditory perceptual objects as generative models: setting the stage for communication by sound. *Brain Lang* 148:1–22. <https://doi.org/10.1016/j.bandl.2015.05.003>
- Zs V, Ragó A, Honbolygó F, Csépe V (2021) Disrupted or delayed? Stress discrimination among preterm as compared to full-term infants during the first year of life. *Infant Behav Dev* 62:101520. <https://doi.org/10.1016/j.infbeh.2020.101520>