HHS Public Access

Author manuscript

Int J Psychophysiol. Author manuscript; available in PMC 2016 February 01.

Published in final edited form as:

Int J Psychophysiol. 2015 February ; 95(2): 175–183. doi:10.1016/j.ijpsycho.2014.06.010.

The role of temporal structure in the investigation of sensory memory, auditory scene analysis, and speech perception: A healthy-aging perspective

Johanna Maria Rimmele, PhD,
Department of Neurophysiology and Pathophysiology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

Elyse Sussman, PhD, and
Albert Einstein College of Medicine, Dominick P. Purpura Department of Neuroscience, Bronx, NY

David Poeppel, PhD
Department of Psychology and Center for Neural Science, New York University, New York, NY and Max-Planck Institute for Empirical Aesthetics, Frankfurt, Germany

Abstract

Listening situations with multiple talkers or background noise are common in everyday communication and are particularly demanding for older adults. Here we review current research on auditory perception in aging individuals in order to gain insights into the challenges of listening under noisy conditions.

Informationally rich temporal structure in auditory signals - over a range of time scales from milliseconds to seconds - renders temporal processing central to perception in the auditory domain. We discuss the role of temporal structure in auditory processing, in particular from a perspective relevant for hearing in background noise, and focusing on sensory memory, auditory scene analysis, and speech perception.

Interestingly, these auditory processes, usually studied in an independent manner, show considerable overlap of processing time scales, even though each has its own ‚privileged‘ temporal regimes. By integrating perspectives on temporal structure processing in these three areas of investigation, we aim to highlight similarities typically not recognized.

Keywords

Auditory cortex; oscillations; MMN; prediction; time

Corresponding author: Johanna Maria Rimmele, PhD, Department of Neurophysiology and Pathophysiology, University Medical Center Hamburg-Eppendorf, Martinistraβe 52, D - 20246 Hamburg, phone: +49 (40) 7410 - 0, j.rimmele@uke.de.

[1]In this review we discuss evidence for age-related deficits in temporal structure processing at relatively slow (~100 ms to several seconds) time scales. An important characteristic of any extended acoustic signal is that it evolves over time, rendering auditory processing inherently temporal. Auditory perception requires the recognition of temporal structure, that is, a patterned organization of the stimulation over time. By containing repeating patterns, a temporally structured stimulus can be, in a broader sense, 'regular,' in contrast, to a temporally unstructured stimulus that contains no systematic regularities repeated over time. Temporal information occurs at multiple time scales, associated with different perceptual phenomena (Fig. 1). For example, in the case of speech, the most compelling temporal regularity is seen in the modulation spectrum, which demonstrates (across languages) an amplitude modulation with a peak at ~5 Hz, corresponding to the mean syllabic rate. Phonemic (sub-syllabic) information is, by necessity, associated with a higher modulation rate; analogously, intonation contours at the phrasal or sentence level reflect slow regularities (typically variation of the fundamental frequency) over hundreds to thousands of milliseconds. Other signals, too, contain more 'local,' short time scale variation and more 'global,' longer scale temporal structure. Age-related deficits in temporal processing are certainly not constrained to relatively slow time scales (100 ms to several seconds). However, this review will focus on these longer time scales as they are fundamentally involved in processing temporal structure at the level of auditory cortex, including speech perception, and auditory scene analysis (ASA) (Fig. 2). These time scales are relevant for processing speech (Fig. 2b) and other natural sounds (e.g., such as a dog barking) that involve auditory sensory memory (ASM; sounds typically used to research ASM, Fig. 2a). When we listen to one out of multiple simultaneous speakers (or other sound sources), temporal characteristics of separate signals overlap (Fig. 2c). We discuss research that suggests temporal characteristics of the sound input, can aid the segregation of multiple speakers, or the process of auditory stream segregation.

We focus on age-related deficits in speech perception and ASA that may be arise from deficits of temporal structure processing in ASM.

Temporal structure processing in auditory sensory memory

ASM, the retention of sound information after the physical input has ceased, plays an important role in the ability to integrate successive sounds over time into meaningful auditory events (Cowan, 1995; Neisser, 1967). The transient storage of auditory information can last up to 30 seconds (Sams et al., 1993; Winkler et al., 2002), depending upon how ASM is defined. However, theories differ in how they propose that temporal information is accessed at different time scales. A classical approach to the study of ASM (Cowan, 1988, 1984; Massaro, 1972; Näätänen, 1992, 1990; Näätänen et al., 1978) proposes two different types of sensory memory, one for short storage (called "short auditory store/storage") and another for longer storage (called "long auditory store/storage") (Cowan, 1988, 1984; Massaro, 1972). The short auditory store refers to processes occurring at a faster time scale

[1]**Abbreviations:** ASA, auditory scene analysis; ASM, auditory sensory memory; ECoG, electrocorticography; EEG, electroencephalography; ERP, event-related potential; fMRI, functional magnetic resonance imaging; ITPC, inter-trial phase coherence; MEG, magnetoencephalography; MMN, mismatch negativity; SPIN test, Speech Perception in Noise test; STG, superior temporal gyrus; TFS, temporal fine structure;

(up to several hundreds of milliseconds), and the long auditory store to those at a slower time scale (up to several seconds). The two sensory stores have been related to different phenomena that suggest that the temporal processing mechanims in both stores differ (Cowan, 1988; Massaro, 1972). For example, the short auditory store is associated with perceptual integration, including phenomena such as loudness summation (Zwislocki, 1969) and backward masking (Massaro, 1975). Whereas the short auditory store holds representations of sound features that are integrated over time, the long auditory store is argued to represent information about the temporal order of sound segments. In an extension of the classical approach, a „predictive processing" approach of ASM provides a description of how temporal structure is accessed (Schröger, 1997; Schröger et al., 2013; Winkler and Czigler, 2012; Winkler, 2007; Winkler et al., 2009). Some form of "chunking" of single element representations that keeps the temporal order of elements, as already proposed by the classical approach, is now explained in more detail.[2] The predictive approach emphazises the role of regularity extraction and predictive processing in ASM ("regularity-violation" explanation). Regularities are extracted from the auditory stimulation and represented in ASM based on representations of the relations of successive sounds (Schröger, 1997). Regularity representations are used to predict the incoming auditory stimulation. Incoming stimulations that match the prediction are integrated into a predictive model of a tone sequence or an otherwise temporally evolving auditory object (by changing the predictive confidence; Winkler et al., 2009)[3]. In the case of a deviation from the prediction, a „mismatch response" is elicited that indicates potential changes in an object (possibly it is elicited when the weight of the corresponding regularity representation decreases; Winkler et al., 2009). Winkler and Czigler (2012) related predictive processing in ASM to predictive coding theory (e.g., Friston, 2005), proposing that it occurs at an intermediate level of a perceptual processing hierarchy (see also: Garrido et al., 2009; Wacongne et al., 2012).

Temporal processing that relies on ASM can be tested with electroencephalographam (EEG) or magnetoencephalographam (MEG) recordings of the „mismatch response" elicited by regularity violations. The mismatch negativity component (MMN) of the ERP is computed as the difference between ERPs elicited to regularity-confirming "standard" and regularity-violating „deviant" stimuli (deviant-minus-standard ERP; Fig. 3; Note: Specific designs allow to measure the "true MMN", that is, to control for physical stimulus differences and differences in refractoriness; for review: Kujala et al., 2007; Näätänen et al., 2011b; Schröger, 2005, 1997; Schröger et al., 2013). The predictive approach suggests that MMN elicitation depends on both, the representation of single sound elements at a fast time scale (corresponding to the short auditory store of the classical approach) and regularity representations, that is temporal structure processing on a slower time scale (corresponding

[2]Note that Schröger and colleagues avoid the term „auditory sensory memory". Instead they propose a so-called „Auditory Event Representation System (AERS)" and use the term "auditory stimulus event representations" that corresponds to the classical term „auditory sensory memory representation" (Schröger et al., 2013).

[3]The term "auditory object" is controversially discussed (Kubovy and van Valkenburg, 2001; Schröger et al., 2013; Shinn-Cunningham, 2008; Winkler et al., 2009). Auditory objects can be defined spatially, refering to the discrete physical source of sounds (Shinn-Cunningham, 2008). In this review we will refer to a more basic definition of "perceptual objects". Schröger and colleagues define a perceptual auditory object as a sound pattern, which is separable from other sounds, binds events over time, integrates different features, generalizes across different instances, and extrapolates to not represented parts belonging to this perceptual object (see also: (Winkler et al., 2009).

to the long auditory store). Evidence for the latter has been provided by studies that show MMN to deviations of more complex tone patterns, such as deviations of discrete tone patterns (Schröger et al., 1992; Sussman et al., 1999, 1998), continuous tone patterns (Alain et al., 1994; Schröger, 1994; Tervaniemi et al., 1994; Winkler and Schröger, 1995), or abstract rules (Bendixen and Schröger, 2008; Saarinen et al., 1992; for review: Paavilainen, 2013). It is worth noting that there is an additional manner in which regularity extraction can be studied, by using “change responses” that do not requite the canonical mismatch design. Change detection is contingent on detecting a statistical regularity, in spectrotemporal properties of an ongoing stimulus. This approach has been effective at identifying what sorts of changes the auditory system is sensitive to when an MMN design is not fortuitous (see, e.g. Chait et al., 2008, 2007).

The role of temporal structure in auditory scene analysis

The ability to differentiate between talkers in multi-talker situations, in addition to the “within-stream” integration of successive elements in ASM, requires processes that segregate the sound mixture into discrete streams (Sussman, 2005). An important question, particularly in light of age-related difficulties in multi-talker or noisy listening situations, is how within-stream integration in ASM influences stream segregation.

The segregation of relevant sounds from background noise can be based on concurrent sound segregation, using cues such as harmonicity and common onset, as well as on sequential segregation, which may be based on spectral separation of sequential sounds (or timbre differences, spatial location, and stimulus presentation rate) (Bregman, 1990; Moore and Gockel, 2002, 2012). It has been suggested that the within-stream integration of sequential sound elements into perceptual units in ASM takes place in the context of already formed streams (Sussman et al., 1999; Sussman, 2005; Winkler et al., 2003; Sussman et al., 1998, 2013; Sussman, 2007; Yabe et al., 2001). In an MMN study, Sussman (2005) investigated the relationship between the segregation and integration of sounds. They found that segregation of sounds to distinct streams, based on spectral separation of sequential sounds, occurs prior to integration of within-stream sound elements in ASM. Furthermore, temporal integration processes in ASM occured in several sound streams simultaneously.

However, several theories propose that the analysis of temporal structure is more important to stream segregation (Elhilali and Shamma, 2008; Elhilali et al., 2009; Micheyl et al., 2013; Schröger et al., 2013; Shamma et al., 2011; Teki et al., 2013, 2011; Winkler, 2010, 2007; Winkler et al., 2012; for an early approach see: Jones and Boltz, 1989; Jones, 1976). According to the predictive processing approach of ASM, the impact of temporal structure on stream segregation can be explained by a more dynamic interplay between ASM processing and ASA (Schröger et al., 2013; Winkler, 2010, 2007; Winkler et al., 2012). Claims about a dynamic interplay have been supported by accumulating evidence from behavioral studies that found regular patterns can serve as dynamic cues in auditory scene analysis (Andreou et al., 2011; Bendixen et al., 2014a, 2013, 2012, 2010; Devergie et al., 2010; Rimmele et al., 2012a; for review: Bendixen, 2014). While two studies found that regular patterns only stabilize already-formed streams (Bendixen et al., 2013, 2010), findings from a recent study suggest that regular patterns can affect the initial grouping of simultaneously occurring sounds, and thus can serve as a primary cue for stream segregation (Bendixen et al., 2014a). Support for the assumption that effects of regular patterns on ASA are related to temporal structure processing in ASM comes from a recent MMN study (Bendixen et al., 2012). Sound sequences overlapped in frequency range and other acoustic parameters and could only be distinguished based on the different underlying regularities. MMN elicitation suggested that ASM processing can underlie sound segregation when no other cues are available.

Elhilali and colleagues (Elhilali and Shamma, 2008; Elhilali et al., 2009a; Micheyl et al., 2013; Shamma et al., 2011; Teki et al., 2013, 2011) also suggest an important role for one form of temporal structure processing in stream segregation. They describe a computational model wherein, temporal coherence between different elements in an auditory scene as primary criterion impacts stream segregation. Sounds that largely differed in spectral characteristics were perceived as segregated when they were consecutively presented (not temporally coherent), but these sounds were perceived as integrated when presented in a temporally coherent manner (Elhilali et al., 2009a). Neurophysiologically, the impact of temporal coherence on auditory scene analysis might be based on phase-locking of cortical responses to temporally coherent stimuli (Elhilali et al., 2009b; for review: Shamma et al., 2011). While Shamma and colleagues (Shamma et al., 2011) proposed that phase-locking to temporally coherent stimuli is based on stimulus features at a rather slow time scale similar to the syllabic time scale in speech (< 20 Hz), effects of temporal coherence have also been shown with higher frequency stimuli (Teki et al., 2011). Interestingly, phase-locking at a syllabic time scale has been proposed to be important in speech processing. A different neurophysiological approach suggests phase-locking of neural ocillations to be involved in temporal segmentation of speech. Neural oscillatory activity differentially phase-locked to simultaneously presented, temporally incoherent speech, of different speakers (Ding and Simon, 2012; Mesgarani and Chang, 2012; Zion Golumbic et al., 2012).

In summary, several theories emphazise the role of some form of temporal structure in ASA. Although, the temporal coherence and the predictive processing approach make similar claims, they describe different underlying neurophysiological mechanisms.

Age-related deficits in sensory memory processing

Evidence for age-related deficits in central auditory processing comes from behavioral and electrophysiological (e.g., MMN) studies that controlled for peripheral hearing loss, by either excluding participants with mild to moderate hearing loss, by comparing groups with different degrees of hearing loss, or by presenting sounds above individual hearing thresholds. Numerous studies provide evidence for age-related deficits in temporal processing at the level of ASM (for review: Cheng et al., 2013; Näätänen et al., 2011a; Pekkonen, 2000). Older adults show impaired ASM processing for detection of changes in the durations of tones (Karayanidis et al., 1995; Kiang et al., 2009; Pekkonen et al., 1996; Schroeder et al., 1995; Woods, 1992); interstimulus-intervals (Kisley et al., 2005); and short gaps in tones (Alain et al., 2004; Bertoli et al., 2002). These deficits are generally manifest in reduced MMN amplitude in older compared to younger adults. Concordantly, older adults show performance decreases in tasks that require temporally accurate processing, such as an increased gap detection threshold (Fitzgibbons and Gordon-Salant, 1994; Fitzgibbons et al., 2007; Grose et al., 2006; Schneider and Hamstra, 1999; Schneider et al., 1994; Snell et al., 2002; Strouse et al., 1998). Deficits in gap detection tasks might be related to ASM decline, however, decreased temporal acuity might also partly arise from deficits in more basic processing, such as a loss in neural synchrony at the level of the auditory brainstem (Backoff and Caspary, 1994; for a similar argument, see Pichora-Fuller et al., 2007). Studies that found reduced MMN to deviations of two- and three- tone patterns provide more direct evidence for age-related deficits in temporal structure processing in ASM (Alain and Woods, 1999; Rimmele et al., 2012b, 2014). Impaired temporal structure processing in ASM might be related to performance decline in tasks that require temporal processing of sequences, such as discriminating the temporal order of sounds (Fitzgibbons and Gordon-Salant, 1998; Fitzgibbons et al., 2006; Gordon-Salant and Fitzgibbons, 1999; Shrivastav et al., 2008; Trainor and Trehub, 1989).

An advantage of MMN studies is that MMN is elicited even when participants direct their attention away from sounds (or otherwise do not attend to the sounds), allowing the ability to assess the more „automatic“ temporal structure processing in ASM (for review: Näätänen et al., 2010). Thus, effects of age-related decline due to higher cognitive processes, such as explicit memory or attention related deficits, can be minimized by not requiring a task to be performed with the sounds.

Temporal structure processing in speech

It is important to understand how deficits in temporal structure processing impact older adults in more natural listening situations, which often include speech processing. Speech contains acoustic temporal structure at multiple time scales, including the featural and phonemic (10s of msec), syllabic (hundreds of msec), and intonational (~1000 msec) time scales (Giraud and Poeppel, 2012; Stevens, 2002; Zion Golumbic et al., 2012). Speech is initially processed by generic auditory processing mechanisms, including all those early operations that feed into ASM (Davis and Johnsrude, 2007; Holt and Lotto, 2010, 2008; Macken et al., 1999; Näätänen et al., 2011b; Poeppel et al., 2008; Steinschneider et al., 2013). Concordantly, MMN studies found that the temporal structure of speech is processed in ASM at several time scales (Bendixen et al., 2014b; Peter et al., 2012; for review: Näätänen, 2001; Pulvermuller and Shtyrov, 2006). For example, Bendixen and colleagues showed that the temporal structure of syllables was represented in ASM. The omission of the final segment of a monosyllabic word elicited a larger MMN, when participants had a specific expectation about the combination of segments (i.e., either the word „Lachs“ or the word „Latz“ was presented as standard, “la“ was presented as deviant), compared to when they had no specific expectation (i.e., the word „Lachs“ and „Latz“ were both presented as standard).

Despite many MMN studies on speech processing, relatively few studies have used MMN to test temporal structure processing in continuous speech. Temporal structure processing of continuous speech has been more typically investigated by measuring neural oscillations with EEG, MEG or electrocorticography (ECoG) (for review: Giraud and Poeppel, 2012; Giraud et al., 2007a; Zion Golumbic et al., 2012). Giraud and Poeppel (2012) propose that phase modulations of ongoing oscillations provide a mechanism for simultaneous temporal integration of speech signals at the phonemic time scale (integration of information within short time windows; oscillations in the low gamma frequency range; say ~30 Hz and below), the syllabic time scale (integration of information within an about 125–250 ms window; oscillations in the theta frequency range; 4–8 Hz) and at the intonational time scale (integration of information within an about 500–1000 ms window; oscillations in the delta frequency range; 1–3 Hz) (Fig. 4; see also: Doelling et al., 2013; Giraud et al., 2007; Howard and Poeppel, 2010; Luo and Poeppel, 2007; Peelle et al., 2012). The inter-trial phase coherence (ITPC) of neural theta oscillations can be used to measure temporal structure processing at the slower syllabic time scales. ITPC indicates the precision of phase-locking between the stimulus structure and the neural processes (the ITPC is higher for several „same sentence“ stimuli compared to several „different sentence“ stimuli; Giraud and Poeppel, 2012).

Functionally, phase-locking at the syllabic time scale has been related to a predictive timing mechanism, probably based on neural entrainment by (quasi-) rhythmic stimulation, (Arnal and Giraud, 2012). Temporal information at the syllabic time scale seems to be crucial for speech intelligibility (Doelling et al., 2013; Ghitza et al., 2013; Giraud and Poeppel, 2012; Shannon et al., 1995; Zion Golumbic et al., 2012). Luo and Poeppel (2007) found a higher ITPC for clear speech compared to degraded speech (see also: Peelle et al., 2012). Importantly, speech tracking also occurs for unintelligible, time-inverted speech, and thus seems to indicate spectro-temporal structure processing or temporal segmentation and not intelligibility processing per se (Cogan and Poeppel, 2011; Howard and Poeppel, 2010; Pena and Melloni, 2012).

The „oscillatory speech-tracking approach“ provides indicators to test speech structure processing at the level of auditory cortex at similar time scales to those underlying ASM processing. In comparison to the MMN component, these indicators reflect temporal structure tracking more directly. Although, there is some evidence that oscillatory activity in the delta and theta range might reflect similar processes than MMN (Karakaş et al., 2000), overall it is not clear how oscillatory speech-tracking relates to temporal structure processing in ASM.

Age-related deficits in speech processing

Several studies suggest relatively preserved speech understanding, when speech is presented in silence (Frisina and Frisina, 1997; Larsby et al., 2005). However, older adults show deficits in specific aspects of speech processing. For example, performance is reduced in tasks that require rapid speech processing (Wingfield et al., 2003), or temporally acurate speech processing (Tremblay et al., 2003). Importantly, several studies reported age-related performance decline in determining the temporal order of speech (e.g., syllables; Fogerty et al., 2012, 2010; Humes et al., 2013). Although Bellis et al. (2000) did not find age-related changes in MMN elicited by deviant syllables, hemispheric asymmetry of speech processing was reduced in older adults compared to younger adults at earlier processing levels (P1, N1 ERP; Note: MMN was only recorded at temporal electrode sides; Bellis et al., 2000). In a functional magnetic resonance imaging study (fMRI; Peelle et al., 2010), older adults showed activation of a core sentence processing network (including left inferior and middle frontal gyri, left inferior parietal cortex, left middle temporal gyrus) when listening to complex sentences, whereas activity in inferior frontal areas was reduced compared to younger adults. Additionally, older adults showed increased activation of areas outside of the core sentence processing network. Findings have been related to compensatory activity, which might explain older adults' relatively preserved speech processing in easy listening situations, while a decreased activity in areas specialized in speech processing might results in pronounced deficits when listening becomes more challenging.

In summary, in spite of some evidence for age-related deficits in temporal structure processing in speech, mainly coming from behavioral studies, there are few neurophysiological studies that have directly investigated aging effects on temporal structure processing of speech at the auditory cortex level (both using the MMN and the oscillatory speech-tracking approach).

Listening in background noise - Aging and auditory scene analysis

An important question is how deficits in temporal structure processing in ASM impact older adults' abilities to listen in background noise. Because temporal structure has been proposed to serve as a dynamic cue, processing deficits should particularly affect sequential stream segregation. Although, older adults show deficits in concurrent stream segregation (Alain et al., 2001; Snyder and Alain, 2007, 2005), behavioral and electrophysiological findings are more controversial for sequential stream segregation because many studies do not find deficits (Alain et al., 1996; Snyder and Alain, 2007; Trainor and Trehub, 1989; for review: Alain et al., 2006), while some studies do (Grimault et al., 2001). Grimault and colleagues show reduced streaming of complex tones in older adults with and without hearing loss comapred to younger adults. They attributed their findings to deficits in peripheral frequency selectivity. However, studies on sequential streaming tested older adults' ability to use simple frequency cues (e.g., Δf between two tone streams; Alain et al., 1996; Snyder and Alain, 2007), or simple temporal cues (e.g., presentation rate; Trainor and Trehub, 1989). A recent study found an age-related deficit in the use of more dynamic cues for ASA, that is, older adults were less able to use regular patterns to stabilize stream segregation (Rimmele et al., 2012a). Aging effects were only present when the regular patterns were ignored (i.e., presented in the unattended stream), possibly, indicating older adults ability to compensate deficits when the sounds were attended (for a similar argumentation: Alain et al., 2004). However, findings require further clarification, as deficits were only observed for certain types of tone patterns, and neurophysiological mechanisms were not accessed in that study. Thus, so far the relationship to temporal structure processing in ASM is still speculative.

Speech understanding difficulties get more pronounced in challenging listening situations, particularly when speech is presented with background noise (Dubno et al., 1984; Frisina and Frisina, 1997; Helfer and Freyman, 2008; Larsby et al., 2005; Rajan and Cainer, 2008; Stuart and Phillips, 1996). Interestingly, age-related deficits in speech perception in background noise have been reported to be strongest when background noise contains temporal segmentation similar to that in speech (Larsby et al., 2008, 2005; Rajan and Cainer, 2008). Rajan and Cainer tested age effects in a „speech perception in noise" test

(SPIN; Kalikow et al., 1977). Participants had to identify three keywords of a sentence in correct order, while it was presented with different types of background noise. Speech reception thresholds (signal-to-noise ratio for 50% detection of sentences in noise) were higher for older adults compared to younger adults, only when speech was presented with mulit-talker babble noise in the background, but not with speech weighted noise (i.e., noise with a long-term average spectrum equal to that of the sentences). As both types of noise cause masking effects due to overlap of competing signals in time and frequency ("energetic masking"), findings suggest that in this study age-related decline was not due to energetic masking. The authors consider that babble noise has several additional effects related to temporal structure and/or higher order processing of speech that can explain the age-related decline. Although this study could not specify the reason for the effects of babble noise, findings from another study support the assumption that age-related hearing difficulties in background noise are related to temporal structure processing. Larsby and colleagues (2008) found increased age-related deficits not only when speech, but also when unintelligible, reversed speech was presented in the background. A possible explanation of these findings is that effects are due to age-related impairment of temporal structure processing of speech in auditory cortex. An alternative explanation is that older adults deficit in processing the temporal fine structure (TFS) of speech results in a reduced ability to listen in the dips of background noise (e.g., Hopkins and Moore, 2011). Further evidence for age-related impairment of temporal structure processing of speech comes from a study that tested age effects on the ability to use the first-formant (F1) structure of phonemes to integrate streams (Hutka et al., 2013). Stream integration/segregation was accessed by a serial order task (vowel pattern recognition), whereas integrated percepts result in a higher task performance. A regularly changing first-formant structure facilitated stream integration in younger, but not in older adults. Perfomance in conditions with a discontinuous formant structure was similar in both age groups. These findings are in line with what has been shown for regular tone patterns (Rimmele et al., 2012a).

Furthermore, older adults' difficulties, when speech is presented in noise, are not necessarily related to factors of higher cognitive processing, e.g. older adults showed no deficits in using cues such as "speaker familiarity" to enhance speech processing (Johnsrude et al., 2013).

Perspective – Neurophysiological measures

Despite accumulating evidence for a more dynamic interplay of temporal structure processing and stream segregation, and evidence for age-related deficits in temporal structure processing, there is a lack of research that directly investigates these processes. In particular, there is little knowledge about the neurophysiological mechanisms that underlie the difficulties in more challenging listening conditions. Further research is needed to clarify age-related deficits, and better distinguish them from auditory processing deficits at earlier processing stages, as well as deficits in higher cognitive processing. In this review, we introduced two neurophysiological approaches that may be suitable to advance research in this field (MMN, oscillatory speech-tracking approach). Next, we outline recent research in younger adults using the oscillatory speech-tracking approach.

Several studies used this approach to disentangle effects of spectro-temporal structure processing and selective attention in multi-talker environments (Ding and Simon, 2012; Kerlin et al., 2010; Mesgarani and Chang, 2012; Zion Golumbic et al., 2013; for review: Zion Golumbic et al., 2012). Using ECoG in surgical epilepsy patients, Zion Golumbic at al. (2013) found that the interplay between temporal structure processing and selective attention differs at different processing stages. Participants attended one out of two simultaneous presented talkers. Speech-tracking at the syllabic rate (theta) occurred for the attended, as well as the unattended stream „lower" in the auditory processing hierarchy (superior temporal gyrus, STG). In higher processing areas, it resulted in selective tracking of only the attended stream (broader distributed across brain areas). In an EEG study, Kerlin et al. (2010) found that spatial selective attention (indicated by contralateral alpha wave supression) had an early effect on speech processing in noise, while the time course of temporal structure processing (indicated by oscillatory activity in the theta range) peaked later during sentence processing.

In summary, not only the MMN approach (Sussman, 2005; Yabe et al., 2001), but also the oscillatory speech-tracking approach can be used to test temporal structure processing of several simultaneous speech streams and allows one to distinguish between temporal structure processing and selective attention effects on speech processing in noise.

Conclusions

Findings from basic auditory and speech research suggest that age-related deficits in hearing in background noise might be partially due to impaired temporal structure processing at relatively slow time scales at the level of auditory cortex. While, there is evidence for age-related deficits in temporal structure processing in ASM (particularly using non-speech stimuli), further clarification is necessary on how these deficits relate to other findings in speech research, and particularly to older adults performance in noisy listening situations. We advocate for an integrative approach, which investigates temporal structure processing in basic auditory and speech processes in parallel. We consider analyses of MMN and the oscillatory speech-tracking approach as a promising avenue to gaining an increased understanding of the underlying neurophysiological mechanisms.

Acknowledgements

This research was supported by the German Research Foundation SFB936/B2 (J. M. R.), the National Institute of Deafness and Other Communication Disorders (DC004263, E.S.).

References

Alain, C.; Dyson, BJ.; Snyder, JS. Aging and the Perceptual Organization of Sounds: A change of Scene?. In: Conn, PM., editor. Handbook of Models for Human Aging. New York, NY: Academic Press; 2006. p. 759-770.

Alain C, McDonald KL, Ostroff JM, Schneider B. Age-related changes in detecting a mistuned harmonic. J. Acoust. Soc. Am. 2001; 109

Alain C, McDonald KL, Ostroff JM, Schneider B. Aging: A Switch From Automatic to Controlled Processing of Sounds? Psychol. Aging. 2004; 19:125–133. [PubMed: 15065936]

Alain C, Ogawa KH, Woods DL. Aging and the Segregation of Auditory Stimulus Sequences. J. Gerontol. B. Psychol. Sci. Soc. Sci. 1996; 51B:P91–P93. [PubMed: 8785691]

Alain C, Woods DL. Age-related changes in processing auditory stimuli during visual attention: Evidence for deficits in inhibitory control and sensory memory. Psychol. Aging. 1999; 14:507–519. [PubMed: 10509703]

Alain C, Woods DL, Ogawa KH. Brain indices of automatic pattern processing. NeuroReport. 1994; 6:140–144. [PubMed: 7703403]

Andreou L-V, Kashino M, Chait M. The role of temporal regularity in auditory segregation. Hear. Res. 2011; 280:228–235. [PubMed: 21683778]

Arnal LH, Giraud A-L. Cortical oscillations and sensory predictions. Trends Cogn. Sci. 2012; 16:390–398. [PubMed: 22682813]

Backoff PM, Caspary DM. Age-related changes in auditory brainstem responses in fischer 344 rats: effects of rate and intensity. Hear. Res. 1994; 73:163–172. [PubMed: 8188545]

Bellis TJ, Nicol T, Kraus N. Aging Affects Hemispheric Asymmetry in the Neural Representation of Speech Sounds. J. Neurosci. 2000; 20:791–797. [PubMed: 10632608]

Bendixen. Predictability effects in auditory scene analysis: A review. Front. Neurosci. 2014 in press.

Bendixen A, Bőhm TM, Szalárdy O, Mill R, Denham SL, Winkler I. Different roles of similarity and predictability in auditory stream segregation. Learn. Percept. 2013; 5:37–54.

Bendixen A, Denham SL, Gyimesi K, Winkler I. Regular patterns stabilize auditory streams. J. Acoust. Soc. Am. 2010; 128:3658–3666. [PubMed: 21218898]

Bendixen A, Denham SL, Winkler I. Feature predictability flexibly supports auditory stream segregation or integration. Acta Acust. United Acust. 2014a in revision.

Bendixen A, Scharinger M, Strauβ A, Obleser J. Prediction in the service of comprehension: Modulated early brain responses to omitted speech segments. Cortex. 2014b; 53:9–26. [PubMed: 24561233]

Bendixen A, Schröger E. Memory trace formation for abstract auditory features and its consequences in different attentional contexts. Biol. Psychol. 2008; 78:231–241. [PubMed: 18439740]

Bendixen A, Schröger E, Ritter W, Winkler I. Regularityextractionfromnonadjacentsounds. Front. Psychol. 2012; 3:3–12. doi: [PubMed: 22291676]

Bertoli S, Smurzynski J, Probst R. Temporal resolution in young and elderly subjects as measured by mismatch negativity and a psychoacoustic gap detection task. Clin. Neurophysiol. 2002; 113:396–406. [PubMed: 11897540]

Bregman, A. Auditory scene analysis: the perceptual organization of sound. Cambridge, MA: Cambridge. The MIT Press; 1990.

Chait M, Poeppel D, de Cheveigne A, Simon JZ. Processing asymmetry of transitions between order and disorder in human auditory cortex. J. Neurosci. Off. J. Soc. Neurosci. 2007; 27:5207–5214.

Chait M, Poeppel D, Simon JZ. Auditory temporal edge detection in human auditory cortex. Brain Res. 2008; 1213:78–90. [PubMed: 18455707]

Cheng C-H, Hsu W-Y, Lin Y-Y. Effects of physiological aging on mismatch negativity: A meta-analysis. Int. J. Psychophysiol. 2013; 90:165–171. [PubMed: 23831479]

Cogan GB, Poeppel D. A mutual information analysis of neural coding of speech by low-frequency MEG phase information. J. Neurophysiol. 2011; 106:554–563. [PubMed: 21562190]

Cowan N. On short and long auditory stores. Psychol. Bull. 1984; 96:341–370. [PubMed: 6385047]

Cowan N. Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. Psychol Bull. 1988; 104:163–191. [PubMed: 3054993]

Cowan, N. Attention and Memory. Oxford University Press; 1995.

Davis MH, Johnsrude IS. Hearing speech sounds: Top-down influences on the interface between audition and speech perception. Audit. Cortex 2006 - List. Brain. 2007; 229:132–147.

Devergie A, Grimault N, Tillmann B, Berthommier F. Effect of rhythmic attention on the segregation of interleaved melodies. J. Acoust. Soc. Am. 2010; 128:EL1–EL7. [PubMed: 20649182]

Ding N, Simon JZ. Neural coding of continuous speech in auditory cortex during monaural and dichotic listening. J. Neurophysiol. 2012; 107:78–89. [PubMed: 21975452]

Doelling, KB.; Arnal, LH.; Ghitza, O.; Poeppel, D. Acoustic landmarks drive delta–theta oscillations to enable speech comprehension by facilitating perceptual parsing. NeuroImage in press; 2013.

Dubno JR, Dirks DD, Morgan DE. Effects of age and mild hearing loss on speech recognition in noise. J. Acoust. Soc. Am. 1984; 76:87–96. [PubMed: 6747116]

Elhilali M, Ma L, Micheyl C, Oxenham AJ, Shamma SA. Temporal Coherence in the Perceptual Organization and Cortical Representation of Auditory Scenes. Neuron. 2009a; 61:317–329. [PubMed: 19186172]

Elhilali M, Shamma SA. A cocktail party with a cortical twist: How cortical mechanisms contribute to sound segregation. J. Acoust. Soc. Am. 2008; 124:3751–3771. [PubMed: 19206802]

Elhilali M, Xiang J, Shamma SA, Simon JZ. Interaction between Attention and Bottom-Up Saliency Mediates the Representation of Foreground and Background in an Auditory Scene. PLoS Biol. 2009b; 7:e1000129. [PubMed: 19529760]

Fitzgibbons PJ, Gordon-Salant S. Age Effects on Measures of Auditory Duration Discrimination. J. Speech Hear. Res. 1994; 37:662–670. [PubMed: 8084196]

Fitzgibbons PJ, Gordon-Salant S. Auditory Temporal Order Perception in Younger and Older Adults. J Speech Lang Hear Res. 1998; 41:1052–1060. [PubMed: 9771628]

Fitzgibbons PJ, Gordon-Salant S, Barrett J. Age-related differences in discrimination of an interval separating onsets of successive tone bursts as a function of interval duration. J. Acoust. Soc. Am. 2007; 122:458–466. [PubMed: 17614503]

Fitzgibbons PJ, Gordon-Salant S, Friedman SA. Effects of age and sequence presentation rate on temporal order recognition. J. Acoust. Soc. Am. 2006; 120:991–999. [PubMed: 16938986]

Fogerty D, Humes LE, Kewley-Port D. Auditory temporal-order processing of vowel sequences by young and elderly listeners. J. Acoust. Soc. Am. 2010; 127:2509–2520. [PubMed: 20370033]

Fogerty D, Kewley-Port D, Humes LE. Temporal offset judgments for concurrent vowels by young, middle-aged, and older adults. J. Acoust. Soc. Am. 2012; 131:EL499–EL505. [PubMed: 22713028]

Frisina DR, Frisina RD. Speech recognition in noise and presbycusis: relations to possible neural mechanisms. Hear. Res. 1997; 106:95–104. [PubMed: 9112109]

Friston K. A theory of cortical responses. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 2005; 360:815–836. [PubMed: 15937014]

Garrido MI, Kilner JM, Stephan KE, Friston KJ. The mismatch negativity: A review of underlying mechanisms. Clin. Neurophysiol. 2009; 120:453–463. [PubMed: 19181570]

Ghitza O, Giraud A-L, Poeppel D. Neuronal oscillations and speech perception: critical-band temporal envelopes are the essence. Front. Hum. Neurosci. 2013; 6:340. [PubMed: 23316150]

Giraud A-L, Kleinschmidt A, Poeppel D, Lund TE, Frackowiak RSJ, Laufs H. Endogenous Cortical Rhythms Determine Cerebral Specialization for Speech Perception and Production. Neuron. 2007; 56:1127–1134. [PubMed: 18093532]

Giraud A-L, Poeppel D. Cortical oscillations and speech processing: emerging computational principles and operations. Nat. Neurosci. 2012; 15:511–517. [PubMed: 22426255]

Gordon-Salant S, Fitzgibbons PJ. Profile of Auditory Temporal Processing in Older Listeners. J. Speech Lang. Hear. Res. 1999; 42:300–311. [PubMed: 10229448]

Grimault N, Micheyl C, Carlyon RP, Arthaud P, Collet L. Perceptual auditory stream segregation of sequences of complex sounds in subjects with normal and impaired hearing. Br. J. Audiol. 2001; 35:173–182. [PubMed: 11548044]

Grose JH, Hall JW III, Buss E. Temporal processing deficits in the pre-senescent auditory system. J. Acoust. Soc. Am. 2006; 119:2305–2315. [PubMed: 16642844]

Helfer KS, Freyman RL. Aging and speech-on-speech masking. Ear Hear. 2008; 29

Hirsch IJ, Sherrick CEJ. Perceived order in different sense modalities. J. Exp. Psychol. 1961; 62:423–432. [PubMed: 13907740]

Holt L, Lotto A. Speech perception as categorization. Atten. Percept. Psychophys. 2010; 72:1218–1227. [PubMed: 20601702]

Holt LL, Lotto AJ. Speech Perception Within an Auditory Cognitive Science Framework. Curr. Dir. Psychol. Sci. 2008; 17:42–46. [PubMed: 19060961]

Hopkins K, Moore BCJ. The effects of age and cochlear hearing loss on temporal fine structure sensitivity, frequency selectivity, and speech reception in noise. J. Acoust. Soc. Am. 2011; 130:334–349. [PubMed: 21786903]

Howard MF, Poeppel D. Discrimination of Speech Stimuli Based on Neuronal Response Phase Patterns Depends on Acoustics But Not Comprehension. J. Neurophysiol. 2010; 104:2500–2511. [PubMed: 20484530]

Humes L, Busey T, Craig J, Kewley-Port D. Are age-related changes in cognitive function driven by age-related changes in sensory processing? Atten. Percept. Psychophys. 2013; 75:508–524. [PubMed: 23254452]

Hutka SA, Alain C, Binns MA, Bidelman GM. Age-related differences in the sequential organization of speech sounds. J. Acoust. Soc. Am. 2013; 133:4177–4187. [PubMed: 23742369]

Johnsrude IS, Mackey A, Hakyemez H, Alexander E, Trang HP, Carlyon RP. Swinging at a Cocktail Party: Voice Familiarity Aids Speech Perception in the Presence of a Competing Voice. Psychol. Sci. 2013

Jones MR. Time, our lost dimension: toward a new theory of perception, attention, and memory. Psychol. Rev. 1976; 83:323–355. [PubMed: 794904]

Jones MR, Boltz M. Dynamic attending and responses to time. Psychol. Rev. 1989; 96:459–491. [PubMed: 2756068]

Kalikow DN, Stevens KN, Elliott LL. Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. J. Acoust. Soc. Am. 1977; 61:1337–1351. [PubMed: 881487]

Karakaş S, Erzengin ÖU, Başar E. The genesis of human event-related responses explained through the theory of oscillatory neural assemblies. Neurosci. Lett. 2000; 285:45–48. [PubMed: 10788704]

Karayanidis F, Andrews S, Ward P, Michie PT. ERP indices of auditory selective attention in aging and Parkinson’s disease. Psychophysiology. 1995; 32:335–350. [PubMed: 7652110]

Kerlin JR, Shahin AJ, Miller LM. Attentional Gain Control of Ongoing Cortical Speech Representations in a “Cocktail Party”. J. Neurosci. 2010; 30:620–628. [PubMed: 20071526]

Kiang M, Braff DL, Sprock J, Light GA. The relationship between preattentive sensory processing deficits and age in schizophrenia patients. Clin. Neurophysiol. 2009; 120:1949–1957. [PubMed: 19786365]

Kisley MA, Davalos DB, Engleman LL, Guinther PM, Davis HP. Age-related change in neural processing of time-dependent stimulus features. Cogn. Brain Res. 2005; 25:913–925.

Kubovy M, van Valkenburg D. Auditory and visual objects. Objects Atten. 2001; 80:97–126.

Kujala T, Tervaniemi M, Schröger E. The mismatch negativity in cognitive and clinical neuroscience: Theoretical and methodological considerations. Biol. Psychol. 2007; 74:1–19. [PubMed: 16844278]

Lalor EC, Foxe JJ. Neural responses to uninterrupted natural speech can be extracted with precise temporal resolution. Eur. J. Neurosci. 2010; 31:189–193. [PubMed: 20092565]

Larsby B, Hällgren M, Lyxell B. The interference of different background noises on speech processing in elderly hearing impaired subjects. Int. J. Audiol. 2008; 47:S83–S90. [PubMed: 19012115]

Larsby B, Hällgren M, Lyxell B, Arlinger S. Cognitive performance and perceived effort in speech processing tasks: Effects of different noise backgrounds in normal-hearing and hearing-impaired subjects. Int. J. Audiol. 2005; 44:131–143. [PubMed: 15916114]

Luo H, Poeppel D. Phase Patterns of Neuronal Responses Reliably Discriminate Speech in Human Auditory Cortex. Neuron. 2007; 54:1001–1010. [PubMed: 17582338]

Macken W, Tremblay S, Alford D, Jones D. Attentional Selectivity in Short-term Memory: Similarity of Process, Not Similarity of Content, Determines Disruption. Int. J. Psychol. 1999; 34:322–327.

Massaro DW. Preperceptual images, processing time and perceptual units in auditory perception. Psychol. Rev. 1972; 79:124–145. [PubMed: 5024158]

Massaro, DW. Experimental psychology and information processing. Chicago: Rand McNally College Publishing Company; 1975.

Mesgarani, N.; Chang, EF. Selective cortical representation of attended speaker in multi-talker speech perception. Nature advance online publication; 2012.

Micheyl C, Hanson C, Demany L, Shamma S, Oxenham AJ. Auditory stream segregation for alternating and synchronous tones. J. Exp. Psychol. Hum. Percept. Perform. 2013; 39:1568–1580. [PubMed: 23544676]

Moore BCJ, Gockel H. Factors Influencing Sequential Stream Segregation. Acta Acust. United Acust. 2002; 88:320–333.

Moore BCJ, Gockel HE. Properties of auditory stream formation. Philos. Trans. R. Soc. B Biol. Sci. 2012; 367:919–931.

Näätänen R. The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. Behav. Brain Sci. 1990; 13:201–288.

Näätänen, R. Attention and brain function. Hillsdale, NJ: Erlbaum; 1992.

Näätänen R. The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). Psychophysiology. 2001; 38:1–21. [PubMed: 11321610]

Näätänen R, Astikainen P, Ruusuvirta T, Huotilainen M. Automatic auditory intelligence: An expression of the sensory-cognitive core of cognitive processes. Brain Res. Rev. 2010; 64:123–136. [PubMed: 20298716]

Näätänen R, Gaillard AWK, Mäntysalo S. Early selective-attention effect on evoked potential reinterpreted. Acta Psychol. (Amst.). 1978; 42:313–329. [PubMed: 685709]

Näätänen R, Kujala T, Kreegipuu K, Carlson S, Escera C, Baldeweg T, Ponton C. The mismatch negativity: an index of cognitive decline in neuropsychiatric and neurological diseases and in ageing. Brain. 2011a; 134:1–19. [PubMed: 21186261]

Näätänen R, Kujala T, Winkler I. Auditory processing that leads to conscious perception: A unique window to central auditory processing opened by the mismatch negativity and related responses. Psychophysiology. 2011b; 48:4–22. [PubMed: 20880261]

Neisser U. Cognitive psychology. Appleton-Century-Crofts. 1967

Paavilainen P. The mismatch-negativity (MMN) component of the auditory eventrelated potential to violations of abstract regularities: A review. Int. J. Psychophysiol. 2013; 88:109–123. [PubMed: 23542165]

Peelle JE, Gross J, Davis MH. Phase-Locked Responses to Speech in Human Auditory Cortex are Enhanced During Comprehension. Cereb. Cortex. 2012

Peelle JE, Troiani V, Wingfield A, Grossman M. Neural Processing during Older Adults' Comprehension of Spoken Sentences: Age Differences in Resource Allocation and Connectivity. Cereb. Cortex. 2010; 20:773–782. [PubMed: 19666829]

Pekkonen E. Mismatch Negativity in Aging and in Alzheimer's and Parkinson's Diseases. Audiol Neurootol. 2000; 5:216–224. [PubMed: 10859416]

Pekkonen E, Rinne T, Reinikainen K, Kujala T, Alho K, Naatanen R. Aging Effects on Auditory Processing: An Event-Related Potential Study. Exp. Aging Res. Int. J. Devoted Sci. Study Aging Process. 1996; 22:171–184.

Pena M, Melloni L. Brain oscillations during spoken sentence processing. J. Cogn. Neurosci. 2012; 24:1149–1164. [PubMed: 21981666]

Peter V, Mcarthur G, Thompson WF. Discrimination of stress in speech and music: A mismatch negativity (MMN) study. Psychophysiology. 2012; 49:1590–1600. [PubMed: 23066846]

Pichora-Fuller MK, Schneider BA, MacDonald E, Pass HE, Brown S. Temporal jitter disrupts speech intelligibility: A simulation of auditory aging. Hear. Res. 2007; 223:114–121. [PubMed: 17157462]

Poeppel D, Idsardi WJ, van Wassenhove V. Speech perception at the interface of neurobiology and linguistics. Philos. Trans. R. Soc. B Biol. Sci. 2008; 363:1071–1086.

Pulvermuller F, Shtyrov Y. Language outside the focus of attention: the mismatch negativity as a tool for studying higher cognitive processes. Prog. Neurobiol. 2006; 79

Rajan R, Cainer KE. Ageing without hearing loss or cognitive impairment causes a decrease in speech intelligibility only in informational maskers. Neuroscience. 2008; 154:784–795. [PubMed: 18485606]

Rimmele J, Schröger E, Bendixen A. Age-related changes in the use of regular patterns for auditory scene analysis. Hear. Res. 2012a; 289:98–107. [PubMed: 22543088]

Rimmele J, Sussman ES, Keitel K, Jacobsen T, Schröger E. Electrophysiological Evidence for Age Effects on Sensory Memory Processing of Tonal Patterns. Psychol. Aging. 2012b; 27:384–398. [PubMed: 21823798]

Rimmele JM, Sussman E, Jacobsen T, Schröger E. Consequences of age-related changes in tone sequence processing in sensory memory. Unpubl. Manuscr. 2014

Saarinen J, Paavilainen P, Schröger E, Tervaniemi M, Näätänenen R. Representation of abstract attributes of auditory stimuli in the human brain. NeuroReport. 1992; 3

Sams M, Hari R, Rif J, Knuutila J. The human auditory sensory memory trace persists about 10 sec: Neuromagnetic evidence. J Cogn. Neurosci. 1993; 5:363–370. [PubMed: 23972223]

Schneider BA, Hamstra SJ. Gap detection thresholds as a function of tonal duration for younger and older listeners. J. Acoust. Soc. Am. 1999; 106:371–380. [PubMed: 10420628]

Schneider BA, Pichora-Fuller MK, Kowalchuk D, Lamb M. Gap detection and the precedence effect in young and old adults. J. Acoust. Soc. Am. 1994; 95:980–991. [PubMed: 8132912]

Schroeder MM, Ritter W, Vaughan HG. The Mismatch Negativity to Novel Stimuli Reflects Cognitive Declinea. Ann. N. Y. Acad. Sci. 1995; 769:399–401. [PubMed: 8595042]

Schröger E. An event-related potential study of sensory representations of unfamiliar tonal patterns. Psychophysiology. 1994; 31:175–181. [PubMed: 8153253]

Schröger E. On the detection of auditory deviations: A pre-attentive activation model. Psychophysiology. 1997; 34:245–257. [PubMed: 9175439]

Schröger E. The Mismatch Negativity as a Tool to Study Auditory Processing. Acta Acust. United Acust. 2005; 91:490–501.

Schröger E, Bendixen A, Denham S, Mill R, Bőhm T, Winkler I. Predictive Regularity Representations in Violation Detection and Auditory Stream Segregation: From Conceptual to Computational Models. Brain Topogr. 2013:1–13.

Schröger E, Näätänen R, Paavilainen P. Event-related potentials reveal how non-attended complex sound patterns are represented by the human brain. Neurosci. Lett. 1992; 146:183–186. [PubMed: 1491786]

Shamma SA, Elhilali M, Micheyl C. Temporal coherence and attention in auditory scene analysis. Trends Neurosci. 2011; 34:114–123. [PubMed: 21196054]

Shannon RV, Zeng F-G, Kamath V, Wygonski J, Ekelid M. Speech Recognition with Primarily Temporal Cues. Science. 1995; 270:303–304. [PubMed: 7569981]

Shinn-Cunningham BG. Object-based auditory and visual attention. Trends Cogn. Sci. 2008; 12:182–186. [PubMed: 18396091]

Shrivastav MN, Humes LE, Aylsworth L. Temporal order discrimination of tonal sequences by younger and older adults: The role of duration and rate. J. Acoust. Soc. Am. 2008; 124:462–471. [PubMed: 18646990]

Snell KB, Mapes FM, Hickman ED, Frisina DR. Word recognition in competing babble and the effects of age, temporal processing, and absolute sensitivity. J. Acoust. Soc. Am. 2002; 112:720–727. [PubMed: 12186051]

Snyder JS, Alain C. Age-related changes in neural activity associated with concurrent vowel segregation. Cogn. Brain Res. 2005; 24:492–499.

Snyder JS, Alain C. Sequential Auditory Scene Analysis Is Preserved in Normal Aging Adults. Cereb. Cortex. 2007; 17:501–512. [PubMed: 16581981]

Steinschneider M, Nourski KV, Fishman YI. Representation of speech in human auditory cortex: Is it special? Hear. Res. 2013 in press.

Stevens KN. Toward a model for lexical access based on acoustic landmarks and distinctive features. J. Acoust. Soc. Am. 2002; 111:1872–1891. [PubMed: 12002871]

Strouse A, Ashmead DH, Ohde RN, Grantham DW. Temporal processing in the aging auditory system. J. Acoust. Soc. Am. 1998; 104:2385–2399. [PubMed: 10491702]

Stuart A, Phillips DP. Word Recognition in Continuous and Interrupted Broadband Noise by Young Normal-Hearing, Older Normal-Hearing, and Presbyacusic Listeners. Ear Hear. 1996; 17

Sussman E, Ritter W, Vaughan HG. Attention affects the organization of auditory input associated with the mismatch negativity system. Brain Res. 1998; 789:130–138. [PubMed: 9602095]

Sussman E, Ritter W, Vaughan HG. An investigation of the auditory streaming effect using event-related brain potentials. Psychophysiology. 1999; 36:22–34. [PubMed: 10098377]

Sussman ES. Integration and segregation in auditory scene analysis. J. Acoust. Soc. Am. 2005; 117:1285–1298. [PubMed: 15807017]

Sussman ES. A New View on the MMN and Attention Debate. The Role of Context in Processing Auditory Events. J. Psychophysiol. 2007; 21:164–175.

Sussman ES, Chen S, Sussman-Fort J, Dinces E. The Five Myths of MMN: Redefining How to Use MMN in Basic and Clinical Research. Brain Topogr. 2013:1–12.

Teki S, Chait M, Kumar S, Shamma S, Griffiths TD, Angelaki D. Segregation of complex acoustic scenes based on temporal coherence. eLife. 2013; 2

Teki S, Chait M, Kumar S, von Kriegstein K, Griffiths TD. Brain Bases for Auditory Stimulus-Driven Figure–Ground Segregation. J. Neurosci. 2011; 31:164–171. [PubMed: 21209201]

Tervaniemi M, Maury S, Näätänen R. Neural representations of abstract stimulus features in the human brain as reflected by the mismatch negativity. NeuroReport. 1994; 5:844–846. [PubMed: 8018861]

Trainor LJ, Trehub SE. Aging and auditory temporal sequencing: Ordering the elements of repeating tone patterns. Percept. Psychophys. 1989; 45:4117–4426.

Tremblay KL, Piskosz M, Souza P. Effects of age and age-related hearing loss on the neural representation of speech cues. Clin. Neurophysiol. 2003; 114:1332–1343. [PubMed: 12842732]

Wacongne C, Changeux J-P, Dehaene S. A Neuronal Model of Predictive Coding Accounting for the Mismatch Negativity. J. Neurosci. 2012; 32:3665–3678. [PubMed: 22423089]

Wingfield A, Peelle JE, Grossman M. Speech Rate and Syntactic Complexity as Multiplicative Factors in Speech Comprehension by Young and Older Adults. Aging Neuropsychol. Cogn. 2003; 10:310–322.

Winkler I. Interpreting the Mismatch Negativity. J. Psychophysiol. 2007; 21:147–163.

Winkler I. Predictive models in auditory stream segregation. Proc. 15TH WORLD Congr. Psychophysiol. Int. Organ. Psychophysiol. IOP Bp. Hung. 2010 Sep 1–4.77:215–215. 2010.

Winkler I, Czigler I. Evidence from auditory and visual event-related potential (ERP) studies of deviance detection (MMN and vMMN) linking predictive coding theories and perceptual object representations. Predict. Inf. Process. Brain Princ. Neural Mech. Models. 2012; 83:132–143.

Winkler I, Denham S, Mill R, Bőhm TM, Bendixen A. Multistability in auditory stream segregation: a predictive coding view. Philos. Trans. R. Soc. B Biol. Sci. 2012; 367:1001–1012.

Winkler I, Denham SL, Nelken I. Modeling the auditory scene: predictive regularity representations and perceptual objects. Trends Cogn. Sci. 2009; 13:532–540. [PubMed: 19828357]

Winkler I, Korzyukov O, Gumenyuk V, Cowan N, Linkenkaer-Hansen K, Ilmoniemi RJ, Alho K, Näätänen R. Temporary and longer term retention of acoustic information. Psychophysiology. 2002; 39:530–534. [PubMed: 12212645]

Winkler I, Schröger E. Neural representation for the temporal structure of sound patterns. NeuroReport. 1995; 6:690–694. [PubMed: 7605929]

Winkler I, Sussman E, Tervaniemi M, Horváth J, Ritter W, Näätänen R. Preattentive auditory context effects. Cogn. Affect. Behav. Neurosci. 2003; 3:57–77. [PubMed: 12822599]

Woods DL. Auditory selective attention in middle-aged and elderly subjects: an event-related brain potential study. Electroencephalogr. Clin. Neurophysiol. 1992; 84:456–468. [PubMed: 1382955]

Yabe H, Winkler I, Czigler I, Koyama S, Kakigi R, Sutoh T, Hiruma T, Kaneko S. Organizing sound sequences in the human brain: the interplay of auditory streaming and temporal integration. Brain Res. 2001; 897:222–227. [PubMed: 11282382]

Zion Golumbic EM, Ding N, Bickel S, Lakatos P, Schevon CA, McKhann GM, Goodman RR, Emerson R, Mehta AD, Simon JZ, Poeppel D, Schroeder CE. Mechanisms Underlying Selective Neuronal Tracking of Attended Speech at a “Cocktail Party”. Neuron. 2013; 77:980–991. [PubMed: 23473326]

Zion Golumbic EM, Poeppel D, Schroeder CE. Temporal context in speech processing and attentional stream selection: A behavioral and neural perspective. Neurobiol. Lang. 2010 Neurobiol. Lang. Conf. 2012; 122:151–161.

Zwislocki JJ. Temporal Summation of Loudness: An Analysis. J. Acoust. Soc. Am. 1969; 46:431–441. [PubMed: 5804115]

Highlights

- We review the relevance of temporal structure for hearing in background noise
- We review age-related deficits in temporal structure processing in auditory cortex
- Neurophysiological measures advance aging research on hearing in background noise

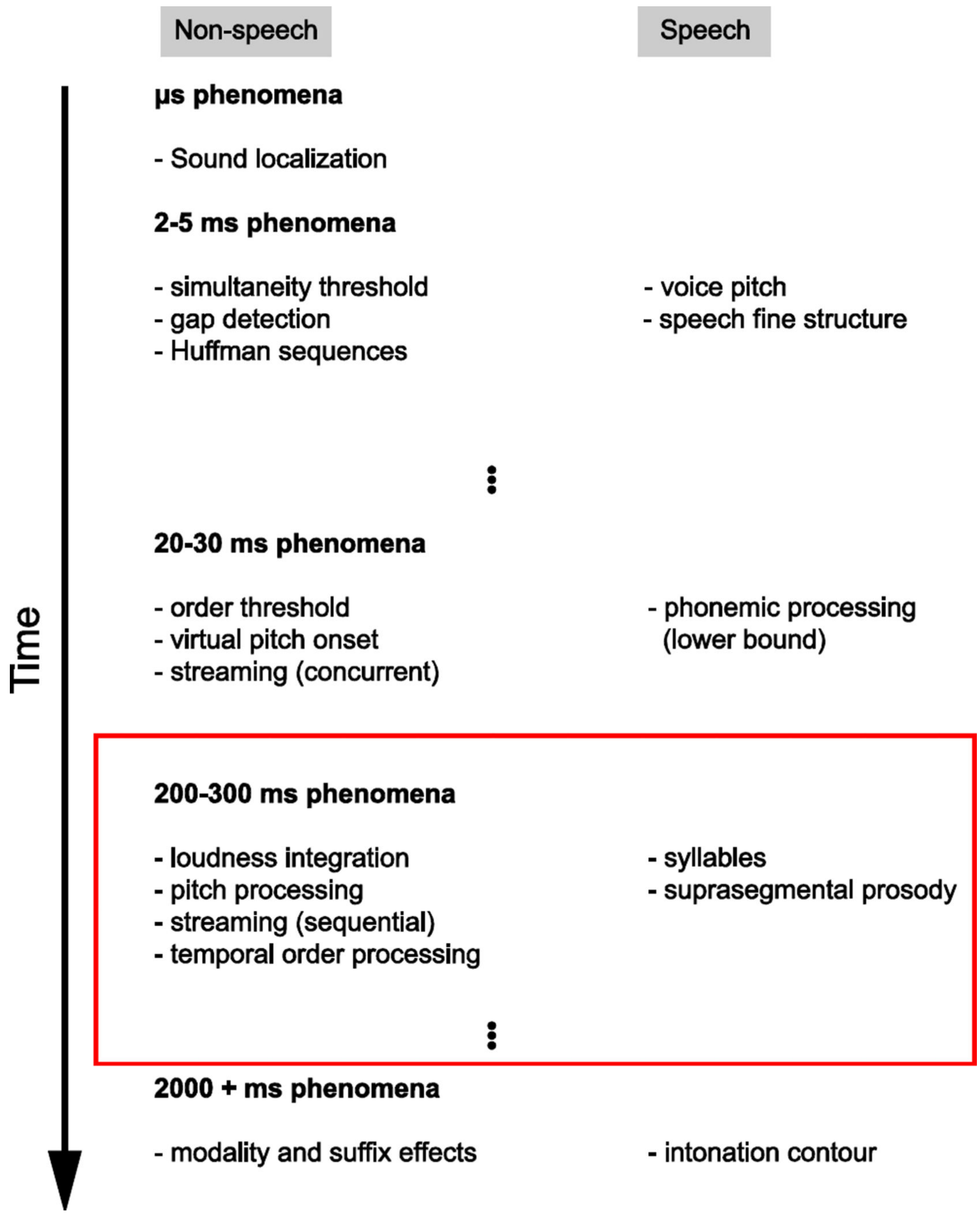

Figure 1. Sound contains temporal structure at multiple time scales

Non-speech perceptual phenomena at multiple time scales are displayed in the left column; speech perceptual phenomena are displayed in the right column (Time bar: short duration phenomena are displayed at the top and longer duration at the bottom). The review discusses temporal processing at a time scale from about 200 ms to several seconds (red box).

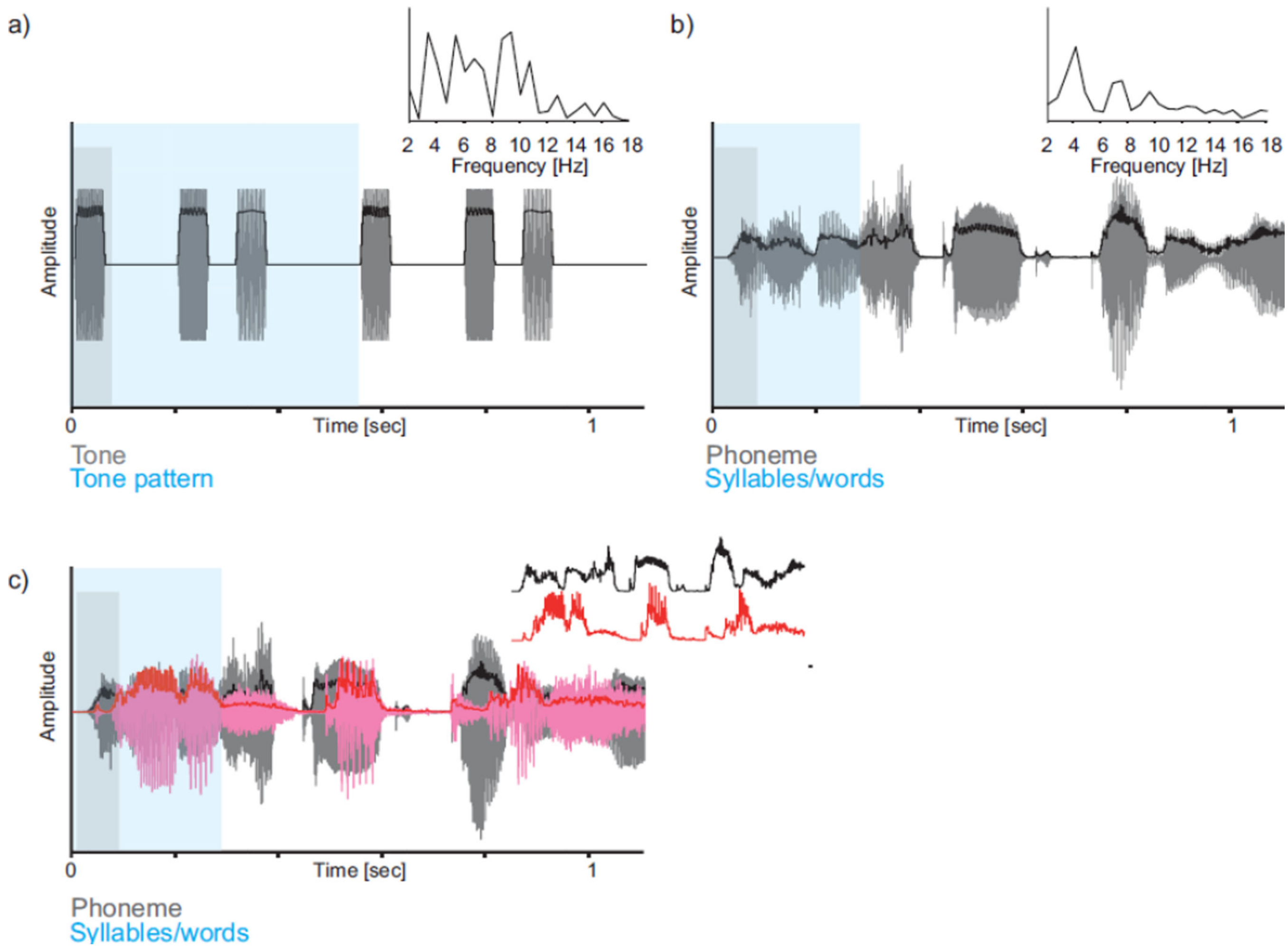

Figure 2. Temporal structure in speech and non-speech stimuli

a) Temporal structure of pure tone patterns. Single tones can be processed at a faster time scale (grey box), while the perception of tone patterns (blue box) occurs at slower time scales. Tone patterns, as displayed here, are typically used in sensory memory research. The envelope (displayed in black) of the tone pattern waveform (grey) illustrates the temporal structure of the tone patterns at a slower time scale. The frequency spectrum of the envelope of the tone pattern waveform is displayed on the right above the waveform (calculated as the RMS, as described in Lalor and Foxe [2010]. The frequency spectrum of the envelope shows peaks in the theta and alpha range. **b)** Temporal structure of speech stimuli (displayed as in a)). The frequency spectrum of the envelope shows typical peaks in the theta frequency range. **c)** Temporal structure of two simultaneously presented sentences (grey/black and pink/red waveform). (displayed as in **a)**). The envelopes of both sentences are displayed separately on the right top of the waveform.

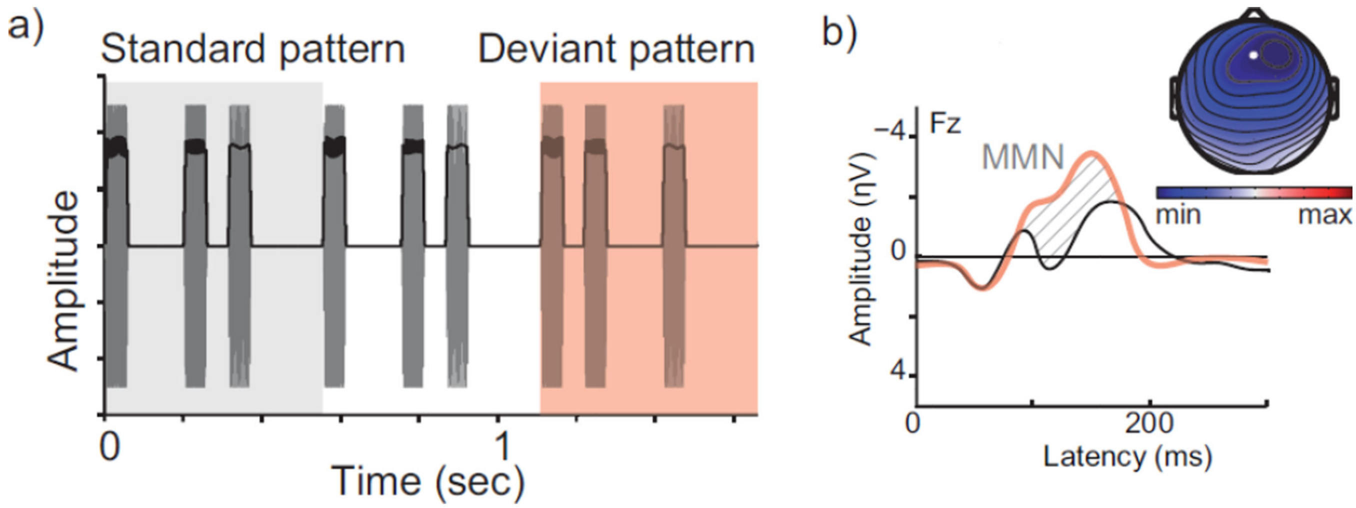

Figure 3. Mismatch negativity (MMN) as indicator for temporal structure processing in sensory memory

a) The soundwave of a repeating three tone pattern is displayed in grey, with the temporal envelope displayed in black (calculated as the RMS, as described in Lalor & Foxe [2010]). A standard tone pattern is underlayed with a grey box, a deviant tone pattern (two interstimulus-intervals are interchanged) with a red box. **b)** Illustration of event-related potentials (ERPs) that are elicited by standard (black line) and by deviant patterns (red line) at Fz electrode, are illustrated. MMN component is calculated as deviant-minus-standard ERP (grey shaded plane under the curve). The fronto-central topography of MMN is illustrated (spline interpolated isocontour voltage maps of the grand average of the MMN component); Fz electrode displayed as white circle).

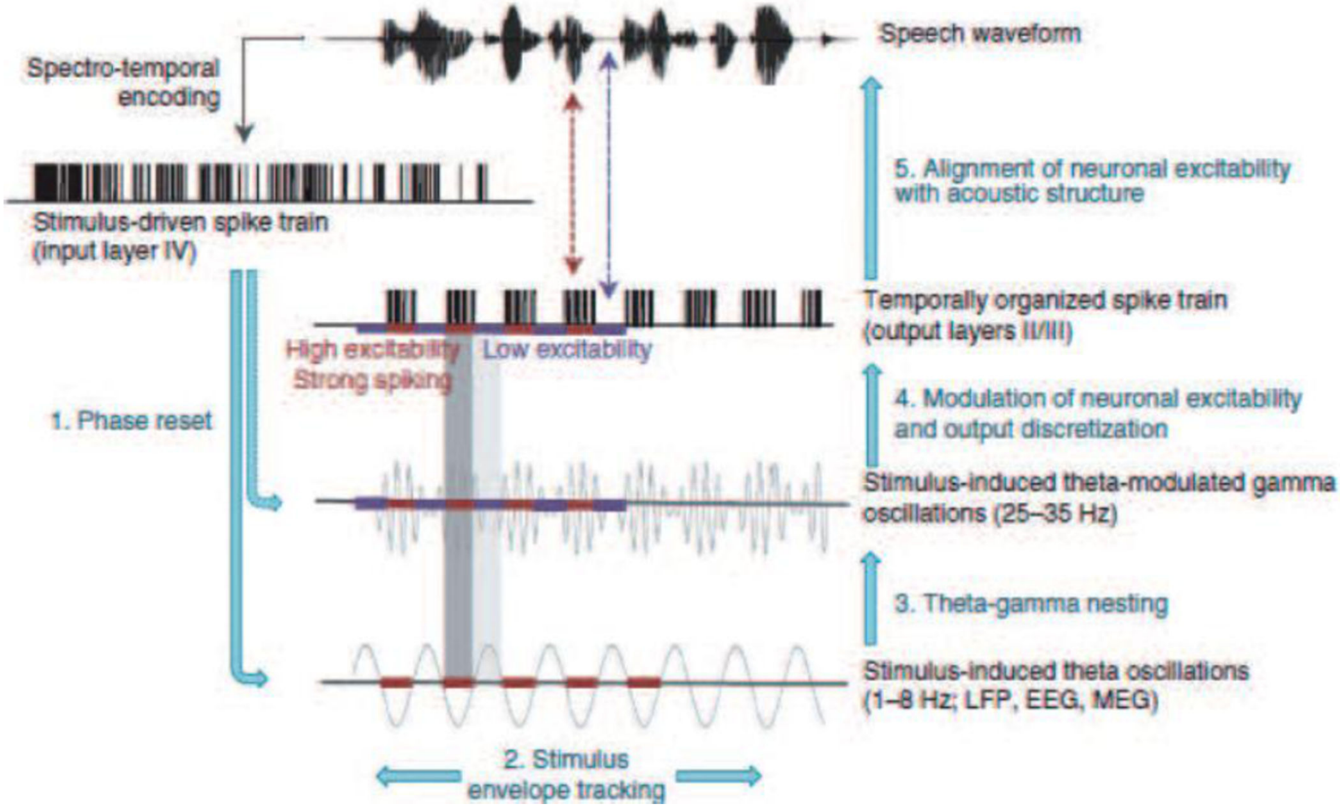

Figure 4. Neuroal oscillations encode temporal structure of speech (figure by Giraud & Poeppel, 2013)

Giraud & Poeppel propose a theory of early oscillation-based speech processing. Speech is tracked by cortical theta and gamma oscillations through five operations. Speech-tracking results in a high-resolution spectro-temporal representation of speech in primary auditory cortex. Most of the layer IV cortical neurons phase-lock to amplitude modulations in the speech signal. The onset of this response results in phase reset of theta oscillations in superficial layers where auditory cortex output is generated (step 1). Theta oscillations now track the speech envelope (step 2). In a third step (3), theta reset results in gamma reset and thus frequency nesting. Neural excitability phase of neurons that generate the feedforward signal from A1 to higher order areas is controlled by gamma power (step 4). This results in an alignment of neuronal excitability phase with the acoustic structure (step 5).