Hearing Research 305 (2013) 10–18

Contents lists available at ScienceDirect

Hearing Research

journal homepage: www.elsevier.com/locate/heares

Review

Attention to natural auditory signals

Emily Caporello Bluvas [a,b,*], Timothy Q. Gentner [a,c,d]

[a] Neurosciences Graduate Program, University of California, San Diego, USA
[b] Institute For Neural Computation, University of California, San Diego, USA
[c] Department of Psychology, University of California, San Diego, USA
[d] Kavli Institute for Brain and Mind, La Jolla, CA, USA

A R T I C L E I N F O

Article history:
Received 12 April 2013
Received in revised form
13 August 2013
Accepted 17 August 2013
Available online 24 August 2013

A B S T R A C T

The challenge of understanding how the brain processes natural signals is compounded by the fact that such signals are often tied closely to specific natural behaviors and natural environments. This added complexity is especially true for auditory communication signals that can carry information at multiple hierarchical levels, and often occur in the context of other competing communication signals. Selective attention provides a mechanism to focus processing resources on specific components of auditory signals, and simultaneously suppress responses to unwanted signals or noise. Although selective auditory attention has been well-studied behaviorally, very little is known about how selective auditory attention shapes the processing on natural auditory signals, and how the mechanisms of auditory attention are implemented in single neurons or neural circuits. Here we review the role of selective attention in modulating auditory responses to complex natural stimuli in humans. We then suggest how the current understanding can be applied to the study of selective auditory attention in the context natural signal processing at the level of single neurons and populations in animal models amenable to invasive neuroscience techniques.

This article is part of a Special Issue entitled "Communication Sounds and the Brain: New Directions and Perspectives".

1. Introduction

Processing natural stimuli is not a passive, deterministic event. As the complexity of signals increases, so too does the number ways that they can be perceived and thus related to different behaviors. For example, the acoustic environments in which natural signals, such as speech, occur are often contaminated by background noise and other potentially informative auditory signals, including other speech. The simple act of conversing with a friend on a crowded bus, for example, requires one to "tune out" the road, engine and other ambient noise, and "focus on" the stream of speech being produced by one's friend rather than that from the talkative stranger nearby. The challenge of attending selectively to a stream of speech amid one or more competing streams of speech is referred to as the cocktail party problem (CPP; Cherry, 1953). Humans and non-humans can solve species-typical versions of the CPP, but how these solutions are implemented in the activity patterns of single neurons, neural populations, and well-defined neural circuits is not known. Gaining such knowledge requires animal models that permit direct control over attention driven behaviors and invasive experimental techniques unavailable in humans. Here we review literature on selective attention to speech, and introduce relevant animal studies of natural communication and auditory attention with the broader goal of understanding neural mechanisms that control natural signal processing in realistic environments.

2. Auditory object formation and attention

Sensory signals can be understood at elementary levels, e.g. photons striking the retina, but it is the patterns in these simple elements and their organization into objects that guides most behaviors. We "see" faces and trees, etc. not photons. The same holds for audition where the ability to parse acoustic signals into objects is crucial for vocal communication and most other natural behaviors. Parsing the acoustic environment into distinct objects, known as auditory scene analysis (ASA; Bregman, 1990), depends on the immediate acoustic properties of the auditory scene and expectations of object parameters (for recent a review see Moore and

---

* Corresponding author. Department of Psychology, University of California, San Diego, La Jolla, CA 92093, USA. Tel.: +1 858 822 1867.
E-mail addresses: emilycaporello@gmail.com (E. Caporello Bluvas), tgentner@ucsd.edu (T.Q. Gentner).

0378-5955/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.heares.2013.08.007

Gockel, 2012). Attention, in turn, helps to allocate limited processing resources to objects relevant for one behavior or another. It can be captured by low-level acoustic features in a “bottom–up” manner, or in the case of the cocktail party problem, selectively directed to objects of interest, such as speech, in a “top–down” manner. Whereas “bottom up” attention is often driven by salient differences between objects (or between an object and the background), solving the cocktail party problem requires that objects of interest are segregated from competing objects, typically of the same kind (e.g., other speech). Often times, the stimulus-driven “bottom–up” and selective “top–down” processes interact, such as when a speaker is separated from other sound sources in one or more dimensions of acoustic parameter space, e.g., location (Cherry, 1953) or pitch (Brokx and Nooteboom, 1982; Darwin, 1997; Darwin and Hukin, 2000a; Spieth et al., 1954).

Scene analysis also interacts with attention. There exists a long history of research examining how ignored (unattended) auditory objects are processed, and what information is available to the listener from outside of the attended stream. Consistent with the idea that acoustics play a strong role in stream formation, abrupt changes in the low-level acoustic properties of unattended streams, such as adding tones or altering pitch (e.g. by switching speaker sex; Cherry, 1953), are detected by most listeners. Likewise, the inclusion of behaviorally relevant sounds in the ignored speech streams, such as one’s name (Moray, 1959) or words related to content in the attended stream (Cherry, 1953; Treisman, 1960) can be later recalled by listeners. Even when explicit content of the ignored speech cannot be recalled, performance related to the attended stream can be negatively affected when the ignored speech includes content that is semantically related to the attended stream (Lewis, 1970; Treisman et al., 1974). Increasing the number of distracter streams – as opposed to increasing the acoustic content within a single ignored stream – can also affect performance related to the attended stream (Bregman and Rudnicky, 1975; Macken et al., 2003) and the magnitude of attention-related electroencephalography (EEG) potentials (Alain et al., 1993). Indeed, many acoustic deviations (i.e. expectancy violations) in ignored streams, including streams composed of linguistic stimuli (Coch et al., 2005), produce detectable changes in (EEG) signals (Alain et al., 1994; Paavilainen et al., 1995; Winkler et al., 2005; Sussman et al., 2005) Thus, auditory stream (object) formation can occur without direct attention, and information in un-attended streams is subject to substantial high-level processing.

Despite these considerations, many less salient aspects of unattended streams do fall outside of the awareness of the listener. In the earliest CPP experiments, subjects easily reported the content and detailed talker attributes from target streams. Yet, they were unable to recall specific words or phrases from non-target streams, and failed altogether to notice more dramatic alterations in unattended streams such as a change in the language (Cherry, 1953) or reversal of the speech (Wood and Cowan, 1995), so long as low-level acoustics expectations remained intact. Additionally, while attention may not be necessary for object formation, it can speed up the segmentation process (see Snyder et al., 2012 for a comprehensive review). For instance, in a well known ASA stimulus paradigm where triplets of tones can be perceived as one or two separate streams (“galloping tone” stimuli Alain and Bernstein, 2008; Moore and Gockel, 2012), the time period required to build-up to a percept of two streams (Bregman, 1978) can be reduced by directing attention to the auditory scene (Carlyon et al., 2001, 2003; Snyder et al., 2006) or reset by diverting attention away from the auditory scene (Carlyon et al., 2003; Thompson et al., 2011). Whether or not attention can affect the time course of scene analysis/object formation in the context of speech or other natural stimuli is undetermined.

Although it is a necessary component, ASA alone cannot provide a complete solution to CPP. Like other vocal communication signals, the continuous speech signal produced by a single talker can vary greatly (and rapidly) along multiple perceptual dimensions including pitch, amplitude and spatial location. Overlap with other streams may occur, along with the addition of extraneous noises that may make it temporarily impossible to track a particular stream of speech. Yet behaviorally, brief interruptions in speech and non-verbal stimuli (Warren and Verbrugge, 1984) can go entirely unnoticed by the listener. This phenomenon, referred to as temporal induction (TI), can, given adequate acoustic and semantic information, provide the illusion of continuous speech over several hundred milliseconds (Bashford and Warren, 1987; Bashford et al., 1992). Like ASA, TI may contribute to formation and/or the ongoing maintenance of an auditory object/stream. In particular, the TI phenomenon is evidence that brief portions of an auditory stream can go unheard without behavioral consequence, suggesting that attention may not need to be allocated uniformly and continuously across time to be effective.

3. Broad-scale neural correlates of auditory attention

Consistent with the understanding that selective attention can modulate auditory responses at multiple processing stages, it is not surprising to find that neural correlates of attention, measured with EEG and MEG (magnetoencephalography), appear throughout the sensory stream. Attention-dependent modulation of EEG signals can be observed as soon as 20 ms after stimulus onset (Woldorff and Hillyard, 1991; Woldorff et al., 1987), an effect thought to originate from primary auditory cortex (Woldorff et al., 1993). MEG, functional magnetic resonance imaging (fMRI), and positron emission tomography (PET) data further support early cortical gain in the contralateral temporal lobe of attended tones and speech (Ding and Simon, 2012; Grady et al., 1997; O’Leary et al., 1996), both within primary (Fujiwara et al., 1998; Jäncke et al., 1999) and secondary auditory regions (Okamoto et al., 2011; Petkov et al., 2004; Woods et al., 2009). Early dichotic listening studies, where subjects attend to tone streams in one ear while ignoring input to the other ear, indicated that selective attention can modify auditory event-related potentials (ERPs). Specifically, the N1 negative component (80–110 ms after sound onset), P2 component (150–200 ms after onset) and P3 component (250–450 ms after onset) of the ERP waveform are larger when auditory stimuli presented to the contralateral ear are attended versus ignored (Hillyard et al., 1973; Picton and Hillyard, 1974).

The canonical ERP waveform components differ in their sources and their sensitivities to different aspects of attention tasks. N1 and P2 components can be modified by the amplitude, location, and stimulus parameters of acoustic stimuli (Hillyard et al., 1973). The N1 component is thought to originate at least in part within the primary auditory cortex (Knight et al., 1981; Giard et al., 1988), while the P2 component is thought to originate in secondary cortices anterior to A1 (Rif et al., 1991). Attention modifies the magnitude of the N1 and P2 components elicited by all stimuli (targets and distracters) in the attended channel (Hansen and Hillyard, 1978; Hillyard et al., 1973; Picton and Hillyard, 1974). The P3 component, however, which is modified only by behaviorally relevant stimuli, is largely invariant to changes in stimulus parameters (Hillyard et al., 1973); it occurs in multiple sensory modalities (Barlow, 1969; Klinke et al., 1968), and is elicited even during behaviorally relevant stimulus omissions (Picton and Hillyard, 1974). The P3 component is thought to originate in parietal regions (Downar et al., 2000; Picton and Hillyard, 1974), and to reflect modality-general decision making, arousal, or motor response (reviewed in Pritchard, 1981).

Selective auditory attention can also modify neural activity in subcortical sensory regions such as the inferior colliculus (Rinne et al., 2008) and possibly the dorsal cochlear nucleus (Kaltenbach, 2006), presumably through the descending corticofugal projections from auditory cortex (Winer, 2005). There is debate over whether selective attention effects can be observed in the medial olivocochlear bundle (recently reviewed in Tanedo and Bhagat, 2008), but recent studies confirm the presence of attention-dependent modulations in the magnitude of evoked otoacoustic emissions (EOAEs; De Boer and Thornton, 2007; Maison et al., 2001; Smith et al., 2012). Regardless, it is reasonable to conclude that most levels of the central auditory system can be modulated by attention. One potential advantage to modifying responses at the level of the brainstem, or even the cochlea, is that speech harmonics or other first order features useful for scene segregation may be differentially enhanced (Anderson and Kraus, 2011).

Task demands also affect attention-related activity. Both location and pitch can be used to solve the cocktail party problem (Bregman, 1990; Darwin and Hukin, 2000a, 2000b; Spieth et al., 1954; Woods et al., 2001), and attending to either the spectral features of speech or its location activate distinct brain regions. Attention to location results in increased activity in regions associated with spatial processing, such as posterior temporal cortex and posterior parietal regions, while attention to spectral features of speech leads to increased activity in auditory- and speech-specific regions such as anterior and superior temporal cortex (Ahveninen et al., 2006; Hill and Miller, 2010). The distinct activity patterns elicited by spatial and non-spatial attention to speech also correspond to similar “what” vs. “where” pathway activity seen in auditory discrimination tasks with non-linguistic stimuli (Alain et al., 2001; Degerman et al., 2006; Maeder et al., 2001; Paltoglou et al., 2011; Woods and Alain, 2001). Different brain regions are affected even when altering the discrimination between non-spatial auditory features, such as duration vs. pitch (Brechmann and Scheich, 2005) high vs. low frequency (Paltoglou et al., 2009), or pitch vs. loudness (Rinne et al., 2007). A similar effect is seen in bimodal selective attention studies, where subjects are presented with both visual and auditory stimuli and perform discriminations in one modality or another (Petkov et al., 2004; Johnson and Zatorre, 2005, 2006; Woods et al., 2009; Degerman et al., 2007; Salmi et al., 2007). In each case, the regions showing increased activation above baseline correlate to the features that are behaviorally relevant for the task, or, in tasks where attention is not being explicitly measured, the features that are likely being attended. Similar effects have been observed when measuring the strength of gamma oscillations over the temporal lobe during auditory (versus visual) attention (Sokolov et al., 2004). Increased gamma activity has been associated with attention in other modalities, such as vision (Tallon-Baudry et al., 1997; Tiitinen et al., 1993).

In line with these results, attending to speech, as opposed to non-linguistic stimuli, activates regions such as the superior temporal sulcus and gyri that are known to be involved in language processing (Alho et al., 2006; Belin et al., 2000; Hari et al., 1989; Pasley et al., 2012; Rämä et al., 2012), often more prominently within the left hemisphere (Jäncke et al., 1999; Hugdahl et al., 2003; Coch et al., 2005; Jäncke et al., 2001; Hill and Miller, 2010; Alho et al., 2003; but see Ding and Simon, 2012). An early electrocorticography (ECog) study in humans similarly found a gain in gamma band power within the left hemisphere secondary auditory cortex during a phoneme discrimination task (Crone et al., 2001), and a more recent EEG study found increased gamma band activity in the left temporal lobe during the correct detection of degraded words (Hannemann et al., 2007). This well-documented left-hemisphere specialization for speech processing is thought to underlie the frequently observed bias in subjects to attend to speech presented to the right ear (“right ear advantage”, see Hugdahl et al., 2009 for a review, Kompus et al., 2012), since auditory information is processed in the hemisphere contralateral to peripheral stimulation.

Within specific brain regions, attention improves discriminability of speech-evoked responses. A recent electrocorticography study measured activity in the posterior superior temporal lobe during selective attention to one of two simultaneously presented speakers (Mesgarani and Chang, 2012). In this region, local field potentials (LFP) in the high gamma range (75–150 Hz) were not only modified by attention, but the evoked response was more likely to resemble the response to the attended speaker alone. Compellingly, this effect was absent during error trials, indicating that poorly discriminable neural responses in cortex are associated with an inability to detect a lingual target in a noisy environment. Combined with other recent studies which have observed gamma-band modulation during associative auditory learning (Headley and Weinberger, 2011), selective auditory attention tasks (Crone et al., 2001; Debener et al., 2003; Doesburg et al., 2012), and several selective visual attention tasks (Fries, 2001; Gregoriou et al., 2009; Jensen et al., 2007; Siegel et al., 2008), it seems clear that gamma band frequency oscillations in sensory cortices, particularly in regions that encode the attended target, are tied closely to attentional state. Enhanced power within the gamma band is thought to reflect synchronous firing at gamma frequencies, which, when localized to receptive fields overlapping with the attended stimulus, may provide a mechanism for robust encoding of the target (see Jensen et al., 2007 review). However, the neural mechanisms that underlie these broad changes, like many observed effects of attention, remain largely unknown.

The foregoing evidence indicates that attention to auditory objects via top–down mechanisms tends to most greatly affect the encoding of object parameters which are most useful for segregation from background acoustics, suggesting that the mechanisms of attention are affected not only by the attended content, but by the background content as well. Different brain regions may be differentially modulated by auditory attention depending on the nature of specific stimuli and task demands, e.g., detecting a difference in space or in pitch. Other studies also note, however, consistent attention-dependent modulation of the *same* regions regardless of task parameters including superior parietal cortex, dorsolateral frontal cortex, the pre-central sulcus, and middle and superior temporal cortex (Degerman et al., 2006; Falkenberg et al., 2011; Hill and Miller, 2010; Paus et al., 1997; Salmi et al., 2009; Westerhausen et al., 2010; Wu et al., 2007; Zatorre et al., 1999). A recent study of spatial attention using high-density electroencephalography found that alpha band activity (8–12 Hz) originating in posterior parietal regions could be used to determine the identity of an attended speech source among two distracters (Kerlin et al., 2010). Similar activation of superior parietal cortex has been reported for fMRI data during attentional switching in visual tasks (Sauseng et al., 2005; Shomstein and Yantis, 2006), suggesting a role for parietal cortex in spatial attention across modalities (Giesbrecht et al., 2003; Shomstein and Yantis, 2006; Sauseng et al., 2005; Rihs et al., 2009).

Overall, the evidence supports a flexible model of attention where the loci of attentional modulation are highly affected by the acoustic parameters and behavioral relevance of the stimuli, as well as the ultimate task-related goals of the animal. It is likely that what we refer to by one term, “attention”, reflects the effects of network-level actions to modulate the modality-specific or modality-general processing capabilities in different brain regions. Although the net effect of these modulations share many similarities, and many are simultaneously engaged by complex stimuli, the implementation of “attention” in any auditory region at the single neuron or network-level is unknown, and so the so mechanistic similarities between regions (or sensory modalities), or the existence of a general “attention network”, remain open questions. To gain more detailed insight into the neural circuitry modulated by attention, we turn to studies in animal models.

4. Animal models

Recent years have seen an increasing interest in investigating auditory streaming and attention in animal models. Besides allowing for more invasive experimental techniques that can address the neural mechanisms underlying attention, many animal species regularly encounter difficult cocktail-party-like environments in which selection of a relevant target stream would likely require attention (see Bee, 2008 for a review). For instance, king penguins reside in large noisy colonies of several thousand vocal birds, yet young penguins can identify the distant call of their parent at volume levels 6 dB below background noise (Aubin and Jouventin, 1998). The difficulty of this recognition task is made clear by the many strategies employed by the penguins to solve it quickly, such as modifying acoustic parameters of the call to separate it from background noise, and limiting calls to geographically specific “rendezvous areas” (Aubin, 2004). Frogs, like humans, can use the frequency separation between simultaneous communication calls as a perceptual cue for stream segregation (Nityananda and Bee, 2011). In somewhat more disturbing example, Vampire bats, who will often feed on the same prey animal several nights in a row, can detect faint breathing sounds and use them to identify individuals (Gröger and Wiegrebe, 2006). Controlled laboratory experiments on European starlings, a songbird species, indicate that subjects can learn to recognize an individual song even when it is always simultaneously presented with another song, and that subjects can recognize learned songs masked by a loud songbird chorus (Hulse et al., 1997). Both ferrets and European starlings have been trained to perform stream segregation tasks similar to those used to study the process in humans (Ma et al., 2010; MacDougall-Shackleton et al., 1998). Even singing insects, such as crickets, face a version of the cocktail party problem in their natural environments. Crickets rely on several adaptations to identify conspecifics in noisy environments, including neural gain control mechanisms that reduce the influence of background noise from relevant call signals (Schmidt and Römer, 2011). As in the case of human speech detection, animal models are likely to employ both ASA (cited above) and TI (Petkov et al., 2003, 2007; but see Seeba et al., 2010).

5. The advantages to studying animals

One benefit of developing non-human models of auditory attention is the application of techniques that allow researchers to investigate the neural mechanisms underlying attention, such as extracellular and intracellular electrophysiology, calcium and voltage imaging, viral transfection, and transgenetics. These techniques could help us understand more about the attentional effects observed in human auditory attention studies, such as the role of corticofugal pathways underlying subcortical attention-dependent changes in neural activity (Ryan et al., 1984; Suga and Ma, 2003) the existence of an anatomical hierarchy of auditory attention modulation as reported in the visual domain (Yin et al., 2008), and the role of neurotransmitters such as acetylcholine (Bakin and Weinberger, 1996; Kilgard and Merzenich, 1998a, 1998b; Metherate and Weinberger, 1989; Metherate et al., 2012). Single neuron data can also allow us to investigate the applicability of models proposed to solve the cocktail party problem and selective attention, such as those involving sparse coding (Asari et al., 2006), normalization (Reynolds and Heeger, 2009), receptive field modulation (Mesgarani et al., 2010), and others (Jääskeläinen et al., 2011; Sagi et al., 2001), and to understand how the powerful algorithms that underlie these models are implemented in detailed biological networks.

6. Moving forward in auditory attention research

The application of extracellular electrophysiology in auditory attention research originated decades ago in the classic work of Hubel et al. (1959). In a largely anecdotal study, the authors report data from neurons in secondary auditory cortices of awake unrestrained cats that responded to complex sounds only when the cat appeared to be attending to it (often inferred from gaze orientation). The observed results were intriguing, but at the time little was known about the processing of complex sounds in high-level cortical regions or controlling attention in non-human species. As the authors write at the end of their paper, “attention is an elusive variable that no one has as yet been able to quantify.” Perhaps for these reasons, follow-up work on auditory processing in awake, behaving animals came only years later and focused primarily on simpler stimuli (such as clicks, tones, and white noise) driving neural responses in regions such as A1 that were better understood functionally and anatomically. These studies consistently reported an increase in the single-neuron response to a behaviorally-relevant stimulus during task engagement compared to passive or ignored conditions (Miller et al., 1972; Hocherman et al., 1976; Pfingst et al., 1977; Benson and Hienz, 1978; for a comprehensive review see Fritz et al., 2007a,b). Decades of research on learning-dependent plasticity in the auditory cortex has provided us with ample evidence that the response properties of auditory neurons are capable of changing to better respond to relevant stimuli (for reviews see Pienkowski and Eggermont, 2011; Jääskeläinen et al., 2007). A well-known example of this plasticity is the reorganization of the A1 tonotopy that leads to an over-representation of behaviorally relevant frequencies following discrimination training (Recanzone et al., 1993).

The enhanced response to an attended stimulus over an unattended stimulus could be accomplished in a variety of ways. A theory developed early on as a result of the first EEG dichotic listening studies involved a specific auditory attention network that would represent a “trace” of the attended target, against which other auditory inputs would be compared to (Näätänen, 1992; Näätänen and Michie, 1979). This was in contrast to a gain model, also proposed early on in auditory attention research (Hillyard et al., 1973; Woldorff et al., 1993) and similar to those proposed in visual attention (Desimone and Duncan, 1995; Hillyard et al., 1998), in which the firing rates of cells that respond to the attended stimulus increase, while the firing rates of cells that respond to the ignored stimulus decrease. Another more recently proposed version is the tuning model, where attention also modulates the tuning properties of individual cells (Jääskeläinen et al., 2007) in addition to general gain-model modulation. This model shares similarity with a recently proposed model of visual attention (Reynolds and Heeger, 2009). In support of the tuning model, Fritz and colleagues developed experiments that demonstrated rapid changes in individual neuron response properties during a difficult tone detection task (Fritz et al., 2003). Shortly after the onset of behavior, the receptive fields of ferret A1 neurons changed to enhance task-relevant frequencies. These changes developed rapidly and could be reshaped by introducing a new behaviorally relevant frequency. In follow-up experiments using chord targets with several relevant frequencies, the authors were able to show that the cells' receptive fields could change to enhance responses to several discrete, behaviorally relevant frequencies. Stimulus reconstruction data from neural activity in these ferrets indicate that behavioral state (and likely attention) affects the tuning properties of A1 neurons (Mesgarani et al., 2008), and recent human studies using EEG and fMRI further support state-dependent tuning modulation (Ahveninen et al., 2011; Kauramäki et al., 2007; Okamoto et al., 2007). Interestingly, Fritz and colleagues not only saw enhancement of relevant frequencies, but also suppression of irrelevant frequencies (Fritz et al., 2005, 2007b). This suppression of the single-unit response to irrelevant sounds has been confirmed in other experiments (Otazu et al., 2009) and is consistent with imaging data in humans (Paltoglou et al., 2009). Most recently, researchers have shown that the magnitude of suppression in the receptive field is strongest in low signal-to-noise environments (Atiani et al., 2009; Neelon et al., 2011), consistent with the interpretation that suppression is increasing along with attentional load.

The foregoing studies highlight the point that selective auditory attention should not be operationalized in only one way. Attention modulates responses (both neural and behavioral) at multiple levels in processing pathways and in multiple ways depending on the interaction of behaviorally relevant stimuli, task demands, experience, and expectations. Further development of animal models for attention therefore necessitate the incorporation of natural, behaviorally relevant stimuli into behavioral designs the control and isolate specific aspects of attention so that underlying neural circuits can be studied and their modulation by attention determined.

7. Development of behavioral tools to study the rapid, flexible allocation of attention

To date, the predominant means of manipulating attention has been to measure neural responses to the same stimuli presented during passive listening and while the animal is actively using them in some kind of auditory task (Fritz et al., 2003; Otazu et al., 2009). While attention almost certainly varies across these two conditions, other behavioral-state variables, such as motivation, arousal, reward expectation, and non-attention driven task demands vary as well. Indeed, rapid shifts in receptive field gain commonly attributed to attention (Fritz et al., 2003), can be recapitulated by varying task demands (David et al., 2012).

Dissociating the effects of attention from those of reward, working memory, arousal, and motor control is difficult. Researchers in the visual system have spent years carefully developing tasks that can control for and measure the allocation of attention (Yantis, 1998). The hallmark of selective attention is that in a highly stimulating environment, the amount of sensory input is simply too much to process simultaneously (James, 1890). Targets are chosen from the scene and attended to, while other parts of the scene are ignored. Task difficulty, captured in measures such as percent accuracy, is often used as evidence that the task requires attention to solve, and attention-related effects have been shown to scale with difficulty (Cate et al., 2009; Falkenberg et al., 2011). Another, perhaps more foolproof design includes “catch trials”, on which the experimenter tests the subject’s knowledge of the ignored parts of the scene. This is the animal behavior equivalent to asking human subjects if they noticed words or other aspects of the ignored speech in a dichotic listening task, noted by Hillyard and colleagues in their 1973 study as a crucial part of attention-related experimental design (Hillyard et al., 1973). If subjects cannot accurately report on the distracter stream, but can report on the target stream, then, by deduction, they are selectively attending to the target. Recently presented experiments using well-controlled auditory tasks (Rodgers and DeWeese, 2012; Yin et al., 2012), including those using natural communication signals (Caporello and Gentner, 2012), indicate promise in more direct studies of auditory attention in animal models.

8. Identifying the source: the prefrontal cortex and selective auditory attention

One important step for understanding auditory attention, particularly top–down selective attention, is localization of its source. The established role of prefrontal cortex ‘higher-level’ processes such as working memory, cognitive control, and decision making (see Duncan, 2001; Gold and Shadlen, 2007; Matsumoto and Tanaka, 2004; Miller and Cohen, 2001 for reviews), and in visual attention (Kincade et al., 2005; Vossel et al., 2006), make it a target in auditory attention as well. Consistent with such a role, selective attention can enhance gamma frequency oscillations between prefrontal and sensory cortices in both the visual (Buschman and Miller, 2007; Gregoriou et al., 2009; Siegel et al., 2008) and auditory (Doesburg et al., 2012; Fritz et al., 2007b) domains, and this synchrony may indicate a directional flow from prefrontal to sensory regions (Gregoriou et al., 2009). Recent imaging data reveals similar activation of frontal (Teshiba et al., 2013) and insular cortex during selective attention that is absent during bottom–up attentional capture (Huang et al., 2012). BOLD activation in frontal cortex has also been tied to task difficulty and/or attentional load (Altmann et al., 2008; Bidet-Caulet et al., 2010), although. Single neuron responses measured during auditory target detection (Fritz et al., 2010; Rodgers and DeWeese, 2012) and memory tasks (Yin et al., 2012) found that frontal cortex neurons respond robustly to behaviorally informative stimuli, sometimes up to 100 min following task performance (Fritz et al., 2010). Curiously, some studies have observed a lack of prefrontal cortex modulation during attention to continuous speech (Alho et al., 2003, 2006). This may reflect the fact that humans are very experienced in attending to continuous speech, and so it represents a relatively low attentional load, or a more substantial gap in our understanding.

One idea worth pursuing is that the prefrontal cortex is essential for the suppression of ignored auditory stimuli. Patients with compromised prefrontal function have difficulty ignoring noisy stimuli (reviewed in Knight et al., 1999), and anatomical work indicates that prefrontal regions project strongly onto inhibitory neuron populations in the auditory cortex (Barbas et al., 2013). Anatomically, visual and auditory regions connect with distinct regions of prefrontal cortex, and the prefrontal cortex is more richly connected to the associative auditory cortex than any other modality (Reviewed in detail in Barbas et al., 2013). This robust interconnectivity between prefrontal and auditory cortices may be an indicator that the auditory modality is a particularly strong candidate for selective attention research.

9. The unique nature of auditory stimuli

Although we have drawn parallels between selective attention and other sensory modalities such as vision (and we contend there is likely great translational profit in identifying shared mechanisms of attention), some aspects of auditory attention are none-the-less unique. Perhaps the most strikingly unique aspect is the critical role of time in auditory processing. While visual objects, defined along spatial dimensions, may remain constant over time, auditory objects are defined only across time. Unlike many artificial auditory stimuli such as tones and noise, which may remain constant through time, the features of natural auditory stimuli such as speech (pitch, amplitude, spectral contrast) fluctuate across time and it is precisely these fluctuations that carry the information in natural signals. We have already noted that expectations can influence auditory stream segregation, but it is worth noting that in the context of many vocal communication signals, including speech and language, the time scales for the expectations are dynamic. That is, the signal occurring at any given time may set up expectations about future signals at a range of possible times. While stimulus predictability has been shown to effect single unit processing of natural signals in songbird auditory regions (Gill et al., 2008), in general very little is known about the influence of expectation on attention mechanisms. It is reasonable to predict, however, that attention may bias temporal expectation so that particularly informative portions of natural auditory signals are encoded differently than others. Recent work supporting these ideas shows that adults (Sanders et al., 2002) and children as young as three years old (Astheimer and Sanders, 2012) attend most carefully to word onsets, which are highly informative about speech content. Moreover, attention probes presented during speech elicit larger N1 ERPs during word onsets (Astheimer and Sanders, 2009), than during the least predictable portions of speech (Astheimer and Sanders, 2011). The predictability of the attended stream also affects the low frequency oscillations obtained with ECog arrays, consistent with a role for temporal dynamics in attention (Besle et al., 2011).

Because of the dynamic nature of auditory objects, the factors that are most useful for separating targets from distracters could oscillate throughout the listening period – from pitch to space to volume, etc. Furthermore, the predictability of speech fluctuates in time, allowing for periods where attentional modulation may be more or less necessary for successful signal detection, discrimination, and recognition. The complexity and variable nature of speech makes it unlikely that attentional modulation is consistent over the stimulus duration in either its magnitude or in the brain regions it affects.

Conversely, peripheral coding of auditory signals lacks the spatial resolution of vision – while spatially separate visual objects are distinctly encoded in the retina, the nature of cochlear encoding requires that the brain reconstruct the spatial locations of sound sources. Indeed, evidence shows that cortical spatial coding differs in nature between the two modalities. Visual areas in the parietal cortex show clearly defined spatial maps, but parietal auditory regions that are sensitive to spatial location show only rough, if any, maplike structure (Kong et al., 2012). The process of separating auditory objects is made all the more difficult in noisy environments where sounds overlap in frequency space, and therefore in cochlear activation. These inherent differences in spatial representation should be kept in mind when comparing auditory and visual spatial attention data.

10. The importance of natural stimuli

To date, the majority of non-human auditory attention and task engagement experiments have employed simple or synthetic stimuli due to their easy interpretability and analytical benefits. Understanding behavioral relevance is crucial, however, to understanding the neural encoding of complex signals. In the songbird system, a well-established model of vocal communication and a one of the strongest animal models of human speech perception (Abe and Watanabe, 2011; Kiggins et al., 2012; Knudsen and Gentner, 2010), secondary regions of the auditory cortex selectively respond to conspecific birdsong (Grace et al., 2003) and behaviorally relevant songs elicit stronger responses from these regions than unfamiliar songs (Gentner and Margoliash, 2003; Gentner et al., 2004; Woolley and Doupe, 2008). Recently, data from our lab has demonstrated that songbirds can flexibly allocate auditory attention to one of two simultaneously presented birdsongs, and that this selective allocation of attention modifies high-level sensory neurons (Caporello and Gentner, 2012). In another animal model, the mouse, pup calls presented to mothers evoke stronger inhibition of frequencies below the call pitch than in virgin mice (Galindo-Leon et al., 2009). Overall, these results suggest that experience leads to changes in neural processing that enhance the neural responses to behaviorally relevant natural signals.

Additionally, recent studies in these regions have found that neural responses to song features are tolerant to natural variation (Meliza and Margoliash, 2012), amplitude (Billimoria et al., 2008), and spectro-temporal modulation (Moore et al., 2013), which may relate to mechanisms of stream segregation. In neurons affected by masking, the specific changes observed in the response to natural signals may indicate targets of auditory attention. For instance, in primary auditory regions, masking leads to an addition of spikes during gaps in natural signals and a loss of spikes during informative portions of the signal (Narayan et al., 2007). Perhaps the time points where these changes occur would be likely targets of auditory attention in noisy environments.

Although the acoustic complexity of natural signals presents a range of technical challenges, they also bestow several advantages. Natural communication signals, in particular, allow researchers the opportunity to investigate more complex and ethologically relevant auditory behaviors, such as individual recognition or mate choice, that have adapted in the context of natural acoustic scenes. In principle, these natural functions of communication signals can be exploited for the study of attention.

Language is a uniquely human trait, unparalleled in its rich semantics and complexity, and there may be mechanisms evolved specifically to process it. Although attending to speech likely involves language processing, speech perception must also be closely constrained by lower-level acoustic features (Nygaard et al., 1994; Pisoni et al., 1985; reviewed in Hickok and Poeppel, 2007). The problem of identifying, selecting, and acting upon the acoustics of complex conspecific vocalizations (like speech) in noisy environments is shared by many species (Knudsen and Gentner, 2010). This shared function, in turn, leads to the hypothesis that many aspects of selective auditory attention are conserved. By taking advantage of the behavioral abilities and technical advantages of animal models it both mechanisms of selective auditory attention at the level of single neurons and well-defined neural networks. Such an understanding would have tremendous benefit for treatment in a host of human attention-related disabilities.

References

Abe, K., Watanabe, D., 2011. Songbirds possess the spontaneous ability to discriminate syntactic rules. Nat. Neurosci. 14, 1067–1074.

Ahveninen, J., Jääskeläinen, I.P., Raij, T., Bonmassar, G., Devore, S., Hämäläinen, M., Levänen, S., Lin, F.-H., Sams, M., Shinn-Cunningham, B.G., et al., 2006. Task-modulated "what" and "where" pathways in human auditory cortex. Proc. Natl. Acad. Sci. U. S. A. 103, 14608–14613.

Ahveninen, J., Hämäläinen, M., Jääskeläinen, I.P., Ahlfors, S.P., Huang, S., Lin, F.-H., Raij, T., Sams, M., Vasios, C.E., Belliveau, J.W., 2011. Attention-driven auditory cortex short-term plasticity helps segregate relevant sounds from noise. Proc. Natl. Acad. Sci. U. S. A. 108, 4182–4187.

Alain, C., Bernstein, L.J., 2008. From sounds to meaning: the role of attention during auditory scene analysis. Curr. Opin. Otolaryngol. Head Neck Surg. 16, 485–489.

Alain, C., Achim, A., Richer, F., 1993. Perceptual context and the selective attention effect on auditory event-related brain potentials. Psychophysiology 30, 572–580.

Alain, C., Woods, D.L., Ogawa, K.H., 1994. Brain indices of automatic pattern processing. NeuroReport 6, 140–144.

Alain, C., Arnott, S.R., Hevenor, S., Graham, S., Grady, C.L., 2001. "What" and "where" in the human auditory system. Proc. Natl. Acad. Sci. U. S. A. 98, 12301–12306.

Alho, K., Vorobyev, V.A., Medvedev, S.V., Pakhomov, S.V., Roudas, M.S., Tervaniemi, M., Van Zuijen, T., Näätänen, R., 2003. Hemispheric lateralization of cerebral blood-flow changes during selective listening to dichotically presented continuous speech. Cogn. Brain Res. 17, 201–211.

Alho, K., Vorobyev, V.A., Medvedev, S.V., Pakhomov, S.V., Starchenko, M.G., Tervaniemi, M., Näätänen, R., 2006. Selective attention to human voice enhances brain activity bilaterally in the superior temporal sulcus. Brain Res. 1075, 142–150.

Altmann, C.F., Henning, M., Döring, M.K., Kaiser, J., 2008. Effects of feature-selective attention on auditory pattern and location processing. NeuroImage 41, 69–79.

Anderson, S., Kraus, N., 2011. Sensory-Cognitive interaction in the neural encoding of speech in noise: a review. J. Am. Acad. Audiol. 21, 575–585.

Asari, H., Pearlmutter, B.A., Zador, A.M., 2006. Sparse representations for the cocktail party problem. J. Neurosci. 26, 7477–7490.

Astheimer, L.B., Sanders, L.D., 2009. Listeners modulate temporally selective attention during natural speech processing. Biol. Psychol. 80, 23–34.

Astheimer, L.B., Sanders, L.D., 2011. Predictability affects early perceptual processing of word onsets in continuous speech. Neuropsychologia 49, 3512–3516.

Astheimer, L.B., Sanders, L.D., 2012. Temporally selective attention supports speech processing in 3- to 5-year-old children. Dev. Cogn. Neurosci. 2, 120–128.

Atiani, S., Elhilali, M., David, S.V., Fritz, J.B., Shamma, S.A., University of M, 2009. Task difficulty and performance induce diverse adaptive patterns in gain and shape of primary auditory cortical receptive fields. Neuron 61, 467–480.

Aubin, T., 2004. Penguins and their noisy world. An. Acad. Bras. Ciênc. 76, 279–283.

Aubin, T., Jouventin, P., 1998. Cocktail-party effect in king penguin colonies. Philos. Trans. R. Soc. Lond. B Biol. Sci. 265, 1665–1673.

Bakin, J.S., Weinberger, N.M., 1996. Induction of a physiological memory in the cerebral cortex by stimulation of the nucleus basalis. Proc. Natl. Acad. Sci. U. S. A. 93, 11219–11224.

Barbas, H., Bunce, J.G., Medella, M., 2013. Prefrontal pathways that control attention. In: Stuss, D.T., Knight, R.T. (Eds.), Principles of Frontal Lobe Function. Oxford University Press, Oxford, pp. 31–48.

Barlow, J.S., 1969. Some observations on the electrophysiology of timing in the nervous system. Electroencephalogr. Clin. Neurophysiol. 27, 545.

Bashford, J.A., Warren, R.M., 1987. Multiple phonemic restorations follow the rules for auditory induction. Percept. Psychophys. 42, 114–121.

Bashford, J.A., Riener, K.R., Warren, R.M., 1992. Increasing the intelligibility of speech through multiple phonemic restorations. Percept. Psychophys. 51, 211–217.

Bee, M.A., 2008. Finding a mate at a cocktail party: spatial release from masking improves acoustic mate recognition in grey treefrogs. Anim. Behav. 75, 1781–1791.

Belin, P., Zatorre, R.J., Lafaille, P., Ahad, P., Pike, B., 2000. Voice-selective areas in human auditory cortex. Nature 403, 309–312.

Benson, D.A., Hienz, R.D., 1978. Single-unit activity in the auditory cortex of monkeys selectivity attending left vs. right ear stimuli. Brain Res. 159, 307–320.

Besle, J., Schevon, C.A., Mehta, A.D., Lakatos, P., Goodman, R.R., McKhann, G.M., Emerson, R.G., Schroeder, C.E., 2011. Tuning of the human neocortex to the temporal dynamics of attended events. J. Neurosci. 31, 3176–3185.

Bidet-Caulet, A., Mikyska, C., Knight, R.T., 2010. Load effects in auditory selective attention: evidence for distinct facilitation and inhibition mechanisms. NeuroImage 50, 277–284.

Billimoria, C.P., Kraus, B.J., Narayan, R., Maddox, R.K., Sen, K., 2008. Invariance and sensitivity to intensity in neural discrimination of natural sounds. J. Neurosci. 28, 6304–6308.

De Boer, J., Thornton, A.R.D., 2007. Effect of subject task on contralateral suppression of click evoked otoacoustic emissions. Hear Res. 233, 117–123.

Brechmann, A., Scheich, H., 2005. Hemispheric shifts of sound representation in auditory cortex with conceptual listening. Cereb. Cortex 15, 578–587.

Bregman, A.S., 1978. Auditory streaming is cumulative. J. Exp. Psychol. Hum. Percept. Perform. 4, 380–387.

Bregman, A.S., 1990. Auditory Scene Analysis: the Perceptual Organization of Sound. MIT Press.

Bregman, A.S., Rudnicky, A.I., 1975. Auditory segregation: stream or streams? J. Exp. Psychol. Hum. Percept. Perform. 1, 263–267.

Brokx, J.P.L., Nooteboom, S.G., 1982. Intonation and the perceptual separation of simultaneous voices. J. Phon. 10, 23–36.

Buschman, T.J., Miller, E.K., 2007. Top–down versus bottom–up control of attention in the prefrontal and posterior parietal cortices. Science 315, 1860–1862.

Caporello, E.L., Gentner, T.Q., 2012. Selective attention to natural auditory streams in single cortical neurons of the European starling. In: Society for Neuroscience Meeting, (New Orleans, LA), p. 368.29.

Carlyon, R.P., Cusack, R., Foxton, J.M., Robertson, I.H., 2001. Effects of attention and unilateral neglect on auditory stream segregation. J. Exp. Psychol. Hum. Percept. Perform. 27, 115–127.

Carlyon, R.P., Plack, C.J., Fantini, D.A., Cusack, R., 2003. Cross-modal and non-sensory influences on auditory streaming. Perception 32, 1393–1402.

Cate, A.D., Herron, T.J., Yund, E.W., Stecker, G.C., Rinne, T., Kang, X., Petkov, C.I., Disbrow, E.A., Woods, D.L., 2009. Auditory attention activates peripheral visual cortex. PloS One 4, e4645.

Cherry, E.C., 1953. Some experiments on the recognition of speech, with one and two ears. J. Acoust. Soc. Am. 25, 975–979.

Coch, D., Sanders, L.D., Neville, H.J., 2005. An event-related potential study of selective auditory attention in children and adults. J. Cogn. Neurosci. 17, 605–622.

Crone, N.E., Boatman, D., Gordon, B., Hao, L., 2001. Induced electrocorticographic gamma activity during auditory perception. Clin. Neurophysiol. 112, 565–582.

Darwin, C.J., 1997. Auditory grouping. Trends Cogn. Sci. 1, 327–333.

Darwin, C.J., Hukin, R.W., 2000a. Effectiveness of spatial cues, prosody, and talker characteristics. J. Acoust. Soc. Am. 107, 970–977.

Darwin, C.J., Hukin, R.W., 2000b. Effects of reverberation on spatial, prosodic, and vocal-tract size cues to selective attention. J. Acoust. Soc. Am. 108, 335–342.

David, S.V., Fritz, J.B., Shamma, S.A., University of M, 2012. Task reward structure shapes rapid receptive field plasticity in auditory cortex. Proc. Natl. Acad. Sci. U. S. A. 109, 2144–2149.

Debener, S., Herrmann, C.S., Kranczioch, C., Gembris, D., Engel, A.K., 2003. Top–down attentional processing enhances auditory evoked gamma band activity. NeuroReport 14, 683–686.

Degerman, A., Rinne, T., Salmi, J., Salonen, O., Alho, K., 2006. Selective attention to sound location or pitch studied with fMRI. Brain Res. 1077, 123–134.

Degerman, A., Rinne, T., Pekkola, J., Autti, T., Jääskeläinen, I.P., Sams, M., Alho, K., 2007. Human brain activity associated with audiovisual perception and attention. NeuroImage 34, 1683–1691.

Desimone, R., Duncan, J., 1995. Neural mechanisms of selective visual attention. Annu. Rev. Neurosci. 18, 193–222.

Ding, N., Simon, J.Z., 2012. Neural coding of continuous speech in auditory cortex during monaural and dichotic listening. J. Neurophysiol. 107, 78–89.

Doesburg, S.M., Green, J.J., McDonald, J.J., Ward, L.M., 2012. Theta modulation of inter-regional gamma synchronization during auditory attention control. Brain Res. 1431, 77–85.

Downar, J., Crawley, A.P., Mikulis, D.J., Davis, K.D., 2000. A multimodal cortical network for the detection of changes in the sensory environment. Nat. Neurosci. 3, 277–283.

Duncan, J., 2001. An adaptive coding model of neural function in prefrontal cortex. Nat. Rev. Neurosci. 2, 820–829.

Falkenberg, L.E., Specht, K., Westerhausen, R., 2011. Attention and cognitive control networks assessed in a dichotic listening fMRI study. Brain Cogn. 76, 276–285.

Fries, P., 2001. Modulation of oscillatory neuronal synchronization by selective visual attention. Science 291, 1560–1563.

Fritz, J.B., Shamma, S.A., University of M, Elhilali, M., Klein, D.J., 2003. Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. Nat. Neurosci. 6, 1216–1223.

Fritz, J.B., Elhilali, M., Shamma, S.A., University of M, 2005. Differential dynamic plasticity of A1 receptive fields during multiple spectral tasks. J. Neurosci. 25, 7623–7635.

Fritz, J.B., Elhilali, M., David, S.V., Shamma, S.A., University of M, 2007a. Auditory attention – focusing the searchlight on sound. Curr. Opin. Neurobiol. 17, 437–455.

Fritz, J.B., Elhilali, M., Shamma, S.A., University of M, 2007b. Adaptive changes in cortical receptive fields induced by attention to complex sounds. J. Neurophysiol. 98, 2337–2346.

Fritz, J.B., David, S.V., Radtke-Schuller, S., Yin, P., Shamma, S.A., University of M, 2010. Adaptive, behaviorally gated, persistent encoding of task-relevant auditory information in ferret frontal cortex. Nat. Neurosci. 13, 1011–1019.

Fujiwara, N., Nagamine, T., Imai, M., Tanaka, T., Shibasaki, H., 1998. Role of the primary auditory cortex in auditory selective attention studied by whole-head neuromagnetometer. Cogn. Brain Res. 7, 99–109.

Galindo-Leon, E.E., Lin, F.G., Liu, R.C., 2009. Inhibitory plasticity in a lateral band improves cortical detection of natural vocalizations. Neuron 62, 705–716.

Gentner, T.Q., Margoliash, D., 2003. Neuronal populations and single cells representing learned auditory objects. Nature 424, 669–674.

Gentner, T.Q., Hulse, S.H., Ball, G.F., 2004. Functional differences in forebrain auditory regions during learned vocal recognition in songbirds. J. Comp. Physiol. A. Neuroethol. Sens. Neural. Behav. Physiol. 190, 1001–1010.

Giard, M.-H., Perrin, F., Pernier, J., Peronnet, F., 1988. Several attention-related wave forms in auditory areas: a topographic study. Electroencephalogr. Clin. Neurophysiol. 69, 371–384.

Giesbrecht, B., Woldorff, M.G., Song, A.W., Mangun, G.R., 2003. Neural mechanisms of top–down control during spatial and feature attention. NeuroImage 19, 496–512.

Gill, P.R., Woolley, S.M.N., Fremouw, T.E., Theunissen, F.E., 2008. What's that sound? Auditory area CLM encodes stimulus surprise, not intensity or intensity changes. J. Neurophysiol. 99, 2809–2820.

Gold, J.I., Shadlen, M.N., 2007. The neural basis of decision making. Annu. Rev. Neurosci. 30, 535–574.

Grace, J.A., Amin, N., Singh, N.C., Theunissen, F.E., 2003. Selectivity for conspecific song in the zebra finch auditory forebrain. J. Neurophysiol. 89, 472–487.

Grady, C.L., Meter, J.W. Van, Maisog, J.M., Pietrini, P., Krasuski, J., Rauschecker, J.P., 1997. Modulation of activity in primary and secondary auditory cortex. NeuroReport 8, 2511–2516.

Gregoriou, G.G., Gotts, S.J., Zhou, H., Desimone, R., 2009. High-frequency, long-range coupling between prefrontal and visual cortex during attention. Science (New York, N.Y.) 324, 1207–1210.

Gröger, U., Wiegrebe, L., 2006. Classification of human breathing sounds by the common vampire bat, *Desmodus rotundus*. BMC Biol. 4, 18.

Hannemann, R., Obleser, J., Eulitz, C., 2007. Top–down knowledge supports the retrieval of lexical information from degraded speech. Brain Res. 1153, 134–143.

Hansen, J.C., Hillyard, S.A., 1978. Endogenous brain potentials associated with selective auditory attention. Electroencephalogr. Clin. Neurophysiol., 277–290.

Hari, R., Hfimfilfiinen, M., Kaukoranta, E., Mfikelfi, J., Tiihonen, J., 1989. Selective listening modifies activity of the human cortex. Exp. Brain Res. 74, 463–470.

Headley, D.B., Weinberger, N.M., 2011. Gamma-band activation predicts both associative memory and cortical plasticity. J. Neurosci. 31, 12748–12758.

Hickok, G., Poeppel, D., 2007. The cortical organization of speech processing. Nat. Rev. Neurosci. 8, 393–402.

Hill, K.T., Miller, L.M., 2010. Auditory attentional control and selection during cocktail party listening. Cereb. Cortex 20, 583–590.

Hillyard, S.A., Hink, R.F., Schwent, V.L., Picton, T.W., 1973. Electrical signs of selective attention in the human brain. Science 182, 177–180.

Hillyard, S.A., Vogel, E.K., Luck, S.J., 1998. Sensory gain control (amplification) as a mechanism of selective attention: electrophysiological and neuroimaging evidence. Philos. Trans. R. Soc. Lond. B Biol. Sci. 353, 1257–1270.

Hocherman, S., Benson, D.A., Goldstein, M.H., Heffner, H.E., Hienz, R.D., 1976. Evoked unit activity in auditory cortex of monkeys performing a selective attention task. Brain Res. 117, 51–68.

Huang, S., Belliveau, J.W., Tengshe, C., Ahveninen, J., 2012. Brain networks of novelty-driven involuntary and cued voluntary auditory attention shifting. PloS One 7, e44062.

Hubel, D.H., Henson, C.O., Rupert, A., Galambos, R., Series, N., May, N., 1959. “Attention” units in the auditory cortex. Science 129, 1279–1280.

Hugdahl, K., Thomsen, T., Ersland, L., Morten Rimol, L., Niemi, J., 2003. The effects of attention on speech perception: an fMRI study. Brain Lang. 85, 37–48.

Hugdahl, K., Westerhausen, R., Alho, K., Medvedev, S.V., Laine, M., Hämäläinen, H., 2009. Attention and cognitive control: unfolding the dichotic listening story. Scand. J. Psychol. 50, 11–22.

Hulse, S.H., MacDougall-Shackleton, S.A., Wisniewski, A.B., 1997. Auditory scene analysis by songbirds: stream segregation of birdsong by European starlings (*Sturnus vulgaris*). J. Comp. Psychol. 111, 3–13.

Jääskeläinen, I.P., Ahveninen, J., Belliveau, J.W., Raij, T., Sams, M., 2007. Short-term plasticity in auditory cognition. Trends Neurosci. 30, 653–661.

Jääskeläinen, I.P., Ahveninen, J., Andermann, M.L., Belliveau, J.W., Raij, T., Sams, M., 2011. Short-term plasticity as a neural mechanism supporting memory and attentional functions. Brain Res. 1422, 66–81.

James, W., 1890. The Principles of Psychology. H. Holt and Company.

Jäncke, L., Mirzazade, S., Shah, N.J., 1999. Attention modulates activity in the primary and the secondary auditory cortex: a functional magnetic resonance imaging study in human subjects. Neurosci. Lett. 266, 125–128.

Jäncke, L., Buchanan, T.W., Lutz, K., Shah, N.J., 2001. Focused and nonfocused attention in verbal and emotional dichotic listening: an FMRI study. Brain Lang. 78, 349–363.

Jensen, O., Kaiser, J., Lachaux, J.-P., 2007. Human gamma-frequency oscillations associated with attention and memory. Trends Neurosci. 30, 317–324.

Johnson, J.A., Zatorre, R.J., 2005. Attention to simultaneous unrelated auditory and visual events: behavioral and neural correlates. Cereb. Cortex 15, 1609–1620.

Johnson, J.A., Zatorre, R.J., 2006. Neural substrates for dividing and focusing attention between simultaneous auditory and visual events. NeuroImage 31, 1673–1681.

Kaltenbach, J.A., 2006. The dorsal cochlear nucleus as a participant in the auditory, attentional and emotional components of tinnitus. Hear Res. 216–217, 224–234.

Kauramäki, J., Jääskeläinen, I.P., Sams, M., 2007. Selective attention increases both gain and feature selectivity of the human auditory cortex. PloS One 2, e909.

Kerlin, J.R., Shahin, A.J., Miller, L.M., 2010. Attentional gain control of ongoing cortical speech representations in a “cocktail party”. J. Neurosci. 30, 620–628.

Kiggins, J.T., Comins, J.A., Gentner, T.Q., 2012. Targets for a comparative neurobiology of language. Front. Evolut. Neurosci. 4, 6.

Kilgard, M.P., Merzenich, M.M., 1998a. Cortical map reorganization enabled by nucleus basalis activity. Science 279, 1714–1718.

Kilgard, M.P., Merzenich, M.M., 1998b. Plasticity of temporal information processing in the primary auditory cortex. Nat. Neurosci. 1, 727–731.

Kincade, J.M., Abrams, R.A., Astafiev, S.V., Shulman, G.L., Corbetta, M., 2005. An event-related functional magnetic resonance imaging study of voluntary and stimulus-driven orienting of attention. J. Neurosci. 25, 4593–4604.

Klinke, R., Fruhstorfer, H., Finkenzeller, P., 1968. Evoked responses as a function of external and stored information. Electroencephalogr. Clin. Neurophysiol. 25, 119–122.

Knight, R.T., Hillyard, S.A., Woods, D.L., Neville, H.J., 1981. The effects of frontal cortex lesions on event-related potentials during auditory selective attention. Electroencephalogr. Clin. Neurophysiol. 52, 571–582.

Knight, R.T., Staines, W.R., Swick, D., Chao, L.L., 1999. Prefrontal cortex regulates inhibition and excitation in distributed neural networks. Acta. Psychol. 101, 159–178.

Knudsen, D.P., Gentner, T.Q., 2010. Mechanisms of song perception in oscine birds. Brain Lang. 115, 59–68.

Kompus, K., Specht, K., Ersland, L., Juvodden, H.T., Van Wageningen, H., Hugdahl, K., Westerhausen, R., 2012. A forced-attention dichotic listening fMRI study on 113 subjects. Brain Lang. 121, 240–247.

Kong, L., Michalka, S.W., Rosen, M.L., Sheremata, S.L., Swisher, J.D., Shinn-Cunningham, B.G., Somers, D.C., 2012. Auditory spatial attention representations in the human cerebral cortex. Cereb. Cortex, 1–12.

Lewis, J.L., 1970. Semantic processing of unattended messages using dichotic listening. J. Exp. Psychol. 85, 225–228.

Ma, L., Micheyl, C., Yin, P., Oxenham, A.J., Shamma, S.A., University of M, 2010. Behavioral measures of auditory streaming in ferrets (*Mustela putorius*). J. Comp. Psychol. 124, 317–330.

MacDougall-Shackleton, S.A., Hulse, S.H., Gentner, T.Q., White, W., 1998. Auditory scene analysis by European starlings (*Sturnus vulgaris*): perceptual segregation of tone sequences. J. Acoust. Soc. Am. 103, 3581–3587.

Macken, W.J., Tremblay, S., Houghton, R.J., Nicholls, A.P., Jones, D.M., 2003. Does auditory streaming require attention? Evidence from attentional selectivity in short-term memory. J. Exp. Psychol. Hum. Percept. Perform. 29, 43–51.

Maeder, P.P., Meuli, R.A., Adriani, M., Thiran, A.B., Fornari, E., Thiran, J.P., Pittet, A., Clarke, S., 2001. Distinct pathways involved in sound recognition and localization: a human fMRI study. NeuroImage 14, 802–816.

Maison, S., Micheyl, C., Collet, L., 2001. Influence of focused auditory attention on cochlear activity in humans. Psychophysiology 38, 35–40.

Matsumoto, K., Tanaka, K., 2004. The role of the medial prefrontal cortex in achieving goals. Curr. Opin. Neurobiol. 14, 178–185.

Meliza, C.D., Margoliash, D., 2012. Emergence of selectivity and tolerance in the avian auditory cortex. J. Neurosci. 32, 15158–15168.

Mesgarani, N., Chang, E.F., 2012. Selective cortical representation of attended speaker in multi-talker speech perception. Nature 485, 1–5.

Mesgarani, N., David, S.V., Fritz, J.B., Shamma, S.A., University of M, 2008. Phoneme representation and classification in primary auditory cortex. J. Acoust. Soc. Am. 123, 899–909.

Mesgarani, N., Fritz, J.B., Shamma, S.A., University of M, 2010. A computational model of rapid task-related plasticity of auditory cortical receptive fields. J. Comput. Neurosci. 28, 19–27.

Metherate, R., Weinberger, N.M., 1989. Acetylcholine produces stimulus-specific receptive field alterations in cat auditory cortex. Brain Res. 480, 372–377.

Metherate, R., Intskirveli, I., Kawai, H.D., 2012. Nicotinic filtering of sensory processing in auditory cortex. Front. Behav. Neurosci. 6, 44.

Miller, E.K., Cohen, J.D., 2001. An integrative theory of prefrontal cortex function. Annu. Rev. Neurosci. 24, 167–202.

Miller, J.M., Sutton, D., Pfingst, B.E., Ryan, A.F., Beaton, R.D., Gourevitch, G., 1972. Single cell activity in the auditory cortex of rhesus monkeys: behavioral dependency. Science 177, 449–451.

Moore, B.C.J., Gockel, H.E., 2012. Properties of auditory stream formation. Philos. Trans. R. Soc. Lond. B Biol. Sci. 367, 919–931.

Moore, R.C., Lee, T., Theunissen, F.E., 2013. Noise-invariant neurons in the avian auditory cortex: hearing the song in noise. PLoS Comput. Biol. 9, e1002942.

Moray, N., 1959. Attention in dichotic listening: affective cues and the influence of instructions. J. Exp. Psychol. 9, 56–60.

Näätänen, R., 1992. Attention and Brain Function. Lawrence Erlbaum Associates, Hillsdale, NJ.

Näätänen, R., Michie, P.T., 1979. Early selective-attention effects on the evoked potential: a critical review and reinterpretation. Biol. Psychol. 8, 81–136.

Narayan, R., Best, V., Ozmeral, E.J., McClaine, E., Dent, M., Shinn-Cunningham, B.G., Sen, K., 2007. Cortical interference effects in the cocktail party problem. Nat. Neurosci. 10, 1601–1607.

Neelon, M.F., Williams, J., Garell, P.C., 2011. Elastic attention: enhanced, then sharpened response to auditory input as attentional load increases. Front. Human Neurosci. 5, 41.

Nityananda, V., Bee, M.A., 2011. Finding your mate at a cocktail party: frequency separation promotes auditory stream segregation of concurrent voices in multi-species frog choruses. PloS One 6, e21191.

Nygaard, L.C., Sommers, M.S., Pisoni, D.B., 1994. Speech perception as a talker-contingent process. Psychol. Sci. 5, 42–46.

O’Leary, D.S., Andreason, N.C., Hurtig, R.R., Hichwa, R.D., Watkins, G.L., Ponto, L.L., Rogers, M., Kirchner, P.T., 1996. A positron emission tomography study of binaurally and dichotically presented stimuli: effects of level of language and directed attention. Brain Lang. 53, 20–39.

Okamoto, H., Stracke, H., Wolters, C.H., Schmael, F., Pantev, C., 2007. Attention improves population-level frequency tuning in human auditory cortex. J. Neurosci. 27, 10383–10390.

Okamoto, H., Stracke, H., Bermudez, P., Pantev, C., 2011. Sound processing hierarchy within human auditory cortex. J. Cogn. Neurosci. 23, 1855–1863.

Otazu, G.H., Tai, L.-H., Yang, Y., Zador, A.M., 2009. Engaging in an auditory task suppresses responses in auditory cortex. Nat. Neurosci. 12, 646–654.

Paavilainen, P., Saarinen, J., Tervaniemi, M., Näätänen, R., 1995. Mismatch negativity to changes in abstract sound features during dichotic listening. J. Psychophysiol. 9, 243–249.

Paltoglou, A.E., Sumner, C.J., Hall, D.A., 2009. Examining the role of frequency specificity in the enhancement and suppression of human cortical activity by auditory selective attention. Hear Res. 257, 106–118.

Paltoglou, A.E., Sumner, C.J., Hall, D.A., 2011. Mapping feature-sensitivity and attentional modulation in human auditory cortex with functional magnetic resonance imaging. Eur. J. Neurosci. 33, 1733–1741.

Pasley, B.N., David, S.V., Mesgarani, N., Flinker, A., Shamma, S.A., University of M, Crone, N.E., Knight, R.T., Chang, E.F., 2012. Reconstructing speech from human auditory cortex. PLoS Biol. 10, e1001251.

Paus, T., Zatorre, R.J., Hofle, N., Caramanos, Z., Gotman, J., Petrides, M., Evans, A.C., 1997. Time-related changes in neural systems underlying attention and arousal during the performance of an auditory vigilance task. J. Cogn. Neurosci. 9, 392–408.

Petkov, C.I., O’Connor, K.N., Sutter, M.L., 2003. Illusory sound perception in macaque monkeys. J. Neurosci. 23, 9155–9161.

Petkov, C.I., Kang, X., Alho, K., Bertrand, O., Yund, E.W., Woods, D.L., 2004. Attentional modulation of human auditory cortex. Nat. Neurosci. 7, 658–663.

Petkov, C.I., O’Connor, K.N., Sutter, M.L., 2007. Encoding of illusory continuity in primary auditory cortex. Neuron 54, 153–165.

Pfingst, B.E., Connor, T.A.O., Miller, J.M., 1977. Brain response plasticity of neurons in auditory cortex of the rhesus monkey. Exp. Brain Res. 29, 393–404.

Picton, T.W., Hillyard, S.A., 1974. Human auditory evoked potentials. II: effects of attention. Electroencephalogr. Clin. Neurophysiol. 36, 191–199.

Pienkowski, M., Eggermont, J.J., 2011. Cortical tonotopic map plasticity and behavior. Neurosci. Biobehav. Rev. 35, 2117–2128.

Pisoni, D.B., Nusbaum, H.C., Luce, P.A., Slowiaczek, L.M., 1985. Speech perception, word recognition and the structure of the lexicon. Speech Commun. 4, 75–95.

Pritchard, W.S., 1981. Psychophysiology of P300. Psychol. Bull. 89, 506–540.

Rämä, P., Relander-Syrjänen, K., Carlson, S., Salonen, O., Kujala, T., 2012. Attention and semantic processing during speech: an fMRI study. Brain Lang. 122, 114–119.

Recanzone, G.H., Schreiner, C.E., Merzenich, M.M., 1993. Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. J. Neurosci. 13, 87–103.

Reynolds, J.H., Heeger, D.J., 2009. The normalization model of attention. Neuron 61, 168–185.

Rif, J., Hari, R., Matti, S.H., Sams, M., 1991. Auditory attention affects two different areas in the human supratemporal cortex. Electroencephalogr. Clin. Neurophysiol. 79, 464–472.

Rihs, T.A., Michel, C.M., Thut, G., 2009. A bias for posterior alpha-band power suppression versus enhancement during shifting versus maintenance of spatial attention. NeuroImage 44, 190–199.

Rinne, T., Kirjavainen, S., Salonen, O., Degerman, A., Kang, X., Woods, D.L., Alho, K., 2007. Distributed cortical networks for focused auditory attention and distraction. Neurosci. Lett. 416, 247–251.

Rinne, T., Balk, M.H., Koistinen, S., Autti, T., Alho, K., Sams, M., 2008. Auditory selective attention modulates activation of human inferior colliculus. J. Neurophysiol. 100, 3323–3327.

Rodgers, C., DeWeese, M.R., 2012. Neuronal correlates of selective auditory attention in rat prefrontal and sensory cortices. In: Society for Neuroscience Meeting, (New Orleans, LA), p. 599.20.

Ryan, A.F., Miller, J.M., Pfingst, B.E., Martin, G.K., 1984. Effects of reaction time performance on single-unit activity in the central auditory pathway of the rhesus macaque. J. Neurosci. 4, 298–308.

Sagi, B., Nemat-Nasser, S.C., Kerr, R., Hayek, R., Downing, C., Hecht-Nielsen, R., 2001. A biologically motivated solution to the cocktail party problem. J. Comput. Neurosci. 13, 1575–1602.

Salmi, J., Rinne, T., Degerman, A., Salonen, O., Alho, K., 2007. Orienting and maintenance of spatial attention in audition and vision: multimodal and modality-specific brain activations. Brain Struct. Funct. 212, 181–194.

Salmi, J., Rinne, T., Koistinen, S., Salonen, O., Alho, K., 2009. Brain networks of bottom–up triggered and top–down controlled shifting of auditory attention. Brain Res. 1286, 155–164.

Sanders, L.D., Newport, E.L., Neville, H.J., 2002. Segmenting nonsense: an event-related potential index of perceived onsets in continuous speech. Nat. Neurosci. 5, 700–703.

Sauseng, P., Klimesch, W., Stadler, W., Schabus, M., Doppelmayr, M., Hanslmayr, S., Gruber, W.R., Birbaumer, N., 2005. A shift of visual spatial attention is selectively associated with human EEG alpha activity. Eur. J. Neurosci. 22, 2917–2926.

Schmidt, A.K.D., Römer, H., 2011. Solutions to the cocktail party problem in insects: selective filters, spatial release from masking and gain control in tropical crickets. PloS One 6, e28593.

Seeba, F., Schwartz, J.J., Bee, M.A., 2010. Testing an auditory illusion in frogs: perceptual restoration or sensory bias? Anim. Behav. 79, 1317–1328.

Shomstein, S., Yantis, S., 2006. Parietal cortex mediates voluntary control of spatial and nonspatial auditory attention. J. Neurosci. 26, 435–439.

Siegel, M., Donner, T.H., Oostenveld, R., Fries, P., Engel, A.K., 2008. Neuronal synchronization along the dorsal visual pathway reflects the focus of spatial attention. Neuron 60, 709–719.

Smith, D.W., Aouad, R.K., Keil, A., 2012. Cognitive task demands modulate the sensitivity of the human cochlea. Front. Psychol. 3, 30.

Snyder, J.S., Alain, C., Picton, T.W., 2006. Effects of attention on neuroelectric correlates of auditory stream segregation. J. Cogn. Neurosci. 18, 1–13.

Snyder, J.S., Gregg, M.K., Weintraub, D.M., Alain, C., 2012. Attention, awareness, and the perception of auditory scenes. Front. Psychol. 3, 15.

Sokolov, A., Pavlova, M., Lutzenberger, W., Birbaumer, N., 2004. Reciprocal modulation of neuromagnetic induced gamma activity by attention in the human visual and auditory cortex. NeuroImage 22, 521–529.

Spieth, W., Curtis, J.F., Webster, J.C., 1954. Responding to one of two simultaneous messages. J. Acoust. Soc. Am. 26, 391–396.

Suga, N., Ma, X., 2003. Multiparametric corticofugal modulation and plasticity in the auditory system. Nat. Rev. Neurosci. 4, 783–794.

Sussman, E.S., Bregman, A.S., Wang, W.J., Khan, F.J., 2005. Attentional modulation of electrophysiological activity in auditory cortex for unattended sounds within multistream auditory environments. Cogn. Affect. Behav. Neurosci. 5, 93–110.

Tallon-Baudry, C., Bertrand, O., Delpuech, C., Permier, J., 1997. Oscillatory gamma-band (30–70 Hz) activity induced by a visual search task in humans. J. Neurosci. 17, 722–734.

Tanedo, J.R., Bhagat, S.P., 2008. Attentional control of cochlear micromechanics: the role of the medial olivocochlear efferent system. Perspect. Hear Hear Disord. Res. Diagn. 12, 17–21.

Teshiba, T.M., Ling, J., Ruhl, D.A., Bedrick, B.S., Peña, A., Mayer, A.R., 2013. Evoked and intrinsic asymmetries during auditory attention: implications for the contralateral and neglect models of functioning. Cereb. Cortex (New York, N.Y.: 1991) 23, 560–569.

Thompson, S.K., Carlyon, R.P., Cusack, R., 2011. An objective measurement of the build-up of auditory streaming and of its modulation by attention. J. Exp. Psychol. Hum. Percept. Perform. 37, 1253–1262.

Tiitinen, H., Sinkkonen, J., Reinikainen, K., Alho, K., Lavikainen, J., Näätänen, R., 1993. Selective attention enhances the auditory 40-Hz transient response in humans. Nature 364, 59–60.

Treisman, A.M., 1960. Contextual cues in selective listening. Q. J. Exp. Psychol. 12, 242–248.

Treisman, A.M., Squire, R., Green, J., 1974. Semantic processing in dichotic listening? A replication. Mem. Cogn. 2, 641–646.

Vossel, S., Thiel, C.M., Fink, G.R., 2006. Cue validity modulates the neural correlates of covert endogenous orienting of attention in parietal and frontal cortex. NeuroImage 32, 1257–1264.

Warren, W.H., Verbrugge, R.R., 1984. Auditory perception of breaking and bouncing events: a case study in ecological acoustics. J. Exp. Psychol. Hum. Percept. Perform. 10, 704–712.

Westerhausen, R., Moosmann, M., Alho, K., Belsby, S.-O., Hämäläinen, H., Medvedev, S.V., Specht, K., Hugdahl, K., 2010. Identification of attention and cognitive control networks in a parametric auditory fMRI study. Neuropsychologia 48, 2075–2081.

Winer, J.A., 2005. Decoding the auditory corticofugal systems. Hear Res. 207, 1–9.

Winkler, I., Czigler, I., Sussman, E.S., Horváth, J., Balázs, L., 2005. Preattentive binding of auditory and visual stimulus features. J. Cogn. Neurosci. 17, 320–339.

Woldorff, M.G., Hillyard, S.A., 1991. Modulation of early auditory processing during selective listening to rapidly presented tones. Electroencephalogr. Clin. Neurophysiol. 79, 170–191.

Woldorff, M.G., Hansen, J.C., Hillyard, S.A., 1987. Evidence for effects of selective attention in the mid-latency range of the human auditory event-related potential. Electroencephalogr. Clin. Neurophysiol. Suppl. 40, 146–154.

Woldorff, M.G., Gallen, C.C., Hampson, S.A., Hillyard, S.A., Pantev, C., Sobel, D., Bloom, F.E., 1993. Modulation of early sensory processing in human auditory cortex during auditory selective attention. Proc. Natl. Acad. Sci. U. S. A. 90, 8722–8726.

Wood, N.L., Cowan, N., 1995. The cocktail party phenomenon revisited: attention and memory in the classic selective listening procedure of Cherry (1953). J. Exp. Psychol. Gen. 124, 243–262.

Woods, D.L., Alain, C., 2001. Conjoining three auditory features: an event-related brain potential study. J. Cogn. Neurosci. 13, 492–509.

Woods, D.L., Alain, C., Diaz, R., Rhodes, D., Ogawa, K.H., 2001. Location and frequency cues in auditory selective attention. J. Exp. Psychol. 27, 65–74.

Woods, D.L., Stecker, G.C., Rinne, T., Herron, T.J., Cate, A.D., Yund, E.W., Liao, I., Kang, X., 2009. Functional maps of human auditory cortex: effects of acoustic features and attention. PloS One 4, e5183.

Woolley, S.C., Doupe, A.J., 2008. Social context-induced song variation affects female behavior and gene expression. PLoS Biol. 6, e62.

Wu, C.-T., Weissman, D.H., Roberts, K.C., Woldorff, M.G., 2007. The neural circuitry underlying the executive control of auditory spatial attention. Brain Res. 1134, 187–198.

Yantis, S., 1998. Control of visual attention. In: Pashler, H.E. (Ed.), Attention. Psychology Press, pp. 223–292.

Yin, P., Mishkin, M., Sutter, M.L., Fritz, J.B., 2008. Early stages of melody processing: stimulus-sequence and task-dependent neuronal activity in monkey auditory cortical fields A1 and R. J. Neurophysiol. 100, 3009–3029.

Yin, P., Fritz, J.B., Shamma, S.A., University of M, 2012. Responses of ferret frontal cortex neurons in a long-term auditory memory task. In: Society for Neuroscience Meeting, (New Orleans, LA), p. 460.29.

Zatorre, R.J., Mondor, T.A., Evans, A.C., 1999. Auditory attention to space and frequency activates similar cerebral systems. NeuroImage 10, 544–554.