Temporal processing in the auditory core: transformation or segregation?

Elliot H. Smith

Department of Bioengineering, University of Utah, Salt Lake City, Utah

Submitted 16 March 2011; accepted in final form 26 July 2011

Smith EH. Temporal processing in the auditory core: transformation or segregation? J Neurophysiol 106: 2791–2793, 2011. First published July 27, 2011; doi:10.1152/jn.00233.2011.—Scott et al. (J Neurophysiol 105: 712–730, 2011) examine temporal processing in auditory Core fields AI and R in the macaque, arguing that temporal properties are the most salient functional differences between these areas. How these results imply transformation of information between AI and R or whether these fields may be processing divergent information somewhat separately is discussed.

auditory cortex; processing stream; transformation of information

The classic view of primary auditory cortex is of a “whiteboard” (Nelken 2008), containing a complete, yet deconstructed, representation of an incoming soundwave. This information is then synthesized through the auditory cortical circuit to form auditory perception. It has since become clear that neural responses in early auditory cortex are more than the sum of their spectrotemporal parts, being influenced by stimulus or task context (Ghazanfar et al. 2005; Otazu et al. 2009). Since most deconstruction of acoustic information has been accomplished at the stage of the inferior colliculus (IC), the IC could be thought of as the auditory analog to primary visual cortex (Nelken 2008). This begs the question, what roles do early auditory cortical areas play in representing or in developing auditory perception, and how do these roles differ among neuroanatomically defined areas of auditory cortex? This neuroforum article examines how Scott et al. (2011) address this question by studying temporal processing among several early auditory cortical areas.

Macaque auditory cortex is separated into three main regions (Fig. 1): the core, belt, and parabelt. The core field contains three subfields, the primary auditory cortex (AI), the rostral field (R), and the rostrotemporal field (RT). These fields are arranged tonotopically, that is, by their selectivity to stimulus frequency. Reversals in tonotopic gradients have mainly been used to define borders between AI, R, and RT. The belt field contains eight subfields and the parabelt field consists of broadly defined rostral and caudal aspects (Hackett 2011). It has been hypothesized that auditory processing is organized into two streams, analogous to the dorsal and ventral processing streams in the visual system (Rauschecker and Scott 2009). That is, hierarchical auditory processing for sound location is directed caudally from the primary auditory cortex to the parietal lobe, and information about sound identity is directed rostrally along the superior temporal gyrus (STG). This hypothesis is supported by neuroanatomical studies showing that adjacent belt areas are more densely connected and mostly feed-forward (Hackett 2011). Anterior belt and parabelt then connect to anterior temporal lobe and ventrolateral prefrontal areas (Romanski et al. 1999).

There is also neuroanatomical evidence for a hierarchical processing stream extending along the superior temporal plane (STP), inside the lateral fissure (Galaburda and Pandya 1983). Most of the physiological evidence for hierarchical processing in this stream comes from studies of core and belt areas in the marmoset that show less temporal synchronization and more rate-based responses at successively rostral core areas (Bendor and Wang 2008; Wang et al. 2008). A recent study that sampled further along the STP characterized more rostral areas as exhibiting increasingly greater latencies and more inhibition than excitation, traits that are characteristic of areas further along in a processing stream (Kikuchi et al. 2010). Does this mean there are two “ventral” processing streams, one on the STG and one on the STP, or are these two streams elements of a larger processing stream for stimulus quality or identity?

Studies examining early stages of auditory cortical processing, such as Scott et al. (2011) and Rauschecker et al. (1997) in the macaque and Bendor and Wang (2008) in the marmoset, have led some researchers to hypothesize that temporal processing occurs hierarchically along the STP, whereas spectral elements of a sound are processed hierarchically from core to belt to parabelt (Bendor and Wang 2008). How these divergent streams or substreams integrate spectrotemporal features of sound stimuli from lower brain areas into complete auditory perceptions remains a fundamental question in auditory neurophysiology.

Scott et al. (2011) argue for the hypothesis that temporal elements of a sound are processed caudorostrally inside the lateral fissure. This study focuses on core subfields AI and R, placing the two areas in the context of a hierarchical processing stream that extends along the STP and providing preliminary evidence for how these subfields might function within that stream. The paper leaves the reader wanting additional discussion of what greater latencies mean, and how temporal synchronization of firing rates may transform into a rate code as information passes between cortical areas in a processing stream.

Scott et al. (2011) show that AI and R are ostensibly very similar, having similar selectivity for tone frequency and intensity, similar baseline, threshold, and driven firing rates, and similar, yet reversed, tonotopic organization. The salient difference between AI and R, argue the authors, is temporal processing. Showing that R has 13-ms greater median minimum latencies than AI, and that responses in R are less correlated with temporal fluctuations in the stimuli, supports this argument. Yet while timing is a salient difference in the change in processing between AI and R, the authors’ argument about temporal processing is accompanied by a description of R as having a “qualitatively different representation” from AI...
Instead of natural stimuli, the study uses sinusoidally am-plitude-modulated (SAM) and sinusoidally frequency-modu-lated (SFM) tones to address change in processing quality of stimuli that may closer mimic the modulations found in one area from the other. Examining the optimization factors they used to create the frozen noise (central frequency, sound pressure level, and bandwidth) would be a good first step towards understanding how AI and R respond differently to spectrotemporally similar stimuli. Examining the qualitative change in response to features of more spectrotemporally complex stimuli, as has been done with artificial or modified conspecific vocalizations (Rauschecker et al. 1995), could be a very informative next step in determining how temporal transformations occur for stimuli to which the auditory cortical system has evolved to respond.

Instead of natural stimuli, the study uses sinusoidally am-plitude-modulated (SAM) and sinusoidally frequency-modu-lated (SFM) tones to address change in processing quality between AI and R. SAM and SFM are used as simple, general-ized stimuli that may closer mimic the modulations found in natural stimuli. The authors describe these stimuli as “sufficient to identify gross differences in temporal processing” (Scott et al. 2011). Again, AI is better at synchronizing to these stimuli than R, although both areas modulate their firing rates in response to amplitude- and frequency-modulated tones. This result suggests that AI integrates sounds more quickly than R, providing support for a difference in temporal processing between the two areas. It is, however, difficult to see how this result indicates a change in the quality of processing, or a transformation of information, between areas.

While temporal processing is certainly a major difference between AI and R, there are differences between responses in the two areas that don’t accord with the theory that AI and R are linked in a processing hierarchy. The first is the difference in median minimum latency between the two areas. Scott et al. (2011) report a 13-ms difference in minimum latencies between AI and R, which is much longer than synaptic delay for adjacent cortical areas. Similar difference in minimum latency has been reported several times previously (Bendor and Wang 2008; Kusmierek and Rauschecker 2009; Recanzone et al. 2000). This difference in latency between AI and R has been argued to be indicative that the two areas are not in the same processing hierarchy (Kusmierek and Rauschecker 2009) and could indicate that a processing stream for temporal information diverges in the thalamus or lower brain areas. AI and R both receive thalamic input from the ventral division of the medial geniculate nucleus of the thalamus, although likely from different populations of cells (Winer and Lee 2007). This evidence for AI and R representing divergent streams of information is further supported by the result that ablating AI has little effect on R and even seems to enhance its response (Rauschecker et al. 1997).

Despite the 13-ms difference in latencies between AI and R, Scott et al. (2011) show a continuous gradient of latencies from the caudal aspect of AI to the rostral aspect of R. If timing is the defining characteristic between these two areas, the observation that latencies increase continuously along the caudorostral extent of the core auditory cortex, much as frequencies do in AI, might suggest that AI and R are not functionally distinct. Figure 9 from Scott et al. (2011) shows a continuous caudorostral increase in minimum latency from AI to R. While it is clear, however, that overall minimum latencies are greater in R, it is not clear that there is a caudorostral latency gradient. Regressions through both areas were not significant in one hemisphere and barely significant in another. Second, while minimum latencies appear greater in R, the variance among cells is clearly greater in R, which would increase aggregate latencies as measured by a regression line. It is also not clear whether slopes of regression lines for only the cells in R would mimic slopes from both areas or AI alone. Plotting regression lines for the two areas individually as well as regression lines for the areas together would have made explicit whether there is indeed a continuous latency gradient across these two areas, as opposed to increased variance in one area due to several cells with very long latency. Evidence against a caudorostral latency gradient would help account for the long delay between areas and help temporally disambiguate the areas as functionally distinct.

One way to observe information being transferred between cortical areas along a processing stream is for precise temporal coding of stimuli to change to rate-based coding at successive cortical areas (Romanski et al. 1999; Russ et al. 2008). Bendor and Wang (2008) show R and RT to rely on rate-based coding more than neurons in AI. Scott et al. (2011), however, report that fewer neurons in R have rate-based responses than those in
AI, contradicting the data from analogous marmoset areas. Scott et al. (2011) mention that this may be due to disparities in methods between Scott et al. and Bendor and Wang. This topic needs more scrutiny in the macaque, as an understanding of temporal transformations between areas should address what the transformation of information looks like, and ultimately how the transformation is occurring.

Scott et al. (2011) propose that the differences in latency and temporal integration between AI and R might serve to allow the auditory cortex to function on different time scales. While the authors claim the windows of temporal integration for AI and R are suitable for processing formant transitions and syllables, respectively, they provide little evidence for the claim. Furthermore, the large variability in synchrony cutoff shown in Fig. 12 (of Scott et al. 2011) doesn’t support the idea. This organization might be appropriate for hierarchical processing in human language areas, yet is difficult to apply to early auditory cortical areas in a monkey. Decodes of vocalizations in STG and auditory prefrontal neurons actually perform better with smaller windows of time, indicating that a larger temporal window may not be as important for processing increasingly longer sounds (Russ et al. 2008). As with vision, working memory likely maintains perceptual elements over periods of time necessary to understand longer auditory perceptions such as words or sentences. This organizing principle therefore need not be an element of auditory cortical processing, especially at early stages in a primate.

Scott et al. (2011) follow the work of Bendor and Wang (2008) and Wang et al. (2008), characterizing temporal responses in two cortical areas at the beginning of a processing hierarchy responsible for auditory object identity in the macaque. Several interesting questions arise regarding the long delay observed between fields and differences in synchronization capability between the two fields, as well as the lack of rate-based coding observed in the macaque that has been observed in the marmoset. Whether the results from Scott et al. (2011), Wang et al. (2008), Kusmierek and Rauschecker (2009), and Rauschecker et al. (1997) provide evidence for a transformation in information processing between AI and R or support somewhat independent processing in the two areas requires further scrutiny. Understanding how and where information is transferred in these cortical areas will be an integral element in understanding how a deconstructed neural representation is transformed into auditory perception.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

REFERENCES


