Latency of the Auditory Evoked Neuromagnetic Field Components: Stimulus Dependence and Insights Toward Perception

Timothy P.L. Roberts, Paul Ferrari, Steven M. Stufflebeam, and *David Poeppel

Biomagnetic Imaging Laboratory, Department of Radiology, University of California, San Francisco, California, and *Department of Linguistics, University of Maryland, College Park, Maryland, U.S.A.

Summary: This review will focus on investigations of the auditory evoked neuromagnetic field component, the M100, detectable in the magnetoencephalogram recorded during presentation of auditory stimuli, approximately 100 milliseconds after stimulus onset. In particular, the dependence of M100 latency on attributes of the stimulus, such as intensity, pitch and timbre will be discussed, along with evidence relating M100 latency observations to perceptual features of the stimuli. Comparison with investigation of the analogous electrical potential component, the N1, will be made. Parametric development of stimuli from pure tones through complex tones to speech elements will be made, allowing the influence of spectral pitch, virtual pitch and perceptual categorization to be delineated and suggesting implications for the role of such latency observations in the study of speech processing. The final section will deal with potential clinical applications offered by M100 latency measurements, as objective indices of normal and abnormal cortical processing. Key Words: Latency—Magnetoencephalography—Temporal encoding—Stimulus attributes—Perception—Auditory.

Over the past decade, technological advances in functional neuroimaging, particularly positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) have given rise to a plethora of studies mapping spatial organization of brain functional centers associated with sensory and cognitive processes. Notwithstanding the tremendous insights being gained by these endeavors, both in controlled neuroscientific paradigms and in clinical cases, on an individual basis, several considerations must be recognized. First, the physiologic basis for both fMRI and PET lies in the hemodynamic correlates of neuronal activity consequent to stimulation, or related to task performance, and thus the fMRI and PET measures are somewhat indirect assays of neuronal activity. Second, despite recent advances in event-related fMRI, hemodynamically based functional neuroimaging techniques tend to require either sustained activation responses (block design methods) or at least can not separate individual temporal components intrinsic to an evoked response. Third, and most important, in the context of this manuscript, PET and fMRI are limited in temporal resolution. While it has been demonstrated that deconvolution can in some cases allow separation of responses associated with separate stimulus events with short interstimulus intervals, the real temporal resolution of these methods is typically measured in hundreds of milliseconds (ms), or even seconds.

On the other hand, timing is known to be a critical
element of neuronal activity. There are several manifestations of this at various levels of electrophysiological study. Recent advances in systems neuroscience provide evidence for dynamic, temporal coding strategies that are thought necessary to supplement spatiotopic processing (Arieli et al., 1996; Yu and Margoliash, 1996). Popular new concepts include using the information (1) encoded in neuronal spike trains (Rieke et al., 1997), (2) encoded in oscillatory activity (e.g., 10Hz, 40Hz), or (3) encoded in latency shifts of specific event-related peaks (Eysel, 1996; Heil and Irvine, 1996; Roberts and Poeppel, 1996; Wörgötter et al., 1996). Particularly motivating for this review are findings that latency can be used as a dimension to encode stimulus attributes. For example, Wörgötter and colleagues (1996) have found that a larger temporal separation arising from stimulus-related latency differences (e.g., contrast in a visual stimulus) can help segregate information and aid discernability in vision. In auditory cortex, Heil has recently found (Heil and Irvine, 1996) that the first-spike latency of auditory cortical neurons is strongly determined by the spectral transition rate of the stimulus, and that latency might be used in such a context as a rate-of-change detector.

Imaging, or sensing, methods sensitive to the electrical activity of neurons during an evoked response may allow elucidation of some of these temporal features and thus provide an additional dimension of insight into brain function. As such EEG and its magnetic counterpart, magnetoencephalography (MEG), offer attractive avenues of investigation, with sensitivity to neuronal activity and submillisecond temporal resolution.

Attempts to combine spatial and temporal data from, for example, fMRI and MEG, will doubtless yield advances, on the one hand, in understanding the sequence of activation of multiple spatial sites indicated as “involved” by fMRI (i.e., use the time resolution of MEG to resolve the temporal sequence ambiguity of multiple areas of activity inherently time-averaged by the fMRI process) and, on the other hand, in improving strategies for modeling the source of the observed magnetoencephalogram, for example by exploiting the foci of activity identified by fMRI as putative sources in a multiple source algorithm such as MUSIC (Mosher et al., 1992; Liu et al., 1998).

However, it is the premise of this work that there is intrinsic value in the information content of the time domain of the observed evoked response activity (measured in the form of the electroencephalogram, or more particularly, the magnetoencephalogram) and that this information may reflect encoding of stimulus attributes and the groundwork for the formation of the sensory percept.

Data to support this argument will be drawn from a much studied evoked response component, namely that occurring approximately 100 ms after onset of auditory stimulation and described in the EEG literature as the N1, or in the MEG literature as the N1m, or alternatively the N100m, or simply the M100 (Fig. 1). In fact, in the light of the following data and discussion, this latter nomenclature (with inbuilt implications of “100” milliseconds) may appear inappropriate as latency variations can arise for a number of important reasons.
The auditory N1 response measured by EEG has been described as a vertex negative deflection consisting of several subcomponents, with a major subcomponent arising from primary auditory cortex, A1 (Picton et al., 1977; Näätänen and Picton, 1987; Virtanen et al., 1998). According to Hari (1980), the corresponding N1m peak detectable in the MEG reflects the tangential part of the supratemporal component of the N1. As such, it is in principle, a simpler more focused target for parametric investigation. Even so, source modeling has revealed several spatially distinct sources for the N1m signal; however, most are reported to lie in or near auditory cortex and are often considered as an ensemble and represented by a single equivalent current dipole (Näätänen, 1992; Teale et al., 1996a; Huotiilainen et al., 1998; Teale et al., 1998).

While spatial organization of brain function is an appealing product of fMRI and PET methods, source modeling the MEG signal may also reveal spatial features of cortical organization. For example, Romani (1982) reported that a tonotopic organization of the sources of the N1m was resolvable, with mediolateral frequency dependence as anticipated from electrophysiological studies in nonhuman primates and cat (Merzenich and Brugge, 1973; Schreiner, 1984). Pantev (1988, 1995) discusses tonotopy of different temporal components, suggesting resolution of tonotopic organizations in primary and secondary cortical areas.

However, despite ever-increasing sensor density of MEG detection systems, and improving anatomic knowledge of the individual head shape and even electrical conductivity, from high-resolution MRI and impedance tomography, the challenge of source modeling the MEG signal persists and descriptions of underlying neuronal activity derived from external recordings of the EEG or MEG remain, at best, estimates.

The body of this work explores the possibility that detected variations in the M100 latency reflect underlying encoding and processing activity of auditory cortex neurons. Latency variation has been explored parametrically as a function of stimulus attributes, such as duration, intensity, frequency and spectral structure. Of particular interest in developing applications of this work is the extension of such “sound” stimuli to include complex sounds eliciting the percept of a particular vowel, allowing us to traverse the border from sound to speech and speculate about the debate crystallized in the statement “speech is special.”

Even if a level of spatial organization, such as tonotopy is accepted, it seems likely that space alone does not provide adequate resolution for decoding and identifying the many auditory stimuli of our experience. The suggestion of additional mechanisms including forms of temporal encoding (as reflected in the M100 latency) brought the comment (Eysel, 1996) that temporal coding could “improve segmentation and synchronization,” broadening the number of “objects” that could be perceived separately.

**MAGNETOEENCEPHALOGRAPHIC MEASUREMENTS OF M100 LATENCY IN RESPONSE TO VARYING STIMULUS ATTRIBUTES**

Systems such as the spatially tonotopic organization of neuronal populations have been proposed to represent frequency or pitch information in terms of spatial location within the cortex (Romani et al., 1982; Lauter et al., 1985; Pantev et al., 1988). Spatial descriptions alone, however, may be insufficient. A number of studies have additionally investigated the temporal signature of neuronal responses to auditory stimuli. Of particular interest is the influence of stimulus attributes, such as intensity, pitch and timbre on the latency of the long latency components of the auditory evoked response, the electrical N1 and its magnetic counterpart, the M100.

**Evidence Regarding Latency of Responses to Sinusoidal Tones of Different Frequency**

In his seminal review of the electrical potential, N1, Picton (1977) discusses the amplitude response of the N1 components to stimulus attribute variation. He also acknowledges the latency dependence on stimulus intensity (being longer at low presentation levels) and mentions that slow stimulus onset rates, or ramp times, lead to latency prolongation. In fact, in 1966, Rapin et al. while studying the effects of stimulus intensity on N1 latency, also observed a prolongation of N1 latency for low frequencies (250 Hz) compared to higher frequencies (1 kHz, 6 kHz). This was especially apparent below presentation levels of 35dB SL (sensation level).

Jacobson (1992) shows evidence for frequency-dependence of the N1 latency and identifies the importance of considering the details of the time-domain of evoked responses. Consistent with other investigators, this study also reports increased peak amplitude for low frequencies—perhaps because sources of responses to low-frequency stimulation are located more superficially in cortex (Pantev et al., 1988). Recognizing that the magnitude of observed latency variation could not be accounted for by mere auditory periphery arguments, the latency effects were attributed to the combination of peripheral (3 ms) and central conduction delays.
Probing lower frequencies in terms of click rate for streams of auditory clicks, Forss et al. (1993) also demonstrate a latency dependence in the neuromagnetic evoked field component, N100m (also referred to as M100). With click rates as low as 40 Hz, the authors find latency prolongation on the order of the click interval. This is taken to imply that the auditory cortex and in particular the N100m generator integrates auditory input over at least two successive clicks, leading to “temporal integration times” for the N100m of the order of 20 to 25 ms.

Woods et al. (1993) compared evoked potential responses to 250 Hz versus 4 kHz tones. At latencies around 100 ms, corresponding to the electrical N1, he observes a 20 ms latency prolongation at low frequencies compared with higher frequencies. Evaluating other components of the evoked response, the author notes that analogous latency differences are also observable, but that their magnitude appears to increase with the absolute latency: at the brainstem, the latency difference is 1 ms, for middle latency auditory evoked potential components occurring 20 to 50 ms after stimulus onset, the latency difference is 3 to 6 ms. Verkindt (1995) also confirms prolongation of the N1 latency for low frequencies (250 Hz) compared with higher frequencies (500 Hz and above). The magnitude of the observed latency prolongation is 10 to 20 ms.

Roberts and Poeppel (1996) report an observation of M100 latency dependence on frequency with low frequencies (100 Hz) associated with latency prolongation of the order of 30 ms compared to high frequencies. Above 500 Hz to 1 kHz, no further latency reduction is observed (Fig. 2). This observation can be described in terms of a “fixed cost” of approximately 100 ms, plus a period-dependent time approximately equal to three periods of the sinusoidal waveform (30 ms for 100 Hz, 3 ms for 1 kHz). These observations prompted Greenberg to discuss a model of central expansion of differences evident in the cochlear traveling wave, linking peripheral and central components (Greenberg et al., 1998).

These observations provide an interesting opportunity for studying the processes of auditory temporal integration. If as suggested by Forss, the relevant period of “temporal integration” for the N100m is approximately 20 to 25 ms, or as implied by the studies of Roberts and Poeppel (1996), and the model of Greenberg (1998), and psychophysical data from Kay (1982) to be related to several cycles of the stimulus, one is led to a prediction that if the frequency of an auditory stimulus changes after the temporal integration window has elapsed, the latency of the resultant M100 evoked field will in fact be determined by the state of the stimulus during the temporal integration window, i.e., the first few tens of milliseconds after tone onset. We probed this issue, both psychophysically and with MEG recordings in healthy volunteers, using stimuli in which a transition from 100 Hz to 1 kHz (or vice versa) was inserted at various times (10 ms, 20 ms, 30 ms . . .) into a 400 ms duration stimulus. In the psychophysics portion, subjects were required to identify the presence of two versus one token in the stimulus. For all subjects, boundaries were found at 20 to 30 ms for the stimuli transitioning from high (1 kHz) to low (100 Hz). For the reversed transitions, the boundary was shifted (35 to 60 ms), suggesting that longer durations of the 100 Hz initial element were required for its adequate identification than for the 1 kHz element (Fig. 3A). This observation is entirely in accord with the “several cycles” argument. M100 latencies were also observed to transition sharply from plateaus of frequency no. 2 (when tone no. 1 was of insufficient duration, <20 to 30 ms) to plateaus corresponding to frequency no. 1 (when tone no. 1 duration exceeded 20 to 30 ms) despite a subsequent transition to frequency no. 2 (Fig. 3B).
Evidence Regarding Latency of Responses to Stimuli of Different Intensity

More extensively studied than the influence of frequency on N1, or M100, latency is the role of stimulus presentation level, or intensity. As early as 1981, Elberling showed that decreasing stimulus intensity led to decreasing evoked response amplitude, and increasing response latency of the neuromagnetic evoked response component (Elberling et al., 1981). Thus stimulus intensity and frequency both have a part in determining response component latency, and provide potentially ambiguous stimulus attribute pairings. This ambiguity can be resolved in part, by considering the stimulus intensity/evoked response amplitude relationship.

Pantev (1989a) proposes an amplitopic spatial organization in auditory cortex in which areas responding to low intensity stimuli are deeper and areas responding to high intensity stimulation are more superficial. Vasama (1995) observed that N100m latency decreased with increasing intensity, while response amplitude increased. In contrast to the study of Pantev, no clear amplitopic organization was observed. The author also concludes that if evoked fields are to be used in a quantitative or clinical application, to "probe the integrity of the auditory pathways," the measured quantities, amplitude, and latency do not suffer from the variability associated with source modeling, using a simple single equivalent dipole and two dipole approaches.

Stufflebeam et al. (1998) report similar findings of increasing M100 latency with decreasing stimulus presentation level, but also make the observation that at any given stimulus latency level, the shape of the latency dependence on stimulus tone frequency is preserved. That is to say, that if stimulus intensity is known, then M100 latency can still be interpreted in terms of stimulus attributes, particularly frequency. Low frequency tones were associated with relatively prolonged M100 latencies compared to higher frequency tones at the same presentation level; furthermore, the absolute prolongation increased at lower presentation levels, scaling with absolute latency.
**Beyond Sinusoids: Effect of Harmonic Content**

Most natural sounds are not, however, well approximated by sinusoidal tones of a particular frequency. Thus it is of considerable interest (especially in the study of speech perception) to investigate spectraly more complex sounds and their evoked responses.

Mäkelä et al. (1988) showed N100m latency prolongation for 125 Hz compared with higher frequencies for square wave stimuli. Consistent with other reports, the author notes little latency dependence on intensity at the intensities used: 40 to 70 dB sound pressure level.

Pantev (1989b, 1991, 1996) concludes that the M100 reflects the perception of “virtual” pitch in harmonic tones, rather than spectral pitch, since the source localization of the evoked field colocalizes with the source of a pure tone with a frequency equal to the perceived pitch, not the actual spectral energy distribution. While this uses spatial or source localization arguments to link activity 100 ms after stimulus with the processing of virtual pitch rather than the accumulation of spectral pitch, our findings suggest temporal evidence leading to a similar conclusion. In particular we find that for amplitude-modulated tones (200% modulated, suppressed carrier modulation, with modulation frequency, f_m = 100Hz), a prolongation of the M100 latency is observed compared to responses elicited by pure sinusoids at and around the carrier frequency. The magnitude of this prolongation appears relatively independent of carrier frequency and is on average approximately 5 ms (Fig. 4A). Nonetheless, its statistical significance and reproducible occurrence points to involvement of “secondary spectral analysis” in the generation of the M100 evoked response component. This is quite consistent with Pantev’s conclusion (1989b) that the M100 relates to the virtual pitch (which might be derived from such secondary spectral analysis).

In addition to replicating the finding of latency dependence on pure tone frequency, Langner (1997) also investigates encoding of periodicity information intrinsic to harmonic tone complexes. He proposes a spatial “periodotopy,” orthogonal to the tonotopic axis. While the spatial details of this are not entirely consistent with Pantev’s arguments (1989b), enough methodological differences between the studies exist allow this issue to remain unresolved. Not inconsistent with either set of studies, we propose an analogous reflection of periodicity encoding in the time domain, perhaps describable as “periodochrony” (Roberts, 1998b). In particular we find that the magnitude of the latency prolongation observed with amplitude-modulated tones might in fact be a func-
tion of the modulation frequency, or interovertone frequency spacing for harmonic tones. This might be best demonstrated by considering triangular and square waveforms, with a fundamental frequency of 100 Hz. Where the sinusoid has but a single spectral line at the fundamental frequency, and thus has intrinsic ambiguity in the concepts of spectral versus virtual pitch, the other waveforms have harmonic components in addition to the fundamental. The triangular wave has spectral components present at all harmonics (2Fo, 3Fo, 4Fo . . .) whereas the square wave has energy present at only odd harmonics (3Fo, 5Fo, 7Fo . . .). Thus the interovertone spacing is Fo for triangular wave and 2Fo for the square wave. Interestingly since the harmonics are not evenly weighted, the median frequency of the spectral energy distribution is approximately equal for both waveforms and therefore would not predict differences in M100 evoked response component latency for the two waveforms, although both would be expected to have shorter latencies than the pure 100 Hz sinusoid, by virtue of the additive higher frequency (shorter latency) energy form the harmonic content. Put another way, the M100 latency might be predicted primarily from the median frequency of the spectral energy distribution. However, the above “periodochronic” encoding model would suggest a secondary consideration: namely that the evoked response to the triangular waveform with its lower frequency periodicity would exhibit a prolonged M100 latency compared to the square wave. As can be seen from Fig. 4B, the latency prolongation for a 100 Hz square wave is indeed less than that observed for the triangular wave. Both are less than that observed for 100 Hz sinusoidal tones (which have energy only at 100 Hz).

In an event-related potential (ERP) study, Ragot and Lepaul-Ercele (1996) demonstrate a variation in N1 latency with stimulus frequency parameters somewhat analogous to the findings of Roberts and Poeppel (1996) with MEG. The stimuli they used consist of synthesized vowels but can be characterized primarily in terms of pitch (speaker voicing frequency) and formant position (which guides phonetic percept, and actual spectral energy distribution). They observe that latency decreases with increasing speaker pitch, but is independent of phonetic character. They tend to overconclude that latency does not vary with spectral center and is rather dominated by perceived pitch. This conclusion is warranted only in the range of vowels they studied, since formant positions varied only in the range from 700 Hz to 2300 Hz. In the study of Roberts and Poeppel (1996), latencies of responses evoked by tones in this frequency range were approximately constant. Latency variation was most pronounced in the range from 100 Hz to 500 Hz. Consequently, based on the choice of stimuli, Ragot's experiment seems especially sensitized to latency variations of the secondary type, attributable to pitch extraction, or secondary spectral analysis of periodicity, consistent with the observations of the groups of Pantev, Langner and Roberts discussed above (Pantev et al., 1996; Langner et al., 1997; Roberts, 1998a). Nonetheless for periodicities ranging from 100 Hz to 330 Hz, they observe a considerable latency variations of approximately 15 ms.

**Perception**

The clear sensitivity of M100 latency to secondary features of the acoustic spectrum, such as periodicity, allows us to embark on the road from stimulus attribute accumulation, through feature extraction and toward the formation of the auditory percept. Consider the amplitude-modulated tones. While all spectral energy is located at frequencies at or near the carrier, f0, a clear percept of the lower modulation frequency, fM, is nonetheless elicited. The observation that the M100 evoked by such a stimulus has a latency prolongation (in the direction of the low frequency content of the modulation frequency) provides temptation to suggest that either the M100 latency reflects the formation of the percept, or is modulated by perceptual feedback. In fact, the correlation between M100 latency prolongation and subjective perception of the virtual pitch may simply reflect the process of secondary spectral analysis, periodicity determination and pitch extraction. Nonetheless, since these are undoubtedly component elements of the process of percept formation, perhaps both “perception-based” and “analysis-based” arguments are equivalent. In either case, no causality can be unequivocally determined; that is the M100 latency may reflect or correlate with periodicity determination or perception, but it remains to be established if it is a cause or a consequence.

Further evidence probing the shady ground of percept formation and correlated observations in M100 latency is provided by the study of Roberts et al. (1998a). Building on the simple observation that low-frequency (100 Hz) tones are associated with longer latencies than high-frequency (1 kHz) tones, this study used additive mixtures of sinusoids of both frequency. The relative content of each component in the stimulus varied along the continuum from "pure" 100 Hz to "pure" 1 kHz. In the psychophysics component, subjects were required to categorize each stimulus in the continuum as "predominantly low frequency" or "predominantly high frequency." A categorical boundary could be clearly defined. In the MEG component, M100 latencies were
The M100 latency tracks the categorical nature of perception. A continuum of stimuli were generated by addition of low-frequency (100 Hz) and high-frequency (1 kHz) sinusoids in various amplitude mixtures. Subjects were asked to identify a preponderance of "low" versus "high" frequency content (two alternative forced choice). Magnetoencephalographic recordings were made with each stimulus presented at least 100 times and M100 latency was measured; detection thresholds were determined separately for all stimuli and stimuli were presented at -40 dB SL (sensation level). Latencies for pure 100 Hz tokens were typically prolonged 20 to 30 milliseconds (ms) compared to pure 1 kHz stimuli. Psychophysical data from a single representative subject are shown in open squares—a clear categorical boundary is defined with the "point of ambiguity" occurring at a 10.3 dB bias in favor of the low frequency (100 Hz) contribution. This corresponded with the subject's higher detection threshold for 100 Hz versus 1 kHz. M100 latency values do not show a smooth transition that might be expected from the linear superposition of 100 Hz and 1 kHz responses, but rather track the behavioral data, showing a "categorical" boundary between plateau values typical of high-frequency and low-frequency pure tones, with a 9.8 dB bias. In this subject the dynamic range (plateau separation) is found by hyperbolic tangent (tanh) curve fitting to be 33.7 ms. The "rate constant" of the tanh function is 0.7 for the psychophysical data and 0.5 for the M100 latency behavior.

From complex sounds, with associated harmonic content, we are able to make the transition to speech elements, and to begin to assess early aspects of speech perception. Acoustically, the simplest vehicle for such a transition is the vowel and thus vowels form the next category of stimuli. Vowels have perceptual attributes associated with the articulatory gestures that form the basis for their production, e.g., tongue height (e.g., /a/ versus /i/), frontness (i.e., /i/versus /u/), and lip rounding (/a/ versus /u/). The acoustic features that give vowels their unique perceptual identity (vowel quality) are the bands of energy peaks at particular frequencies, the "formants," F1, F2, F3 etc. The distribution of formants in a vowel is largely a function of the articulatory gestures. The first formant F1 is targeted as the manipulated parameter, based on observations of the sensitivity of MEG measures (specifically, evoked response component latencies) to variations in frequency in the range 100 Hz to 1 kHz (i.e., the range occupied by F1). In using simple vowel elements, and arguing for the assessment of "speech" perception, some studies invoke "linguistically relevant" tasks to give the isolated sounds a speech context. Additionally, since many aspects of "language" are believed to be lateralized in one cerebral hemisphere (generally the left hemisphere for right-handed and most left-handed individuals), these investigations tend to probe separate hemispheric responses for manifestations of differential responses.

Results from our own work using MEG also show patterns of activity in auditory cortex which correlate with linguistically relevant perceptual features. We have observed hemispheric asymmetries in auditory cortex that are conditioned by the tasks subjects are asked to execute (Poeppel and Roberts, 1996). In particular, using vowels of different pitch and phonetic identity, Poeppel and Roberts show task-related increases in M100 amplitude, bilaterally, while showing pitch-dependent latency shifts that are independent of any task demands, and are also exhibited bilaterally. They speculate that M100 latency manifestations of temporal encoding of stimulus attributes precede specific phonetic analysis and rather reflect acoustic properties.

In another study using the vowels /a/ and /u/, Poeppel...
Cumulatively, the data clearly point to the involvement of superior temporal cortex, including the transverse gyri of Heschl, in early speech sound processing. However, it remains to be established whether the observed observations in M100 response reflect any speech-specific processing since most data can be accounted for by spectral arguments and behavior is consistent with responses from simple and complex nonspeech sounds. It is possible that activity at this stage and location is common to the processing of both complex nonspeech and speech sounds and that distinction of the latter is apparent only in later components of the evoked response (Eulitz et al., 1995).

In a related fMRI/MEG investigation, we used simple vowel stimuli (male and female utterances of /a/ and /u/), presented bilaterally, and examined responses under three behavioral conditions: passive listening, phonetic discrimination, and speaker pitch discrimination (Roberts et al., 1998b). In short, the fMRI results demonstrate bilaterally symmetric activation of auditory cortex in the passive condition, with a significant bilateral increase of area of activation in the task conditions. This increase showed a tendency toward greater magnitude in the left hemisphere; however, the activation pattern derived could best be described as “partially-lateralized,” exhibiting left to right areas of activation in the ratio approximately 4:3. The corresponding MEG results demonstrated similar bilaterally symmetric M100 amplitudes in the passive state, with slight (10% to 20%), but symmetric increases in peak amplitude in the task conditions, i.e., demonstrating no task-related asymmetry in the 100ms response. However, consistent with the observations of Szymanski et al. (1999) and Papanicolaou (1999), asymmetries in dipolar activity were indeed evident in the later aspects of the evoked field (150 to 400 ms) post stimulus. This also is consistent with the observations of Eulitz (1995) who concluded that speech perception related influences in the evoked field were not evident in the M100 response, but only in the later, sustained field. Interestingly, such multimodal approaches yield additional insight into the individual measurement techniques themselves: if fMRI is considered to “average” over an entire evoked response, then partial lateralization in response to linguistically-relevant stimulation might be anticipated from a simple summation of early symmetric activity (as indicated by the M100) and later strongly asymmetric activity (as indicated by the sustained field response).

Effects at 50 ms and 200 ms

Is M100 latency reflecting the groundwork of acoustic perception formation or is it in fact showing perceptual
feedback? Does it in fact relate to temporal coding at a neuronal level, to synchrony across neuronal populations, or is it merely an epiphenomenon? While these questions remain open, some additional insights might be gained by examining the behavior of earlier (middle latency responses) and later (sustained field responses).

Analogous to the multiple observations of M100 latency responses being slightly shorter in the right hemisphere than the left (see also Fig. 2), Huotilainen (1998) shows shorter latency effects in the right hemisphere at both 50 ms (labeled Pam, the magnetic counterpart of middle latency potential Pa) as well as 100 ms (N1m).

Analogous to the main M100 latency dependence on stimulus frequency, Woods (1995) shows middle auditory evoked potential latencies 13% to 18% shorter for high frequencies (4 kHz versus 250 Hz). This is in line with slightly frequency dependent brainstem evoked response latencies. Thus, the absolute magnitude of the latency dependence appears to increase with absolute latency. It may not, however, be a linear relationship - since M100 latency differences (often called the “dynamic range of the M100 latency effect”) are of the order 20% to 30%. Interestingly, in contrast to Pentev (1995), Woods shows no clear evidence of tonotopy for sources of middle latency responses. This may be attributable to differences in magnetic (Pentev) versus electrical (Woods) recording techniques. In fact, Pentev shows a different tonotopic organization for Pam, compared to the later M100 response and suggests that the Pam activity is arising from primary auditory cortex and that the M100 activity reflects secondary activity from associated areas (with reversed medio-lateral tonotopic organization).

In contrast to the earlier evoked responses, the later activity (sustained potential, or sustained field), occurring several hundred milliseconds after stimulus onset may additionally be strongly influenced by features other than mere stimulus attributes. Indeed although Pentev (1994) reports a sustained field dependence on stimulus frequency, many others (e.g., Picton et al., 1978; Hari et al., 1989) have reported additional influence of subject attention on responses at later latencies. Of relevance to the study of language, Papanicolaou’s group have demonstrated characteristics of the late evoked activity showing strong asymmetries, leftward in most subjects, but reversed in patients confirmed by Wada testing as right hemisphere dominant for language (Simos et al., 1998; Papanicolaou et al., 1999). Szymanski also shows that the late field elicited by simple vowel stimuli (not generally recognized as optimal probes of hemispheric asymmetries) nevertheless exhibit analogous leftward asymmetry, which reverses in right hemisphere dominant patients (Szymanski et al., 1999). Thus, considerable support is available for the contention that late evoked activity reflects processing streams, not active in the 50 ms or 100 ms latency ranges.

Why Not Use Event-related Potentials?

Current availability of large array MEG systems is somewhat limited, while EEG equipment is more widespread and less expensive. Since many of the latency observations reported for the M100 response may have detectable correlates in the electrical potential, N1, is the additional expense of MEG measurement justified for this application and can useful investigations be performed with scalp recorded EEG?

According to Näätänen and Picton (1987), the N1 electrical potential can be described in terms of at least 6 subcomponents, which may respond differently to differing attributes of a stimulus. Additionally, EEG recordings are sensitive to many other aspects of background electrical activity or “brain noise.” The MEG technique has in some sense, diminished sensitivity, both spatially and in terms of the source orientation. However, for identification of the 100 ms component of an evoked response, and for parametric study thereof, this may be advantageous. In fact, while the M100 has been modeled as having two subcomponents, it is quite commonly describable using a single source model.

Several factors cause differences in the timing of the neuromagnetic response of the M100 and the electrical response of the N1, which reflect different measurements of the latency of the auditory evoked response. Simultaneous EEG and MEG measurement may yield potentially new results regarding the fine temporal evoked waveforms (Pentev et al., 1995; Tallon-Baudry et al., 1997; Virtanen et al., 1998).

There is some speculation and debate as to whether MEG is a more accurate measurement of the timing of neural activity as compared to EEG/ERPs. There are some theoretical reasons suggesting that MEG may be a superior method, but recent direct comparisons of simultaneous measurements do not confirm that they have practical implications, at least to statistical significance. First, there are technical differences in recording the evoked response with EEG and MEG, including the electrode montage. The standard electrode montage for EEG consists of approximately 19 electrodes over the scalp. The N1 latency is measured by recording the maximal ERP over a time window of approximately 50 to 160 ms. One drawback of this technique is that phase delays across the electrodes will alter the exact timing of
the maximal evoked potential. This problem can be overcome somewhat by using the maximal activity using a model, such as a dipole model. However, with a standard array of electrodes, modeling the ERP as a dipole is difficult. Precise measurement of the latency is better and more reliably measured with a high-density array of sensors. Magnetoencephalographic responses are typically recorded with a large number of sensors, and typically have a superior signal-to-noise ratio compared to evoked potentials (Virtanen et al., 1998).

Event-related potentials are affected by the propagation of the electrical waves, which are conducted over longer distances in a conductive medium, such as the meninges, skull, and scalp. The skull has low conductivity to electricity, but is transparent to magnetic fields, which leads to a more pure measure of the evoked latency (Hämäläinen et al., 1993). Depending on the conductance of the material, the electrical potential may have a relative phase shift caused by conductivity and impedance differences. This is not as much of a problem for magnetic fields, as they are not affected by the electrical impedance. Recently, however, using theoretical models, the improved sensitivity to measuring magnetic sources versus electrical sources has been disputed recently (Malinvuo et al., 1997). The problem of electrode density can be ameliorated by using a high-density electrode montage (e.g., Gevins et al., 1994).

Experimentally, recent studies report similar localizations and latency measurements of the M100 and the N1. Virtanen (1998) used simultaneous high-density 64-channel EEG and 122-channel MEG recordings during auditory stimulation. They found that the latencies and magnitudes of calculated electrical and magnetic dipoles were statistically equivalent. The EEG standard deviations, however, were greater. The authors conclude that due to the higher noise level with EEG, MEG may be superior in some circumstances.

A final, but perhaps fundamental, reason for differences in the M100 and N1 is that they are the additive result of multiple different generators. It is generally accepted that there are at least three, and as many as six, separate cortical generators that comprise N1 (Woods, 1995), referred to as N1', N1a, N1b, N1c, among others. Woods reports that the N1' generator is the largest and demonstrates strong tonotopy. The generators have different neurophysiologic properties such as tonotopy and temporal integration times (Alain et al., 1997), which lead to additive difference in the peak activity of a compound electrical potential or magnetic field.

The M100 also has been shown to have multiple sources (Hämäläinen et al., 1993; Moran et al., 1993). Unlike EEG, in MEG the orientation of the generators makes some of them invisible to current sensor arrays. For example, MEG does not detect radial sources, whereas EEG is insensitive to source orientation (Hämäläinen et al., 1993). The N1'/P90 component has a tangential orientation and the largest magnitude, and may contribute the largest component to the M100 waveform. The N1b has a tangential source orientation and is likely involved in the generation of the M100. The N1a and N1c components have radial oriented electrical source, and therefore do not likely contribute to the M100. Magnetoencephalography does not detect deeper or distant cortical sources because of the rapid decrease of the magnetic field with distance. Therefore, using MEG to measure the latency of the M100 may be the result of fewer overlapping fields and generators, making the measurement potentially more robust.

Reite (1982) also shows that 100 ms components of auditory evoked fields and auditory evoked potentials show a different amplitude dependence on stimulus class (clicks versus tones versus noise bursts), suggesting different influences on auditory evoked potential generators compared to auditory evoked field sources.

**CLINICAL APPLICATIONS OF LATENCY MEASUREMENTS**

Is there a role for latency measurements in clinical assessment of neurologic diseases and disorders? Despite tremendous advances in magnetic resonance imaging (MRI), many neurologic disorders and diseases are distinctly devoid of apparent structural abnormalities and are thus difficult to confirm diagnosis, or monitor treatment response, using conventional neuroimaging techniques. As indicators of, for example, temporal processing, MEG and ERP measures offer possible objective quantitation of a functionally relevant parameter: time. The following paragraphs illustrate some potential clinical applications of latency measures in a variety of disease states.

**Parkinson’s Disease**

Parkinson’s disease is a neurodegenerative disorder that results in atrophy of the basal ganglia. Clinically, it is characterized by motor findings (resting tremor, hypokinesia, and rigidity) as well as general cognitive decline. Pekkonen et al. (1998) studied the disorder using MEG and evoked auditory neuromagnetic fields. They found both a reduction in the amplitude of the M100 and a prolongation in its latency. Thus, MEG may be used in diagnosing the disease. Furthermore, it confirms other
studies indicating that the basal ganglia are involved in auditory processing, specifically in the timing.

Alzheimer's Disease

Alzheimer's disease is a leading cause of dementia worldwide. It results in general atrophy of the brain, but most strikingly so in the temporal lobes. It follows that auditory processing is likely to be affected at some level, given that the primary auditory cortex is in the temporal lobe. Pekkonen et al. (1996) studied 12 patients with Alzheimer's disease measuring the M100 responses. Affected individuals demonstrated marked prolongation of the M100 latency (40 to 50 ms) compared to neurologically healthy subjects. The P50m was also markedly prolonged. Alzheimer's disease thus appears to impair the temporal processing of auditory information, and at an early cortical stage.

Schizophrenia

The pathogenesis of schizophrenia remains unsolved. Evoked responses have lead to some insight into possible mechanisms of the disorder. It has been widely reported that the N1 is delayed in latency for schizophrenic patients (Jacobson, 1994). Measurements of the M100 have confirmed these results and have extended them. Reite et al. (1989) reports that the N1m was asymmetric when comparing responses from contralateral stimulation in neurologically normal subjects, but more symmetric in schizophrenic subjects. Specifically, the dipole was more anterior in the left hemisphere and oriented more posteriorly in male schizophrenics than in male controls. Further studies by Reite (1997) confirm anomalous cerebral lateralization in schizophrenics, but found that it is sex specific. In female schizophrenics the hemispheric asymmetry is more asymmetric. These results suggest that schizophrenia manifests itself as a disorder of cerebral lateralization, and affects the sexes specifically. Hajek et al. (1997a,b) confirm these MEG findings, but also found that the left temporal lobe gray matter was decreased in volume compared to normal controls. There is thus both structural and functional data in the left temporal lobe in schizophrenics.

Evoked potentials have reported similar reports in alterations in the timing of the N1. Additionally, schizophrenics also have altered timing in other portions of the late evoked potentials, including prolonged N200 (Bahramali et al., 1998) and P300 (Laurent et al., 1999) latencies.

Tinnitus

Tinnitus is a major problem and affects over 30 million people worldwide. It results in the perception of a high pitched sound in the absence of auditory stimulation. An objective measurement would be of wide clinical utility. Pantev et al. (1989c) reported that the M100 latency was longer for patients with tinnitus. Pantev and Hoke found that the P2m amplitude was diminished and the M100 amplitude was increased in patients with unilateral tinnitus. The authors propose a ratio of amplitude intensity of the 200 ms component versus the 100 ms component as an objective index of the disorder. The effect on the P2m and contradictory results was reviewed by Jacobson (1994).

A more recent study, Hoke et al. (1998) found that patients with tinnitus exhibited a more complicated response, as manifested by changes in the spectrotemporal pattern of the neuromagnetic response, which depended on the type of tinnitus (tonal versus noisiform; monaural versus binaural). Further research is necessary with careful selection of types of tinnitus to find any changes in the latency of the M100.

Drawing on the basic results presented above suggests that an interesting opportunity for objective tinnitus severity quantification might be to observe the M100 latency in response to a presented low frequency tone (e.g., 100 Hz). In healthy volunteers this might be expected to be prolonged 30 ms compared to a tone of 1 kHz frequency. However, in the presence of spectral energy at higher frequencies (e.g., square wave described above), this prolongation is reduced. It is tempting to speculate that the perceived high frequency sound, hallmarking the subjective sensations of some forms of tinnitus, might also lead to a reduction in M100 latency prolongation for a low frequency presented tone. If proven true, this could in principle lead to a quantitative measure of the cortical component of the tinnitus sensation.

Normal Brain Development

Several reports have found developmental changes in the M100. Rojas et al. (1998) found that preadolescents had longer M100 latencies. This was also shown by Paetau (1995). Rojas et al. (1998) speculate that this may relate to prolonged conduction velocities between the cochlea and Heschl’s gyrus due to myelination of thalamocortical and cortico-cortical fibers that occur during adolescence. Furthermore, they found a decrease in the refractory period in older children. This may explain why ERP N1 studies report a lack of M100 response in
young children, as the interstimulus intervals were too short. The mismatch field is also an important part of the analysis of development because it does not require the attention of the subject. The above described latency dependencies on stimulus attributes such as pitch and timbre have recently been confirmed in 13 to 16 year-old adolescents (unpublished data).

At the other spectrum of development, the aging brain also shows alterations in the timing of the M100. Pekkonen et al. (1995) found that in older subjects the interhemispheric difference of the M100 was longer than in younger patients, although the P50m interhemispheric difference was the same. The locations of the dipole sources were not altered. This may relate to age related loss of white matter fibers crossing through the corpus callosum.

These considerations should be borne in mind when attempting to use evoked field component latencies to characterize abnormalities of development, e.g., attention deficit hyperactivity disorder, autism spectrum disorders and dyslexia, as well as studies of dementia and aging.

Brain Tumors

Nakasato et al. (1997) studied patients with temporal lobe tumors, both before and after surgical removal. Seven of the patients had M100 absence or latency prolongation, corresponding to either direct tumor involvement or vasogenic edema. In two out of the four patients studied postoperatively, the M100 latencies recovered to within normal limits. Thus, MEG can be used noninvasively to study functional recovery post surgery.

Sleep

Naka et al. (1999) recently studied the effects of various levels of sleep on the auditory evoked neuromagnetic fields. Using pure tone bursts at a variety of frequencies (250, 1000, 4000 Hz), the authors found substantial alterations in both the timing and dipole locations during sleep. Specifically, the M100 equivalent current dipole had a prolonged latency and was located more anterior and superior than in the awake state. This may have implications for clinical and research sleep studies.

Cochlear Implants

There are reports of both evoked potentials and fields from cochlear implant patients. The M100 and N1 latency in subjects with cochlear implants has been studied in a number of reports, with disparate results. The M100 latency in each of the reports about the evoked field varies from normal, prolonged, or early. Hari et al. (1988) studied a cochlear implant recipient who had suffered adult-onset bilateral hearing loss and found that the M100 had a normal latency, magnitude, and localization. Hoke et al. (1989) reported that the M100 was both prolonged and decreased in amplitude by one third. The reason for the diversity of results may relate to the effectiveness of the cochlear implants, as illustrated by a recent ERP study. Groenen et al. (1997) report that cochlear implant users who performed poorly on speech recognition tasks demonstrated more diversity in the amplitude and latency of the middle latency response. More studies are necessary to know if the same holds true for the M100 timing.

SUMMARY

Where spatial tonotopy and, perhaps, periodotopy provide mechanisms for organizing primary and secondary spectral attributes of the auditory input in spatial coordinates, other temporal mechanisms of encoding may be present. Tonochrony and periodochrony may supplement, augment, and in some cases stand in place of spatial organizational principles. Having this additional dimension to encode input materials could "improve segmentation and synchronization," broadening the number of "objects" that could be perceived separately (Eyssel, 1996). Some indication that temporal encoding strategies are being implemented may be inferred from the latency of the N1 component of the auditory evoked electrical potential, or its magnetic counterpart, the M100.

M100 latency reflects a powerful encoding mechanism, sensitive to physical attributes of the stimulus, but also to some perceptual qualities. The degree to which the perceptual quality is predictable from the spectral features is central to the interpretation of M100 latency variation—either as an indicator of accumulated acoustic parameters, in the preperceptual pathway, or rather as an evoked response, modulated by perceptual feedback.

Although some prior data suggests the M100 latency is dominated by the spectral energy distribution of complex tones, it has been shown that the M100 latency is further modulated according to a secondary spectral analysis. This explanation accounts for differing latency prolongations observed with triangle and square wave stimuli of the same fundamental component and similar spectral center of gravity. Also, this explanation allows for the prolongation of M100 latency in response to amplitude modulated tones relative to sinusoids of the
carrier frequency. Interestingly, this spectral substructure analysis corresponds to a subjective perception of the low modulation frequency. So, based on data from triangles, square waves, and amplitude-modulated tones, it is unclear whether the M100 latency is determined by an acoustic analysis (allowing contributions of both center of gravity as well as substructure or periodicity) or by a perceptual processing stage which recognizes a sound’s timbre or the presence of the “missing fundamental,” respectively.

Data from the two-tone additive mix (Roberts et al., 1998a) supports the involvement of further perceptual processes having a role in M100 latency determination, since the psychophysical phenomenon of categorical perception is paralleled by a “categorical” shift of M100 latency between two extreme plateaus.

The nature of the perceptual specification of the electrophysiological response latency remains unclear, and the early appearance (100 ms) of such involvement may seem somewhat surprising. Further investigation along these lines must evaluate the relative importance of perceptual versus purely acoustic contributions to the M100 component of the evoked response. Analogous investigation of earlier and later evoked response components (e.g., the M50/M60 and M200 peaks) may also offer insight into the timing of the feedback and/or spectral analysis processes. Magnetoencephalography offers the appropriate temporal resolution to document these studies, while also offering an opportunity for resolving differences in spatial origin of evoked response contributions.

Furthermore, the available evidence points to common processing and analysis of complex nonspeech and speech tokens at the 100 ms poststimulus timeframe. Both spatial localizations from source modeling and latency response behavior can be accounted for based on primary and secondary spectral analysis. These steps, relating to spectral center of gravity, and periodicity (and the corresponding percept of pitch) are likely necessary elements of stimulus recognition, identification and pitch extraction, and precede subsequent characterization in terms of phonetic category.

As alluded to in the above paragraphs, acknowledging the functional relevance of the time domain and using brain potentials/field component latencies as objective indices of functional impairment opens up a spectrum of potential clinical applications. As yet, nonspecific observations (e.g., M100 latency prolongation) might be made more specific by using an array of simple and complex stimuli and observing patterns of latency response behavior, probing breakdown at levels of “detection” versus “primary identification” versus “processing of sec-

ondary spectral attributes.” It is our belief that this methodology offers great promise in the 4-dimensional description of brain function.

Acknowledgments: The authors thank Howard Rowley, MD, Heidi Roberts, MD, Nicole Gage, MS, Michael Szymanski, Ph.D., Steven Greenberg, Ph.D., and Christoph Schreiner, MD, Ph.D. for their collaboration and insightful comments. The authors also thank Susanne Homma, RT and Jeff Walker, BS for their excellent experimental assistance.

REFERENCES


Hari R, Hämäläinen M, Kaukoranta E, Mäkelä JP, Jouisiemsi SL,


Reite M, Teale P, Goldstein L, Whalen J, Linnville S. Late auditory magnetic sources may differ in the left hemisphere of schizophrenic patients. *Arch Gen Psychiatry* 1989;46:565–72.