Cortical encoding of aperiodic and periodic speech sounds: Evidence for distinct neural populations

Santeri Yrttiaho a,b,c,* , Patrick J.C. May b,c, Hannu Tiitinen b,c, Paavo Alku a

a Department of Signal Processing and Acoustics, Aalto University School of Electrical Engineering, P.O. Box 13000, FI-00076 AALTO, Finland
b Department of Biomedical Engineering and Computational Science, Aalto University School of Science, P.O. Box 12200, FI-00076 AALTO, Finland
c BioMag Laboratory, HUSLAB, Hospital District of Helsinki and Uusimaa, Helsinki University Central Hospital, P.O. Box 340, FI-00029 HUS, Finland

ARTICLE INFO

Article history:
Received 4 October 2010
Revised 1 December 2010
Accepted 28 December 2010
Available online 6 January 2011

Keywords:
Magnetoencephalography
Auditory evoked field
N1m
Stimulus-specific adaptation
Interstimulus interval
Pitch

ABSTRACT

Most speech sounds are periodic due to the vibration of the vocal folds. Non-invasive studies of the human brain have revealed a periodicity-sensitive population in the auditory cortex which might contribute to the encoding of speech periodicity. Since the periodicity of natural speech varies from (almost) periodic to aperiodic, one may argue that speech aperiodicity could similarly be represented by a dedicated neuron population. In the current magnetoencephalography study, cortical sensitivity to periodicity was probed with normal natural vowels and their aperiodic counterparts in a stimulus-specific adaptation paradigm. The effects of intervening adaptor stimuli on the N1m elicited by the probe stimuli (the actual effective stimuli) were studied under interstimulus intervals (ISIs) of 800 and 200 ms. The results indicated a periodicity-dependent release from adaptation which was observed for aperiodic probes alternating with periodic adaptors under both ISIs. Such release from adaptation can be attributed to the activation of a distinct neural population responsive to aperiodic (probe) but not to periodic (adaptor) stimuli. Thus, the current results suggest that the aperiodicity of speech sounds may be represented not only by decreased activation of the periodicity-sensitive population but, additionally, by the activation of a distinct cortical population responsive to speech aperiodicity.

© 2011 Elsevier Inc. All rights reserved.

Introduction

Speech communication involves both periodic and aperiodic sounds which are characterized, respectively, by the presence and the absence of the periodic acoustical excitation produced by the vibration of the vocal folds in the human larynx. Depending on the vocal fold function, the degree of periodicity may vary even within speech sounds belonging to the same phonemic category. That is, many speech sounds, such as vowels, that typically occur in periodic forms can also be produced without periodic vibration of the vocal folds (e.g., in whispered speech). The degree of periodicity also varies in vowels of natural speech due to aspiration noise and the irregular oscillation of the vocal folds that takes place, for example, in the production of speech sounds which are characterized by the presence or the absence of the periodic vocal fold vibration. Furthermore, given the inherent contextual nature of the processing of natural connected speech, the investigation of the representation of speech periodicity should take into account the ways that these representations may be influenced by the preceding acoustic context.

The cortical processing of sound periodicity has been studied non-invasively by several magnetoencephalographic (MEG) studies. The studies focusing on the N1m, a prominent response elicited around 100 ms after stimulus onset and generated in the auditory cortex (May and Tiitinen, 2010), have revealed a larger response amplitude for auditory stimulation with a periodic as opposed to an aperiodic structure (Hertrich et al., 2000; Alku et al., 2001; Gutchalk et al., 2004; Tiitinen et al., 2005; Lütkenhöner et al., 2006; Yrttiaho et al., 2008, 2009, 2010; Soeta et al., 2005; Soeta and Nakagawa, 2008). Such an amplitude increase in cortical population-level responses could be interpreted in terms of a larger neural population activated by the periodic than by the aperiodic stimulus. In this view, the results seem to indicate the presence of a periodicity-sensitive population that is specifically activated in the periodic condition. This amplitude effect has typically been coupled with a more anterior source location of the response elicited by the periodic sound than that elicited by the aperiodic sound (Alku et al., 2001;
Gutschalk et al., 2004; Tiitinen et al., 2005; Yrttiaho et al., 2008, 2009, 2010), providing further evidence for the distinction between the underlying neural generators of the responses elicited by varying degrees of stimulus periodicity. Imaging studies with positron emission tomography (Griffiths et al., 1998) and functional magnetic resonance imaging (Griffiths et al., 2001; Patterson et al., 2002; Penagos et al., 2004; Hall et al., 2006; Hall and Plack, 2009; Lewis et al., 2009; Garcia et al., 2010; Puschmann et al., 2010) along with invasive electrode recordings from human patients (Schönwiesner and Zatorre, 2008; Griffiths et al., 2010) have corroborated these results by showing that cortical areas are activated by sound periodicity. According to Griffiths et al. (2010), periodicity-related responses occur in neural ensembles in the primary auditory cortex and in an adjacent area rather than in one subregion only. The results of the previous studies, thus, suggest that there is a cortical population or a set of populations that is sensitive to sound periodicity, the activation of which could play a key role in the processing of the periodicity of speech sounds.

Given that prominent N1m responses can be elicited by aperiodic stimuli as well, the auditory cortex seems to contain a substantial population that can be activated by sounds without a periodic structure. Accordingly, Gutschalk et al. (2004) used equivalent current dipole analyses to isolate two distinct sources of neuromagnetic responses: an anterior source activated only by sound periodicity and a posterior source activated by periodic and aperiodic stimuli alike. Furthermore, contrasting the fMRI BOLD-signal in periodic and aperiodic conditions separately to silence indicated widespread activation that was shared between the periodic and the aperiodic condition (Patterson et al., 2002). Thus, there seems to be a considerable overlap between the cortical populations activated by periodic and aperiodic stimuli. However, based on the existing data, it is very difficult to estimate the extent of the overlap between these populations.

The contribution to the N1m response of a specific population with selectivity to a distinct feature may be isolated by carefully planned experimental paradigms that minimize the contribution from generators that are not related to the processing of that particular feature. A procedure that has been previously used for this purpose is the paradigm of stimulus-specific adaptation (SSA) introduced in an electroencephalographic study of the N1 wave (Butler, 1968). The paradigm exploits the adaptability of the most prominent auditory event-related response, recorded magnetically as the N1m (e.g., Lu et al., 1992; for a review, see May and Tiitinen, 2010). The adaptation of the N1m generators is a relatively slow process that takes around 300–400 ms to build up (Sable et al., 2004). Importantly, the adaptation is stimulus-specific: the attenuation of the N1m elicted by the probe (the actual effective stimulus) depends on the similarity of the probe to its preceding context of adaptor stimuli (Butler, 1968; Jääskeläinen et al., 2004; Salminen et al., 2009, 2010). Therefore, in the SSA paradigm, the attenuation caused by a specific adaptor is measured with reference to the condition where the adaptor is identical to the probe (thus, yielding an activation pattern that best matches that elicited by the probe). As the stimulus-specificity of adaptation arises presumably due to feature-selectivity of different cortical populations, the paradigm can be used to make reverse inferences about feature-selectivity when observing stimulus-specific release from adaptation. That is, if there is a release from adaptation for a specific adaptor–probe combination against the reference condition, the probe can be suggested to activate, at least partly, a different neural population than the adaptor. This view is supported by results from animal models (Ulanovsky et al., 2003, 2004) showing stimulus-specific adaptation in the auditory cortex.

In studies of cortical processing of sound periodicity, a version of the SSA paradigm has been implemented by using stimulation where the adaptor and the probe stimuli are presented as a continuous sequence, that is, without an intervening silent interstimulus gap. Here the stimulus-specific release from adaptation can be observed as a feature-change response that is elicited by a change in the degree of stimulus periodicity. Such a response elicited by the onset of periodicity after continuous noise has been reported by Martin and Boothroyd (1999), Krumholz et al. (2003) and Gutschalk et al. (2004). This result could be attributed to the activation of a periodicity-sensitive population that remains unadapted during the aperiodic stimulation. Interestingly, the results on such changes in the opposite direction, that is, from periodic to aperiodic, are conflicting. While Krumholz et al. (2003) and Gutschalk et al. (2004) found no response to be triggered to the onset of aperiodicity, a substantial response elicited by the onset of aperiodic noise after a continuous periodic stimulus was observed by Martin and Boothroyd (1999). Importantly, a possible explanation for such an aperiodicity-related response is the activation of a population that is responsive to aperiodic but not to periodic stimuli. Alternatively, however, the response may reflect activation of neurons tuned to spectral regions falling between the harmonics of the periodic stimulus irrespective of the degree of stimulus periodicity. Recently, von Kriegstein et al. (2010) investigated the processing sites of voiced and whispered syllables in an fMRI study. They found aperiodic whispered syllables to elicit activation in a distinct region located immediately adjacent to the primary auditory cortex. Comparable results have not been found in previous fMRI studies using noise burst stimuli (Griffiths et al., 2001; Patterson et al., 2002; von Kriegstein et al., 2006).

Previous studies, thus, lend contradictory results regarding the hypothesis that a dedicated population might contribute to the processing of sound aperiodicity. To further test this hypothesis, and to complement the hemodynamic findings by von Kriegstein et al. (2010) with an MEG approach, the focus of the current investigation was in the cortical sensitivity to the degree of speech aperiodicity. Special emphasis was placed on the processing of vowel aperiodicity through using spectrally matched vowel stimuli. In natural continuous speech, periodic and aperiodic sounds alternate rapidly and are, thus, encountered in variable acoustic contexts. Accordingly, the secondary goal of this study was to investigate the processing of speech periodicity not just in isolation but under the influence of the preceding stimulation context (cf., Seither-Preisler et al., 2004). To this end, the SSA paradigm was used: adaptor–probe pairs were constructed by using four combinations made from periodic and aperiodic vowel stimuli. In these combinations both periodic and aperiodic vowel stimuli were used as probes and adaptors. The interstimulus interval (ISI) was, further, varied to allow mapping the cortical activation both before (ISI = 200 ms) and after (ISI = 800 ms) the adaptive processes related to the immediately preceding stimulus presumably become fully operational (cf., Sable et al., 2004). According to the assumptions of the paradigm, release from adaptation for periodic probes against aperiodic adaptors could be taken as evidence for the existence of a periodicity-specific population. Conversely, release from adaptation for aperiodic probes against periodic adaptors would suggest the existence of an aperiodicity-sensitive population in the auditory cortex.

Methods

Subjects

Ten healthy right-handed subjects (6 males; average age 24 years, standard deviation 4 years) participated in the study with written informed consent. The experiment was approved by the Ethical committee of the Helsinki University Central Hospital. During the experiment the subjects sat upright, with their head placed inside the MEG sensor array, watching a silent movie and were under instruction not to pay attention to the auditory stimuli and to avoid excessive eye-movements.

Stimuli

The stimuli consisted of a periodic and an aperiodic version of the vowel /a/. The periodic vowel was recorded in an anechoic chamber...
with a high-quality condenser microphone. The recording was made from an adult male uttering the vowel sound with a fundamental frequency of 108 Hz. The aperiodic vowel was constructed using semi-synthetic speech generation (Alku et al., 1999). This allowed us to manipulate the degree of sound periodicity without altering the overall spectral envelope of the vowel. The synthesis consisted of extracting the spectral envelope of the original vowel and using it to separate the glottal excitation from the vocal tract filter function by means of inverse filtering. The aperiodic excitation was produced by using a random white noise sequence. This noise excitation was then passed through the extracted vocal tract filter to produce an aperiodic sound with a spectral profile matching the original periodic vowel.

In natural speech communication, aperiodic vowels are prevalent particularly in whispered speech. The aperiodic stimuli of the present study, however, do not represent whispered vowels because they were generated by affecting only the excitation (glottal pulseform vs. random noise) of the speech production mechanism. In contrast, the production of whispered vowels by real speakers is known to differ from the generation of normal speech with respect to three major features in the frequency domain (Tartter, 1989; Ito et al., 2005; Vestergaard and Patterson, 2009). Firstly, and most importantly, the harmonic structure, which is present in normal vowels, is absent in whispered sounds due to the lack of the periodic glottal excitation. Secondly, increased coupling of the vocal tract and trachea results in higher values of the formant frequencies in whispers. Thirdly, the spectral flatness of whispered speech is larger than that of normal vowels. Correspondingly, the decay of the spectrum towards higher frequencies is more gradual in whispered than in normal vowels. Here, the objective in the resynthesis of the aperiodic vowel stimuli was not to mimic whispering as such, but to produce aperiodic versions of the original vowel while matching the stimuli with respect to spectral envelope and sound pressure level. The current aperiodic vowel stimuli, consequently, resemble a voice type that is extremely breathy and occurs, for example, in natural utterances produced in voice pathologies. While the aperiodic vowels had an altered vocal quality, both the periodic and the aperiodic stimuli could easily be perceived as the vowel /a/. The vowel sounds were trimmed to the duration of 200 ms using 5-ms Hann windows to smooth the stimulus onsets and offsets. The stimuli are shown in Fig. 1. Both the periodic and the aperiodic vowels were presented to the subject via plastic tube headphones at 75 dB(A) SPL.

Fig. 1. Stimulus waveforms and spectra. The waveforms (left) and the spectra (right) of both aperiodic (top) and periodic (bottom) vowel stimuli are shown. The aperiodic vowel was generated semisynthetically by replacing the natural glottal excitation pattern with an aperiodic random sequence. The periodic vowel was uttered by a male speaker and is characterized with a fundamental frequency of 108 Hz. The vocal spectral envelope, presentation level (75 dB(A) SPL), and duration (200 ms) were identical for both stimulus types.

**Stimulus-specific adaptation paradigm**

In the stimulus-specific adaptation paradigm, the effects of an adaptor stimulus on the response to a probe stimulus are measured. Due to the stimulus-specific nature of the adaptation of the N1(m) (Butler, 1968; Jääskeläinen et al., 2004; Salminen et al., 2009, 2010), a key factor in the paradigm is the difference in the acoustic features between the probe and the adaptor. The effects of the adaptor in one condition are compared against those in the reference condition where the same stimulus is used as both the adaptor and the probe. In the current study, both the periodic and the aperiodic vowels were used as adaptors and probes thus forming four adaptor–probe combinations. Larger responses, relative to the reference condition, elicited by a specific probe are assumed to indicate the activation of a distinct population responsive to the degree of probe periodicity. Such a periodicity-specific release from adaptation is, thus, operationalized in the current study as the difference between the N1m in conditions where the probe periodicity was either same or different from the probe periodicity.

In addition to the acoustic features of the probe and the adaptor, a parameter that largely determines the amount of adaptation of the N1m response is the ISI (e.g., Lu et al., 1992; for a review, see May and Tiitinen, 2010). In the current experiment, two different ISI settings were used. In the first setting, the ISI was kept constant at 800 ms and, in the second setting, it alternated between 200 ms and 800 ms. That is, the second setting comprised stimulus pairs with short (200 ms) intra- and longer (800 ms) inter-pair silent intervals. The condition with alternating ISIs can be analyzed by either considering the first stimulus in a pair as the probe and the second stimulus of the preceding pair as the adaptor or by considering the second stimulus in the pair as the probe and the first stimulus in the pair as the adaptor. These two ways of analyzing the data yield adaptor–probe ISIs of 800 ms and 200 ms, respectively. Thus, three ISI conditions emerge from the constant and alternating ISI settings where adaptor–probe offset–onsets are: 1) constantly 800 ms, 2) 800 ms alternating with 200 ms, and 3) 200 ms alternating with 800 ms. As the N1(m) for stimuli at constant ISIs shorter than 300 ms becomes minute and difficult to observe (May and Tiitinen, 2010), a condition with a constant 200-ms ISI was not included in the current experiment. Importantly, neither of the ISIs used are long enough to allow full recovery of the N1(m) amplitude once initial adaptation has occurred (cf., Lu et al., 1992). Thus, release from adaptation can feasibly be investigated in both 800-ms and 200-ms ISI conditions.

The experiment, thus, consisted of seven recording blocks defined by the adaptor–probe combinations and ISI settings (yielding effectively 12 different experimental conditions). The order of presentation of these blocks was counterbalanced across subjects by a Latin-square design. All 12 resultant conditions arising from using two different stimuli with two different ISI settings are depicted in Fig. 2.

**MEG acquisition**

The data was acquired with a 306 channel MEG system (Elekta Neuromag, Helsinki, Finland). The head position of the subject was localized before each recording with coils placed on the subject’s scalp. The MEG data was recorded with a bandwidth of 0.1–200 Hz and at a sampling rate of 600 Hz. On-line averaging of MEG data was synchronized to a trigger signal which was coupled to the onset of each stimulus or stimulus pair in the constant and alternating ISI condition, respectively. The averaged epochs contained a 100-ms pre-trigger baseline and post-stimulus activation extending 100 ms beyond stimulus offset. In each condition, 150 artifact-free averaged epochs were acquired. The epochs containing electro-oculogram values exceeding [150]μV or MEG values in the excess of [3000]fT/cm were rejected online.
MEG analysis

The N1m responses were determined as the most prominent deflection in the averaged MEG gradiometer data in the 85–175 ms post-stimulus latency range relative to the silent baseline interval. In the constant ISI condition a single stimulus and, thus, a single N1m to be averaged was present for each trial. In this case the baseline was set to a time range starting 100 ms before the onset of the stimulus and ending at stimulus presentation. In the experimental conditions characterized by the alternating ISI, the stimuli were presented as one stimulus pair per trial. Here a single baseline starting 100 ms before the first stimulus of the pair was used for the two N1m responses elicited by the stimulus pair.

The source level N1m amplitude (nAm) and source location were determined by fitting an equivalent current dipole (ECD) to the maximum amplitude (in fT/cm units) time-point in the data from a subset of 44 planar gradiometers in the left and in the right hemisphere separately (Fig. 3). Prior to the ECD fitting, the data was band-pass filtered at 1–20 Hz with filter tails having widths of 1 and 20 Hz for the high- and low-pass cutoff frequencies, respectively.

Statistical analysis

The differences of the mean values of the N1m latency, amplitude and source location (in three-dimensional coordinates) were analyzed with a repeated measures analysis of variance (ANOVA) with factors: probe (periodic and aperiodic), adaptor (periodic and aperiodic), ISI (800 ms constant, 800 ms alternating with 200 ms, and 200 ms alternating with 800 ms), and hemisphere (left and right). All statistically significant ANOVA results are reported. For post-hoc comparisons, Newman–Keuls tests were used. The reliability of the analyses was verified by sequentially repeating the ANOVAs for data subsets where the data from one single subject was omitted at a time. The ANOVA effects remained statistically significant for all such data subsets, which confirms the reliability of these effects.

Results

Prominent N1m responses were observed in all experimental conditions and in both cortical hemispheres (Fig. 4). The relationship between stimulation condition and the N1m amplitude is depicted in Fig. 5 and the results on N1m latency and anteroposterior source location are shown in Fig. 6. The ANOVA results on the amplitude, latency, and source location of the N1m response are indicated in Table 1.

Both the periodic and the aperiodic vowel stimuli were presented in the currently used stimulus-specific adaptation paradigm. The periodic probe elicited a larger N1m response (Fig. 5) than the aperiodic probe [F(1, 8) = 22.76, p < 0.01]. The dipole moments in the periodic and the aperiodic condition were, respectively, 15.0 nAm and 8.7 nAm in the left hemisphere and 14.6 nAm and 12.0 nAm in the right hemisphere. The source location of the N1m (Fig. 6) elicited by the periodic probe was, furthermore, more anterior than that observed in the case of the aperiodic probe [F(1, 8) = 12.41, p < 0.01]: the ECDs in the periodic versus aperiodic condition were located −4.0 mm vs. −9.2 mm and 3.3 mm vs. −0.4 mm anterior to the center of the head in the left and in the right hemisphere, respectively.

Stimulus-specific, that is, periodicity-specific adaptation was observed in the current results where the effect of probe periodicity on N1m
amplitude depended on the periodicity of the adaptor \([F(1,8) = 8.58, p < 0.05]\). More precisely, the N1m amplitude for the aperiodic probe was larger in the case of the adaptor being periodic than when it was aperiodic \((12.2 \text{ nAm vs. } 8.5 \text{ nAm}, p < 0.05)\). For the periodic probe, a larger N1m was observed when the adaptor was aperiodic than when it was periodic, although this difference failed to reach statistical significance \((15.9 \text{ nAm vs. } 13.7 \text{ nAm}, p = 0.16)\). The effect of the adaptor on the N1m amplitude also interacted with the hemisphere \([F(1,8) = 11.93, p < 0.01]\). However, the stimulus-specific adaptation was, importantly, equally manifested in both cortical hemispheres.

Stimulus-specific adaptation was studied over two different time-scales specified by the offset-onset intervals of 800 ms and 200 ms, the former being used as either a constant ISI or as alternating with the latter. The stimulus timing proved to have several consequences for the N1m response. First, the N1m latency (Fig. 6) was intensely modulated by the ISI \([F(2,16) = 134.26, p < 0.001]\): the latency in the 200-ms ISI condition \((145.0 \text{ ms})\) was delayed by approximately 20 ms \((p < 0.001)\) relative to the condition with the 800-ms ISI \((126.3 \text{ ms and } 126.0 \text{ ms for the constant and the alternating condition, respectively})\). The ISI affected the N1m amplitude as well \([F(2,16) = 13.92, p < 0.001]\): the amplitude in the condition with the constant ISI of 800 ms \((16.2 \text{ nAm})\) was larger than in the condition where the adaptor–probe ISI was 200 ms \((9.7 \text{ nAm}, p < 0.001)\) or where the 800-ms adaptor–probe ISI alternated with an ISI of 200 ms \((11.6 \text{ nAm}, p < 0.01)\). The third consequence of shortening the ISI to 200 ms was that the periodicity-specificity indicated by the larger N1m response to the periodic than to the aperiodic probe was abolished \((10.5 \text{ nAm vs. } 8.9 \text{ nAm}, p = \text{n.s.})\). The largest difference between the periodic and the aperiodic vowel conditions was observed for the 800-ms constant ISI \((20.0 \text{ nAm vs. } 12.4 \text{ nAm})\). However, the periodicity-specific adaptation of the N1m (Fig. 5, indicated by the dependency of the effect of probe periodicity on N1m amplitude on the periodicity of the adaptor), and a shift in the N1m source location (Fig. 6) were observed in all ISI conditions including the 200-ms ISI condition.

Discussion

Previous studies suggest the activation of a periodicity-sensitive cortical population during the presence of periodic speech sounds as evidenced by larger N1m amplitudes and more anterior source locations of the N1m response elicited by periodic than by aperiodic vowel stimuli \((\text{Alku et al., 2001; Tiihinen et al., 2005; Yrttiaho et al.})\).
In the current study we further investigated the possibility that a population sensitive to sound aperiodicity might contribute to the N1m elicited by aperiodic speech stimuli. This hypothesis was tested with the SSA paradigm where the N1m elicited by a probe stimulus was measured in the context of different adaptor stimuli. We assumed that tuning to different degrees of periodicity among distinct neural populations might be reflected in a periodicity-specific release from adaptation of the N1m (cf., Butler, 1968; May and Tiitinen, 2010). Therefore, both periodic and aperiodic vowels were used as probes and as adaptors in the current experiment. As in the previous studies, sensitivity to periodicity was revealed by larger and more anteriorly located N1m responses elicited by the periodic than by the aperiodic probes. Importantly, a periodicity-specific release from adaptation for the aperiodic probe preceded by the periodic adaptor was observed, suggesting that the aperiodic vowels may activate a distinct aperiodicity-sensitive population. To study how both the previously observed periodicity-sensitivity and the currently revealed periodicity-specific adaptation are affected by the temporal relations between the adaptors and the probes, we varied the ISI between two values (800 ms and 200 ms). The results indicated that the periodicity-specific increase in the N1m amplitude for the periodic relative to the aperiodic probe condition was abolished for the shorter (200 ms) ISI. The periodicity-specific adaptation and the periodicity-specific displacement in the N1m source location were, however, observed for both the 800 ms- and the 200-ms ISI.

Several studies addressing the encoding of sound periodicity have found evidence for a specialized cortical population or populations responsive to sound periodicity (Griffiths et al., 1998, 2001; Hertrich et al., 2000; Alku et al., 2001; Patterson et al., 2002; Gutschalk et al., 2004; Penagos et al., 2004; Tiitinen et al., 2005; Soeta et al., 2005; Soeta and Nakagawa, 2008; Lütkenhöner et al., 2006; Hall et al., 2006; Hall and Plack, 2009; Yrttiaho et al., 2008, 2009, 2010; Lewis et al., 2009; Schönwiesner and Zatorre, 2008; Griffiths et al., 2010). However, the few data sets that pertain to whether aperiodicity is also encoded by the activation of a specialized population do not seem to converge on the issue. While Martin and Boothroyd (1999) reported stimulus-specific release from adaptation for aperiodic noise probes following periodic tone adaptors, studies using iterated ripple noise (IRN) (Krumholz et al., 2003) and click-train stimuli (Gutschalk et al., 2004) found no evidence for such an effect. In fMRI studies, evidence for cortical sensitivity to sound aperiodicity has been found for whispered syllables (von Kriegstein et al., 2010) but not for non-speech noise bursts (Griffiths et al., 2001; Patterson et al., 2002; von Kriegstein et al., 2006). The recent investigation by von Kriegstein et al. (2010) revealed cortical sensitivity to aperiodicity in whispered syllables, which is a result interestingly similar to the one observed in the current investigation. Even though both investigations involved speech sound stimuli of an aperiodic structure, it is worth emphasizing that there are clear differences in stimulus characteristics between the two studies. The aperiodic stimuli of the present study were generated by affecting only the excitation (glottal pulseform vs. random noise) of the speech production mechanism. Hence, the periodic and aperiodic vowel stimuli had matching formant frequencies and spectral envelopes. In contrast, acoustical cues present in whispered vowels of natural speech involve several factors (e.g., shifted formant frequencies, altered spectral tilt) and only one of them, the absence of harmonic structure, is directly linked to periodicity. Thus, even though the results found by von Kriegstein et al. (2010) on cortical sensitivity to aperiodicity in speech were in line with those reported in the present investigation, the two studies approach the topic from different perspectives. Importantly, recent imaging results show that the choice of stimulus material (harmonic tones vs. IRN) leads to different estimates of the location of periodicity-sensitive cortical regions (Hall and Plack, 2009). Therefore, in the current study, the periodicity of a natural vowel sound was manipulated while preserving the spectral envelope of the stimuli. The current results indicate cortical sensitivity to the periodicity of vowel stimuli both through SSA and in terms of direct differences in the N1m amplitude and source location to periodic and aperiodic...
vowel probes. The larger amplitude coupled with a more anterior source location in the periodic than in the aperiodic condition was similar to that found in several studies indicating cortical sensitivity to sound periodicity (Hertrich et al., 2000; Alku et al., 2001; Tiitinen et al., 2005; Lütkenhöner et al., 2006; Soeta and Nakagawa, 2008; Yritysalo et al., 2008, 2009, 2010). This result might be taken to suggest that the population activated by the periodic stimulus is larger than that activated by the aperiodic stimulus and, thus, includes a distinct periodicity-sensitive subpopulation. In the current study, the SSA paradigm further allowed isolating the contribution of an aperiodicity-sensitive neural activation to the N1m by adapting the response to the aperiodic probe by a periodic adaptor. Thus, the SSA paradigm revealed cortical activation presumably elicited by vowel aperiodicity that could not be observed directly by using only constant isolated stimuli.

Previous MEG results indicate that the cortical sensitivity to sound periodicity is modulated by the ISI (Seither-Preisler et al., 2004). To study how such sensitivity in the case of speech sounds depends on the stimulation rate, the current study compared the cortical processing of vowel periodicity at a short ISI of 200 ms to that observed at an 800-ms ISI. N1m responses were observed for both ISIs albeit with a delayed latency in the 200-ms ISI condition. In this sense, our results resemble earlier data showing that prominent N1m responses can be elicited already after short intervals (down to 70 ms, Loveless et al., 1996) if these alternate with longer ISIs (Budd and Michie, 1994; Loveless et al., 1996; Nishimura et al., 2003). A delayed latency of the N1 at an ISI of 100–200 ms compared to that observed for ISI values between 400 and 1000 ms has also been reported by Wang et al. (2008). The effect might be akin to forward masking, since (continuous) masking has been shown to both decrease the N1m latency of the N1 at an ISI of 100 ms (Wang et al., 2008) if these alternate with longer ISIs (Budd and Michie, 1994). The effect might be akin to forward masking, since (continuous) masking has been shown to both decrease the N1m amplitude and to increase its latency (Hari and Mäkelä, 1988).

Interestingly, the previously reported amplitude difference between the periodic and the aperiodic probes (cf., Hertrich et al., 2000; Alku et al., 2001; Tiitinen et al., 2005; Yritysalo et al., 2008, 2009, 2010) observed here at an ISI of 800 ms was abolished in the 200-ms ISI condition. Seither-Preisler et al. (2004) varied the onset-to-onset ISI between IRN probes and aperiodic noise adaptors presented without an intervening silent gap. Their results indicated an increase of 5.9 nAm of the amplitude of the N1m elicited by the periodic probe for every doubling of the ISI (in the 500–4000 ms range). While a precise comparison of the current relationship between periodicity-sensitivity and ISI to that observed by Seither-Preisler et al. (2004) is difficult due to differences in the experimental design, the observation of decreased periodicity-sensitivity for decreased ISI appears to be qualitatively similar in both studies. However, in the current results, the source location of the response persisted to be more anterior for the periodic than for the aperiodic condition in the 200-ms ISI case as well. Furthermore, the periodicity-specific release from adaptation was observed at both ISIs. Thus, the periodicity-specific activation was observed not only for the relatively long ISIs typically used in N1(m) studies, but also at a shorter ISI closer to the rate of natural speech. These results provide evidence for the role of periodicity- and aperiodicity-sensitive neural populations in encoding the rapid alternations in the periodicity of natural connected speech.

Perceptually, the degree of periodicity of vowel sounds relates closely to their voice quality. Aperiodic excitation further characterizes consonant sounds such as unvoiced fricatives (e.g., /ʃ/, /θ/, and /ʃ/). In the light of the current results, the encoding of these features of speech might be supported by the activation of the periodicity-sensitive and aperiodicity-sensitive populations in the auditory cortex. Given that the vowel /a/ was used as the stimulus in the current study, the results would best apply to the encoding of the degree of periodicity of vowel sounds. One might, in addition, tentatively suggest that the aperiodic nature of fricative sounds might be represented with cortical activation that is selective to sound aperiodicity. However, further experiments using such consonant sounds would be required to confirm this hypothesis.

In conclusion, the current results suggest that the periodicity of speech sounds is represented by the context-dependent activation of neural populations that are selective to particular degrees of sound periodicity. Importantly, evidence supporting the hypothesis of a distinct aperiodicity-sensitive population that encodes the aperiodicity of speech sounds, corroborating the fMRI results of von Kriegstein et al. (2010), was found through periodicity-specific release from adaptation.

Acknowledgments

Funding. This work was supported by the Graduate School in Electronics, Telecommunications and Automation (GETA), and by the Academy of Finland (Research Programme Lastu: project nos. 135003, 135009).

References


