



HAL
open science

The development of auditory temporal processing during the first year of life

Laurianne Cabrera, Bonnie Lau

► **To cite this version:**

Laurianne Cabrera, Bonnie Lau. The development of auditory temporal processing during the first year of life. *Hearing, Balance and Communication*, 2022, 20 (3), pp.155-165. 10.1080/21695717.2022.2029092 . hal-03900035

HAL Id: hal-03900035

<https://hal.science/hal-03900035v1>

Submitted on 15 Dec 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The development of auditory temporal processing during the first year of life

Laurianne Cabrera ¹, Bonnie K. Lau ²

¹ Université de Paris, INCC UMR 8002, CNRS,
45 rue des saints-pères, F-75006 Paris, France

² Department of Otolaryngology – Head & Neck Surgery, University of Washington
1701 NE Columbia Rd, Box 257923, Seattle, WA 98195

Corresponding author email: blau@uw.edu

Acknowledgements

This work was partially supported by NIH grant R00 DC016640 to BKL
and by ANR grant ANR-17-CE28-008 to LC.

Conflict of Interest

The authors declare no conflict of interests.

Abstract

Objectives. The processing of auditory temporal information is important for the extraction of voice pitch, linguistic information, as well as the overall temporal structure of speech. However, many aspects of its early development remain poorly understood. This paper reviews the development of auditory temporal processing during the first year of life when infants are acquiring their native language.

Methods. First, potential mechanisms of neural immaturity are discussed in the context of neurophysiological studies. Next, what is known about infant auditory capabilities is considered with a focus on psychophysical studies involving non-speech stimuli to investigate the perception of temporal fine structure and envelope cues. This is followed by a review of studies involving speech stimuli, including those that present vocoded signals as a method of degrading the spectro-temporal information available to infant listeners.

Results/Conclusion. This review suggests that temporal resolution may be well developed in the first postnatal months, but that the ability to use and process the temporal information in an efficient way along the entire auditory pathway is longer to develop. Those findings have crucial implications for the development of language abilities, especially for infants with hearing impairment who are using cochlear implants.

Keywords: temporal processing, auditory development, infant perception

Infant's ability to perceive the rapid intensity and frequency changes in sound over time is important for speech perception and early spoken language acquisition. The fast temporal information, known as temporal fine structure (TFS), and the slow temporal fluctuations in amplitude, known as temporal envelope, are found in each speech utterance and convey key linguistic information. The fast temporal fluctuations (> 500 Hz) signal fine phonetic contrasts, and voice-pitch information, and the slow temporal information (< 50 Hz) conveys rhythm and syllabicity (1). Temporal processing has garnered much research interest because accurate coding of temporal acoustic cues is required for robust speech perception in noisy real-life listening conditions (e.g., 2). Moreover, whether temporal processing relates to dyslexia and language impairment is often debated (e.g., 3–5). Understanding the role of temporal envelope perception is also of particular importance for cochlear implant users who have access only to the temporal envelope of sounds and not to the TFS through their prosthesis.

This paper reviews the development of auditory temporal processing during the first year of life with a focus on auditory capabilities as determined by psychophysical and electrophysiological measures. While prior studies indicate that the cochlea is functionally and structurally well developed by the first postnatal months (6–9), the central auditory pathway and primary auditory processes involved in frequency, temporal and intensity coding are not adult-like at birth and continue to mature into childhood (10, 11). Despite a wealth of studies on infant auditory development (reviewed in 12, 13) as well as infant speech perception (reviewed in 14), our understanding of how well temporal information is encoded in the immature auditory system and how well infants can make use of this information to listen in their everyday acoustic environments remains far from complete. This question is even more

important for infants with sensorineural hearing loss who are eligible for cochlear implantation. A cochlear implant (CI) is a surgically implanted auditory prosthesis that can provide a sense of sound for individuals with severe hearing loss. They convey limited spectral information and present primarily the slow temporal modulations in the temporal envelope of sound. Even though hearing and speech understanding can be partially restored with CIs, children who use CIs show significant variability in auditory skills and spoken language outcomes that remains difficult to predict (e.g., 15, 16). Understanding the processing of auditory temporal cues during the early years of development is thus essential from both a theoretical and clinical perspective.

Spectral and temporal decomposition by the auditory system

Audition begins with the conduction of sound waves by the outer and middle ear followed by cochlear processing involving the basilar membrane, inner and outer hair cells, and auditory nerve fibers. Numerous biomechanical processes take place within the peripheral auditory system, but of particular interest in this review are those involved in the temporal analysis of incoming sounds. The basilar membrane decomposes complex waveforms into their frequency components, with high frequencies producing greater displacement at the base of the cochlea and low frequencies producing greater displacement at the apex. This organization of frequency by place of excitation, or tonotopy, is often modeled as a bank of narrowly tuned bandpass filters (17, 18). Each band can then be considered as a sinusoidal carrier with superimposed amplitude modulation (AM) and frequency modulation (FM) that the auditory system is able to decompose (19–22). The FM or TFS is determined by the dominant frequencies in the signal that fall close to the center frequency of the band. The AM or

temporal envelope corresponds to the relatively slow fluctuations in amplitude below 500 Hz superimposed onto the carrier. These two temporal variations are illustrated in Figure 1 in three different frequency bands.

What makes studying temporal processing so challenging?

Temporal characteristics of a sound can be represented in both spectral and temporal domains, making it difficult to distinguish between spectral and temporal processing. Two peripheral processes can also be used to code the temporal characteristics of a sound. First, the spectrum is represented in a place code with the place of excitation on the basilar membrane providing a code for frequency and the firing rate of auditory nerve fibers at each place providing a code for intensity. Second, representation of the time waveform, or the temporal code, relies on phase locking. This synchronous firing of auditory nerve fibers to the time waveform provides a code for the TFS of low-frequency components and modulation in firing rate provides a code for envelope fluctuations at all frequencies.

Although both envelope and TFS are encoded by phase locking, physiological studies suggest some dissociation in the representation of envelope and TFS in the early stages of the auditory system. For instance, studies conducted in animal models suggest that the accuracy of phase locking to the TFS is robust up to 1-2 kHz and then declines to be no longer detectable at about 4-6 kHz (23–26). In contrast, there is some evidence to suggest that neural phase locking to slower envelope cues remains accurate for higher carrier frequencies (27–29). The upper limits of phase locking, however, is currently an active topic of debate (30).

Why might we expect temporal processing to mature in the first year of life?

Mature auditory temporal processing requires an accurate representation of the time waveform of sound in the auditory nerve response. Furthermore, once an accurate representation of a sound's acoustic features is established in the cochlea and auditory nerve response, it must be maintained at least to the level of the brainstem for a listener to access this information. There is evidence that temporal aspects of the neural response take longer to develop than does frequency resolution, with at least three likely sources of immaturity: changes in phase locking, response latency, and neural adaptation.

The maturation of phase locking in infancy has primarily been investigated using electrophysiological methods. The ability for a neuron to phase lock to a sound stimulus depends on synaptic efficiency and the myelination of nerve fibers. The frequency following response (FFR), a scalp-recorded evoked response that occurs at the same frequency of its stimulation tone, is thought to reflect phase locking in the auditory nerve and brainstem. The term FFR describes the response that encodes the TFS of a stimulus while the term envelope following response (EFR) describes the response that encodes the periodicity of the temporal envelope. Levi and colleagues (31) measured FFRs and EFRs in 1-month-olds and adults to investigate temporal coding. The stimuli they presented were amplitude modulated tones with carrier frequencies of 500, 1000, or 2000 Hz with a modulation depth of 100%. Infants were tested with a modulation frequency of 80 Hz while adults were tested with a modulation frequency of 40 and 80 Hz, as these were previously determined to be the best modulation frequency for each age group (32). They found no difference between infant and adult FFR thresholds. Moreover, EFR thresholds at each group's best modulation frequency were also the same at 500 and 1000 Hz carrier frequencies. However, infant EFR thresholds were worse than

adults for the highest carrier frequency, suggesting that although robust phase locking to both envelope and TFS is observed by 1 month of age, there are some age-related differences in EFRs depending on the carrier frequency and modulation frequency.

Besides FFRs to AM tones, there have been numerous studies measuring the FFR to speech stimuli. The review of these studies is beyond the scope of this paper but an important point to note is that the recording parameters of infant speech-evoked FFRs including stimulus type and duration, as well as stimulus intensity, polarity, presentation rate and number, all influence the FFRs measured (reviewed in 33). Likewise, there have been many studies which use EFRs, also referred to as auditory steady state responses, elicited by amplitude modulated tones at modulation frequencies between 70-100 Hz as a method of hearing threshold estimation, suggesting that robust phase locking is observed at these rates in infants (34–40).

Auditory brainstem responses (ABR) have also been used as an indirect measure of phase locking since for an ABR to be evoked, synchronous firing of neurons as well as synchronous transmission in the brainstem is required. The ABR, also a scalp-recorded response, is composed by a series of positive waves of which Wave I through V are most typically analyzed (41). The maturation of ABR waveform morphology showing increasingly distinct waves with age is thought to reflect the maturation of neural synchronization (42, 43). Similarly, the later maturation of cortical evoked potentials can also be interpreted to indicate that synchrony in the cortex takes long to mature (44, 45). The exact age when phase locking in young listeners becomes comparable to adults is still unknown.

Response latency is another temporal aspect of the neural response that improves with development. Analyses of ABR wave latency in preterm infants shows that the latency of waves

get shorter with age (46). Comparison of decreases in ABR wave latencies shows a more significant decrease in wave V latency than in wave I, suggesting greater improvements in transmission through the brainstem compared with transmission to the auditory nerve. Maturation of ABR wave I latency is reached by term birth while both wave V and the wave I-V interpeak latency continues to improve until about 24 months (47), demonstrating that maturation of synchronous responses as well as synchronous transmission in the brainstem takes longer than in the auditory periphery. Likewise, greater decreases in latency of cortical evoked potentials, more central responses, have been observed during development further suggesting that mature response latency takes longer to develop in more central parts of the auditory system (44, 45). Developmental changes in latency are thought to result from myelination, axonal and dendritic maturation, as well as synaptic development (44, 45).

To investigate synaptic development, Ponton and colleagues generated a model which separated the contribution of myelination versus synaptic efficiency by measuring changes in ABR interpeak latency (48). According to their model, the wave III-IV interpeak interval is asynaptic and relates to myelination, wave III-V reflects both myelination and synaptic efficiency, and wave IV-V reflects synaptic efficiency. Measures of interpeak interval in infants showed that the wave III-IV interval matures prior to wave IV-V suggesting that myelination matures before synaptic efficiency. Moreover, the wave IV-V interval does not reach adult-like levels until 1 year of age, suggesting that synaptic immaturity is a contributor to response latency and may relate to maturation of temporal processing in infancy.

Finally, adaptation, a decrease in neural response with continued stimulation, is another source of age-related change observed in temporal processing. Adaptation in human infants

has been investigated by looking at rate effects and forward masking with ABRs. Increasing the rate of clicks used to elicit an ABR results in increased ABR latencies, evidence of neural adaptation (49). In premature infants at 28 weeks conceptional age, greater increase in response latency is seen than in adults, suggesting a greater adaptation effect. This effect decreases with age until 3 months where changes in response latencies with increasing rate are the same as for adults (49). When comparing ABR to a click and to a click followed by a broadband forward masker, the difference in ABR thresholds was found to be higher for newborns than adults suggesting more masking in newborns (49). This increased susceptibility to masking is indicative of low-level immaturity, which may be explained by greater neural adaptation in newborns.

Thus, electrophysiological studies in infants suggest that neural phase locking, subcortical and cortical response latency and neural adaptation are not adult-like during the first few months of life. The consequences of this neural maturation on temporal processing have been assessed using psychophysical tasks to evaluate infants' perception of temporal cues.

Perceptual performance on temporal processing tasks

Psychophysical studies investigating infant temporal processing have primarily utilized simple sounds such as pure tones and noises to measure the developing auditory system's ability to resolve changes in sound over time. A greater focus has been placed in studying envelope perception with very few studies assessing the detection of change in FM, or TFS.

From the few FM studies that have been conducted, we know that at 4 months of age, infants can discriminate FM sweeps but show greater sensitivity to larger frequency sweeps that involves higher frequencies, corresponding to the frequency range of infant-directed speech (IDS). Using a habituation paradigm, Colombo et al. (50) habituated 4-month-olds to white noise then presented one of three different signals: 1) a 150 Hz pure tone, 2) a FM sweep from 150 to 275 Hz back to 150 Hz over a 1s period corresponding to the intonation frequency range of adult-directed speech (ADS), and 3) a FM sweep from 150 to 550 Hz back to 150 Hz over a 1s period corresponding to the intonation frequency range of IDS. They only found evidence that 4-month-olds could detect the change from white noise to the 550 Hz IDS FM sweep but not the pure tone or the 275 Hz ADS FM sweep. In contrast, in their companion study they did find that 4-month-olds can discriminate between the two 275 Hz ADS and 550 Hz IDS FM sweeps (51).

Leibold and Werner (52) hypothesized that perhaps due to the habituation paradigm 4-month-olds seem to respond to the FM sweep modelled after the IDS contour because it was more interesting but not because they could not discriminate the sounds in the other conditions. Using an observer-based psychophysical procedure instead, which is a method designed to test infant sensitivity (53), they measured the detection threshold for the same ADS and IDS FM sweeps used by Colombo and colleagues, also in 4-month-olds. In this study, infants were required to ignore no-signal trials of only broadband noise and respond when they heard a FM sweep embedded in the noise. Thresholds for the 275 Hz ADS FM sweep was 3 dB higher (worse) than thresholds for the 550 Hz IDS FM sweep. They also measured detection thresholds for a pure tone condition and found that thresholds for the 275 Hz pure tone was also 3 dB

higher (worse) than the 550 Hz pure tone. The results from this study show that infants can indeed discriminate both 275 Hz and 550 Hz FM sweeps but also show slightly greater sensitivity to the 550 Hz IDS FM sweep because the frequency excursion of this sweep extends into higher frequency range than the 275 Hz sweep. Even by 6 months of age, infants continue to require larger frequency sweeps to detect a FM change in a 1-kHz tone compared to adults (54). Overall, these studies show that infants are able to detect differences in TFS in the first 6 months of life, and thus, that neural phase locking is functional. However, the development of efficiency, or how well they are able to use the available temporal information remains unclear.

The temporal modulation transfer function (TMTF; 55) is often considered as the most reliable measure of temporal resolution. The TMTF describes the relationship between modulation rate and the modulation depth required to detect the target AM. For a continuous noise carrier, the TMTF is a lowpass function with detection thresholds worsening as modulation frequency increases. The cutoff frequency of this function is thought to reflect the limit of auditory temporal resolution, while the overall height is thought to reflect processing efficiency, that is the capacity to make efficient use of the available information. To obtain a TMTF, listeners are asked to determine which of two sounds is amplitude modulated. Modulation depth is varied to define a threshold for different modulation frequencies. If the auditory system can resolve the fluctuations in a sound's amplitude, only a small modulation depth would be required. However, if the auditory system cannot resolve the fluctuations in a sound, modulation depth needs to be increased. In a recent study, Walker and colleagues showed that 3-month-olds did require about 10 dB more of AM depth to detect a 10 Hz modulation than adults. However, the cutoff frequency of the TMTF was not different between

infants and adults (56). Similar patterns of results are observed until late childhood (57–59) suggesting that TMTFs may not relate only to the maturation of sensory coding of AM information, but also to the maturation of how well the available AM information is used (60).

Various additional methods have been employed to assess temporal resolution. One commonly used method is gap detection, which typically involves determining the shortest gap duration that can be detected in a sound. For broadband noise stimuli, adult gap detection thresholds are about 3-5 ms, indicating exquisite temporal resolution in the mature auditory system (e.g., 61) while infant gap detection thresholds have been found to be much worse (e.g., 62, 63). Werner and colleagues (62) tested the ability of 3-, 6-, and 12-month-old infants as well as adults to detect gaps of different durations in a broadband noise. They manipulated the frequencies available in the broadband noise by presenting an additional high pass filtered noise masker with different cut-off frequencies - 500, 2000, or 8000 Hz - to mask frequencies above the specified cutoffs. In all frequency cutoff conditions, infants across age groups performed much worse than adults and had thresholds that were 40-60 ms higher. Participants in all age groups, however, showed better gap detection thresholds with increasing frequency cutoff, meaning that they performed better on the task when higher frequency information was made available. This similar effect of frequency on gap detection suggests that although infant performance is much worse, the mechanisms contributing to temporal resolution are similar in infants and adults.

A point to note, gap detection thresholds have also been measured with electrophysiological methods, which tend to show better infant thresholds (64, 65). Trainor and colleagues (64) measured the mismatch negativity in response to deviant stimuli that consisted

of gaussian-modulated pure tones with silent gaps of 4, 8, or 12 ms and found that 6-month-olds' gap detection was comparable to adults in all three conditions. Werner and colleagues (65) obtained both psychophysical and electrophysiological gap detection thresholds to gaps in broadband noise bursts. In 3-month-olds, they found poor psychophysical gap detection thresholds, consistent with the psychophysical thresholds they found previously (62), but thresholds measured by ABRs were comparable to adults. However, the results of such gap detection tasks may also involve spectral processing and intensity resolution, making the mechanisms underlying gap detection and its relevance to temporal processing unclear.

Forward masking is another paradigm used to study temporal processing. In forward masking paradigms, a masker is presented a short duration prior to a target tone. The closer the masker is to the target in time, the higher the threshold should be. By increasing the time between the masker and the target, the threshold for when the target returns to its unmasked value can be obtained, which indicates the duration of the auditory nerve fiber recovery period. If immature neurons require a longer recovery period, infant forward masking thresholds would be higher than adult thresholds. Werner (66) measured forward masking thresholds in 3-month-olds, 6-month-olds and adults at several different target-masker delays - 5, 10, 25, and 200 ms - using a 1000-Hz pure tone and a 100-ms broadband noise masker. Three-month-olds showed more forward masking than adults at all delays while 6-month-old thresholds were comparable to adult thresholds. These results suggest immature neural adaptation at 3 months of age (66).

Finally, pitch discrimination can also rely on temporal processing, depending on the stimulus presented. The results of many studies show that infants are sensitive to pitch

information in both speech and music in the first few months of life (e.g., 67–73). By 3 months of age, infants have been shown to detect changes in fundamental frequency (F0; pitch) in the presence of random variations in spectral content as well as adults with musical training and even surpass adults without musical training (74). However, as the stimuli presented in this study and most other infant pitch discrimination studies contain both place and temporal information for F0, it is not possible to say for certain that these pitch measures are reflective of temporal processing. Nevertheless, infants have also been shown to discriminate the pitch of stimuli which do not contain spectral cues for pitch, thus isolating temporal information only, but perhaps with greater difficulty. High frequency unresolved harmonics do not produce discernible peaks in the cochlear excitation pattern and thus rely on the phase locked firing of auditory nerve fibers to the envelope periodicity as a code for pitch. Clarkson and Rogers found that 7- and 8-month-old infants had difficulty categorizing the pitch of unresolved harmonics suggesting that infants may not have access to the temporal information for pitch in these complexes (75).

Butler and colleagues approached the question using different stimuli, high pass filtered iterated ripple noise (IRN), which also does not have spectral cues for pitch. They tested 8-month-olds' ability to discriminate a change from 167 to 200 Hz in the F0 of the IRN and found that performance varied significantly across infants leading the authors to conclude that while infants were able to discriminate the pitch of IRN, it is likely a difficult task for them (76). However, Lau and Werner (77) found that infants were able to categorize the pitch of unresolved harmonics in a low frequency region, 2500–4500 Hz, and a high frequency region, 4000–6000 Hz. They found that nearly all infants tested at both 3 and 7 months of age were

able to perceive the pitch of the unresolved harmonics in both spectral regions, suggesting that by 3 months of age, infants do have access to envelope periodicity encoded by phase locking. Together, the results from these pitch perception studies suggest that while infants can perceive pitch via temporal information alone, it is likely that the presence of spectral information provides a more salient pitch percept, as is the case for adult listeners.

Coding of temporal information for speech signal

A more recent approach used to assess the cortical tracking of temporal information involves using electroencephalography (EEG) to measure brain responses to natural speech. At birth and at 6 months, infant neural responses have been shown to track the envelope phase of speech sentences (78). Moreover, 7-month-olds show stronger cortical tracking of the speech envelope of IDS than ADS (79). However, as IDS differs from ADS on many acoustic parameters such as slower rate, higher pitch, and more exaggerated pitch contours, it is difficult to determine what drove the stronger envelope tracking.

Telkemeyer and colleagues (80) measured newborns' neural responses to temporal modulations for non-speech sounds mimicking the temporal structure of speech using simultaneous EEG and near-infrared spectroscopy (NIRS). Four kinds of temporally varying stimuli were generated by concatenating noise segments of varying lengths, 12, 25, 160, and 300 ms, forming 9-s-long stimuli with different modulation frequency patterns. No difference across conditions in EEG was observed, however, the vascular response recorded with NIRS did show differences across conditions with the greatest response elicited by the 25 ms modulation. The authors point out that this is the most relevant modulation frequency for

speech processing at the phonetic level. They interpret their findings to suggest that the newborn auditory cortex is sensitive to the temporal structure of sound and the brain is tuned to critical temporal properties of linguistic signals at birth.

Another method used to assess the processing of TFS and envelope information in speech, is the use of vocoders, speech analysis-synthesis systems that can selectively degrade the spectral and temporal components of original speech utterances (19, 20). A recent NIRS-EEG study with newborns presented two syllables with varying consonants, /pa/ vs /ta/, that were vocoded to replace the original speech TFS by pure tones in 32 frequency bands. Two conditions were designed to also manipulate the envelope information; In the first condition, the TFS is degraded and the original temporal envelope is preserved while in the second condition, the TFS is degraded and the faster temporal envelope is filtered out in order to preserve only the fluctuations below 8 Hz. The syllables were presented using an oddball paradigm, with a standard syllable presented 80% of time and a deviant syllable 20% of the time, within stimulus blocks. The EEG results revealed a mismatch response between standard and deviant syllables in both conditions. This suggests that at birth, the brain is able to detect a consonant change in syllables when only the slowest envelope cues (under 8 Hz) are preserved (81). More notably, the vascular responses recorded using NIRS revealed a different time course as well as a different region of activation between the two vocoder conditions, suggesting that the processing of fast and slow speech envelope may not rely on the same neural mechanisms.

These results are consistent with behavioral findings showing that while 3- and 6-month-olds are able to discriminate consonant contrasts on the sole basis of slow temporal

envelope cues, they showed greater difficulty when speech contained only the slowest envelope cues compared to when the faster envelope cues are preserved (82–84). In adults, a different pattern is observed behaviorally; that is, adults show near perfect discrimination performance when only the slowest envelope cues are available (84). The fact that adults do not require faster envelope modulation for accurate speech perception in quiet has been demonstrated using different tasks and different speech stimuli (19, 20). Altogether, the results from these infant vocoder studies suggest that before 6 months of age, infants are able to encode and use both the TFS and envelope cues found in speech but that they require the faster (> 8 Hz) envelope cues whereas adults do not.

The crucial coding of temporal information for infants using cochlear implants

A CI converts sound into electrical impulses that are delivered to the auditory nerve by an electrode array inserted into the inner ear. While a CI does not reproduce the auditory nerve response that sound elicits in acoustic hearing, implants have been successful at restoring speech perception in quiet (22). Nevertheless, limitations in the encoding of sound leave CI users with an impoverished signal resulting in continued auditory deficits. CIs process incoming sounds and deliver the AM cues of the original signal via a limited number of relatively broad frequency channels (see 100 for further details). TFS cues of the original signal are not transmitted by CI processors but are replaced by a fixed train of pulses which are amplitude modulated by the original envelope cues. Although the first CIs offered only one single electrode, conveying no place information, adult users were able to recognize speech in quiet

using only the AM cues transmitted by their device. These first clinical results highlighted the critical role of temporal information and envelope cues for speech perception (86)

In adults, a CI delivers an impoverished signal to a mature language processing system. In the case of congenitally deaf young children, a CI delivers the same impoverished signal to an immature auditory system as well as an immature language system. Nevertheless, numerous studies show that oftentimes, children fitted with CIs before the age of 3 years of age can develop good receptive and expressive language skills (87, 88). It has also been demonstrated that an earlier age at implantation leads to better language skills although large individual variability is observed (89–94). What is not well studied however, is how congenitally deaf infants learn the properties of their native language through the impoverished spectrotemporal information delivered by the CI processors. As far as we know, there have been no studies conducted thus far investigating temporal processing in infants who use CIs. In children, there has been one study looking at AM detection in a group of 10 prelingually deafened CI users between 7 and 16 years of age who received their implants before age 3. This group of children showed poorer AM detection thresholds and greater individual variability in AM sensitivity compared to a group of chronologically age-matched normal-hearing peers (95). In postlingually deafened adult CI users, speech perception in quiet and in noise has been related to their ability to use AM information (96–98). Since infants are currently implanted at increasingly younger ages, measures of auditory temporal processing in this population may facilitate the search for better early predictors of speech and language development.

Conclusion

The auditory system's ability to encode and process temporal variations is of fundamental importance to speech perception and spoken language acquisition. Neurophysiological studies suggest that while many neural mechanisms supporting temporal processing are immature at birth, they are functional in the first postnatal months and develop rapidly in infancy. In contrast, psychophysical studies show that by 6 months of age, although infants are able to detect both TFS and envelope fluctuations in sound, they perform poorly in comparison to adults on most temporal processing tasks. One possible explanation for these seemingly contradictory findings is that perhaps temporal resolution is well developed in the first postnatal months, so that the encoding of temporal information by the auditory nerve is comparable to adults. However, the ability to use and process this temporal information efficiently along the entire auditory pathway may have a longer developmental trajectory. Moreover, the contribution of nonsensory factors such as attention, memory, internal noise or listening strategies, to infant-adult differences on temporal processing tasks is an important consideration (see 58, 59, 99, 100). Nevertheless, continued maturation of the central auditory neural pathway beyond infancy likely improves efficiency in temporal processing until late childhood. This extended period of maturation may have important implications for the development of language abilities, especially for individuals who rely heavily on temporal cues for speech perception, such as cochlear implant users.

References

1. S. Rosen, Temporal information in speech: Acoustic, auditory and linguistic aspects. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **336**, 367–373 (1992).

2. M. A. Stone, C. Füllgrabe, B. C. Moore, Notionally steady background noise acts primarily as a modulation masker of speech. *The Journal of the Acoustical Society of America* **132**, 317–326 (2012).
3. D. V. Bishop, R. P. Carlyon, J. M. Deeks, S. J. Bishop, Auditory temporal processing impairment: neither necessary nor sufficient for causing language impairment in children. *J Speech Lang Hear Res* **42**, 1295–1310 (1999).
4. A. A. Benasich, P. Tallal, Infant discrimination of rapid auditory cues predicts later language impairment. *Behav Brain Res* **136**, 31–49 (2002).
5. U. Goswami, V. Leong, Speech rhythm and temporal structure: Converging perspectives? *Laboratory Phonology* **4**, 67–92 (2013).
6. C. Abdala, Distortion product otoacoustic emission (2 f- f) amplitude as a function of f/f frequency ratio and primary tone level separation in human adults and neonates. *The Journal of the Acoustical Society of America* **100**, 3726–3740 (1996).
7. J. Y. Bargones, E. M. Burns, Suppression tuning curves for spontaneous otoacoustic emissions in infants and adults. *The Journal of the Acoustical Society of America* **83**, 1809–1816 (1988).
8. T. Morlet, *et al.*, Spontaneous otoacoustic emissions in preterm neonates: prevalence and gender effects. *Hearing research* **90**, 44–54 (1995).
9. R. Pujol, M. Lavigne-Rebillard, Development of neurosensory structures in the human cochlea. *Acta oto-laryngologica* **112**, 259–264 (1992).
10. J. K. Moore, Maturation of human auditory cortex: Implications for speech perception. *Ann Otol Rhinol Laryngol Suppl* **189**, 7–10 (2002).
11. J. R. Saffran, J. F. Werker, L. A. Werner, “The infant’s auditory world: Hearing, speech, and the beginnings of language” in *Handbook of Child Psychology*, D. Kuhn & R. Siegler, (2006), pp. 58–108.
12. L. A. Werner, Infant auditory capabilities. *Current Opinion in Otolaryngology & Head & Neck Surgery* **10**, 398–402 (2002).
13. L. A. Werner, R. R. Fay, A. N. Popper, *Human auditory development* (Springer, 2012).
14. P. K. Kuhl, Early language acquisition: cracking the speech code. *Nature Reviews Neuroscience* **5**, 831–843 (2004).
15. M. A. Svirsky, A. M. Robbins, K. I. Kirk, D. B. Pisoni, R. T. Miyamoto, Language development in profoundly deaf children with cochlear implants. *Psychol Sci* **11**, 153–158 (2000).

16. J. K. Niparko, *et al.*, Spoken Language Development in Children Following Cochlear Implantation. *JAMA* **303**, 1498–1506 (2010).
17. B. R. Glasberg, B. C. J. Moore, Derivation of auditory filter shapes from notched-noise data. *Hearing Research* **47**, 103–138 (1990).
18. B. C. J. Moore, Speech processing for the hearing-impaired: successes, failures, and implications for speech mechanisms. *Speech Communication* **41**, 81–91 (2003).
19. R. Drullman, Temporal envelope and fine structure cues for speech intelligibility. *J. Acoust. Soc. Am.* **97**, 585–592 (1995).
20. R. V. Shannon, F. G. Zeng, V. Kamath, J. Wygonski, M. Ekelid, Speech recognition with primarily temporal cues. *Science* **270**, 303–304 (1995).
21. Z. M. Smith, B. Delgutte, A. J. Oxenham, Chimaeric sounds reveal dichotomies in auditory perception. *Nature* **416**, 87–90 (2002).
22. F.-G. Zeng, *et al.*, Speech recognition with amplitude and frequency modulations. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 2293–2298 (2005).
23. D. H. Johnson, The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones. *The Journal of the Acoustical Society of America* **68**, 1115–1132 (1980).
24. N. Y. Kiang, R. R. Pfeiffer, W. B. Warr, A. S. Backus, Stimulus coding in the cochlear nucleus. *Trans Am Otol Soc* **53**, 35–58 (1965).
25. A. R. Palmer, I. M. Winter, C. J. Darwin, The representation of steady-state vowel sounds in the temporal discharge patterns of the guinea pig cochlear nerve and primarylike cochlear nucleus neurons. *J. Acoust. Soc. Am.* **79**, 100–113 (1986).
26. J. E. Rose, J. F. Brugge, D. J. Anderson, J. E. Hind, Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey. *J. Neurophysiol.* **30**, 769–793 (1967).
27. P. X. Joris, C. E. Schreiner, A. Rees, Neural processing of amplitude-modulated sounds. *Physiol. Rev.* **84**, 541–577 (2004).
28. P. X. Joris, T. C. Yin, Responses to amplitude-modulated tones in the auditory nerve of the cat. *J. Acoust. Soc. Am.* **91**, 215–232 (1992).
29. S. Kale, M. G. Heinz, Envelope coding in auditory nerve fibers following noise-induced hearing loss. *J. Assoc. Res. Otolaryngol.* **11**, 657–673 (2010).
30. E. Verschooten, E. Verschooten, S. Shamma, A.J. Oxenham, B.C. Moore, P.X. Joris, M. G. Heinz, & C.J. Plack, The upper frequency limit for the use of phase locking to code

- temporal fine structure in humans: A compilation of viewpoints. *Hearing Research* **377**, 109–121 (2019).
31. E. C. Levi, R. C. Folsom, R. A. Dobie, Coherence analysis of envelope-following responses (EFRs) and frequency-following responses (FFRs) in infants and adults. *Hearing Research* **89**, 21–27 (1995).
 32. E. C. Levi, R. C. Folsom, R. A. Dobie, Amplitude-modulation following response (AMFR): Effects of modulation rate, carrier frequency, age, and state. *Hearing Research* **68**, 42–52 (1993).
 33. F.A. Lemos, A.D. da Silva Nunes, C.K. de Souza Evangelista, C. Escera, K. V. M. Taveira, & S. A. Balen, Frequency-following response in newborns and infants: a systematic review of acquisition parameters. *Journal of Speech, Language, and Hearing Research* **64**, 2085–2102 (2021).
 34. B., Cone-Wesson, F. Rickards, C. Poulis, J. Praker, L. Tan, & J. Pollard, The auditory steady-state response: Clinical observations and applications in infants and children. *J Am Acad Audiol* **13**, 270–282 (2002).
 35. G. Rance, R. Roper, L. Symons, L.J. Moody, C. Poulis, M. Dourlay, & T. Kelly, Hearing threshold estimation in infants using auditory steady-state responses. *J Am Acad Audiol* **16**, 291–300 (2005).
 36. S. A. Small, D. R. Stapells, Multiple auditory steady-state response thresholds to bone-conduction stimuli in young infants with normal hearing. *Ear and Hearing* **27**, 219–228 (2006).
 37. A. Van Maanen, D. R. Stapells, Normal multiple auditory steady-state response thresholds to air-conducted stimuli in infants. *J Am Acad Audiol* **20**, 196–207 (2009).
 38. O.G. Lins, T. W. Picton, B. L. Boucher, A. Durieux-Smith, S. C. Champagne, L. M. Moran, ... & G. Savio, Frequency-specific audiometry using steady-state responses. *Ear and Hearing* **17**, 81–96 (1996).
 39. F. W. Rickards, *et al.*, Auditory steady-state evoked potential in newborns. *Br J Audiol* **28**, 327–337 (1994).
 40. G. Savio, J. Cárdenas, M. Pérez Abalo, A. González, J. Valdés, The low and high frequency auditory steady state responses mature at different rates. *Audiol Neurootol* **6**, 279–287 (2001).
 41. R. F. Burkard, J. J. Eggermont, M. Don, *Auditory evoked potentials: Basic principles and clinical application* (Lippincott Williams & Wilkins, 2007).
 42. J. J. Eggermont, A. Salamy, Maturation time course for the ABR in preterm and full term infants. *Hear Res* **33**, 35–47 (1988).

43. A. Salamy, Maturation of the auditory brainstem response from birth through early childhood. *J Clin Neurophysiol* **1**, 293–329 (1984).
44. J. L. Wunderlich, B. K. Cone-Wesson, Maturation of CAEP in infants and children: A review. *Hear. Res.* **212**, 212–223 (2006).
45. J. J. Eggermont, J. K. Moore, “Morphological and functional development of the auditory nervous system” in *Human Auditory Development*, Springer Handbook of Auditory Research., L. Werner, R. R. Fay, A. N. Popper, Eds. (Springer New York, 2012), pp. 61–105.
46. M. P. Gorga, J. R. Kaminski, K. A. Beauchaine, Auditory brain stem responses from graduates of an intensive care nursery using an insert earphone. *Ear Hear* **9**, 144–147 (1988).
47. M. P. Gorga, J. R. Kaminski, K. L. Beauchaine, W. Jesteadt, S. T. Neely, Auditory brainstem responses from children three months to three years of age: normal patterns of response. II. *J Speech Hear Res* **32**, 281–288 (1989).
48. C. W. Ponton, J. K. Moore, J. J. Eggermont, Auditory brain stem response generation by parallel pathways: differential maturation of axonal conduction time and synaptic transmission. *Ear and Hearing* **17**, 402–410 (1996).
49. R. E. Lasky, The effects of rate and forward masking on human adult and newborn auditory evoked brainstem response thresholds. *Dev Psychobiol* **24**, 51–64 (1991).
50. J. Colombo, J. E. Frick, J. S. Ryther, J. T. Coldren, D. W. Mitchell, Infants’ Detection of Analogs of “Motherese” in Noise. *Merrill-Palmer Quarterly* **41**, 104–113 (1995).
51. J. Colombo, F. D. Horowitz, Infants’ attentional responses to frequency modulated sweeps. *Child Dev* **57**, 287–291 (1986).
52. L. J. Leibold, L. A. Werner, Infant auditory sensitivity to pure tones and frequency-modulated tones. *Infancy* **12**, 225–233 (2007).
53. L. A. Werner, “Observer-based approaches to human infant psychoacoustics” in *Methods in Comparative Psychoacoustics*, BioMethods., G. M. Klump, R. J. Dooling, R. R. Fay, W. C. Stebbins, Eds. (Birkhäuser Basel, 1995), pp. 135–146.
54. R. N. Aslin, Discrimination of frequency transitions by human infants. *J Acoust Soc Am* **86**, 582–590 (1989).
55. N. F. Viemeister, Temporal modulation transfer functions based upon modulation thresholds. *The Journal of the Acoustical Society of America* **66**, 1364–1380 (1979).
56. B. A. Walker, C. M. Gerhards, L. A. Werner, D. L. Horn, Amplitude modulation detection and temporal modulation cutoff frequency in normal hearing infants. *The Journal of the Acoustical Society of America* **145**, 3667–3674 (2019).

57. J. W. Hall, J. H. Grose, Development of temporal resolution in children as measured by the temporal modulation transfer function. *The Journal of the Acoustical Society of America* **96**, 150–154 (1994).
58. E. Buss, C. Lorenzi, L. Cabrera, L. J. Leibold, J. H. Grose, Amplitude modulation detection and modulation masking in school-age children and adults. *The Journal of the Acoustical Society of America* **145**, 2565–2575 (2019).
59. L. Cabrera, L. Varnet, E. Buss, S. Rosen, C. Lorenzi, Development of temporal auditory processing in childhood: Changes in efficiency rather than temporal-modulation selectivity. *The Journal of the Acoustical Society of America* **146**, 2415–2429 (2019).
60. L. A. Werner, K. Boike, Infants' sensitivity to broadband noise. *The Journal of the Acoustical Society of America* **109**, 2103–2111 (2001).
61. A. G. Samelli, E. Schochat, The gaps-in-noise test: Gap detection thresholds in normal-hearing young adults. *International Journal of Audiology* **47**, 238–245 (2008).
62. L. A. Werner, G. C. Marean, C. F. Halpin, N. B. Spetner, J. M. Gillenwater, Infant auditory temporal acuity: gap detection. *Child Dev* **63**, 260–272 (1992).
63. S. E. Trehub, B. A. Schneider, J. L. Henderson, Gap detection in infants, children, and adults. *The Journal of the Acoustical Society of America* **98**, 2532–2541 (1995).
64. L. J. Trainor, S. S. Samuel, R. N. Desjardins, R. R. Sonnadara, Measuring temporal resolution in infants using mismatch negativity. *Neuroreport* **12**, 2443–2448 (2001).
65. L. A. Werner, R. C. Folsom, L. R. Mancl, C. L. Syapin, Human auditory brainstem response to temporal gaps in noise. *J Speech Lang Hear Res* **44**, 737–750 (2001).
66. L. A. Werner, Forward masking among infant and adult listeners. *J Acoust Soc Am* **105**, 2445–2453 (1999).
67. R. G. Karzon, J. G. Nicholas, Syllabic pitch perception in 2- to 3-month-old infants. *Perception & Psychophysics* **45**, 10–14 (1989).
68. T. Nazzi, C. Floccia, J. Bertoni, Discrimination of pitch contours by neonates. *Infant Behavior and Development* **21**, 779–784 (1998).
69. L. J. Trainor, C. A. Zacharias, Infants prefer higher-pitched singing. *Infant Behavior and Development* **21**, 799–805 (1998).
70. K. Mattock, M. Molnar, L. Polka, D. Burnham, The developmental course of lexical tone perception in the first year of life. *Cognition* **106**, 1367–1381 (2008).
71. J. Plantinga, L. J. Trainor, Melody recognition by two-month-old infants. *The Journal of the Acoustical Society of America* **125**, EL58–EL62 (2009).

72. B. K. Lau, L. A. Werner, Perception of missing fundamental pitch by 3- and 4-month-old human infants. *J Acoust Soc Am* **132**, 3874–3882 (2012).
73. B. K. Lau, K. Lalonde, M.-M. Oster, L. A. Werner, Infant pitch perception: Missing fundamental melody discrimination. *The Journal of the Acoustical Society of America* **141**, 65–72 (2017).
74. B. K. Lau, A. J. Oxenham, L. A. Werner, Infant pitch and timbre discrimination in the presence of variation in the other dimension (Submitted).
75. M. G. Clarkson, E. C. Rogers, Infants require low-frequency energy to hear the pitch of the missing fundamental. *J. Acoust. Soc. Am.* **98**, 148–154 (1995).
76. B. E. Butler, N. A. Folland, L. J. Trainor, Development of pitch processing: Infants’ discrimination of iterated rippled noise stimuli with unresolved spectral content. *Hearing Research* **304**, 1–6 (2013).
77. B. K. Lau, L. A. Werner, Perception of the pitch of unresolved harmonics by 3- and 7-month-old human infants. *The Journal of the Acoustical Society of America* **136**, 760–767 (2014).
78. M. C. Ortiz Barajas, R. Guevara, J. Gervain, The origins and development of speech envelope tracking during the first months of life. *Dev Cogn Neurosci* **48**, 100915 (2021).
79. M. Kalashnikova, V. Peter, G. M. Di Liberto, E. C. Lalor, D. Burnham, Infant-directed speech facilitates seven-month-old infants’ cortical tracking of speech. *Sci Rep* **8**, 13745 (2018).
80. S. Telkemeyer, S. Rossi, S.P. Koch, T. Nierhaus, J. Steinbrink, D. Poeppel, ... & I. Wartenburger, Sensitivity of newborn auditory cortex to the temporal structure of sounds. *J. Neurosci.* **29**, 14726–14733 (2009).
81. L. Cabrera, J. Gervain, Speech perception at birth: The brain encodes fast and slow temporal information. *Science advances* **6**, eaba7830 (2020).
82. L. Cabrera, J. Bertocini, C. Lorenzi, Perception of Speech Modulation Cues by 6-Month-Old Infants. *Journal of Speech, Language, and Hearing Research* **56**, 1733–1744 (2013).
83. L. Cabrera, C. Lorenzi, J. Bertocini, Infants discriminate voicing and place of articulation with reduced spectral and temporal modulation cues. *Journal of Speech, Language, and Hearing Research* **58**, 1033–1042 (2015).
84. L. Cabrera, L. Werner, Infants’ and adults’ use of temporal cues in consonant discrimination. *Ear Hear* **38**, 497–506 (2017).
85. R. V. Shannon, Advances in auditory prostheses. *Curr. Opin. Neurol.* **25**, 61–66 (2012).

86. R. V. Shannon, Q.-J. Fu, J. Galvin, L. Friesen, "Speech Perception with Cochlear Implants" in *Cochlear Implants: Auditory Protheses and Electric Hearing*, Springer Handbook of Auditory Research., F.-G. Zeng, A. N. Popper, R. R. Fay, Eds. (Springer, 2004), pp. 334–376.
87. S. Bouton, W. Serniclaes, J. Bertoncini, P. Colé, Perception of speech features by French-speaking children with cochlear implants. *J. Speech Lang. Hear. Res.* **55**, 139–153 (2012).
88. R. T. Miyamoto, D. M. Houston, K. I. Kirk, A. E. Perdew, M. A. Svirsky, Language development in deaf infants following cochlear implantation. *Acta oto-laryngologica* **123**, 241–244 (2003).
89. C. M. Connor, T. A. Zwolan, Examining multiple sources of influence on the reading comprehension skills of children who use cochlear implants. *Journal of Speech, Language and Hearing Research* **47**, 509–525 (2004).
90. S. J. Dettman, D. Pinder, R. J. Briggs, R. C. Dowell, J. R. Leigh, Communication development in children who receive the cochlear implant younger than 12 months: risks versus benefits. *Ear and hearing* **28**, 11S–18S (2007).
91. J. G. Nicholas, A. E. Geers, Will they catch up? The role of age at cochlear implantation in the spoken language development of children with severe to profound hearing loss. *Journal of Speech, Language and Hearing Research* **50**, 1048–1062 (2007).
92. B. Richter, S. Eißele, R. Laszig, E. Löhle, Receptive and expressive language skills of 106 children with a minimum of 2 years' experience in hearing with a cochlear implant. *International journal of pediatric otorhinolaryngology* **64**, 111–125 (2002).
93. M. A. Svirsky, S.-W. Teoh, H. Neuburger, Development of language and speech perception in congenitally, profoundly deaf children as a function of age at cochlear implantation. *Audiology and Neurotology* **9**, 224–233 (2004).
94. J. B. Tomblin, B. A. Barker, L. J. Spencer, X. Zhang, B. J. Gantz, The effect of age at cochlear implant initial stimulation on expressive language growth in infants and toddlers. *Journal of Speech, Language and Hearing Research* **48**, 853–867 (2005).
95. M.-H. Park, J. H. Won, D. L. Horn, J. T. Rubinstein, Acoustic temporal modulation detection in normal-hearing and cochlear implanted listeners: Effects of hearing mechanism and development. *Journal of the Association for Research in Otolaryngology* **16**, 389–399 (2015).
96. Y. Cazals, M. Pelizzone, O. Saudan, C. Boex, Low-pass filtering in amplitude modulation detection associated with vowel and consonant identification in subjects with cochlear implants. *The Journal of the Acoustical Society of America* **96**, 2048 (1994).
97. Q.-J. Fu, Temporal processing and speech recognition in cochlear implant users. *Neuroreport* **13**, 1635–1639 (2002).

98. D. Gnansia, D.S. Lazard, A. C. Léger, C. Fugain, D. Lancelin, B. Meyer, & C. Lorenzi, Role of slow temporal modulations in speech identification for cochlear implant users. *Int J Audiol* 53, 48–54 (2014).
99. L. A. Werner, Issues in human auditory development. *J Commun Disord* 40, 275–283 (2007).
100. E. Buss, J. W. Hall 3rd, J. H. Grose, M. B. Dev, Development of adult-like performance in backward, simultaneous, and forward masking. *J. Speech Lang. Hear. Res.* 42, 844–849 (1999).

Figures

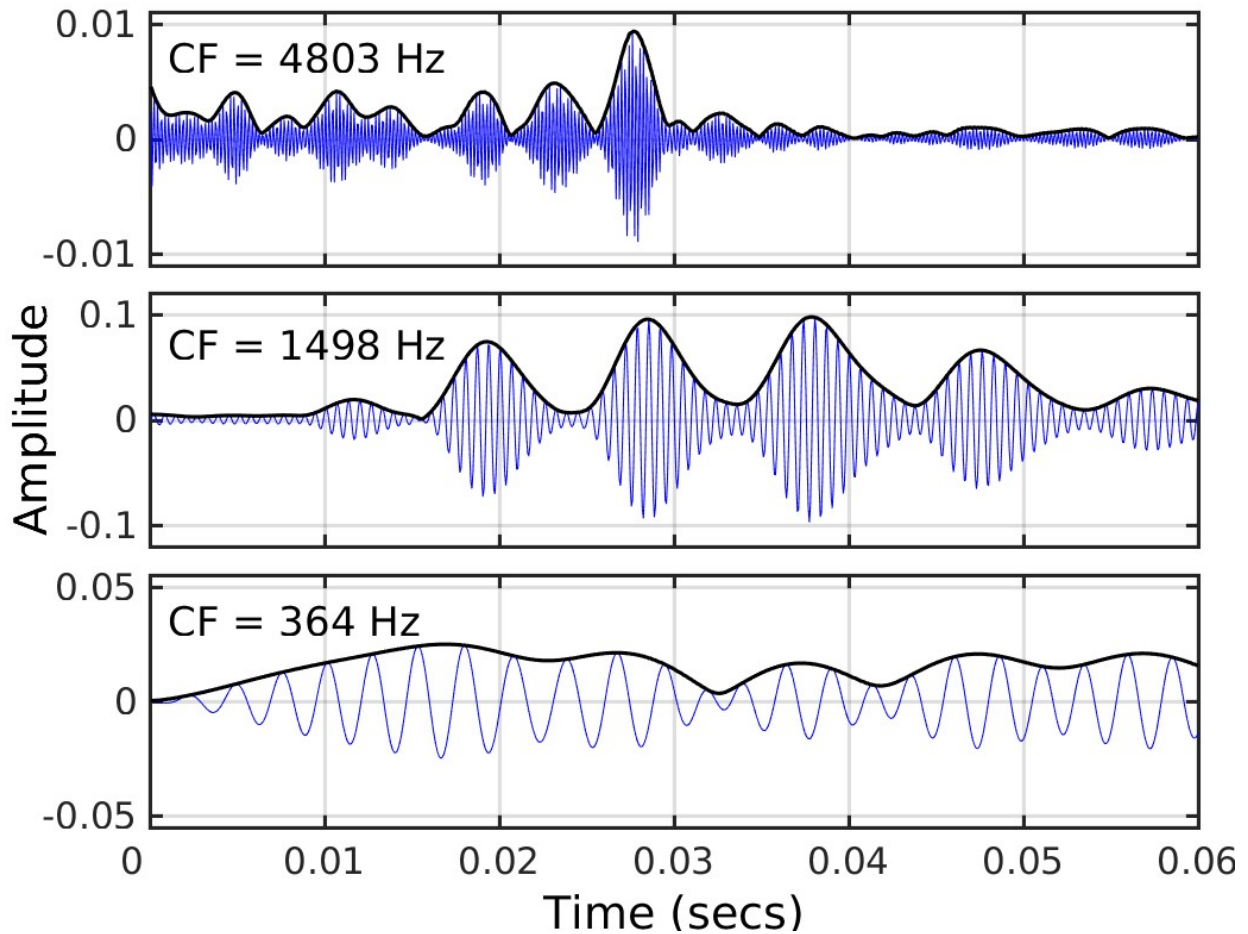


Figure 1. Outputs of simulated cochlear filters centred at 364, 1498 and 4803 Hz (from bottom to top) in response to a segment of a speech signal, the sound “en” in “sense”. These filter outputs are similar to the waveforms that would be observed at places on the basilar membrane tuned to 364, 1498 and 4803 Hz. For each centre frequency, the signal can be considered as a slowly-varying envelope imposed on a more rapid temporal fine structure. The envelope for each band signal is shown by the thick line. Figure and caption by Michael Stone - Own work, CC BY-SA 4.0 <https://commons.wikimedia.org/w/index.php?curid=65997712>