Fachgebiet für Bioanaloge
Informationsverarbeitung

Technichal University of Munich
School of Computation, Information and Technology
Bio-inspired Information Processing

Hauptseminar Bioanaloge Informationsverarbeitung

Parameters influencing the pitch evoked by amplitude-modulated pulse trains in cochlear implant users

Written by:
Natascha Niessen

Munich, August 20, 2023
SS 2023

Abstract

This review aims to give a systematic overview on the influence of the following parameters of amplitude-modulated pulse trains on the pitch-perception of cochlear implant (CI) users: carrier rate, modulation frequency, modulation depth, level and modulation shape. Pitch is important for music perception, distinguishing different voices and understanding words in tonal languages. Carrier rate and modulation frequency are both important for pitch perception as can be shown with two-dimensional stimulus spaces constructed with a multidimensional scaling procedure. In order to influence pitch by varying the modulation frequency, it is beneficial to use deeply modulated, harmonic stimuli with carrier rates above 800 pps. Furthermore, stimulating with sharper modulation functions such as exponential decay modulation (EDM) and at Mid-Level also makes it easier to control pitch with the modulation frequency. Polyphonic experiments show that CI users can tell how many amplitude-modulated stimuli are simultaneously applied to different electrodes.

1 Introduction

One important factor for improving the hearing experience of cochlear implant (CI) users is pitch. Pitch refers to the acoustic property of a sound allowing to rank sounds on a scale from low to high and is often associated to musical notes. However, improved pitch perception not only enables CI users to enjoy music. It also plays a role in distinguishing different voices and thus being able to focus on one speaker in an environment where multiple speakers are talking, e.g. in a restaurant. Furthermore, in tonal languages such as Mandarin words that might sound the same to a person unfamiliar with this language can have a completely different meaning depending on the pitch height that the speaker is using.

In normal hearing, there are several ways to encode pitch. One of them is based on frequency-location mapping along the basilar membrane, i.e. tonotopy. High frequency sounds cause the basilar membrane to vibrate near the oval window, in the basal region of the cochlear. Lower frequencies travel further along the membrane. This phenomenon is called place pitch and is imitated in CIs where the auditory nerve fibers are directly excited by an electrical stimulation. The nerve fibers are connected to the inner hair cells which are embedded in the organ of corti lying on top of the basilar membrane. The electrode location is used to evoke different pitch percepts while the stimulation rate is kept constant. Another way to evoke pitch in CI exists. In rate pitch, the stimulation rate is varied while stimulating at the same electrode.

Furthermore, pitch can be varied by changing the modulation frequency when stimulating with amplitude-modulated signals.

The studies that this literature review is based on analyze how pitch perception is influenced by varying the following parameters of amplitude-modulated signals: carrier rate, modulation frequency, modulation depth, level and modulation shape. As an outlook, a paper on polyphony with different electrode locations using amplitude-modulated signals was included.

2 Amplitude-modulated signals

The amplitude-modulated signals used to evoke pitch in the studies analyzed in this review consist of biphasic pulses of a certain carrier rate whose amplitude is being modulated according to a certain modulation frequency (see Figure 1). This type of modulation is called sinusoidal amplitude modulation (SAM). For the carrier signal we talk about rate in pulses per second (pps) because the current pulses are discrete, while for the modulation frequency we talk about frequency in Hz because the modulation of the amplitudes of the carrier signal is continuous. For amplitude-modulated signals the smoothed curve outlining the extremes of the amplitude-modulated signal is referred to as the envelope. In contrast, the term fine structure is used to describe the individual pulses of the amplitude-modulated signal.

Another important parameter of amplitude-modulated signals is the modulation depth. The modulation depth is a percentage representing the depth of the amplitude modulation. At 100% modulation depth, all amplitude values of the electrical dynamic range (EDR) are reached. At 50% modulation depth the lowest amplitude that is reached corresponds to half of the EDR. The modulation depth can also be expressed as the difference between the highest and lowest amplitudes in the pulse train in dB. In this case, for example $-6\ dB$ correspond to a modulation depth of 50% (see Figure 2).

3 Relative importance of carrier rate and modulation frequency

Before analyzing the effects of varying carrier rate and modulation frequency, the importance of these two parameters on pitch perception is described based on the findings of McKay and Carlyon (1999) who aimed to find out how many dimensions are sufficient to explain the pitch percept elicited by amplitude-modulated signals.

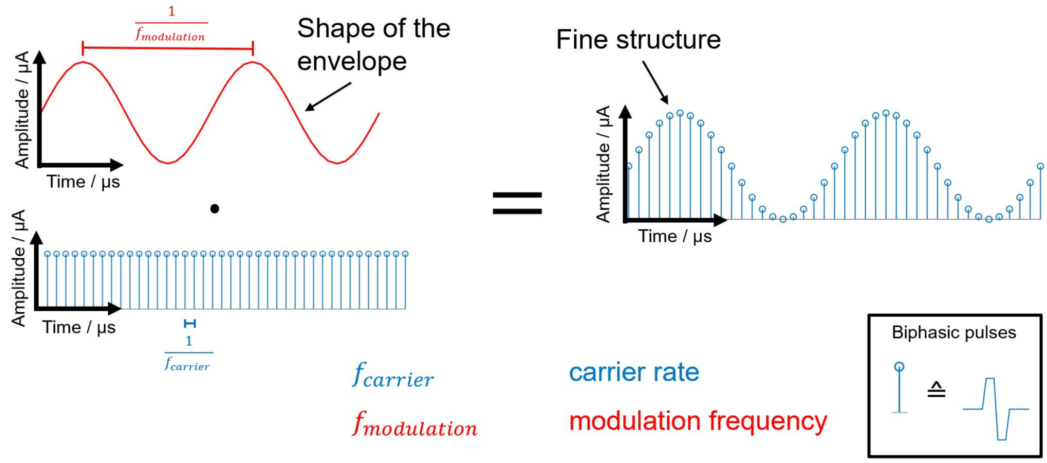

FIGURE 1 – Parameters and terms used to describe amplitude-modulated signals. The amplitude of the carrier signal (blue, bottom left), having a certain carrier rate $f_{carrier}$, is modulated with a certain modulation frequency $f_{modulation}$ (red, top left). In reality the pulses used in the present studies for the carrier are biphasic. For simplicity they are visualized as monophasic pulses. The smoothed curve outlining the extremes of the amplitude-modulated signal is called envelope. The actual shape of the signal with the individual pulses is called fine structure.

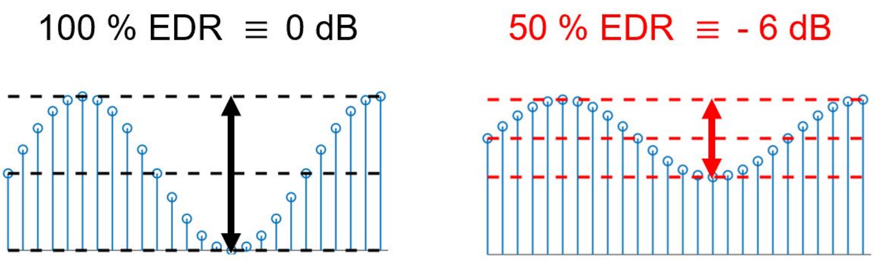

FIGURE 2 – Modulation depth of amplitude-modulated signals in percent and dB. The modulation depth is a percentage representing the amplitude modulation. At 100% modulation depth, all amplitude values of the electrical dynamic range (EDR) are reached. At 50% modulation depth the lowest amplitude that is reached corresponds to half of the EDR. The modulation depth can also be expressed as the difference between the highest and lowest amplitudes in the pulse train in dB. In this case, for example $-6\ dB$ correspond to a modulation depth of 50%

3.1 Parameters influencing pitch perception

In order to find the parameters influencing pitch perception, subjects were presented with a set of ten loudness balanced amplitude-modulated stimuli having carrier rates between 140 and 300 pps and modulation frequencies between 63 and 150 Hz. The two parameters were chosen such that the stimuli were all harmonic, i.e. the carrier rate is an integer multiple of the modulation frequency. This choice allows for more regularities in pitch perception and its dependence on carrier rate as explained later on in this review (McKay, McDermott, and Clark, 1994). All stimuli were loudness balanced prior to the experiments.

Subjects were presented with two stimuli from the previously described set one after the other. Subsequently, they were asked to determine how similar the stimuli sounded. The perceived dissimilarity had to be expressed on a scale from "Both stimuli sound exactly the same" to "The stimuli sound very dissimilar" by moving a cursor. This was repeated until all 10 stimuli had been compared with one another. The collected dissimilarity data was then analyzed with a multidimensional scaling procedure (ALSCAL, alternating least-squares algorithm) in order to attempt to find a structure in the set of dissimilarity measures. If e.g. two stimuli were rated by the subject as being very dissimilar, they would be placed very far apart in the multidimensional space that is constructed with the help of the ALSCAL algorithm. That way, the stimuli are assigned to specific locations in space such that the distances between the stimuli in this space match the measured dissimilarities as closely as possible.

The ALSCAL algorithm uses an iterative procedure developed based on the Kruskal algorithm and minimizes a so called stress function. In order to determine the appropriate number of dimensions influencing pitch perception, supposedly corresponding to different parameters in this scenario, the average reduction in stress across CI subjects can be compared when adding further dimensions to the constructed stimulus space. The reduction in stress from one to two dimensions was 0.16, compared to only 0.08 from two to three dimensions. This shows that one dimension, and therefore one parameter, is not enough to describe the dissimilarity distribution among the stimuli. However, adding a third dimension does not significantly reduce the stress. Thus, two-dimensional stimulus spaces were constructed for all CI users. The R^2 value expresses how much of the variance in the dissimilarity data is being accounted for by the two dimensions and can be used for comparisons across subjects. Other parameters such as modulation depth, electrode location, level and modulation shape are not relevant here as modulation depth, modulation shape and electrode location were kept constant and all stimuli were loudness balanced.

In Figure 3 such a stimulus space is displayed for one CI user. The different carrier rates are denoted by the symbol type, while the numbers above the symbols indicate the modulation frequencies in Hz. One can see that the stimuli are almost perfectly ranked based on the carrier rate in the vertical direction and based on the modulation frequency in the horizontal direction. However, some irregularities appear such as e.g. the stimulus with a carrier rate of 300 pps and a modulation frequency of 150 Hz that is located lower in the second dimension of the two-dimensional space than the other stimuli with the same carrier rate (filled squares, $f_{carrier}$ = 300 pps) and also lower than stimuli having lower carrier rates (open squares, $f_{carrier}$ = 252 pps).

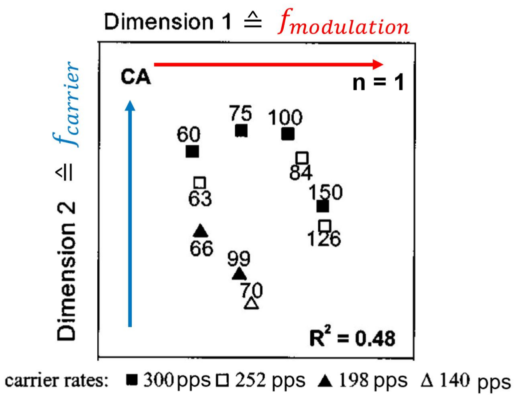

FIGURE 3 – Two-dimensional stimulus space for one cochlear implant user (subject code: CA) (modified from McKay and Carlyon (1999)). Different carrier rates are denoted by the symbol type, numbers above the symbols indicate the modulation frequencies. The stimulus space was constructed with the ALSCAL algorithm that assigns the stimuli to specific locations in space such that the distances between the stimuli in this space match the dissimilarities between stimulus pairs as closely as possible.

A ranking based on the carrier rate in the vertical direction and based on the modulation frequency in the horizontal direction can be observed in Figure 4 for all four CI users participating in this study. All together, the findings show the importance of both carrier rate and modulation frequency for pitch perception. Increasing the carrier rate leads to a higher pitch perception. The same holds for increasing the modulation frequency.

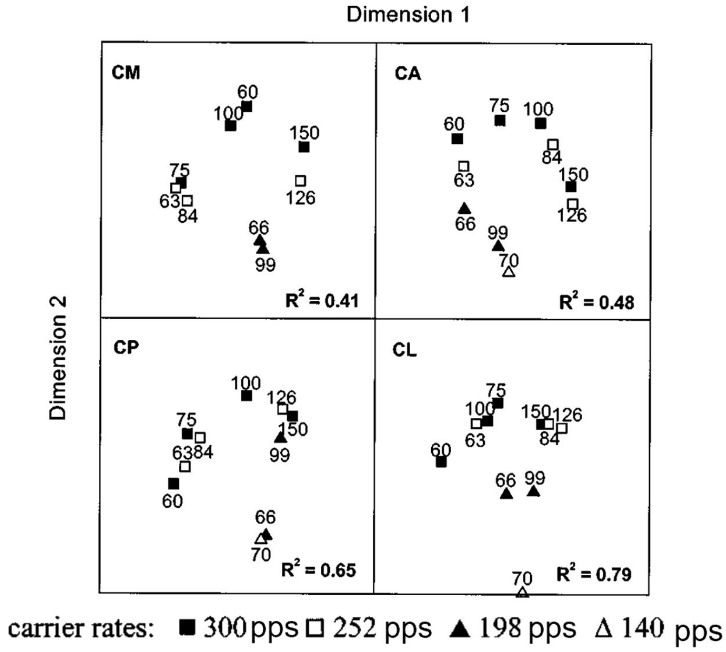

FIGURE 4 – Two-dimensional stimulus spaces for four cochlear implant users (McKay and Carlyon, 1999). Different carrier rates are denoted by the symbol type, numbers above the symbols indicate the modulation frequencies. The stimulus spaces were constructed with the ALSCAL algorithm that assigns the stimuli to specific locations in space such that the distances between the stimuli in this space match the dissimilarities as closely as possible.

3.2 Perceptual weighting of carrier rate and modulation frequency

The data for each individual subject was then evaluated compared to the other subjects. First, a group analysis was performed as in Figure 5 (top). Then, the relative importance of the two dimensions was determined for each individual subject based on the individual stimulus spaces and the one from the group analysis. In Figure 5 (bottom) the relative importance of the two dimensions is represented by the angle of the subject weighting vector. If the weighting vector of a subject is directed rather in the direction of the first dimension this indicates that the subject mostly used the modulation frequency to rank the perceptual differences among the stimuli. In contrast, if the vector is aligned more with the second dimension, the subject mostly used the carrier rate for pitch perception. For instance, subject CA uses predominantly the carrier rate for pitch perception rather than the modulation frequency as its vector in Figure 5 (bottom) is more aligned with the second dimension of the two-dimensional stimulus space. The parameter weighting vectors show a high inter-subjective variability.

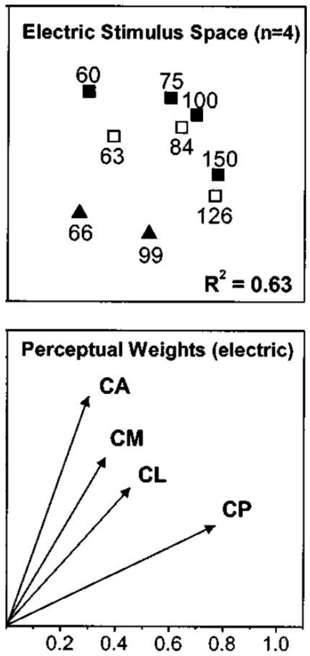

FIGURE 5 – Two-dimensional stimulus space for the group analysis (top) and perceptual weightings for all CI users (bottom) (McKay and Carlyon, 1999). The relative importance of the two dimensions was determined for each individual subject based on the individual stimulus spaces and the one from the group analysis. If the weighting vector of a subject is directed rather in the direction of the first dimension, the subject mostly used the modulation frequency to rank the perceptual differences among the stimuli. If the vector is aligned more with the second dimension, the subject mostly used the carrier rate for pitch perception. E.g. subject CA uses predominantly the carrier rate for pitch perception rather than the modulation frequency as its vector is more aligned with the second dimension of the two-dimensional stimulus space. The parameter weighting vectors show a high inter-subjective variability.

4 Carrier rate

As previously described, McKay and Carlyon (1999) have shown that the carrier rate is a parameter influencing the pitch perception of amplitude-modulated signals. McKay, McDermott, and Clark (1994) further analyzed the influence of the carrier rate on the subject´s ability to distinguish between sounds with a different modulation frequency. This is important in a scenario where pitch is evoked through electrical stimulation by varying the modulation frequency and keeping the carrier rate constant. The question is which carrier rate is suited best for evoking pitch through the modulation frequency.

4.1 Optimal carrier rate for evoking pitch through the modulation frequency

In order to find the optimal carrier rate for evoking pitch, McKay, McDermott, and Clark (1994) carried out a pitch-ranking experiment during which the subjects were presented with two stimuli one after the other and then had to decide which of the two was higher in pitch. The modulation frequency of one of the stimuli was always 150 Hz and the other 200 Hz. The carrier rate was the same for both stimuli and only varied among the trials. The order of the two stimuli differing only in modulation frequency was random. This pitch-ranking was then repeated several times in order to obtain the percentage of responses that the 200 Hz stimulus was ranked higher in pitch than the 150 Hz stimulus. This experiment was then repeated with different carrier rates to obtain the responses for all carrier rates from 400 pps to 12 kpps. All stimuli were loudness balanced prior to the experiments.

Figure 6 shows the performance of five subjects in this pitch-ranking experiment. A response percentage of 100% means that the subject always identified the stimulus with a modulation frequency of 200 Hz as higher than the one with a modulation frequency of 150 Hz. The following analysis is based on the performance of subject 1 displayed in the upper left panel. For carrier rates below 800 pps the subject´s responses show that the modulation frequency can hardly be used as an indicator for pitch at low carrier rates. For carrier rates of 500 pps the ranking is even reversed meaning that the subject always identified the 150 Hz stimulus as higher in pitch. The 200 Hz stimulus was mostly detected as being higher in pitch than the 150 Hz stimulus for carrier rates greater than 800 pps. These results were confirmed for four out of the five subjects tested. Except for subject 3 that was unable to distinguish between the two modulation frequencies regardless of the carrier rate, all subjects were able to rank a modulation frequency of 200 Hz higher in pitch than a modulation frequency of 150 Hz for carrier rates above 800 pps.

In summary, this pitch-ranking experiment suggests that for electrical stimulation with amplitude-modulated signals it is best to use carrier rates above 800 pps as for lower carrier rates the ranking is strongly dependent on the carrier rate. One possible interpretation is that for low carrier rates the envelope is only sparsely sampled and thus the envelope is less salient compared to the fine structure of the signal.

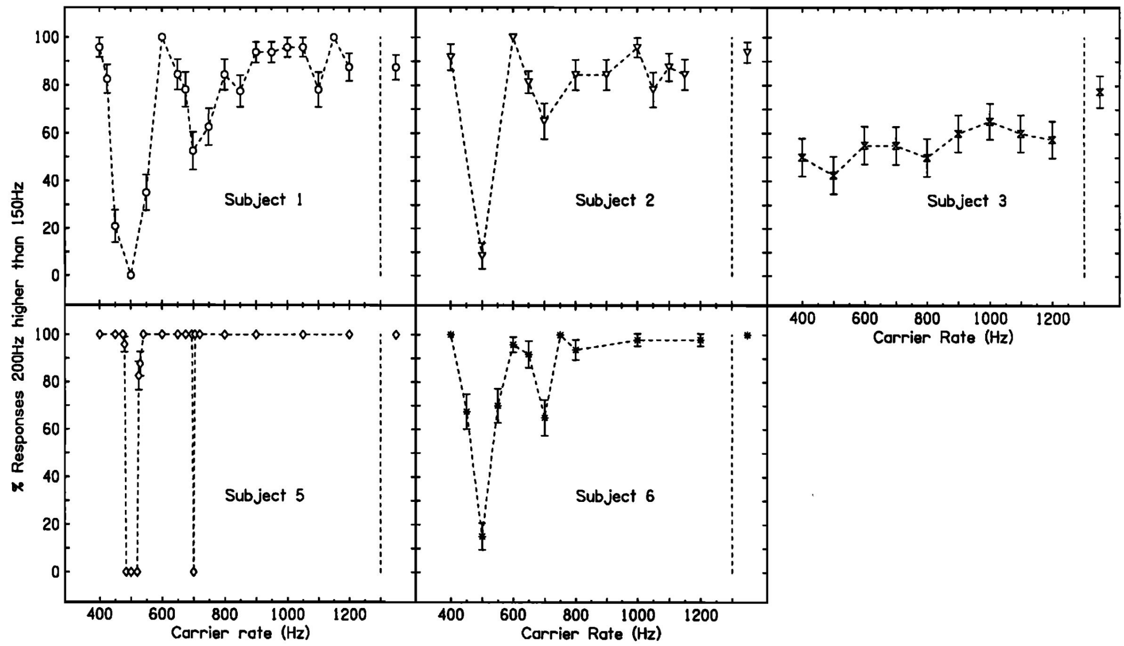

FIGURE 6 – Performance of five subjects in the pitch-ranking experiment for different carrier rates (McKay, McDermott, and Clark, 1994). A response percentage of 100% means that the subject always identified the stimulus with a modulation frequency of 200 Hz as higher than the one with a modulation frequency of 150 Hz.

4.2 Harmonic and inharmonic relationships between carrier rate and modulation frequency

Based on the results of the pitch-ranking experiment, McKay, McDermott, and Clark (1994) have put forward the hypothesis that for low carrier rates the ranking may be affected by harmonic or inharmonic relationships between carrier rate and modulation frequency. Harmonic stimuli are characterized by a carrier rate that is an integer multiple of the modulation frequency. In these stimuli, every period of the envelope, that is determined by the modulation frequency, has the same fine structure. In contrast, the fine structure of inharmonic stimuli exhibits additional periodicities as the fine structure in each modulation period is different from the preceding one. Thus, the hypothesis is that if the fine structure, and thus not only the envelope, is important for evoking pitch, then the pitch of harmonic and inharmonic stimuli should be significantly different.

In order to compare the pitch of harmonic and inharmonic stimuli, both were compared to unmodulated stimuli in a pitch-matching experiment. For this purpose, the subjects were presented with a modulated stimulus that was either harmonic or inharmonic and an unmodulated stimulus whose rate was varied until the subject reported that both sounds were equal in pitch. The rate for which they were equal in pitch is referred to in the following as the "Rate of Equivalent Pitch". In Figure 7 the Rate of Equivalent Pitch is plotted on the vertical axis against the carrier rate plotted on the horizontal axis. The upper dotted line represents a Rate of Equivalent Pitch of 200 Hz and the bottom dotted line a Rate of Equivalent Pitch of 100 Hz which are subsequently used for comparison.

The modulation frequency was, at first, chosen as 200 Hz and the carrier rates were chosen accordingly to result in an harmonic or inharmonic stimulus. Each point represents one modulated stimulus. The carrier rate of the stimulus is shown on the horizontal axis and the modulation frequency is denoted by the symbol type. It can be seen that the 200 Hz modulated harmonic stimuli (filled triangles) are matched to pitch rates equal or higher than their modulation frequency of 200 Hz. In contrast, inharmonic stimuli (empty triangles) are matched to other, mostly lower, pitch rates, some even being close to 100 Hz which corresponds to half of their modulation frequency of 200 Hz.

This might be due to a phenomenon called "beating of aliased components" which reveals that an inharmonic stimulus exhibits an additional 100 Hz-periodicity which can be referred to as pseudoperiod and which seems to be the frequency that is used for pitch-matching. In order to support this hypothesis, harmonic stimuli with a modulation frequency of 100 Hz and carrier rates of 500 and 700 pps were also matched (filled circles). The results in Figure 7 show that especially for the 500 pps inharmonic stimulus (leftmost unfilled triangle for each subject), the matched pitch rate is very close to the harmonic stimulus with a modulation frequency of 100 Hz at a carrier rate of 500 pps (leftmost filled circle). The 700 pps inharmonic stimulus (rightmost unfilled triangle) is matched to a rate intermediate to the two Rates of Equivalent Pitch for the 100 Hz (rightmost filled circle) and 200 Hz harmonic stimuli (rightmost filled triangle). Thus, this beating of aliased components seems to be more salient for a carrier rate of 500 pps. These findings show that it is rather not only the envelope of the stimulus but mainly the temporal fine structure that is used to effect the pitch-match and that harmonic stimuli show more regularities in how pitch is affected by changing the carrier rate.

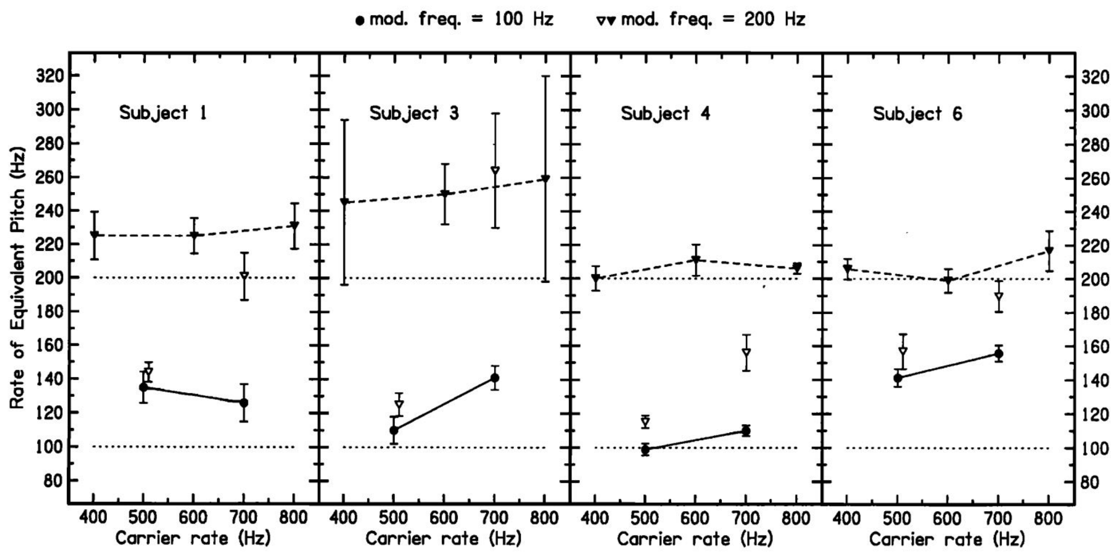

FIGURE 7 – Performance of four subjects in the pitch-matching experiment for harmonic and inharmonic stimuli (McKay, McDermott, and Clark, 1994). The subjects were presented with a modulated stimulus that was either harmonic or inharmonic and an unmodulated stimulus whose rate was varied until the subject reported that both sounds were equal in pitch. The rate for which they were equal in pitch is referred to as the "Rate of Equivalent Pitch" and is plotted on the vertical axis. Each point represents one modulated stimulus. The carrier rate of the stimulus is shown on the horizontal axis and the modulation frequency is denoted by the symbol type. The upper dotted line represents a Rate of Equivalent Pitch of 200 Hz and the bottom dotted line a Rate of Equivalent Pitch of 100 Hz.

4.3 Effect of the carrier rate on pitch-matching

In order to analyze the effect of the carrier rate on the previously described pitch-matching performance, McKay, McDermott, and Clark (1994) conducted the same pitch-matching experiment this time with each stimulus having a modulation frequency of 100 Hz and a carrier rate harmonically related to this modulation frequency. Thus, multiples of 100 pps between 200 and 1000 pps were used. For each modulated stimulus, the Rate of Equivalent Pitch was determined as before. Figure 8 shows that the effect of the carrier rate is not monotonic and that there are significant differences among subjects in how the carrier rate affects the Rate of Equivalent Pitch even for harmonic stimuli. This suggests that the perception of pitch does not only depend on the modulation frequency or a single subharmonic of the stimulus. Indeed, Figure 8 shows examples in which the matched rate does not correspond to any subharmonic of the carrier rate.

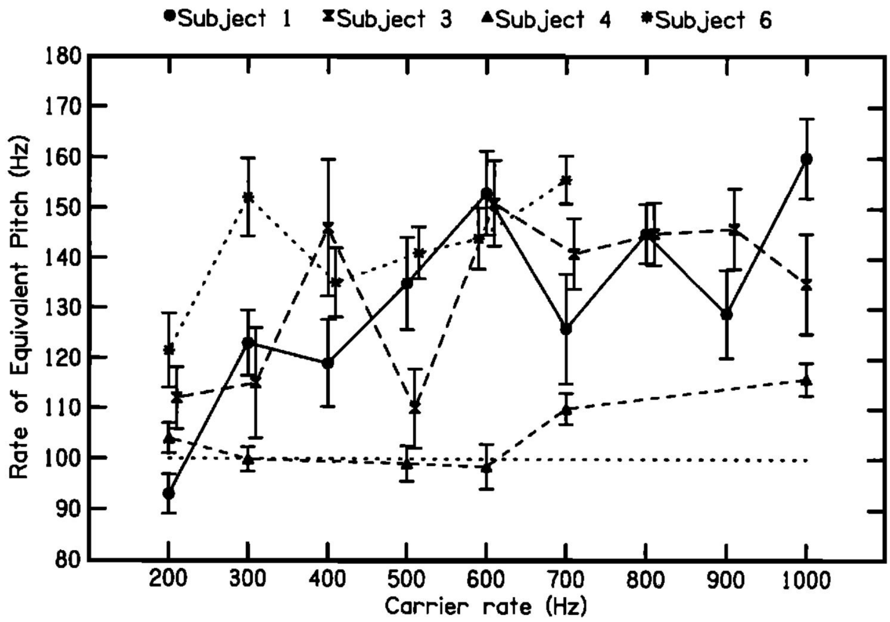

FIGURE 8 – Performance of four subjects in the pitch-matching experiment for different carrier rates (McKay, McDermott, and Clark, 1994). Each stimulus has a modulation frequency of 100 Hz and a carrier rate harmonically related to this modulation frequency. For each modulated stimulus, the Rate of Equivalent Pitch was determined.

In summary, in order to use the modulation frequency to evoke pitch in CI users it is beneficial to use carrier rates above 800 Hz that are harmonically related to the modulation frequency. Furthermore, pitch corresponds to an average of subharmonics of the modulation frequency or pseudoperiods appearing in the amplitude-modulated signals for inharmonic stimuli and thus not only on the modulation frequency.

5 Modulation frequency

McKay and Carlyon (1999) have shown the influence of the modulation frequency on pitch perception with the two-dimensional stimulus spaces described above. The way in which the modulation frequency impacts pitch has been further analyzed by Vandali et al. (2013). In an experiment similar to the previously described pitch-matching they presented two stimuli to the subject, the first being an amplitude-modulated stimulus with a certain carrier rate and modulation frequency and the second stimulus being an unmodulated stimulus with, initially, a rate equal to the modulation frequency of the modulated stimulus.

Subsequently, the rate of the unmodulated stimulus was decreased and increased in order to find the Rate of Equivalent Pitch at which the subject reported both stimuli to evoke the same pitch. Once this rate had been found, the Equivalent Pitch Rate Offset (EP-rate offset) was calculated as the difference between the Rate of Equivalent Pitch and the modulation frequency of the modulated stimulus. For instance, if the unmodulated stimulus initially sounded lower then its rate was increased until it matched the pitch of the modulated stimulus and the EP-rate offset was positive. All stimuli were loudness balanced prior to the experiments.

Figure 9 shows the EP-rate offsets for a modulation frequency of 100 Hz and a modulation frequency of 200 Hz. The different symbols (filled and unfilled circles) correspond to different modulation shapes. On the vertical axis the EP-rate offsets are plotted in semitones relative to the respective modulation frequencies. The dotted line represents an EP-rate offset of zero. For both modulation frequencies the EP-rate offsets are positive which means that the amplitude-modulated stimulus sounds higher than the unmodulated stimulus with a rate equal to the modulation frequency.

When comparing the EP-rate offsets between modulation frequencies of 100 and 200 Hz it becomes obvious that the EP-rate offset is higher for the lower modulation frequency. A similar trend could be observed for a modulation frequency of 300 Hz. In summary, this study shows that the amplitude-modulated stimulus sounds higher in pitch than the unmodulated stimulus with a rate equal to the modulation frequency. This difference in pitch decreases with increasing modulation frequency.

6 Modulation depth

Vandali et al. (2013) also analyzed the influence of modulation depth on the EP-rate offset as an indicator for pitch perception.

Figure 10 shows how the EP-rate offset varies with modulation depth here defined as a percentage of the electrical dynamic range (EDR) at a modulation frequency of 100 Hz. For the lowest modulation depth tested (12.5%), the EP-rate offsets are a lot higher than for deeper modulated signals. For very deep modulation, the rate of the unmodulated pulse that produces an equivalent pitch is nearly equal to the modulation frequency.

McKay and Carlyon (1999) analyzed the influence of modulation depth with

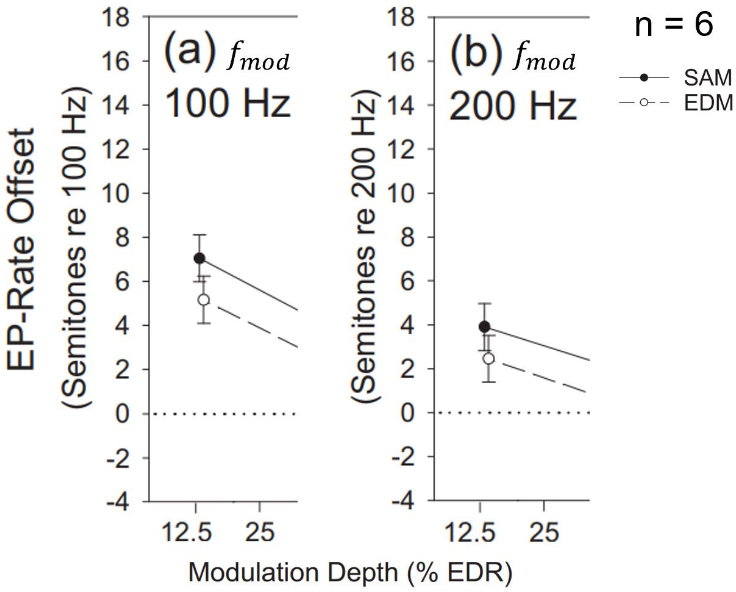

FIGURE 9 – Equivalent Pitch Rate Offsets (EP-rate offsets) in semitones relative to modulation frequencies of 100 Hz (left) and 200 Hz (right) (modified from Vandali et al. (2013)). The dotted line represents an EP-rate offset of zero. The symbols represent different modulation shapes (SAM: sinusoidal amplitude-modulation, EDM: exponential decay modulation).

the help of the previously described stimulus spaces and perceptual weight vectors (see section 3 and Figures 3, 4 and 5) showing which of the two parameters, carrier rate and modulation frequency, was more dominant in the subject´s pitch perception. The impact of modulation depth was only examined for one normal hearing subject for whom the two-dimensional stimulus spaces had been constructed in the same way as for the CI users only with acoustical instead of electrical stimuli. As the overall findings, stating the influence of carrier rate and modulation frequency on pitch, are similar for the normal hearing and the CI user groups, the findings on modulation depth for a normal hearing subject are included in this review.

In the two-dimensional stimulus space with the first dimension corresponding to the modulation frequency and the second dimension corresponding to the carrier rate in Figure 11, the 0 dB vector corresponds to the original, deep modulation depth used for the initial experiments on stimulus spaces. The modulation depth increases from the -6 dB vector to the 0 dB vector. The influence of the modulation depth on the perceptual weighting of carrier rate and modulation frequency is clearly visible. With increasing modulation depth the

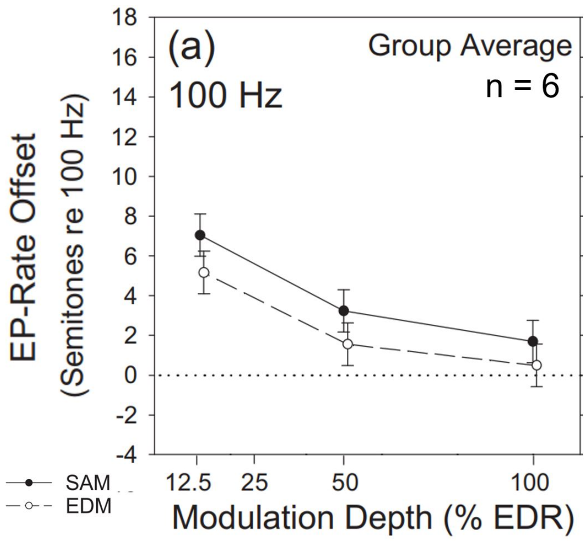

FIGURE 10 – Equivalent Pitch Rate Offsets (EP-rate offsets) for different modulation depths defined as a percentage of the electrical dynamic range (EDR) with a modulation frequency of 100 Hz averaged for six subjects. (Vandali et al., 2013). The symbols represent different modulation shapes (SAM: sinusoidal amplitude-modulation, EDM: exponential decay modulation).

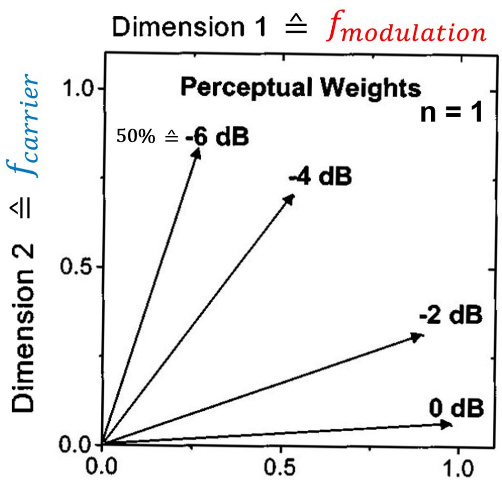

FIGURE 11 – Perceptual weight vectors of one normal hearing subject for different modulation depths in dB (McKay and Carlyon, 1999). The first dimension of the two-dimensional stimulus space is related to the modulation frequency and the second dimension to the carrier rate. The modulation depth increases from the -6 dB vector to the 0 dB vector.

subject relies more on modulation frequency than on carrier rate to compare the stimuli in pitch. This confirms the findings of Vandali et al. (2013) according to which the pitch of an amplitude-modulated signal converges to the pitch elicited by an unmodulated pulse train with a rate equal to the modulation frequency with increasing modulation depth.

All results considered, it is possible to state that the pitch of the modulated stimulus converges towards the pitch of the unmodulated stimulus with increasing modulation depth. It is thus beneficial to use deep modulation depths in order to modify the pitch mainly with the modulation frequency and less with the carrier rate.

7 Level

In the context of electrical stimulation, level is often used to refer to the loudness perceived by the subject but is also related to the current levels used for stimulation. The effect of presentation level was studied by Vandali et al. (2013) who compared the EP-rate offsets at Mid-Level ("mid-loudness) and C-Level ("comfortable loudness"). All their previously described experiments were conducted at C-Level which had been found through adaptively adjusting the level of the stimuli in a preliminary experiment. The Mid-Level was determined by reducing the stimulation level until the subject reported the loudness to be halved compared to C-Level.

The results in Figure 12 show that the EP-rate offsets at Mid-Level are overall smaller than the EP-rate offsets at C-Level. This shows that level indeed influences pitch perception and that stimulating at Mid-Level is beneficial for influencing pitch with the modulation frequency. Therefore, all stimuli were loudness balanced prior to the experiments on modulation frequency and modulation depth.

8 Modulation shape

The numerous existing CI coding strategies use different modulation functions. One modulation shape used for instance in the eTone strategy is called exponential decay modulation (EDM) and can be taken as an example for a sharper, narrower modulation function. The current versus time plot of the EDM is depicted in Figure 13. All experiments by Vandali et al. (2013) were conducted with EDM and sinusoidal amplitude modulation (SAM). Figures 9, 10 and 12 show that the sharper modulation function EDM evokes a lower pitch than SAM with EP-rate offsets for EDM being approximately one to two semitones

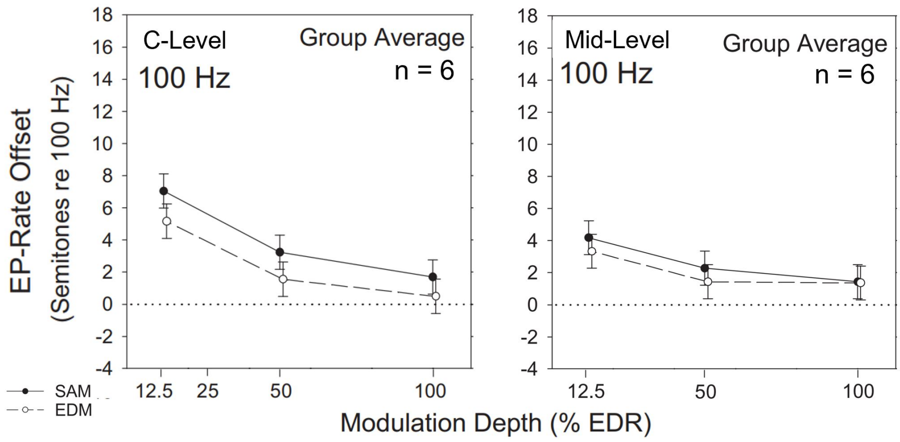

FIGURE 12 – Equivalent Pitch Rate Offsets (EP-rate offsets) for different modulation depths with a modulation frequency of 100 Hz at C-level (left) and Mid-Level (right) (Vandali et al., 2013). The Mid-Level was determined by reducing the stimulation level until the subject reported the loudness to be halved compared to C-Level (comfortable level) that had been determined in a previous experiment. The symbols represent different modulation shapes (SAM: sinusoidal amplitude-modulation, EDM: exponential decay modulation).

smaller than for SAM across all modulation frequencies, modulation depths and at Mid-Level as well as at C-Level. Thus, sharper modulation functions seem to be more beneficial to control pitch with the help of the modulation frequency.

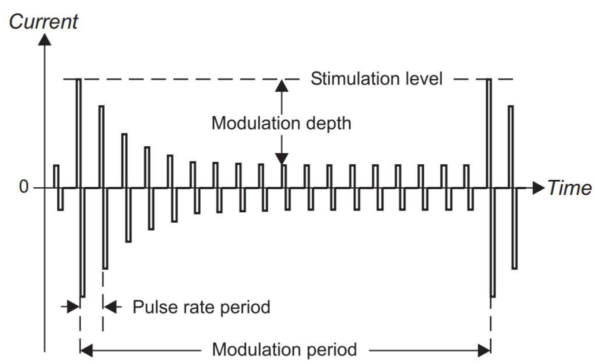

FIGURE 13 – The exponential decay modulation (EDM) plotted as current versus time (Vandali et al., 2013).

9 Electrode location

The use of amplitude-modulated stimuli at different electrode locations for evoking pitch has not yet been analyzed in detail. Penninger et al. (2014) published findings on CI users´ability to tell how many electrodes are being stimulated simultaneously using amplitude-modulated stimuli. Among other things, this ability is important for music perception where multiple streams played at the same time combine to a harmony which is not transmittable with current CIs. The ability to perceive several modulation frequencies applied at different electrode locations was tested in different conditions where one, two or three pitches were evoked simultaneously. The subject was then asked to identify how many pitches he perceived. The percentage of correct responses for each condition was documented. For all conditions a carrier rate of 5000 pps was sinusoidally modulated with 13 different modulation frequencies ranging from 261.63 to 523.25 Hz corresponding to musical notes C4-C5. The results in Figure 14 are averaged across the repetitions conducted for the different modulation frequencies. The stimulated electrodes were: Basal (E4), Middle (E11), Apical (E18).

In the 1-pitch condition only one electrode was stimulated at a time. As can be seen in Figure 14 the subjects performed better at identifying a single pitch for locations closer to the apical end of the cochlea. This may be due to the smaller difference between the rate pitch of the frequencies tested and the place pitch in the apical region compared with basal and middle electrodes. In general, apical neurons transmit pitch better than basal neurons in the low-frequency range (Middlebrooks and Snyder, 2010). For the 2-pitch condition the performance was significantly improved when the two stimulated electrodes had the greatest distance between each other (basal and apical). When the two electrodes are further apart the overlap in the neural populations stimulated by each electrode is reduced and thus it is easier to identify two separate pitches. Performance for the 3-pitch condition was as good as for the apical region stimulated with only one pitch. Thus the subjects were very well able to identify whenever three pitches were played simultaneously.

The experimental data was further analyzed for the 1-pitch condition to find if the modulation frequency had any impact on the subject´s performance. Similarly, for the 2-pitch condition the performance was compared for different modulation frequency pairs. Performance was quite similar across all modulation frequencies and frequency pairs. As rate and place pitch are deeply confounded in this additional polyphonic analyis it remains unclear if the subjects based their decision on place or rate pitch cues.

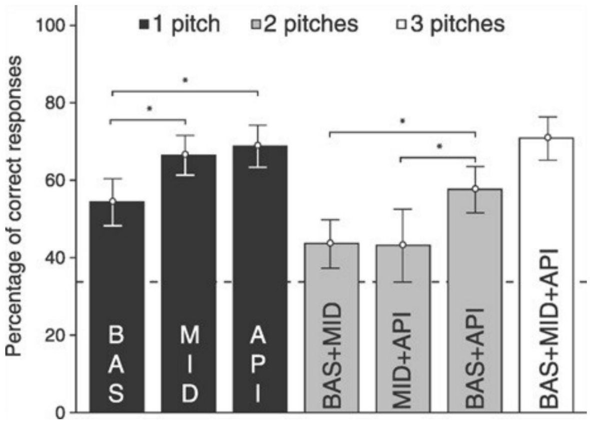

FIGURE 14 – Mean performance accuracy across subjects and combinations for the 1-,2- and 3-pitch condition testing CI users´ability to tell how many electrodes are being stimulated simultaneously using amplitude-modulated stimuli (Penninger et al., 2014). The stimulated electrodes were: Basal (E4), Middle (E11), Apical (E18). In each condition one, two or three electrodes were stimulated simultaneously. The subjects performed better at identifying the 1-pitch conditions for locations closer to the apical end of the cochlea. For the 2-pitch condition the performance was significantly improved when the two stimulated electrodes had the greatest distance between each other (basal and apical). Performance for the 3-pitch condition was as good as for the apical region stimulated with only one pitch.

10 Conclusion

The findings presented show the potential of amplitude-modulated stimuli to convey pitch. However, as explained in the section on level, loudness plays an important role in pitch perception and may therefore have influenced the results of the studies even though all studies included preliminary experiments for loudness balancing the stimuli used.

In general, it is hard to convey a low pitch with a cochlear implant as the insertion depth is limited and thus, the locations where low frequencies would be perceived are beyond the region maximally reached by the electrode array (Penninger et al., 2014). The reason for these short electrode arrays is the intersubject variability in cochlear anatomies. To eliminate this effect, Hochmair et al. (2015) proposed a new flexible long straight electrode that permits deep insertion reaching the apical region. This enables a better approximation of the tonotopically correct stimulation of sounds and allows to cover a wider range of cochlear locations. Furthermore, the wider spacing of channels helps to reduce channel interaction. This promising electrode array allows for improved speech understanding especially in noisy environments.

For further improving pitch perception, amplitude-modulated signals can be used to improve pitch resolution for CI users. Hereby, carrier rate and modulation frequency are both important for pitch perception as can be shown with two-dimensional stimulus spaces constructed with a multidimensional scaling procedure. In order to influence pitch by varying the modulation frequency, it is beneficial to use deeply modulated, harmonic stimuli with carrier rates above 800 pps. Furthermore, stimulating with sharper modulation functions such as EDM and at Mid-Level also makes it easier to control pitch with the modulation frequency. Polyphonic experiments show that CI users can tell how many amplitude-modulated stimuli are simultaneously applied to different electrodes. Such findings are of importance as current CIs are mainly focusing on speech processing and thus do not allow for perception of pitch necessary for listening to music or distinguishing different voices.

References

Hochmair, I, E Hochmair, P Nopp, M Waller, and C Jolly (2015). “Deep electrode insertion and sound coding in cochlear implants”. In: *Hearing Research*.

McKay, Colette M. and Robert P. Carlyon (1999). “Dual temporal pitch percepts from acoustic and electric amplitude-modulated pulse trains”. In: *Journal of the Acoustical Society of America*.

McKay, Colette M., Hugh J. McDermott, and Graeme M. Clark (1994). “Pitch percepts associated with amplitude-modulated current pulse trains in cochlear implantees”. In: *Journal of the Acoustical Society of America*.

Middlebrooks, JC and RL Snyder (2010). “Selective electrical stimulation of the auditory nerve activates a pathway specialized for high temporal acuity”. In: *Journal of Neuroscience*.

Penninger, RT, EKludt, CJ Limb, M Leman, I Dhooge, and A Buechner (2014). “Perception of polyphony with cochlear implants for 2 and 3 simultaneous pitches”. In: *Otology Neurotology*.

Vandali, A, D Sly, R Cowan, and R van Hoesel (2013). “Pitch and loudness matching of unmodulated and modulated stimuli in cochlear implantees”. In: *Hearing Research*.