Exploring rate-coded pitch perception in CI users vs. NH listeners to wavelet vocoder using the Oticon Medical Research Platform

Bradford C. Backus[1], Tobias Herzke[2]

1. Oticon Medical, Vallauris, France, 2. Hörzentrum Oldenburg gGmbH and Cluster of Excellence "Hearing4all", Germany

CONCLUSIONS

- Adding temporal jitter to wavelet vocoders can degrade rate-pitch performance to individual CI user levels.
- Adding jitter also solves the problem of hearing the dominant fixed-rate pitch.
- Wavelet vocoder may help develop TFS strategies for CI users in the future.

ABSTRACT

Pitch can be transmitted to CI users by varying stimulation rate as well as by changing the stimulating electrode [11]. However, the commercial rollout of such variable rate strategies has not yet capitalized on this opportunity. This is partially due implant limitations and partially due to the lack of development tools with which to realistically evaluate rate-coding ideas.

Vocoders are used give normal hearing people an idea of what listening through a CI is like. To be useful the performance of a CI user vs. vocoder listener must be aligned—this is difficult for rate-coding vocoders due to the complex action of the basilar membrane and the resulting resolved vs. unresolved harmonics [4]. We compare CI vs. vocoder performance when using a novel rate-and-place-coded music strategy through the Oticon Medical Research Platform (OMRP). The platform can (1) directly stimulate a CI user via a MIDI keyboard and (2) present a wavelet vocoded version (like the Bräcker vocoder [1]) of those electrograms to a normal listener.

AIM

Use the Oticon Medical Research Platform (OMRP) to quickly explore a variable-rate music CI coding strategy. Determine how well we can predict the CI performance on musical pitch tasks with a vocoder designed to do this.

BACKGROUND

There are 4 classes of vocoder available to CI researchers: (1) tone, (2) noise, (3) spiral [3], and (4) wavelet [1]. Each one has advantages and disadvantages, and each has different aptitudes for aiding CI research (Table 1).

The wavelet vocoder described here assigns one gammatone wavelet prototype to each electrode and places a wavelet of appropriate intensity at each electrode pulse event. Although conceptually simple, this conecpt produces overlapping wavelets when the stimulation rate is greater than ~1/10 the wavelet centre frequency. The interference can be constructive or destructive and requires additional level correction to compensate for the interference. This was done by amplifying each wavelet with a factor depending on the time that has elapsed since the last pulse on the same electrode was auralized. A table of correction factors is precomputed on startup.

Since each electrode is assigned a specific wavelet prototype, current spread can be simulated by changing the wavelet's damping parameter. But this method of simulating current spread is complicated by any overlapping wavelets. This problem was already addressed using machine learning techniques [2] and is the subject of previous work. Here we look more closely at using the wavelet vocoder to approach CI-user's pitch performance.

The wavelet vocoder can be a useful tool for CI researchers--one which lends itself to areas of vocoder investigation that were previously intractable.

Table 1: Investigations that can be carried out with different classes of vocoder

Investigation	Vocoder Type			
	Tone	Noise	Spiral**	Wavelet
Place Pitch		✓	✓	✓*
Temporal Pitch				✓*
Current Spread	✓*	✓	✓	✓*
Envelope ITD	✓	✓	✓	✓
Pulse ITD				✓
ILD	✓	✓	✓	✓

*** Investigation are non-ideal but possible with vocoder adjustment or advanced experimental paradigms or both. For example, temporal pitch investigations could be undertaken with the wavelet vocoder using high frequency electrodes under low stimulation rates (a factor of 10 below the centre frequency) and energy, acoustic masking of the stimulation rate's pitch place. Current spread has been investigated in tone vocoders using a technique to mix analysis modulations in neighboring carriers [4].**

**** Updates have been made to the spiral vocoder to transform it from a tonal version to wavelet which is not captured here [9].**

OMRP + MIDI

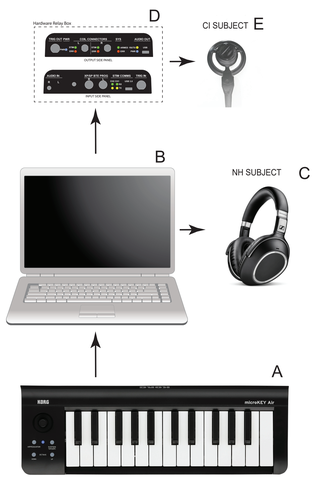


Figure 1: Panels A-E showing the hardware connections for the experiments allowing testing of CI subjects and normal hearing subject using the same system. Midi input from a keyboard (A) is processed with a music CI processing strategy by a computer (B) to produce electrodogram output which is auralized for normal hearing subjects using the wavelet vodocoder presented through headphone (C). The same electrodogram is sent through the hardware relay (D) box to an antenna coil (E) connected to a CI subject.

SYSTEM DIAGRAM

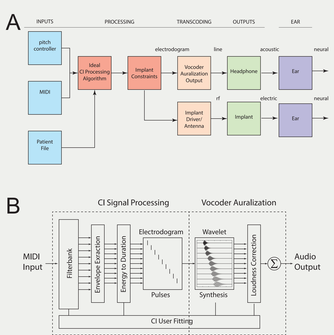


Figure 2: Architectural blocks of the experiment setup (Panel A), showing the inputs, CI processing and outputs. An electrodogram provides the input to the vocoder. Panel B shows how the electrodogram input is used to produce a pre-specified wavelet at each electrode pulse location (convolution). The CI user fitting is used in forward and reverse modes to enforce within band and energy equivalance between the input and output.

WAVELET VOCODER

$$g(t) = at^2 e^{-2\pi bt} cos(2\pi ft + \emptyset)\Big|_{t\geq 0}$$

$$BW(b) = \frac{3\pi}{8}$$

Table 2: **Oticon Medical Spiky CI fitting parameters**

E#	f_c (Hz)	f_{low} (Hz)	f_{high} (Hz)	BW (Hz)	ch (µs)	d_c (µs)
1	7422	5836	8008	1172	23.5	26.5
2	6315	4964	6836	1042	25	49.5
3	5374	4483	6264	891	27.5	60.5
4	4492	4152	4883	731	30	60.5
5	3776	3451	4152	651	35	63
6	3125	2799	3451	652	37	64
7	2604	2409	2799	390	39	67
8	2213	2018	2409	391	40	69
9	1888	1758	2018	260	41	70
10	1627	1497	1758	260	41	75
11	1432	1367	1497	130	39	74
12	1302	1237	1367	130	39	75
13	1172	1107	1237	130	37	74
14	1042	977	1107	130	38	74
15	911	846	977	131	38	72
16	781	716	846	130	37	70
17	651	586	716	130	38	68
18	520	456	586	130	36	64
19	380	326	456	130	33	56
20	260	195	326	131	28	48

EXPLANATION

On an individual wavelet basis, the parameter 'b' controls bandwidth and could be used to simulate a form current spread, however, the electric field decays exponentially (Ernst's function) not like a gammatone and additionally, when trains of wavelets are combined two confounds arise: (1) overlapping wavelet interference reduces the bandwidth and (2) ringing of auditory filters in response to the periodic pulsatile-like signals act--in a similar way--to produces a localized narrow-band signal, thereby converting any strong temporal periodicity into a sharp place-pitch--unnoticeably.

Q: Why do normal hearing people listening through a vocoder hear the stimulation rate while CI users never complain about it?

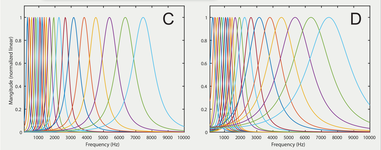


Figure 3: Panel A shows the gammatone wavelets assigned to each electrode described in Table 1. Panel B shows the wavelets when each bandwidth was increased by a factor of 2. Panels C and D show the frequency responses for A and B repectively. The frequency content can be controlled to simulate current spread for individual pulses, but when pulses are combined into a pulse train the spectrum is narrowed again (locally) by the action of the basilar membrane (see below section on Jitter).

CONTROLLING PITCH

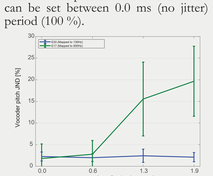
Each wavelet's frequency content is generally chosen to match its electrode frequency mapping (Fig. 4), but other mappings (e.g. the physiological place of the electrode or current spread effects are possible). Playback of individual non-overlapping wavelets evoke a pitch percept corresponding to their frequency content as 'place pitch'.

Variable Stimulation Rates

Repeated stimulation of the same electrode can evoke a rate pitch percept in cochlear impant users (for rates < ~300 Hz). Existing literature suggests that one may be able to use this idea by adapting the stimulation rate to the acoustic pitch for low-frequency electrodes. On receiving such variable-rate electrodogram input, the wavelet vocoder will place the wavelets at times driven by the variable rate. If the wavelets overlap, interference alters the carrier frequency so that the vocoder auralizes the rate-related pitch. But while rate and place-pitch are independent for CI users, it is not possible for a vocoder to preserve this separation for NH listeners.

Fixed Stimulation Rates

Traditional CI coding strategies stimulate implants with a fixed repetition rate. In consequence, some electrodes (but not all) are stimulated with this repetition rate. A corresponding pitch is perceived by listeners of the vocoder output, but not by CI users (it may be too fast or simply ignored). Our vocoder supports adding temporal jitter to the wavelets (Fig 5). A random delay for each wavelet placement between 0.0ms and t_max, where t_max can be set between 0.0 ms (no jitter) and the fixed repetition period (100 %).

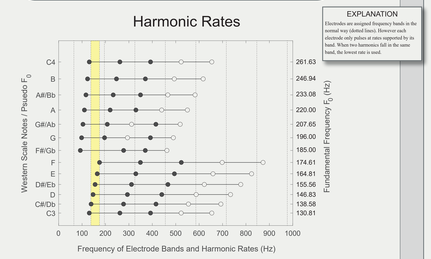


EXPLANATION

To investigate if adding temporal jitter can degrade the stimulation-rate pitch, a 3-AFC JND experiment with 5 NH listeners was conducted. Subjects had to identify the interval in which the average rate deviated from the other two intervals. We adapted the deviant's average-rate step to determine the 'pitch' JND [8]. This was performed for base two stimulation rates (130Hz and 500Hz) on their corresponding electrodes each with 4 jitter t_max settings (Jitter was taken by random sampling of a uniform distribution between 0 and t_max)

Figure 4: Adding jitter to the electrodogram before vocoder auralization can degrade rate-related pitch perception (JNDs) in NH listeners. The JND is a function of the amount of jitter, the original fixed stimulation rate and the electrode to which it is applied (wavelet).

HARMONIC RATE CODING



EXPLANATION

Electrodes are assigned frequency bands in the normal way (dotted lines). However each electrode only pulses at rates supported by its band. When two harmonics fall in the same band, the lowest rate is used.

Figure 5: In harmonic rate coding, each electrode is assigned range of stimulation rates (dotted lines). Normally this is chosen to coincide with that electrodes frequency mapping. For music, the harmonic stimulation rates are taken directly as the harmonics (or overtones) of the notes themselves (circles). In this experiment only the first 3 harmonics (closed circles) were used. In cases where the fundamental is too high to be represented by the first electrode, an octave shift is introduced (e.g. between F and F#) on the basis of octave equivalence in music. For fixed rate coding no such shift is used. A yellow highlight shows where rate sensitivity occurs for S1, meaning rate cues are likely restricted to the first harmonic represented on the most apical stimulating electrode for S1.

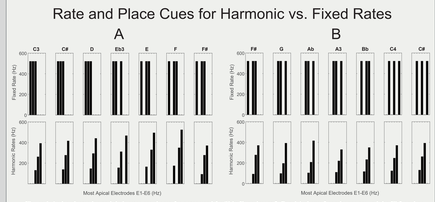
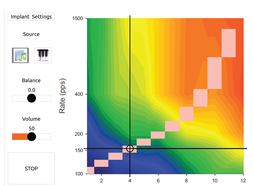


Figure 6: Active electrodes and their rates are shown for notes used for S1 in Experiment 2. Panel A shows 3 semitones around the Eb3 anchor and Panel B shows the same for the A3 anchor. Bars depict the rate and place cues available to S1 in Experiment 2 (e.g. no place cues exist for harmonic coding in B). Note: a 2 electrode shift is shown for the harmonic-rate coding because S1 did not have rate sensitivity in her most apical electrode E20 (see Experiment 1) and reported strange sensations in E19. Relative loudness was set as: H1 = 100%, H2 = 75%, H3=50%.

EXPERIMENT 1



EXPLANATION

Subjects interact directly with the stimulation parameters as they are being presented. This allows very fast determination of where they can perceive rate discrimition in acute testing. We used this information to orient ourselves to the other results.

Figure 7: OMRP Realtime touchscreen control panel allowing subjects to change the stimulating electrode number (x-axis) and its rate (y-axis) in order to determine sensitivity to rate in each of 4 apical electrodes.

Single Electrode Rate Perceptions

There was a bifurcation of sensation depending on stimulation rate for CI subject S1. For stimulation rates below 200 Hz, S1 reported the sensation of 'vibration' that was very low frequency. S1 reported that between 130-150 Hz there was a rapid change in the 'pitch' of this vibration but little change from 100Hz-130Hz, and almost no perceptible change for rates beyond 200Hz. The normal hearing listener perceived all rate changes from 100-1000 Hz as pitch changes.

Table 2: Self-reported sensitivity to rate during initial single electrode acute testing at the 6 most apical electrodes for CI subject S1 (top table) and NH subject S2. For S1 E2B did not consistently test as hearable or working and therefore a 2-electrode shift was introduced for the music discrimination experiment for the CI user S1.

***NH + Jitter showed a artifact problem at higher frequencies**

		E20	E19	E18	E17	E16	E15
CI	L	n/a	140	160	130	130	110
	H	n/a	160	186	180	190	200
NH	L	100	100	100	100	100	1000
	H	1000	1000	1000	1000	1000	1000
NH+ Jitter	L	150	200	240	280	500	300
	H*	1000	1000	1000	1000	1000	1000

EXPLANATION

Subjects manipulated the stimulation rate for a given electrode (by sliding their finger up and down on the Landsberger-style [11] test panel). They reported the low-high thresholds between which they could hear differences in the stimuli due to this rate manipulation. The CI subject S1 had a very small range of rates over which she could detect these changes compared with NH subject S2. The basilar membrane can convert intended time cues into place cues showing the challenge for vocoders to simulate what CI users hear. But we can add jitter to come closer to CI users experiences.

EXPERIMENT 2

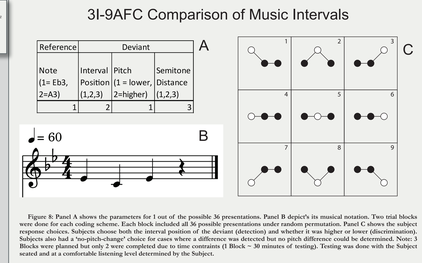
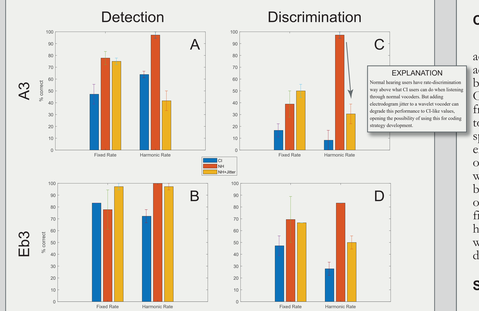


Figure 8: Panel A shows the parameters for 1 out of the possible 36 presentations. Panel B depice's its musical notation. Two trial blocks were done for each coding scheme. Each block included all 36 possible presentations under random permutation. Panel C shows the subject response choices. Subjects choose both the interval position of the deviant (detection) and whether it was higher or lower (discrimination). Subjects also had a 'no-pitch-change' choice for cases where a difference was detected but no pitch difference could be determined. Note: 3 Blocks were planned but only 2 were completed due to time contraints (1 Block ~ 30 minutes of testing). Testing was done with the Subject seated and at a comfortable listening level determined by the Subject.



EXPLANATION

Normal hearing users have rate-discrimination way above what CI users can do when listening through normal vocoders. But adding electrodogram jitter to a wavelet vocoder can degrade this performance to CI-like values, opening the possibility of using this for coding strategy development.

Figure 9: Figure shows how adding 40% jitter to the electrodogram before auralization--40% jitter means that interpulse intervals are chosen at random from uniform distribution of intervals extending up to +/- 40% of the original stimulation interval--reduces both detection and discrimination for NH users to near CI performance. This is especially clear when rate is the only available cue, e.g. A3 (Panels A and C). Adding jitter has no effect on fixed frame rate coding, just as one would expect as these codes include only electrode or place-based cues. Note: error bars are not true error bars but show the values for the 1st and 2nd block runs (N=18) with the coloured bars showing the average (N=36). Where there are no error bars the 2 runs produced the exact same % correct values.

EXPERIMENT 3

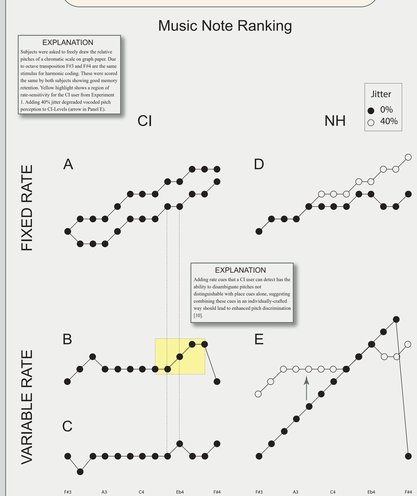


Figure 10: Panel A showing perceived pitch changes drawn by the subject S1 while playing chromatic scale from note F3# to F4# listening with a standard 520 Hz fixed rate frequency for all electrodes (H1 = 100%, H2 = 75%, H3=50%). Panel B shows the same type of plot, but the stimulation was produced using harmonic rate coding, including a shift of 2 electrodes toward the base (H1=100%, H2=75%, H3=50%). Note that due to octave transposition of this coding scheme, F3# and F4# produce the same stimulation pattern and were also scored identically. Panel C shows the same plot without electrode shifting for comparison. Panels D and E show the ranking for a normal hearing subject listening through a vocoder to the fixed-rate and harmonic-rate schemes repectively. Open circles show where 40% jitter was added. The yellow highlighted area indicates where the CI Subject has rate-sensitivity according to Experiment 1.

DISCUSSION

Pitch and Jitter

Investigations of CI temporal pitch or temporal fine structure (TFS) coding strategies is challenging using vocoders and NH listeners. This is due to the conversion of temporal-pitch cues to place-pitch cues by the action of the basilar membrane. Here we demonstrate that we can degrade vocoder pitch performance in NH listeners to CI levels by adding jitter to a wavelet vocoder. This technique necessarily degrades both the temporal and temporally generated place cues simultaneously, but it does so in a meaningful way.

Although this vocoder produces isolated temporal pitch output, the place cues that co-arise in the basilar membrane would have to be effectively masked (which is possible) to isolate the temporal ones to produce direct comparison to a CI user, so there is still no good way to study temporal CI pitch alone with vocoders without very careful thought [9].

Current Spread

Several techniques have been employed in vocoders to account for electrical current spread. Fu and Nogaki 2004 [5] added current spread to a noise vocoder by using widened bands with slow roll-off triangular and overlapping slopes. Oxenham and Kreft 2017 cross-mixed modulation outputs from neighboring analysis bands into the tones carriers of a tone vocoder to simulate an 8 dB/Oct monopolar current spread [4]. And Grange, Culling et al (2017) invented an entirely new class of vocoder with decoupled analysis and output stages that utilized many separate tone carriers into which a smaller number of analysis modulation outputs could be mixed [9]. The wavelet vocoder also gives some control over the current spread but our current choice of gammatone filter shape doesn't mimic the current spread and 8 dB/Oct is hard to achieve here. In the future, we may use a more apt wavelet or explore mixing the analysis outputs as others have done to simulate current spread better.

Supporting New Vocoder Research Areas

The wavelet vocoder presented here delivers new capabilities to CI researchers. For example, in binaural (stereo) form, it presents pulse-by-pulse or carrier ITDs on high frequency electrodes at standard pulse rates and for lower frequency electrodes this can be achieved with reduced stimulation rates. Studying reduced-rate strategies is a hot area. CI research on the topic of ITDs consistently shows that the best ITD sensitivity is found with pulse trains of 50 to 100 pulses-per-second (pps), with star CI listener's having detection thresholds of 100 to 200 µs--although subject variability is large [6]. Developing low-rate ITD-enabled binaural coding strategies is one area where the wavelet vocoder could provide insight on how to incorporate carrier ITDs with envelope ITDs [7].

This vocoder has been developed as part of the Oticon Medical Research Platform (OMRP) available now to researchers.

ACKNOWLEDGEMENT

Many thank to Andrew Oxenham, Jacques Grange and our colleagues at Oticon Medical, Audio3 Ltd, and Hörzentrum Oldenburg for their inspiration and inputs.

REFERENCES

1. T. Brücker, V. Hohmann, B. Kollmeier, M. Schulte, Simulation und Vergleich von Sprachkodierungsstrategien in Cochlea-Implantaten [Simulation and comparison of speech coding strategies in cochlear implants]. Zeitschrift für Audiologie / Audiological Acoustics, 48(4), 158–169, (2009).

2. C. E. Rasmussen and C. K. I. Williams, Gaussian Processes for Machine Learning, MIT Press, ISBN 0-262-18253-X, (2006).

3. J. Grange, J. Culling, et al. Cochlear implant simulator with independent representation of the full spiral ganglion. J. Acous. Soc. Am 142, EL484; doi: 10.1121/1.5009602. (2017)

4. A. Oxenham and H. Kreft. Speech Perception in Tones and Noise via Cochlear Implants Reveals Influence of Spectral Resolution on Temporal Processing. Trends in Hearing, Vol. 18: 1–14, (2014)

5 Q. Fu and G. Nogaki. Noise Susceptibility of Cochlear Implant Users: The Role of Spectral Resolution and Smearing. JARO 6: 19–27, (2004).

6. van Hoesel, R. J., Jones, G. L., and Litovsky, R. Y. "Interaural time delay sensitivity in bilateral cochlear implant users: Effects of pulse rate, modulation rate, and place of stimulation," J. Assoc. Res. Otolaryngol. 10, 557–567, (2009).

7. V. Noel and D. Eddington. Sensitivity of bilateral cochlear implant users to fine-structure and envelope interaural time difference. J. Acous. Soc. Am 133-4, (2013).

8. Ewert, S. "AFC - A modular framework for running psychoacoustic experiments and computational perception models," in Proceedings of the International Conference on Acoustics AIA-DAGA Merano, Italy, pp. 1326-1329, (2013).

9. Grange J. A., Culling J. F. "Decoding the Auditory Nerve. Implications for Cochlear Implant Technology" Poster 1215 in CIAP 2019 Lake Tahoe, CA, USA, (2019).

10 Susan R.S. Bissmeyer, Raymond L. Goldsworthy, Combining Place and Rate of Stimulation Improves Frequency Discrimination in Cochlear Implant Users, Hearing Research, Volume 424, (2022).

11. Landsberger DM, Vermeire K, Claes A, Van Rompaey V, Van de Heyning P. Qualities of Single Electrode Stimulation as a Function of Rate and Place of Stimulation with a Cochlear Implant. Ear Hear. May-Jun;37(3):e149-59 (2016) .