# *The (re-)invention of the modern cochlear implant? (Circa1965-1990)*

Mark White

The following pages contain the Aug. 19th 2017 draft. This manuscript will be periodically updated. The most recent draft can be downloaded from: http://cochlearimplant.us/some_of_sciences_contributions_with_excerpts.pdf

# *The (re-)invention of the modern cochlear implant? (circa1965-1990)*

Mark White

## ABSTRACT

How have the sciences helped us create modern cochlear implants? Answering this question should help guide our future efforts to improve cochlear implants. It should motivate increased efforts in the basic and applied sciences. It also should improve our awareness of how many researchers played key roles in the invention of modern-day cochlear implants.

This manuscript only describes a subset of scientific research that directly led to the creation current-day cochlear implants. Here we examine the science that led directly to the following inventions: (1) The Spectral Peak Picker processing strategy, (2) The "Virtual Channel" and "Current Steering" processing strategies, and (3) The Continuous Interleaved Sampling (CIS) processing strategy.

## INTRODUCTION

The sciences were essential in the creation of modern cochlear implants. Many researchers from around the world have played key roles in the invention of the modern cochlear implant. Here we argue that a strong focus on research is essential for significant improvements in cochlear implant performance.

By understanding the essential importance of past research in the development of current implant technology, the importance of future research will become more evident. The research leading to the creation of the following 3 inventions will be examined in this manuscript:[1]

1. The Spectral- or Channel-Peak-Picker processing strategy.
2. The "Virtual Channel" and "Current Steering" processing strategy.
3. The Continuous Interleaved Sampling (CIS) processing strategy.

This manuscript describes only a subset of scientific research that led to modern-day cochlear implants.[2] In addition, some more recent and important areas of cochlear

---

[1] Over the last 25 years B Wilson has stated that he invented these strategies. Wilson states [Wilson2015a], [Wilson2016]: "Prof. Wilson is the inventor of many of the speech processing strategies used with the present-day cochlear implant, including the continuous interleaved sampling (CIS), spectral peak picking (e.g., "n-of-m"), and virtual channel strategies, among others."

[2] For example, the contributions to the design of very early cochlear implants by scientists and clinicians has been documented by [MuMi2013] and others.

implant research will not be described – even though these areas are as important, and are likely to become more important.[3]

Most scientific literature, such as historical summaries and primary source documents, have some inaccuracies. In very rare cases the inaccuracies are severe. Because of such severe inaccuracies, many researchers are largely unaware of the science that led to current cochlear implant processing strategies. -- and the scientific contributions made by many of their colleagues.

There are a number of reasons why such severe inaccuracies become part of the scientific literature: (1) Authors of secondary publications may have "taken-at-face-value" source documents that contain such severe inaccuracies (e.g., the source document's authors failed to "credit" or "reference" their sources). Clearly, the author(s) of such source documents are far more culpable than are the authors of the secondary referring publications. Authors of such secondary publications, or summaries, may have been quite unaware of prior research (a) because they were not active in the research area during the period in question, or (b) because they may not have consulted a diverse-enough set of researchers that were active during the period, or (c) because relevant historical documents were not readily available or accessible.

Access-to and analysis-of primary source documents is fundamental to accurately understanding the history of cochlear implants: The full-text of all the documents referenced in this manuscript can be directly accessed via web links in the "References" section. – except that the full-text of some books/monographs are not included.

> Excerpt from an email from Philip Loizou's sent to Mark White in 2011: ...your website that contains a wealth of historical information about the development of CIS... and wish I had access to it when I wrote my introductory survey article in IEEE Signal Proc. Magazine (1998).

In addition to examining the scientific origins of these inventions, we will examine why many researchers and clinicians have been unaware of the origins of modern-day cochlear implants.[4]

---

[3] Fortunately, there are useful summaries of some of these other areas. In some cases, a significant literature exists. For other laboratories, the historical literature is sparse. For example, an overview of some of UCSF's early contributions has only recently and belatedly been published: "Early UCSF contributions to the development of multiple-channel cochlear implants" [Me2015]. I believe it is interesting and worthwhile to read, although I believe it contains some errors.

[4] The evidence presented here indicates that the 3 cochlear implant processing strategies were not invented by any single person or group but were invented by many researchers from around the world. It is hoped that the contributions of these researchers will be effectively "unmasked" here.

## PROCESSING STRATEGIES

### 1. Spectral or Channel Peak Picker

In 1957, Peterson and Cooper [PeCo57] proposed the “channel peak-picker” strategy for improving channel vocoders. White [Wh78] proposed using their idea for cochlear implant processing. Early-on the “channel peak-picker” or “spectral peak-picker” strategy (also later called the “N-of-M” strategy) was described by 2 cochlear implant groups ([ToDoBlCl83], [Wh83b]). This illustrates how basic studies of speech perception and communications led to the invention of “channel peak-pickers” in cochlear implants. Channel peak-picking in cochlear implants has been used in an effort to reduce the "information-load" on the CNS by emphasizing or picking the most salient channels for stimulation. It has also been used for the purpose of reducing channel interactions. Excerpts:

> [Wh78] p. 80: If necessary, further reductions in the acoustic signal's redundancy can be obtained. The spectrum channel vocoder's channel signals are not completely independent. One channel vocoder variation called a peak-picking vocoder, attempts to reduce this dependence (Peterson and Cooper, 1957; Flanagan, 1972). It operates by transmitting three to five or more channel signals which, at any instant, represent local maxima of the short time spectrum. The identities of the "picked" maximum channels and their amplitudes are signaled to a conventional multi-channel vocoder synthesizer. Thus, at any one time only a few channels of the synthesizer are activated. This version of a peak-picking channel vocoder is very similar to a formant vocoder with rather coarse formant frequency quantization. Formant vocoders generally utilize knowledge about the speech production mechanism to a greater extent than the other vocoder realizations. For instance, information about the filtering capacity of the human vocal tract is utilized to determine the number of formants and the number of poles required to model the tract for formant trajectory calculations. Also, formant trajectories are known to be reasonably continuous functions because of the continuity in the motion of the vocal tract. These and other constraints are used in formant vocoders to reduce channel transmission bandwidth requirements without severely compromising the intelligibility of the speech. However, the formant vocoder does strip some useful information from the input signal. ...
>
> [ToDoBlCl83] p. 994: The results indicate that, in addition to the presentation of the second formant by activation of only one [of the 10] electrode[s] during a stimulus period in a speech processor, as described by Tong et al (5), speech information such as the first formant may be presented to implant patients by two-electrode stimulation.
>
> [Wh83b] last page: With the compression systems illustrated in figures 5 and 6 it is possible to "expand" or "contract" the across-channel spectral representation. In one extreme example of this expansion, only those channels "centered" at the spectral peaks (e.g., at the formant frequencies of the speech signal) would receive suprathresho1d stimulation. Such a compression system could be implemented by using a large amount of cross-coupling between "nearby" channels. The cross-coupling would be very highly "inhibitory". The channel receiving the largest amplitude signal would effectively "turnoff" all of the other channels which are strongly and "negatively" coupled to this channel. This type of processing is similar to that used in peak-picking channel vocoders described by Peterson and Cooper (1957). This type of processing is also very analogous to the visual edge enhancement process first described by Mach.

The channel peak picker was also described by Wilson in his “best and final offer” submitted in response to NIH's July 5, 1983 RFP. At that time he called it a “formant peak picker.” In this same document, Wilson referenced the [ToDoBlCl83] paper.

> [Wilson1985/83] page A-3-12: In effect the formant "peak picker" of the second design enhances the signal-to-noise ratio for transmission of speech parameters essential to intelligibility. ... The formant-vocoder approach of the second design would also permit a further reduction in the number of electrodes required to encode the speech parameters.

Additional information about the history of the "Channel Peak Picker" can be found here.

*Details: Analysis of references and attribution of invention.* If the Spectral Peak Picker processing strategy was not invented by Wilson, had he simply re-invented it? – because he had not been aware of the previous research. Were the prior publications (described in the immediately preceding paragraphs) referenced by Wilson at, or before, his first description of the invention? If so, it's unlikely that Wilson re-invented the Spectral Peak Picker strategy.

In that same submission in which Wilson first described "spectral peak picking," Wilson referenced the [ToDoBlCl83] Science article by saying:

> [Wilson1985/83] from 1983 appendix, page A-3-6: "When the frequency and intensity of electrical stimuli are held constant, but the site at which stimuli are delivered is varied, distinct tonal sensations are evoked that can be ranked according to the electrode's position along the cochlear partition. ( ... Tong et al., 1982 and 1983)."

This text only describes percepts elicited when one electrode channel is stimulated. Wilson did not reference or describe the Melbourne group's proposal for a multi-electrode, multi-formant-picking strategy, nor their significant psychophysical research that lead directly to their proposal.

In 1978, Wilson visited the UCSF cochlear implant laboratory for 1 or 2 days.[5] After I handed Wilson my PhD thesis, Wilson stated: "You don't know how much this means to me." Wilson's only referral, ever, to any part of the thesis [Wh78] was:

> [Wilson84a] page 30: "The RTI Patient Interface is essentially a redesign of the existing UCSF Interface described in Mark White's thesis, Chapter 5. A redesign was initiated to take advantage of newer technology and to provide expanded system flexibility." [The identical text was repeated in Appendix-1-1 of Wilson's 1985 contract proposal.]

Wilson has never referenced or credited any other content in [Wh78] in any of his later writings.

## 2. Virtual Channels and Current Steering strategies

UCSF suggested using channel interactions to "shift the stimulus focus" [MeWh77] and then was extended in [Wh78]: A large number of logical channels can be defined, each

---

[5] See [Wilson1985/83] bottom of the 187th sequential page. Use your PDF reader's "navigation mode" to access this unnumbered page.

logical channel being defined by the proportion of stimulating current that is to pass through each one of the 16 electrode contacts. Excerpts:

> [MeWh77] p. 336: Inter-electrode interactions occur when nearby channels are simultaneously stimulated. They can be circumvented by appropriate sequencing of stimuli. Interactions could conceivably be taken advantage of to shift the stimulus focus between adjacent bipolar electrode pairs.
>
> [Wh78] p. 203: Each logical channel is defined by specifying the portion of stimulating current that is to pass through each one of the 16 electrode contacts. For most logical channel specifications, only two to four nearby electrode contacts would be driven by non-zero proportions. For example, logical channel #30 might be defined by giving electrode contact #8 a +1.0 proportion, electrode contact #7 a -0.5 proportion, and contact #9 a -0.5
>
> [Wh78] p. 37: Individual excitation patterns are altered when two or more electrode pairs are stimulated simultaneously... Potentially, field interactions might move centers of excitation.
>
> [Wh78] p. 60: However, field interactions may be useful in increasing the repertoire of excitation patterns that may be elicited with electrical stimulation. With such "multipolar electrode channels" (i.e., more than two electrode contacts comprise an electrode channel), appropriate electrode contact spacing and array geometry may play a key role in the effective use of such interactions.
>
> [Wh78] p. 64 (bottom): An electrode channel can be composed of two or more electrode contacts or poles. Controlling currents at these multiple poles is used to change the spatial and temporal aspects of the excitation patterns.
>
> [Wh78] p. 99 (bottom): Changes in spatial excitation patterns... are elicited by changing one or more of the following stimulus parameters... relative amplitudes of the currents generated at the channel's multiple electrode poles. This allows one to manipulate electrical field interactions, the positions of current sources and sinks, and the temporal summation of eighth-nerve membrane charge.

As before, basic science played an important role here: It's unlikely that “current steering” strategies would have been proposed before the discovery of electric-field interactions occurring between cochlear electrodes that are simultaneously stimulated. The Utah group [Ed76], [EdDoBrMlPa78] first measured and described such “simultaneous electric field interactions.”

In 1984-85, in a study possibly related to current-steering, UCSF demonstrated the ability to “move” categorical boundaries in a speech recognition task by changing the relative stimulus intensities of 2 nearby electrode channels [WhOcMeSc90].

In the mid-1980s the Stanford cochlear implant research group extended the early proposals from the 1970's: They optimized linear transformation matrices — in order to compute electrode-array currents necessary to elicit any “specified cochlear excitation pattern” [Va85a], [Va85b]. They subsequently refined and customized the system for each patient — and performed behavioral threshold and pitch-ranking measurements that were encouraging [ToWh87]. Since then 2 cochlear implant companies, as well as university research groups, have continued to develop, refine, customize, and test such systems. Excerpts:

> [VA85b] p. 900: If the current spreading function is known, however, so is its inverse; and it is possible, using "current deconvolution," to compute the current pattern required at the electrodes to produce the desired pattern at the neurons. This paper will deal with the potential and the limitations of current deconvolution techniques as applied to scala tympani cochlear prostheses.
>
> [ToWh87] p. 891: These sharpened stimuli exhibited lower interaction and were pitch ranked with greater consistency than either monopolar or bipolar stimuli.

Wilson's earliest description of “virtual channels” was in 1992:

> [Wilson92] page 4: Design of VCIS Processors: A possible refinement and extension of the CIS approach is illustrated in Fig. 3 . Here adjacent electrodes may be stimulated simultaneously to shift the perceived pitch in any direction with respect to the corresponding single-electrode percepts.

*Details: Analysis of references and attribution of invention.* If the Virtual Channels processing strategy was not invented by Wilson, had he simply re-invented it? We determined that there were prior publications and communications from research groups that described the Virtual Channels strategy prior to Wilson's first description. If these same publications and communications were referenced by Wilson prior to his “invention” of this processing strategy, it's unlikely that Wilson re-invented the Virtual Channels strategy.

Considerably before Wilson described a “current steering” or “virtual channel” strategy, he had referenced earlier publications ([MeWh77], [Wh78]) that had described the strategy. Excerpts from Wilson's referring documents:

> “The major assumption of the spiral-plane model is that the characteristics of tissue in the plane of computation are homogeneous. An obvious method for evaluating this and other assumptions in the model is to compare model predictions with the results of animal experiments in which direct measurements of stimulus-response fields can be made. One such set of experiments was performed at UCSF in the late 70's with an array of aligned bipolar electrodes placed in the scala tympani of adult cats (see, e.g., Merzenich and White, 1977).” Text from page 13 & 16 of [Wilson84b] that refers to [MeWh77].
>
> “The RTI Patient Interface is essentially a redesign of the existing UCSF Interface described in Mark White's thesis, Chapter 5. A redesign was initiated to take advantage of newer technology and to provide expanded system flexibility.” Text from page 30 of [Wilson84a] that refers to [Wh78].

In neither progress report did Wilson credit the “current-steering strategy” described in [MeWh77] and in [Wh78]. – nor has he in any subsequent publications.

In the same conference and session in which Wilson and a colleague made presentations, van Compernolle of Stanford presented: “Current Spreading and Current Deconvolution in Scala Tympani Prostheses.” The entire abstract of van Compernolle's presentation can be found in the “[VA85b] p. 900” paragraph” above, and on page 64 of [Wilson85c]. In his 8th quarterly progress report, Wilson copied a page of the 1985 IEEE EMBS Conference program to document his presentations:

> [Wilson85c], page 64: The excerpt below has been shortened by omitting the abstracts. Only the presentation titles, presenters, and session title are listed:
>
> - *Speech Processors for Auditory Prostheses*, Blake S Wilson and Charles C Finley
> - *Speech Processing for Cochlear Implants*, ELV Wallenberg, IJ Hochmair-Desoyer, and ES Hochmair
>
> Session C13 – Cochlear Prostheses:
> - *Implications of Speech Encoding in the Normal Cochlea for a Cochlear Prosthesis* , Ben M. Clopton
> - *Current Spreading and Current Deconvolution in Scala Tympani Prostheses, Dirk van Compernolle*
> - *A Simple Finite-Difference Model of ... Bipolar Electrodes of the UCSF Array, CC Finley and BS Wilson*

## 3. CIS strategy

The early history of the cochlear implant has been summarized elsewhere (e.g., [MuMi2013]). This section describes some of the subsequent research that lead to the modern cochlear implant, the CIS, and related processing strategies.

This section is relatively brief! A more detailed description of the science, invention, and development of the CIS processing strategy can be found at: http://cochlearimplant.us

***The <u>core</u> processing models: (1) 8th-Nerve-Mimicking, (2) Vocoders, and (3) "Information" Models. Subsequent, improved models were "built-on-top" of one or more of these core models.***

- The "place-coding theory" of cochlear processing, attributed to Helmholtz and Bekesy, has been fundamental to the success of the cochlear implant. Using processing strategies that communicate only temporal information to the cochlear nerve (i.e., the use of the "volley-coding" alone) have performed poorly. For almost all patients, performance in speech tests was poor. By the early 1980's, most research groups felt it necessary to include "place-coding/spectral" information within their processing system: They used a set of appropriately-tuned bandpass filters to differentially drive multiple, appropriately-placed, cochlear electrodes. If the electronic technology had existed at the time, it is very likely that Helmholtz and Bekesy would have proposed this most-basic, and important, strategy.

- F Blair Simmons and colleagues' very early and sophisticated psychophysical experiments and their analyses were truly ground-breaking. Additionally, F Blair Simmons and colleagues in 1965 [Si.etal65], [Me2015, p. 41] were interested in ultimately implementing some form of mimicking or vocoder-based cochlear implant.

> [Si.etal65] p. 106: An attempt to separate the speech spectrum into frequency bands and to process each band so as to produce pulse stimuli which would make more efficient use of the characteristic "pitches" of the various electrodes also failed to produce any discrimination among speech-derived signals.

- The MIT and UCSF groups [KiMo72], [MeWh77], [Wh78] proposed the same 2 core models for cochlear implants. One cochlear implant model would attempt to directly "mimic" 8th nerve responses to acoustic stimulation. The other model was similar, except that a channel-vocoder would be "inserted" in between the acoustic source and the cochlea. Vocoder research provided evidence that a relatively small number of channels/electrodes might be sufficient to communicate speech in low noise environments — each channel transmitting low-frequency envelope and voicing information.

> [KiMo72] p. 723: In principle, a prosthetic device could mimic the action of the cochlea so effectively that the activity in the auditory nerve would be indistinguishable from the activity normally present in the nerve when the cochlea is intact.
>
> [KiMo72] p. 726: It is tempting to speculate that by analogy with the vocoder, a small number of frequency selective stimulus channels properly attached to the nerve might encode speech signals sufficiently well for adequate intelligibility.
>
> [MeWh77] p. 336: One models the excitation of the auditory nerve array by normal sound. That is, the multichannel array effects spatially and temporally patterned input that is the closest possible facsimile to that generated by normal sound in a normal cochlea. The second model is based on a voice excited channel vocoder. In the third model perceptually important information is delivered so that the processing employs the best information-bearing modes of stimulation.
>
> [Wh78] p. 104: Mimicking approaches: The "mimicking approach" is a general title for those processing techniques which are designed to reproduce certain feature(s) of normal hearing function. For example, one might attempt to replicate the eighth-nerve excitation patterns. One could attempt to replicate those neural patterns generated by the normal acoustic speech waveform; or one could attempt to replicate those excitation patterns elicited by acoustic stimuli which are generated by an analysis-synthesis vocoder....

- A third, complementary, model ("The information model") was described by UCSF in the belief that the patient's CNS could at the very least partially "fill-in" for the researchers' ample ignorance!

> [Wh78] p. 72: Initially, because of the humans' extremely adaptive perceptive capacities, as compared to those of current, automated speech recognition systems, stimulus processing algorithms probably will not be designed to substitute for the higher levels of the speech analysis "chain" such as phonetic analysis. Initially, the stimulus processing algorithms probably will be designed to effectively communicate parameters extracted from the speech waveform. In some cases, these parameters will be chosen to directly approximate or mimic the function of the outer, middle, and inner ear functions. In other cases, other choices of parameters and stimulation methods will be used because it is believed that the subject may readily learn to use such stimuli (House, Stevens, Sandel and Arnold, 1962).
>
> [Wh78] p. 96: Furthermore, it appears that the human is able to adapt to many alterations in the speech waveform (Licklider, 1946) and speech spectrum. For example, speech perception can be relearned even if the speech spectrum has been inverted (Blesser, 1969). Speech recognition is easily accomplished over a wide range of speakers; even though the formant frequencies vary widely across speakers for the same word (Ladefoged and Broadbent, 1957; Broad and Shoup, 1975). People adapt to speakers with strong accents; and, with further practice, people can learn foreign languages.
>
> [Me.etal79] p. 202: In our own research, basic psychophysical experiments are being conducted that are designed to reveal how aspects of speech signals critical to intelligibility

> might be least equivocally coded across implanted multielectrode arrays. In this basic, non-simulation approach, these elemental components of speech sounds are then mapped across the multielectrode (and hence the auditory nerve) array (see White, 1978).

### *The science directly responsible for the low-rate and the high-rate interleaved-pulses (also known as the CIS) processing strategies.*

A. With single-channel stimulation, there was "at least some" evidence of the transmission of Fine-Grain Temporal Information to the CNS in animals and patients. [Si.etal65], [Si69], [Cl.etal72], [KiMo72], and [Me.etal73], using basic psychophysical and neurophysiology experiments, found that some relatively fine-grain temporal information could be communicated to the CNS using a single electrode. However, not nearly enough information could be transmitted to the CNS for speech understanding. Thus, it was strongly argued by almost all research groups that communicating 'place/spectral' information would necessarily be an essential part of an effective cochlear implant system.

In a speech perception task, UCSF found direct evidence for the significant value of communicating fine-grain temporal information in a cochlear implant: A patient using only a single electrode-channel could utilize relatively fine-grain temporal information to "recognize" (albeit imperfectly) synthetic vowels differing only in their 1st formant frequency [Wh83a]. In contrast, the same patient performed at chance-levels when synthetic vowels differed only in their 2nd formant frequency.

B. The Utah group [Ed76], [EdDoBrMlPa78] first measured and described electric field inter-electrode interactions in cochlear implants. The interactions occurred during simultaneous stimulation of 2 electrode channels. They proposed temporally-interleaving stimulus pulses across channels to eliminate such interactions. It was clear from Utah's human and UCSF's subsequent animal measurements that interleaving pulses between nearby electrode channels reduced channel interactions [Ed76], [MeWh77], [EdDoBrMlPa78], [Wh78].

> [EdDoBrMlPa78] p. 23: The fact that no changes in threshold were observed when the simultaneous pulses of two electrodes were interlaced means that a relatively high degree of flexibility exists in minimizing interactions between input channels at the site of cochlear stimulation.
>
> [Wh78] p. 44: This and much similar data reveals that sub-threshold biphasic stimuli presented as little as 50 usec. apart in one channel cannot alter the threshold of excitation of an adjacent channel. This data is very compatible with Hill's model for neural excitation (Figure 3.7).
>
> [Wh78] p. 61: If electrode pairs are simultaneously stimulated, excitation patterns cannot be predicted merely by spatially summing each channel's independent response.

C. Simple variations of the "core" vocoder cochlear implant design were suggested by UCSF:

> [Wh78] p. 81-82: The spectrum channel vocoder and the formant vocoder analysis stage suggest stimulus processing models for a cochlear prosthesis. The spectrum channel vocoder

> stimulus processing might involve stimulating n electrode channels; each electrode channel might represent one vocoder analysis channel. Each analysis channel's output magnitude might be conveyed by changing stimulus amplitude and/or by changing stimulus timing (e.g. frequency) and/or by changing stimulating electrodes and/or by utilizing electric field interactions between three of more electrode contacts.

D. Channel vocoders and formant vocoders (to an even greater extent) can cause the loss of useful information [Fl72] — by reducing the transmission of “seemingly redundant information.” For example, speech information that is “effectively redundant” in low noise listening environments can become “anything but redundant” in high-noise listening environments.

> [Wh78] p. 75: Vocoder analysis stages generally reduce the information rate required to convey speech. Redundant and irrelevant information is discarded in the analysis stage.
>
> [Wh78] p. 81: As the channel bandwidth is reduced, more useful information is lost. Also, if the speech signal is in a noisy environment, the formant vocoder's performance is degraded more than the channel vocoder's performance in a similar environment.

E. In an effort to communicate such potentially useful information, UCSF developed an augmentation of the “core” channel-vocoder model for cochlear implants — by adding finer-grain temporal information in the stimulation of each electrode-channel of a cochlear implant processor [Wh78, pp. 85-88, 57, 93-94], [Wh83c]. UCSF proposed to communicate finer-grain temporal information, in addition to the “slower” temporal envelope information, by merging the core channel-vocoder model with a simple “8th-nerve-mimicking model for stimulating fibers near each electrode.”

> [Wh78] p. 62: In one processing scheme, one would use a bank of bandpass filters to analyze speech. The output of each filter would be further processed and then used to excite a small population of acoustic nerve fibers. One might attempt to electrically excite each small population of neurons such that the composite post stimulus time histogram of this small population of neurons is similar in shape to the envelope of a bandpass filtered version of the speech signal (Kiang and Moxon, 1972)...
>
> One might attempt to excite the nerve with more than low frequency envelope information (0-25 Hz.). Higher frequency envelope information (50-300 Hz.) might help to convey voicing information. Still higher frequency "phase-locking" information might be used by the nervous system (Evans, 1977; Moller, 1977). Electrical stimuli, phase-locked to the spectral peak(s) within the bandpass filter's passband, might be useful to the nervous system.
>
> One variation of this stimulation method would involve the temporal interlacing (within each electrode channel) of a range of biphasic pulse amplitudes. The higher amplitude pulses would occur relatively infrequently compared to the total number of biphasic pulses generated. Such a stimulus sequence could generate relatively low discharge rates on the perimeter of each electrode channel's excitatory domain. At distances closer to the particular electrode channel, discharge rates would increase. Such a pattern would more closely mimick the excitation pattern generated by a low SPL acoustic tone. Channel interactions, proper stimulus amplitude control, and electrode sequencing might be used to approximate the asymetrical spatial-temporal patterns observed, particularly when mimicking higher stimulus levels (see Figure 1 of Pfeiffer, 1975; and Figure 7 of Evans, 1975b).

F. To communicate this finer-grain temporal information would require higher pulse rates. When interleaving pulses across electrode-channels, would there be inter-channel "temporal" interactions when using such short inter-pulse intervals? Because UCSF believed strongly in the potential of communicating finer-grain temporal information to each electrode channel, such "temporal" inter-channel interactions were measured at very short inter-pulse intervals between 2 channels in neurophysiology experiments in cat [MeWh77], [Wh78]:

> [MeWh77] p. 334: These and much similar data reveal that stimuli presented as little as 75 usec apart in one channel cannot alter the threshold of excitation of an adjacent channel.
>
> [Wh78] p. 44: This and much similar data reveals that subthreshold biphasic stimuli presented as little as 50 usec. apart in one channel cannot alter the threshold of excitation of an adjacent channel. This data is very compatible with Hill's model for neural excitation (Figure 3.7).

G. UCSF extended these studies to humans, using both psychophysical and electrophysiology measures. As before, the goal was to determine how much temporally-interleaved pulses on nearby channels would interact [WhMeGa84]. Interactions were measured as a function of inter-pulse interval, inter-electrode distances, implant subject, electrode-type, and relative electrode location (base->apex). In some subjects, there were no measurable interactions at very short inter-pulse intervals under any condition. For other subjects, there were significant interactions, particularly between nearby monopolar electrode channels at short inter-pulse intervals. However, even in those cases, "temporal interactions" were much smaller than those for simultaneous stimulation. For more information, go to this link.

> [Me2015] p. 43: In studies largely led by Mark White, we had directly documented interference [due to non-simultaneous stimuli] parametrically in both animal and human models ... and well understood the conditions for minimizing or avoiding it.
>
> [WhMeGa84] p. 501: Nonsimultaneous stimulation of electrode channels should eliminate the most severe component of channel interaction.

H. With high pulse-rate stimulation, many researchers had assumed that (1) neurons would discharge at rates above their normal ("acoustically-driven") maximum-discharge-rate and therefore might be damaged, and/or (2) the neurons' responses might become so "distorted" that the CNS would be unable to decode the responses. Most researchers had not envisioned the possibility that many or all 8th-nerve fibers might be driven within their normal-discharge-rate range, even for the higher loudnesses.

- The stochastic hypothesis: One of the primary concerns of early researchers, and still to this day, was the belief that electrical stimuli would cause responding neurons to discharge in "lock-step." In other words, it was presumed that 8th nerve fibers would respond to electrical stimulation in a deterministic manner. This is in stark contrast to what happens in acoustic stimulation, where neurons do not respond in lock-step, but respond in a highly-stochastic manner [Wh78 p. 61].

- Stochastic input-output behavior to electric stimuli was observed by Kiang and Moxon [KiMo72] and by Merzenich et. al. [Me.etal73] in auditory nerve fibers and in inferior colliculus neurons, respectively. As a consequence of their work, White incorporated a stochastic component into an extended Hill neuron model [Wh78, p 54-55].

- ○ UCSF believed in the value of understanding [Wh78, pp. 47, 54-55], [Wh84a, p.100], [Wh84b, pp.397-499], [Wh87b] and potentially enlisting this seemingly "limited stochastic response" to electrical stimulation — for the purpose of potentially improving the communications of information to the CNS. [For more information about "Stochastic Neural Response," see Appendix 1.]

***Implementations of the low-rate IP and the high-rate IP processing strategies:***

- Univ. of Melbourne: In 1982 the Melbourne group described [ToDoBlCl83] interleaving pulses between 2 channels to convey both F1 and F2 information in order to avoid "current summation in the cochlea." As in their earlier 1-formant "F2" design, during voiced speech, the channel pulse-rate was approximately equal to F0. -- on each of the 2 stimulated channels . By 1984, Melbourne researchers had developed an acoustic model of this processing strategy. [Blamey.etal85] used the acoustic model with normal-hearing subjects to test the new 2-formant strategy and compared its performance with the earlier 1-formant (F2) strategy. The results were encouraging and by 1986 Melbourne researchers [Blamey.etal87] had implemented and tested it in cochlear implant patients with very positive results.

  > [ToDoBlCl83] p. 994: The results indicate that, in addition to the presentation of the second formant by activation of only one [of the 10] electrode[s] during a stimulus period in a speech processor, as described by Tong et al (5), speech information such as the first formant may be presented to implant patients by two-electrode stimulation.
  >
  > [ToDoBlCl83] p. 993: The pulse trains delivered to the electrode pair were at the same repetition rate, with the pulses on the more apical electrode leading by 0.5 msec. There was therefore, no temporal overlap between pulses on the two electrodes.

- Paris group: In contrast to their 1970's strategy, their 1980's [ChFuMeLa83] stimulation protocol was an interleaved-pulses (IP) strategy, where pulses were delivered sequentially across all electrode channels. Each temporally-interleaved channel was driven at 300 pps. They state: "Only one electrode is stimulated at a time, and during this stimulation all the other electrodes are grounded." They state that the repetition rate and sequencing was for the purpose of simulating the "cochlear traveling wave." They do not indicate that the non-simultaneous stimulation was for the purpose of reducing channel interactions.

  The Paris group modulated pulse-width to control the level of neural excitation at each channel. However, it is unclear whether the electrode channels were electrically isolated.[6] It is also unclear whether the electrodes were driven with controlled-current sources.[7]

- UCSF: In 1977, UCSF received the first-ever NIH grant to design, implement, and test multichannel cochlear implants in humans. Immediately after Lindsay Vurek had completed (1) the take-home portable and (2) the bench-top real-time compressed analog multichannel processor in 1980-1982, he started implementing the bench-top real-time high-rate interleaved pulses (IP) processor, sometime in 1980-1982.

  The initial system had 4 channels (but was designed for expansion to 8 channels) and was to be able to stimulate at any rate up to 1000 pps per channel. Each channel's envelope-detector included a low-pass filter that was to be adjustable from 25 Hz up to 1000 Hz — with an option to bypass the filter. The basic science (discussed earlier) indicated that patients could use some of the finer-grain temporal information in speech. Thus, UCSF believed that at least some F0/voicing and F1 information could be accessed and utilized by patients in recognizing speech ***if*** channel pulse-rates and cut-offs of the envelope detectors' low-pass filters enabled the transmission of this information. UCSF did not include an F0/voicing extractor in the design of the IP processor. Nor did UCSF include a channel-peak-picker in the design. Of course, if initial tests of this system indicated our assumptions were incorrect, these processing elements would need to be added.

  > Proceedings of the 1983 SF Cochlear Implant conference, Merzenich [Me85] p. 127: Further Device Development: ... Two computer-based systems are being constructed that would allow for evaluation of almost any imaginable coding scheme on the same implant patient, given a temporary trans-cutaneous [percutaneous] link via our electrode cable. One is a real-time microprocessor-based model of our next generation of cochlear implant sound processors...

  The real-time processor was completed in April-May 1985, and met all specifications except that the maximum rate per channel was only slightly higher than 600 pps per channel. – with a total pulse-rate of 2400 pps across all 4-channels. Because our grant was not continued, this processor was never tested on patients. The processor's design specification, including a listing of the program for interleaving of pulses across channels, is available here.

---

[6] If the electrodes in the array were not electrically isolated (e.g., by using controlled-current sources), an given electrode's stimulating current could end-up being "delivered to unintended ground electrodes:" The un-isolated electrodes, that have the lowest contact impedances, would likely receive a significant fraction of the stimulus current.

[7] If current sources were not used, an electrode's driving-source impedance becomes particularly important. (Note: The electrode's-contact-impedance is often the largest part of this driving-source impedance.) The driving-source impedances determine the time-constants of the currents exiting the electrodes into the cochlear tissue. If these time-constants are not short enough, the current-waveforms across electrode channels may temporally overlap. Such temporal-overlap could cause undesirable electric-field interactions within the cochlea.

> [Me2015] p. 43: …Those and other observations led to our development, in 1984-86, of an initial model of an interleaved pulse-processor (IPP) CI model, designed to drive an 8-channel device, operating at rates up to about 600 Hz.
>
> [Me2015] p. 44: ...the UCSF research engineering team was focused on implementing a next-generation CI model, an 8-channel digital CI using an interleaved pulse processing coding strategy. In 1984, we described this device development strategy in detail in an NIH program project grant. Unfortunately, this grant received critical technical reviews that specifically questioned the wisdom of IPP coding! ... our own IPP speech processing model was never applied in a human subject.

- UCSF/Wilson collaboration: Wilson was contracted by NIH in late 1983 to implement a highly flexible, non-real-time simulator and provide it for use by UCSF researchers for simulating and testing implant processor designs from UCSF and from Wilson.

> Merzenich [Me85] p. 127 of the Proceedings of the 1983 SF Cochlear Implant conference: Further Device Development: ... Two computer-based systems are being constructed that would allow for evaluation of almost any imaginable coding scheme on the same implant patient, given a temporary trans-cutaneous [percutaneous] link via our electrode cable. ... The second is a powerful non-real-time model being developed by a collaborative team at the Research Triangle Institute in North Carolina.
>
> [Wilson83] p. 4: Everyone involved in the discussions on patient testing agreed that RTI 's original plan, of just building the computer simulator and providing it for use by UCSF personnel, is flawed for the reasons mentioned above.
>
> [Wilson83] p. 9: Pending approval of the project officer for this contract...
>
> [Wilson83] p. 4: RTI personnel should be involved in patient testing to an extent much greater than that originally indicated in RTI's proposal for this project. ... both RTI and UCSF should prepare lists of speech processors and psychophysical experiments well before the end of February, 1984, when the next patient is tentatively scheduled for surgery at UCSF. The groups should then meet a few weeks prior to surgery to decide on the final set of tests to be conducted and on the personnel requirements for each test.

This was to be a collaboration between Wilson, (BSEE 1974, Duke; PhD 2016, Duke), who had just begun work in cochlear implants [Wilson85d]. UCSF had 2 academic and 2 clinical faculty members, each having more than 10 years of experience in cochlear implant research. UCSF also had 7 professional staff members, each with experience in cochlear implant research. At the beginning of Wilson's first cochlear implant contract, during a week in November 1983, a total of 11 members of the UCSF team briefed Wilson on UCSF's designs, implementations, and protocols (see [Wilson83] p. 3). UCSF's research and engineering was transferred to Wilson from 1978 to 1989. [8] For example, Wilson's [Wilson84a] proposed computer hardware interface and block-diagram signal processing software were partially derived from those developed in chapter 4 and 5 of White's UC-Berkeley thesis [Wh78]. Wilson augmented the software design with additional components such as UCSF's compression classes [Wh83b] and UCSF's channel peak picker:

---

[8] Much of this transferred UCSF research is documented throughout this manuscript.

> [Wilson84a] page 30: "The RTI Patient Interface is essentially a redesign of the existing UCSF Interface described in Mark White's thesis, Chapter 5. A redesign was initiated to take advantage of newer technology and to provide expanded system flexibility."
>
> [Wilson84a] page 51: "FUNCTIONS OF BLOCKS FOR BLOCK-DIAGRAM COMPILER... compressors (specify type and parameters look at Mark White's paper for classes)" [See [Wh83b] to view White's compressor paper.]
>
> [Wilson84a] page 51: "FUNCTIONS OF BLOCKS FOR BLOCK-DIAGRAM COMPILER... peak picker (specify analysis interval)"

In 1983 UCSF asked Wilson to implement the same processing in his planned non-real-time simulator as had been specified for UCSF's real-time high-rate IP processor. In addition, UCSF requested that Wilson implement the compressed analog processor. – for the purpose of comparing it with UCSF's existing compressed analog system which had already been tested-on patients. Excerpts:

> [Wilson83] p. 3: I spent the week of November 21 [1983] in San Francisco... Meetings were held to discuss... engineering considerations in the design and fabrication of … real-time speech processors (primarily with Vurek …)
>
> [Wilson84a] p. 28: ... we plan to begin testing of implanted patients at UCSF, which is a new objective recently approved by the technical monitor and contracting officer for this project. Now that the tools for design and evaluation of speech processors for multichannel auditory prostheses are nearly complete, we are anxious to put them to the best-possible use in collaboration with our talented colleagues at UCSF.

In March, 1985, three months before he was to deliver the non-real-time simulator to UCSF, Wilson made a list of the speech processing algorithms he planned to test:

> [Wilson85] p. 36: Major Goals of San Francisco Trips, June, 1985... Evaluate multichannel coding strategies : (a) present, 4-channel UCSF processor; (b) 8-channel UCSF processor; (c) 4-channel processor that presents a "multipulse" excitation signal (from the linear-prediction residual) to the configured electrode channels according to the frequency and bandwidth of F2; (d) 8-channel version of the above.

UCSF was positive about Wilson's new "multipulse excitation" processing strategy, but reminded RTI of the importance of implementing and testing the high-rate IP processor — as had been specified to Wilson at the beginning of his contract.

By July, 1985 the non-real-time simulator had been implemented and delivered by Wilson to UCSF for testing. The high-rate IP strategy that Wilson had been asked to implement did not meet UCSF's specifications: The simulator did not deliver high enough pulse-rates on each channel. Only a low-rate version[9] of the IP processor was implemented and tested at UCSF [Wilson85b]. Thus, little or no fine-grain temporal information was conveyed to the patient. Fortunately, even with these deficits, preliminary tests indicated that the low-rate IP processor appeared to perform

---

[9] This processor included the channel-peak-picker "add-on." The pulse rate was 312 pps per channel for the 2 channels that had the highest energy during the 10 msec analysis time-frame. There were a total of 6 channels from which the 2 channels were "picked from" during each time-frame.

somewhat better than: (1) the "multipulse excitation" strategy and (2) the compressed-analog strategy.

After UCSF's NIH grant ended in 1985, White moved to Raleigh to work at NC State University in 1986 as an Associate Professor. White continued to collaborate with Wilson until 1990 and was a sub-contractor for Wilson from 1988-1990. The processing strategy originally specified for UCSF's real-time high-pulse-rate IP processor had still not been tested! From 1986, it took another 2 years for White to convince Wilson to implement the high-rate version of the interleaved-pulses processor [Wilson87a] pp. 34-39. In 1989-90, the high-rate IP processor was finally tested on patients [Wilson89b].[10]

> [Wilson88] Contract Proposal, pp. 139-140: Since his move to North Carolina two years ago, Dr. White has collaborated with us in (a) the development of new signal processing strategies for cochlear implants; (b) the investigation of possible mechanisms of intracochlear stimulation using models of the electric field patterns produced by intracochlear electrodes and of the resulting neural responses to the imposed electric fields; and (c) the preparation of several manuscripts for publication. His insight and experience have been invaluable. .... Because Dr. White's contribution to the project will be multifaceted and large, we have arranged a subcontract with N.C. State University for 20 percent of his time during the academic year.

More about the history of the science, invention, and development of the CIS processor is available at http://cochlearimplant.us

*Details: Analysis of references and attribution of invention.* If the CIS processing strategy was not invented by Wilson, had he simply re-invented it? To answer this question, we will examine the NIH-funded collaboration between UCSF and Wilson. – and what UCSF had communicated to Wilson. We will also determine whether, and in what manner, Wilson referred-to or formally referenced other researchers publications. – publications that had described the IP and high-rate IP processing strategies prior to Wilson's first description of the high-rate IP processing strategy.

The NIH-funded collaboration between UCSF and Wilson:

As documented above in the “CIS” section, the primary purpose for Wilson's first NIH contract was to implement a general-purpose, mini-computer-based, non-real-time cochlear implant simulator. – and deliver it to UCSF. The purpose of the simulator was to significantly “simplify” the implementation of different processing strategies. UCSF and Wilson would both propose processing strategies that would be tested at UCSF, using the simulator.

---

[10] The high-rate IP processor was renamed 2 times: Wilson's name for the high-rate IP processing strategy (i.e., that did not use an F0 extractor for pulse-timing) was the “max-rate IP processor” or the “maximum-rate IP processor” [e.g., Wilson89a, p.23 & p. 33]. However, within 3 months after using channel pulse-rates significantly above 300 pps, he re-branded the same processor as the “super-sampler” and no longer referred to it as an “IP processor” [Wilson89b, p. 4]. Within another 3 months, he again renamed it, this time as the “CIS” processor and still did not refer to it as an “IP processor” [Wilson90, p. 4].

The real-time, bench-top, high-rate IP processor that Lindsay Vurek was implementing was first described to Wilson during Wilson's week-long visit to UCSF in November 1983. – at the beginning of Wilson's NIH contract. Wilson was to implement the high-rate IP processor in his non-real-time system that he was building. This was the priority. In addition, UCSF requested that Wilson implement UCSF's compressed analog processor strategy. – for comparison with UCSF's real-time implementation of that strategy. UCSF's real-time implementation of the compressed analog strategy had already been constructed and successfully tested with patients at UCSF.

Wilson and colleagues visited UCSF at least 6 times from 1983 to 1985. The high-rate IP processing strategy and its rationale were described to Wilson during those visits. – and thereafter by White (after having moved to Raleigh, North Carolina). -- until Wilson finally decided to implement and test the high-rate IP processor in the late 80's.

Amazingly, in all of Wilson's progress reports and publications, he only attributes one processing strategy to UCSF: the compressed analog processing strategy. – which had already been implemented and tested with patients at UCSF and Utah/MIT before Wilson even started his first NIH contract. Neither UCSF's proposed high-rate IP processing strategy, nor Vurek's real-time bench-top implementation of it, were ever credited to UCSF by Wilson!

***Prior publications, references, and attribution of credit***

In Question 1, we determined that there were prior publications from research groups that described the IP and one that described the high-rate IP processing strategy. Were these same publications referenced by Wilson prior to his "invention" of these processing strategies? If so, it's unlikely that Wilson re-invented either the IP or the high-rate IP processing strategy.

Before we examine Wilson's pattern of referencing of the "IP publications," it should be noted that other research groups have also made errors in referencing the IP processing strategy: For example, UCSF [MeWh77] failed to reference the personal communications from Eddington [Ed76] to Mark White. UCSF [WhMeGa84] referenced the Utah group [EdDoBrMlPa78] for their measurements of electric field interactions, but failed to reference [EdDoBrMlPa78] for their proposal to use non-simultaneous stimulation to reduce such channel interactions.

Before Wilson first described the "interleaved pulses" (IP) processing strategy in his 1985 7th quarterly progress report [Wilson85b], Wilson had "referenced" earlier publications ([MeWh77], [Wh78], [EdDoBrMlPa78], [ToDoBlCl83], [WhMeGa84]) that had already described the IP processing strategy. One of those publications, [Wh78], described the high-rate IP processing strategy as well as other processing strategies.

With one very interesting exception, Wilson's text descriptions referencing those prior publications never mentioned any of the IP processing proposals described in those publications. As before, Wilson referred to unrelated content in those prior publications. Excerpts from Wilson's referring documents:

> [Wilson84b] page 16: One such set of experiments was performed at UCSF in the late 70's with an array of aligned bipolar electrodes placed in the scala tympani of adult cats (see, e.g., Merzenich and White, 1977). ... Superimposed on the effective stimulus field is the curve of exponential falloff in the response fields measured for "well-positioned" electrodes in the cat (space constant = .87 mm.; see Merzenich and White for details).
>
> [Wilson84a] page 30: The RTI Patient Interface is essentially a redesign of the existing UCSF Interface described in Mark White's thesis, Chapter 5. A redesign was initiated to take advantage of newer technology and to provide expanded system flexibility.
>
> [Wilson1985/83] from 1983 appendix, page A-3-6: Also, results of scaling and matching experiments indicate that, while pitch corresponds to rate of stimulus frequencies up to about 200 Hz, it either accelerates to very high values (Eddington et al, 1978...
>
> [Wilson1985/83] from 1983 appendix, page A-3-9 (1983): In rereading Eddington et al.' s 1978 paper, though, we were reminded of their finding that pulses delivered to single electrodes in a scala-tympani array are perceived as higher in pitch when pulse duration is decreased.
>
> [Wilson1985b] p. 36: ... may explain how patients implanted with monopolar arrays can rank their electrodes (see, e.g., Eddington et al, 1978).
>
> [Wilson1985/83] from 1983 appendix, page A-3-6: "When the frequency and intensity of electrical stimuli are held constant, but the site at which stimuli are delivered is varied, distinct tonal sensations are evoked that can be ranked according to the electrode's position along the cochlear partition. ( ... Tong et al., 1982 and 1983)."
>
> [Wilson1985/83] unnumbered page located immediately before the page labeled "210": The page's title is "References mistakenly left out of the original proposals." On that page, one of the references listed is: White, M. W., Merzenich, M. M. and Gardi, J. N., Multichannel cochlear implants, Arch. Otolaryngol., 110 (1984) 493-501. Note: In an effort to obtain Wilson's text that referred to this UCSF publication, I filed a Freedom of Information Act (FOIA) request to NIH to obtain the "original proposals." NIH responded: "No records responsive to those items of your request were located."

There is one very interesting exception to Wilson's referencing practices: Wilson referenced [WhMeGa84] as describing the use of interleaved-pulses to greatly reduce electric field interactions. -- But only for the *obscure 2-channel Brewer-Plomp processor*!

> [Wilson87b] p. 9: Pulsatile stimuli always are presented nonsimultaneously to the two channels, greatly reducing electric field interactions that might compromise perceived distinctions between channels (Wilson et al., 1987a; White et al., 1984). Variations of the Breeuwer/Plomp processors are produced with different choices of parameters for the post processor.

This is the only reference by Wilson crediting any research group for proposing the interleaving of pulses across channels to "greatly reduce electric field interactions."

Section Summary: Two consistent and complementary lines of evidence help us to answer question 2: (1) Wilson referenced 3rd party publications that had already described the IP and the high-rate IP processing strategies. (2) The high-rate IP, and the supporting research for it, were repeatedly described in detail to Wilson by UCSF as part of their

NIH-funded collaboration. -- over a 4-5 year period, starting in 1983. This clearly indicates that Wilson did not “re-invent” either the low-rate IP or the high-rate IP processing strategies.

For a short analysis of Wilson's most recent publications as they relate to the invention of the high-rate IP processing strategy, see Appendix 2.

## CONCLUSION

Relatively basic research (i.e., developing deeper understanding) was and is necessary for significant improvements in cochlear implants. By examining the historical importance of such research, it should become clear that future progress will likely require “much more of the same!”

Many researchers from around the world have played key roles in the invention and development of the cochlear implant. This manuscript has credited only a subset of these researchers and their contributions.

Secondarily, the evidence presented here indicates that the 3 cochlear implant processing strategies were not invented, nor reinvented, by any single researcher or group. It is an injustice to many researchers to state otherwise [Wilson2015a], [Wilson2016]. It is hoped that this manuscript has effectively “unmasked” some of the contributions made by a good number of scientists.

**REFERENCES**

The full text of each publication can be downloaded by clicking the corresponding web link.

[Blamey.etal85] Blamey PJ, Martin LFA, Clark GM, "A comparison of three speech coding strategies using an acoustic model of a cochlear implant," J Acoust Soc Am, 77(1):209-217, 1985.

[Blamey.etal87] Blamey PJ, Sowell RC, Brown AM, Clark GM, Seligman PM, "Vowel and consonant recognition of cochlear implant patients using formant-estimating speech processors," J Acoust Soc Am, 82(1):48-57, 1987.

[Bruce.etal99a] I.C. Bruce, White, M. W., Irlicht, L. S., O'Leary, S. J., Dynes, S., Javel, E., Clark, G. M. "A Stochastic Model of the Electrically Stimulated Auditory Nerve: Single-Pulse Response." IEEE Transactions on Biomedical Engineering, 46(6):617-629, 1999.

[Bruce.etal99b] I.C. Bruce, White, M. W., Irlicht, L. S., O'Leary, Clark, G. M. "The Effects of Stochastic Neural Activity in a Model Predicting Loudness Perception with Cochlear Implants." IEEE Transactions on Biomedical Engineering, 46(12):1393-1404, 1999.

[Cl.etal72] Clark, G. M., Nathar, J. M., Kranz, H. G., & Maritz, J. S. A behavioral study on electrical stimulation of the cochlea and central auditory pathways of the cat. Experimental neurology, 36(2), 350-361, 1972.

[ChFuMeLa83] Chouard CH, Fugain C, Meyer B, Lacombe H. Long-term results of the multichannel cochlear implant. Ann N Y Acad Sci. 1983; 405:387-411.

[DoDoTu62] Doyle, J.H., Doyle, J.B. Jr., Turnbull, F.M. Jr., Electrical stimulation of eighth cranial nerve. Arch. Otolaryngol. 80, 388-391, 1964.

[Ed76] Personal Communications From Don Eddington to Mark White, midwinter meeting of the ARO, 1976. For more information about this important contribution, click here.

[EdDoBrMlPa78] Eddington DK, Dobelle WH, Brackmann DE, Mladejovsky MG, Parkin JL, "Auditory prostheses research with multiple channel intracochlear stimulation in man," The Annals of Otology, Rhinology, and Laryngology 87(6 Pt 2):1-39, 1978.

[Fl72] Flanagan JL, Speech Analysis Synthesis and Perception, (second edition), Springer-Verlag, New York, 1972.

[Ki65] Kiang NYS, "Discharge Patterns of Single Fibers in the Cats Auditory Nerve," Research Monograph No. 35, MIT Press, Cambridge, MA, 1965.

[KiMo72] Kiang NYS, Moxon EC, "Physiological Considerations in Artificial Stimulation of the Inner Ear," Ann Otol 81:714-730, 1972.

[Me.etal73] Merzenich MM, Michelson RP, Pettit CR, Schindler RA, Reid M Neural encoding of sound sensation evoked by electrical stimulation of the acoustic nerve. Ann Otol Rhinol Laryngol 82:486-503, 1973.

[MeWh77] Merzenich, M.M., and White, M.W.: Cochlear Implant: The interface problem. In: Functional Electrical Stimulation. (Hambrecht, F.T., and Reswich, J.B. eds.), Marcel Dekker, Inc., Vol 3, pp. 321-340, 1977.

[Me.etal79] Michael M. Merzenich, Mark White, Michael C. Vivion, Patricia A. Leake-Jones & Sheila Walsh (1979) Some Considerations of Multichannel Electrical Stimulation of the Auditory Nerve in the Profoundly Deaf; Interfacing Electrode Arrays with the Auditory Nerve Array, Acta Oto-Laryngologica, 87:3-6, 196-203, 1979.

[Me85] 1983 Proceedings of the San Francisco Cochlear Implant conference: Merzenich, M.M., 1985. University of California San Francisco cochlear implant device. In: Cochlear Implants. Raven Press, New York, pp. 121-130.

[Me2015] Merzenich, M.M. Early UCSF contributions to the development of multiple-channel cochlear implants. Hearing Research 322 (2015) 39-46, 2015.

[Miller.etal2008] CA Miller, N Hu, F Zhang, BK Robinson, and PJ Abbas, "Changes Across Time in the Temporal Responses of Auditory Nerve Fibers Stimulated by Electric Pulse Trains." JARO 9: 122–137 (2008).

[MuMi2013] A Mudry, M Mills, "The Early History of the Cochlear Implant A Retrospective," JAMA Otolaryngol Head Neck Surg. 139(5): 446-453 (2013).

[PeCo57] Peterson, E., & Cooper, F. S. (1957). Peakpicker: A Band- Width Compression Device. The Journal of the Acoustical Society of America, 29(6), p. 777 M3.

[Si.etal65] Simmons, F. B., Epley, J. M., Lummis, R. C., Guttman, N., Frishkopf, L. S., Harmon, L. D., & Zwicker, E. (1965). Auditory nerve: electrical stimulation in man. Science, 148(3666), 104-106.

[Si69] Simmons, F. B. (1969). Cochlear implants. Archives of Otolaryngology, 89(1), pp. 61-69.

[ToDoBlCl83] Tong, YC, Dowell, RC, Blamey, PJ and Clark, G. M., Two-component hearing sensations produced by two-electrode stimulation in the cochlea of a deaf patient, Science, 219 (Feb 25, 1983) 993-994.

[ToWh87] Townshend B, White RL, Reduction of Electrical Interaction in Auditory Prostheses. IEEE Transactions on Biomedical Engineering, BME-34(11), pp. 891-897, November, 1987.

[Ty.etal97] M Tykocinski, RK Shepherd, GM Clark, “Reduction in excitability of the auditory nerve following electrical stimulation at high stimulus rates. II. Comparison of fixed amplitude with amplitude modulated stimuli.” Hearing Research, 112: 1–2, pp. 147-157, 1997.

[Va85a] Van Compernolle, Dirk. "A computational model of the cochlea used with cochlear prosthesis patients." Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '85. Vol. 10, pp 427-429, 1985.

[Va85b] Van Compernolle D, White RL, "Current Spreading and Current Deconvolution in Scala Tympani Prostheses" (Abstract) IEEE Trans on Biomedical Engineering. Vol. BME-32, No 10, p. 900, Annual Meeting, Chicago, October, 1985.

[Wh78] White, M.W., Design considerations of a prosthesis for the profoundly deaf, Ph.D. dissertation, Univ. California, Berkeley, 1978.

[Wh83a] White, M.W.: Formant frequency discrimination and recognition in subjects implanted with an Intracochlear stimulating electrode. Annals of the New York Academy of Science 405:348-359, 1983.

[Wh83b] White, M.W.: Compression Systems for Cochlear Prostheses. In: Mechanisms of Hearing. (W.R. Webster and L.M. Aitkin, eds.) Monash Press, Clayton, Victoria, Australia, pp 184-189, Aug. 23-27, 1983 (conference paper).

[Wh83c] White, M.W.: The cochlear implant: the coding of auditory information. Proceedings of the First Vienna International Workshop on Functional Electrostimulation, edited by Dr. H. Thoma, University of Vienna, Vienna, Austria, p. 240-243, 1983.

[WhMeGa84] White, M.W., Merzenich, M.M., Gardi, J.N.: Multichannel cochlear implants: channel interactions and processor design. Archives of Otolaryngology 110:493-501, 1984.

[Wh84a] White, M.W.: Psychophysical and Neurophysiological Considerations in the design of a cochlear prosthesis. Audiologia Italiana 1:77-117, 1984.

[Wh84b] White, M.W.: The Multichannel Cochlear Prosthesis: Channel Interactions. Proceedings of the Sixth Annual Conference of the IEEE Engineering in Medicine and Biology Society " Frontiers of Engineering and Computing in Health Care," Los Angeles, California, pp. 396-400, 1984.

[Wh86] White, M.W.: Compression systems for hearing aids and cochlear prostheses. Journal of Rehabilitation Research and Development. 23:25-39, 1986.

[Wh87a] White, M.W.: U.S. Patent 4,701,953, "Signal Compression System," October 20, 1987.

[Wh87b] White, M.W.: Neurophysiological and Psychophysical Considerations in the Design of a Cochlear Prosthesis. Annals of Otology, Rhinology and Larngology, 96:1, Part 2 (Supplement 128) p. 42, January-February, 1987.

[WhFiWi87] White, M.W., Finley, C.C., Wilson, B.S.: Electrical Stimulation Model of the Auditory Nerve: Stochastic Response Characteristics. Proceedings of the Ninth Annual Conference of the IEEE Engineering in Medicine and Biology Society, 4:1906-1907, 1987.

[WhOcMeSc90] White M.W., M.T. Ochs, M.M. Merzenich, E.D. Schubert, "Speech Recognition in Analog Multichannel Cochlear Prostheses: Initial Experiments in Controlling Classifications," IEEE Transactions on Biomedical Engineering 37:1002-1010, 1990.

[Wilson83] December 27, 1983. First quarterly progress report, NIH contract N01-NS-3-2356.

[Wilson84a] March 26, 1984. Second quarterly progress report, NIH contract N01-NS-3-2356.

[Wilson84b] Dec. 26, 1984. 5th quarterly progress report, NIH contract N01-NS-3-2356.

[Wilson1985/83] This proposal was submitted by Wilson on May 15, 1985 in response to RFP No. NIH-NINCDS-85-09. Appendix 3 of this proposal contains text from Wilson's 1983 proposal's "best and final offer" in response to RFP No. NIH-NINCDS-83-06.

[Wilson85a] March 26, 1985. 6th quarterly progress report, NIH contract N01-NS-3-2356.

[Wilson85b] June 26, 1985. 7th quarterly progress report, NIH contract N01-NS-3-2356.

[Wilson85c] Sept. 26, 1985. 8th quarterly progress report, NIH contract N01-NS-3-2356.

[Wilson85d] May 15, 1985, "Chronological record of projects directed by Blake Wilson," from Wilson's 1985 resume contained in [Wilson1985/83].

[Wilson87a] June 26, 1987. 7th quarterly progress report, NIH contract N01-NS-5-2396. Note: pp. 38-39 were missing from NIH website.

[Wilson87b] Sept. 26, 1987. 8th quarterly progress report, NIH contract N01-NS-5-2396.

[Wilson89a] July 31, 1989, 1st quarterly progress report, NIH contract N01-DC-9-2401.

[Wilson89b]  Oct. 31, 1989, 2nd quarterly progress report, NIH contract N01-DC-9-2401.

[Wilson90]  Jan. 31, 1990, 3rd quarterly progress report, NIH contract N01-DC-9-2401.

[Wilson92]  First quarterly progress report, NIH contract N01-DC-2-2401.

[Wilson2008]  Wilson BS, Dorman MF, Cochlear implants: a remarkable past and a brilliant future. Hearing research, 242(1), 3-21, (2008)

[Wilson2015a] Scroll-down to Wilson's Adjunct faculty web page to access his claim of inventions.  Invention claims are quoted verbatim from Wilson's faculty web page as of March 23, 2015.

[Wilson2015b]  Wilson BS, Getting a decent (but sparse) signal to the brain for users of cochlear implants, Hearing Research, Vol 322:24–38, April 2015.

[Wilson2016] Russ Prize, biography with statement of inventions by Wilson.  The same "statement of invention" is quoted verbatim from this web page within the Russ Prize website.

[Xu.etal97] J Xu, RK Shepherd, RE Millard, GM Clark, "Chronic electrical stimulation of the auditory nerve at high stimulus rates: a physiological and histopathological study." Hearing Research, Vol 105: 1–2, March 1997, pp. 1-29

## Appendix 1

### *Use of the naturally occurring stochastic behavior of 8th Nerve*

Why did UCSF have an interest in the seemingly unimportant "narrow edge of the operating range" where stochastic behavior was observed? Stochastic responsiveness became far more interesting when very surprising psychophysical evidence accumulated — indicating that many multi-channel cochlear implant patients may indeed be operating in that "edge of the operating range." [Wh84a], [Wh84b], [Wh87b]. In addition, UCSF saw the potential of developing stimulus processing strategies and electrode systems to further move the entire operating region of all patients into this "graded, stochastic-response region." Amusingly, it was later discovered that this "narrow edge or transition" was not so narrow after all! [Bruce.et.al99 a,b].

***How we are using our knowledge of the auditory nerve's stochastic behavior:***

- The concept of stimulating the auditory nerve at rates above 200-300 pps appeared quite unreasonable because the normal auditory nerve does not discharge at much higher rates than that — even for moderately short time-windows. The strong concern was that such "extreme stimulation" would put the fibers into saturation and could actually cause damage to the fibers. And, in fact, the Melbourne group [Ty.etal97], [Xu.etal97] found that at high pulse-rates, and at very high stimulus amplitudes, auditory neurons could be damaged — particularly with continuous, unmodulated or uninterrupted, stimulation. By stimulating at high pulse-rates, but at lower stimulus amplitudes, discharge-probabilities and corresponding discharge-rates could be kept within the normal range.

- Furthermore, accurate transmission of fine-grain temporal information on the auditory nerve becomes far more likely if most fibers do not fire at "saturated" high discharge rates, where strong refractory effects could significantly reduce the transmission of information about the stimulus. If, instead, most fibers were driven at low enough amplitudes (but at high stimulus pulse-rates), those fibers would would not be "saturated." As a consequence, those fiber's discharge probability would be a relatively strong function of pulse amplitude. In other words, it should be possible to effectively communicate information about the amplitude of each stimulus pulse to the CNS.

- Because of loudness summation in multichannel stimulation, each channels' stimulus level is set to a lower value than would be required for single channel stimulation to obtain the same loudness percept. [WhMeGa84, pp. 498-499] Cochlear implant companies normally need to reduce the stimulus amplitude on all channels when, at the end of "fitting" each channel, they turn all channels on for normal operation. Lowering stimulus levels can further "push" stimulation levels down into the stochastic operating region.

- Psychophysical evidence indicated that the use of higher pulse rates resulted in larger dynamic ranges, possibly due to an "effective broadening" of the region of stochastic response [Wh84a], [Wh84b], [Bruce.etal98b]. Data from neurophysiology experiments gave investigators insight into mechanisms that might be responsible for (1) the increased dynamic range and for (2) psychophysical measures of temporal response such as gap detection, MDT, and temporal release from masking. Possible mechanisms include (1) neural accommodation's "amplitude compression effect" and (2) "facilitation's augmentation" of the accommodation response at very high pulse-rates [Miller.etal2008], [Bruce.etal98 a], [Wh84a].

- UCSF's early electrophysiology was particularly useful in opening our minds to the possibility that the transmission of information at such high stimulus pulse-rates was possible. In electrophysiological data obtained from peripheral recording sites other than the acoustic nerve UCSF [Me.etal73] reported that the periods of low frequency sinusoidal stimuli were reproduced in the post-stimulus-time histograms for frequencies up to 400 to 700 Hz. when 50 stimulus repetitions were averaged. This data was qualitatively similar to the periodic discharges obtained with periodic acoustic stimulation [Ki65].

- The spread of the electric field around an electrode: For a given loudness and stimulus, a monopolar electrode (particularly when at a substantial distance from excitable-neural-tissue) will only need to elicit a low "discharge-probability per pulse" across many fibers — to achieve a given "sum of spikes." In contrast, focused stimulation (e.g., when a bipolar electrode is quite close-to-excitable-neural-tissue) will elicit a much larger "discharge-probability per pulse" in a small number of fibers to achieve the same "sum of spikes" [Miller.etal2008], [Bruce.etal98 a,b], [Wh84a].

- As current steering and/or electrode fabrication techniques are further improved, the number of effective electrode channels may be increased significantly. This leads to the possibility that the small number of fibers driven by each of these many "virtual channels" will be driven at even lower discharge probabilities — to elicit a given loudness. In other words, the work-load of the nerve fibers will be more evenly distributed across all fibers in the cochlea. If this is accurate, increasing the number of effective electrode-channels may "naturally improve the mimicking of acoustic-driven stochastic response" without any additional design effort!

  [Click here for additional information about UCSF's investigation of the "Stochastic Response of the 8th Nerve"]

## Appendix 2: Recent Publications of Wilson

In 2008, Wilson describes the essential concepts behind the high-rate-IP / CIS strategy:

> [Wilson2008] p. 10: There is a broad equivalence among contemporary signal processing strategies, e.g., CIS and ACE, in terms of the levels of performance achieved by patients on standard tests of speech understanding. This equivalence indicates, among other things, that there is no obvious advantage in either selecting or not selecting subsets of channels for each frame of stimulation. What seems to be important instead is the many features shared by the strategies, e.g., nonsimultaneous stimulation across electrodes; rates of stimulation near or above 1000 pulses/s/stimulated electrode; the same or highly-similar processing in each bandpass channel for extracting and mapping envelope signals; a low-pass cutoff in the envelope detector (or its equivalent) of at least 200 Hz; and at least 6–8 active (selected or always utilized) channels and corresponding sites of stimulation, to match or slightly exceed the number of effective channels that can be supported with present-day ST implants and the current processing strategies.

In 2015, Wilson, describes something seemingly quite different!

> [Wilson2015b] p. 33, column 2: I note that the gains in performance with CIS have sometimes been attributed to the nonsimultaneous stimulation across electrodes. However, the gains were produced with the discovery of the combination of many elements and not just nonsimultaneous stimulation, which had been used before (e.g., Doyle et al., 1964) but not in conjunction with the other elements. The breakthrough was in: (1) the combination; (2) exactly how the parts were put together; and (3) the details in the implementation of each part.

> [Wilson2015b] p. 33, column 1: CIS was a unique combination of new and prior elements, including but not limited to [sic]: (1) a full representation of energies in multiple frequency bands spanning a wide range of frequencies; (2) no further analysis of, or "feature extraction" from, this or other information; (3) a logarithmic spacing of center and corner frequencies for the bandpass filters; (4) a logarithmic or power law transformation of band energies into pulse amplitudes (or pulse charges); (5) customization of the transformation for each of the utilized electrodes in a multi-electrode implant, for each patient; (6) nonsimultaneous stimulation with charge-balanced biphasic pulses across the electrodes; (7) stimulation at relatively high rates at each of the electrodes; (8) stimulation of all of the electrodes at the same, fixed rate; (9) use of cutoff frequencies in the energy detectors that include most or all of the F0s and F0 variations in human speech; (10) use of those same cutoff frequencies to include most or all of the frequencies below the pitch saturation limits for implant patients; (11) use of the "4x oversampling" rule for determining minimum rates of stimulation; (12) use of current sources rather than the relatively uncontrolled voltage sources that had been used in some prior implant systems; and (13) a relatively high number of processing channels and associated electrodes (at least four but generally higher and not limited in number).

In my opinion, the relatively new elements were: 6, 7, 9, 10, 11* which redundantly represent the 2 fundamental advancements that this manuscript has documented in the "CIS strategy" section: They were: (1) non-simultaneous stimulation to reduce electric-field channel interactions, and (2) substantially increasing channel pulse-rate and the channel's envelope-filter cutoff frequency to enable the transmission of finer-grain temporal information to the 8th nerve.

In my opinion, the prior elements were: 1-5, 8**, 12-13 These items I consider to be "prior" because: Before 1985 one or more of the cochlear implant groups had already incorporated one or more of these elements in one of their processing strategies. In other words, those elements were already considered useful for designing some types of processors. -- and therefore were a part of the "standard conceptual tool-set" of many "practicing" cochlear implant scientists and engineers.

** Element 8 is commonly-used, effective, and easily implemented! However, it is not a "necessary constraint." – as long as the other relevant processing constraints are met (e.g., non-simultaneous stimulation of nearby channels).

* Element 11 is backed by little research or theory. For example, in quantitative questions regarding the coding and communications of information via electrical stimulation of the 8th nerve, it might be helpful to know something about how the 8th nerve responds to the expected "range" of amplitude-modulated high-rate pulsatile stimulation. – e.g., "range" of the spectra to be communicated. Currently, there is very little data available.

---

–

Wilson has continually failed to credit the research experiments and processing strategies described by the Utah/MIT, Melbourne, and UCSF groups that led directly to the high-rate IP (aka CIS) processor. – which UCSF initially proposed and implemented. In his recent Hearing Research article [Wilson2015b], Wilson failed to reference any publications from either the Utah/MIT or the UCSF research groups!

As indicated above, Wilson stated:

> [Wilson2015b] p. 33, column 2: I note that the gains in performance with CIS have sometimes been attributed to the nonsimultaneous stimulation across electrodes. However, the gains were produced with the discovery of the combination of many elements and not just nonsimultaneous stimulation, **which had been used before (e.g., Doyle et al., 1964)** ...

It is remarkable that Wilson chose to reference Doyle [DoDoTu64] rather than [EdDoBrMlPa78], [MeWh77], and/or [WhMeGa84]. Wilson's reference to Doyle is clearly incorrect. In 1973 UCSF [Me.etal73] referenced Doyle's work. Doyle's work was not useful to UCSF because (1) Doyle's goal was not to reduce electric field interactions between electrode-channels. Doyle did not indicate any awareness of electric field interactions. (2) Doyle's processing was a "pure volley theory" strategy that did not communicate any place/spectral information. It was designed for the purpose of keeping each channel's pulse rate at less than ~200 pps.

Specifically, Doyle's strategy was to temporally-sample the entire unfiltered speech waveform, and then temporally-distribute (demultiplex) the samples across multiple electrode channels. – in order to limit the maximum pulse rate delivered to each channel to no more than 200 pps. – for the purpose of preventing neural discharge-rate saturation.

Wilson’s reference to Doyle is deceptive, at worst, and completely irrelevant, at best. If Wilson had wanted to reference early papers describing the value-of (and methods-for) reducing electric field interactions within the cochlea, he would have referenced [EdDoBrMlPa78], [MeWh77], and/or [WhMeGa84].