

## Original Article

## Evaluation of a cochlear-implant processing strategy incorporating phantom stimulation and asymmetric pulses

Robert P. Carlyon\*, Jolijn Monstrey\*, John M. Deeks\* & Olivier Macherey\*<sup>†</sup>\*MRC Cognition & Brain Sciences Unit, Cambridge, UK, and <sup>†</sup>LMA, CNRS, UPR 7051, Aix-Marseille Univ., Marseille, France

The British Society of Audiology



The International Society of Audiology



## Abstract

**Objective:** To evaluate a speech-processing strategy in which the lowest frequency channel is conveyed using an asymmetric pulse shape and “phantom stimulation”, where current is injected into one intra-cochlear electrode and where the return current is shared between an intra-cochlear and an extra-cochlear electrode. This strategy is expected to provide more selective excitation of the cochlear apex, compared to a standard strategy where the lowest-frequency channel is conveyed by symmetric pulses in monopolar mode. In both strategies all other channels were conveyed by monopolar stimulation. **Design:** Within-subjects comparison between the two strategies. Four experiments: (1) discrimination between the strategies, controlling for loudness differences, (2) consonant identification, (3) recognition of lowpass-filtered sentences in quiet, (4) sentence recognition in the presence of a competing speaker. **Study sample:** Eight users of the Advanced Bionics CII/Hi-Res 90k cochlear implant. **Results:** Listeners could easily discriminate between the two strategies but no consistent differences in performance were observed. **Conclusions:** The proposed method does not improve speech perception, at least in the short term.

**Key Words:** Cochlear implants; asymmetric pulses; phantom stimulation; speech perception

Multi-channel cochlear implants (CIs) encode the frequency spectrum of sound by controlling the place of excitation along the auditory nerve (AN) array. This control is potentially limited both by the finite number of implanted electrodes and, in many devices, by the fact that the electrode array does not span the whole length of the cochlea. Recently, two techniques have been proposed that have the potential partially to overcome these limitations and to excite apical AN fibres more selectively than is possible with conventional stimulation methods.

“Phantom stimulation” (Wilson et al, 1992; Saoji & Litvak, 2010) refers to the situation where current is injected by one intra-cochlear electrode, and where a proportion  $\sigma$  is returned via a second intra-cochlear electrode, with the remainder returned by an extra-cochlear electrode. As shown in Figure 1a, this is equivalent to presenting two pulses of opposite polarity and differing amplitudes to each of the two intra-cochlear electrodes. This polarity difference results in the centres of gravity of the two voltage distributions corresponding to the stimulation at the two electrodes, “pushing away” from each other (cf. Macherey & Carlyon, 2012). This effect can also be clearly seen when  $\sigma = 1$  (Figure 1b), which is identical to standard bipolar

stimulation, and where the distance between the centres of the voltage distributions is greater than that between the two electrodes (as indicated by the dashed lines). Hence it is expected that there are maxima of excitation close to, but not exactly centred on, each of the two stimulating electrodes. However, for some smaller values of  $\sigma$ , the current delivered by one electrode produces only sub-threshold excitation, but still affects the pattern of current distribution. On average, a value of  $\sigma = 0.75$  produces close to the maximum shift in place-pitch perception, although this value varies across listeners (Saoji & Litvak, 2010; Macherey & Carlyon, 2012). Recently, Saoji et al (2013) used a forward masking paradigm to show that this place-pitch shift corresponded to a change in the centre of gravity of the excitation pattern. The only commercially available implants capable of producing phantom stimulation are those manufactured by Advanced Bionics.

An alternative technique was recently described in publications from our laboratory (Macherey et al, 2011; Macherey & Carlyon, 2012). As mentioned above, bipolar stimulation ( $\sigma = 1$ ) will normally produce two local sites of stimulation, close to each of the stimulated electrodes. However, this can be avoided by the use of

Correspondence: Robert P. Carlyon, MRC Cognition & Brain Sciences Unit, 15 Chaucer Rd, Cambridge CB2 7EF, UK. E-mail: bob.carlyon@mrc-cbu.cam.ac.uk

This is an open-access article distributed under the terms of the CC-BY-NC-ND 3.0 License which permits users to download and share the article for non-commercial purposes, so long as the article is reproduced in the whole without changes, and provided the original source is credited.

(Received 13 February 2014; accepted 1 June 2014)

ISSN 1499-2027 print/ISSN 1708-8186 online © 2014 British Society of Audiology, International Society of Audiology, and Nordic Audiological Society  
DOI: 10.3109/14992027.2014.932024

**Abbreviations**

AN	Auditory nerve
F0	Fundamental frequency
MCL	Most comfortable level
PS	Pseudomonophasic
SRT	Speech reception threshold
SYM	Symmetric

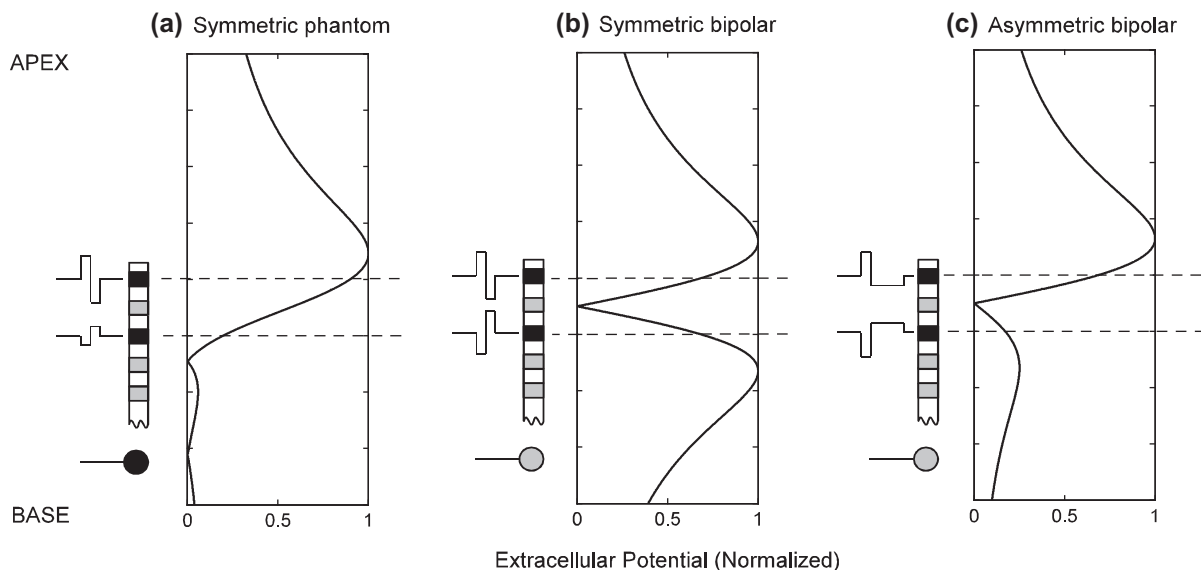
asymmetric pulse shapes (Figure 1c). It is now known that CI users are preferentially sensitive to anodic stimulation (Macherey et al, 2008). As a result, the place of excitation produced by an asymmetric waveform in bipolar mode depends on the electrode at which the anodic current is concentrated into a short high-amplitude portion (Macherey et al, 2010, 2011; Macherey & Carlyon, 2012; Carlyon et al, 2013). Macherey et al (2011) have shown that this technique can shift the locus of excitation by about 1 mm more apically than is possible with symmetric pulses in either monopolar or bipolar mode. Furthermore, the combination of asymmetric pulses and phantom stimulation (the “asymmetric phantom”) produces more consistent shifts in place-of-excitation than either asymmetric or phantom stimulation alone (Macherey & Carlyon, 2012). The use of asymmetric pulses in bipolar mode can produce place-of-excitation shifts in patients implanted with either Advanced Bionics or Cochlear corporation devices (Carlyon et al, 2013). It has also been shown that selective excitation of the apex, obtained using this technique, can improve temporal processing—as measured by an increase in the range over which increasing pulse rate produces an increase in pitch (Macherey et al, 2011).

There are several possible situations in which more selective activation of the apex might benefit speech perception. One, discussed previously (Carlyon et al, 2013), is specific to patients having residual low-frequency hearing in the implanted ear, and who have been fit-

ted with a short electrode array in order to preserve that remaining auditory function. This approach involves a trade-off between being able to stimulate as wide a range of AN fibres as possible, without the electrode array encroaching upon, and possibly damaging, the cochlear region responsible for the residual acoustic hearing. The ability to steer excitation more apically without increasing insertion depth might provide an advantage in this trade-off. Here, however, we examine situations more relevant to the majority of CI patients who rely entirely on their device for auditory sensation. For such patients, presenting the lowest frequency channel to a more apical set of AN fibres, that are more distinct from those excited by the remainder of the spectrum, might enhance the transmission of low-frequency speech cues. For a single speaker, such cues might include phonetic features such as nasality and the presence of voicing (Miller & Nicely, 1955). In addition, when two competing speakers have different fundamental frequencies (F0s), excitation in this low-frequency region may be more strongly correlated with the voice having the lower F0. The ability to group low-frequency excitation with that elicited by higher-frequency parts of the same source might be enhanced when the low-frequency excitation is more distinct from that produced by the rest of the mixture. We test these ideas by comparing performance in two speech-processing strategies, one of which has the lowest frequency channel presented to apical regions of the cochlea using the asymmetric phantom technique. We used a range of tests that were sensitive to information conveyed by low frequencies and that were designed to maximize the possibility of observing such an advantage. Our results showed that, although the two strategies were easily discriminable, no such advantages were observed.

**Study design and general methods**

A total of eight post-lingually deafened users of the Advanced Bionics Hi-Res90k device took part, although not all subjects took



**Figure 1.** The large panels show extracellular potential as a function of cochlear place for three methods of stimulation: (a) symmetric phantom ( $\sigma = 0.75$ ), (b) symmetric bipolar ( $\sigma = 1.0$ ), (c) asymmetric pulses in bipolar mode ( $\sigma = 1.0$ ). The smaller panels, to the left of each large panel, show the current at each of two intra-cochlear electrodes. As in Macherey & Carlyon (2012, Figure 1) and for illustration purposes, neural elements are assumed to be arranged along an axis parallel to the electrode array and located 3 mm away from it. The medium is assumed to be homogeneous. The dashed lines show the locations of the stimulating electrodes.

part in every experiment. All used the Fidelity 120 program with extended low filter in their daily program map. Further details of each subject are given in Table 1. Each experiment contrasted two strategies, one of which was a 15-channel HiRes strategy (Koch et al, 2004), with anodic-first symmetric biphasic pulses presented in monopolar mode to electrodes 1–15, in an apical-to-basal order. Pulse rate was 860 pps/channel with a phase duration of 24  $\mu$ s, except for subject AB3 for whom a rate of 719 pps and a phase duration of 32  $\mu$ s was used. The frequency-to-channel map was the “extended low” option available clinically in Advanced Bionics devices; this has lower filter settings for channel 1 of 250–421 Hz, compared to 350–421 Hz for the standard map and with a 15-channel strategy. This is labelled SYM in the remainder of the text. The other strategy was identical except that channel 1 was mapped onto electrodes 1 and 3 in “asymmetric phantom” mode. Specifically, electrode 1 was stimulated with a pseudomonophasic (PS) pulse in which the first phase was anodic and had a duration of 97  $\mu$ s, followed by a cathodic phase that was four times longer and that had a quarter of the amplitude of the leading phase. Seventy-five percent of the current was returned on electrode 3, which therefore received a polarity-inverted and three-quarter-sized version of the stimulus applied to electrode 1; the remaining current was returned via the (extra-cochlear) case. The two strategies are referred to as SYM and PS respectively, to describe the shape of the pulses for channel 1. Note that both strategies differed from that used clinically in that only 15 channels were used and that the pulse rate per channel was substantially lower than the 2900 pps used clinically. The first of these changes was made primarily because of technical limitations in the research software used, whereas the second was necessitated by the long duration of the pulses on channel 1 of the PS strategy.

For both strategies the inter-phase gap was always zero. The inter-pulse interval (the time interval between the offset of a pulse and the onset of the following pulse on another electrode) was always zero for PS. It was also zero for SYM except for the interval between the monopolar pulses of electrodes 1 and 2, which was 421  $\mu$ s for AB3 and 437  $\mu$ s for all other subjects. This was necessary because the asymmetric phantom pulse in the PS strategy was much longer than the monopolar pulses, and we needed to add such a silent gap to maintain the same rate across strategies.

In all cases stimulation involved pulse trains sent directly to the patient’s device using a laboratory processor, bypassing the clinical speech-processing algorithm. This was achieved using research software (BEDCS, BEPS, HRStream) provided by Advanced Bionics, together with, for some experiments, the APEX software platform (Laneau et al, 2005). The APEX software was modified from the publicly available version in order to provide greater flexibility and to interface with the Advanced Bionics device, and

served as a “wrapper” around BEDCS and HRStream. For BEDCS and BEPS, the hardware consisted of the clinical programming interface (CPI) connected to a Platinum Sound Processor (PSP). For HRStream, we used a USB-connected streaming interface board (SIB), provided by Advanced Bionics, which was connected to a PSP.

When speech stimuli were presented, a processing strategy was designed using the HRStream research interface and the Matlab programming language; this then processed the required .wav files and produced pulse-train sequences that were then sent to the patient’s device during testing. For both strategies, the signal processing was performed in Matlab and aimed to mimic the clinical HiRes strategy except that (1) the stimuli were not pre-emphasized and were not passed through the automatic gain control, and (2) the pulse rate was lower (719 pps or 820 pps per channel compared to the 2900-pps rate commonly used in HiRes). Input waveforms had a resolution of 16 bits and were directly passed through a bank of 6th order Butterworth bandpass filters similar to those used in the HiRes strategy (Nogueira et al, 2009). Envelopes were extracted by half-wave rectifying each filter’s output and computing the time integral of the half-wave rectified signal over each stimulation period. The 15 envelopes were further converted to dB units and saved to a 15-channel wav file that was written to the hard disk of the experimental computer. This 15-channel .wav file was finally converted to electrical stimulation patterns by HRStream according to the strategy that was tested (SYM or PS) and to the threshold and comfort levels measured in a given patient for that specific strategy. The conversion from envelope values in dB and electrical levels in  $\mu$ A followed a linear function. The range of envelope values mapped between T (threshold) and M (most comfortable) levels was 60 dB. For all subjects, T and M levels were measured for each channel and each strategy separately and saved in a fitting file. Impedances were obtained for all electrodes prior to the start of each experiment; these values were input to our experimental software which prevented current levels that would lead to a requested voltage that would exceed compliance, defined as 7 V. All stimuli were calibrated using a test implant connected to a digital storage oscilloscope. Impedances of all electrodes were checked using clinical software at the start and end of every testing session.

## Experiment 1: Discriminating between strategies

### *Rationale and Method*

A pre-requisite for any advantage of an experimental strategy is, of course, that it should be discriminable from the reference strategy. Because the SYM and PS strategies differed only in the stimulation applied to one of the 15 channels, we considered it important to

**Table 1.** Details of the eight subjects, listing age at testing, aetiology of deafness (where known), duration of deafness, and the duration of implant use.

<i>Subject</i>	<i>Age at testing (years)</i>	<i>Aetiology</i>	<i>Duration of deafness (years)</i>	<i>Implant use (years)</i>
AB1	79	Not known	Not known	3.0
AB2	79	Possibly noise induced	23	3.1
AB3	66	Unknown progressive	40	3.8
AB4	68	Otosclerosis	18	0.8
AB5	52	Ototoxicity	18	3.1
AB6	31	Unknown	> 5	1.1
AB7	64	Otosclerosis	25	3.0
AB8	57	Unknown	> 20	6.4

determine how easily our listeners could tell them apart. A trivial reason why two strategies might be discriminable but not differ in their usefulness would occur if they differed in overall loudness. The aim was that the two strategies would differ in the spatial distribution of neural excitation, which might be expected to correspond to a difference in pitch or timbre. We therefore loudness-balanced stimuli processed using the two strategies, and then required subjects to discriminate between them. To control for any residual loudness differences that remained despite this loudness balancing procedure, listeners discriminated between a SYM stimulus and several PS stimuli that spanned a range of closely-spaced levels. The idea was that, if listeners could only discriminate between the strategies using loudness cues, then there should be some level of the PS stimulus where discrimination is at or close to chance.

In this experiment only, the pulse rate per channel was reduced to 687 pps and the phase duration for the symmetric pulses (equal to the duration of the first phase for the PS pulses) was increased to 32  $\mu$ s. This phase duration was imposed by a technical requirement of the BEDCS software. Another feature particular to this experiment is that, rather than presenting speech sounds, the stimulation applied to each electrode was constant throughout each 400-ms stimulus. This was equivalent to the output of each strategy to a steady stimulus whose energy in each analysis filter was the same across all filters.

Prior to the main discrimination experiment, we needed to set the levels of the stimuli in the two strategies to be equal. This was achieved by performing the following steps: (1) The same current level was applied to the pulses on channels 2–15, and the level of this multi-electrode stimulus was adjusted by the listener to be comfortably loud (level 6 on the Advanced Bionics loudness chart; described here as the most comfortable level, or MCL). (2) The current level for channel 1 (presented in isolation) was also adjusted to be comfortably loud, separately for the SYM and PS strategies. (3) Channels 1 and 2–15 were then combined, keeping the same dB difference between them as obtained in stages (1) and (2), and, for each strategy separately, the level of the entire 15 channels was then adjusted to be comfortably loud. Throughout this adjustment the relative levels of all channels were constant in dB. (4) The level of the SYM pulses on channel 1 was set to 1.5 dB lower than that obtained in stage (3), and the PS pulse train on channel 1 was loudness balanced to this level. For this loudness balancing the listener performed four adjustments, two with the SYM level fixed and with the PS stimulus varying, and two with the PS stimulus fixed and the SYM stimulus varying (for further details see McKay & McDermott, 1998; Carlyon et al, 2010). The resulting levels of the SYM and PS pulse trains on channel 1 were then each combined with channels 2–15, with the level of channels 2–15 set to 1.5 dB below that obtained in stage (3).

In the main experiment listeners discriminated between the SYM stimulus at the level obtained in the preliminary stages described above, and the PS stimuli at the loudness-matched level  $\pm 0, 0.3, 0.6, 0.9,$  and  $1.2$  dB. Based on preliminary measures, an exception was made for listener AB5, who was tested at relative levels of  $-0.6, -0.3, 0, +0.3, +0.6, +0.9, +1.2, +1.4,$  and  $+1.7$  dB. In each trial listeners were presented with two instances of the SYM stimulus and one of the PS stimulus, with the latter occurring with equal probability in either the second or third interval. No feedback was provided. Listeners were instructed to identify the interval containing the sound that was different from that in the other two intervals. Each block consisted of six trials at each of the nine levels of the PS stimulus. Data were typically obtained from the average of five blocks, leading to a total of 30 trials per point for each listener.

Exceptions were listener AB5 who completed eight blocks, and listener AB3 who completed four.

### Results

Discrimination scores are plotted as a function of the level of the PS stimulus in Figure 2. It can be seen that all scores are at or close to 100%, with no sign of any range of levels where performance approaches chance. Hence the two strategies were easily discriminable, even in the absence of any loudness difference. A caveat is that the relative levels of stimulation on the different channels differed between this experiment and experiments 2–4. Here, the level of channel 1 was set so as to be approximately equally as loud as channels 2–15 combined, and the levels of those channels were the same in both strategies. In experiments 2–4, which each channel produced, in isolation, approximately the same loudness for a given strategy, but the levels on channels 2–15 could have differed between strategies.

## Experiment 2: Consonant identification

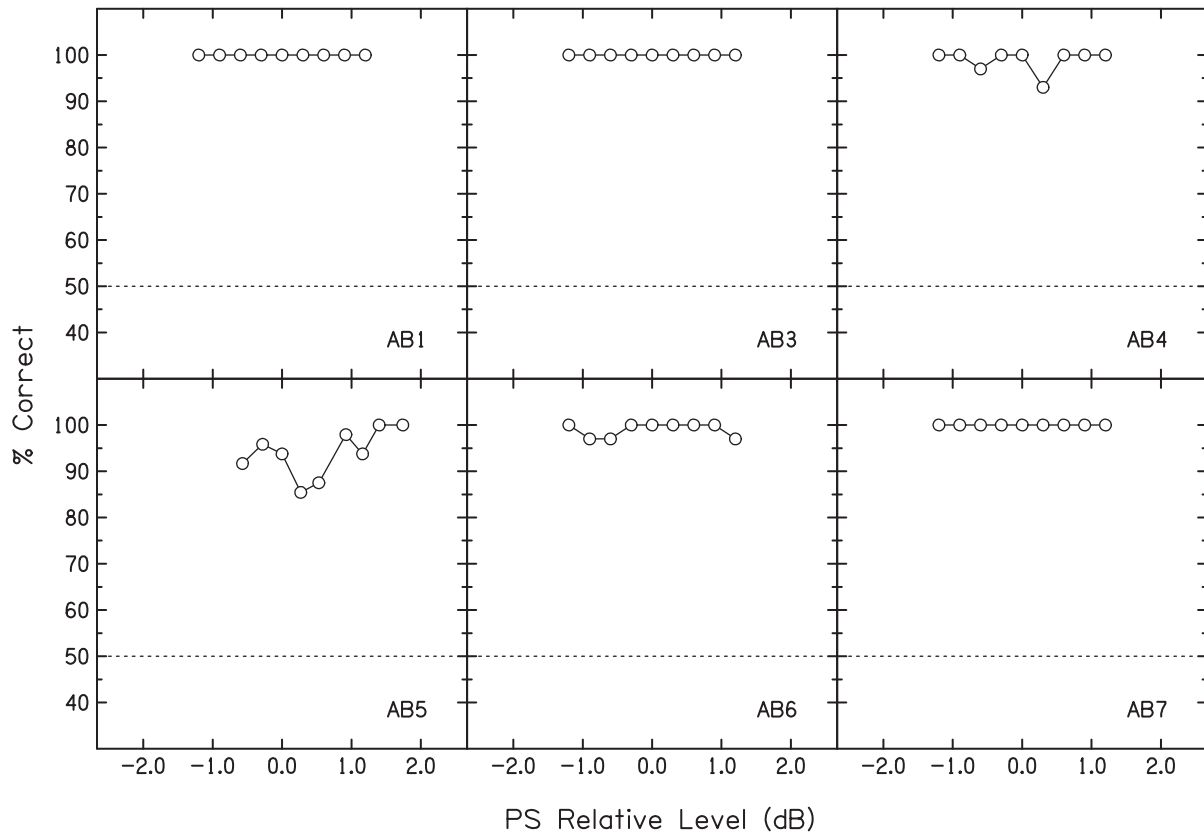
### Method

Experiment 2 measured identification of consonants in a vowel-consonant-vowel context, presented using the SYM and PS strategies. We used a set of consonants, pairs of which differed in only one phonetic feature (Table 2), allowing us to perform an information transmission analysis (Miller & Nicely, 1955). Because the two strategies differed only in the lowest-frequency channel, it was important for us to be able separately to analyse transmission of those features that are conveyed by low-frequency energy. For example, transmission of frication would not be expected to differ between the two strategies, whereas both nasality and voicing, which are conveyed by low-frequency information, might be more effectively transmitted by the PS strategy.

The ten consonants /b/, /d/, /g/, /m/, /n/, /ŋ/, /s/, /z/, /f/ and /v/ were recorded in a vac context, both with the vowel (v) equal to /I/ and /a/. They were spoken by a native speaker of southern British English (author RPC) and recorded (44 100 Hz sampling rate; 16-bit resolution) in a double-walled sound-insulating room with an AKG microphone, model C1000S, and a Marantz Portable Solid State Recorder, model PMD670. The root mean square levels of all tokens were then adjusted in software to be equal. Each stimulus was then processed with the SYM and PS strategies. Subjects adjusted the level of the stimuli having the central consonant /b/ to their MCL by gradually adjusting the level of the input sound waveform in steps of  $\pm 1, 3,$  and  $5$  dB by clicking on one of six virtual buttons presented on a computer screen. All other stimuli with the same vowel were set to the same level.

In the training part of the experiment the stimuli were those recorded with the /I/ vowels, and rough English transcriptions of each possible syllable were presented as virtual buttons on the computer screen: “eebee”, “eemee”, “eedee”, “eenee”, “eegee”, “eengee”, “eese”, “eezee”, “eefee”, and “eevee”. It was explained that the consonant in “eegee” was a plosive as in the word “good”. The listener could click on each button and listen to the corresponding sound as often as required.

The main part of the experiment used syllables recorded with the /a/ vowel. Rough transcriptions of the ten possible sounds (“aba”, “ama” etc.) were presented on virtual response buttons. The different sounds were presented in random order and the listener was required to identify that sound by clicking on the appropriate button.



**Figure 2.** Results of experiment 1. Each panel shows one listener’s percent correct as a function of the level of the PS stimulus relative to the loudness-balanced level.

No feedback was given. Each run consisted of three presentations of each syllable processed using one strategy (SYM or PS). There were nine runs per condition leading to a total of 27 presentations of each syllable per strategy and listener.

*Results*

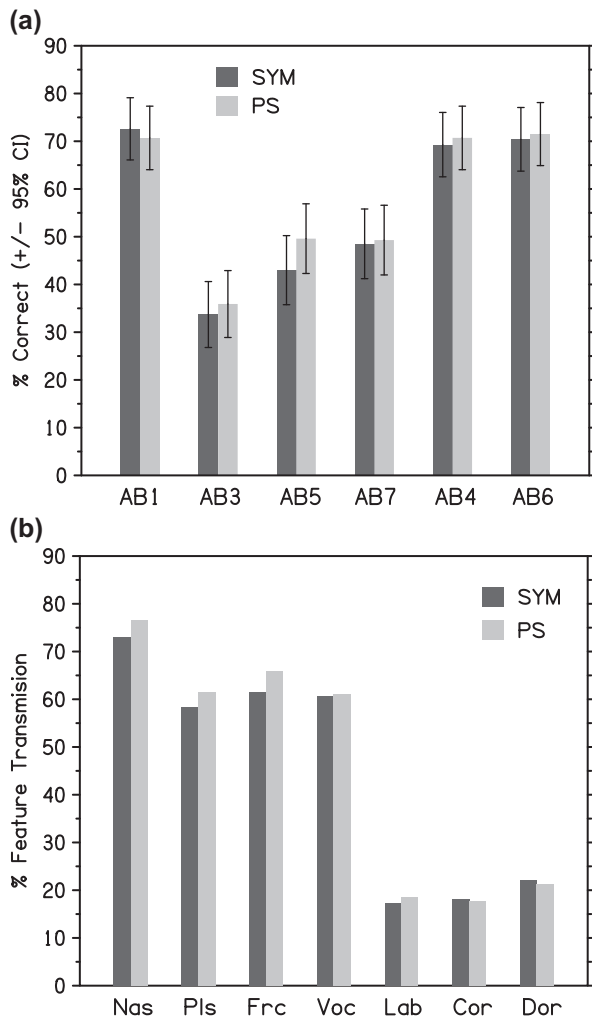
Percent correct is shown, together with 95% confidence intervals, for the two strategies in Figure 3a. It can be seen that all listeners performed significantly above the chance level of 10% for both strategies, but that performance for the two strategies was similar. This was confirmed using a paired-samples t-test ( $t(5) = 1.52$ ;  $p = 0.19$ ).

A more detailed analysis of the results can be obtained by inspecting the confusion matrices, and, in particular, by examining the results of the feature transmission analysis. The confusion matrix for each strategy, averaged across listeners, is shown in

**Table 2.** Phonetic features for the ten consonants used in experiment 2. Features are: manner of articulation (plosive (1), fricative (2), nasal (3)), place of articulation (labial (1), coronal (2), and dorsal (3)), and voicing (voiced (1) or unvoiced (2)).

	<i>b</i>	<i>m</i>	<i>d</i>	<i>n</i>	<i>g</i>	<i>ng</i>	<i>s</i>	<i>z</i>	<i>f</i>	<i>v</i>
Manner	1	3	1	3	1	3	2	2	2	2
Place	1	1	2	2	3	3	2	2	1	1
Voicing	1	1	1	1	1	1	2	1	2	1

Table 3a and 3b, and the transmission index for each feature is shown for the two strategies in Figure 3b. The results are consistent with previous studies with CI users in that there was better transmission of manner (plosive, fricative, nasal) and voicing than for place-of-excitation (e.g. Munson et al, 2003). However, there are no features for which we observed a statistically significant improvement for the PS, compared to the SYM strategy. This was true even for features for which low-frequency energy is important. Inspection of Table 3a and 3b reveals that listeners made very few confusions between pairs of stimuli that differed only in nasality (/b/ vs /m/, /d/ vs /n/, and /g/ vs /ŋ/; mean confusions 1.25%, 1.5%, and 1.5% respectively). Hence one reason for the lack of an effect may have been that the transmission of this feature was close to ceiling for most listeners. Table 3c shows that transmission of nasality for two listeners, AB3 and AB7, was well below ceiling for both strategies, but substantially better for the PS than for the SYM strategy (AB3: 40% vs. 22%, AB7: 47% vs. 32%). We can therefore not rule out the possibility that the PS strategy enhances the transmission of nasality, but that this benefit was obscured by ceiling effects for most listeners. However, the consequences of this benefit are likely to be negligible, given that the transmission of nasality was good for most listeners with both strategies, and because no listeners correctly reported a significantly larger proportion of consonants with the PS than with the SYM strategy. In addition, there was no consistent benefit for the transmission of voicing, which also depends on low-frequency energy during the consonant (e.g. /s/ vs /z/), and which was well below ceiling for most listeners.



**Figure 3.** Results of the consonant identification task of experiment 2. (a) Percent correct for each listener for the SYM and PS strategies. (b) Transmission of each phonetic feature averaged across listeners. Nas = manner (nasality), Pls = manner (plosive), Frc = manner (fricative), Voc = voicing, Lab = place (labial), Cor = place (coronal), Dor = place (dorsal).

### Experiment 3: Lowpass filtered speech

#### Rationale and Method

Because the SYM and PS strategies differ only in the lowest-frequency channel, a difference in their effectiveness may be more

clearly demonstrated in conditions where low-frequency energy is essential for good performance. Experiment 3 therefore tested the identification of low-pass filtered sentences. Specifically, the stimuli were taken from the IHR sentence lists (MacLeod & Summerfield, 1990) and lowpass filtered with a 6th order Butterworth having a 3-dB-down point at 750 Hz. This might be viewed as an extreme version of the situation where the speaker is close to the unimplanted ear of a CI user, and where the head shadow attenuates higher-frequency components.

Two of the processed sentences (“They moved the furniture” and “He tore his shirt”) were adjusted by each listener for both strategies in order to obtain MCL, as in experiment 2. The lowest MCL from these two sentences, for each strategy, was then used in the main experiment for all sentences. The IHR sentence lists each contain 15 sentences with three keywords per sentence. Participants completed four lists per strategy, and scores were defined as the total number

**Table 3.** (a) and (b). Confusion matrices, averaged across listeners, for the vCv stimuli used in experiment. (c) Feature analysis results for listeners AB1 to AB7. (a) SYM strategy confusion matrix. Correct responses are shown in bold.

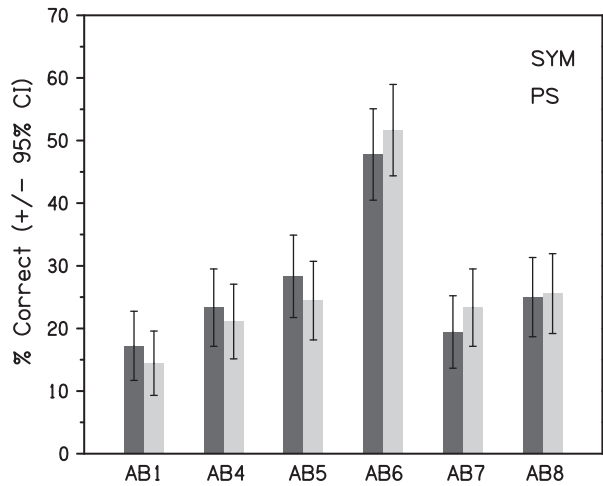
	aba	ama	ada	ana	aga	anga	asa	aza	afa	ava
aba	<b>62%</b>	2%	3%	0%	7%	5%	2%	7%	1%	10%
ama	0%	<b>52%</b>	0%	27%	2%	2%	2%	1%	1%	13%
ada	1%	1%	<b>51%</b>	3%	30%	5%	2%	6%	0%	2%
ana	1%	5%	0%	<b>77%</b>	1%	14%	1%	0%	0%	2%
aga	0%	0%	1%	0%	<b>96%</b>	3%	0%	1%	0%	0%
anga	1%	32%	0%	38%	1%	<b>14%</b>	0%	2%	2%	11%
asa	0%	0%	0%	0%	0%	0%	<b>58%</b>	7%	32%	3%
aza	0%	1%	1%	1%	0%	0%	12%	<b>69%</b>	0%	17%
afa	0%	0%	0%	0%	0%	0%	22%	16%	<b>56%</b>	6%
ava	0%	0%	20%	0%	9%	0%	7%	33%	1%	<b>28%</b>

(b) PS strategy confusion matrix.

	aba	ama	ada	ana	aga	anga	asa	aza	afa	ava
aba	<b>69%</b>	2%	3%	2%	9%	4%	2%	4%	0%	6%
ama	1%	<b>64%</b>	0%	20%	2%	2%	0%	1%	1%	10%
ada	1%	0%	<b>53%</b>	2%	31%	4%	1%	6%	0%	1%
ana	1%	6%	1%	<b>82%</b>	0%	8%	0%	0%	0%	2%
aga	0%	0%	1%	1%	<b>96%</b>	2%	0%	1%	0%	0%
anga	0%	33%	1%	48%	0%	<b>13%</b>	1%	0%	0%	5%
asa	0%	0%	0%	0%	0%	0%	<b>52%</b>	8%	37%	2%
aza	0%	0%	0%	0%	1%	0%	6%	<b>79%</b>	0%	14%
afa	1%	0%	0%	0%	0%	0%	25%	19%	<b>48%</b>	7%
ava	4%	0%	21%	0%	7%	1%	10%	34%	0%	<b>24%</b>

(c) Feature analysis results.

	AB1		AB3		AB4		AB5		AB6		AB7	
	SYM	PS	SYM	PS	SYM	PS	SYM	PS	SYM	PS	SYM	PS
manner: nasality	100%	88%	23%	40%	100%	96%	83%	93%	100%	96%	32%	47%
manner: plosive	72%	75%	34%	38%	72%	56%	36%	56%	68%	64%	68%	81%
manner: fricative	59%	63%	63%	81%	79%	65%	54%	67%	57%	56%	57%	65%
voicing	88%	93%	20%	17%	70%	68%	54%	49%	63%	71%	68%	68%
place: labial	15%	13%	2%	3%	29%	41%	4%	4%	33%	32%	20%	17%
place: coronal	41%	35%	3%	3%	21%	26%	2%	5%	35%	36%	7%	2%
place: dorsal	30%	27%	8%	12%	43%	32%	16%	19%	21%	26%	14%	12%



**Figure 4.** Percent correct for each listener and strategy in the sentence test of experiment 3. Error bars show 95% confidence intervals.

of correct keywords out of 180. The sentence lists were selected from lists 3–10 inclusive. Listeners AB1, AB4, AB6, and AB7 heard the odd-numbered lists processed using the SYM strategy and the even-numbered lists processed with the PS strategy, whereas the opposite was true for listeners AB8 and AB5. Unfortunately two other scheduled listeners were unavailable meaning that the assignment of lists to strategy was not perfectly counterbalanced; therefore the list-to-strategy assignment was entered as a fixed factor when analysing the results (see below). Prior to data collection listeners practiced with one list per strategy, selected from lists 1 and 2. They were warned that the task might prove difficult, and were instructed to guess if not sure what words they had heard.

### Results

The percentage of keywords correctly reported by each listener and for each strategy are shown in Figure 4. Performance was generally quite low, presumably due to the lowpass filtering applied to the stimuli. It is clear that there was no consistent difference between the strategies. This was confirmed by an analysis of variance, with strategy as the within-subject factor and the list-to-strategy assignment as the between-subjects factor. Neither the main effects nor the interaction were significant (Strategy:  $F(1,4) = 0.49$ ,  $p = 0.52$ . Assignment:  $F(1,4) = 0.014$ ,  $p = 0.92$ . Interaction:  $F(1,4) = 0.05$ ,  $p = 0.83$ ).

## Experiment 4: Speech perception with a concurrent speaker

### Rationale and Method

The aim of this experiment was to test whether a presumably greater independence between the first two channels of the PS strategy would allow subjects to better follow the fundamental frequency of a speaker in the presence of a concurrent voice. For this experiment, the cut-off frequencies of the first four analysis filters were modified and encompassed a larger range than in previous experiments (100–164, 164–270, 270–444, and 444–730; compared to 250–421, 421–505, 505–607, and 607–730 for channels 1, 2, 3, and 4, respectively). They had, therefore, also a larger bandwidth. Higher-frequency filters remained unchanged.

Target sentences were obtained from a coordinate response measure corpus recorded with British English speakers (Kitterick & Summerfield, 2007). All sentences had the following structure “ready <call sign>, go to <colour> <number> now.” The subject’s task was to correctly identify the “colour” and “number” key-words by clicking on virtual buttons displayed on a computer screen. Utterances from one male speaker (M1) and one female speaker (F3) of the corpus were used. Before presentation, target sentences were mixed with a speech masker. The maskers were IEEE sentences played backwards to limit the amount of informational masking. To maximize F0 differences between target and masker, male target sentences were mixed with female masker sentences and vice-versa.

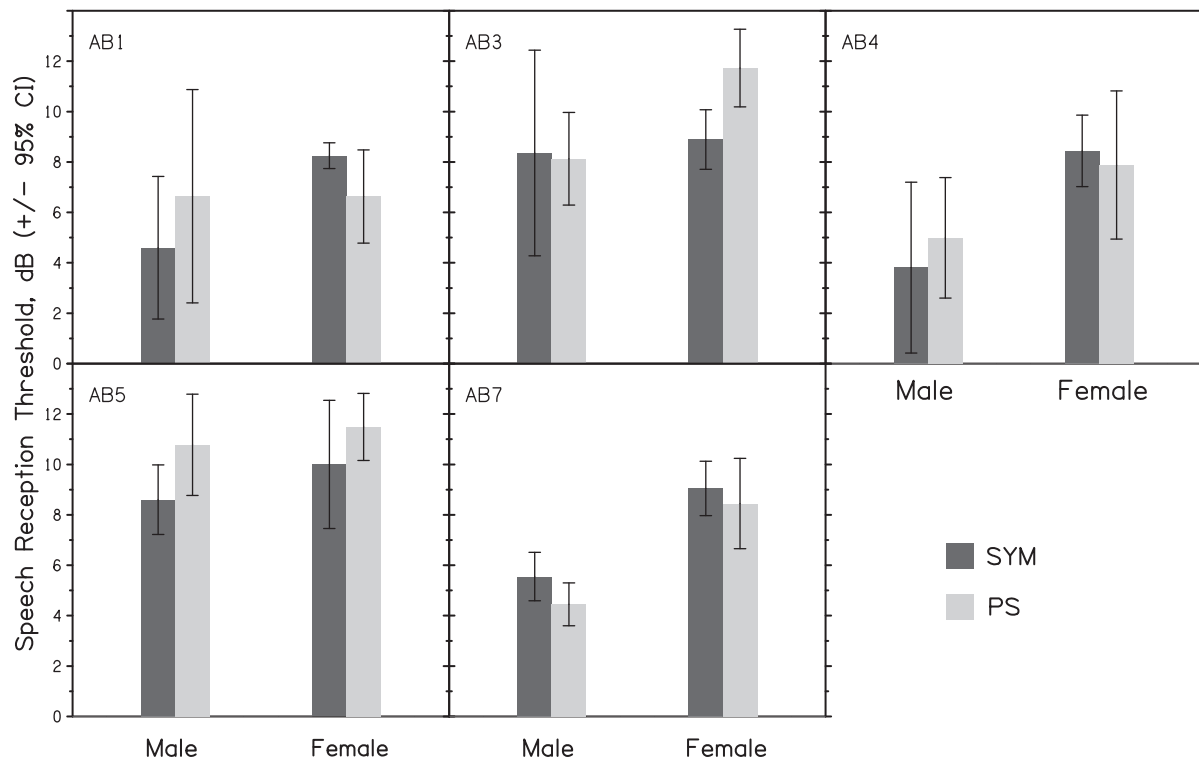
The level of the target sentence was fixed and that of the masker was varied. To determine the level of the target sentences to be used, a target sentence processed through SYM or PS was first presented to the subject in quiet. The subject was asked to adjust the level of this input file so that it sounded most comfortable separately for each strategy. Subsequently, the same level adjustment was repeated in the presence of a masker using a target-to-masker ratio (TMR) of +15 dB. The minimum of the four level values, i.e. two strategies (SYM and PS) in two conditions (quiet or noise) minus 15 dB was used as the level of the target sentences in the remainder of the experiment.

The training phase consisted of two stages. First, two blocks of 50 sentences were presented in quiet to the subject, one block of each strategy (PS and SYM). At the same time the sentences were played, they were also displayed on the screen so that the subject knew what was being said. This has been shown to speed up training of vocoded sentences in normal-hearing subjects (Davis et al, 2005). In the second stage, two additional blocks of 50 sentences were also presented in quiet to the subject. The subject had to indicate the “colour” and “number” keywords and feedback was provided. The gender of the target sentences changed from step 1 to step 2. In each stage, only one target gender was used for each subject and strategy, but the allocation of target gender to strategy was different for the two stages. For example, a listener might hear the PS strategy with the male target speaker and the SYM strategy with the female target in stage 1, and then the female speaker/SYM strategy and male speaker/PS strategy in stage 2. In this way all subjects experienced each strategy with both target genders prior to the main experiment.

The test phase was an adaptive measure of speech reception threshold (SRT). The initial target to-masker-ratio (TMR) was +15 dB. The TMR changed in 5-dB steps for the first four reversals and 2.5 dB steps for the last eight reversals. SRT was calculated as the average of the last six reversals. There were four conditions differing in the strategy (PS or SYM) and in the gender of the target sentences (Male and Female). SRTs for each of these four conditions were measured in turn, and this procedure was repeated so that there were four measures per condition. Five subjects took part.

### Results

Figure 5 shows the SRTs obtained for the five subjects in each of the four conditions. A two-way repeated-measures analysis of variance [strategy  $\times$  speaker gender] revealed no difference between the two strategies ( $F(1,4) = 1.5$ ,  $p = 0.28$ ). The male speaker was more intelligible (yielding smaller SRTs) than the female speaker, as shown by a significant effect of gender ( $F(1,4) = 21.7$ ,  $p = 0.01$ ). However, this effect of gender did not depend on which strategy was used, as



**Figure 5.** Speech reception thresholds for the identification of sentences in the presence of a competing talker (experiment 4). Each panel shows the results for one listener, separately for the male and female target talkers. Error bars show 95% confidence intervals.

shown by a non-significant interaction between strategy and gender ( $F(1,4) = 0.2$ ,  $p = 0.6$ ).

## Discussion

Our first experiment showed that listeners could easily discriminate between the two strategies. They typically reported the difference in terms of pitch or timbre, as would be expected from previous single-channel studies showing that the form of asymmetric phantom stimulation used here produces a more apical locus of excitation than occurs when stimulating the most apical electrode with symmetric pulses in monopolar mode (Macherey et al, 2011; Macherey & Carlyon, 2012). However, no advantage of the strategy was observed for any of the speech tests used.

Some previous studies have reported an advantage for processing strategies that are designed to enhance the transmission of low-frequency information. However, it is not usually the case that this aspect of the sound encoding is changed in isolation. For example, Koch et al (2004) compared the HiRes strategy to the listener's preferred "conventional" strategy (including continuous interleaved sampling, CIS) and found superior scores for HiRes on a number of speech tests. One way in which the HiRes strategy differs from CIS is that it more accurately represents the fine structure at the output of low-frequency analysis filters. However, as the authors acknowledged, the two strategies differed both in the number of channels (16 for HiRes, 8 for the "conventional" strategy) and in pulse rate. The comparison of strategies that differ in multiple dimensions is ubiquitous in cochlear implant research, and stretches back at least as far as Wilson's (1991) influential comparison of a six-channel CIS strategy with a four-channel analogue strategy.

A limitation of the present study is that performance on the two strategies was measured acutely, rather than after an extended period during which listeners could adapt to the new strategies. It is well-known that performance on a range of speech tests improves during the months after implantation, and that the time course of this improvement can be quite extended. For example, Tyler et al (1997) reported that speech perception improved significantly during the first month after implantation, and there was also a significant improvement between months 9 and 30 post-implantation. More relevant to the present study are those experiments that have investigated the effects of changing from one processing strategy to another. In particular, there is evidence both from experiments with CI listeners (Fu et al, 2002) and with simulations of CI hearing presented to normal hearing listeners (Rosen et al, 1999) that listeners take time to adapt to a new frequency-to-place map. This is relevant to the present study because our two strategies differed primarily in the place of excitation elicited by the most apical channel, and because it could be argued that the PS strategy would take longer to "learn" than the SYM strategy. A difference is that most previous studies compared a new map to the clinical map, rather than investigating adaptation to two new maps, as done here. An exception is the study by Henshall & McKay (2002) who studied the effects of several new maps, including two different 10-electrode maps. They found that listeners differed markedly in which of these two maps yielded the best performance, but that the pattern of results could not be accounted for by the similarity of each map to the listener's clinical map.

Litvak et al (2011) have presented preliminary results from a study that implemented a strategy similar to ours but with symmetric pulses (the "symmetric phantom"). They reported an improvement compared to the standard HiRes strategy after three months of take-home experience. This result differs from the



lack of improvement observed here, and the discrepancy could possibly be due to the extended exposure to the new strategy in their study. However their strategy also differed from HiRes in that the frequency range of the input to the lowest channel was shifted downwards, and so the improvement could have been due to this increased frequency range rather than to the phantom stimulation. Clearly the only way to resolve this issue would be to perform a take-home study with the experimental and standard strategies counterbalanced, and in which the same input frequency range was used for both. However we should note that the complete lack of an advantage reported here occurred despite the fact that we included stimuli and analyses (lowpass filtered speech; feature analysis) that were especially sensitive to the transmission of low-frequency information.

### Acknowledgements

This research was supported by a grant from MRC Technology's Development Gap Fund. Author OM acknowledges support from the ANR (Project ANR-11-PDOC-0022).

**Declaration of interest:** The authors declare no conflict of interest.

### References

- Carlyon R.P., Deeks J.M. & McKay C.M. 2010. The upper limit of temporal pitch: Stimulus duration, conditioner pulses, and the number of electrodes stimulated. *J Acoust Soc Am*, 127, 1469–1478.
- Carlyon R.P., Deeks J.M. & Macherey O. 2013. Polarity effects on place pitch and loudness for three cochlear-implant designs and at different cochlear sites. *J Acoust Soc Am*, 134, 503–509.
- Davis M.H., Johnsrude I.S., Hervais-Adelman A., Taylor K. & McGettigan C. 2005. Lexical information drives perceptual learning of distorted speech: Evidence from the comprehension of noise-vocoded sentences. *J Exp Psychol Gen*, 134, 222–241.
- Fu Q.J., Shannon R.V. & Galvin III, J.J. 2002. Perceptual learning following changes in the frequency-to-electrode assignment with the Nucleus-22 cochlear implant. *J Acoust Soc Am*, 112, 1664–1674.
- Henshall K. & McKay C.M. 2002. Frequency-to-electrode allocation and speech perception with cochlear implants. *J Acoust Soc Am*, 111, 1036–1044.
- Kitterick P.T. & Summerfield A.Q. 2007. The role of attention in the spatial perception of speech. *Assoc Res Otolaryngol Abstr*, 30, 423.
- Koch D.B., Osberger M.J., Segel P. & Kessler D. 2004. HiResolution™ and conventional sound processing in the HiResolution™ bionic ear: Using appropriate outcome measures to assess speech recognition ability. *Audiology and Neuro-Otology*, 9, 214–223.
- Laneau J., Boets B., Moonen M., van Wieringen A. & Wouters J. 2005. A flexible auditory research platform using acoustic or electric stimuli for adults and young children. *Journal of Neuroscience Methods*, 142, 131–136.
- Litvak L.M., Saoji A.I., Bhattacharya A., Agrawal S., Nogueira W. et al. 2011. *Functionally Extending The Electrode Array Longitudinally Using Phantom Stimulation Protocols*. Conference on Implantable Auditory Prostheses; Asilomar, USA.
- Macherey O., Carlyon R.P., Deeks J.M., van Wieringen A. & Wouters J. 2008. Higher sensitivity of human auditory nerve fibers to positive electrical currents. *J Assoc Res Otolaryngol*, 9, 241–251.
- Macherey O., van Wieringen A., Carlyon R.P., Dhooge I. & Wouters J. 2010. Forward-masking patterns produced by symmetric and asymmetric pulse shapes in electric hearing. *J Acoust Soc Am*, 127, 326–338.
- Macherey O., Deeks J.M. & Carlyon R.P. 2011. Extending the limits of place and temporal pitch perception in cochlear implant users. *J Assoc Res Otolaryngol*, 12, 233–251.
- Macherey O. & Carlyon R.P. 2012. Place-pitch manipulations with cochlear implants. *J Acoust Soc Am*, 131, 2225–2236.
- MacLeod A. & Summerfield Q. 1990. A procedure for measuring auditory and audio-visual speech-reception thresholds for sentences in noise: Rationale, evaluation, and recommendations for use. *Brit J Audiol*, 24, 29–43.
- McKay C.M. & McDermott H.J. 1998. Loudness perception with pulsatile electrical stimulation: The effect of interpulse intervals. *J Acoust Soc Am*, 104, 1061–1074.
- Miller G.A. & Nicely P.E. 1955. An analysis of perceptual confusions among some English consonants. *J Acoust Soc Am*, 27, 338–352.
- Munson B., Donaldson G.S., Allen S.L., Collison E.A. & Nelson D.A. 2003. Patterns of phoneme perception errors by listeners with cochlear implants as a function of overall speech perception ability. *J Acoust Soc Am*, 113, 925–935.
- Nogueira W., Litvak L., Edler B., Ostermann J. & Buchner A. 2009. Signal processing strategies for cochlear implants using current steering. *EURASIP Journal on Advances in Signal Processing*.
- Rosen S., Faulkner A. & Wilkinson L. 1999. Adaptation by normal listeners to upward spectral shifts of speech: Implications for cochlear implants. *J Acoust Soc Am*, 106, 3629–3636.
- Saoji A.A. & Litvak L.M. 2010. Use of 'phantom electrode' technique to extend the range of pitches available through a cochlear implant. *Ear Hear*, 31, 693–701.
- Saoji A.A., Landsberger D.M., Padilla M. & Litvak L.M. 2013. Masking patterns for monopolar and phantom electrode stimulation in cochlear implants. *Hearing Research*, 298, 109–116.
- Tyler R.S., Parkinson A.J., Woodworth G.G., Lowder M.W. & Gantz B.J. 1997. Performance over time of adult patients using the Ineraid or Nucleus cochlear implant. *J Acoust Soc Am*, 102, 508–522.
- Wilson B., Lawson D., Zerbi M. & Finley C. 1992. Speech processors for auditory prostheses (First Quarterly Progress Report). NIH. project number N01-DC-2-2401.
- Wilson B.S., Finley C.C., Lawson D.T., Wolford R.D., Eddington D.K. et al. 1991. Better speech recognition with cochlear implants. *Nature*, 352, 236–238.