

Differential-Rate Sound Processing for Cochlear Implants

David B Grayden^{1,2}, Sylvia Tari^{2,3}, Rodney D Hollow³

¹National ICT Australia, c/- Electrical & Electronic Engineering,
The University of Melbourne Vic 3010

²The Bionic Ear Institute, East Melbourne Vic 3002

³The Royal Victorian Eye & Ear Hospital, East Melbourne Vic 3002
d.grayden@ee.unimelb.edu.au

Abstract

The aim of this study was to evaluate the Differential-Rate Sound Processing (DRSP) strategy for speech perception by users of the Nucleus cochlear implant. The DRSP strategy uses both low and high rates of stimulation – applying a low rate of stimulation to apical, lower-frequency electrodes and a high rate of stimulation to basal, higher-frequency electrodes. Ten patients implanted with the Nucleus CI-24M cochlear implant were evaluated with three speech processing strategies: 250 pulses per second (pps) over all electrodes, 1500 pps over all electrodes, and DRSP with 250 pps on electrodes up to 2400 Hz and 1500 pps for the remaining electrodes. There was no difference found between strategies across research subjects for consonant, vowel, CVC words and sentences-in-noise tests. However, there were significant individual subject differences confirming that there is still a case for considering low-rate stimulation when Mapping patients. The DRSP strategy demonstrated a compromise between the fixed low-rate and high-rate strategies that can perform equal to or better than either of these strategies for most subjects.

1. Introduction

An open question that still remains for cochlear implants is deciding what rate of stimulation will be most effective for each user. The first successful multi-channel cochlear implant was implanted in 1978 (for a review of cochlear implant design and development, see Grayden & Clark, 2006). An early speech processing strategy extracted the second formant (F2) frequency and fundamental frequency (F0) to control electrode stimulation (Tong *et al.*, 1979). The frequency of F2 controlled the location of electrode stimulation and F0 controlled the rate of stimulation. Improvements were made by also extracting the first formant (F1) frequency and adding a second stimulated electrode for each pitch period (Dowell, Seligman, Blamey & Clark, 1987). The MULTIPLEAK stimulation strategy added stimulation of a number of fixed electrodes to better represent high-frequency information (Dowell, Whitford, Seligman, Franz & Clark, 1990). Each of these strategies was characterized by the property that they used varying rates of stimulation controlled by F0.

The next strategies developed were the Spectral Maxima Sound Processor (SMSP) strategy (McDermott, McKay, & Vandali, 1992) and the Continuous Interleaved Sampler (CIS) strategy (Wilson, Lawson, Zerbi & Finley, 1992). These were a departure from the earlier strategies as they used a fixed stimulation rate, relying on modulation of the electrical stimuli to convey temporal information such as the fundamental frequency. The SMSP strategy stimulated electrodes that corresponded to maxima in the sound spectra, while CIS stimulated all of a small number of electrodes each

cycle. The SMSP-type strategies have been named SPEAK, ACE and N-of-M strategies in commercial implementations.

There are many different parameters to set when Mapping these fixed-rate strategies for individual users, such as threshold (T) and comfort (C) stimulation levels for each electrode, number of maxima, loudness growth function, electrode frequency allocations, gain functions and rate of stimulation. It is the latter that has received the most attention and research, especially as cochlear implant hardware has been made more capable of increased rates of stimulation.

The rate of stimulation refers to the number of stimulation pulses delivered per second per electrode. Thus an SMSP-type strategy using 6 maxima with a stimulation rate of 250 pps on each channel will stimulate up to 6 electrodes 250 times per second. As these strategies use sequential, interleaved stimulation across electrodes, this example corresponds to a total rate of stimulation of $6 \times 250 = 1500$ Hz. Typical clinical rates of stimulation range from 250 pps up to 1800 pps.

Since the modulation of the envelope of the signal within each frequency band is important for understanding speech, it is natural to assume that using a higher rate of stimulation will result in higher speech perception scores. However, this has not been the case in many studies that have examined the effects of rate of stimulation. For example, using the Nucleus® 24 cochlear implant with the ACE strategy, Vandali, Whitford, Plant & Clark (2000) demonstrated that there was no significant difference in performance across research participants for CVC words and sentences-in-noise recognition between 250, 807 and 1615 pps, except for sentences at low signal-to-noise ratios (SNR) where the 1615

pps rate showed the worst performance. Using the same rates of stimulation in the Nucleus® 24 implant, Grayden & Clark (2000) demonstrated no significant difference in vowel and consonant recognition performance. Psarros, Plant, Decker, Whitford & Cowan (2002) found that, in children, there was very little difference in rate of stimulation between 250 pps and 900 pps. Friesen, Shannon & Cruz (2005) compared a wide range of rates of stimulation using the Clarion® C1, Clarion® C2 and Nucleus® 24 cochlear implants. Again it was found that there was no improvement in speech perception scores with increasing rate of stimulation.

All of the above studies have demonstrated that there is not a significant improvement in speech perception with different rates of stimulation across subjects. However, there are differences observed between individual users (Vandali *et al.*, 2000; Grayden & Clark, 2000; Psarros *et al.*, 2002; Skinner, Arndt & Staller, 2002). In particular, Skinner *et al.* (2002) found that more users subjectively preferred high-rate strategies (68%) over a low-rate strategy (23%) and that subjects generally performed better when using their preferred strategy. However, there were still a large number of users who preferred, and performed better with, a low rate of stimulation.

Information transmission analyses for phonetic features performed by Grayden & Clark (2000) found that the lack of benefit provided by higher rates may be due to different phonetic cues being better transmitted at different rates. In particular, it was apparent that high rates of stimulation (807 pps and 1615 pps) provided improved information about temporal information and frication in speech, but that a low rate (250 pps) provided more spectral detail required for the perception of place of articulation.

The results of the Grayden & Clark (2000) study suggested the application of a coding strategy where two stimulation rates are used simultaneously. A low rate of stimulation is used in the 0-2400 Hz range, which encode the first two formants and most place of articulation information. The remaining electrodes, 2400-8000 Hz, are stimulated using a high rate of stimulation to provide improved presentation of temporal information and frication. Representation of frication will benefit from the more stochastic nature of neural

firing, which results from high-rate stimulation (Clark, 1998). Temporal events, such as onset of pitch periods and plosive bursts, tend to be present across the full spectrum and so there should be information transmission benefit when presenting high rate stimulation on the high-frequency electrodes.

This paper describes the Differential-Rate Sound Processing (DRSP) strategy for speech perception by users of the Nucleus® 24 cochlear implant. The DRSP strategy uses both low and high rates of stimulation – applying a low rate of stimulation to apical, lower-frequency electrodes and a high rate of stimulation to basal, higher-frequency electrodes.

2. Method

2.1. The DRSP strategy

The Differential-Rate Sound Processing (DRSP) strategy was implemented on the Spear 3 research processor¹. A standard ACE program was modified to work at two different rates of stimulation: 250 pps and 1500 pps. It was implemented by assigning a number of electrodes to work at 250 pps (electrodes corresponding to frequencies at and below 2400 Hz) and the remainder to work at 1500 pps (electrodes above 2400 Hz). The analysis rate used was 1500 windows per second as would be the case for a standard ACE program. However, after maxima selection, the 250 pps electrodes were only stimulated if they were the first out of every six analysis windows or, conversely, five out of every six cycles were discarded for the 250 pps electrodes. Thus the rate of stimulation for the low-frequency electrodes was $1500 / 6 = 250$ pps.

The strategy was set up for each user by first Mapping the electrodes for 250 pps and for 1500 pps. The two Maps were loudness balanced for each user by alternating between them while listening to 8-speaker multi-talker babble noise; the threshold (T) and comfort (C) levels were adjusted to make the two rates sound the same loudness. A hybrid Map was created for the DRSP strategy by assigning the 250 pps T and C levels to the low frequency electrodes and the 1500 pps T and C levels to the high frequency electrodes.

Subject	Duration of CI Use	Aetiology	Number of Electrodes	Clinical Rate (Hz)	Vowels SNR (dB)	CUNY SNR (dB)
1	1 yr, 1 mo	Unknown	22	900	0	10
2	1 yr, 3 mo	Unknown	22	250	0	10
3	7 mo	Unknown	22	250	0	10
4	2 yr, 1 mo	Menieres	17	900	Quiet	Quiet
5	1 yr, 2 mo	Otosclerosis	20	250	0	10
6	3 yr, 7 mo	Meningitis	20	500	0	10
7	6 mo	Unknown	20	900	5	10
8	4 mo	Unknown	22	900	0	10
9	6 mo	Unknown	20	900	5	10
10	5 mo	Unknown	22	1200	0	10

Table 1: Details of research subjects

2.2. Research subjects

The participants in the study were ten hearing impaired subjects who had received a Nucleus® CI-24M cochlear implant². Subjects differed in their amount of experience with the cochlear implant prior to commencing the study.

Table 1 lists the details of the subjects who participated in the study. “Duration of CI Use” provides the number of years and months since implantation to the start of the study. “Number of Electrodes” provides the number of electrodes used in the Map. “Clinical Rate” gives each user’s normal rate of stimulation used in their processor. The final columns give signal-to-noise ratios (SNR) used for the vowel and sentences tests – details of these are described below.

The average age of subjects was 68 years with standard deviation of 9 years (range 57-82). The average duration of CI use was 1 year, 2 months, with standard deviation of 1 year.

2.3. Evaluation materials

The following recorded test material was used in the study:

- Consonants uttered in /aCa/ context in a closed set of 16 spoken by a female speaker and presented in quiet. In each testing session, each consonant was presented 3 times per list and 3 lists were presented giving a total of 9 presentations per consonant. Each strategy was tested in 4 sessions, so each consonant was presented 36 times for each strategy.
- Vowels uttered in /hVd/ context in a closed set of 11 spoken by a female speaker and presented in speech-shaped random noise. Background noise was used in this test to avoid ceiling effects. For each subject, the signal-to-noise ratio (SNR) was set to 15, 10, 5 or 0 dB such that they obtained a score in the range 20%-50% using their clinical strategy. There were 4 presentations of each phoneme per list and 2 lists per session giving a total of 32 presentations of each phoneme per strategy.
- Open-set CVC words uttered by a female speaker with 50 words per list and presented in quiet. These were scored by number of words, phonemes, vowels and consonants correct. There were 2 lists per session giving a total of 400 words per strategy.
- CUNY-like sentences uttered by a female speaker with 102 words per list, scored by number of words correct, presented in 8-speaker multi-talker babble. For each subject the noise level was set to quiet, 15, 10 or 5 dB SNR such that they obtained a score in the range 20%-70% using their clinical strategy, in order to avoid ceiling effects. There were 2 lists presented in each session, giving a total of 816 words per strategy.

2.4. Evaluation procedure

In the first session, subjects were programmed with a 250 pps strategy (hereafter called the ‘LOW’ strategy) and with a 1500 pps strategy (‘HIGH’) in the first session. All strategies selected 8 maxima for stimulation each cycle. For each strategy, the T and C levels were found using standard audiological procedures and the appropriate rates of stimulation on each electrode. The C levels were loudness balanced by sweeping through electrodes in groups of three

and asking the subject to comment on which were louder or softer; if one was louder or softer, the C level for that electrode was changed accordingly and the procedure repeated. The T levels were balanced after C level balancing by sweeping across groups of three electrodes and loudness matching at 50% of the dynamic range; any electrode found to be louder or softer was adjusted by changing the T level appropriately. Each strategy was also presented to the subject to determine if any overall adjustments were required to change loudness levels for speech sounds and the listener’s own voice. At the end of this first session, the strategy closest to the patient’s own (clinical) strategy was set up on the Spear 3 processor and the subject took it away with them for one week of familiarisation.

In the second session, one week later, the subjects were given practice with the testing material using their clinical strategy in their own processor and the most appropriate signal-to-noise ratio was established for the vowel and sentence tests. The results of this for each patient are shown in Table 1. The patient was then given the other strategy (LOW or HIGH) that they had not trialled on the Spear 3 processor for take-home experience.

In the third session, one week later, the LOW and HIGH strategies were loudness balanced. This was performed to reduce the possibility of increased performance for one strategy that might arise if it was louder than the others. Loudness balancing was achieved by playing multi-talker babble at 65 dB SPL from a loudspeaker while the processor was switched between LOW and HIGH at a fixed sensitivity setting. If one strategy was reported louder than the other, C levels were globally adjusted to reduce the overall volume for that strategy. The DRSP strategy was then set up using portions of LOW and HIGH as appropriate. The subject was given the first strategy of their evaluation protocol to take home for three weeks of familiarisation experience.

The design of the study was ABCCBA, where A, B, or C were LOW, HIGH or DRSP. The assignment of A, B and C to strategies (and thus the order of testing) was balanced across patients. The reversed repeat protocol was used to balance out learning effects to a first approximation. After each change of strategy (and between the two C instances) there were at least three weeks of familiarisation time allowed before undertaking two testing sessions that were spaced one week apart.

3. Results

3.1. Closed-set consonants

Figure 1 shows the results for each subject for closed-set consonant recognition. Analysis of variance (ANOVA) for the score found no significant difference for Strategy (fixed factor: $F = 0.291$, $p = 0.751$), a significant difference between Subjects (random factor: $F = 102$, $p < 0.001$) and a significant interaction between Strategy and Patient ($F = 2.54$, $p = 0.001$). For all tests and material, there were significant differences found between subjects and significant interaction effects between Subject and Strategy at $p < 0.01$, so they will no longer be reported for each speech perception test below.

ANOVA applied to each subject’s results separately found significant differences between Strategy for Subject 6 ($F = 10.3$, $p < 0.001$), Subject 7 ($F = 3.69$, $p = 0.036$) and Subject

10 ($F = 4.12$, $p = 0.025$). For these subjects, Tukey's Post Hoc pair-wise comparisons were made and found that for Subject 6 HIGH and DRSP performed significantly better than LOW, for Subject 7 HIGH performed significantly better than LOW and for Subject 10 DRSP performed significantly better than HIGH. Only significant differences will be reported for the speech perception tests, so any further statements below that compare strategies will drop the word "significantly".

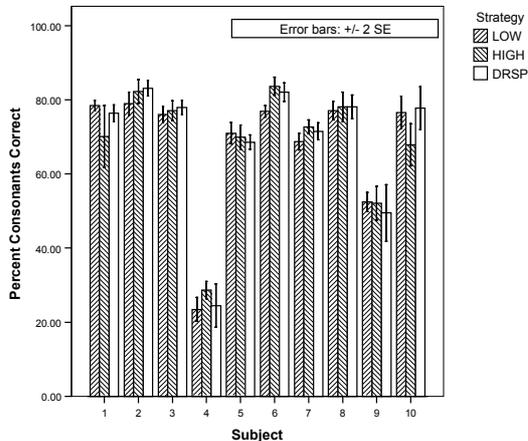


Figure 1: Closed-set consonant recognition results.

3.2. Closed-set vowels in noise

Figure 2 shows the results for each subject for closed-set vowel recognition in noise. ANOVA found no significant difference for Strategy ($F = 0.147$, $p = 0.864$). ANOVA for each subject individually found significant differences between Strategy for Subject 1 ($F = 7.15$, $p = 0.004$), Subject 2 ($F = 3.46$, $p = 0.05$) and Subject 6 ($F = 5.62$, $p = 0.011$). For these subjects, Tukey's Post Hoc pair-wise comparisons were made and found that for Subject 1 DRSP performed better than HIGH, for Subject 2 LOW performed better than HIGH and for Subject 6 HIGH performed better than DRSP and LOW.

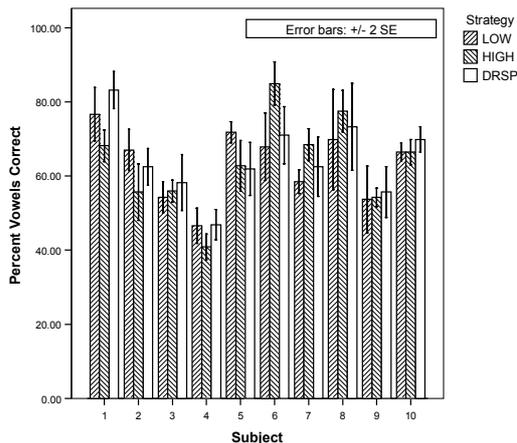


Figure 2: Closed-set vowel recognition results.

3.3. CVC words

Figure 3 shows the results for each subject for CVC word recognition. Figures 4, 5, and 6 display individual subject results for CVC phonemes, vowels and consonants, respectively. ANOVA found no significant difference for Strategy for any of these scores across subjects.

ANOVA of word score for each subject found significant differences between Strategy for Subject 1 ($F = 9.14$, $p = 0.001$), Subject 6 ($F = 3.93$, $p = 0.035$), Subject 8 ($F = 7.91$, $p = 0.003$) and Subject 9 ($F = 4.11$, $p = 0.031$). Tukey's Post Hoc pair-wise comparisons found that for Subject 1 DRSP and LOW performed better than HIGH, for Subject 6 HIGH performed better than LOW, for Subject 8 DRSP performed better than HIGH and LOW and for Subject 9 HIGH performed better than LOW.

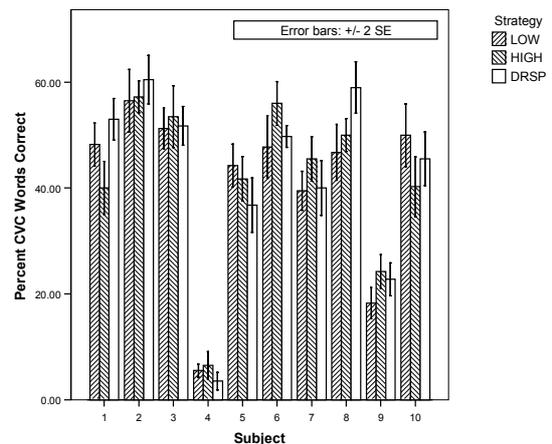


Figure 3: CVC words recognition results.

ANOVA of CVC words scored by phonemes correct for each subject found significant differences between Strategy for Subject 1 ($F = 7.84$, $p = 0.003$), Subject 5 ($F = 6.80$, $p = 0.005$) and Subject 8 ($F = 6.72$, $p = 0.006$). Tukey's Post Hoc pair-wise comparisons found that for Subject 1 DRSP performed better than HIGH, for Subject 5 LOW performed better than HIGH and DRSP and for Subject 8 DRSP performed better than HIGH and LOW.

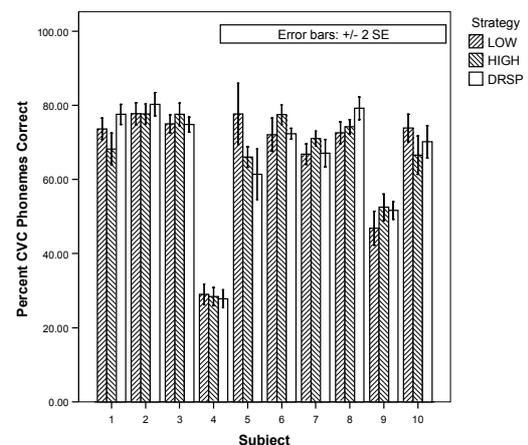


Figure 4: CVC phonemes recognition results.

ANOVA of CVC word vowel score for each subject found significant differences between Strategy for Subject 1 ($F = 5.65$, $p = 0.011$), Subject 2 ($F = 4.94$, $p = 0.017$), Subject 5 ($F = 5.95$, $p = 0.009$) and Subject 6 ($F = 4.32$, $p = 0.027$). Tukey's Post Hoc pair-wise comparisons found that for Subject 1 DRSP and LOW performed better than HIGH, for Subject 2 DRSP performed better than LOW and HIGH, for Subject 5 LOW performed better than DRSP and for Subject 6 HIGH performed better than DRSP.

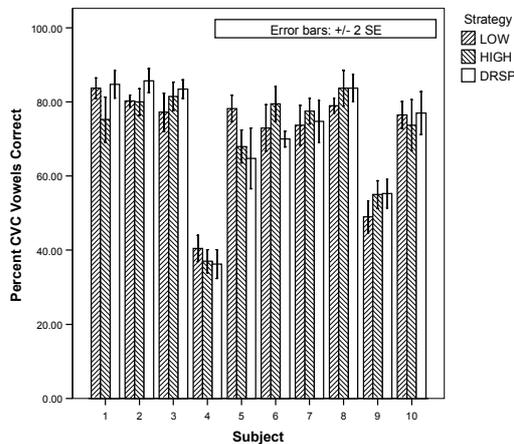


Figure 5: CVC vowels recognition results.

ANOVA of CVC word consonant score for each subject found significant differences between Strategy for Subject 1 ($F = 6.95$, $p = 0.005$), Subject 8 ($F = 7.74$, $p = 0.003$) and Subject 10 ($F = 4.41$, $p = 0.025$). Tukey's Post Hoc pair-wise comparisons found that for Subject 1 DRSP performed better than HIGH, for Subject 8 DRSP performed better than HIGH and LOW and for Subject 10 LOW performed better than HIGH.

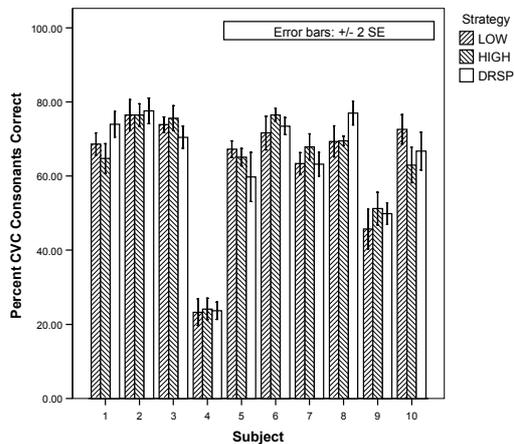


Figure 6: CVC consonants recognition results.

3.4. CUNY-like sentences in noise

Figure 7 shows the results for each subject for CUNY-like sentence recognition in noise scored by words correct. ANOVA found no significant difference for Strategy. ANOVA of word score for each subject found significant

differences between Strategy for Subject 3 ($F = 13.35$, $p < 0.001$), Subject 6 ($F = 11.58$, $p < 0.001$) and Subject 7 ($F = 3.65$, $p = 0.043$). Tukey's Post Hoc pair-wise comparisons found that for Subject 3 HIGH performed better than DRSP and LOW, for Subject 6 HIGH and DRSP performed better than LOW and for Subject 7 HIGH performed better than DRSP.

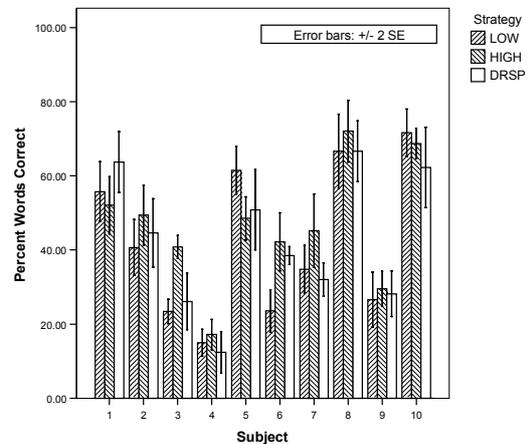


Figure 7: CUNY-like sentence word recognition results.

4. Discussion

Overall there was no difference in performance across subjects between the strategies: LOW, HIGH and DRSP. This is consistent with other studies that have found very little to separate low-rate and high-rate strategies.

However, there were differences in performance for individual results. In particular, Subjects 1 and 8 showed a clear preference for DRSP over most of the tests. There were two subjects, Subjects 2 and 10, who showed some degree of preference for DRSP and LOW over HIGH. Two subjects, Subjects 3 and 7, showed a preference for HIGH on some tests. Subject 6 also showed a preference for HIGH over LOW but only performed better than DRSP for some tests. Subject 5 showed some preference for LOW. The remaining subjects, Subjects 4 and 9, did not show much difference at all between the strategies.

Counting up the number of significant results for each subject there were 5 tests where DRSP was amongst the worst performers. HIGH was amongst the worst for 13 tests. LOW was amongst the worst for 11 tests. While this is only a very rough indication of performance, it illustrates the potential for DRSP to be a compromise strategy that could well be worth considering as an alternative to single-rate strategies.

HIGH seemed more likely to be the most effective for speech perception in noise. There was no result in this test where HIGH was the worst performer. This shows some disagreement with results in the literature. However, there was quite a wide amount of variation in sentence scores between lists for each subject, illustrated by the large error bars in Figure 7, so further testing may be necessary.

What was not borne out in this study was the anticipated improvement in vowel recognition with LOW and DRSP that would be expected for low rates in the first and second formant regions. What seems most apparent from these results

is that when there was a subject preference for a low or high rate then this was true for most of the test material. Further study with a large subject pool and increased test material may highlight the results more clearly.

5. Conclusions

This study confirmed that there is still a case for considering low-rate stimulation when Mapping patients. The Differential-Rate Sound Processing strategy demonstrates a compromise between the fixed low-rate and high-rate strategies that can perform equal to, sometimes better than, low and high rate strategies for most subjects.

6. Footnotes

¹ The SPEAR3 processor and research system was developed by the Cooperative Research Centre for Cochlear Implant and Hearing Aid Innovation. A product brief, "SPEAR3 Product Brief – 3rd Generation Speech Processor for Electrical and Acoustic Research", can be obtained from HearWorks, 384 Albert Street, East Melbourne, 3002, Australia, or online from <http://www.hearworks.com.au>.

² This study was approved by the Royal Victorian Eye and Ear Hospital Human Research Ethics Committee (Approval Number 97/318H/02). Research subjects were reimbursed their travelling expenses.

7. Acknowledgements

Thank you to the research volunteers who gave up so much of their time to be involved in this project. Thank you to Mr Andrew Vandali, Dr Robert Fearn and the CRC for Cochlear Implant and Hearing Aid Innovation who provided the Spear 3 processors and implemented the DRSP modifications.

This work was supported by The Bionic Ear Institute and the Cochlear Implant Clinic of the Royal Victorian Eye & Ear Hospital.

National ICT Australia is funded by the Australian Government's Backing Australia's Ability initiative, in part through the Australian Research Council.

8. References

- Clark, G. M. (1998). Recent advances for cochlear implants. *Auris Nasus Larynx*, 25, 73-87.
- Dowell, R. C., Seligman, P. M., Blamey, P. J. & Clark, G. M. (1987). Speech perception using a two-formant 22-electrode cochlear prosthesis in quiet and in noise. *Acta Oto-Laryngologica*, 104, 439-446.
- Dowell, R. C., Whitford, L. A., Seligman, P. M., Franz, B. K.-H. & Clark, G. M. (1990). Preliminary results with a miniature speech processor for the 22-electrode/Cochlear hearing prosthesis. In T. Sacristan (Ed.), *Otorhinolaryngology, head and neck surgery* (pp. 1167-1173). Amsterdam: Kugler and Ghedini.
- Friesen, L. M., Shannon, R. V. & Cruz, R. J. (2005). Effects of stimulation rate on speech recognition with cochlear implants. *Audiology & Neurotology*, 10, 169-184.
- Grayden, D. B. & Clark, G. M. (2000). The effect of rate of stimulation of the auditory nerve on phoneme recognition. In *Proceedings of the Eighth Australian International Conference on Speech Science & Technology, SST-2000*, Canberra, 4-7 December 2000, 356-361.
- Grayden, D. B. & Clark, G. M. (2006). Implant design and development. In H. R. Cooper & L. C. Craddock (Eds.), *Cochlear Implants: A Practical Guide* (pp. 1-20). London: Whurr Publishers Limited.
- McDermott, H. J., McKay, C. M. & Vandali, A. E. (1992). A new portable sound processor of the University of Melbourne/Nucleus Limited multielectrode cochlear implant. *Journal of the Acoustical Society of America*, 91, 3367-3371.
- Psarros, C. E., Plant, K. L., Decker, J. A., Whitford, L. A. & Cowan, R. S. C. (2002). Conversion from the SPEAK to the ACE strategy in children using the Nucleus 24 cochlear implant system: Speech perception and speech production outcomes. *Ear & Hearing*, 23, 18S-27S.
- Skinner, M. W., Arndt, P. L. & Staller, S. J. (2002). Nucleus® 24 Advanced Encoder conversion study: Performance versus preference. *Ear & Hearing*, 23, 2S-17S.
- Tong, Y. C., Black, R. C., Clark, G. M., Forster, I. C., Millar, J. B., O'Loughlin, B. J. & Patrick, J. F. (1979). A preliminary report on a multiple-channel cochlear implant operation. *Journal of Laryngology and Otology*, 93, 679-695.
- Vandali, A. E., Whitford, L. A., Plant, K. L. & Clark, G. M. (2000). Speech perception as a function of electrical stimulation rate: Using the Nucleus 24 cochlear implant system. *Ear & Hearing*, 21, 608-624.
- Wilson, B. S., Lawson, D. T., Zerbi, M. & Finley, C. C. (1992). Speech processors for auditory prostheses. *Fifth quarterly progress report, April 1992. NIH contract N01-DC-9-2401. Research Triangle Institute*.