

Trends in Hearing

Listening to Music Through Hearing Aids: Potential Lessons for Cochlear Implants

Journal:	<i>Trends in Hearing</i>
Manuscript ID	TIH-21-0115.R1
Manuscript Type:	Perspective
Date Submitted by the Author:	n/a
Complete List of Authors:	Moore, Brian; University of Cambridge, ;
Keywords:	hearing aids, cochlear implants, music, automatic gain control, dynamic range

Moore

Lessons from hearing aids

1

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Listening to Music Through Hearing Aids: Potential Lessons for Cochlear Implants

Brian C. J. Moore

Cambridge Hearing Group, Department of Psychology, University of Cambridge,

Downing Street, Cambridge CB2 3EB, England

email: bcjm@cam.ac.uk

For Peer Review

1
2
3
4 14 **Abstract**
5
6 15
7
8 16 Some of the problems experienced by users of hearing aids (HAs) when listening to music are
9
10 17 relevant to cochlear implants (CIs). One problem is related to the high peak levels (up to 120 dB
11
12 18 SPL) that occur in live music. Some HAs and CIs overload at such levels, because of the limited
13
14 19 dynamic range of the microphones and analog-to-digital converters (ADCs), leading to perceived
15
16 20 distortion. Potential solutions are to use 24-bit ADCs or to include an adjustable gain between the
17
18 21 microphones and the ADCs. A related problem is how to squeeze the wide dynamic range of music
19
20 22 into the limited dynamic range of the user, which can be only 6-20 dB for CI users. In HAs, this is
21
22 23 usually done via multi-channel amplitude compression (automatic gain control, AGC). In CIs, a
23
24 24 single-channel front-end AGC is applied to the broadband input signal or a control signal derived
25
26 25 from a running average of the broadband signal level is used to control the mapping of the channel
27
28 26 envelope magnitude to an electrical signal. This introduces several problems: (1) an intense
29
30 27 narrowband signal (e.g. a strong bass sound) reduces the level for *all* frequency components,
31
32 28 making some parts of the music harder to hear; (2) the AGC introduces cross-modulation effects
33
34 29 that can make a steady sound (e.g. sustained strings or a sung note) appear to fluctuate in level.
35
36 30 Potential solutions are to use several frequency channels to create slowly varying gain-control
37
38 31 signals and to use slow-acting (or dual time-constant) AGC rather than fast-acting AGC.
39
40 32
41
42 33 **Keywords:** music, dynamic range, automatic gain control, hearing aids, cochlear implants,
43
44 34 compression speed
45
46 35
47
48
49
50
51
52
53
54
55
56
57
58
59
60

37 Introduction

38 Users of both hearing aids (HA) and cochlear implants (CI) experience problems when
39 listening to music (McDermott, 2004; Looi, McDermott, McKay & Hickson, 2008; Chasin &
40 Hockley, 2014; Madsen & Moore, 2014; Moore, 2016). Some of these problems are related to the
41 reduced resolution and processing capacity of the impaired auditory system (Shannon, 1983;
42 Moore, 2003; Moore, 2007). For example, both HA users and CI users have reduced spectral
43 resolution (Pick, Evans & Wilson, 1977; Glasberg & Moore, 1986; Friesen, Shannon, Baskent &
44 Wang, 2001). For CI users, this reduction is very substantial because of the spread of current within
45 the cochlea, and so it is likely to severely limit the ability to “hear out” one instrument or voice in
46 the presence of another instrument or voice (Mehta & Oxenham, 2017). This in turn will limit the
47 enjoyment of any music that includes multiple instruments and voices (Limb & Roy, 2014) and this
48 limitation is unlikely to be alleviated by changes in CI processing, unless a way can be found of
49 making the electrical stimulation much more place selective. However, other problems are related
50 to the design of the HAs and CIs and these problems can potentially be reduced by improvements
51 in design. This paper considers some of the lessons that have been learned from studies of music
52 perception via HAs and describes how those lessons might be applied to improving the design of
53 CIs.

54 In both HAs and CIs, the broadband signals picked up by the microphone(s) are passed
55 through an array of bandpass filters to create channel signals corresponding to the outputs of filters
56 with different center frequencies. In HAs, the channel signals are processed in various ways before
57 being combined to create a broadband signal that is delivered to a miniature loudspeaker (called a
58 receiver) that generates the output sound. In CIs the channel signals are used to create electrical
59 signals that are delivered to the individual electrodes in the array implanted within the cochlea.
60 Hence, in CIs, the number of channels is usually equal to the number of implanted electrodes,
61 although the effective number of independent channels is smaller than the number of electrodes,
62 because of current spread in the cochlea (Friesen et al., 2001). In a CI, the signal in the channel
63 with the lowest center frequency is usually used to derive the electrical signal delivered to the most
64 apical electrode in the cochlea, while the signal in the channel with the highest center frequency is
65 usually used to drive the most basal electrode, with a continuous gradation in between. This

66 represents an attempt to recreate the tonotopic mapping of frequency to place that occurs within a
67 normal cochlea (von Békésy, 1960). Thus, information about the short-term spectrum of sounds is
68 conveyed by the relative strength of the electrical signals across the electrode array.

70 **Requirements for Music Listening With CIs**

71 It is helpful first to consider the properties that a CI should have in order to improve the
72 experience of listening to music. The key properties are:

73 (1) Envelope cues in different frequency regions, which are the main source of auditory information
74 provided by CIs (Wilson, Finley, Lawson, Wolford, Eddington & Rabinowitz, 1991; Clark, Tong &
75 Patrick, 1990; Zeng, Popper & Fay, 2003), should be preserved and coded as faithfully as possible.

76 (2) Nonlinear distortion occurring prior to filtering of the signal into frequency bands or channels
77 should be minimal. This is because nonlinear distortion leads both to changes in the shape of the
78 waveform and to frequency components in the short-term spectrum that are not present in the
79 original signal, giving a misleading representation of the spectral shape of a sound (Tan, Moore &
80 Zacharov, 2003). Nonlinear distortion introduced after the signal is filtered into frequency channels
81 in a CI does not have the same effect on the representation of the spectrum of a sound, but it can
82 distort the representation of the envelope in each channel.

83 (3) The wide range of sound levels that occur in music, especially live music, must be compressed
84 into the narrow range of electrical current values between the detection threshold and the highest
85 comfortable level. This must be done while preserving the representation of the envelope of the
86 sound in each frequency channel as well as possible.

87 As is discussed below, these properties are not easy to achieve, and the current state of the
88 art is far from optimal in this respect.

90 **Sources of Envelope Distortion**

91 In this section, I consider some (but not all) of the sources of envelope distortion in HAs
92 and CIs. Envelope distortion introduced by automatic gain control systems is discussed in a later
93 section. Most HAs and CIs incorporate various forms of adaptive signal processing, i.e. signal
94 processing that changes over time in response to changes in the input signal. Examples are noise

1
2
3
4 95 reduction and adaptive directional processing. These forms of signal processing, because they are
5
6 96 time-varying, inevitably distort the envelopes of the channel signals in HAs and CIs.

7
8 97 Noise-reduction systems can be applied to the signal from a single microphone. They have
9
10 98 been designed to improve the ability to understand speech in the presence of background noise.
11
12 99 They generally work by estimating the momentary speech-to-noise ratio in each channel and
13
14 100 applying attenuation to the channels with the poorest estimated speech-to-noise ratio (Holube,
15
16 101 Hamacher & Wesselkamp, 1999; Hamacher, Fischer, Kornagel & Puder, 2006; Launer, Zakis &
17
18 102 Moore, 2016; Yang & Fu, 2005). In more recent approaches, artificial neural networks have been
19
20 103 used to process speech in background sounds for application both to HAs (Keshavarzi, Goehring,
21
22 104 Turner & Moore, 2019; Healy, Tan, Johnson & Wang, 2021) and CIs (Goehring, Bolner,
23
24 105 Monaghan, van Dijk, Zarowski & Bleeck, 2017; Goehring, Keshavarzi, Carlyon & Moore, 2019).

25
26 106 All single-microphone noise-reduction systems involve a trade-off; the more the
27
28 107 background noise is reduced relative to the speech, the more distortion there is, including envelope
29
30 108 distortion. There appear to be large individual differences in preferences for noise-reduction
31
32 109 systems and in preferences for the amount of noise reduction, some people being “noise haters” and
33
34 110 others being “distortion haters” (Brons, Houben & Dreschler, 2012). In any case, since these noise-
35
36 111 reduction systems have been designed to improve the perception of speech in noise, they are
37
38 112 unlikely to be of any benefit when listening to music. Rather, the envelope distortion that they
39
40 113 introduce is likely to degrade the perception of music by users of HAs and CIs. Therefore, it is
41
42 114 recommended that a CI is set up with a dedicated music program, as is often done for HAs, and in
43
44 115 the music program any adaptive noise-reduction processing is disabled.

45
46 116 A second form of adaptive processing, directional processing, is used when two or more
47
48 117 microphones are available, as is the case with most HAs and CIs. Some (but not all) such systems
49
50 118 are based on the assumption that the “target” sound that the user wishes to hear comes from the
51
52 119 front. They attempt to estimate, for each frequency channel, the direction of the most prominent
53
54 120 interfering sound coming from the sides or back, and to create a null in the directional response so
55
56 121 as to attenuate that interfering sound (Launer et al., 2016). This creates time-varying changes in the
57
58 122 effective frequency response of the HA or CI for sounds coming from the sides or back, and also
59
60 123 distorts the envelope representation of such sounds. Again, this is likely to impair music perception,

1
2
3
4 124 since music often involves spatially distributed sound sources. Even for directional processing
5
6 125 systems that attempt to preserve sounds of interest from several directions, the processing is
7
8 126 adaptive and time varying, and is likely to result in the introduction of spurious amplitude
9
10 127 modulation. As for noise-reduction processing, it is recommended that a CI is set up with a
11
12 128 dedicated music program and that any adaptive directional processing is disabled for that program.
13

14 129

15 16 130 **The Problem of the Dynamic Range of the Input Signal**

17
18 131 Most CIs and some HAs use 16-bit analog-to-digital converters (ADCs) to digitize the
19
20 132 microphone signals (Launer et al., 2016; Zakis, 2016). In theory this can code a 96-dB range of
21
22 133 input sound levels (6 dB per bit). In practice the achieved range is typically 85-90 dB, because of
23
24 134 microphone noise, noise in the ADCs themselves, and noise in the electronic circuitry. Although in
25
26 135 principle the dynamic range can be increased by the application of a low-level spectrally shaped
27
28 136 noise called dither (Vanderkooy & Lipshitz, 1984), this is not to my knowledge applied in HAs or
29
30 137 CIs. In most HAs and CIs, the gain of the pre-amplifier between the microphone and the ADC is set
31
32 138 so that the lowest sound level that can be coded is about 15 dB SPL, which means that the highest
33
34 139 sound level that can be coded is 100-105 dB SPL. When listening to music in the home, peak sound
35
36 140 levels rarely exceed 95 dB SPL and most CIs and HAs can handle this without significant
37
38 141 distortion. However, when listening to live music, or amplified music in a club or discotheque, peak
39
40 142 sound levels can reach 115-120 dB SPL (Hockley, Bahlmann & Fulton, 2012; Chasin & Hockley,
41
42 143 2014). Sound at such levels can be unpleasant and may appear distorted even for people with
43
44 144 normal hearing, because of nonlinear distortion produced in the outer and middle ear (Price & Kalb,
45
46 145 1991) and because of the effects of upward spread of masking (Studebaker, Sherbecoe, McDaniel
47
48 146 & Gwaltney, 1999). However, for users of HAs and CI, such high sound levels can also lead to
49
50 147 overload (peak clipping) that reduces the perceived sound quality, at least for users of HAs (Tan &
51
52 148 Moore, 2008). Apart from introducing spectral changes in the signal, peak clipping results in a
53
54 149 distortion of the envelope cues that are important for CI users.

55
56 150 There are several solutions to this problem. One is to use 24-bit ADCs, a solution that has
57
58 151 been adopted by several manufacturers of HAs, even though it decreases battery life. Another
59
60 152 solution is to include an adjustable gain between the microphones and the ADCs; this gain

1
2
3
4 153 adjustment can be compensated for in the subsequent digital-processing stages of the HA or CI
5
6 154 (Hockley et al., 2012; Zakis, 2016). This is done in some but not all HAs and CIs. It is
7
8 155 recommended, therefore, that the range of signal levels that can be handled by CIs without
9
10 156 distortion is increased, either via the use of 24-bit ADCs or via the use of an adjustable gain
11
12 157 between the microphones and the ADCs.
13
14 158

15 159 **Squeezing Music Into the Limited Dynamic Range of the CI** 16 17 18 160 **User**

19 20 161 *The Perceptual Dynamic Range for Users of HAs and CIs*

21
22 162 Much recorded music, and music that is broadcast or transmitted via the internet, is
23
24 163 subjected to some form of amplitude compression to reduce its dynamic range (the difference
25
26 164 between the highest level and lowest level in the music) (Croghan, Arehart & Kates, 2014;
27
28 165 Hjortkjær & Walther-Hansen, 2014). This is especially true for “pop” music. However, live music,
29
30 166 especially classical music, can have a very wide range dynamic range; the peaks occurring during a
31
32 167 *fortississimo* (*fff*) passage may be 70 dB or more above the dips in level in a *pianississimo* (*ppp*)
33
34 168 passage (Hockley et al., 2012; Kirchberger & Russo, 2016). The perceptual dynamic range –
35
36 169 defined here as the range between the detection threshold and the level at which sound starts to
37
38 170 become uncomfortably loud – is typically at least 95-100 dB for people with normal hearing
39
40 171 (Moore, 2012). This is enough to allow the *pianississimo* passages to be heard without the
41
42 172 *fortississimo* passages being uncomfortably loud, except perhaps when a *pianississimo* passage
43
44 173 occurs immediately after a *fortississimo* passage. However, hearing loss usually results in loudness
45
46 174 recruitment (a more rapid than normal growth of loudness with increasing sound level, once the
47
48 175 sound level exceeds the elevated detection threshold) and a reduced perceptual dynamic range
49
50 176 (Fowler, 1936; Steinberg & Gardner, 1937). For a hearing loss of, say, 70 dB, the perceptual
51
52 177 dynamic range is typically only about 30 dB (Miskolczy-Fodor, 1960; Moore, 2004; Moore &
53
54 178 Glasberg, 2004). For electrical pulses delivered directly to an individual electrode in a CI, the
55
56 179 perceptual dynamic range is typically only 6-20 dB (Fourcin et al., 1979; Zeng, 2004). Hence some
57
58 180 form of amplitude compression is essential to squeeze the wide range of sound levels encountered
59
60 181 during performances of live music into the limited perceptual dynamic range of users of HAs and

Moore

Lessons from hearing aids

8

182 CIs.

183

184 *Amplitude Compression in HAs*

185 In HAs, amplitude compression is usually achieved via multi-channel automatic gain
186 control (AGC). The digitized microphone signal is filtered into several frequency channels
187 (typically between 3 and 24) and independent AGC is applied in each channel. A basic
188 characteristic of an AGC system is its input-output function, a plot of the output level as a function
189 of the input level, both in dB. For many AGC systems, the gain (output level minus input level) is
190 independent of the input level for low input levels. Hence, the input-output function has a slope of
191 one. This is called linear amplification. At higher input levels, the gain decreases progressively with
192 increasing input level, and the input-output function has a slope of less than one. The compression
193 threshold is defined as the input level at which the gain is reduced by 2 dB, relative to the gain
194 applied in the region of linear amplification (ANSI, 2014). One reason for having a compression
195 threshold is that it is impractical to continue to increase the gain indefinitely as the input level
196 decreases because this would make microphone noise or low-level environmental noises intrusive.
197 Indeed, for very low input levels, the gain in many HAs is reduced to prevent such noises from
198 being audible; this is called expansion, and it is associated with a slope of the input-output function
199 greater than one.

200 Over the range of input levels where compression is applied, the “amount” of compression
201 is specified by the compression ratio, the change in input level (in dB) required to achieve a 1-dB
202 change in output level (for an input level above the compression threshold). For example, a
203 compression ratio of 3 means that each 3-dB increase in input level leads to a 1-dB increase in
204 output level, so the input-output function has a slope of 1/3. In most HAs the compression ratio is
205 set independently for each channel depending on the perceptual dynamic range at the center
206 frequency of that channel, which is often estimated from the audiometric threshold (Keidser,
207 Dillon, Carter & O'Brien, 2012; Moore, Glasberg & Stone, 2010).

208 AGC systems vary in how rapidly they respond to a sudden change in the input sound level.
209 The speed of the AGC systems in HAs is typically measured using a sound whose input level
210 changes abruptly between 55 and 90 dB SPL (ANSI, 2014). When the sound level increases, the

gain decreases. The time taken for the output to get within 3 dB of its steady value is called the attack time. When the sound level decreases, the gain increases. The time taken for the output to increase to within 4 dB of its steady value is called the recovery time or release time. For HAs, there is no general consensus about what compression speed is “best” (Moore, 2008). Some manufacturers use fast-acting compression, typically with attack times in the range 0.5-20 ms and release times in the range 5-150 ms. Some manufacturers use slow-acting compression, typically with attack times in the range 20-2000 ms and release times in the range 500-5000 ms or more. The attack time is usually chosen to be smaller than the release time to reduce the experience of uncomfortable loudness when sudden increases in sound level occur. Some manufacturers use systems with two or more speeds (Moore & Glasberg, 1988; Moore, Glasberg & Stone, 1991; Stöbich, Zierhofer & Hochmair, 1999; Stone, Moore, Alcántara & Glasberg, 1999; Nordqvist & Leijon, 2004). In these systems, the gain is mostly determined by a slow-acting AGC system, but a fast-acting system takes over and rapidly reduces the gain if there is a sudden increase in sound level. If the increase in sound level is brief (for example the sound of a door slamming or a knife being dropped on a plate), the gain returns to the value set by the slow-acting system.

There have been several studies of preferences for compression speed for HA users when listening to music. Hansen (2002) used a simulated 15-channel hearing aid and a paired-comparison procedure. The gain for medium input sound levels was adjusted for each subject using the NAL-R fitting rule (Byrne & Dillon, 1986), but the compression ratio was fixed at 2 for all channels and subjects, rather than being tailored to suit the individual hearing losses. The stimuli included classical and pop music presented at an input level of 80 dB SPL. The longest release time used (4000 ms) was significantly preferred over the two shorter release times used (400 and 40 ms).

Moore, Füllgrabe and Stone (2011) used a paired-comparison procedure and a simulated five-channel HA to compare preferences for slow versus fast compression, using both speech and music stimuli and using three mean levels for the input signal, 50, 65 and 80 dB SPL. The simulated HA was programmed to suit the individual hearing losses, which were mild, using the CAM2 procedure (Moore et al., 2010). They found slight mean preferences for slow compression for both speech and music, but the effect was only clear for an input level of 80 dB SPL, and not for input levels of 50 and 65 dB SPL. They also assessed the effect of slightly delaying the audio signal

relative to the gain-control signal (by 2.55 to 15 ms). This delay, referred to as the “alignment delay”, allowed the gain to be reduced just *before* any sudden increase in sound level (Robinson & Huntington, 1973; Verschuure et al., 1993), thus increasing the effectiveness of the AGC system in preventing loudness discomfort. The effect of the alignment delay was small except for a highly percussive sound (solo xylophone) combined with the fast compression system, for which alignment delays of 2.5, 5, and 7.5 ms were preferred over no delay.

Croghan et al. (2014) used a paired-comparison procedure and a simulated HA with either 3 or 18 channels to compare preferences for slow versus fast compression. They used both rock and classical music, and the stimuli were subjected to various amounts of compression *prior to* the simulated HA processing, to simulate the compression that is often applied to recorded music. The stimuli were presented at a nominal level of 65 dB SPL prior to the simulated HA processing. The HAs were individually fitted using the NAL-NL1 prescription method (Byrne, Dillon, Ching, Katsch & Keidser, 2001). For stimuli that were not compressed prior to HA processing, slow compression was preferred over fast compression for both the 3- and 18-channel systems and for both classical and rock music.

Moore and Sęk (2016b) used a paired-comparison procedure and a simulated five-channel HA to compare preferences for slow versus fast compression. Three mean levels were used for the input signal, 50, 65 and 80 dB SPL. They tested a hypothesis put forward by Moore (2008) that reduced sensitivity to the temporal fine structure (TFS) of sounds may be associated with a preference for slow compression. This hypothesis is based on the observation that high sensitivity to TFS is associated with a better ability to understand speech at low speech-to-background ratios (Füllgrabe, Moore & Stone, 2015), perhaps because TFS information is useful for the perceptual segregation of the target and background sounds (Lunner, Hietkamp, Andersen, Hopkins & Moore, 2012). Also, high sensitivity to TFS may be associated with less reliance on envelope cues in different frequency channels and with more resistance to the envelope distortion produced by fast-acting compression (Stone & Moore, 1992; 2004; Souza, 2002; Stone, Moore, Füllgrabe & Hinton, 2009). On the other hand, people with poor sensitivity to TFS, including users of CIs, may rely mainly on envelope cues in different frequency channels, and for them it is important to avoid the envelope distortion that is introduced by fast compression. Hence, slow compression may be the

1
2
3
4 269 preferred option for HA users and CI users with poor sensitivity to TFS.

5
6 270 Moore and Søk (2016b) tested subjects with moderate-to-severe sensorineural hearing loss
7
8 271 and assessed their sensitivity to TFS indirectly by measuring difference limens for frequency
9
10 272 (DLFs) for a pure tone signal centered at 2000 Hz, based on the (not universally accepted)
11
12 273 assumption that frequency discrimination depends on the use of TFS information for frequencies up
13
14 274 to a few kHz (Moore, 1973; Heinz, Colburn & Carney, 2001a; Heinz, Colburn & Carney, 2001b;
15
16 275 Moore & Ernst, 2012). The simulated HA was programmed to suit the individual hearing losses
17
18 276 using the CAM2 prescription method (Moore et al., 2010), modified slightly as described in Moore
19
20 277 and Søk (2013) and Moore and Søk (2016a). Only one ear of each subject was tested.

21
22 278 For all three input levels used, there were small overall preferences for slow over fast
23
24 279 compression. Figure 1 shows the individual and mean preferences averaged across the three levels
25
26 280 for listening to music. Preference scores above 0 indicate a preference for slow compression. Nine
27
28 281 subjects showed a preference for slow compression of 0.5 scale units or more (the scale went from
29
30 282 -3 to +3, with 0 indicating no preference), and the rest showed very small preferences for slow
31
32 283 compression or no clear preference. Thus, at least for the music stimuli used in the study (jazz
33
34 284 music, classical music and a man singing with a guitar), slow-acting compression seems to be a
35
36 285 “safe” option. The logarithms of the DLFs for the test ears were weakly but significantly correlated
37
38 286 with preference scores for music, based on a one-tailed test: $r = 0.39, p < 0.05$. This provides some
39
40 287 weak support for the hypothesis that poorer sensitivity to TFS (indicated by larger DLFs) is
41
42 288 associated with stronger preferences for slow compression.

43
44 289 In summary, studies of hearing-impaired subjects using simulated HAs suggest that slow
45
46 290 compression is preferable to fast compression for listening to music. However, it should be
47
48 291 remembered that while the studies described above used a reasonably wide range of input levels
49
50 292 (50-80 dB SPL), the range was not as wide as might be encountered when listening to live music.
51
52 293 Also, the stimuli did not include conditions where the level changed abruptly from a high to a low
53
54 294 value, as might occur in music. When the level decreases abruptly, the gain in a slow-acting
55
56 295 compression system takes some time to increase; for example, the recovery time of the slow-acting
57
58 296 compression used by Moore and Søk (2016b) was 3000 ms. For users of HAs the subjective effect
59
60 297 of an abrupt decrease in level is that the HA has become “dead” for a second or two, which is

298 clearly undesirable.

299

300 *Amplitude Compression in CIs*

301 Because of the very small perceptual dynamic range for electrical signals applied to the
302 auditory nerve, all CIs incorporate some form of amplitude compression. Figure 2 is a simplified
303 schematic of the signal processing that is performed in many CIs. The analog microphone signal is
304 subjected to analog-to-digital conversion (ADC) (often after pre-amplification and filtering to
305 reduce the influence of strong low-frequency components) and the broadband digital signal is
306 subjected to AGC. The amplitude-compressed signal is then filtered into frequency channels using
307 an array of bandpass filters (BPF, 1 to n), the channel signals are rectified and lowpass filtered
308 (LPF) to extract the envelope, the envelope signals are subjected to instantaneous compression
309 (often referred to as mapping), and the compressed envelope is used to modulate the amplitude or
310 width of the biphasic current pulses delivered to each electrode. Thus, there are two stages of
311 amplitude compression, produced by the front-end AGC and by the mapping for the individual
312 channels. The CIs produced by three of the main manufacturers, Cochlear Corporation, Med-El,
313 and Advanced Bionics, all have this general structure, although they differ in the speed of the front-
314 end AGC, the number of channels, the sharpness of the filters used to create the channel signals,
315 and the method of envelope extraction (Vaerenberg, Govaerts, Stainsby, Nopp, Gault & Gnansia,
316 2014). The early devices made by Cochlear Corporation have a front-end slow-acting “automatic
317 sensitivity control” (Seligman & Whitford, 1995) followed by a fast-acting AGC system, with a
318 very high compression ratio (called a compression limiter) (Khing, Swanson & Ambikairajah,
319 2013). The current CIs manufactured by Med-El, Advanced Bionics, and Cochlear Corporation
320 incorporate a dual time-constant system (Stöbich et al., 1999; Boyle, Buchner, Stone, Lenarz &
321 Moore, 2009; Khing et al., 2013) similar to that described by Stone et al. (1999). The CIs
322 manufactured by Cochlear Corporation also include the option of a multi-channel slow-acting AGC
323 system called adaptive dynamic range optimization (ADRO, Blamey, 2005) which is applied
324 separately to the envelope signal in each channel (Wolfe, Schafer, John & Hudson, 2011; Khing et
325 al., 2013). To my knowledge, there are no published evaluations of the ADRO system for music
326 listening.

1
2
3
4 327 When listening to music, there are several problems associated with the use of single-
5
6 328 channel AGC applied to the broadband signal:
7
8 329 (1) The input-output functions of the front-end AGCs used in CIs have been designed to work well
9
10 330 for the typical speech signals encountered in everyday life but may not be optimal for music. For
11
12 331 example, the Advanced Bionics system uses a compression threshold of about 63 dB SPL
13
14 332 combined with a high compression ratio (about 12). This means that many music signals would be
15
16 333 subjected to strong (albeit slow-acting) amplitude compression, reducing the impression of dynamic
17
18 334 changes in music, such as the contrast between a *forte* passage and a *piano* passage. In comparison,
19
20 335 the multi-channel AGC systems used in HAs typically have compression thresholds in the range 20
21
22 336 to 45 dB SPL and compression ratios in the range 1.5 to 3 (Moore, Stone & Alcántara, 2001).
23
24 337 (2) A relatively intense narrowband signal (e.g. a strong bass sound) will reduce the level for *all*
25
26 338 frequency components, making some parts of the music harder to hear.
27
28 339 (3) The AGC introduces cross-modulation effects (Stone & Moore, 2004; Stone & Moore, 2008): a
29
30 340 brief strong signal from one source (e.g. a drum beat) leads to a reduction in gain for another sound
31
32 341 source (e.g. strings). This can make a steady sound (e.g. sustained strings or a sung note) appear to
33
34 342 fluctuate in level and it may make it harder to segregate sound sources. Such effects are strongest
35
36 343 for fast-acting compression, but they occur to some extent even when slow compression is used,
37
38 344 unless it is very slow indeed.

39
40 345 Improvements might be produced by:

41
42 346 (1) Changing the input-output function of the front-end AGC. There are very few systematic studies
43
44 347 exploring the effect on music perception of varying the compression threshold or the compression
45
46 348 ratio of the front-end AGC in a CI. Gilbert, Deroche, Jiradejvong, Chan Barrett and Limb (2021)
47
48 349 explored the effect of varying the compression ratio of the front-end AGC in the Med-El CI,
49
50 350 keeping the compression threshold fixed at 48 dB SPL, and found no significant effect on music
51
52 351 preferences. However, they varied the compression ratio only over a small range (2.5 to 3.5) and
53
54 352 they used samples of music with a single fixed input level of 65 dBA. They did not explore the
55
56 353 effect of simultaneously varying the compression threshold and the compression ratio. It seems
57
58 354 plausible that a lower compression threshold and a lower compression ratio than currently used
59
60 355 would give a better impression of the dynamic changes in a piece of music.

356 (2) Using a multi-channel rather than single-channel front-end AGC system. This would prevent an
357 intense narrowband signal from reducing the level of all frequency components, and it would also,
358 to some extent, reduce the cross-modulation effects described above.

359 (3) For dual time-constant AGC systems, making the slow time constants even slower. However,
360 experimental studies are required to assess whether or not this is beneficial and to determine the
361 optimal compression speed, if there is such a thing.

362 As noted above, in addition to the front-end AGC, the CIs of Cochlear Corporation, Med-El,
363 and Advanced Bionics include a stage of instantaneous compression in the transformation from the
364 envelope magnitude in a given channel to the electrical signal applied to the electrode that is driven
365 by that channel. Figure 3 is a schematic illustration of the mapping system used in each channel of
366 an Advanced Bionics CI. The x-axis shows the input sound level. The y-axis shows the magnitude
367 of the electrical signal delivered to the electrode in current units (CU). In this example, the
368 threshold (T) level for that channel is 20 CU and the most comfortable level (M) is 200 CU. The
369 input dynamic range (IDR) is the range of levels that the sound processor codes into electrical
370 stimulation current in the range between the T level and the M level (20 to 200 CU) for each
371 channel. The default IDR is 60 dB, but the IDR can be adjusted via a “sensitivity” control. Figure 3
372 shows schematic input-output functions for IDRs of 50, 65, and 80 dB. Levels above the top of the
373 IDR are mapped to a very small range of CUs from the M level (200 CU) to an upper limit of about
374 220 CU (depending on the electrical pulse width and rate). Levels below the bottom of the IDR are
375 mapped to an inaudible current, below the T level. The mapping for input levels above the
376 compression threshold of 63 dB SPL depends on the time-varying gain of the front-end AGC.

377 In Figure 3, the input-output functions are straight lines when plotted as CU in linear units
378 against effective level in dB. This represents a compressive form of mapping. The shapes of the
379 input output-functions are loosely based on studies of the relationship between electrical signals
380 applied to a single electrode and loudness (Zeng & Shannon, 1995; Zeng, Galvin & Zhang, 1998).
381 The idea is that loudness fluctuations over time conveyed by changes in the electrical signal applied
382 to a given electrode should resemble the loudness fluctuations that would occur in a normal ear in
383 response to fluctuations in sound level in a local frequency region. In practice, however, the shapes
384 of the functions relating loudness to the electrical signal applied to an electrode can vary from one

1
2
3
4 385 CI user to another and can vary across electrodes within a CI user. Also, for input signals with
5
6 386 levels above the compression threshold (63 dB SPL in this example), which would occur often for
7
8 387 live music, much stronger compression is applied. Thus, the mapping process inevitably results in
9
10 388 some distortion of the internal representation of the envelope for CI users. Finally, the shapes of the
11
12 389 functions mapping envelope magnitude to electrical signal vary across manufacturers (Vaerenberg
13
14 390 et al., 2014), and it is not clear what form of mapping gives the most faithful internal representation
15
16 391 of the envelope in each channel.

17
18 392 One manufacturer of CIs, Oticon Medical, uses a different system, which is illustrated in
19
20 393 Figure 4, adapted from Langner, Buchner and Nogueira (2020). This system does not use a front-
21
22 394 end AGC. Rather all amplitude compression is applied in the mapping from envelope magnitude in
23
24 395 a given channel to the current applied to the corresponding electrode (the electrical pulse width is
25
26 396 modulated). This system is called “Voice Guard” since it was designed to convey the envelope
27
28 397 fluctuations of speech signals as faithfully as possible. The mapping from envelope amplitude to
29
30 398 electrical pulse width is compressive and it is also adaptive. The root-mean-square (RMS) level of
31
32 399 the input signal is averaged over the previous 2 s. The estimate of the RMS level is updated every 2
33
34 400 ms. This slowly-changing RMS level estimate is used to control the “knee-point” of the mapping
35
36 401 function (similar to the compression threshold) in 3-dB steps, as illustrated in Figure 5. This is done
37
38 402 separately for each of four frequency ranges. For a “Low” RMS input level (60 dB SPL), the
39
40 403 kneepoint is set to its lowest value (ranging from 52 dB SPL for low center frequencies to 41 dB
41
42 404 SPL for high center frequencies). For a “High” RMS input level (80 dB SPL), the kneepoint is set
43
44 405 to its highest value (ranging from 70 dB SPL for low center frequencies to 58 dB SPL for high
45
46 406 center frequencies). The system incorporates a “hysteresis” function that prevents the knee-point
47
48 407 from fluctuating wildly when the input level is halfway between the values corresponding to two
49
50 408 knee-points. The output electrical signal at the kneepoint is always set to 75% of the electrical
51
52 409 dynamic range between the threshold (T) and the highest comfortable level (C) for each channel.

53
54 410 As noted above, the Voice Guard system was designed with the intention of improving the
55
56 411 perception of speech, especially speech in background noise. As far as I know, it has not been
57
58 412 evaluated or optimized for music listening. However, since the mapping in the individual channels
59
60 413 is controlled by the slowly-changing RMS level of the broadband signal, the Voice Guard system is

1
2
3
4 414 likely to suffer from some of the same problems as for the front-end AGCs used in other CI
5
6 415 systems. Specifically, if the RMS level of the broadband signal increases, the electrical output
7
8 416 decreases for all channels, regardless of whether the increase in RMS level was produced by a
9
10 417 narrowband signal (e.g. a strong low-frequency note) or by a broadband signal (e.g. a crescendo in
11
12 418 the music). Thus the system will suffer from cross-modulation effects. Also, in the Voice Guard
13
14 419 system all of the compression of dynamic range is performed in the instantaneous mapping from
15
16 420 channel envelope magnitude to electrical output. The research reviewed above for HAs suggest that
17
18 421 for music listening slow compression is preferred over fast compression. Therefore, performing all
19
20 422 of the compression via instantaneous mapping is unlikely to be optimal for music listening. Finally,
21
22 423 because all of the compression is performed via the mapping, very strong compression has to be
23
24 424 used in the mapping. For example, for a high RMS input level of 80 dB SPL, the kneepoint for a
25
26 425 channel channel centered close to 2 kHz would be 66 dB. This means that all levels above 66 dB in
27
28 426 that channel would be mapped into only 25% of the electrical dynamic range. Potentially, this could
29
30 427 reduce the detectability of small envelope fluctuations, although the ability to detect small changes
31
32 428 in electric current is usually best towards the upper end of the electrical dynamic range (Shannon,
33
34 429 1983; McKay, Henshall, Farrell & McDermott, 2003).

35
36 430 Improvements in the Voice Guard system for music listening might be produced by:

- 37
38 431 (1) Controlling the mapping in individual channels by estimating the running RMS level in
39
40 432 different frequency regions and using that to control the mapping in sub-groups of electrodes, rather
41
42 433 than by controlling the mapping in all channels by the running RMS level of the broadband signal.
43
44 434 For example, the running RMS level at low frequencies would be used to control the mapping for
45
46 435 the most apical electrodes.
- 47
48 436 (2) Incorporating a multi-channel slow-acting or dual-time-constant AGC system prior to the
49
50 437 instantaneous channel mapping.

51
52 438
53
54 439 *Approaches Based on Loudness Models*

55
56 440 Although not implemented in current CI devices, some researchers have explored schemes
57
58 441 for CI processing based on loudness models, with aim of restoring the perception of loudness
59
60 442 approximately to “normal” for a wide range of input levels and spectra (McKay et al., 2003;

1
2
3
4 443 McDermott, McKay, Richardson & Henshall, 2003; Francart & McDermott, 2012); for a review,
5
6 444 see Wouters, McDermott and Francart (2015). Such schemes are based on loudness models for
7
8 445 normal hearing and use the concept of specific loudness, which is the loudness density as a function
9
10 446 of center frequency (Moore, 2014). The aim is to restore the specific loudness pattern to “normal”.
11
12 447 The loudness-based schemes have shown promising results for speech in quiet and in noise and for
13
14 448 artificial test signals, but they have not, to my knowledge, been comprehensively evaluated using
15
16 449 music.

17
18 450

19 451 **Issues Associated with Binaural CIs**

22 452 Many people with bilateral severe or profound hearing loss are fitted with bilateral CIs. The
23
24 453 use of bilateral CIs can improve the ability to understand speech in the presence of background
25
26 454 sounds and can also improve the ability to localize sounds (Dunn, Tyler, Oakley, Gantz & Noble,
27
28 455 2008; Litovsky, Parkinson & Arcaroli, 2009; Culling, Jelfs, Talbert, Grange & Backhouse, 2012).
29
30 456 The improvement in localization appears to depend largely on the use of interaural level difference
31
32 457 (ILD) cues (Seeber & Fastl, 2008; Litovsky et al., 2009). Currently, the slowly-adapting AGC
33
34 458 systems that are used in CIs operate independently at the two ears. The use of independent AGC at
35
36 459 the two ears distorts interaural level difference (ILD) cues for sound localization (Wiggins &
37
38 460 Seeber, 2011), especially when the head is moved (Archer-Boyd & Carlyon, 2019; Archer-Boyd &
39
40 461 Carlyon, 2021). The magnitude of the effect of head movement varies in a complex way with the
41
42 462 attack and release time of the AGC system and with the speed and type of movement of the head of
43
44 463 the CI user. This disrupts ILD cues for sound localization, especially dynamic cues.

46 464 In principle, the distortion of dynamic ILD cues can be reduced by linking the AGC across
47
48 465 ears, as is done in some HAs. When the AGC is linked, the same gain-control signal is applied to
49
50 466 the signal at each ear, preserving ILD cues. The evidence for benefits of linked multi-channel AGC
51
52 467 in HAs is mixed, and when slow-acting AGC is used the benefits appear to be minimal (Wiggins &
53
54 468 Seeber, 2013; Moore, Kolarik, Stone & Lee, 2016). One study has reported benefits of
55
56 469 synchronization of the front-end AGC in experimental versions of bilaterally fitted Advanced
57
58 470 Bionics CIs for the localization of both static and moving sound sources (Dwyer, Chen, Hehrmann,
59
60 471 Dwyer & Gifford, 2021). However, in that study the participants did not move their heads. Also, it

1
2
3
4 472 is not clear whether the benefits found were a result of better preservation of ILD cues because of
5
6 473 the synchronization, or whether they resulted from the fact that synchronization of AGC systems
7
8 474 generally results in less overall compression and more slowly changing compression. Finally, it
9
10 475 should be noted that linking of AGC systems distorts the trajectory of the changes in level at each
11
12 476 ear as the user moves their head (Archer-Boyd & Carlyon, 2021); in other words, monaural
13
14 477 envelope cues are distorted. Clearly, more research is needed into the potential benefits of linking
15
16 478 the AGC systems in bilaterally fitted CIs, especially in the context of music perception.
17
18 479

480 **Conclusions**

481 This paper has reviewed the AGC systems used in HAs and has considered how some of the
482 lessons learned from studies of the AGC systems in HAs might be applied to improving music
483 listening for users of CIs. The following recommendations are made:

- 484 1) A CI should be set up with a dedicated music program and any adaptive directional processing
485 and noise reduction should be disabled for that program.
 - 486 2) The dynamic range of the input stage of CIs should be increased, either by the use of ADCs with
487 more resolution than 16 bits or via the use of a controllable gain applied to the microphone
488 signal(s) before analog-to-digital conversion.
 - 489 3) There is a need for systematic studies exploring the effect on music perception of varying the
490 compression threshold and the compression ratio of the front-end AGC used in the CIs of several
491 manufacturers. It seems plausible that a lower compression threshold and a lower compression
492 ratio than currently used would give a better impression of the dynamic changes in a piece of
493 music, but this remains to be assessed, using both recorded and live music.
 - 494 4) Experimental studies comparing multi-channel and single-channel front-end AGC systems for
495 music listening should be conducted. In principle the use of multi-channel AGC would prevent
496 an intense narrowband signal from reducing the effective level of all frequency components, and
497 it would also, to some extent, reduce cross-modulation effects between different sound sources.
 - 498 5) For dual time-constant AGC systems, the effect of making the slow time constants even slower
499 should be explored.
- 50
51
52
53
54
55
56
57
58
59
60

- 1
2
3
4 500 6) For the system in which the running RMS level of input signal is used to control the kneepoint
5
6 501 of the mapping function, the potential benefits should be explored of controlling the mapping in
7
8 502 individual channels by estimating the running RMS level in different frequency regions and
9
10 503 using that to control the mapping in sub-groups of electrodes, rather than by controlling the
11
12 504 mapping in all channels by the running RMS level of the broadband signal. As for multi-channel
13
14 505 AGC, this should prevent an intense narrowband signal from reducing the effective level of all
15
16 506 frequency components, and it should also, to some extent, reduce cross-modulation effects
17
18 507 between different sound sources.
- 19
20 508 7) For the the system in which the running RMS level of input signal is used to control the
21
22 509 kneepoint of the mapping function, the potential benefits should be explored of incorporating a
23
24 510 multi-channel slow-acting or dual-time-constant AGC system prior to the instantaneous channel
25
26 511 mapping. This would reduce the amount of instantaneous compression required for the mapping
27
28 512 in each channel, helping to preserve short-term intensity contrasts and amplitude-modulation
29
30 513 patterns.
- 31
32 514 8) Research is needed into the potential benefits of linking the AGC systems in bilaterally fitted
33
34 515 CIs, especially in the context of music perception.
- 35
36 516 9) Users of bilateral CIs should be cautioned that rapid head movements might affect the apparent
37
38 517 positions of sounds in space. They should be advised to move their heads only slowly when
39
40 518 listening to music.

41
42 519

43 520 **Acknowledgments**

44 521 I thank Paddy Boyle, Alan Archer-Boyd, Andrew Oxenham, and two reviewers for helpful
45
46 522 comments on an earlier version of this paper.

47
48 523

49 524 **Declaration of Conflicting Interests**

50 525 The author has no conflicting interests to declare.

51
52 526

53 527 **Funding**

54 528 The author disclosed receipt of the following financial support for the research, authorship,

529 and/or publication of this article: This work was supported by the Medical Research Council (grant
530 G0701870).

531

532 **References**

533

534 ANSI (2014). *ANSI S3.22-2014, Specification of hearing aid characteristics*. New York: American
535 National Standards Institute.

536 Archer-Boyd, A. & Carlyon, R. P. (2021). Further simulations of the effect of cochlear-implant pre-
537 processing and head movement on interaural level differences. *The Journal of the Acoustical*
538 *Society of America*, 150, 506-525. doi: 10.1121/10.0005647

539 Archer-Boyd, A. W. & Carlyon, R. P. (2019). Simulations of the effect of unlinked cochlear-
540 implant automatic gain control and head movement on interaural level differences. *The Journal*
541 *of the Acoustical Society of America*, 145, 1389. doi: 10.1121/1.5093623

542 Blamey, P. J. (2005). Adaptive dynamic range optimization (ADRO): a digital amplification
543 strategy for hearing aids and cochlear implants. *Trends in Amplification*, 9, 77-98. doi:
544 10.1177/108471380500900203

545 Boyle, P. J., Buchner, A., Stone, M. A., Lenarz, T. & Moore, B. C. J. (2009). Comparison of dual-
546 time-constant and fast-acting automatic gain control (AGC) systems in cochlear implants.
547 *International Journal of Audiology*, 48, 211-21. doi: 10.1080/14992020802581982

548 Brons, I., Houben, R. & Dreschler, W. A. (2012). Perceptual effects of noise reduction by time-
549 frequency masking of noisy speech. *The Journal of the Acoustical Society of America*, 132,
550 2690-2699. doi: 10.1121/1.4747006

551 Byrne, D. & Dillon, H. (1986). The National Acoustic Laboratories' (NAL) new procedure for
552 selecting the gain and frequency response of a hearing aid. *Ear and Hearing*, 7, 257-265. doi:
553 10.1097/00003446-198608000-00007

554 Byrne, D., Dillon, H., Ching, T., Katsch, R. & Keidser, G. (2001). NAL-NL1 procedure for fitting
555 nonlinear hearing aids: characteristics and comparisons with other procedures. *Journal of the*
556 *American Academy of Audiology*, 12, 37-51.

557 Chasin, M. & Hockley, N. S. (2014). Some characteristics of amplified music through hearing aids.
558 *Hearing Research*, 308, 2-12. doi: 10.1016/j.heares.2013.07.003

559 Clark, G. M., Tong, Y. C. & Patrick, J. F. (1990). *Cochlear Prostheses*. Edinburgh: Churchill
560 Livingston.

- 1
2
3 561 Croghan, N. B., Arehart, K. H. & Kates, J. M. (2014). Music preferences with hearing aids: effects
4 562 of signal properties, compression settings, and listener characteristics. *Ear and Hearing*, 35,
5 563 e170-184. doi: 10.1097/AUD.0000000000000056
6
7
8 564 Culling, J. F., Jelfs, S., Talbert, A., Grange, J. A. & Backhouse, S. S. (2012). The benefit of
9 565 bilateral versus unilateral cochlear implantation to speech intelligibility in noise. *Ear and*
10 566 *Hearing*, 33, 673-682. doi: 10.1097/AUD.0b013e3182587356
11
12
13 567 Dunn, C. C., Tyler, R. S., Oakley, S., Gantz, B. J. & Noble, W. (2008). Comparison of speech
14 568 recognition and localization performance in bilateral and unilateral cochlear implant users
15 569 matched on duration of deafness and age at implantation. *Ear and Hearing*, 29, 352-359. doi:
16 570 10.1097/AUD.0b013e318167b870
17
18
19 571 Dwyer, R. T., Chen, C., Hehrmann, P., Dwyer, N. C. & Gifford, R. H. (2021). Synchronized
20 572 automatic gain control in bilateral cochlear implant recipients yields significant benefit in static
21 573 and dynamic listening conditions. *Trends in Hearing*, 25, 23312165211014139. doi:
22 574 10.1177/23312165211014139
23
24
25 575 Fourcin, A. J., Rosen, S. M., Moore, B. C. J. et al. (1979). External electrical stimulation of the
26 576 cochlea: clinical, psychophysical, speech-perceptual and histological findings. *British Journal*
27 577 *of Audiology*, 13, 85-107. doi: 10.3109/03005367909078883
28
29
30 578 Fowler, E. P. (1936). A method for the early detection of otosclerosis. *Archives of Otolaryngology*,
31 579 24, 731-741.
32
33
34 580 Francart, T. & McDermott, H. (2012). Speech perception and localisation with SCORE bimodal: a
35 581 loudness normalisation strategy for combined cochlear implant and hearing aid stimulation.
36 582 *PLoS One*, 7, e45385. doi: 10.1371/journal.pone.0045385
37
38
39 583 Friesen, L. M., Shannon, R. V., Baskent, D. & Wang, X. (2001). Speech recognition in noise as a
40 584 function of the number of spectral channels: comparison of acoustic hearing and cochlear
41 585 implants. *The Journal of the Acoustical Society of America*, 110, 1150-1163. doi:
42 586 10.1121/1.1381538
43
44
45 587 Füllgrabe, C., Moore, B. C. J. & Stone, M. A. (2015). Age-group differences in speech
46 588 identification despite matched audiometrically normal hearing: Contributions from auditory
47 589 temporal processing and cognition. *Frontiers in Aging Neuroscience*, 6, Article 347, 1-25. doi:
48 590 10.3389/fnagi.2014.00347
49
50
51 591 Gilbert, M. L., Deroche, M. L. D., Jiradejvong, P., Chan Barrett, K. & Limb, C. J. (2021). Cochlear
52 592 implant compression optimization for musical sound quality in MED-EL users. *Ear and*
53 593 *Hearing*, (in press). doi: 10.1097/AUD.0000000000001145
54
55
56
57
58
59
60

Moore Lessons from hearing aids 22

- 1
2
3 594 Glasberg, B. R. & Moore, B. C. J. (1986). Auditory filter shapes in subjects with unilateral and
4 595 bilateral cochlear impairments. *The Journal of the Acoustical Society of America*, 79, 1020-
5 596 1033. doi: 10.1121/1.393374
6
7
8 597 Goehring, T., Keshavarzi, M., Carlyon, R. P. & Moore, B. C. J. (2019). Using recurrent neural
9 598 networks to improve the perception of speech in non-stationary noise by people with cochlear
10 599 implants. *The Journal of the Acoustical Society of America*, 146, 705-718. doi:
11 600 10.1121/1.5119226
12
13 601 Goehring, T., Bolner, F., Monaghan, J. J., van Dijk, B., Zarowski, A. & Bleeck, S. (2017). Speech
14 602 enhancement based on neural networks improves speech intelligibility in noise for cochlear
15 603 implant users. *Hearing Research*, 344, 183-194. doi: 10.1016/j.heares.2016.11.012
16
17 604 Hamacher, V., Fischer, E., Kornagel, U. & Puder, H. (2006). Applications of adaptive signal
18 605 processing methods in high-end hearing instruments. In E. Hänsler & G. Schmidt (Eds.).
19 606 *Topics in Acoustic Echo and Noise Control: Selected Methods for the Cancellation of*
20 607 *Acoustical Echoes, the Reduction of Background Noise, and Speech Processing* (pp. 599-636).
21 608 New York: Springer.
22
23 609 Hansen, M. (2002). Effects of multi-channel compression time constants on subjectively perceived
24 610 sound quality and speech intelligibility. *Ear and Hearing*, 23, 369-380. doi:
25 611 10.1097/00003446-200208000-00012
26
27 612 Healy, E. W., Tan, K., Johnson, E. M. & Wang, D. (2021). An effectively causal deep learning
28 613 algorithm to increase intelligibility in untrained noises for hearing-impaired listeners. *The*
29 614 *Journal of the Acoustical Society of America*, 149, 3943-3953. doi: 10.1121/10.0005089
30
31 615 Heinz, M. G., Colburn, H. S. & Carney, L. H. (2001a). Evaluating auditory performance limits: I.
32 616 One-parameter discrimination using a computational model for the auditory nerve. *Neural*
33 617 *Computation*, 13, 2273-2316. doi: 10.1162/089976601750541804
34
35 618 Heinz, M. G., Colburn, H. S. & Carney, L. H. (2001b). Evaluating auditory performance limits: II.
36 619 One-parameter discrimination with random-level variation. *Neural Computation*, 13, 2317-
37 620 2338. doi: 10.1162/089976601750541813
38
39 621 Hjortkjær, J. & Walther-Hansen, M. (2014). Perceptual effects of dynamic range compression in
40 622 popular music recordings. *Journal of the Audio Engineering Society*, 62, 37-41. doi:
41 623 10.17743/jaes.2014.0003
42
43 624 Hockley, N. S., Bahlmann, F. & Fulton, B. (2012). Analog-to-digital conversion to accommodate
44 625 the dynamics of live music in hearing instruments. *Trends in Amplification*, 16, 146-158. doi:
45 626 10.1177/1084713812471906
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 627 Holube, I., Hamacher, V. & Wesselkamp, M. (1999). Hearing Instruments: noise reduction
4 628 strategies. *Proc. 18th Danavox Symposium: Auditory Models and Non-linear Hearing*
5 629 *Instruments*, 359-377.
- 6
7
8 630 Keidser, G., Dillon, H., Carter, L. & O'Brien, A. (2012). NAL-NL2 Empirical Adjustments. *Trends*
9 631 *in Amplification*, 16, 211-223. doi: 10.1177/1084713812468511
- 10
11
12 632 Keshavarzi, M., Goehring, T., Turner, R. E. & Moore, B. C. J. (2019). Comparison of effects on
13 633 subjective speech intelligibility and sound quality of two algorithms for reducing babble: A
14 634 deep recurrent neural network and spectral subtraction. *The Journal of the Acoustical Society*
15 635 *of America*, 145, 1493-1503. doi: 10.1121/1.5094765
- 16
17
18 636 Khing, P. P., Swanson, B. A. & Ambikairajah, E. (2013). The effect of automatic gain control
19 637 structure and release time on cochlear implant speech intelligibility. *PLoS One*, 8, e82263. doi:
20 638 10.1371/journal.pone.0082263
- 21
22
23 639 Kirchberger, M. & Russo, F. A. (2016). Dynamic range across music genres and the perception of
24 640 dynamic compression in hearing-impaired listeners. *Trends in Hearing*, 20, 1-16. doi:
25 641 10.1177/2331216516630549
- 26
27
28 642 Langner, F., Buchner, A. & Nogueira, W. (2020). Evaluation of an adaptive dynamic compensation
29 643 system in cochlear implant listeners. *Trends in Hearing*, 24, 2331216520970349. doi:
30 644 10.1177/2331216520970349
- 31
32
33 645 Launer, S., Zakis, J. & Moore, B. C. J. (2016). Hearing aid signal processing. In G. R. Popelka, B.
34 646 C. J. Moore, A. N. Popper & R. R. Fay (Eds.). *Hearing Aids* (pp. 93-130). New York:
35 647 Springer.
- 36
37
38 648 Limb, C. J. & Roy, A. T. (2014). Technological, biological, and acoustical constraints to music
39 649 perception in cochlear implant users. *Hearing Research*, 308, 13-26. doi:
40 650 10.1016/j.heares.2013.04.009
- 41
42
43 651 Litovsky, R. Y., Parkinson, A. & Arcaroli, J. (2009). Spatial hearing and speech intelligibility in
44 652 bilateral cochlear implant users. *Ear and Hearing*, 30, 419-431. doi:
45 653 10.1097/AUD.0b013e3181a165be
- 46
47
48 654 Looi, V., McDermott, H., McKay, C. & Hickson, L. (2008). Music perception of cochlear implant
49 655 users compared with that of hearing aid users. *Ear and Hearing*, 29, 421-434. doi:
50 656 10.1097/AUD.0b013e31816a0d0b
- 51
52
53 657 Lunner, T., Hietkamp, R. K., Andersen, M. R., Hopkins, K. & Moore, B. C. J. (2012). Effect of
54 658 speech material on the benefit of temporal fine structure information in speech for young
55 659 normal-hearing and older hearing-impaired participants. *Ear and Hearing*, 33, 377-388. doi:
56 660 10.1097/AUD.0b013e3182387a8c

Moore Lessons from hearing aids 24

- 661 Madsen, S. M. K. & Moore, B. C. J. (2014). Music and hearing aids. *Trends in Hearing*, 18, 1-29.
662 doi: 10.1177/2331216514558271
- 663 McDermott, H. J. (2004). Music perception with cochlear implants: a review. *Trends in*
664 *Amplification*, 8, 49-82. doi: 10.1177/108471380400800203
- 665 McDermott, H. J., McKay, C. M., Richardson, L. M. & Henshall, K. R. (2003). Application of
666 loudness models to sound processing for cochlear implants. *The Journal of the Acoustical*
667 *Society of America*, 114, 2190-2197. doi: 10.1121/1.1612488
- 668 McKay, C. M., Henshall, K. R., Farrell, R. J. & McDermott, H. J. (2003). A practical method of
669 predicting the loudness of complex electrical stimuli. *The Journal of the Acoustical Society of*
670 *America*, 113, 2054-2063. doi: 10.1121/1.1558378
- 671 Mehta, A. H. & Oxenham, A. J. (2017). Vocoder simulations explain complex pitch perception
672 limitations experienced by cochlear implant users. *Journal of the Association for Research in*
673 *Otolaryngology*, 18, 789-802. doi: 10.1007/s10162-017-0632-x
- 674 Miskolczy-Fodor, F. (1960). Relation between loudness and duration of tonal pulses. III. Response
675 in cases of abnormal loudness function. *The Journal of the Acoustical Society of America*, 32,
676 486-492. doi: 10.1121/1.1908107
- 677 Moore, B. C. J. (1973). Frequency difference limens for short-duration tones. *The Journal of the*
678 *Acoustical Society of America*, 54, 610-619. doi: 10.1121/1.1913640
- 679 Moore, B. C. J. (2003). Coding of sounds in the auditory system and its relevance to signal
680 processing and coding in cochlear implants. *Otology & Neurotology*, 24, 243-254. doi:
681 10.1097/00129492-200303000-00019
- 682 Moore, B. C. J. (2004). Testing the concept of softness imperception: Loudness near threshold for
683 hearing-impaired ears. *The Journal of the Acoustical Society of America*, 115, 3103-3111. doi:
684 10.1121/1.1738839
- 685 Moore, B. C. J. (2007). *Cochlear Hearing Loss: Physiological, Psychological and Technical*
686 *Issues, 2nd Ed.* Chichester: Wiley.
- 687 Moore, B. C. J. (2008). The choice of compression speed in hearing aids: Theoretical and practical
688 considerations, and the role of individual differences. *Trends in Amplification*, 12, 103-112.
689 doi: 10.1177/1084713808317819
- 690 Moore, B. C. J. (2012). *An Introduction to the Psychology of Hearing, 6th Ed.* Leiden, The
691 Netherlands: Brill.
- 692 Moore, B. C. J. (2014). Development and current status of the "Cambridge" loudness models.
693 *Trends in Hearing*, 18, 1-29.

- 1
2
3 694 Moore, B. C. J. (2016). Effects of sound-induced hearing loss and hearing aids on the perception of
4 695 music. *Journal of the Audio Engineering Society*, 64, 112-123. doi: 10.17743/jaes.2015.0081
5
6 696 Moore, B. C. J. & Glasberg, B. R. (1988). A comparison of four methods of implementing
7 697 automatic gain control (AGC) in hearing aids. *British Journal of Audiology*, 22, 93-104. doi:
8 698 10.3109/03005368809077803
9
10 699 Moore, B. C. J. & Glasberg, B. R. (2004). A revised model of loudness perception applied to
11 700 cochlear hearing loss. *Hearing Research*, 188, 70-88. doi: 10.1016/S0378-5955(03)00347-2
12 701 Moore, B. C. J. & Ernst, S. M. A. (2012). Frequency difference limens at high frequencies:
13 702 Evidence for a transition from a temporal to a place code. *The Journal of the Acoustical*
14 703 *Society of America*, 132, 1542-1547. doi: 10.1121/1.4739444
15 704 Moore, B. C. J. & Şek, A. (2013). Comparison of the CAM2 and NAL-NL2 hearing-aid fitting
16 705 methods. *Ear and Hearing*, 34, 83-95. doi: 10.1097/AUD.0b013e3182650adf
17 706 Moore, B. C. J. & Şek, A. (2016a). Comparison of the CAM2A and NAL-NL2 hearing-aid fitting
18 707 methods for participants with a wide range of hearing losses. *International Journal of*
19 708 *Audiology*, 55, 93-100. doi: 10.3109/14992027.2015.1095360
20 709 Moore, B. C. J. & Şek, A. P. (2016b). Preferred compression speed for speech and music and its
21 710 relationship to sensitivity to temporal fine structure. *Trends in Hearing*, 20, 1-15. doi:
22 711 10.1177/2331216516640486
23 712 Moore, B. C. J., Glasberg, B. R. & Stone, M. A. (1991). Optimization of a slow-acting automatic
24 713 gain control system for use in hearing aids. *British Journal of Audiology*, 25, 171-182. doi:
25 714 10.3109/03005369109079851
26 715 Moore, B. C. J., Stone, M. A. & Alcántara, J. I. (2001). Comparison of the electroacoustic
27 716 characteristics of five hearing aids. *British Journal of Audiology*, 35, 307-325. doi:
28 717 10.1080/00305364.2001.11745249
29 718 Moore, B. C. J., Glasberg, B. R. & Stone, M. A. (2010). Development of a new method for deriving
30 719 initial fittings for hearing aids with multi-channel compression: CAMEQ2-HF. *International*
31 720 *Journal of Audiology*, 49, 216-227. doi: 10.3109/14992020903296746
32 721 Moore, B. C. J., Füllgrabe, C. & Stone, M. A. (2011). Determination of preferred parameters for
33 722 multi-channel compression using individually fitted simulated hearing aids and paired
34 723 comparisons. *Ear and Hearing*, 32, 556-568. doi: 10.1097/AUD.0b013e31820b5f4c
35 724 Moore, B. C. J., Kolarik, A., Stone, M. A. & Lee, Y.-W. (2016). Evaluation of a method for
36 725 enhancing interaural level differences at low frequencies. *The Journal of the Acoustical Society*
37 726 *of America*, 140, 2817-2828. doi: 10.1121/1.4965299
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Moore Lessons from hearing aids 26

- 1
2
3 727 Nordqvist, P. & Leijon, A. (2004). Hearing-aid automatic gain control adapting to two sound
4 728 sources in the environment, using three time constants. *The Journal of the Acoustical Society of*
5 729 *America*, 116, 3152-3155. doi: 10.1121/1.1793207
- 6
7
8 730 Pick, G., Evans, E. F. & Wilson, J. P. (1977). Frequency resolution in patients with hearing loss of
9 731 cochlear origin. In E. F. Evans & J. P. Wilson (Eds.). *Psychophysics and Physiology of*
10 732 *Hearing* (pp. 273-281). London: Academic Press.
- 11
12
13 733 Price, G. R. & Kalb, J. T. (1991). Insights into hazard from intense impulses from a mathematical
14 734 model of the ear. *The Journal of the Acoustical Society of America*, 90, 219-227. doi:
15 735 10.1121/1.401291
- 16
17
18 736 Robinson, C. E. & Huntington, D. A. (1973). The intelligibility of speech processed by delayed
19 737 long-term averaged compression amplification. *The Journal of the Acoustical Society of*
20 738 *America*, 54, 314.
- 21
22
23 739 Seeber, B. U. & Fastl, H. (2008). Localization cues with bilateral cochlear implants. *The Journal of*
24 740 *the Acoustical Society of America*, 123, 1030-1042. doi: 10.1121/1.2821965
- 25
26
27 741 Seligman, P. & Whitford, L. (1995). Adjustment of appropriate signal levels in the Spectra 22 and
28 742 mini speech processors. *Annals of Otology, Rhinology and Laryngology*, Suppl. 166, 172-175.
- 29
30
31 743 Shannon, R. V. (1983). Multichannel electrical stimulation of the auditory nerve in man. I. Basic
32 744 psychophysics. *Hearing Research*, 11, 157-189. doi: 10.1016/0378-5955(83)90077-1
- 33
34
35 745 Souza, P. E. (2002). Effects of compression on speech acoustics, intelligibility, and sound quality.
36 746 *Trends in Amplification*, 6, 131-165. doi: 10.1177/108471380200600402
- 37
38
39 747 Steinberg, J. C. & Gardner, M. B. (1937). The dependence of hearing impairment on sound
40 748 intensity. *The Journal of the Acoustical Society of America*, 9, 11-23. doi: 10.1121/1.1915905
- 41
42
43 749 Stöbich, B., Zierhofer, C. M. & Hochmair, E. S. (1999). Influence of automatic gain control
44 750 parameter settings on speech understanding of cochlear implant users employing the
45 751 continuous interleaved sampling strategy. *Ear and Hearing*, 20, 104-116. doi:
46 752 10.1097/00003446-199904000-00002
- 47
48
49 753 Stone, M. A. & Moore, B. C. J. (1992). Syllabic compression: Effective compression ratios for
50 754 signals modulated at different rates. *British Journal of Audiology*, 26, 351-361. doi:
51 755 10.3109/03005369209076659
- 52
53
54 756 Stone, M. A. & Moore, B. C. J. (2004). Side effects of fast-acting dynamic range compression that
55 757 affect intelligibility in a competing speech task. *The Journal of the Acoustical Society of*
56 758 *America*, 116, 2311-2323. doi: 10.1121/1.1784447
- 57
58
59
60

- 1
2
3 759 Stone, M. A. & Moore, B. C. J. (2008). Effects of spectro-temporal modulation changes produced
4 760 by multi-channel compression on intelligibility in a competing-speech task. *The Journal of the*
5 761 *Acoustical Society of America*, 123, 1063-1076. doi: 10.1121/1.2821969
6
7 762 Stone, M. A., Moore, B. C. J., Alcántara, J. I. & Glasberg, B. R. (1999). Comparison of different
8 763 forms of compression using wearable digital hearing aids. *The Journal of the Acoustical*
9 764 *Society of America*, 106, 3603-3619. doi: 10.1121/1.428213
10
11 765 Stone, M. A., Moore, B. C. J., Füllgrabe, C. & Hinton, A. C. (2009). Multi-channel fast-acting
12 766 dynamic-range compression hinders performance by young, normal-hearing listeners in a two-
13 767 talker separation task. *Journal of the Audio Engineering Society*, 57, 532-546.
14
15 768 Studebaker, G. A., Sherbecoe, R. L., McDaniel, D. M. & Gwaltney, C. A. (1999). Monosyllabic
16 769 word recognition at higher-than-normal speech and noise levels. *The Journal of the Acoustical*
17 770 *Society of America*, 105, 2431-2444. doi: 10.1121/1.426848
18
19 771 Tan, C. T. & Moore, B. C. J. (2008). Perception of nonlinear distortion by hearing-impaired people.
20 772 *International Journal of Audiology*, 47, 246-256. doi: 10.1080/14992020801945493
21
22 773 Tan, C. T., Moore, B. C. J. & Zacharov, N. (2003). The effect of nonlinear distortion on the
23 774 perceived quality of music and speech signals. *Journal of the Audio Engineering Society*, 51,
24 775 1012-1031.
25
26 776 Vaerenberg, B., Govaerts, P. J., Stainsby, T., Nopp, P., Gault, A. & Gnansia, D. (2014). A uniform
27 777 graphical representation of intensity coding in current-generation cochlear implant systems.
28 778 *Ear and Hearing*, 35, 533-543. doi: 10.1097/AUD.0000000000000039
29
30 779 Vanderkooy, J. & Lipshitz, S. P. (1984). Resolution below the least significant bit in digital systems
31 780 with dither. *Journal of the Audio Engineering Society*, 32, 106-113.
32
33 781 Verschuure, J., Dreschler, W. A., de Haan, E. H. et al. (1993). Syllabic compression and speech
34 782 intelligibility in hearing impaired listeners. *Scandinavian Audiology*, 22, Suppl. 38, 92-100.
35
36 783 von Békésy, G. (1960). *Experiments in Hearing*. New York: McGraw-Hill.
37
38 784 Wiggins, I. M. & Seeber, B. U. (2011). Dynamic-range compression affects the lateral position of
39 785 sounds. *The Journal of the Acoustical Society of America*, 130, 3939-3953. doi:
40 786 10.1121/1.3652887
41
42 787 Wiggins, I. M. & Seeber, B. U. (2013). Linking dynamic-range compression across the ears can
43 788 improve speech intelligibility in spatially separated noise. *The Journal of the Acoustical*
44 789 *Society of America*, 133, 1004-1016. doi: 10.1121/1.4773862
45
46 790 Wilson, B. S., Finley, C. C., Lawson, D. T., Wolford, R. D., Eddington, D. K. & Rabinowitz, W.
47 791 M. (1991). Better speech recognition with cochlear implants. *Nature*, 352, 236-238. doi:
48 792 10.1038/352236a0

- 1
2
3 793 Wolfe, J., Schafer, E. C., John, A. & Hudson, M. (2011). The effect of front-end processing on
4 794 cochlear implant performance of children. *Otology & Neurotology*, 32, 533-538. doi:
5 795 10.1097/MAO.0b013e318210b6ec
6
7
8 796 Wouters, J., McDermott, H. J. & Francart, T. (2015). Sound coding in cochlear implants: From
9 797 electric pulses to hearing. *IEEE Signal Processing Magazine*, 32, 67-80. doi:
10 798 10.1109/MSP.2014.2371671
11
12 799 Yang, L. P. & Fu, Q. J. (2005). Spectral subtraction-based speech enhancement for cochlear
13 800 implant patients in background noise. *The Journal of the Acoustical Society of America*, 117,
14 801 1001-1004. doi: 10.1121/1.1852873
15
16 802 Zakis, J. (2016). Music perception and hearing aids. In G. R. Popelka, B. C. J. Moore, A. N. Popper
17 803 & R. R. Fay (Eds.). *Hearing Aids* (pp. 217-252). New York: Springer.
18
19 804 Zeng, F.-G. (2004). Compression and cochlear implants. In S. P. Bacon, A. N. Popper & R. R. Fay
20 805 (Eds.). *Compression: From Cochlea to Cochlear Implants* (pp. 184-220). New York: Springer.
21
22 806 Zeng, F.-G., Popper, A. N. & Fay, R. R. (2003). *Auditory Protheses*. New York: Springer-Verlag.
23
24 807 Zeng, F. G. & Shannon, R. V. (1995). Loudness of simple and complex stimuli in electrical
25 808 hearing. *Annals of Otology, Rhinology and Laryngology*, 104, Suppl. 166, 235-238.
26
27 809 Zeng, F. G., Galvin, J. J. & Zhang, C. Y. (1998). Encoding loudness by electric stimulation of the
28 810 auditory nerve. *Neuroreport*, 9, 1845-1848. doi: 10.1097/00001756-199806010-00033
29
30
31
32
33
34 811
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 812 Figure captions
4

5 813 **Figure 1.** Preferences for slow versus fast compression for individual subjects and for the mean
6
7 814 across subjects, as obtained by Moore and Søk (2016b) using a simulated HA. Bars falling above
8
9 815 the horizontal line at the center of the panel indicate preferences for slow compression over fast
10
11 816 compression. Redrawn from Moore and Søk (2016b).

12
13 817 **Figure 2.** Simplified schematic diagram of the signal processing performed in several CIs.

14
15 818 **Figure 3.** Schematic input-output functions for the Advanced Bionics CIs for input dynamic ranges
16
17 819 (IDRs) of 50, 65, and 80 dB.

18
19 820 **Figure 4.** Schematic diagram of the “Voice Guard” system used in the Oticon Medical CI.

20
21 821 **Figure 5.** Schematic diagram of the adaptive mapping from the envelope magnitude in a given
22
23 822 channel to the electrical output for that channel, as used in the Oticon Medical Voice Guard system.
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Moore

Lessons from hearing aids

1

1

2

3

4

5 Brian C. J. Moore

6

7 Cambridge Hearing Group, Department of Psychology, University of Cambridge,

8 Downing Street, Cambridge CB2 3EB, England

9

10

11 email: bcjm@cam.ac.uk

12

For Peer Review

Abstract

Some of the problems experienced by users of hearing aids (HAs) when listening to music are relevant to cochlear implants (CIs). One problem is related to the high peak levels (up to 120 dB SPL) that occur in live music. Some HAs and CIs overload at such levels, because of the limited dynamic range of the microphones and analog-to-digital converters (ADCs), leading to perceived distortion. Potential solutions are to use 24-bit ADCs or to include an adjustable gain between the microphones and the ADCs. A related problem is how to squeeze the wide dynamic range of music into the limited dynamic range of the user, which can be only 6-20 dB for CI users. In HAs, this is usually done via multi-channel amplitude compression (automatic gain control, AGC). In CIs, a single-channel front-end AGC is applied to the broadband input signal or a control signal derived from a running average of the broadband signal level is used to control the mapping of the channel envelope magnitude to an electrical signal. This introduces several problems: (1) an intense narrowband signal (e.g. a strong bass sound) reduces the level for *all* frequency components, making some parts of the music harder to hear; (2) the AGC introduces cross-modulation effects that can make a steady sound (e.g. sustained strings or a sung note) appear to fluctuate in level. Potential solutions are to use several frequency channels to create slowly varying gain-control signals and to use slow-acting (or dual time-constant) AGC rather than fast-acting AGC.

Keywords: music, dynamic range, automatic gain control, hearing aids, cochlear implants, compression speed

37 Introduction

38 Users of both hearing aids (HA) and cochlear implants (CI) experience problems when
39 listening to music (McDermott, 2004; Looi, McDermott, McKay & Hickson, 2008; Chasin &
40 Hockley, 2014; Madsen & Moore, 2014; Moore, 2016). Some of these problems are related to the
41 reduced resolution and processing capacity of the impaired auditory system (Shannon, 1983;
42 Moore, 2003; Moore, 2007). For example, both HA users and CI users have reduced spectral
43 resolution (Pick, Evans & Wilson, 1977; Glasberg & Moore, 1986; Friesen, Shannon, Baskent &
44 Wang, 2001). For CI users, this reduction is very substantial because of the spread of current within
45 the cochlea, and this so it is likely to probably severely limits the ability to “hear out” one
46 instrument or voice in the presence of another instrument or voice (Mehta & Oxenham, 2017). This
47 in turn will limit the enjoyment of any music that includes multiple instruments and voices (Limb &
48 Roy, 2014) and this limitation is unlikely to be alleviated by changes in CI processing, unless a way
49 can be found of making the electrical stimulation much more place selective. However, other
50 problems are related to the design of the HAs and CIs and these problems can potentially be
51 reduced by improvements in design. This paper considers some of the lessons that have been
52 learned from studies of music perception via HAs and describes how those lessons might be applied
53 to improving the design of CIs.

54 In both HAs and CIs, the broadband signals picked up by the microphone(s) are passed
55 through an array of bandpass filters to create channel signals corresponding to the outputs of filters
56 with different center frequencies. In HAs, the channel signals are processed in various ways before
57 being combined to create a broadband signal that is delivered to a miniature loudspeaker (called a
58 receiver) that generates the output sound. In CIs the channel signals are used to create electrical
59 signals that are delivered to the individual electrodes in the array implanted within the cochlea.
60 Hence, in CIs, the number of channels is usually equal to the number of implanted electrodes,
61 although the effective number of independent channels is smaller than the number of electrodes,
62 because of current spread in the cochlea (Friesen et al., 2001) (Friesen et al., 2001). In a CI, the
63 signal in the channel with the lowest center frequency is usually used to derive the electrical signal
64 delivered to the most apical electrode in the cochlea, while the signal in the channel with the
65 highest center frequency is usually used to drive the most basal electrode, with a continuous

1
2
3
4 66 gradation in between. This represents an attempt to recreate the tonotopic mapping of frequency to
5
6 67 place that occurs within a normal cochlea (von Békésy, 1960). Thus, information about the short-
7
8 68 term spectrum of sounds is conveyed by the relative strength of the electrical signals across the
9
10 69 electrode array.

71 **Requirements for Music Listening With CIs**

72 It is helpful first to consider the properties that a CI should have in order to improve the
73 experience of listening to music. The key properties are:

74 (1) Envelope cues in different frequency regions, which are the main source of auditory information
75 provided by CIs (Wilson, Finley, Lawson, Wolford, Eddington & Rabinowitz, 1991; Clark, Tong &
76 Patrick, 1990; Zeng, Popper & Fay, 2003), should be preserved and coded as faithfully as possible.

77 (2) Nonlinear distortion occurring prior to filtering of the signal into frequency bands or channels
78 should be minimal. This is because nonlinear distortion leads both to changes in the shape of the
79 waveform and to frequency components in the short-term spectrum that are not present in the
80 original signal, giving a misleading representation of the spectral shape of a sound (Tan, Moore &
81 Zacharov, 2003). Nonlinear distortion introduced after the signal is filtered into frequency channels
82 in a CI does not have the same effect on the representation of the spectrum of a sound, but it can
83 distort the representation of the envelope in each channel.

84 (3) The wide range of sound levels that occur in music, especially live music, must be ~~squashed or~~
85 compressed into the narrow range of electrical current values between the detection threshold and
86 the highest comfortable level. This must be done while preserving the representation of the
87 envelope of the sound in each frequency channel as well as possible.

88 As is discussed below, these properties are not easy to achieve, and the current state of the
89 art is far from optimal in this respect.

91 **Sources of Envelope Distortion**

92 In this section, I consider some (but not all) of the sources of envelope distortion in HAs
93 and CIs. Envelope distortion introduced by automatic gain control systems is discussed in a later
94 section. Most HAs and CIs incorporate various forms of adaptive signal processing, i.e. signal

1
2
3
4 95 processing that changes over time in response to changes in the input signal. Examples are noise
5
6 96 reduction and adaptive directional processing. These forms of signal processing, because they are
7
8 97 time-varying, inevitably distort the envelopes of the channel signals in HAs and CIs.

9
10 98 Noise-reduction systems can be applied to the signal from a single microphone. They have
11
12 99 been designed to improve the ability to understand speech in the presence of background noise.
13
14 100 They generally work by estimating the momentary speech-to-noise ratio in each channel and
15
16 101 applying attenuation to the channels with the poorest estimated speech-to-noise ratio (Holube,
17
18 102 Hamacher & Wesselkamp, 1999; Hamacher, Fischer, Kornagel & Puder, 2006; Launer, Zakis &
19
20 103 Moore, 2016; Yang & Fu, 2005). In more recent approaches, artificial neural networks have been
21
22 104 used to process speech in background sounds for application both to HAs (Keshavarzi, Goehring,
23
24 105 Turner & Moore, 2019; Healy, Tan, Johnson & Wang, 2021) and CIs (Goehring, Bolner,
25
26 106 Monaghan, van Dijk, Zarowski & Bleeck, 2017; Goehring, Keshavarzi, Carlyon & Moore, 2019).

27
28 107 All single-microphone noise-reduction systems involve a trade-off; the more the
29
30 108 background noise is reduced relative to the speech, the more distortion there is, including envelope
31
32 109 distortion. There appear to be large individual differences in preferences for noise-reduction
33
34 110 systems and in preferences for the amount of noise reduction, some people being “noise haters” and
35
36 111 others being “distortion haters” (Brons, Houben & Dreschler, 2012). In any case, since these noise-
37
38 112 reduction systems have been designed to improve the perception of speech in noise, they are
39
40 113 unlikely to be of any benefit when listening to music. Rather, the envelope distortion that they
41
42 114 introduce is likely to degrade the perception of music by users of HAs and CIs. Therefore, it is
43
44 115 recommended that a CI is set up with a dedicated music program, as is often done for HAs, and in
45
46 116 the music program any adaptive noise-reduction processing is disabled.

47
48 117 A second form of adaptive processing, directional processing, is used when two or more
49
50 118 microphones are available, as is the case with most HAs and CIs. Some (but not all) such systems
51
52 119 are based on the assumption that the “target” sound that the user wishes to hear comes from the
53
54 120 front. They attempt to estimate, for each frequency channel, the direction of the most prominent
55
56 121 interfering sound coming from the sides or back, and to create a null in the directional response so
57
58 122 as to attenuate that interfering sound (Launer et al., 2016). This creates time-varying changes in the
59
60 123 effective frequency response of the HA or CI for sounds coming from the sides or back, and also

1
2
3
4 124 distorts the envelope representation of such sounds. Again, this is likely to impair music perception,
5
6 125 since music often involves spatially distributed sound sources. Even for directional processing
7
8 126 systems that attempt to preserve sounds of interest from several directions, the processing is
9
10 127 adaptive and time varying, and is likely to result in the introduction of spurious amplitude
11
12 128 modulation. As for noise-reduction processing, it is recommended that a CI is set up with a
13
14 129 dedicated music program and that any adaptive directional processing is disabled for that program.
15

16 130

17 131 **The Problem of the Dynamic Range of the Input Signal**

18
19
20 132 Most CIs and some HAs use 16-bit analog-to-digital converters (ADCs) to digitize the
21
22 133 microphone signals (Launer et al., 2016; Zakis, 2016). In theory this can code a 96-dB range of
23
24 134 input sound levels (6 dB per bit). In practice the achieved range is typically 85-90 dB, because of
25
26 135 microphone noise, noise in the ADCs themselves, and noise in the electronic circuitry. Although in
27
28 136 principle the dynamic range can be increased by the application of a low-level spectrally shaped
29
30 137 noise called dither (Vanderkooy & Lipshitz, 1984), this is not to my knowledge applied in HAs or
31
32 138 CIs. In most HAs and CIs, the gain of the pre-amplifier between the microphone and the ADC is set
33
34 139 so that the lowest sound level that can be coded is about 15 dB SPL, which means that the highest
35
36 140 sound level that can be coded is 100-105 dB SPL. When listening to music in the home, peak sound
37
38 141 levels rarely exceed 95 dB SPL and most CIs and HAs can handle this without significant
39
40 142 distortion. However, when listening to live music, or amplified music in a club or discotheque, peak
41
42 143 sound levels can reach 115-120 dB SPL (Hockley, Bahlmann & Fulton, 2012; Chasin & Hockley,
43
44 144 2014). Sound at such levels can be unpleasant and may appear distorted even for people with
45
46 145 normal hearing, because of nonlinear distortion produced in the outer and middle ear (Price & Kalb,
47
48 146 1991) and because of the effects of upward spread of masking (Studebaker, Sherbecoe, McDaniel
49
50 147 & Gwaltney, 1999). However, for users of HAs and CI, such high sound levels can also lead to ~~and~~
51
52 148 this can lead to overload (peak clipping) that reduces the perceived sound quality, at least for users
53
54 149 of HAs (Tan & Moore, 2008). Apart from introducing spectral changes in the signal, peak clipping
55
56 150 results in a distortion of the envelope cues that are important for CI users.

57
58 151 There are several solutions to this problem. One is to use 24-bit ADCs, a solution that has
59
60 152 been adopted by several manufacturers of HAs, even though it decreases battery life. Another

153 solution is to include an adjustable gain between the microphones and the ADCs; this gain
154 adjustment can be compensated for in the subsequent digital-processing stages of the HA or CI
155 (Hockley et al., 2012; Zakis, 2016). This is done in some ~~HAs, but not, to my knowledge, in but not~~
156 ~~all HAs and~~ CIs. It is recommended, therefore, that the range of signal levels that can be handled by
157 CIs without distortion is increased, either via the use of 24-bit ADCs or via the use of an adjustable
158 gain between the microphones and the ADCs.

160 **Squeezing Music Into the Limited Dynamic Range of the CI** 161 **User**

162 *The Perceptual Dynamic Range for Users of HAs and CIs*

163 Much recorded music, and music that is broadcast or transmitted via the internet, is
164 subjected to some form of amplitude compression to reduce its dynamic range (the difference
165 between the highest level and lowest level in the music) (Croghan, Arehart & Kates, 2014;
166 Hjortkjær & Walther-Hansen, 2014). This is especially true for “pop” music. However, live music,
167 especially classical music, can have a very wide range dynamic range; the peaks occurring during a
168 *fortississimoforte* (*fff*) passage may be 70 dB or more above the dips in level in a
169 *pianississimopiano* (*ppp*) passage (Hockley et al., 2012; Kirchberger & Russo, 2016). The
170 perceptual dynamic range – defined here as the range between the detection threshold and the level
171 at which sound starts to become uncomfortably loud – is typically at least 95-100 dB for people
172 with normal hearing (Moore, 2012). This is enough to allow the *pianississimopiano* passages to be
173 heard without the *fortississimoforte* passages being uncomfortably loud, except perhaps when a
174 *pianississimo* passage occurs immediately after a *fortississimo* passage. However, hearing loss
175 usually results in loudness recruitment (a more rapid than normal growth of loudness with
176 increasing sound level, once the sound level exceeds the elevated detection threshold) and a
177 reduced perceptual dynamic range (Fowler, 1936; Steinberg & Gardner, 1937). For a hearing loss
178 of, say, 70 dB, the perceptual dynamic range is typically only about 30 dB (Miskolczy-Fodor, 1960;
179 Moore, 2004; Moore & Glasberg, 2004). For electrical pulses delivered directly to an individual
180 electrode in a CI, the perceptual dynamic range is typically only 6-20 dB (Fourcin et al., 1979;
181 Zeng, 2004). Hence some form of amplitude compression is essential to squeeze the wide range of

1
2
3
4 182 sound levels encountered during performances of live music into the limited perceptual dynamic
5
6 183 range of users of HAs and CIs.

7
8 184
9
10 185 *Amplitude Compression in HAs*
11
12 186 In HAs, amplitude compression is usually achieved via multi-channel automatic gain
13
14 187 control (AGC). The digitized microphone signal is filtered into several frequency channels
15
16 188 (typically between 3 and 24) and independent AGC is applied in each channel. A basic
17
18 189 characteristic of an AGC system is its input-output function, a plot of the output level as a function
19
20 190 of the input level, both in dB. For many AGC systems, the gain (output level minus input level) is
21
22 191 independent of the input level for low input levels. Hence, the input-output function has a slope of
23
24 192 one. This is called linear amplification. At higher input levels, the gain decreases progressively with
25
26 193 increasing input level, and the input-output function has a slope of less than one. The compression
27
28 194 threshold is defined as the input level at which the gain is reduced by 2 dB, relative to the gain
29
30 195 applied in the region of linear amplification (ANSI, 2014). One reason for having a compression
31
32 196 threshold is that it is impractical to continue to increase the gain indefinitely as the input level
33
34 197 decreases because this would make microphone noise or low-level environmental noises intrusive.
35
36 198 Indeed, for very low input levels, the gain in many HAs is reduced to prevent such noises from
37
38 199 being audible; this is called expansion, and it is associated with a slope of the input-output function
39
40 200 greater than one.

41
42 201 Over the range of input levels where compression is applied, the “amount” of compression
43
44 202 is specified by the compression ratio, the change in input level (in dB) required to achieve a 1-dB
45
46 203 change in output level (for an input level above the compression threshold). For example, a
47
48 204 compression ratio of 3 means that each 3-dB increase in input level leads to a 1-dB increase in
49
50 205 output level, so the input-output function has a slope of 1/3. In most HAs the compression ratio is
51
52 206 set independently for each channel depending on the perceptual dynamic range at the center
53
54 207 frequency of that channel, which is often estimated from the audiometric threshold (Keidser,
55
56 208 Dillon, Carter & O'Brien, 2012; Moore, Glasberg & Stone, 2010).

57
58 209 AGC systems vary in how rapidly they respond to a sudden change in the input sound level.
59
60 210 The speed of the AGC systems in HAs is typically measured using a sound whose input level

changes abruptly between 55 and 90 dB SPL (ANSI, 2014). When the sound level increases, the gain decreases. The time taken for the output to get within 3 dB of its steady value is called the attack time. When the sound level decreases, the gain increases. The time taken for the output to increase to within 4 dB of its steady value is called the recovery time or release time. For HAs, there is no general consensus about what compression speed is “best” (Moore, 2008). Some manufacturers use fast-acting compression, typically with attack times in the range 0.5-20 ms and release times in the range 5-150 ms. Some manufacturers use slow-acting compression, typically with attack times in the range 20-2000 ms and release times in the range 500-5000 ms or more. The attack time is usually chosen to be smaller than the release time to reduce the experience of uncomfortable loudness when sudden increases in sound level occur. Some manufacturers use systems with two or more speeds (Moore & Glasberg, 1988; Moore, Glasberg & Stone, 1991; Stöbich, Zierhofer & Hochmair, 1999; Stone, Moore, Alcántara & Glasberg, 1999; Nordqvist & Leijon, 2004). In these systems, the gain is mostly determined by a slow-acting AGC system, but a fast-acting system takes over and rapidly reduces the gain if there is a sudden increase in sound level. If the increase in sound level is brief (for example the sound of a door slamming or a knife being dropped on a plate), the gain returns to the value set by the slow-acting system.

There have been several studies of preferences for compression speed for HA users when listening to music. Hansen (2002) used a simulated 15-channel hearing aid and a paired-comparison procedure. The gain for medium input sound levels was adjusted for each subject using the NAL-R fitting rule (Byrne & Dillon, 1986), but the compression ratio was fixed at 2 for all channels and subjects, rather than being tailored to suit the individual hearing losses. The stimuli included classical and pop music presented at an input level of 80 dB SPL. The longest release time used (4000 ms) was significantly preferred over the two shorter release times used (400 and 40 ms).

Moore, Füllgrabe and Stone (2011) used a paired-comparison procedure and a simulated five-channel HA to compare preferences for slow versus fast compression, using both speech and music stimuli and using three mean levels for the input signal, 50, 65 and 80 dB SPL. The simulated HA was programmed to suit the individual hearing losses, which were mild, using the CAM2 procedure (Moore et al., 2010). They found slight mean preferences for slow compression for both speech and music, but the effect was only clear for an input level of 80 dB SPL, and not for

1
2
3
4 240 input levels of 50 and 65 dB SPL. They also assessed the effect of slightly delaying the audio signal
5
6 241 relative to the gain-control signal (by 2.55 to 15 ms). This delay, referred to as the “alignment
7
8 242 delay”, allowed the gain to be reduced just *before* any sudden increase in sound level (Robinson &
9
10 243 Huntington, 1973; Verschuure et al., 1993), thus increasing the effectiveness of the AGC system in
11
12 244 preventing loudness discomfort. The effect of the alignment delay was small except for a highly
13
14 245 percussive sound (solo xylophone) combined with the fast compression system, for which
15
16 246 alignment delays of 2.5, 5, and 7.5 ms were preferred over no delay.

17
18 247 Croghan et al. (2014) used a paired-comparison procedure and a simulated HA with either 3
19
20 248 or 18 channels to compare preferences for slow versus fast compression. They used both rock and
21
22 249 classical music, and the stimuli were subjected to various amounts of compression *prior to* the
23
24 250 simulated HA processing, to simulate the compression that is often applied to recorded music. The
25
26 251 stimuli were presented at a nominal level of 65 dB SPL prior to the simulated HA processing. The
27
28 252 HAs were individually fitted using the NAL-NL1 prescription method (Byrne, Dillon, Ching,
29
30 253 Katsch & Keidser, 2001). For stimuli that were not compressed prior to HA processing, slow
31
32 254 compression was preferred over fast compression for both the 3- and 18-channel systems and for
33
34 255 both classical and rock music.

35
36 256 Moore and Søk (2016b) used a paired-comparison procedure and a simulated five-channel
37
38 257 HA to compare preferences for slow versus fast compression. Three mean levels were used for the
39
40 258 input signal, 50, 65 and 80 dB SPL. They tested a hypothesis put forward by Moore (2008) that
41
42 259 reduced sensitivity to the temporal fine structure (TFS) of sounds may be associated with a
43
44 260 preference for slow compression. This hypothesis is based on the observation that high sensitivity
45
46 261 to TFS is associated with a better ability to understand speech at low speech-to-background ratios
47
48 262 (Füllgrabe, Moore & Stone, 2015), perhaps because TFS information is useful for the perceptual
49
50 263 segregation of the target and background sounds (Lunner, Hietkamp, Andersen, Hopkins & Moore,
51
52 264 2012). ~~The ability to “glimpse” speech in the dips of background sounds, often called dip listening,~~
53
54 265 ~~becomes more important at low speech-to-background ratios (Bernstein & Grant, 2009) and the~~
55
56 266 ~~greater audibility of low-level parts of the signal provided by fast compression (Moore, Peters &~~
57
58 267 ~~Stone, 1999) may be beneficial for HA users with good sensitivity to TFS. More generally~~ Also,
59
60 268 high sensitivity to TFS may be associated with less reliance on envelope cues in different frequency

269 channels and with more resistance to the envelope distortion produced by fast-acting compression
270 (Stone & Moore, 1992; 2004; Souza, 2002; Stone, Moore, Füllgrabe & Hinton, 2009). On the other
271 hand, people with poor sensitivity to TFS, including users of CIs, may rely mainly on envelope
272 cues in different frequency channels, and for them it is important to avoid the envelope distortion
273 that is introduced by fast compression. Hence, slow compression may be the preferred option for
274 HA users and CI users with poor sensitivity to TFS.

275 Moore and Şek (2016b) tested subjects with moderate-to-severe sensorineural hearing loss
276 and assessed their sensitivity to TFS indirectly by measuring difference limens for frequency
277 (DLFs) for a pure tone signal centered at 2000 Hz, based on the (not universally accepted)
278 assumption that frequency discrimination depends on the use of TFS information for frequencies up
279 to a few kHz (Moore, 1973; Heinz, Colburn & Carney, 2001a; Heinz, Colburn & Carney, 2001b;
280 Moore & Ernst, 2012). The simulated HA was programmed to suit the individual hearing losses
281 using the CAM2 prescription method (Moore et al., 2010), modified slightly as described in Moore
282 and Şek (2013) and Moore and Şek (2016a). Only one ear of each subject was tested.

283 For all three input levels used, there were small overall preferences for slow over fast
284 compression. Figure 1 shows the individual and mean preferences averaged across the three levels
285 ~~for the 22 subjects~~ for listening to music. Preference scores above 0 indicate a preference for slow
286 compression. Nine subjects showed a preference for slow compression of 0.5 scale units or more
287 (the scale went from -3 to +3, with 0 indicating no preference), and the rest showed very small
288 preferences for slow compression or no clear preference. Thus, at least for the music stimuli used in
289 the study (jazz music, classical music and a man singing with a guitar), slow-acting compression
290 seems to be a “safe” option. The logarithms of the DLFs for the test ears were weakly but
291 significantly correlated with preference scores for music, based on a one-tailed test: $r = 0.39, p <$
292 0.05 . This provides some weak support for the hypothesis that poorer sensitivity to TFS (indicated
293 by larger DLFs) is associated with stronger preferences for slow compression.

294 In summary, studies of hearing-impaired subjects using simulated HAs suggest that slow
295 compression is preferable to fast compression for listening to music. However, it should be
296 remembered that while the studies described above used a reasonably wide range of input levels
297 (50-80 dB SPL), the range was not as wide as might be encountered when listening to live music.

1
2
3
4 298 Also, the stimuli did not include conditions where the level changed abruptly from a high to a low
5
6 299 value, as might occur in music. When the level decreases abruptly, the gain in a slow-acting
7
8 300 compression system takes some time to increase; for example, the recovery time of the slow-acting
9
10 301 compression used by Moore and Søk (2016b) was 3000 ms. For users of HAs the subjective effect
11
12 302 of an abrupt decrease in level is that the HA has become “dead” for a second or two, which is
13
14 303 clearly undesirable.

15
16 30417
18 305 *Amplitude Compression in CIs*

19
20 306 Because of the very small perceptual dynamic range for electrical signals applied to the
21
22 307 auditory nerve, all CIs incorporate some form of amplitude compression. Figure 2 is a simplified
23
24 308 schematic of the signal processing that is performed in many CIs. The analog microphone signal is
25
26 309 subjected to analog-to-digital conversion (ADC) (often after pre-amplification and filtering to
27
28 310 reduce the influence of strong low-frequency components) and the broadband digital signal is
29
30 311 subjected to AGC. The amplitude-compressed signal is then filtered into frequency channels using
31
32 312 an array of bandpass filters (BPF, 1 to n), the channel signals are rectified and lowpass filtered
33
34 313 (LPF) to extract the envelope, the envelope signals are subjected to instantaneous compression
35
36 314 (often referred to as mapping), and the compressed envelope is used to modulate the amplitude or
37
38 315 width of the biphasic current pulses delivered to each electrode. Thus, there are two stages of
39
40 316 amplitude compression, produced by the front-end AGC and by the mapping for the individual
41
42 317 channels. The CIs produced by three of the main manufacturers, Cochlear Corporation, Med-El,
43
44 318 and Advanced Bionics, all have this general structure, although they differ in the speed of the front-
45
46 319 end AGC, the number of channels, the sharpness of the filters used to create the channel signals,
47
48 320 and the method of envelope extraction (Vaerenberg, Govaerts, Stainsby, Nopp, Gault & Gnansia,
49
50 321 2014). The early devices made by Cochlear Corporation have a front-end slow-acting “automatic
51
52 322 sensitivity control” (Seligman & Whitford, 1995) followed by a fast-acting AGC system, with a
53
54 323 very high compression ratio (called a compression limiter) (Khing, Swanson & Ambikairajah,
55
56 324 2013). The current CIs manufactured by Med-El, Advanced Bionics, and Cochlear Corporation
57
58 325 incorporate a dual time-constant system (Stöbich et al., 1999; Boyle, Buchner, Stone, Lenarz &
59
60 326 Moore, 2009; Khing et al., 2013) similar to that described by Stone et al. (1999).- The CIs

1
2
3
4 327 manufactured by Cochlear Corporation also include the option of a multi-channel slow-acting AGC
5
6 328 system called adaptive dynamic range optimization (ADRO, Blamey, 2005) which is applied
7
8 329 separately to the envelope signal in each channel (Wolfe, Schafer, John & Hudson, 2011; Khing et
9
10 330 al., 2013). To my knowledge, there are no published evaluations of the ADRO system for music
11
12 331 listening.

13
14 332 When listening to music, there are several problems associated with the use of single-
15
16 333 channel AGC applied to the broadband signal:
17
18 334 (1) The input-output functions of the front-end AGCs used in CIs have been designed to work well
19
20 335 for the typical speech signals encountered in everyday life but may not be optimal for music. For
21
22 336 example, the Advanced Bionics system uses a compression threshold of about 63 dB SPL
23
24 337 combined with a high compression ratio (about 12). This means that many music signals would be
25
26 338 subjected to strong (albeit slow-acting) amplitude compression, reducing the impression of dynamic
27
28 339 changes in music, such as the contrast between a *forte* passage and a *piano* passage. In comparison,
29
30 340 the multi-channel AGC systems used in HAs typically have compression thresholds in the range 20
31
32 341 to 45 dB SPL and compression ratios in the range 1.5 to 3 (Moore, Stone & Alcántara, 2001).
33
34 342 (2) A relatively intense narrowband signal (e.g. a strong bass sound) will reduce the level for *all*
35
36 343 frequency components, making some parts of the music harder to hear.
37
38 344 (3) The AGC introduces cross-modulation effects (Stone & Moore, 2004; Stone & Moore, 2008): a
39
40 345 brief strong signal from one source (e.g. a drum beat) leads to a reduction in gain for another sound
41
42 346 source (e.g. strings). This can make a steady sound (e.g. sustained strings or a sung note) appear to
43
44 347 fluctuate in level and it may make it harder to segregate sound sources. Such effects are strongest
45
46 348 for fast-acting compression, but they occur to some extent even when slow compression is used,
47
48 349 unless it is very slow indeed.

49
50 350 Improvements might be produced by:

51
52 351 (1) Changing the input-output function of the front-end AGC. ~~I am not aware of any~~ There are very
53
54 352 few systematic studies exploring the effect on music perception of varying the compression
55
56 353 threshold or the compression ratio of the front-end AGC in a CI. Gilbert, Deroche, Jiradejvong,
57
58 354 Chan Barrett and Limb (2021) explored the effect of varying the compression ratio of the front-end
59
60 355 AGC in the Med-El CI, keeping the compression threshold fixed at 48 dB SPL, and found no

356 significant effect on music preferences. However, they varied the compression ratio only over a
357 small range (2.5 to 3.5) and they used samples of music with a single fixed input level of 65 dBA.
358 They did not explore the effect of simultaneously varying the compression threshold and the
359 compression ratio. ~~but it~~ seems plausible that a lower compression threshold and a lower
360 compression ratio than currently used would give a better impression of the dynamic changes in a
361 piece of music.

362 (2) Using a multi-channel rather than single-channel front-end AGC system. This would prevent an
363 intense narrowband signal from reducing the level of all frequency components, and it would also,
364 to some extent, reduce the cross-modulation effects described above.

365 (3) For dual time-constant AGC systems, making the slow time constants even slower. However,
366 experimental studies are required to assess whether or not this is beneficial and to determine the
367 optimal compression speed, if there is such a thing.

368 As noted above, in addition to the front-end AGC, the CIs of Cochlear Corporation, Med-El,
369 and Advanced Bionics include a stage of instantaneous compression in the transformation from the
370 envelope magnitude in a given channel to the electrical signal applied to the electrode that is driven
371 by that channel. Figure 3 is a schematic illustration of the mapping system used in each channel of
372 an Advanced Bionics CI. The x-axis shows the input sound level. The y-axis shows the magnitude
373 of the electrical signal delivered to the electrode in current units (CU). In this example, the
374 threshold (T) level for that channel is 20 CU and the most comfortable level (M) is 200 CU. The
375 input dynamic range (IDR) is the range of levels that the sound processor codes into electrical
376 stimulation current in the range between the T level and the M level (20 to 200 CU) for each
377 channel. The default IDR is 60 dB, but the IDR can be adjusted via a “sensitivity” control. Figure 3
378 shows schematic input-output functions for IDRs of 50, 65, and 80 dB. Levels above the top of the
379 IDR are mapped to a very small range of CUs from the M level (200 CU) to an upper limit of about
380 220 CU (depending on the electrical pulse width and rate). Levels below the bottom of the IDR are
381 mapped to an inaudible current, below the T level. The mapping for input levels above the
382 compression threshold of 63 dB SPL depends on the time-varying gain of the front-end AGC.

383 In Figure 3, the input-output functions are straight lines when plotted as CU in linear units
384 against effective level in dB. This represents a compressive form of mapping. The shapes of the

1
2
3
4 385 input output-functions are loosely based on studies of the relationship between electrical signals
5
6 386 applied to a single electrode and loudness (Zeng & Shannon, 1995; Zeng, Galvin & Zhang, 1998).
7
8 387 The idea is that loudness fluctuations over time conveyed by changes in the electrical signal applied
9
10 388 to a given electrode should resemble the loudness fluctuations that would occur in a normal ear in
11
12 389 response to fluctuations in sound level in a local frequency region. In practice, however, the shapes
13
14 390 of the functions relating loudness to the electrical signal applied to an electrode can vary from one
15
16 391 CI user to another and can vary across electrodes within a CI user. Also, for input signals with
17
18 392 levels above the compression threshold (63 dB SPL in this example), which would occur often for
19
20 393 live music, much stronger compression is applied. Thus, the mapping process inevitably results in
21
22 394 some distortion of the internal representation of the envelope for CI users. Finally, the shapes of the
23
24 395 functions mapping envelope magnitude to electrical signal vary across manufacturers (Vaerenberg
25
26 396 et al., 2014), and it is not clear what form of mapping gives the most faithful internal representation
27
28 397 of the envelope in each channel.

29
30 398 One manufacturer of CIs, Oticon Medical, uses a different system, which is illustrated in
31
32 399 Figure 4, adapted from Langner, Buchner and Nogueira (2020). This system does not use a front-
33
34 400 end AGC. Rather all amplitude compression is applied in the mapping from envelope magnitude in
35
36 401 a given channel to the current applied to the corresponding electrode (the electrical pulse width is
37
38 402 modulated). This system is called “Voice Guard” since it was designed to convey the envelope
39
40 403 fluctuations of speech signals as faithfully as possible. The mapping from envelope amplitude to
41
42 404 electrical pulse width is compressive and it is also adaptive. The root-mean-square (RMS) level of
43
44 405 the input signal is averaged over the previous 2 s. The estimate of the RMS level is updated every 2
45
46 406 ms. This slowly-changing RMS level estimate is used to control the “knee-point” of the mapping
47
48 407 function (similar to the compression threshold) in 3-dB steps, as illustrated in Figure 5. This is done
49
50 408 separately for each of four frequency ranges. For a “Low” RMS input level (60 dB SPL), the
51
52 409 kneepoint is set to its lowest value (ranging from 52 dB SPL for low center frequencies to 41 dB
53
54 410 SPL for high center frequencies). For a “High” RMS input level (80 dB SPL), the kneepoint is set
55
56 411 to its highest value (ranging from 70 dB SPL for low center frequencies to 58 dB SPL for high
57
58 412 center frequencies). The system incorporates a “hysteresis” function that prevents the knee-point
59
60 413 from fluctuating wildly when the input level is halfway between the values corresponding to two

1
2
3
4 414 knee-points. The output electrical signal at the kneepoint is always set to 75% of the electrical
5
6 415 dynamic range between the threshold (T) and the highest comfortable level (C) for each channel.

7
8 416 As noted above, the Voice Guard system was designed with the intention of improving the
9
10 417 perception of speech, especially speech in background noise. As far as I know, it has not been
11
12 418 evaluated or optimized for music listening. However, since the mapping in the individual channels
13
14 419 is controlled by the slowly-changing RMS level of the broadband signal, the Voice Guard system is
15
16 420 likely to suffer from some of the same problems as for the front-end AGCs used in other CI
17
18 421 systems. Specifically, if the RMS level of the broadband signal increases, the electrical output
19
20 422 decreases for all channels, regardless of whether the increase in RMS level was produced by a
21
22 423 narrowband signal (e.g. a strong low-frequency note) or by a broadband signal (e.g. a crescendo in
23
24 424 the music). Thus the system will suffer from cross-modulation effects. Also, in the Voice Guard
25
26 425 system all of the compression of dynamic range is performed in the instantaneous mapping from
27
28 426 channel envelope magnitude to electrical output. The research reviewed above for HAs suggest that
29
30 427 for music listening slow compression is preferred over fast compression. Therefore, performing all
31
32 428 of the compression via instantaneous mapping is unlikely to be optimal for music listening. Finally,
33
34 429 because all of the compression is performed via the mapping, very strong compression has to be
35
36 430 used in the mapping. For example, for a high RMS input level of 80 dB SPL, the kneepoint for a
37
38 431 channel channel centered close to 2 kHz would be 66 dB. This means that all levels above 66 dB in
39
40 432 that channel would be mapped into only 25% of the electrical dynamic range. Potentially, this could
41
42 433 reduce the detectability of small envelope fluctuations, although the ability to detect small changes
43
44 434 in electric current is usually best towards the upper end of the electrical dynamic range (Shannon,
45
46 435 1983; McKay, Henshall, Farrell & McDermott, 2003).

47
48 436 Improvements in the Voice Guard system for music listening might be produced by:

- 49
50 437 (1) Controlling the mapping in individual channels by estimating the running RMS level in
51
52 438 different frequency regions and using that to control the mapping in sub-groups of electrodes, rather
53
54 439 than by controlling the mapping in all channels by the running RMS level of the broadband signal.
55
56 440 For example, the running RMS level at low frequencies would be used to control the mapping for
57
58 441 the most apical electrodes.
59
60 442 (2) Incorporating a multi-channel slow-acting or dual-time-constant AGC system prior to the

1
2
3
4 443 instantaneous channel mapping.
5

6 444
7
8 445 *Approaches Based on Loudness Models*
9
10 446 Although not implemented in current CI devices, some researchers have explored schemes
11 for CI processing based on loudness models, with aim of restoring the perception of loudness
12 approximately to “normal” for a wide range of input levels and spectra (McKay et al., 2003;
13 McDermott, McKay, Richardson & Henshall, 2003; Francart & McDermott, 2012); for a review,
14 see Wouters, McDermott and Francart (2015). Such schemes are based on loudness models for
15 normal hearing (Moore, 2014 #6652) and use the concept of specific loudness, which is the
16 loudness density as a function of center frequency (Moore, 2014). The aim is to restore the specific
17 loudness pattern to “normal”. The loudness-based schemes have shown promising results for
18 speech in quiet and in noise and for artificial test signals, but they have not, to my knowledge, been
19 comprehensively evaluated using music.
20
21
22
23
24
25
26
27
28
29
30
31

32 456

32 457 **Issues Associated with Binaural CIs**

33
34 458 Many people with bilateral severe or profound hearing loss are fitted with bilateral CIs. The
35
36 459 use of bilateral CIs can improve the ability to understand speech in the presence of background
37
38 460 sounds and can also improve the ability to localize sounds (Dunn, Tyler, Oakley, Gantz & Noble,
39
40 461 2008; Litovsky, Parkinson & Arcaroli, 2009; Culling, Jelfs, Talbert, Grange & Backhouse, 2012).
41
42 462 The improvement in localization appears to depend largely on the use of interaural level difference
43
44 463 (ILD) cues (Seeber & Fastl, 2008; Litovsky et al., 2009). Currently, the slowly-adapting AGC
45
46 464 systems that are used in CIs operate independently at the two ears. The use of independent AGC at
47
48 465 the two ears distorts interaural level difference (ILD) cues for sound localization (Wiggins &
49
50 466 Seeber, 2011), especially when the head is moved (Archer-Boyd & Carlyon, 2019; Archer-Boyd &
51
52 467 Carlyon, 2021). The magnitude of the effect of head movement varies in a complex way with the
53
54 468 attack and release time of the AGC system and with the speed and type of movement of the head of
55
56 469 the CI user. This disrupts ILD cues for sound localization, especially dynamic cues.

57
58 470 In principle, the distortion of dynamic ILD cues can be reduced by linking the AGC across
59
60 471 ears, as is done in some HAs. When the AGC is linked, the same gain-control signal is applied to

1
2
3
4 472 the signal at each ear, preserving ILD cues. The evidence for benefits of linked multi-channel AGC
5
6 473 in HAs is mixed, and when slow-acting AGC is used the benefits appear to be minimal (Wiggins &
7
8 474 Seeber, 2013; Moore, Kolarik, Stone & Lee, 2016). One study has reported benefits of
9
10 475 synchronization of the front-end AGC in experimental versions of bilaterally fitted Advanced
11
12 476 Bionics CIs for the localization of both static and moving sound sources (Dwyer, Chen, Hehrmann,
13
14 477 Dwyer & Gifford, 2021). However, in that study the participants did not move their heads. Also, it
15
16 478 is not clear whether the benefits found were a result of better preservation of ILD cues because of
17
18 479 the synchronization, or whether they resulted from the fact that synchronization of AGC systems
19
20 480 generally results in less overall compression and more slowly changing compression. Finally, it
21
22 481 should be noted that linking of AGC systems distorts the trajectory of the changes in level at each
23
24 482 ear as the user moves their head (Archer-Boyd & Carlyon, 2021); in other words, monaural
25
26 483 envelope cues are distorted. Clearly, more research is needed into the potential benefits of linking
27
28 484 the AGC systems in bilaterally fitted CIs, especially in the context of music perception.
29
30 485

31 486 **Conclusions**

32
33
34 487 This paper has reviewed the AGC systems used in HAs and has considered how some of the
35
36 488 lessons learned from studies of the AGC systems in HAs might be applied to improving music
37
38 489 listening for users of CIs. The following recommendations are made:

- 39
40 490 1) A CI should be set up with a dedicated music program and any adaptive directional processing
41
42 491 and noise reduction should be disabled for that program.
- 43
44 492 2) The dynamic range of the input stage of CIs should be increased, either by the use of ADCs with
45
46 493 more resolution than 16 bits or via the use of a controllable gain applied to the microphone
47
48 494 signal(s) before analog-to-digital conversion.
- 49
50 495 3) There is a need for systematic studies exploring the effect on music perception of varying the
51
52 496 compression threshold and the compression ratio of the front-end AGC used in the CIs of
53
54 497 ~~Cochlear Corporation, Advanced Bionics, and Med-El~~ several manufacturers. It seems plausible
55
56 498 that a lower compression threshold and a lower compression ratio than currently used would
57
58 499 give a better impression of the dynamic changes in a piece of music, but this remains to be
59
60 500 assessed, using both recorded and live music.

- 1
2
3
4 501 4) Experimental studies comparing multi-channel and single-channel front-end AGC systems for
5
6 502 music listening should be conducted. In principle the use of multi-channel AGC would prevent
7
8 503 an intense narrowband signal from reducing the effective level of all frequency components, and
9
10 504 it would also, to some extent, reduce cross-modulation effects between different sound sources.
11
12 505 5) For dual time-constant AGC systems, the effect of making the slow time constants even slower
13
14 506 should be explored.
15
16 507 6) For the ~~Voice-Guard~~ system ~~used by Oticon Medical~~ in which the running RMS level of input
17
18 508 signal is used to control the kneepoint of the mapping function, the potential benefits should be
19
20 509 explored of controlling the mapping in individual channels by estimating the running RMS level
21
22 510 in different frequency regions and using that to control the mapping in sub-groups of electrodes,
23
24 511 rather than by controlling the mapping in all channels by the running RMS level of the
25
26 512 broadband signal. As for multi-channel AGC, this should prevent an intense narrowband signal
27
28 513 from reducing the effective level of all frequency components, and it should also, to some extent,
29
30 514 reduce cross-modulation effects between different sound sources.
31
32 515 7) For the the system in which the running RMS level of input signal is used to control the
33
34 516 kneepoint of the mapping function ~~Voice-Guard system used by Oticon Medical~~, the potential
35
36 517 benefits should be explored of incorporating a multi-channel slow-acting or dual-time-constant
37
38 518 AGC system prior to the instantaneous channel mapping. This would reduce the amount of
39
40 519 instantaneous compression required for the mapping in each channel, helping to preserve short-
41
42 520 term intensity contrasts and amplitude-modulation patterns.
43
44 521 8) Research is needed into the potential benefits of linking the AGC systems in bilaterally fitted
45
46 522 CIs, especially in the context of music perception.
47
48 523 9) Users of bilateral CIs should be cautioned that rapid head movements might affect the apparent
49
50 524 positions of sounds in space. They should be advised to move their heads only slowly when
51
52 525 listening to music.
53
54 526

527 Acknowledgments

528 I thank Paddy Boyle ~~and~~, Alan Archer-Boyd, Andrew Oxenham, and two reviewers for
529 helpful comments on an earlier version of this paper.

530

531 Declaration of Conflicting Interests

532 The author has no conflicting interests to declare.

533

534 Funding

535 The author disclosed receipt of the following financial support for the research, authorship,
536 and/or publication of this article: This work was supported by the Medical Research Council (grant
537 G0701870).

538

539 References

540
541 ANSI (2014). *ANSI S3.22-2014, Specification of hearing aid characteristics*. New York: American
542 National Standards Institute.

543 Archer-Boyd, A. & Carlyon, R. P. (2021). Further simulations of the effect of cochlear-implant pre-
544 processing and head movement on interaural level differences. *The Journal of the Acoustical*
545 *Society of America*, 150, 506-525. doi: 10.1121/10.0005647

546 Archer-Boyd, A. W. & Carlyon, R. P. (2019). Simulations of the effect of unlinked cochlear-
547 implant automatic gain control and head movement on interaural level differences. *The Journal*
548 *of the Acoustical Society of America*, 145, 1389. doi: 10.1121/1.5093623

549 Blamey, P. J. (2005). Adaptive dynamic range optimization (ADRO): a digital amplification
550 strategy for hearing aids and cochlear implants. *Trends in Amplification*, 9, 77-98. doi:
551 10.1177/108471380500900203

552 Boyle, P. J., Buchner, A., Stone, M. A., Lenarz, T. & Moore, B. C. J. (2009). Comparison of dual-
553 time-constant and fast-acting automatic gain control (AGC) systems in cochlear implants.
554 *International Journal of Audiology*, 48, 211-21. doi: 10.1080/14992020802581982

555 Brons, I., Houben, R. & Dreschler, W. A. (2012). Perceptual effects of noise reduction by time-
556 frequency masking of noisy speech. *The Journal of the Acoustical Society of America*, 132,
557 2690-2699. doi: 10.1121/1.4747006

558 Byrne, D. & Dillon, H. (1986). The National Acoustic Laboratories' (NAL) new procedure for
559 selecting the gain and frequency response of a hearing aid. *Ear and Hearing*, 7, 257-265. doi:
560 10.1097/00003446-198608000-00007

Moore Lessons from hearing aids 21

- 1
2
3 561 Byrne, D., Dillon, H., Ching, T., Katsch, R. & Keidser, G. (2001). NAL-NL1 procedure for fitting
4 562 nonlinear hearing aids: characteristics and comparisons with other procedures. *Journal of the*
5 563 *American Academy of Audiology*, 12, 37-51.
6
7 564 Chasin, M. & Hockley, N. S. (2014). Some characteristics of amplified music through hearing aids.
8 565 *Hearing Research*, 308, 2-12. doi: 10.1016/j.heares.2013.07.003
9
10 566 Clark, G. M., Tong, Y. C. & Patrick, J. F. (1990). *Cochlear Prostheses*. Edinburgh: Churchill
11 567 Livingston.
12 568 Croghan, N. B., Arehart, K. H. & Kates, J. M. (2014). Music preferences with hearing aids: effects
13 569 of signal properties, compression settings, and listener characteristics. *Ear and Hearing*, 35,
14 570 e170-184. doi: 10.1097/AUD.0000000000000056
15 571 Culling, J. F., Jelfs, S., Talbert, A., Grange, J. A. & Backhouse, S. S. (2012). The benefit of
16 572 bilateral versus unilateral cochlear implantation to speech intelligibility in noise. *Ear and*
17 573 *Hearing*, 33, 673-682. doi: 10.1097/AUD.0b013e3182587356
18 574 Dunn, C. C., Tyler, R. S., Oakley, S., Gantz, B. J. & Noble, W. (2008). Comparison of speech
19 575 recognition and localization performance in bilateral and unilateral cochlear implant users
20 576 matched on duration of deafness and age at implantation. *Ear and Hearing*, 29, 352-359. doi:
21 577 10.1097/AUD.0b013e318167b870
22 578 Dwyer, R. T., Chen, C., Hehrmann, P., Dwyer, N. C. & Gifford, R. H. (2021). Synchronized
23 579 automatic gain control in bilateral cochlear implant recipients yields significant benefit in static
24 580 and dynamic listening conditions. *Trends in Hearing*, 25, 23312165211014139. doi:
25 581 10.1177/23312165211014139
26 582 Fourcin, A. J., Rosen, S. M., Moore, B. C. J. et al. (1979). External electrical stimulation of the
27 583 cochlea: clinical, psychophysical, speech-perceptual and histological findings. *British Journal*
28 584 *of Audiology*, 13, 85-107. doi: 10.3109/03005367909078883
29 585 Fowler, E. P. (1936). A method for the early detection of otosclerosis. *Archives of Otolaryngology*,
30 586 24, 731-741.
31 587 Francart, T. & McDermott, H. (2012). Speech perception and localisation with SCORE bimodal: a
32 588 loudness normalisation strategy for combined cochlear implant and hearing aid stimulation.
33 589 *PLoS One*, 7, e45385. doi: 10.1371/journal.pone.0045385
34 590 Friesen, L. M., Shannon, R. V., Baskent, D. & Wang, X. (2001). Speech recognition in noise as a
35 591 function of the number of spectral channels: comparison of acoustic hearing and cochlear
36 592 implants. *The Journal of the Acoustical Society of America*, 110, 1150-1163. doi:
37 593 10.1121/1.1381538
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 594 Füllgrabe, C., Moore, B. C. J. & Stone, M. A. (2015). Age-group differences in speech
4 595 identification despite matched audiometrically normal hearing: Contributions from auditory
5 596 temporal processing and cognition. *Frontiers in Aging Neuroscience*, 6, Article 347, 1-25. doi:
6 597 10.3389/fnagi.2014.00347
- 7 598 Gilbert, M. L., Deroche, M. L. D., Jiradejvong, P., Chan Barrett, K. & Limb, C. J. (2021). Cochlear
8 599 implant compression optimization for musical sound quality in MED-EL users. *Ear and*
9 600 *Hearing*, (in press). doi: 10.1097/AUD.0000000000001145
- 10 601 Glasberg, B. R. & Moore, B. C. J. (1986). Auditory filter shapes in subjects with unilateral and
11 602 bilateral cochlear impairments. *The Journal of the Acoustical Society of America*, 79, 1020-
12 603 1033. doi: 10.1121/1.393374
- 13 604 Goehring, T., Keshavarzi, M., Carlyon, R. P. & Moore, B. C. J. (2019). Using recurrent neural
14 605 networks to improve the perception of speech in non-stationary noise by people with cochlear
15 606 implants. *The Journal of the Acoustical Society of America*, 146, 705-718. doi:
16 607 10.1121/1.5119226
- 17 608 Goehring, T., Bolner, F., Monaghan, J. J., van Dijk, B., Zarowski, A. & Bleeck, S. (2017). Speech
18 609 enhancement based on neural networks improves speech intelligibility in noise for cochlear
19 610 implant users. *Hearing Research*, 344, 183-194. doi: 10.1016/j.heares.2016.11.012
- 20 611 Hamacher, V., Fischer, E., Kornagel, U. & Puder, H. (2006). Applications of adaptive signal
21 612 processing methods in high-end hearing instruments. In E. Hänsler & G. Schmidt (Eds.).
22 613 *Topics in Acoustic Echo and Noise Control: Selected Methods for the Cancellation of*
23 614 *Acoustical Echoes, the Reduction of Background Noise, and Speech Processing* (pp. 599-636).
24 615 New York: Springer.
- 25 616 Hansen, M. (2002). Effects of multi-channel compression time constants on subjectively perceived
26 617 sound quality and speech intelligibility. *Ear and Hearing*, 23, 369-380. doi:
27 618 10.1097/00003446-200208000-00012
- 28 619 Healy, E. W., Tan, K., Johnson, E. M. & Wang, D. (2021). An effectively causal deep learning
29 620 algorithm to increase intelligibility in untrained noises for hearing-impaired listeners. *The*
30 621 *Journal of the Acoustical Society of America*, 149, 3943-3953. doi: 10.1121/10.0005089
- 31 622 Heinz, M. G., Colburn, H. S. & Carney, L. H. (2001a). Evaluating auditory performance limits: I.
32 623 One-parameter discrimination using a computational model for the auditory nerve. *Neural*
33 624 *Computation*, 13, 2273-2316. doi: 10.1162/089976601750541804
- 34 625 Heinz, M. G., Colburn, H. S. & Carney, L. H. (2001b). Evaluating auditory performance limits: II.
35 626 One-parameter discrimination with random-level variation. *Neural Computation*, 13, 2317-
36 627 2338. doi: 10.1162/089976601750541813

Moore Lessons from hearing aids 23

- 1
2
3 628 Hjortkjær, J. & Walther-Hansen, M. (2014). Perceptual effects of dynamic range compression in
4 629 popular music recordings. *Journal of the Audio Engineering Society*, 62, 37-41. doi:
5 630 10.17743/jaes.2014.0003
6
7
8 631 Hockley, N. S., Bahlmann, F. & Fulton, B. (2012). Analog-to-digital conversion to accommodate
9 632 the dynamics of live music in hearing instruments. *Trends in Amplification*, 16, 146-158. doi:
10 633 10.1177/1084713812471906
11
12
13 634 Holube, I., Hamacher, V. & Wesselkamp, M. (1999). Hearing Instruments: noise reduction
14 635 strategies. *Proc. 18th Danavox Symposium: Auditory Models and Non-linear Hearing*
15 636 *Instruments*, 359-377.
16
17
18 637 Keidser, G., Dillon, H., Carter, L. & O'Brien, A. (2012). NAL-NL2 Empirical Adjustments. *Trends*
19 638 *in Amplification*, 16, 211-223. doi: 10.1177/1084713812468511
20
21
22 639 Keshavarzi, M., Goehring, T., Turner, R. E. & Moore, B. C. J. (2019). Comparison of effects on
23 640 subjective speech intelligibility and sound quality of two algorithms for reducing babble: A
24 641 deep recurrent neural network and spectral subtraction. *The Journal of the Acoustical Society*
25 642 *of America*, 145, 1493-1503. doi: 10.1121/1.5094765
26
27
28 643 Khing, P. P., Swanson, B. A. & Ambikairajah, E. (2013). The effect of automatic gain control
29 644 structure and release time on cochlear implant speech intelligibility. *PLoS One*, 8, e82263. doi:
30 645 10.1371/journal.pone.0082263
31
32
33 646 Kirchberger, M. & Russo, F. A. (2016). Dynamic range across music genres and the perception of
34 647 dynamic compression in hearing-impaired listeners. *Trends in Hearing*, 20, 1-16. doi:
35 648 10.1177/2331216516630549
36
37
38 649 Langner, F., Buchner, A. & Nogueira, W. (2020). Evaluation of an adaptive dynamic compensation
39 650 system in cochlear implant listeners. *Trends in Hearing*, 24, 2331216520970349. doi:
40 651 10.1177/2331216520970349
41
42
43 652 Launer, S., Zakis, J. & Moore, B. C. J. (2016). Hearing aid signal processing. In G. R. Popelka, B.
44 653 C. J. Moore, A. N. Popper & R. R. Fay (Eds.). *Hearing Aids* (pp. 93-130). New York:
45 654 Springer.
46
47
48 655 Limb, C. J. & Roy, A. T. (2014). Technological, biological, and acoustical constraints to music
49 656 perception in cochlear implant users. *Hearing Research*, 308, 13-26. doi:
50 657 10.1016/j.heares.2013.04.009
51
52
53 658 Litovsky, R. Y., Parkinson, A. & Arcaroli, J. (2009). Spatial hearing and speech intelligibility in
54 659 bilateral cochlear implant users. *Ear and Hearing*, 30, 419-431. doi:
55 660 10.1097/AUD.0b013e3181a165be
56
57
58
59
60

- 1
2
3 661 Looi, V., McDermott, H., McKay, C. & Hickson, L. (2008). Music perception of cochlear implant
4 users compared with that of hearing aid users. *Ear and Hearing*, 29, 421-434. doi:
5 662 10.1097/AUD.0b013e31816a0d0b
6 663
7
8 664 Lunner, T., Hietkamp, R. K., Andersen, M. R., Hopkins, K. & Moore, B. C. J. (2012). Effect of
9 speech material on the benefit of temporal fine structure information in speech for young
10 665 normal-hearing and older hearing-impaired participants. *Ear and Hearing*, 33, 377-388. doi:
11 666 10.1097/AUD.0b013e3182387a8c
12 667
13 668 Madsen, S. M. K. & Moore, B. C. J. (2014). Music and hearing aids. *Trends in Hearing*, 18, 1-29.
14 669 doi: 10.1177/2331216514558271
15
16 670 McDermott, H. J. (2004). Music perception with cochlear implants: a review. *Trends in*
17 671 *Amplification*, 8, 49-82. doi: 10.1177/108471380400800203
18
19 672 McDermott, H. J., McKay, C. M., Richardson, L. M. & Henshall, K. R. (2003). Application of
20 673 loudness models to sound processing for cochlear implants. *The Journal of the Acoustical*
21 674 *Society of America*, 114, 2190-2197. doi: 10.1121/1.1612488
22
23 675 McKay, C. M., Henshall, K. R., Farrell, R. J. & McDermott, H. J. (2003). A practical method of
24 676 predicting the loudness of complex electrical stimuli. *The Journal of the Acoustical Society of*
25 677 *America*, 113, 2054-2063. doi: 10.1121/1.1558378
26
27 678 Mehta, A. H. & Oxenham, A. J. (2017). Vocoder simulations explain complex pitch perception
28 679 limitations experienced by cochlear implant users. *Journal of the Association for Research in*
29 680 *Otolaryngology*, 18, 789-802. doi: 10.1007/s10162-017-0632-x
30
31 681 Miskolczy-Fodor, F. (1960). Relation between loudness and duration of tonal pulses. III. Response
32 682 in cases of abnormal loudness function. *The Journal of the Acoustical Society of America*, 32,
33 683 486-492. doi: 10.1121/1.1908107
34
35 684 Moore, B. C. J. (1973). Frequency difference limens for short-duration tones. *The Journal of the*
36 685 *Acoustical Society of America*, 54, 610-619. doi: 10.1121/1.1913640
37
38 686 Moore, B. C. J. (2003). Coding of sounds in the auditory system and its relevance to signal
39 687 processing and coding in cochlear implants. *Otology & Neurotology*, 24, 243-254. doi:
40 688 10.1097/00129492-200303000-00019
41
42 689 Moore, B. C. J. (2004). Testing the concept of softness imperception: Loudness near threshold for
43 690 hearing-impaired ears. *The Journal of the Acoustical Society of America*, 115, 3103-3111. doi:
44 691 10.1121/1.1738839
45
46 692 Moore, B. C. J. (2007). *Cochlear Hearing Loss: Physiological, Psychological and Technical*
47 693 *Issues, 2nd Ed.* Chichester: Wiley.

Moore Lessons from hearing aids 25

- 1
2
3 694 Moore, B. C. J. (2008). The choice of compression speed in hearing aids: Theoretical and practical
4 considerations, and the role of individual differences. *Trends in Amplification*, 12, 103-112.
5 695 doi: 10.1177/1084713808317819
6 696
7
8 697 Moore, B. C. J. (2012). *An Introduction to the Psychology of Hearing, 6th Ed.* Leiden, The
9 Netherlands: Brill.
10 698
11 699 Moore, B. C. J. (2014). Development and current status of the "Cambridge" loudness models.
12 *Trends in Hearing*, 18, 1-29.
13 700
14 701 Moore, B. C. J. (2016). Effects of sound-induced hearing loss and hearing aids on the perception of
15 music. *Journal of the Audio Engineering Society*, 64, 112-123. doi: 10.17743/jaes.2015.0081
16 702
17 703 Moore, B. C. J. & Glasberg, B. R. (1988). A comparison of four methods of implementing
18 automatic gain control (AGC) in hearing aids. *British Journal of Audiology*, 22, 93-104. doi:
19 704 10.3109/03005368809077803
20 705
21 706 Moore, B. C. J. & Glasberg, B. R. (2004). A revised model of loudness perception applied to
22 cochlear hearing loss. *Hearing Research*, 188, 70-88. doi: 10.1016/S0378-5955(03)00347-2
23 707
24 708 Moore, B. C. J. & Ernst, S. M. A. (2012). Frequency difference limens at high frequencies:
25 Evidence for a transition from a temporal to a place code. *The Journal of the Acoustical*
26 *Society of America*, 132, 1542-1547. doi: 10.1121/1.4739444
27 709
28 710 Moore, B. C. J. & Şek, A. (2013). Comparison of the CAM2 and NAL-NL2 hearing-aid fitting
29 methods. *Ear and Hearing*, 34, 83-95. doi: 10.1097/AUD.0b013e3182650adf
30 711
31 712 Moore, B. C. J. & Şek, A. (2016a). Comparison of the CAM2A and NAL-NL2 hearing-aid fitting
32 methods for participants with a wide range of hearing losses. *International Journal of*
33 *Audiology*, 55, 93-100. doi: 10.3109/14992027.2015.1095360
34 713
35 714 Moore, B. C. J. & Şek, A. P. (2016b). Preferred compression speed for speech and music and its
36 relationship to sensitivity to temporal fine structure. *Trends in Hearing*, 20, 1-15. doi:
37 715 10.1177/2331216516640486
38 716
39 717 Moore, B. C. J., Glasberg, B. R. & Stone, M. A. (1991). Optimization of a slow-acting automatic
40 gain control system for use in hearing aids. *British Journal of Audiology*, 25, 171-182. doi:
41 718 10.3109/03005369109079851
42 719
43 720 Moore, B. C. J., Stone, M. A. & Alcántara, J. I. (2001). Comparison of the electroacoustic
44 characteristics of five hearing aids. *British Journal of Audiology*, 35, 307-325. doi:
45 721 10.1080/00305364.2001.11745249
46 722
47 723 Moore, B. C. J., Glasberg, B. R. & Stone, M. A. (2010). Development of a new method for deriving
48 initial fittings for hearing aids with multi-channel compression: CAMEQ2-HF. *International*
49 *Journal of Audiology*, 49, 216-227. doi: 10.3109/14992020903296746
50 724
51 725
52 726
53 727

- 1
2
3 728 Moore, B. C. J., Füllgrabe, C. & Stone, M. A. (2011). Determination of preferred parameters for
4 729 multi-channel compression using individually fitted simulated hearing aids and paired
5 730 comparisons. *Ear and Hearing*, 32, 556-568. doi: 10.1097/AUD.0b013e31820b5f4c
6
7 731 Moore, B. C. J., Kolarik, A., Stone, M. A. & Lee, Y.-W. (2016). Evaluation of a method for
8 732 enhancing interaural level differences at low frequencies. *The Journal of the Acoustical Society*
9 733 *of America*, 140, 2817-2828. doi: 10.1121/1.4965299
10
11 734 Nordqvist, P. & Leijon, A. (2004). Hearing-aid automatic gain control adapting to two sound
12 735 sources in the environment, using three time constants. *The Journal of the Acoustical Society of*
13 736 *America*, 116, 3152-3155. doi: 10.1121/1.1793207
14
15 737 Pick, G., Evans, E. F. & Wilson, J. P. (1977). Frequency resolution in patients with hearing loss of
16 738 cochlear origin. In E. F. Evans & J. P. Wilson (Eds.). *Psychophysics and Physiology of*
17 739 *Hearing* (pp. 273-281). London: Academic Press.
18
19 740 Price, G. R. & Kalb, J. T. (1991). Insights into hazard from intense impulses from a mathematical
20 741 model of the ear. *The Journal of the Acoustical Society of America*, 90, 219-227. doi:
21 742 10.1121/1.401291
22
23 743 Robinson, C. E. & Huntington, D. A. (1973). The intelligibility of speech processed by delayed
24 744 long-term averaged compression amplification. *The Journal of the Acoustical Society of*
25 745 *America*, 54, 314.
26
27 746 Seeber, B. U. & Fastl, H. (2008). Localization cues with bilateral cochlear implants. *The Journal of*
28 747 *the Acoustical Society of America*, 123, 1030-1042. doi: 10.1121/1.2821965
29
30 748 Seligman, P. & Whitford, L. (1995). Adjustment of appropriate signal levels in the Spectra 22 and
31 749 mini speech processors. *Annals of Otology, Rhinology and Laryngology*, Suppl. 166, 172-175.
32
33 750 Shannon, R. V. (1983). Multichannel electrical stimulation of the auditory nerve in man. I. Basic
34 751 psychophysics. *Hearing Research*, 11, 157-189. doi: 10.1016/0378-5955(83)90077-1
35
36 752 Souza, P. E. (2002). Effects of compression on speech acoustics, intelligibility, and sound quality.
37 753 *Trends in Amplification*, 6, 131-165. doi: 10.1177/108471380200600402
38
39 754 Steinberg, J. C. & Gardner, M. B. (1937). The dependence of hearing impairment on sound
40 755 intensity. *The Journal of the Acoustical Society of America*, 9, 11-23. doi: 10.1121/1.1915905
41
42 756 Stöbich, B., Zierhofer, C. M. & Hochmair, E. S. (1999). Influence of automatic gain control
43 757 parameter settings on speech understanding of cochlear implant users employing the
44 758 continuous interleaved sampling strategy. *Ear and Hearing*, 20, 104-116. doi:
45 759 10.1097/00003446-199904000-00002
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Moore Lessons from hearing aids 27

- 1
2
3 760 Stone, M. A. & Moore, B. C. J. (1992). Syllabic compression: Effective compression ratios for
4 signals modulated at different rates. *British Journal of Audiology*, 26, 351-361. doi:
5 761 10.3109/03005369209076659
6 762
7
8 763 Stone, M. A. & Moore, B. C. J. (2004). Side effects of fast-acting dynamic range compression that
9 affect intelligibility in a competing speech task. *The Journal of the Acoustical Society of*
10 764 *America*, 116, 2311-2323. doi: 10.1121/1.1784447
11 765
12 766 Stone, M. A. & Moore, B. C. J. (2008). Effects of spectro-temporal modulation changes produced
13 by multi-channel compression on intelligibility in a competing-speech task. *The Journal of the*
14 767 *Acoustical Society of America*, 123, 1063-1076. doi: 10.1121/1.2821969
15 768
16 769 Stone, M. A., Moore, B. C. J., Alcántara, J. I. & Glasberg, B. R. (1999). Comparison of different
17 770 forms of compression using wearable digital hearing aids. *The Journal of the Acoustical*
18 771 *Society of America*, 106, 3603-3619. doi: 10.1121/1.428213
19 772
20 772 Stone, M. A., Moore, B. C. J., Füllgrabe, C. & Hinton, A. C. (2009). Multi-channel fast-acting
21 773 dynamic-range compression hinders performance by young, normal-hearing listeners in a two-
22 774 talker separation task. *Journal of the Audio Engineering Society*, 57, 532-546.
23 775
24 775 Studebaker, G. A., Sherbecoe, R. L., McDaniel, D. M. & Gwaltney, C. A. (1999). Monosyllabic
25 776 word recognition at higher-than-normal speech and noise levels. *The Journal of the Acoustical*
26 777 *Society of America*, 105, 2431-2444. doi: 10.1121/1.426848
27 778
28 778 Tan, C. T. & Moore, B. C. J. (2008). Perception of nonlinear distortion by hearing-impaired people.
29 779 *International Journal of Audiology*, 47, 246-256. doi: 10.1080/14992020801945493
30 780
31 780 Tan, C. T., Moore, B. C. J. & Zacharov, N. (2003). The effect of nonlinear distortion on the
32 781 perceived quality of music and speech signals. *Journal of the Audio Engineering Society*, 51,
33 782 1012-1031.
34 783
35 783 Vaerenberg, B., Govaerts, P. J., Stainsby, T., Nopp, P., Gault, A. & Gnansia, D. (2014). A uniform
36 784 graphical representation of intensity coding in current-generation cochlear implant systems.
37 785 *Ear and Hearing*, 35, 533-543. doi: 10.1097/AUD.0000000000000039
38 786
39 786 Vanderkooy, J. & Lipshitz, S. P. (1984). Resolution below the least significant bit in digital systems
40 787 with dither. *Journal of the Audio Engineering Society*, 32, 106-113.
41 788
42 788 Verschuure, J., Dreschler, W. A., de Haan, E. H. et al. (1993). Syllabic compression and speech
43 789 intelligibility in hearing impaired listeners. *Scandinavian Audiology*, 22, Suppl. 38, 92-100.
44 790
45 790 von Békésy, G. (1960). *Experiments in Hearing*. New York: McGraw-Hill.
46 791
47 791 Wiggins, I. M. & Seeber, B. U. (2011). Dynamic-range compression affects the lateral position of
48 792 sounds. *The Journal of the Acoustical Society of America*, 130, 3939-3953. doi:
49 793 10.1121/1.3652887
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 794 Wiggins, I. M. & Seeber, B. U. (2013). Linking dynamic-range compression across the ears can
4 795 improve speech intelligibility in spatially separated noise. *The Journal of the Acoustical*
5 796 *Society of America*, 133, 1004-1016. doi: 10.1121/1.4773862
- 6
7
8 797 Wilson, B. S., Finley, C. C., Lawson, D. T., Wolford, R. D., Eddington, D. K. & Rabinowitz, W.
9 798 M. (1991). Better speech recognition with cochlear implants. *Nature*, 352, 236-238. doi:
10 799 10.1038/352236a0
- 11
12
13 800 Wolfe, J., Schafer, E. C., John, A. & Hudson, M. (2011). The effect of front-end processing on
14 801 cochlear implant performance of children. *Otology & Neurotology*, 32, 533-538. doi:
15 802 10.1097/MAO.0b013e318210b6ec
- 16
17
18 803 Wouters, J., McDermott, H. J. & Francart, T. (2015). Sound coding in cochlear implants: From
19 804 electric pulses to hearing. *IEEE Signal Processing Magazine*, 32, 67-80. doi:
20 805 10.1109/MSP.2014.2371671
- 21
22
23 806 Yang, L. P. & Fu, Q. J. (2005). Spectral subtraction-based speech enhancement for cochlear
24 807 implant patients in background noise. *The Journal of the Acoustical Society of America*, 117,
25 808 1001-1004. doi: 10.1121/1.1852873
- 26
27
28 809 Zakis, J. (2016). Music perception and hearing aids. In G. R. Popelka, B. C. J. Moore, A. N. Popper
29 810 & R. R. Fay (Eds.). *Hearing Aids* (pp. 217-252). New York: Springer.
- 30
31
32 811 Zeng, F.-G. (2004). Compression and cochlear implants. In S. P. Bacon, A. N. Popper & R. R. Fay
33 812 (Eds.). *Compression: From Cochlea to Cochlear Implants* (pp. 184-220). New York: Springer.
- 34
35
36 813 Zeng, F.-G., Popper, A. N. & Fay, R. R. (2003). *Auditory Prostheses*. New York: Springer-Verlag.
- 37
38 814 Zeng, F. G. & Shannon, R. V. (1995). Loudness of simple and complex stimuli in electrical
39 815 hearing. *Annals of Otology, Rhinology and Laryngology*, 104, Suppl. 166, 235-238.
- 40
41 816 Zeng, F. G., Galvin, J. J. & Zhang, C. Y. (1998). Encoding loudness by electric stimulation of the
42 817 auditory nerve. *Neuroreport*, 9, 1845-1848. doi: 10.1097/00001756-199806010-00033

818

819

1
2
3 820 Figure captions
4

5
6 821 **Figure 1.** Preferences for slow versus fast compression for individual subjects and for the mean
7
8 822 across subjects, as obtained by Moore and Søk (2016b) using a simulated HA. Bars falling above
9
10 823 the horizontal line at the center of the panel indicate preferences for slow compression over fast
11
12 824 compression. Redrawn from Moore and Søk (2016b).

13
14 825 **Figure 2.** Simplified schematic diagram of the signal processing performed in several CIs.

15
16 826 **Figure 3.** Schematic input-output functions for the Advanced Bionics CIs for input dynamic ranges
17
18 827 (IDRs) of 50, 65, and 80 dB.

19
20 828 **Figure 4.** Schematic diagram of the “Voice Guard” system used in the Oticon Medical CI.

21
22 829 **Figure 5.** Schematic diagram of the adaptive mapping from the envelope magnitude in a given
23
24 830 channel to the electrical output for that channel, as used in the Oticon Medical Voice Guard system.

25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Peer Review

moderately

preferred

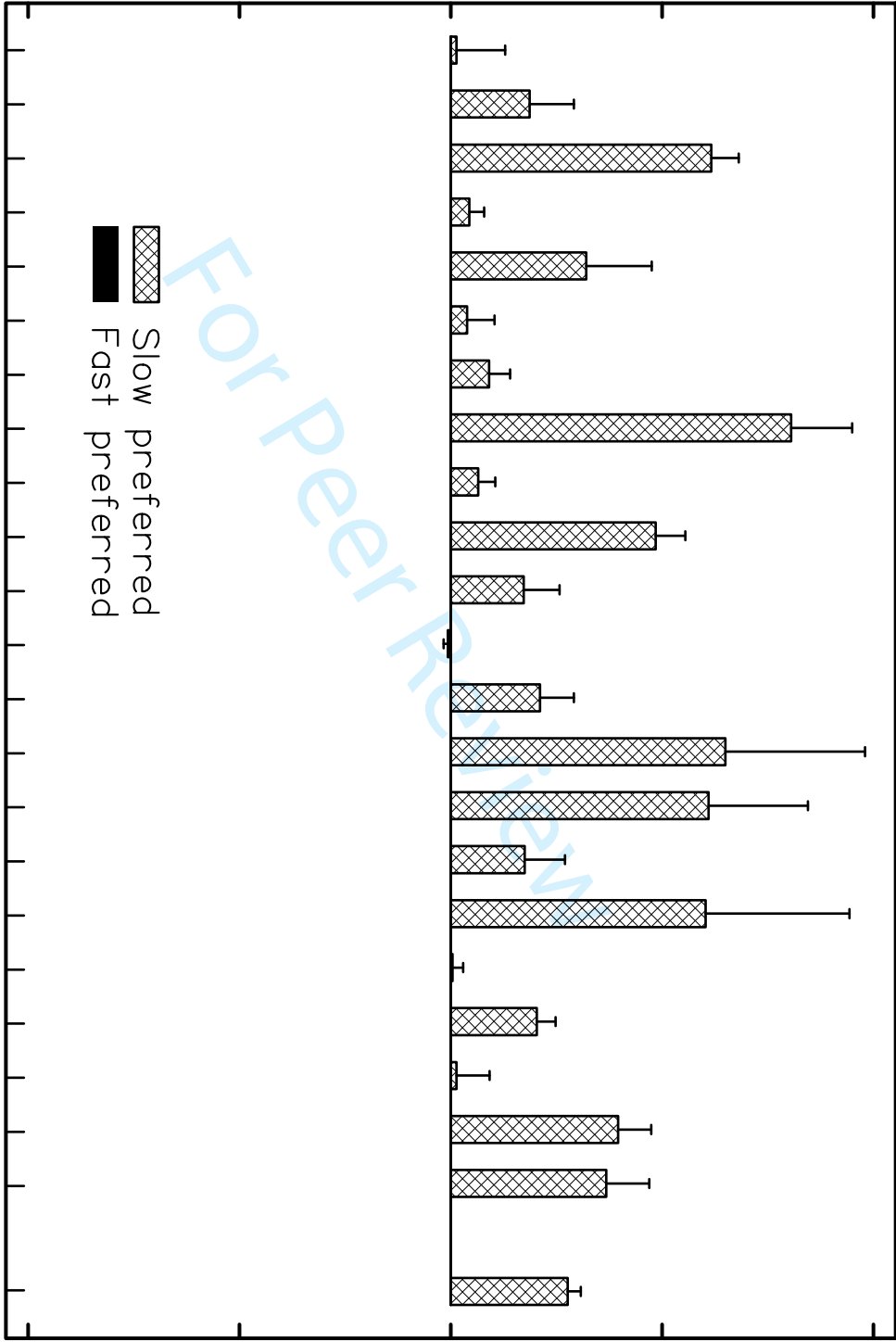
slightly

preferred

slightly

moderately

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54



Preference score

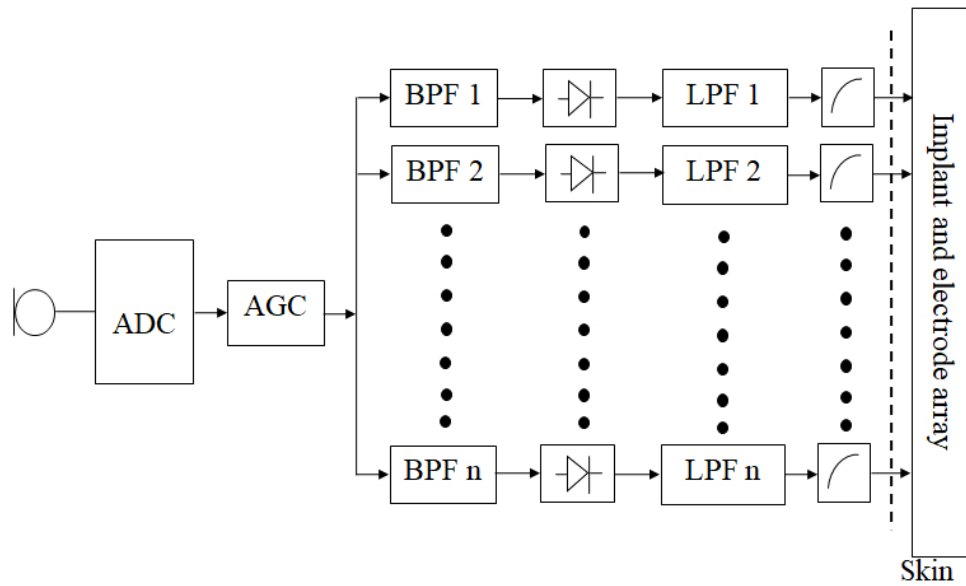


Figure 2. Simplified schematic diagram of the signal processing performed in several CIs.

118x72mm (168 x 168 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

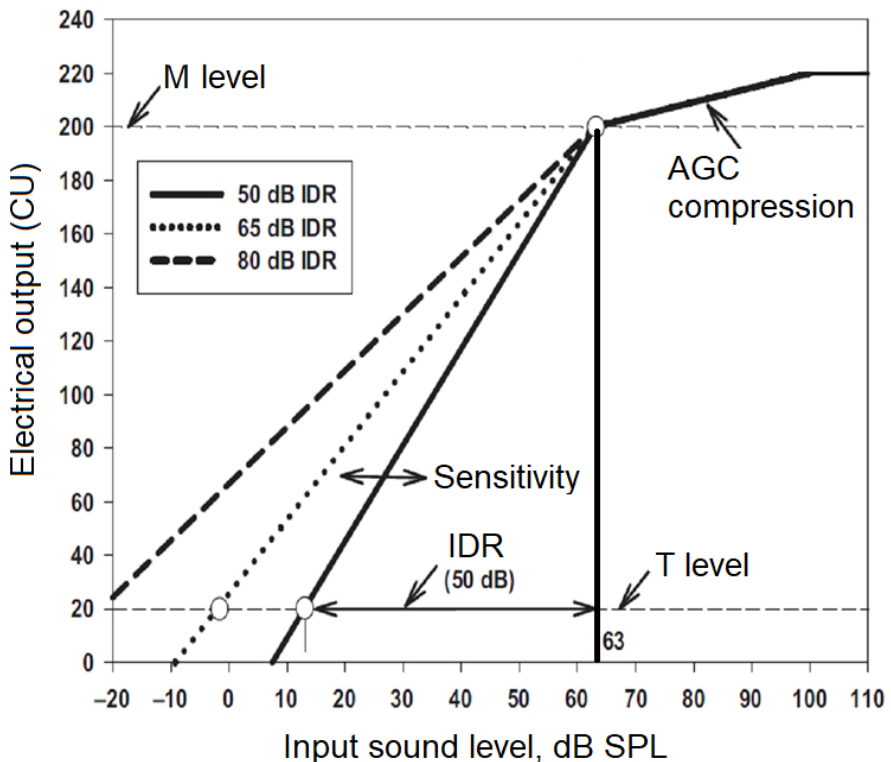


Figure 3. Schematic input-output functions for the Advanced Bionics CIs for input dynamic ranges (IDRs) of 50, 65, and 80 dB.

127x99mm (168 x 168 DPI)

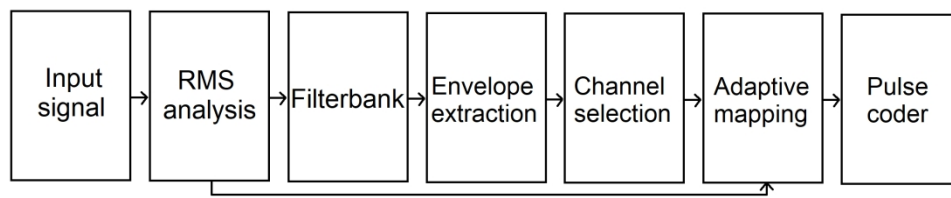


Figure 4. Schematic diagram of the "Voice Guard" system used in the Oticon Medical CI.

377x88mm (168 x 168 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

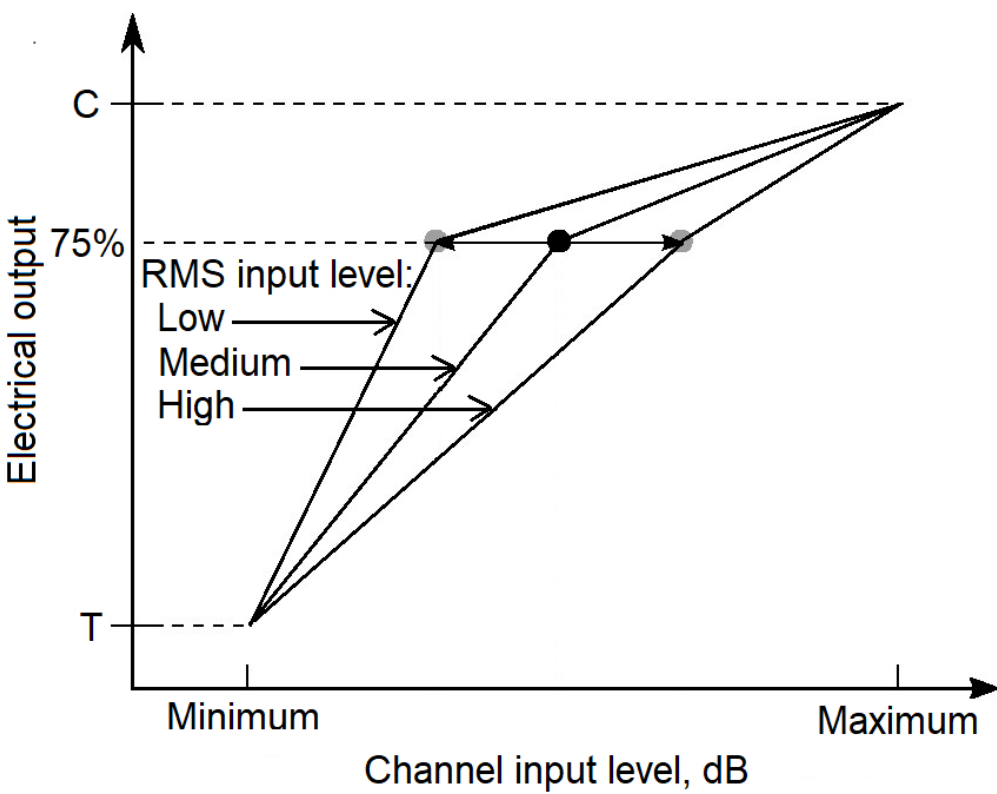


Figure 5. Schematic diagram of the adaptive mapping from the envelope magnitude in a given channel to the electrical output for that channel, as used in the Oticon Medical Voice Guard system.

128x103mm (168 x 168 DPI)