

OPEN

Electro-haptic enhancement of speech-in-noise performance in cochlear implant users

Mark D. Fletcher^{1,3}, Amatullah Hadeedi¹, Tobias Goehring¹  & Sean R. Mills¹ 

Cochlear implant (CI) users receive only limited sound information through their implant, which means that they struggle to understand speech in noisy environments. Recent work has suggested that combining the electrical signal from the CI with a haptic signal that provides crucial missing sound information (“electro-haptic stimulation”; EHS) could improve speech-in-noise performance. The aim of the current study was to test whether EHS could enhance speech-in-noise performance in CI users using: (1) a tactile signal derived using an algorithm that could be applied in real time, (2) a stimulation site appropriate for a real-world application, and (3) a tactile signal that could readily be produced by a compact, portable device. We measured speech intelligibility in multi-talker noise with and without vibro-tactile stimulation of the wrist in CI users, before and after a short training regime. No effect of EHS was found before training, but after training EHS was found to improve the number of words correctly identified by an average of 8.3%-points, with some users improving by more than 20%-points. Our approach could offer an inexpensive and non-invasive means of improving speech-in-noise performance in CI users.

When a sense is impaired or abolished, the brain adapts, relying more heavily on other senses to extract information^{1,2}. This cross-modal plasticity could be exploited to enhance listening in deaf and hearing-impaired individuals fitted with a cochlear implant (CI). A CI is a neural prosthesis that bypasses the damaged or defunct outer and inner ear and electrically stimulates auditory nerve fibres to restore hearing. While, in normal-hearing individuals, speech information is transmitted to the brain by thousands of hair cells, in CI users it is transmitted through, at most, just 22 micro-electrodes. Because of this, much of the key information used to extract speech from a noisy background is not available. CI users therefore struggle to understand speech-in-noise much more than normal-hearing listeners^{3,4}. Results from previous studies suggest that speech amplitude envelope information is a crucial missing feature for CI users listening in noise^{5,6}. We hypothesize that, given the brain's ability to extract missing information through another sense, speech-in-noise performance in CI users may be improved by providing speech envelope information through tactile stimulation. This haptic augmentation of the electrical signal from the CI will be referred to as “electro-haptic stimulation” (EHS).

Several studies have shown examples of sensory substitution, where information no longer transmitted through one sense is instead transmitted through another^{7–10}. For example, researchers have shown that tactile aids (with no concurrent audio signal) can be used to transmit complex speech information^{11–13}. Similarly, when a sense is impaired and therefore all of the critical information cannot be transferred, missing information can be provided through another sense^{2,14,15}. Two recent studies showed examples of this sensory augmentation for CI listening^{16,17}. Fletcher *et al.*¹⁶ found that tactile presentation of speech envelope improved speech intelligibility for normal-hearing subjects listening to CI simulations (NHICIs)¹⁶. Fletcher *et al.* extracted their tactile signal from speech-in-noise and presented it to the index finger. Huang *et al.* found similar results for CI listening in noise when the fundamental frequency was presented to the finger¹⁷, although in their study the tactile signal was extracted from the clean speech, rather than from the speech-in-noise. In the current study, we used an approach that is suitable for a real-world application. Three key features of our approach were: (1) as in Fletcher *et al.*, we extracted the speech envelope from the speech-in-noise signal using a signal processing approach that was

¹Faculty of Engineering and Physical Sciences, University of Southampton, University Road, Southampton, SO17 1BJ, United Kingdom. ²MRC Cognition and Brain Sciences Unit, University of Cambridge, 15 Chaucer Road, Cambridge, CB2 7EF, United Kingdom. ³University of Southampton Auditory Implant Service, University of Southampton, University Road, Southampton, SO17 1BJ, United Kingdom. Correspondence and requests for materials should be addressed to M.D.F. (email: M.D.Fletcher@soton.ac.uk)

Received: 19 December 2018

Accepted: 17 July 2019

Published online: 06 August 2019

computationally lightweight and appropriate for real-time processing; (2) we delivered tactile stimulation to the wrist rather than the finger, which is a more suitable site for real-world use; and (3) we used a vibration intensity that can be produced by a near-silent, inexpensive, compact device with low power consumption. The aim of the current study was to establish whether our approach can be used to improve speech-in-noise performance in CI users.

It is likely that a short training regime will increase the benefit of EHS to speech-in-noise performance. Fletcher *et al.*¹⁶ found that training was important; the benefit of tactile stimulation to speech recognition in noise was found to grow substantially after just 30 minutes of exposure to speech-in-noise and concurrent tactile stimulation. Fletcher *et al.* used NHCI users whereas, in the current study, CI users were tested. Evidence that CI users integrate auditory and lip-reading cues more effectively than NHCI users has been used to suggest that CI users are better at integrating multisensory information than normal-hearing listeners¹⁸. It may therefore be expected that CI users will benefit more from tactile stimulation. However, it is possible that lip-reading represents a special case as lip-reading cues are often used to aid speech understanding in real-world situations, whereas speech envelope information is never presented through touch in real-world settings. The novelty of the tactile stimulus could be distracting at first or less readily integrated with auditory information. It is therefore possible that training or familiarisation will be required to facilitate the multisensory integration needed for CI users to benefit from EHS.

To investigate the effect of EHS, 10 CI users completed a speech testing session, followed by two training sessions, and then a final testing session. In the testing sessions, participants completed a speech-in-noise task, where the percentage of correctly identified keywords in noise was measured both with and without concurrent tactile stimulation of the wrists. Speech testing was conducted with multi-talker noise, in which CI users are known to struggle most^{19,20}. In the training sessions, participants received speech-in-noise training with concurrent tactile stimulation. These training sessions used audiobook material with different talkers from the one used in the testing sessions. The two training sessions totalled just 20 minutes of exposure to speech-in-noise and tactile stimulation. It is anticipated that, after training, EHS will increase the percentage of keywords in noise that CI users are able to correctly identify.

Results

Figure 1 shows the effect of tactile stimulation on speech-in-noise performance (the percentage of key words correctly identified) before and after training. The results were analyzed using a repeated-measures ANOVA, with factors 'Session' (before or after training) and 'Condition' (with or without tactile stimulation). There was a significant interaction effect, indicating that the effect of tactile stimulation in the post-training session was significantly larger than in the pre-training session ($F(1,9) = 14.2, p = 0.004, \eta_p^2 = 0.917$). Main effects of Condition and Session were non-significant.

Paired t-tests (with a Bonferroni corrected alpha of 0.025), revealed a significant effect of Condition in the post-training session ($t(9) = 3.4, p = 0.008, d = 1.08$), but not in the pre-training session ($t(9) = 0.82, p = 0.43$). Before training, in 7 out of 10 participants' performance was worse with tactile stimulation, and individual results were highly variable. The largest improvement in performance with tactile stimulation before training was 12.8%-points (improving from 54.5% to 67.2%, $SE = 5.5, P9$), and the largest reduction in performance was 13.9%-points (decreasing from 36.1% to 22.2%, $SE = 5.8, P7$). There was also substantial variability within participants, with the difference in performance across sentence lists within a single session and condition being as much as 44.4%-points (the condition with tactile stimulation, session 1 for P6). After training, all ten participants correctly identified more keywords with tactile stimulation, with the smallest improvement being 0.5%-points (improving from 67.8% to 68.3%; $SE = 7.3\%, P8$) and the largest improvement being 21.8%-points (33.8% to 55.6%, $SE = 4.2\%, P1$). The mean improvement with tactile stimulation after training was 8.3%-points (improving from 54.3% to 62.5%; $SE = 2.5\%$). It should be noted that, while the order of the sentence lists used for the two conditions (with and without tactile stimulation) was counterbalanced across our 10 participants, the sentence lists used were not fully counterbalanced between the two sessions (before and after training).

Discussion

EHS was found to significantly improve speech-in-noise performance after a short training regime. The regime consisted of just 20 minutes of exposure to speech-in-noise material with concurrent tactile stimulation. Tactile stimulation increased the number of words correctly identified by 8.3%-points on average compared to audio alone, with some individuals improving by more than 20%-points and no individuals decreasing in performance. This benefit was observed for speech in multi-talker noise, where CI users are known to struggle most^{19,20}, and in which noise reduction algorithms typically perform poorly^{21,22}. The tactile signal was delivered at an intensity that could readily be produced by a low-cost wearable device after computationally non-intensive signal processing suitable for a real-time application.

Fletcher *et al.*¹⁶, who presented speech envelope through tactile stimulation to the index finger of NHCI users, found a slightly larger benefit of tactile stimulation for speech-in-noise performance than in the current study (10.8%-points compared to 8.3%-points, on average). The current study deployed tactile stimulation to the wrist, rather than the fingertip. The wrist, while being a more practical site for real-world use, has higher vibro-tactile thresholds than the fingertip²³ and a lower density of mechanoreceptors²⁴. However, researchers have shown better perception of amplitude differences at the wrist than at the fingertip as well as similar gap detection and discrimination of frequency differences²³. It is therefore possible that the wrist could be similarly capable of transferring speech information. Another difference between the studies was that the current study was conducted in CI users who were older than the NHCI users tested by Fletcher *et al.* (61.2 years on average compared to 25.5 years). Tactile sensitivity is known to be reduced in older adults^{25,26} and age also affects integration of information from more than one sense²⁷. These limitations may reduce the benefit of EHS in older CI users. Finally, there was less training (20 minutes compared to 30 minutes) in the current study. Training has been found to be highly

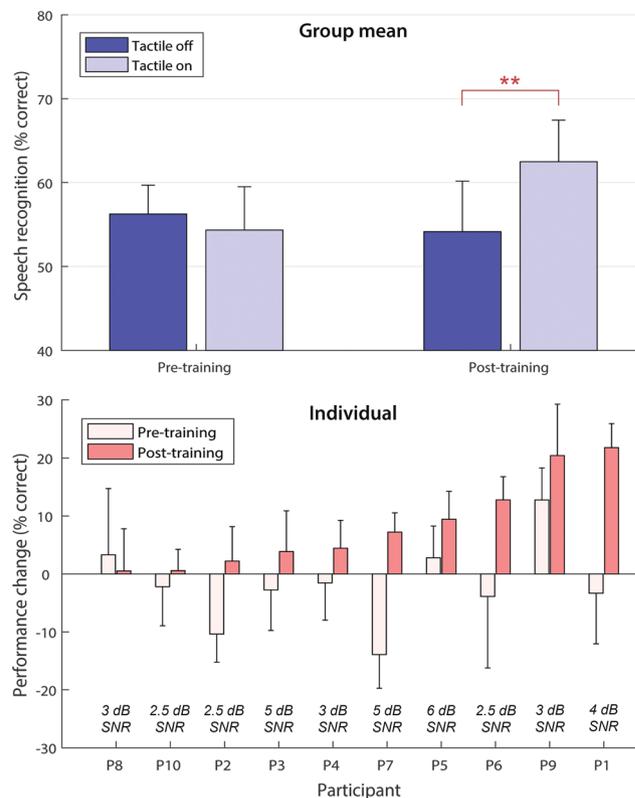


Figure 1. The top panel shows the mean speech-in-noise performance across all participants with and without tactile stimulation, both before and after training. The bottom panel shows the difference in performance with and without tactile stimulation for each individual, before and after training. A positive performance change indicates that performance was better with tactile stimulation. Participants are ordered by the size of their post-training performance change. The signal-to-noise ratio (SNR) at which speech-in-noise performance was measured for each individual is shown in the bottom panel. Error bars show the standard error of the mean (SE).

important in previous work in tactile aids (with no concurrent audio signal¹¹) and large training effects were found both in the current study and in Fletcher *et al.* It therefore seems likely that the smaller amount of training used here may have led to a smaller benefit from tactile stimulation. Another study, by Huang *et al.*¹⁷, reported robust benefits to speech-in-noise performance from EHS without training when the fundamental frequency of speech was derived from the clean speech signal and presented to the fingertip. However, the size of the benefit is difficult to compare directly to the current results because of the different outcome measures and speech corpora used.

There was a great deal of variation in the size of the EHS benefit between individuals, with the improvement in performance after training ranging from 0.5 to 21.8%-points. It is important for future research to establish which factors are most important in determining how much a CI user will benefit from EHS. For example, it should be established whether EHS can provide benefit in congenitally deaf individuals and those who became deaf prelingually. Interestingly, there is evidence of enhanced tactile sensitivity in congenitally deaf individuals²⁸ and that congenitally deaf individuals are able to effectively integrate audio and tactile information²⁹, suggesting that this group may benefit from EHS. It will also be important to establish which factors influence how much training is needed for EHS to be effective. Before training, there was large variability in the effect of EHS, with most participants' performance decreasing with tactile stimulation (by more than 10%-points in two participants; P2 and P7), and one participant's performance improving by 12.8%-points (P9). Both of the participants who had substantial decreases in performance with tactile stimulation before training reported to the experimenter that they found tactile stimulation distracting in the pre-training session. Intriguingly, the participant who improved with tactile stimulation before training had previously learned to play the flute using vibration from the instrument when they had a severe hearing impairment and before they were implanted. This may suggest that learning to use tactile cues for auditory stimuli can generalize from one context (in this case music) to another (in this case speech) and that the ability to use tactile cues can be retained over many years.

In the current study, training consisted of just 20 minutes of exposure to speech-in-noise and concurrent tactile stimulation over the course of two one-hour sessions, and no exposure to speech-in-noise alone. It is possible that the absence of training with speech-in-noise alone created a bias in the post-training testing session towards the condition with EHS. Training without tactile stimulation was not included as an increase in number of visits, a substantial increase in the duration of each visit, or a decrease in the already small amount of training with tactile stimulation would have been required. An increase in the number of visits would have made recruitment

of participants, who often live a significant distance from the research centre, extremely difficult. An increase in the one-hour session time was not deemed appropriate as participants reported finding the training very tiring. Biasing towards the condition with EHS was unlikely to have occurred for three main reasons: (1) participants “train” extensively on speech-in-noise without tactile stimulation in their everyday life outside of the laboratory sessions using the same CI settings as in the experiment; (2) participants trained with different talkers and different material from in the testing sessions, and so were unable to generate a content- or talker-specific bias for EHS; and (3) participant expectation that EHS would be beneficial was controlled by falsely instructing the participants that the experiment compared “audio enhancement” with “vibration enhancement”. This instruction was reinforced in each trial in the testing sessions and the experimenter received no indication that participants were sceptical about the truth of this statement. Furthermore, the experimenter remained blinded to whether EHS was being provided throughout the testing sessions to prevent experimenter bias. In future work, we plan to develop a remote training programme to increase the training time available and thereby allow for further control of bias effects.

As well as allowing further experimental control, the development of a remote training programme would allow several important extensions of the current work. For example, it will be important to establish which training methods provide the most benefit of EHS over long time periods in a real-world setting. The development of interactive, game-like training software with clear goals and rewards, which maximizes user motivation and enjoyment and minimizes the effort required to train, is likely to be important for ensuring maximal uptake of EHS. Once an effective long-term training regime has been established, it will also be highly interesting to investigate whether EHS can be used to accelerate patient acclimatization to a new CI and whether some of the benefits of EHS remain even when tactile stimulation is no longer provided. Finally, as future work moves EHS closer to a real-world application, it will be important to establish the practicality of using a wearable haptic device, exploring, for example, the robustness of EHS to gesticulation and other arm movements.

Future research could also explore whether EHS can enhance spatial hearing. Spatial hearing is severely impaired in CI users, particularly in those with only one implant^{30,31}, who make up 93% of users in the UK³². As well as allowing CI users to better locate auditory objects in space, improved spatial hearing might increase the release from masking that occurs when speech is in a separate spatial location to a masking sound^{33–35}. Remarkably, it has been shown that, using only tactile stimulation, subjects are able to discriminate the location of concurrent sounds whose direction of origin is separated by less than 5°³⁶. Furthermore, other research suggests that the tactile system can make use of tactile intensity differences that originate from sounds whose intensities differ by as little as 1 dB³⁷. Such a difference in sound level between the ears (interaural level difference) would correspond to a horizontal angle change of just a few degrees for sounds with frequency content above ~1 kHz³⁸. The current study was conducted with tactile stimulation to both wrists, a set up that could easily be adapted to provide interaural level difference cues to improve spatial hearing in CI users.

Materials and Methods

Participants. Ten CI users (4 male, 6 female; mean age = 61.2 years, ranging from 41 to 70 years) were recruited through the University of Southampton Auditory Implant Service. Participants were only invited to take part in the experiment if they were: (1) native British English speakers, (2) aged between 18 to 70 years old (inclusive), (3) had been implanted at least 12 months prior to the experiment, (4) had no EAS (i.e. no functional residual low-frequency hearing; >65 dB HL at 250 and 500 Hz), (5) had speech in quiet scores of at least 70% (for Bamford-Kowal-Bench sentences; BKB) measured at their most recent clinical appointment, and (6) had the capacity to give informed consent. Table 1 details participant characteristics. All participants also reported on a screening questionnaire (see Fletcher *et al.*¹⁶) that they had no medical conditions and were taking no medication that might affect their touch perception. Participants gave written informed consent and received an inconvenience allowance.

Vibro-tactile detection thresholds were determined for both the left and right index fingers for sinusoidal vibrations of 31.5 and 125 Hz, using conditions specified in ISO 13091-1:2001. Two participants (P2 and P3) had higher than normal thresholds (>0.7 ms⁻² root mean square [RMS]³⁹) at 125 Hz on the index finger of their left hand, and two other participants (P1 and P7) had elevated thresholds at 125 Hz on the index finger of their right hand. All other measured thresholds at the index finger were at normal levels.

Tactile signal processing. Figure 2 shows a schematic representation of the signal processing chain that was used to convert the audio signal to a tactile signal. The signal processing approach is described in detail in Fletcher *et al.*¹⁶ and is summarized below. The parameters used in the current study were the same as in Fletcher *et al.*, except that a smaller number of tonal carriers were used over a reduced frequency range. The reduced frequency range of the tactile signal (50–230 Hz) means that it can be more easily produced by a low-power wearable device that is near-silent.

To generate the tactile signal, the audio input signal was first downsampled to a sampling frequency of 16 kHz. The signal was then passed through an FIR filter bank with 4 channels that were equally spaced on the ERB scale with upper and lower centre frequencies of 100 and 1000 Hz. The Hilbert envelope for each channel was computed and a first-order low-pass filter was applied with a cut-off frequency of 10 Hz. The channels were then used to modulate the amplitude envelopes of four fixed-phase tonal carriers with center frequencies of 50, 110, 170, and 230 Hz (a frequency range in which the tactile system is highly sensitive⁴⁰). Each of these carrier signals was individually passed through an expander function (as used in Fletcher *et al.*) to amplify temporal modulations and reduce the contribution from the background noise, thereby increasing the salience of speech envelope information. Figure 3 illustrates the effect of the expander by showing the same sentence as: processed clean speech without the expander (panel A), processed speech-in-noise at an SNR of 2.5 dB with the expander turned off (panel B), and the same speech-in-noise signal with the expander turned on (panel C). The enhanced tonal

Participant	Gender	Age	Speech-in-quiet	Type of implant	Time since implantation	Dominant hand
		(years)	(% correct)		(years)	
1	F	65	85	MEDEL Sonata	9.3	Right
2	M	56	89	AB Hi Res 90k	3.4	Right
3	M	69	99	AB Hi Res 90k	1.5	Right
4	F	68	98	Cochlear Nucleus 512 (CA) profile	3.7	Left
5	M	68	94	AB HiRes ultra	1.0	Right
6	F	70	96	Cochlear Nucleus 512 (CA) profile	2.7	Right
7	M	70	92	Cochlear Nucleus Freedom contour	9.1	Right
8	F	44	98	Cochlear nucleus 512 (CA) profile	1.6	Right
9	F	41	93	Cochlear nucleus contour (bilaterally)	7.7	Right
10	F	52	100	AB Hi Res 90k	10.9	Right
<i>Mean</i>		<i>60.3</i>	<i>94.4</i>		<i>5.1</i>	

Table 1. Summary of participant information.

carriers were then summed to form the input signal for tactile presentation to the participant. A 10-ms delay was added, to keep the tactile signal approximately synchronized with the electrical signal from the CI. The tactile signal was presented through two *HVLab* tactile vibrometers. The mean acceleration magnitude of the vibration output for a single sentence was 1.84 ms^{-2} RMS.

Speech and noise stimuli. The speech signal was presented at a level of 65 dB SPL LAeq. Two different speech corpora were used in this study. The BKB Institute of Hearing Research male sentence corpus⁴¹ was used for speech testing. This material was not familiar to participants as clinical audiology test procedures use a different BKB sentence corpus. Training was conducted using speech material from the RealSpeech™ (UK) content library (used with permission of Dr Ian Wiggins and Dr Mark Fletcher), which used two talkers (one male and one female) that are different from the talker used in the BKB sentence corpus. RealSpeech material comprises a set of narratives that cover a variety of general-interest topics recorded under near-anechoic conditions. A non-stationary multi-talker noise recorded by the National Acoustic Laboratories (NAL)⁴² was used in both training and speech testing. The noise was recorded at a party and had a spectrum that matched the international long-term average speech spectrum⁴³.

Apparatus. Participants were seated in a sound-attenuated booth with a background noise level conforming to the 2017 British Society of Audiology recommendations⁴⁴. All stimuli were generated and controlled using custom MATLAB scripts (version R2016a, The MathWorks Inc, Natick, MA, USA). Acoustic stimuli were generated by a laptop located in a separate observation room and played out via an RME Babyface Pro soundcard (Haimhausen, Germany; sample rate of 44.1 kHz and bit depth of 24 bits) and a Genelec 8020 C PM Bi-Amplified Monitor System positioned at head height 1.5 m from the participant. The acoustic stimuli were calibrated at the listening position using a Brüel & Kjær (B&K) G4 type 2250 sound level meter and B&K type 4231 sound calibrator.

Two *HVLab* tactile vibrometers were placed beside the participant's chair and were used to deliver the vibro-tactile signal to the participants' wrists. The vibrometers in this experiment were adapted by the substitution of the standard 6-mm probe with a 10-mm probe, and the removal of the rigid surround. These changes increased the area of skin excitation. A B&K type 4294 calibration exciter was used to calibrate the accelerometers, and the vibration signal was calibrated to provide equal amplitude across the frequency range. A further *HVLab* tactile vibrometer with a 6-mm probe and a rigid surround was used to measure vibro-tactile detection thresholds using *HVLab* diagnostic software.

During testing, the experimenter sat in the control room with no line of sight to the participant. The sentences spoken by the participants were captured using a Shure BG 2.1 dynamic microphone, amplified by a GSI-61 audiometer, and presented to the experimenter through the GSI-61 Operator's Headset.

Procedure. Figure 4 shows a schematic illustration of the experimental procedure. Participants attended four sessions over a two-week period (mean period between first and last sessions = 5.4 days, ranging between 3 and 9 days, with an average gap between any two sessions of 1.8 days). The four sessions consisted of a speech testing session, followed by two training sessions, and then a final testing session. This procedure is similar to the one detailed in Fletcher *et al.*¹⁶, although here only 2 days of training were conducted, totalling around 20 minutes of exposure to speech in noise and concurrent tactile stimulation.

In the pre-training testing session, participants completed a screening questionnaire and had their vibro-tactile detection thresholds measured at the fingertip to check for normal touch perception. Detection thresholds were also measured for each wrist (where tactile stimulation was applied in this study) at 31.5 Hz (Mean = 0.79 ms^{-2} RMS) and 125 Hz (Mean = 0.68 ms^{-2} RMS). Each participant's speech reception threshold (SRT) was then measured without tactile stimulation. Two SRT estimates were made, each using a single BKB sentence list (containing

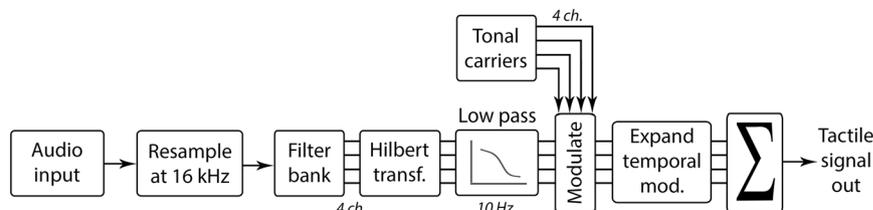


Figure 2. Schematic representation of the signal processing chain for the tactile signal generation.

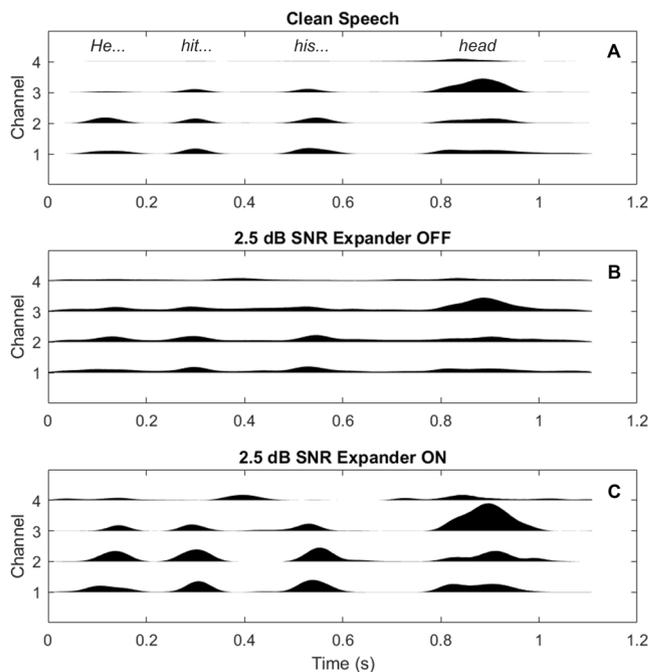


Figure 3. Illustration of the effect of the expander on the tactile signal amplitude for the sentence “He hit his head”. Panels A and B show the tactile signal for clean speech and for speech-in-noise (at 2.5 dB, the lowest SNR used in the study). Panel C shows the same signal as panel B, but with the expander on. The height of each channel waveform corresponds to the amplitude of the signal.

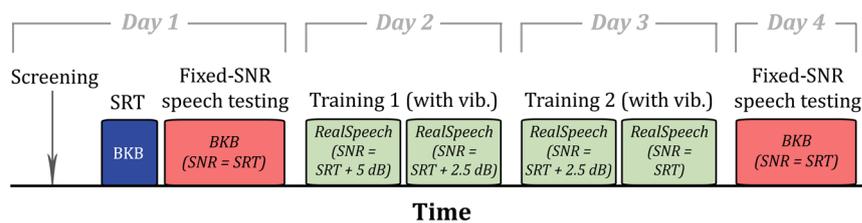


Figure 4. Schematic (not to scale) showing the timeline of the experiment. Speech material was either BKB sentences or audiobook material from RealSpeech.

15 sentences) mixed with multi-talker noise. Trials were marked as correct if participants correctly repeated at least 1 out of 3 keywords. For each SRT estimate, the SNR of the first trial was 5 dB. The sentence used in the first trial was then repeated, with the SNR increased by 2 dB after each repeat, until a trial was marked as correct. A one-up one-down adaptive tracking procedure⁴⁵ with a step size of 2 dB was then followed for the remaining sentences for each SRT. The SRT was calculated as the mean of the last 4 reversals.

In both the pre- and post-training testing sessions, speech-in-noise performance was measured at an SNR equal to the mean of the two SRT estimates made in the first session. Speech testing was conducted with and without concurrent tactile stimulation. In both testing sessions, sentences were played to the participants and they were asked to repeat back what was said. Performance was measured as the proportion of key words that the participants correctly identified. The experimenter was able to use a text display to instruct participants to

repeat their response if it was unclear. Sixteen BKB sentence lists were used in total, eight in each session. Half of the lists were played with concurrent tactile stimulation and half without. The two conditions were alternated in an A-B-A-B pattern across the lists and were counterbalanced across participants. The order of conditions was swapped between the pre- and post-training sessions (i.e. B-A-B-A). During the speech testing, participants were instructed via a text display to either place their wrists on the vibrometer contacts, or to place their arms on the armrests of the chair. The tactile signal was played through the vibrometers in both conditions. When participants had their wrists on the contacts, the message “Vibration enhancement ON. Audio enhancement OFF” was displayed on the screen, and when they had their arms on the armrests the message “Vibration enhancement OFF. Audio enhancement ON” was displayed. This latter message falsely stated that the audio signal had been enhanced in the condition without tactile stimulation. This deceptive cue was included to control for effects of participant expectation that tactile stimulation was intended to improve performance. The experimenter was blinded to which conditions contained tactile stimulation, in order to avoid experimenter bias.

The two training sessions consisted of target speech from the RealSpeech content library. In each session the participant listened to two speech segments, each lasting around five minutes. Half of the segments were read by a female talker and half by a male talker and were presented in a random order. The segments were split into single sentences and mixed with the NAL multi-talker noise. The first segment was presented at an SNR 5 dB above the participants SRT, the second and third at 2.5 dB above, and the final segment at an SNR equal to the participant’s SRT. Participants were asked to repeat each sentence to the experimenter after which the correct sentence text was displayed on the screen. For all training material concurrent tactile stimulation was provided.

The experimental protocol was approved by the University of Southampton Ethics Committee (ERGO ID: 30753) and the UK National Health Service Research Ethics Service (Integrated Research Application System ID: 244309). All research was performed in accordance with the relevant guidelines and regulations.

Data Availability

The dataset from the current study is publicly available through the University of Southampton’s Research Data Management Repository: <https://doi.org/10.5258/SOTON/D1034>.

References

1. Isaiiah, A., Vongpaisal, T., King, A. J. & Hartley, D. E. H. Multisensory Training Improves Auditory Spatial Processing following Bilateral Cochlear Implantation. *J. Neurosci.* **34**, 11119–11130 (2014).
2. Schorr, E. A., Fox, N. A., van Wassenhove, V. & Knudsen, E. I. Auditory-visual fusion in speech perception in children with cochlear implants. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 18748–50 (2005).
3. Spriet, A. *et al.* Speech Understanding in Background Noise with the Two-Microphone Adaptive Beamformer BEAM in the Nucleus Freedom Cochlear Implant System. *Ear and Hearing* **28**, 62–72 (2007).
4. Wouters, J. & Vanden, B. J. Speech recognition in noise for cochlear implantees with a two-microphone monaural adaptive noise reduction system. *Ear and hearing* **22**, 420–30 (2001).
5. Brown, C. A. & Bacon, S. P. Low-frequency speech cues and simulated electric-acoustic hearing. *The Journal of the Acoustical Society of America* **125**, 1658–1665 (2009).
6. Kong, Y.-Y. & Carlyon, R. P. Improved speech recognition in noise in simulated binaurally combined acoustic and electric stimulation. *The Journal of the Acoustical Society of America* **121**, 3717 (2007).
7. Bach-y-Rita, P. *Brain mechanisms in sensory substitution*. (Academic Press Inc, 1972).
8. Bach-y-Rita, P., Carter, C. C., Saunders, F. A., White, B. & Scadden, L. Vision Substitution by Tactile Image Projection. *Nature* **221**, 963–964 (1969).
9. Capelle, C., Trullemans, C., Arno, P. & Veraart, C. A real-time experimental prototype for enhancement of vision rehabilitation using auditory substitution. *IEEE Transactions on Biomedical Engineering* **45**, 1279–1293 (1998).
10. Meijer, P. B. L. An experimental system for auditory image representations. *IEEE Transactions on Biomedical Engineering* **39**, 112–121 (1992).
11. Brooks, P. L., Frost, B. J., Mason, J. L. & Gibson, D. M. Identification of words and manner features of speech by two profoundly deaf teenagers using a tactile vocoder. *The Journal of the Acoustical Society of America* **74**, S105–S105 (1983).
12. Hnath-Chisolm, T. & Kishon-Rabin, L. Tactile presentation of voice fundamental frequency as an aid to the perception of speech pattern contrasts. *Ear and hearing* **9**, 329–334 (1988).
13. Thornton, A. R. D. & Phillips, A. J. A comparative trial of four vibrotactile aids. *Tactile Aids for the Hearing Impaired Ed. IR Summers (London: Whurr) pp* 231–251 (1992).
14. Laurienti, P. J., Burdette, J. H., Maldjian, J. A. & Wallace, M. T. Enhanced multisensory integration in older adults. *Neurobiology of aging* **27**, 1155–1163 (2006).
15. Sumby, W. H. & Pollack, I. Visual contribution to speech intelligibility in noise. *The Journal of the Acoustical Society of America* **26**, 212–215 (1954).
16. Fletcher, M. D., Mills, S. R. & Goehring, T. Vibro-Tactile Enhancement of Speech Intelligibility in Multi-talker Noise for Simulated Cochlear Implant Listening. *Trends in Hearing* **22**, 233121651879783 (2018).
17. Huang, J., Sheffield, B., Lin, P. & Zeng, F.-G. Electro-Tactile Stimulation Enhances Cochlear Implant Speech Recognition in Noise. *Scientific Reports* **7**, 2196 (2017).
18. Rouger, J. *et al.* Evidence that cochlear-implanted deaf patients are better multisensory integrators. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 7295–300 (2007).
19. Oxenham, A. J. & Kreft, H. A. Speech Perception in Tones and Noise via Cochlear Implants Reveals Influence of Spectral Resolution on Temporal Processing. *Trends in Hearing* **18**, 233121651455378 (2014).
20. Stickney, G. S., Zeng, F.-G., Litovsky, R. & Assmann, P. Cochlear implant speech recognition with speech maskers. *The Journal of the Acoustical Society of America* **116**, 1081–1091 (2004).
21. Dawson, P., Mauger, S. & Herbasch, A. Clinical evaluation of signal-to-noise ratio–based noise reduction in Nucleus® cochlear implant recipients. *Ear & Hearing* **32**, 382–90 (2011).
22. Goehring, T. *et al.* Speech enhancement based on neural networks improves speech intelligibility in noise for cochlear implant users. *Hearing Research* **344**, 183–194 (2017).
23. Summers, I. R. *et al.* Tactile information transfer: A comparison of two stimulation sites. *The Journal of the Acoustical Society of America* **118**, 2527–2534 (2005).
24. Vallbo, A. & Johansson, R. The tactile sensory innervation of the glabrous skin of the human hand. in *Active Touch, the Mechanism of Recognition of Objects by Manipulation* (Pergamon Press, 1978).
25. Verrillo, R. T. Effects of aging on the suprathreshold responses to vibration. *Perception & Psychophysics* **32**, 61–68 (1982).

26. Verrillo, R. T., Bolanowski, S. J. & Gescheider, G. A. Effect of aging on the subjective magnitude of vibration. *Somatosensory & Motor Research* **19**, 238–244 (2002).
27. de Dieuleveult, A. L., Siemonsma, P. C., van Erp, J. B. F. & Brouwer, A.-M. Effects of Aging in Multisensory Integration: A Systematic Review. *Front. Aging Neurosci.* **9** (2017).
28. Levänen, S. & Hamdorf, D. Feeling vibrations: enhanced tactile sensitivity in congenitally deaf humans. *Neurosci. Lett.* **301**, 75–77 (2001).
29. Nava, E. *et al.* Audio-Tactile Integration in Congenitally and Late Deaf Cochlear Implant Users. *PLoS One* **9** (2014).
30. Litovsky, R. Y., Johnstone, P. M. & Godar, S. P. Benefits of bilateral cochlear implants and/or hearing aids in children. *International Journal of Audiology* **45**, 78–91 (2006).
31. Verschuur, C. A., Lutman, M. E., Ramsden, R., Greenham, P. & O'Driscoll, M. Auditory Localization Abilities in Bilateral Cochlear Implant Recipients. *Otology & Neurotology* **26**, 965 (2005).
32. British Cochlear Implant Group. Annual UK update. Available at, <https://www.bcig.org.uk/annual-uk-update/>. (Accessed: 5th November 2018) (2018).
33. Freyman, R. L., Helfer, K. S., McCall, D. D. & Clifton, R. K. The role of perceived spatial separation in the unmasking of speech. *The Journal of the Acoustical Society of America* **106**, 3578–3588 (1999).
34. Hirsh, I. J. The Relation between Localization and Intelligibility. *The Journal of the Acoustical Society of America* **22**, 196–200 (1950).
35. Yost, W. A., Dye, R. H. & Sheft, S. A simulated “cocktail party” with up to three sound sources. *Perception & Psychophysics* **58**, 1026–1036 (1996).
36. Richardson, B. L., Wuillemin, D. B. & Saunders, F. J. Tactile discrimination of competing sounds. *Perception & Psychophysics* **24**, 546–550 (1978).
37. Richardson, B. L. & Frost, B. J. Tactile localization of the direction and distance of sounds. *Perception & Psychophysics* **25**, 336–344 (1979).
38. Feddersen, W. E., Sandel, T. T., Teas, D. C. & Jeffress, L. A. Localization of High-Frequency Tones. *The Journal of the Acoustical Society of America* **29**, 988–991 (1957).
39. International Organisation for Standardization. Mechanical vibration–Vibrotactile perception thresholds for the assessment of nerve dysfunction–Part 1: Methods of measurement at the fingertips. *ISO 13091-1:2001* (2001).
40. Verrillo, R. T. Effect of contactor area on the vibrotactile threshold. *The Journal of the Acoustical Society of America* **35**, 1962–1966 (1963).
41. MacLeod, A. & Summerfield, Q. A procedure for measuring auditory and audio-visual speech-reception thresholds for sentences in noise: rationale, evaluation, and recommendations for use. *Br J Audiol* **24**, 29–43 (1990).
42. Keidser, G. *et al.* The National Acoustic Laboratories (NAL) CDs of Speech and Noise for Hearing Aid Evaluation: Normative Data and Potential Applications. *Australian and New Zealand Journal of Audiology, The* **24**, 16 (2002).
43. Byrne, D. *et al.* An international comparison of long-term average speech spectra. *The Journal of the Acoustical Society of America* **96**, 2108–2120 (1994).
44. British Society of Audiology. Recommended procedure: Pure-tone air-conduction and bone-conduction threshold audiometry with and without masking. *British Society of Audiology* 1–36 (2017).
45. Levitt, H. Transformed Up-Down Methods in Psychoacoustics. *The Journal of the Acoustical Society of America* **49**, 467–477 (1971).

Acknowledgements

Thank you to Professor Carl Verschuur for his support and encouragement, and to Dr. Devyane Bele and the dedicated staff at the University of Southampton Auditory Implant Centre for their enthusiastic help with recruitment of participants. We also extend our thanks to Gary Parker and Peter Russell for technical support, to Beepy for heft, and to Dr. Ben Lineton for providing crucial experimental apparatus, spond, and for his general support for the project. We are grateful to each of the participants who took part in this study for giving up their time and for coming back into the lab for four sessions, sometimes travelling a great distance to do so. Lastly, our thanks go to Erin Pegg for her support and guidance. Salary support for author M.D.F was provided by the Oticon Foundation and salary support for author T.G. was provided by Action on Hearing Loss (UK, grant number 82).

Author Contributions

M.D.F., A.H. and S.R.M. designed the experiment. A.H. conducted the experiment. M.D.F. and S.R.M. wrote the main manuscript text. T.G. and M.D.F. designed and implemented the signal processing strategy. All authors reviewed the manuscript.

Additional Information

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019