Evidence that cochlear-implanted deaf patients are better multisensory integrators

J. Rouger*, S. Lagleyre†, B. Fraysse†, S. Deneve§, O. Deguine*, and P. Barone§

*Centre de Recherche Cerveau et Cognition, Université Toulouse 3, Centre National de la Recherche Scientifique, Faculté de Médecine de Rangueil, 31062 Toulouse Cedex 9, France; †Service d’Oto–Rhino–Laryngologie, Hôpital Purpan, FR-31059 Toulouse, France; and §Groupe de Neurosciences Théoriques, École Normale Supérieure, 3, Rue d’Ulm, 75005 Paris, France

Edited by Eric I. Knudsen, Stanford University School of Medicine, Stanford, CA, and approved February 26, 2007 (received for review October 24, 2006)

The cochlear implant (CI) is a neuroprosthesis that allows profoundly deaf patients to recover speech intelligibility. This recovery goes through long-term adaptive processes to build coherent percepts from the coarse information delivered by the implant. Here we analyzed the longitudinal postimplantation evolution of word recognition in a large sample of CI users in unisensory (visual or auditory) and bisensory (visuoauditory) conditions. We found that, despite considerable recovery of auditory performance during the first year postimplantation, CI patients maintain a much higher level of word recognition in speechreading conditions compared with normally hearing subjects, even several years after implantation. Consequently, we show that CI users present higher visuoauditory performance when compared with normally hearing subjects with similar auditory stimuli. This better performance is not only due to greater speechreading performance, but, most importantly, also due to a greater capacity to integrate visual input with the distorted speech signal. Our results suggest that these behavioral changes in CI users might be mediated by a reorganization of the cortical network involved in speech recognition that favors a more specific involvement of visual areas. Furthermore, they provide crucial indications to guide the rehabilitation of CI patients by using visually oriented therapeutic strategies.

The cochlear implant (CI) is a neuroprosthesis that allows profoundly deaf patients to recover speech intelligibility. This recovery goes through long-termadaptive processes to build coherent percepts from the coarse information delivered by the implant. Here we analyzed the longitudinal postimplantation evolution of word recognition in a large sample of CI users in unisensory (visual or auditory) and bisensory (visuoauditory) conditions. We found that, despite considerable recovery of auditory performance during the first year postimplantation, CI patients maintain a much higher level of word recognition in speechreading conditions compared with normally hearing subjects, even several years after implantation. Consequently, we show that CI users present higher visuoauditory performance when compared with normally hearing subjects with similar auditory stimuli. This better performance is not only due to greater speechreading performance, but, most importantly, also due to a greater capacity to integrate visual input with the distorted speech signal. Our results suggest that these behavioral changes in CI users might be mediated by a reorganization of the cortical network involved in speech recognition that favors a more specific involvement of visual areas. Furthermore, they provide crucial indications to guide the rehabilitation of CI patients by using visually oriented therapeutic strategies.

Pre- and Postimplantation Performance of Speech Recognition in CI Patients.

We have analyzed the performance of 97 CI users in disyllabic word recognition by using three modalities (auditory, visual, and visuoauditory) during a longitudinal study that extended over 8 years after implantation (Fig. 1).

Auditory Speech. First, at the time the implant is switched on (T0), CI users obtain a significant recovery of word recognition in auditory modality, with a performance level of 47.1 ± 27.3% SD in quiet conditions. This performance level is much higher than that obtained before implantation by using an external hearing aid (mean 10.4 ± 14.2% correct, P < 0.05). Auditory performance increases significantly during the subsequent months (P < 0.05), before reaching a plateau from about the seventh month on and then showing no significant improvement in the following years (mean 81% over the first year).


The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Abbreviations: CI, cochlear implant; NH, normally hearing; SNR, signal-to-noise ratio.

See Commentary on page 6883.

To whom correspondence should be addressed. E-mail: pascal.barone@cerco.ups-tlse.fr.

© 2007 by The National Academy of Sciences of the USA

www.pnas.org/cgi/doi/10.1073/pnas.0609419104

PNAS | April 24, 2007 | vol. 104 | no. 17 | 7295–7300
Rouger et al. 7296 www.pnas.org/cgi/doi/10.1073/pnas.0609419104

Fig. 1. Word-recognition scores for CI users in the three sensory modalities: auditory only (A only, green), visual only (V only, blue), and bisensory visuo-auditory (AV, red). (A) Longitudinal performance (mean percentage correct ± SD) of the entire cohort of CI users (n = 97) at different times before (preoperatively) and after the cochlear implantation. In the left part of the graph, we have reported the speechreading performance (V only) of NH subjects (n = 42). (B) Individual performance levels for two CI users who have been regularly evaluated during a 3-year period after implantation. Both graphs show the significant recovery of auditory speech recognition during the first year compared with the weak performance before implantation and the stable high values of speechreading at all periods tested. In bisensory conditions, near-maximum scores are achieved.

Speechreading. At T0, speechreading performance in CI users is elevated and impressively higher than that observed for NH subjects tested with the same talker (35.1 ± 14.7% vs. 9.4 ± 7.1%, respectively; P < 0.05). This speechreading ability in CI users is similar to that obtained a few months before implantation (mean 30.1 ± 15.1%; paired test, P = 0.62) and remains unchanged across all postimplantation periods tested (>35%; P > 0.05), although CI users have reached their maximal auditory performance. Furthermore, visual performance at T0 is not correlated with auditory proficiency (r² = 0.001, P = 0.76). At the time of implantation, duration of deafness is not correlated to visual performance (r² = 0.001, P = 0.77). This latter result should be tempered because most of the patients were suffering from a progressive hearing impairment, such that deafness duration could hardly be reliably defined. However, in three CI users who were affected by sudden deafness (such as meningitis) and implanted only 1 year later, speechreading performance levels were similar to that of the CI population (20%, 30%, and 45%, respectively). To strengthen this observation, we have included in our analysis data obtained from five supplementary CI users (not included in the longitudinal retrospective study) suffering from sudden deafness occurring within <1 year of implantation. In this enlarged sample (n = 8), we observed that performance in visual-only conditions was much higher than that observed in NH subjects (mean 27.5 ± 10.7% vs. 9.4%, respectively; P < 0.05). In these CI users, several months or years of auditory recovery postimplantation (auditory-only performance >90% correct) did not affect their speechreading performance. Despite the limited number of observations, this suggests that a high level of speechreading ability can be acquired rapidly during a period of auditory deprivation and then remain at stable values.

Audiovisual Speech. As expected from the classically perceptual benefit derived from multisensory integration (5), prior to implantation, CI users present higher performance in visuo-auditory conditions compared with the auditory-alone conditions (55.8 ± 21.0% vs. 10.4%, respectively; P < 0.05). A similar effect is observed in CI users postimplantation; when compared with unisensory conditions, visuo-auditory integration results in an improvement in word recognition in CI patients tested at T0 (86 ± 17.4% correct; P < 0.05 for both comparisons). From the time of implantation, audiovisual recognition improved slightly (P < 0.05) with practice, allowing CI users to reach near-perfect performance levels (94 ± 12.0%) as early as the second month postimplantation.

We believe that the difference in bisensory performance of CI users when comparing pre- and postimplantation periods (55.8% vs. 86% at T0; P < 0.05) is derived from higher auditory performance provided by the neuroprosthesis. In agreement with this, we saw that, in a limited number of CI users (n = 14) who did not show any improvement in auditory word recognition, the visuo-auditory gain remains unchanged when comparing the two testing periods (mean visuo-auditory benefit preimplantation 0.54 vs. 0.62 at T0; paired test, P = 0.23).

Comparison of Performance of NH and CI Subjects. Our results show that, during the period of deafness, CI patients have developed a specific ability in speechreading that distinguishes them from the poor speechreading skills of NH listeners. We hypothesized that this high visual aptitude might induce in CI users an improvement of the mechanisms of multisensory integration, leading to greater visuo-auditory benefits than those observed in NH subjects. In this scheme, we compared the visuo-auditory gain in CI users at T0 (i.e., without training) to the one obtained in naive NH subjects exposed to a degraded auditory signal. This auditory degradation allows us to make direct comparisons of visuo-auditory performance from both groups at equivalent ranges of nonoptimal auditory performance. To degrade the auditory performance of NH subjects, we first used a masking paradigm with white noise at different SNRs. In this protocol, we observed a higher recognition rate in visuo-auditory versus auditory-only conditions (Fig. 2A), especially at intermediate SNRs at ~15 dB (20). Second, we used a noise-band vocoder paradigm with different frequency bands that simulates the processing strategy of CIs (21). In this simulation, the global temporal and spectral information of the signal is preserved, whereas the fine temporal cues within each spectral component are removed. In this case, performance (auditory and visuo-auditory) decreases rapidly as the number of bands decreases, leading to near-zero values in the two-electrode simulation (1.5% mean recognition in auditory-only presentation). However, whereas bisensory presentation improved NH subjects’ performance (Fig. 2A), the visuo-auditory gain was much lower than the one obtained in the masking protocol at equivalent auditory performance levels, suggesting that visuo-auditory integration mechanisms of speech perception strongly depend on the integration of fine spectrottemporal auditory information. This hypothesis was confirmed by our model (see Are CI Patients Better Multisensory Integrators? When visuo-auditory performance of CI users is compared with that of the NHs exposed to degraded auditory stimuli, we show that the visuo-auditory gain in CI patients is higher than that observed in NH subjects in the simulation or noise-masked conditions (both comparisons, P < 0.001; Fig. 2B). The differences in favor of CI users are especially
large in conditions of low auditory performance, where the range of correct recognition falls to < 30% (CIPs vs. NHs, $P < 0.01$ for both comparisons). For example, a subset of CI users ($n = 13$) unable to perform auditory identification at all (0% correct) showed a high level of performance in the visuoauditory condition (mean 63% correct). In contrast and compared with this subset of CI users, NH subjects showing a similar level of auditory word recognition (0% correct, $n = 19$) due to highly degraded auditory conditions never reached the visuoauditory performance levels (mean 25.4% and 12.5% in masking or vocoder simulation protocols, respectively).

Although in CI patients the high efficiency of bisensory word recognition was not correlated to the level of speechreading ($r^2 = 0.068, P < 0.05$; see ref. 22), we further tested the hypothesis that the difference between CI and NH bisensory integration could be due to differences in absolute levels of visual performance. Consequently, we selected a subgroup of CI users showing low visual performance (lower than 20%; $n = 15$). We found within this group that the visuoauditory gain was still higher than that of NHs engaged in the simulating protocol (0.52 vs. 0.26, $P < 0.001$). In our opinion, this reinforces our conclusion that CI users have acquired a higher bisensory proficiency per se compared with NH subjects.

**Are CI Patients Better Multisensory Integrators?** As mentioned previously, the better performance levels of CI users compared with NHs with simulated implants could be due either to their stronger visual performance or to a better capacity for integrating visuoauditory inputs. Furthermore, electrophysiological studies have challenged whether the rules governing neuronal computing during multisensory interactions are superadditive, additive, or subadditive (23). Does it apply to the performance of speech recognition in bisensory conditions? To evaluate these hypotheses and quantify the multisensory performance, we designed two simple models of word recognition. The first model is a minimal-integration model, in the sense that the integration of auditory and visual cues occurs within the lowest possible level of interaction between both inputs (i.e., probabilistic combination). The second model is an optimal-integration model in which individual spatio- and spectrotemporal audiovisual cues are combined across modalities to minimize the amount of information required for correct word recognition. We fitted a model of optimal multisensory integration to the performance of NHs with masked auditory input (Fig. 3D). We then compared the performance of the model with all subjects’ performance in two other conditions (CI users at T0 and NH subjects with vocoder; Fig. 3A and C). We found that the model fitted very well the performance of CI patients, indicating that at T0 they integrate visuoauditory inputs as efficiently as NHs when their auditory input is degraded by white noise. However, the bisensory performance of NHs with simulated implants was far below the model performance levels (Fig. 3C). Thus, in contrast with CI users, NH subjects did not integrate their visuoauditory input optimally when this auditory input is lacking fine spectrotemporal structure. Furthermore, CI users tested 1 year postimplantation showed a significant improvement of both auditory and visuoauditory performance while keeping a constant high speechreading recognition level. When applying the model to the unisensory performance of CI users at 1 year (Fig. 3B), we found that the evolution of multisensory performance with practice could be entirely explained by their increased auditory performance. This finding suggests that, whereas visual and auditory inputs are integrated optimally from the start, a reorganization of auditory cortices, supporting a better capacity for dealing with distorted auditory inputs, is the main cause for the quasi-perfect multisensory performance reached by CI users after 1 year.

**Discussion**

This study provides a long-term evaluation that shows the impressive benefits of cochlear implantation regarding the recovery of speech recognition because profoundly deaf patients can reach high rates of performance for hearing speech during the first 6 months postimplantation. The present data confirm that a profound hearing loss induces the acquisition of strong speechreading abilities (6, 19, 24, 25), but they represent the first evidence that this skill remains unaffected by the recovery of the auditory functions provided by the prosthesis. CI patients preserve a striking speechreading ability acquired during the period of deafness while they have reached optimal auditory recognition. We interpret this apparently paradoxical strategy developed by CI users as a strategy to maintain through the mechanisms of bisensory integration a high level of speech recognition in a disturbed noisy auditory environment. Previous studies have reported that the performance of CI patients is highly susceptible to noise (16, 26), which is probably due to the...
lead to a powerful utilization of the visual spatiotemporal cues (29) provided by the lip and face movements (10), allowing these patients to reach near-perfect performance in visuoauditory situations. Using our computational model that allows us to avoid ceiling effects in subjects’ performance, we confirmed that the performance of CI patients derived not only from higher efficiency in speechreading, but also from the acquisition of a higher skill level in multisensory integration when visual speech information is matched to an impoverished auditory signal.

Our results provide crucial information on the temporal window during which plastic changes can occur in the cortical network of CI patients during adaptation to the neuroprosthesis. There is now a growing body of evidence showing that sensory deprivation from early developmental stages has an important effect on the remaining sensory modalities (30, 31) through active cross-modal neuromodulatory mechanisms (32, 33). In general, sensory deprivation leads to a compensatory increase in specific skills of the spared modalities that can be observed at both behavioral and neural levels in animal and human subjects (34, 35). However, cochlear implantation constitutes a unique approach to understand the cortical mechanisms that underlie the functional recuperation of the lost sensory modality. First, it has been shown that because CIs provide only a degraded signal that requires specific compensatory strategies, CI users present different levels of activation in auditory areas involved in semantic and/or phonological speech processing (36). Our longitudinal study on a large sample of patients suggests that such changes probably occur during the first 6 months depending on subjects’ performance in speech recognition (37–39) and might remain different from normally hearing listeners (40) even after several years of auditory function recovery. Second, our results highlight that CI users develop a strong visuoauditory perceptive strategy for speech intelligibility while experiencing the reduced spatiotemporal information provided by the implant. This adaptation extends over the first 3 months postimplantation before being stabilized, suggesting that the pattern of brain activity during visuoauditory speech processing in CI users may vary during the corresponding period. Brain-imaging studies in CI deaf subjects have revealed a particular involvement of the low-level visual areas when listening to words (37). This finding corroborates our results of a strong synergy between visual and auditory processing for speech recognition following cochlear implantation. These results, in agreement with our ongoing functional imaging study (41), suggest that the visual activity derived from speechreading could actively influence the activity of the cortical network involved in hearing speech recognition and could participate in the improvement of performance in visuoauditory conditions. The existence of heteromodal connections that link directly unisensory areas in adult primates (42, 43) provides a possible anatomical framework for such direct visuoauditory interactions at low levels of sensory processing (44).

First, at a theoretical level, it has been shown that the fine spatiotemporal auditory information provides important cues for auditory speech recognition (45, 46). Our results broaden the role of the temporal fine structure because it optimizes the audiovisual speech integration leading to a higher multisensory perceptual benefit, in agreement with the actual technological challenge aiming at improving the spectral resolution of CIs. Second, from the clinical point of view, this work provides important cues to adapt the rehabilitation strategy as a function of implant experience. The supranormal skills in multisensory integration observed in CI deaf patients should be used to improve recovery of other auditory functions that are still deficient in CI users. Because visuoauditory training facilitates perceptual learning in a single modality (47, 48), we believe that a strong visually and audiovisually based rehabilitation during the first months postimplantation would significantly improve and fasten the functional recovery of speech intelligibility or
sound localization, which is largely deficient in unilateral CI patients (49, 50).

Materials and Methods

Participants. Our study was based on a retrospective analysis of speech recognition in 97 postlinguistically deafened subjects (mean age 56 years, range 19–82) that received a CI after profound deafness (defined as a hearing loss of ≥90 dB) of diverse etiologies (meningitis, chronic otitis, otosclerosis, neunoma) and durations (mean age 22 years, range 1–57). The clinical implantation criteria included word and open-set sentence auditory-recognition scores <30% under best-aided conditions (i.e., with conventional acoustic hearing aids). All CI patients were recipients of a Nucleus (Cochlear) implant (CI-22 or CI-24) and used a range of different sound-coding strategies. Performance was collected during regular visits to the ear-nose-throat department following a standard rehabilitation program. We restricted our analysis to evaluations performed by the same speech therapist during the 10 years of follow-up and by using exactly the same procedures. First, we collected the performance of CI users tested before the cochlear implantation and by using an external hearing aid. On average, these tests were performed >6 months before the implantation (mean 5.8 months), but on 36 CI users of 97, word-recognition performance was obtained during the last 3 months postimplantation. Then from the day the CI was switched on (T0, usually 1 month postsurgery), CI users were tested at regular intervals during the first year and >8 years postimplantation. Data have been pooled into 12 groups (Fig. 1A) corresponding to the testing period from T0 (n = 91), 3 months (n = 91), 5 months (n = 82), 7 months (n = 77), 1 year (n = 78), 2 years (n = 69), 3 years (n = 41), 4 years (n = 26) 5 years (n = 17), 6 years (n = 11), 7 years (n = 5), and >8 years (n = 4) postimplantation. On average, CI users were tested during a period of 33 months postimplantation (±25), with on average eight sessions per subject (±3). In this postlinguistic deaf adult population, we did not find a relationship between the age of implantation and the performance collected in the uni- or bisensory conditions (all cases P > 0.05). In addition, speech recognition for different sensory modalities was tested in a sample of 163 NH subjects. These control subjects were all native French speakers with self-reported normal or corrected-to-normal vision and without any previously known language or cognitive disorders.

Procedures and Stimuli. All subjects were tested on open-set recognition for French disyllabic words obtained from the classically used French speech therapist list developed by Fournier and presented in visual-only (speechreading, V), auditory-only (A), and visuoauditory (VA) conditions. Only words correctly repeated verbally by subjects were treated as correct responses (% correct score). We calculated the visual contribution to speech recognition by using the method of Sumby and Polack (8) (VA benefit = [(VA−A)/(100−A)]) to normalize for the performance observed in the A condition and thus to be able to compare directly the visuoauditory gain across groups (19). CI users were tested in silence on 20 words in each condition. In NH subjects, we developed three paradigms during which the auditory stimuli (the words pronounced by the speech therapist and recorded onto a PC computer) were differently degraded or presented without alteration. In a masking protocol, in A and VA conditions we additively combined each sound to a masking sound, with the words’ acoustic level shifted to obtain the required SNR (nine SNR conditions: 0, −5, −10, −12, −15, −17, −20, −22, and −25 dB). The mask was a white noise delivered by a pseudorandom number generator and temporally modulated by monophasic sinusoidal lobes (period = 20 msec), with a mean rate of 300 modulations per sec temporally randomly distributed. Gain was 1 at the edge of the lobes and 0.4 at the center. This white noise modulation was carried out to ensure high random temporal fluctuations. Subjects (n = 80) were tested with four lists of 20 words in A or VA conditions, with masking noise at a single SNR condition and the orders for each A or VA sequence being randomized across subjects. At each masking condition in the range from −5 and −22 SNR, data were obtained from a sample of 10 subjects (but only 3 subjects at SNR 0 dB and 7 at SNR −25 dB). In a “simulating” protocol used with a second group of NH subjects (n = 41), for A and VA conditions, we developed noise-band vocoder methods that simulate the signal processing computed in a CI (21). The sound was analyzed through 2, 4, 8, or 16 frequency bands by using sixth-order IIR elliptical analysis filters. The cutoff frequencies of these bands were calculated to ensure equidistance of the corresponding basilar membrane locations of the cochlea according to the human cochlear tonotopic map (51). Spectral analysis was systematically carried out between 250 and 8,000 Hz. Cutoff frequencies were 250, 1,676, and 8,000 Hz for the 2-band condition; 250, 709, 1,676, 3,713, and 8,000 Hz for the 4-band condition; 250, 437, 709, 1,104, 1,676, 2,507, 3,713, 5,462, and 8,000 Hz for the 8-band condition; and 250, 335, 437, 561, 709, 888, 1,104, 1,563, 1,676, 2,053, 2,507, 3,054, 3,713, 4,506, 5,462, 6,613, and 8,000 Hz for the 16-band condition. For each filtered frequency band, signal, temporal envelope was extracted by half-wave rectification and envelope smoothing with a 500-Hz low-pass third-order IIR elliptical filter. The extracted temporal envelope was then used to modulate white noise delivered by a pseudorandom number generator, and the resulting signal was filtered through the same sixth-order IIR elliptical filter that was used for the frequency band selection. Finally, signals obtained from each frequency band were recombined additively, and the overall acoustic level was readjusted to match the original sound level. The performance of at least 10 subjects was analyzed for each band condition. In a last protocol, NH subjects (n = 42) were tested on three lists of 20 disyllabic words presented in V conditions.

In all conditions, the lists of words were equalized for syllabic structure (CV/CVC/CCV), language utilization frequency (Brueis), and anterior–posterior phonemic constitution. The stimuli were uttered by the female French speech therapist, who pronounced each word with even intonation, tempo, and vocal intensity. Utterances were recorded in an anechoic chamber with a professional digital video camera with lights focused on the face such that minimal shadowing occurred. Video was digitized at 720 × 576 pixels at 25 frames per sec, and sound was digitized at 32,000 Hz by using a 16-bit quantization. Audiovisual stimuli with sound degradation were made by using Adobe Premiere Pro 7.0 (Adobe Systems, Mountain View, CA), and temporal coincidence was respected between the original and processed sounds. All stimuli were finally exported in MPEG2 video format with maximum encoding quality.

Visual and Auditory Integration Models. An increase in multisensory performance does not necessarily prove that subjects integrate their visuoauditory inputs. Indeed, being in the presence of two signals rather than one automatically increases the probability of recognition because a word can be recognized from one or the other signal. In model 1 (minimal integration), we suppose that the subjects recognize a word from what they see or hear, or the other signal. In model 2 (minimal integration), we suppose that the subjects recognize a word from what they see or hear, or the other signal. In model 3 (minimal integration), we suppose that the subjects recognize a word from what they see or hear, or the other signal. In model 4 (minimal integration), we suppose that the subjects recognize a word from what they see or hear, or the other signal.
the multisensory enhancement in performance in all conditions (red dotted line in Fig. 3). From this latter result, we can conclude that the visual and auditory modalities are indeed combined in a word-recognition task, albeit to a weaker extent in the case of NH subjects listening to the vocoder simulation.

To quantify multisensory integration, we used a very simplified model where a word is “recognized” when a sufficiently large number of “cues” specific to this word is detected, whether from the visual or auditory input (model 2: optimal multisensory integration). For example, a threshold of six means that six or more specific cues need to be detected to identify a word. These cues could be a specific motion of the mouth or a particular pattern in the time/frequency spectrum of the auditory signal (they do not necessarily correspond to phonemes). Moreover, we assume that the quality of the sensory input controls the average number of “cues” that can be detected in this condition.

The detection of each cue is probabilistic. We suppose that each cue is detected with a particular probability $P$ independently of the other cues. The resulting average number of detected cues is $N = \lambda P$, where $N$ is the total number of cues present in the word and $P$ is the probability of detecting each of them. $P$ depends on the quality of the sensory signal and controls the performance of the model. If $N$ is sufficiently large and $P$ is sufficiently small, the number of detected cues on each trial, $n$, follows approximately a Poisson law: The probability of detecting $k$ cues becomes $P(n = k) = (\lambda e^{-\lambda})^k/k!$. The probability of recognition corresponds to the probability that the number of detected cues will exceed a particular threshold, i.e., $P(n > T)$. Thus, it is a function of both $\lambda$ and $T$. Assuming a fixed threshold $T = 6$ for each subject and condition, we can infer the mean number of cues $\lambda V$, $\lambda A$, and $\lambda VA$ and detect the V, A, and VA sensory inputs for each subject in each condition. For example, $\lambda V$ is the value for which $P_A = \Pi_\lambda(n > T)$ is equal to the observed auditory performance.

If visual and auditory inputs were combined optimally, they should add up together (the total number of detected cues is the sum of the visually and orally detected cues). Thus, the total signal should follow a Poisson law with mean $\lambda VA = \lambda V + \lambda A$. From this, we can infer the performance of an ideal observer: The biseension recognition probability corresponds to the probability that a Poisson-distributed signal with mean $\lambda VA$ will exceed the threshold $T$. Thus, in the case of optimal multisensory integration, we have $P_{VA} = \Pi_{\lambda V + \lambda A}(n > T)$.

Model 2 has one “free” parameter, the threshold $T$. The higher the $T$, the stronger the multisensory enhancement compares to unisensory performance. For $T = 1$, model 2 is equivalent to model 1 (there is no true integration: A word is detected by one or the other sensory modality). $T = 6$ was used for generating the model predictions because it is the best match for both the NH subjects in listening to noise-masked speech (Fig. 3D) and CI users (Fig. 3 A and B). By “best match,” we mean that $T = 6$ minimized the mean squared error between the model predictions and the subject-per-subject performance in these conditions. Although the presence of a free parameter prevents us from proving that multisensory integration is optimal in an absolute sense, it provides a rigorous comparison between multisensory integration performance for different conditions and subject groups.

We thank M. L. Laborde and D. Sauvajon for help collecting the data; and C. James, S. Thorpe, and F. Polleux for comments and corrections on the manuscript. This work was supported by grants from the Action Concertée Incitative Neurosciences Intégratives et Computationnelles (to J.R., O.D., and P.B.), the Fondation pour la Recherche Médicale (to S.L.), and the Action Thémétique et Incitative sur Programmes et Equipes program of the CNRS (to P.B.).