Hearing Research 322 (2015) 163–170

Contents lists available at ScienceDirect

Hearing Research

journal homepage: www.elsevier.com/locate/heares

Research paper

Gradual adaptation to auditory frequency mismatch

Mario A. Svirsky [a, b, *], Thomas M. Talavage [c, d], Shivank Sinha [e], Heidi Neuburger [f], Mahan Azadpour [a]

[a] Dept. of Otolaryngology-HNS, New York University School of Medicine, New York, NY, USA
[b] Center of Neural Science, New York University, New York, NY, USA
[c] ECE, Purdue University, West Lafayette, IN, USA
[d] BME Depts., Purdue University, West Lafayette, IN, USA
[e] EE Dept., Purdue University, Indianapolis, IN, USA
[f] Dept. of Otolaryngology-HNS, Indiana University School of Medicine, Indianapolis, IN, USA

A R T I C L E I N F O

Article history:
Received 13 June 2014
Received in revised form
13 October 2014
Accepted 16 October 2014
Available online 6 November 2014

A B S T R A C T

What is the best way to help humans adapt to a distorted sensory input? Interest in this question is more than academic. The answer may help facilitate auditory learning by people who became deaf after learning language and later received a cochlear implant (a neural prosthesis that restores hearing through direct electrical stimulation of the auditory nerve). There is evidence that some cochlear implants (which provide information that is spectrally degraded to begin with) stimulate neurons with higher characteristic frequency than the acoustic frequency of the original stimulus. In other words, the stimulus is shifted in frequency with respect to what the listener expects to hear. This frequency misalignment may have a negative influence on speech perception by CI users. However, a perfect frequency-place alignment may result in the loss of important low frequency speech information. A trade-off may involve a gradual approach: start with correct frequency-place alignment to allow listeners to adapt to the spectrally degraded signal first, and then gradually increase the frequency shift to allow them to adapt to it over time. We used an acoustic model of a cochlear implant to measure adaptation to a frequency-shifted signal, using either the gradual approach or the "standard" approach (sudden imposition of the frequency shift). Listeners in both groups showed substantial auditory learning, as measured by increases in speech perception scores over the course of fifteen one-hour training sessions. However, the learning process was faster for listeners who were exposed to the gradual approach. These results suggest that gradual rather than sudden exposure may facilitate perceptual learning in the face of a spectrally degraded, frequency-shifted input.

This article is part of a Special Issue entitled <Lasker Award>.

1. Introduction

Cochlear implants are the first successful example of replacing a human sensory end organ with an electronic device. This accomplishment has been recently recognized with the awarding of the Lasker-DeBakey Clinical Medical Research Award to Blake Wilson, Graeme Clark, and Ingeborg Hochmair, three of the pioneers in the field. There are two main reasons why the cochlear implant is considered one of the major medical advances of the twentieth century. One of them is the impact it has had on the clinical treatment of hearing impairment and on the quality of life of hundreds of thousands of patients. These patients include adventitiously hearing impaired adults as well as children who were born profoundly deaf and whose sole auditory input was provided by the implant. The case of children is significant because in their case cochlear implants influence not only their ability to understand speech but also their ability to speak intelligibly and their development of oral language. Additionally, the widespread use of cochlear implants has caused a paradigm shift in a number of related fields by opening a number of new avenues for scientific and clinical pursuit. For example, it seems likely that the development of retinal prostheses has significantly benefitted in many different ways from the prior development of cochlear implants.

* Corresponding author. Dept. of Otolaryngology-HNS, New York University School of Medicine, 550 First Avenue, NBV-5E5, New York, NY 10016, USA. Tel.: +1 646 284 2457.
E-mail address: svirsky@nyu.edu (M.A. Svirsky).

http://dx.doi.org/10.1016/j.heares.2014.10.008
0378-5955/

There are also numerous examples of important scientific questions that could not have been answered were it not for the availability of a large base of cochlear implant users. The study of sensitive periods for speech and language development is one of these examples. The sensitive period hypothesis states that the innate human ability to develop certain skills including language development and intelligible speech production decreases without early exposure to oral language. This is a hypothesis that cannot be easily tested with animal models because animals do not have speech and language. The definitive way to test the hypothesis would be to withdraw auditory input from children at birth and then measure their ability to develop speech language when hearing is restored after different periods of deprivation. This would be obviously unethical. However, cochlear implantation of children who are congenitally and profoundly deaf provides a unique opportunity to conduct an indirect test of the sensitive period hypothesis, and even makes it possible to estimate the length of such sensitive periods for different specific skills. The way this is done is by testing children who have received cochlear implants at different ages and therefore have been deprived of auditory input for different amounts of time (Tomblin et al., 2005; Geers et al., 2011; Tobey, 2013; Niparko et al., 2010; Svirsky et al., 2000, 2004b, 2007). Admittedly, this is not a perfect experiment because cochlear implants do not restore perfect hearing, but it is clear that the existence of cochlear implant technology made it possible to obtain a novel source of data to study the nature and time course of sensitive periods in language development. A second area of scientific study where cochlear implants have opened new doors is the study of the influence of auditory input on speech production. Cochlear implants make it possible to easily manipulate auditory input and measure the subtle effects on the listener's speech production. They also make it possible to completely deprive a listener of sound for periods ranging from minutes to several hours, to measure the changes that occur in speech production parameters in response to the deprivation, and immediately after auditory input is restored (Matthies et al., 1994; Lane et al., 1997; Svirsky and Tobey, 1991; Svirsky et al., 1992). Cochlear implants make it possible to study changes in speech production after implantation in listeners who have been sound deprived for years or even decades (Perkell et al., 1992) and the resulting data can provide insights about underlying mechanisms in normal speech production (Lane et al., 1995; Perkell et al., 1995). Finally, a third area worth mentioning is that cochlear implants have had a profound influence on the study of speech perception using a very impoverished auditory input. The realization that a relatively primitive and distorted signal can be so useful in a pragmatic sense has given renewed impulse to studies that investigate the human brain's ability to understand speech in very adverse situations. Here we present a modest example of a study that does not involve cochlear implants directly but was clearly inspired by thinking about the mechanisms that postlingually deaf cochlear implant users might employ to understand speech.

It is well known that the human brain's ability to recognize patterns in spite of an extremely distorted input is very robust, at least when those distortions occur along a single dimension. In particular, humans can recognize speech in very adverse acoustic conditions that are well beyond the capabilities of the most sophisticated automatic speech recognition systems. Some well known examples include sinewave speech (where each formant frequency is represented by an amplitude and frequency modulated sinewave (Remez et al., 1981)), infinite peak clipping (where speech is converted to a signal that can assume only two values, one positive and one negative, and only the zero crossings of the original signal are preserved (Licklider and Pollack, 1948)), and a four channel noise vocoder (where the input signal is filtered by four adjacent frequency channels and the output of each channel modulates a different noise band (Shannon et al., 1995)). However, there are indications that simultaneous distortion along two dimensions may be a more significant problem. See, for example, Fu and Shannon (1999), Rosen et al. (1999), Başkent and Shannon (2003), and Svirsky et al. (2012). The latter includes audio examples of a clear speech signal along with versions of the same signal that were degraded along one or two different acoustic dimensions.

Because auditory sensory aids impose different kinds of distortions, it is important to understand how listeners perceive degraded speech. This understanding may inform the design of next generations of such aids. For example, cochlear implants impose at least two kinds of signal distortion. One of them is the inevitable spectral degradation that is dictated by a relatively small number of independent stimulation channels. In postlingually deaf adults, cochlear implants may also impose a mismatch between the acoustic frequency of the input signal and the characteristic frequency of the neurons that are stimulated in response to that input. Why does this happen? There is mounting evidence that some models of cochlear implant electrode arrays are not inserted deeply enough to stimulate neurons whose characteristic frequency is under 300 Hz (Harnsberger et al. (2001), Carlyon et al. (2010), Svirsky et al. (2001, 2004a), Zeng et al. (2014)). One particularly compelling data set can be found in McDermott et al. (2009), who tested five users of the Nucleus cochlear implant (model CI24RE(CA)) before they had any experience with the device. These five subjects had some usable acoustic hearing that allowed them to match the percepts elicited by stimulating the most apical electrode to the percepts elicited by pure tones presented to the unimplanted ear. The tones that provided the best match ranged from about 600 Hz to 900 Hz, suggesting that this device imposes an initial basalward frequency shift of one to two octaves (because the most apical electrode is stimulated in response to ambient sounds around 250 Hz but it actually "sounds like" a tone of 600–900 Hz, at least upon initial stimulation). Basalward shift imposes a conflict between the percepts caused by speech sounds and the internal representation of the same speech sounds that was learned by the same (postlingually deaf) listener when he had normal hearing. In other words, this is a discrepancy between novel sensory input and representations of speech sounds stored in long term memory.

Given the possibility of such basalward shift, designers of auditory prostheses are faced with two main options. The first one is to adjust the analysis filters of the speech processor to the estimated characteristic frequency of the stimulated neurons. This option has the advantage that no adaptation is required on the part of the listener but it comes at the expense of losing some low-frequency information that is important for speech perception. Another option is to accept the frequency misalignment and let listeners adapt to it over time. This is what is normally done with users of cochlear implants. This approach has the advantage that the most relevant speech information is delivered to the listener, but it requires adaptation to a frequency shift that can be quite severe in some individuals, and this may lead to incomplete adaptation in those cases. There are a number of studies showing that postlingually deafened cochlear implant users have a remarkable ability to adapt to the standard clinical frequency allocation tables, or even frequency tables that are further shifted from the standard (e.g., Skinner et al. (1995), Svirsky et al. (2001), McKay and Henshall (2002), Svirsky et al. (2004a), Gani et al. (2007)). Other studies have shown that the electric pitch sensation associated with activation of any given electrode adapts with experience, generally becoming more consistent with the frequency information provided by the cochlear implant processor at that electrode's location (Svirsky et al., 2004a; Reiss et al., 2007,