Evaluation of a method to improve perception of voice pitch in users of CIS cochlear implants

Olivier Macherey

Supervised by Stuart Rosen

Rapport de stage de DEA ATIAM
(Paris 6, ENST, Aix Marseille II, UJF Grenoble I)
March - August 2003
Dept of Phonetics & Linguistics
University College of London

French Abstract

Implants cochléaires

Les implants cochléaires permettent à des personnes profondément malentendantes ou sourdes de recouvrer partiellement l'audition. Le principe de fonctionnement de ces prothèses est de stimuler électriquement le nerf auditif à l'aide d'électrodes implantées le long de la cochlée. Le signal acoustique est capté par un microphone puis analysé. Les paramètres extraits de ce signal vont ensuite être codés pour générer des signaux électriques reconstituant le signal d'origine (c.f. fig.1.2). Tous les implants commercialisés de nos jours sont multicanaux: ils permettent d'exciter le nerf auditif à différents endroits de la cochlée, exploitant le codage tonotopique des fréquences. Les recherches en techniques de traitement du signal appliqués aux implants se sont focalisées ces dernières années sur la perception de la parole. Même si celle-ci reste difficile en milieu bruité, les résultats obtenus sont encourageants et les patients peuvent généralement suivre une conversation normale. Les enjeux actuels se tournent vers le codage de la hauteur des sons (ou pitch). En effet, la hauteur des sons a une importance considérable, à la fois dans la perception de la voix parlée (elle renseigne sur le sexe du locuteur, son état émotionnel, son âge ...) mais aussi dans la perception de la musique. Le travail présenté ici s'intéresse à une possibilité d'améliorer le codage de la hauteur des sons dans les implants et particulièrement de la voix parlée.

Principe des méthodes comparées CIS (Continuous Interleaved Sampling) standard et modifiée

La technique CIS consiste à générer des impulsions électriques de façon non simultanée, c'est à dire qu'il n'y a jamais deux canaux qui stimulent le nerf en même temps (c.f. figure 2.3). Ceci a permis de supprimer l'effet d'intéraction entre canaux et les distortions que cela engendrait. Le principe de cette méthode est le suivant: le signal est filtré par huit filtres passe-bande (nous prenons comme exemple l'implant C1 de Clarion, utilisé dans les expériences et qui posséde huit différents canaux). Ces filtres ont des fréquences de coupure en adéquation avec la position des électrodes sur la cochlée (c.f. tab. 4.1). L'enveloppe temporelle de chacun de ces signaux est extraite par rectification et filtrage passe-bas (typiquement à 400 Hz). Ces enveloppes vont ensuite servir à moduler des trains d'impulsions générer à taux fixe, engendrant huit signaux qui sont envoyés aux électrodes correspondantes.

La stratégie testée dans ce rapport s'apparente fortement à la technique CIS classique. Cependant, plutôt que de moduler directement les trains d'impulsions par les enveloppes temporelles, on envisage de pré-moduler ces impulsions par des contours périodiques (de période déduite de la fondamentale du signal). On obtient ainsi des trains d'impulsions périodiques (c.f. fig.3.3). Plusieurs contours différents ont été testés avant le présent travail et l'un deux s'est avéré comme le plus prometteur. On l'appellera "sawsharp" (abréviation de "sharpened sawtooth" signifiant "dent de scie aiguisée"). Le caractère "aiguisé" semble favorable puisqu'il donne une information claire sur la période (par rapport à une sinusoïde par exemple).

Expériences

Les expériences menées ont été réalisées grâce à l'interface CRI de l'industriel Clarion. Cette interface comprend la partie externe de l'implant et permet de générer n'importe quel stimuli (dans la limite du taux de stimulation de l'implant). Les patients ayant passé les tests psychoacoustiques n'avaient donc qu'à changer la partie externe de leur prothèse pour la notre. Cinq patients ont pris part aux tests. Deux expériences ont été effectuées: l'une était un test d'identification de voyelles anglaises et l'autre un test de différenciation entre question et affirmation. Ce dernier test est un des moyens de mesurer la perception de la hauteur dans la parole humaine. Pour chacun des deux tests, les stimuli provenaient de deux locuteurs différents (de sexe masculin et féminin). Le test d'identification de voyelles fut adapté en fonction des performances des patients, trois patients firent un test avec neuf différentes voyelles et deux d'entre eux le firent avec cinq. En effet, du fait de leurs différences (causes de la surdité, nombre de neurones auditives survivantes ...), il existe une grande variabilité de performance entre chacun.

En parallèle de cette étude, des expériences similaires ont été menées avec des normo-entendants (5 sujets) en utilisant un programme de simulation d'implant cochléaire. Le principe en est simple. Les mêmes traitements sont appliqués au signal (filtrage par 8 passe-bandes et extraction des enveloppes) mais on utilise un bruit blanc à la place des trains d'impulsions. Concernant la méthode CIS "sawsharp", on module huit bruits différents par ce même contour. La différence se fait lors de ce que l'on pourrait appeler la resynthèse du signal. En effet, ici, le signal final est un signal acoustique, il est reconstitué en filtrant chacun des signaux par le même passe-bande que pour la première étape puis en les sommant tous. Les tests d'identification de voyelles comportaient ici douze voyelles différentes car les performances des sujets étaient très bonnes. Il faut noter de plus que les stimuli provenaient initialement du même matériel que pour les expériences impliquant les porteurs d'implants.

Resultats

Les résultats des patients pour le test de question/affirmation (c.f. fig. 4.4) indiquent une légère amélioration des performances avec la modulation "sawsharp" (pour les trois sujets dont les résultats sont au-dessus du niveau de chance). Cette amélioration est statistiquement significatrice pour deux patients (LN et PG). Concernant le test d'identification de voyelles (c.f. fig. 4.2), les résultats sont similaires pour les deux méthodes excepté pour un patient (LN) qui montre une baisse de performance avec la pré-modulation.

Les résultats des simulations (c.f. fig. 5.4 et 5.2) sont assez similaires à ceux des patients. La contribution de la modulation "sawsharp" dans le test question/affirmation semble cependant plus importante chez ces normo-entendants. Concernant le test d'identification de voyelles, aucune différence entre les deux méthodes n'est à relever. On note toutefois que les résultats obtenus à partir des stimuli de la locutrice sont meilleurs que ceux provenant du locuteur.

Discussion

La modulation par le contour "sawsharp" semble donc améliorer légèrement la perception des variations de pitch (mesurées dans le test de question/affirmation). Ceci a été clairement observé chez deux patients. Il n'y a cependant pas assez de résultats pour pouvoir tirer une conclusion claire de ce travail. L'amélioration de la perception du pitch avec la modulation "sawsharp" apparait plus clairement dans les simulations, comme cela avait été déja observé dans une précédente étude (c.f. Green [11]).

La conclusion intéressante est que la modulation sawsharp ne semble pas détériorer la perception des voyelles. En effet, aucune amélioration n'était attendue mais on aurait pu penser que le filtrage passe-bas à 32 Hz (au lieu de 400 Hz pour la méthode CIS classique) détériore la perception de la parole. Mais ce n'était pas le cas, on otient donc des résultats consistants avec ceux observés par Drullman chez des normo-entendants ([16]). Un patient a toutefois montré une baisse de performance avec la modulation. Il est de plus intéressant de noter que les patients ont majoritairement trouvé que les voyelles générées avec la modulation "sawsharp" étaient plus difficilement reconnaissables (et cela même si leurs résultats ne le montraient pas). On peut supposer que la méthode dégrade la "qualité du son" ou bien que les patients ont besoin de temps pour s'habituer à un certain type de codage. En effet, ils amélioraient leur performance au fil de l'expérience avec cette méthode de façon plus importante qu'avec la CIS classique.

Perspectives

Refaire le même test avec plus de sujets permettrait peut-être d'avoir une idée plus claire des effets de cette modulation. Des expériences d'identification de consonnes et de phrases complèteraient également la présente étude. Enfin, il serait intéressant de refaire ces mêmes expériences avec un implant plus récent, autorisant des taux de stimulation plus importants. En effet, le contour serait de cette façon codé de manière plus fine.

Introduction

Cochlear implants can restore partial hearing to profoundly deaf people. Improvements have been made in the design of speech processing methods and users of cochlear implants often achieve good performance in speech recognition. They usually report however that music sounds unnatural and is not pleasant to hear. This difficulty to appreciate music can be partly due to the lack of melodic pitch information. Moreover, voice pitch information plays an important role in the perception of speech, providing cues to linguistic features such as word emphasis and question-statement contrats, and also to paralinguistic features such as the age, sex, identity, and emotional state of the speaker ([10]). The present work contributes to a research project of Department of Phonetics and Linguistics (UCL, London), funded by the RNID (Royal National Institute for Deaf people) and aimed at developing speech processing methods that will optimise the availability of voice pitch information.

The first two chapters are an introduction to cochlear implants. Most of the information of these chapters come from reference articles [1], [2], [3] and [4]. The chapter deals with the different parts of a cochlear prostheses and their usefulness whereas the second chapter is specifically related to the signal processing strategies available in implants. The third chapter is a small review on pitch perception, particularly on temporal cues to pitch available in implants. Introducing the strategy evaluated in this report, this chapter states the motivation of the present study. Chapters 4 and 5 describe the psychoacoustical test run during the project, both with cochlear implantees and normally hearing listeners (using a cochlear implant simulator).

Chapter 1

Background

1.1 Hearing

For normally hearing listeners, sound ungoes a serie of transformations as it travels through the outer ear, middle ear, inner ear, auditory nerve and into the brain. The outer ear picks up acoustic waves that are converted to mechanical vibrations by a series of small bones in the middle ear. In the inner ear, the cochlea (a cavity filled with fluid) transforms those mechanical vibrations to vibrations in fluid. Pressure variations within the fluids of the cochlea leads to displacement of a flexible membrane, called the basilar membrane, in the cochlea. The maximum of the vibration's amplitude depends on the frequency of the signal: the higher the frequency is the more the basilar membrane is excited near the base and the lower the frequency is the more the basilar membrane is excited near the apex. Attached to that membrane are hair cells that are bent according to the displacement of the membrane. The bending of the hairs releases a neuro-transmitter causing neurons to fire, signaling the presence of an excitation at a particular site of the cochlea. These neurons communicate with the central nervous system and transmit information about the acoustic signal to the brain. One can see on figure 1.1 a picture of the ear and of its suggested functions.

1.2 Deafness

The normal hearing system presented in the previous section can however experience damage leading to hearing loss or in extreme cases total deafness. Two different types of hearing loss are commonly distinguished.

Conductive hearing loss:
Conductive hearing loss is caused by damage in the outer or the middle ear leading to less efficient transmission of sound. People with conductive deafness tend to speak very quietly because they can usually hear their own voice quite well (due to bone conduction) but cannot hear other voices as well. Conductive deafness is usually helped by a hearing-aid or can be surgically corrected.

Sensorineural hearing loss:
Sensorineural hearing loss is caused by some defect in the cochlea or in the auditory nerve.

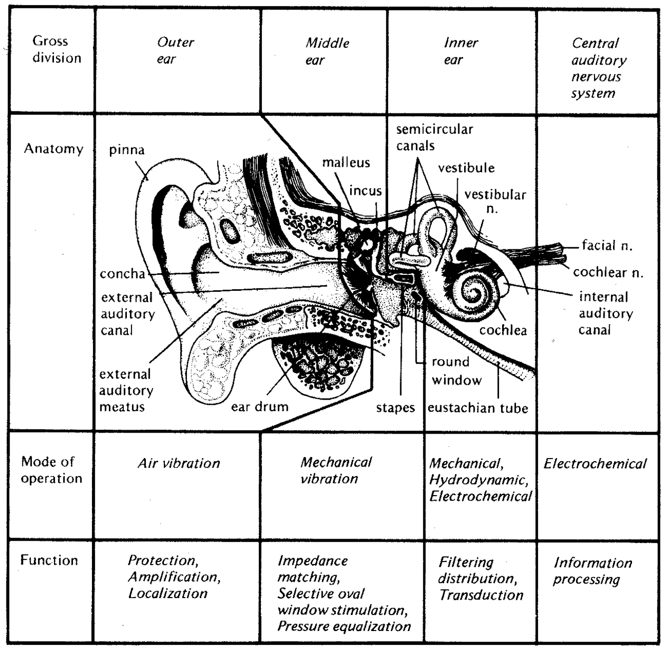

Figure 1.1: The outer, middle and inner-ear and their suggested functions (from Yost [18])

People with sensorineural deafness tend to talk loudly and classical hearing aids are not totally effective. Research has shown that the most common cause of deafness is the loss of hair cells rather than the loss of auditory neurons. If the hair cells are damaged, the auditory system has no way of transforming acoustic pressure waves to neural impulses, and that in turn leads to hearing impairment. Moreover, damaged hair cells can subsequently lead to degeneration of adjacent auditory neurons. In the case of profound or total deafness, cochlear implants have been developed so that the remaining neurons can be excited directly through electrical stimulation, bypassing the normal hearing mechanism (outer, middle and part of the inner ear).

1.3 Cochlear Implants

The aim of this section is to describe the different elements of a cochlear implant. Not much attention will be given to signal processing strategies because the whole third chapter is dedicated to it.

1.3.1 General mechanism

The general mechanism of a cochlear implant is as follows: a microphone picks up the sound, a signal processor converts the sound into electrical signals and a transmission system conveys the electrical signals to an electrode array inserted into the cochlea. A cochlear prosthesis is divided into two parts (as shown on figure 1.2): an external part (microphone and signal processor) and an internal part (receiver and electrode array) which is inserted by an otological surgeon.

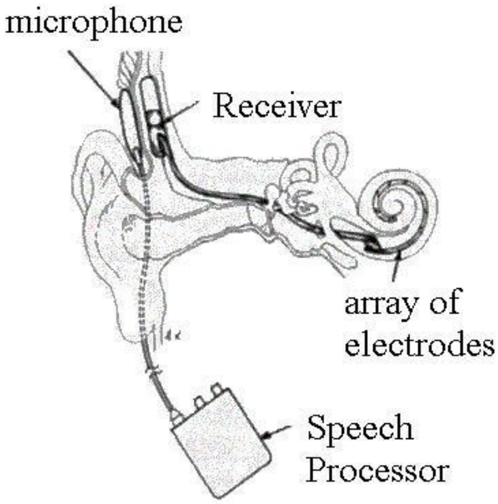

Figure 1.2: External and internal parts of a cochlear implant

1.3.2 Electrode design

The design of electrodes has been the focus of research for over two decades. Several aspects have been taken into account: electrode placement, number of electrodes and spacing of contacts, orientation of electrodes with respect to the excitable tissue and electrode configuration. Most commonly, the electrodes are placed in the scala tympani because it brings the electrodes in close proximity with auditory neurons that lie along the length of the cochlea, mimicking the place coding of a healthy cochlea. The number of electrodes as well as the spacing between the electrodes affects the place resolution for coding frequencies. In principle, the larger the number of electrodes, the finer the place resolution coding frequencies. But this is not necessarily true because of two restricting factors: [1] the number of surviving auditory neurons that can be stimulated at a particular site of the cochlea; [2] the spread of excitation associated with electrical stimulation. The first problem depends mostly on the etiology of deafness. The second problem can be partly solved using non-simultaneous pulsatile stimulation.

1.3.3 Type of stimulation

There are two types of stimulation depending on how information is presented to the electrodes. If the information is presented in analog form (slowly varying waves), then the stimulation is referred to as analog stimulation, and if the stimulation is presented in pulses, then the stimulation is referred to as pulsatile stimulation. Of course, a pulse waveform is also analog because stimulations are electrical but these names are commonly used to differentiate strategies. In analog stimulation, an electrical analog of the acoustic waveform is presented to the electrode. The acoustic signal is bandpass filtered and then, the waveforms are presented simultaneously on every channel. One advantage of this stimulation is that all the information contained in the acoustic waveforms is transmited. But one disadvantage is the problem of channel interactions, due to the simultaneity of stimulation across channels. In pulsatile stimulation, the information is delivered using a set of narrow pulses. The advantage of this stategy is the possibility to stimulate the electrodes in a non-simultaneous fashion, as we will see it in chapter 2 for the CIS strategy.

1.3.4 Transmission

There are two ways of transmitting the signals: through a transcutaneous connection and through a percutaneous connection. The transcutaneous one transmits the stimuli through a radio-frequency link. An external transmitter is used to encode the stimulus information for radio-frequency transmission from an external coil to an implanted coil. The internal receiver decodes the signal and delivers the stimuli to the electrodes. The advantage of this system is that the skin in the scalp is closed after the operation, thus avoiding possible infection. One disadvantage is that the implanted electronics may fail and would require surgery to be replaced. In the percutaneous connection, the stimuli are transmited directly through plug connections. There are no implanted electronics, other than electrodes. The major advantage of this system is his flexibility and signal transparency. Nowadays, all the implants use the transcutaneous connection.

Chapter 2

Signal processing strategies in cochlear implants

The goal of this chapter is to provide an overview of signal processing methods used in cochlear implants. We are not going to cover the whole history of implants but only the different types of methods investigated or commonly used. All of them are now multichannel, that is to say multiple sites in the cochlea are stimulated using an array of electrodes, exploiting the tonotopic coding of frequencies. Different electrodes are stimulated depending on the frequency of the signal. Electrodes near the base of the cochlea are stimulated with high-frequency signals, while electrodes near the apex are stimulated with low-frequency signals.

Several signal processing methods have been developed, differing primarily in the type of information to be transmitted.

2.1 Feature-Extraction Strategies

Feature-extraction strategies were initially used in the multielectrode Nucleus implant manufactured by Nucleus Limited. The processor presents spectral features such as formants obtained by formant-extraction algorithms. The MPEAK strategy of Nucleus is presented on figure 2.1. The fundamental frequency (F0), the first formant (F1) and the second formant (F2) are extracted from the speech signal using zero crossing algorithms. Additional high-frequency information is extracted using envelope detectors from three high-frequency bands. The envelope outputs of the three high frequency bands are delivered to fixed electrodes as indicated. Four electrodes are stimulated at a rate of F0 pulses/sec for voiced sounds, and at a quasi-random rate for unvoiced segments. The high frequency bands (2 kHz) enhance the representation of the second formant and the perception of consonants.

The MPEAK strategy was developed in the late 80's and has been progressively abandoned by Nucleus and replaced by waveform strategies. The main weakness of feature-extraction strategies is in the algorithms used to extract the fundamental and formant frequencies.

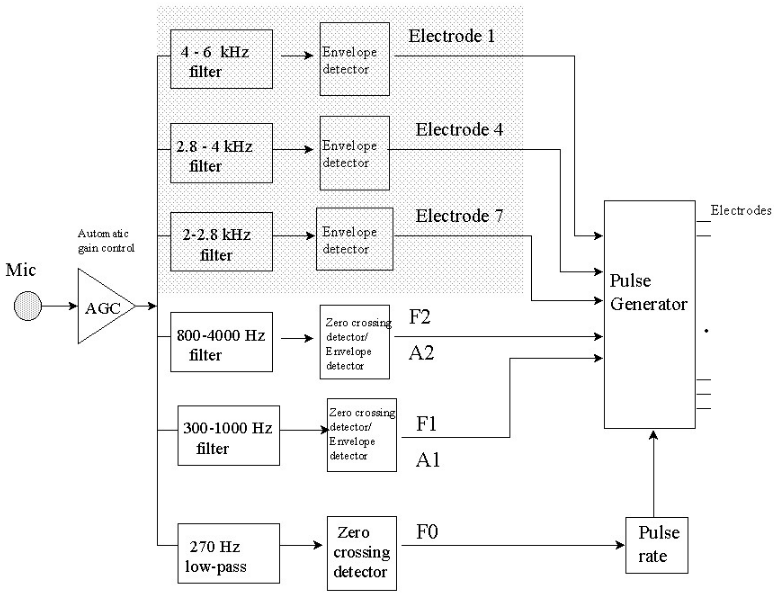

Figure 2.1: The MPEAK strategy (from Loizou [1])

2.2 Waveform Strategies

The type of stimulation in waveform strategies can be either analog or pulsatile. We will refer to the analog case as the Compressed-Analog Approach (CA). For pulsatile stimulation, we will describe the Continuous Interleaved Sampling approach (CIS).

2.2.1 Compressed-Analog Approach (CA)

The signal is first compressed using an automatic gain control and then filtered into several frequency bands. The filtered waveforms go through adjustable gain control and then are sent directly to intracochlear electrodes. The electric signal is delivered simultaneously to the implanted electrodes in analog form.

The main problem with simultaneous simulation is the interaction between channels caused by the summation of electrical fields from individual electrodes. Neural responses to stimuli from one electrode can significantly be distorted by stimuli from other electrodes. Consequently, these interactions can distort speech spectrum information and degrade speech understanding. The SAS (Simultaneous Analog Stimulation) strategy, currently available in the Clarion C1 implant, is similar to the CA approach with some extra compression added.

2.2.2 Continuous Interleaved Sampling Approach (CIS)

This strategy was developed at the Research Triangle Institute (RTI), by Wilson and his colleagues. The channel interaction problem is minimised by using non-simultaneous, interleaved pulses. Results from Wilson [5] show a great improvement in speech perception compared to the analog strategies. Trains of biphasic pulses are delivered to the electrodes in a nonoverlapping way, such that only one electrode is stimulated at a time. The amplitudes of the pulses are derived by extracting the envelopes of the bandpassed waveforms. The CIS strategy is shown on figure 2.2. The signal is first pre-emphasized and then passed through a bank of bandpasss filters. The envelopes of the filtered waveforms are extracted by rectification and low-pass filtering. The rectification can be full-wave or half-wave. The envelope outputs are finally compressed and then used to modulate biphasic pulses. The compression is done by using a logarithmic function in order to fit the patient's dynamic range of electrically evoked hearing. Trains of biphasic pulses, with amplitudes proportional to the envelopes, are delivered to the six electrodes at a constant rate in a nonoverlapping fashion. The rate has been found to have a major impact on speech recognition, high-pulse rate yielding higher performance than low-pulse rate ([6]).

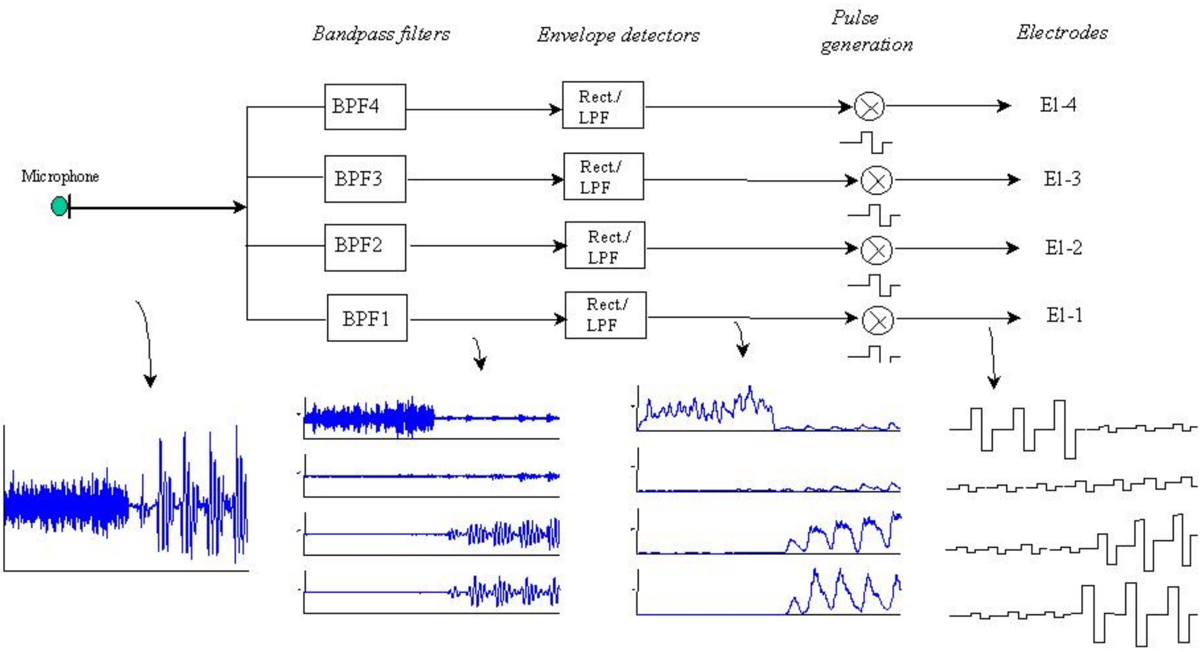

Figure 2.2: the CIS Strategy (from Loizou [1])

The CIS strategy is implemented in several implants: Clarion, Nucleus and Med-EL. The difference between these implants using CIS is mainly the number of channels (8 for Clarion, 22 for Nucleus and 12 for Med-EL). The CIS strategy is not the only pulsatile strategy currently available. Indeed, some others ones have been developed. For example, the PPS (Paired Pulsatile Sampler) strategy by Clarion, as shown on figure 2.3. Rather than stimulating one single electrode, two of them are stimulated at the same time. Pairs are made to avoid channel interaction as much as possible. This allows to increase the rate of stimulation (doubling it compared to CIS). Fu & al. ([6]) have found better phoneme recognition increasing the stimulation rate.

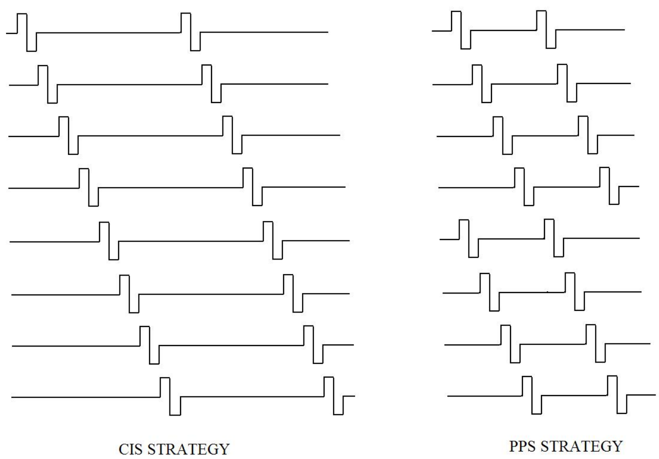

Figure 2.3: Sequences of biphasic pulses for the CIS strategy and for the PPS strategy (from [21])

Chapter 3

Pitch perception in cochlear implants

3.1 Place theory and Rate theory

The notion of pitch is usually correlated with frequency, due to the tonotopic coding of the cochlea (c.f. chapter 1). However, in some cases, normal listeners report perceiving a pitch corresponding to a frequency which doesn't present any energy. One example is the experiment of the missing fundamental. Observers report a 100 Hz pitch associated with a stimulus consisting of a sum of the frequencies of 700, 800, 900, 1000 Hz, without any energy at 100 Hz, as shown in figure 3.1. We notice the 10-msec periodicity of the time waveform. It may be possible that the auditory system perceives the reciprocal of the period which equals 100-Hz. (c.f. [18]).

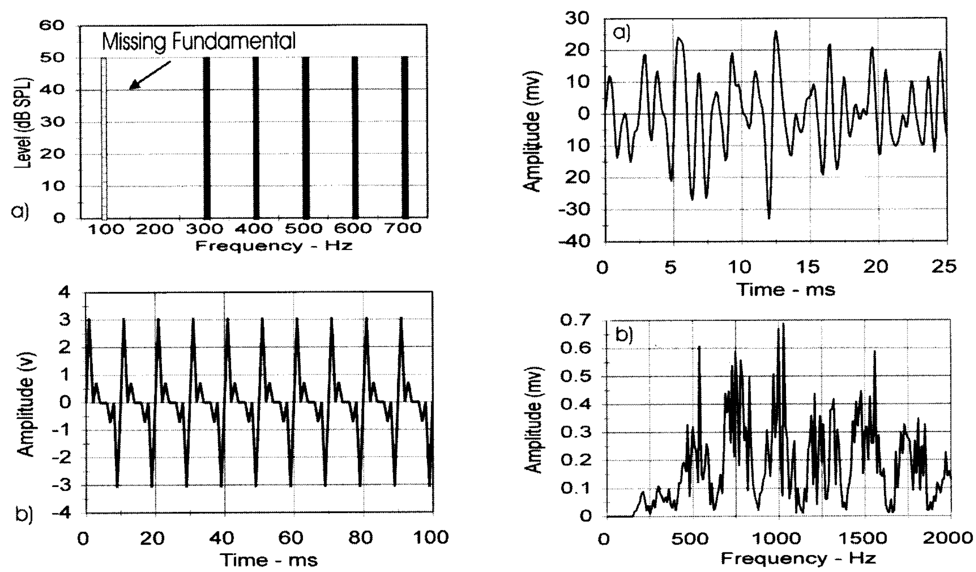

Figure 3.1: The missing fundamental (from Yost [18])

Historically, Ohm was the first to formulate (1843) a relation between Fourier theory and auditory analysis, arguing that a musical sound is determined by the frequency of the lowest Fourier component. The observations of Seebeck (1843) were in contradiction with that statement and Seebeck suggested that pitch was rather determined by the period of the signal's waveform (rate theory). A few years later, Helmholtz (agreeing with Ohm) formulated a theory based on possible resonances in the inner-ear. That was the premise of the place theory for pitch encoding. These two theories (place and rate) are however insufficient to explain the perception of pitch in normal hearing. In the place theory, it is difficult to explain the observed fine resolution of frequency (0.2 %). And the rate theory cannot explain the perception of tones with frequencies greater than the maximum synchronised firing rate of neurons. These two theories have been studied a lot, in order to evaluate both temporal and place cues to pitch.

Cochlear implant patients present a unique opportunity to explore temporal and spectral cues because these two types of information can be disociated allowing the precise exploration of a purely temporal code.

3.2 Temporal cues to pitch in cochlear implants

Several studies have shown temporal cues to pitch are important at low frequencies. The ability to discriminate rates of stimulation is generally limited to rates below about 300 Hz although individual differences extend the limits between 200 and 1000 Hz (Townshend, [12]). That limit of 300 Hz can be found in the recent work of Zeng ([8]). In his experiment, five cochlear implant users were asked to discriminate trains of biphasic pulses presented at different rates. All stimuli were delivered either to the most basal electrode or to the most apical electrode. Results show pitch was dominated by the temporal code at low frequencies, because abilities of patients were similar between apical and basal electrodes. This frequency limit for temporal coding is mainly determined by the refractory period of the axons, which can range from an absolute value of approximately 0.3 ms to a relative value of around 5 ms ([13]).

However, it's not because cochlear implant users cannot discriminate pitch at higher than 300-Hz frequencies that they cannot process temporal information at these frequencies. Psychophysical data show that implant users can detect temporal fluctuations at frequencies as high as 4000 Hz (Shannon, 1992).

To design processing strategies for cochlear implants, it's necessary to know how to combine the place code and the rate code. McKay & al. ([9]) have studied the possible relationship between place and temporal cues in pitch perception. Four users of the Mini System 22 (Nucleus) participated in their study. Difference limens for rate change, place change, and combined rate and place change (with consistent and inconsistent cues) were compared for stimulation at low and high rates. The results suggest place and rate cues are used independently in pulsatile electrical stimulation. This observation tallies with the earlier work of Pijl ([13]). In that study, three subjects were asked to label the interval of a 5^{th} by having a single electrode stimulated. Pulse rate of the lower note of the stimulus intervals was 100 pps and the experiment was repeated three times (on three different electrodes: basal, intermediate and apical). No substantial difference between performance with different electrodes was noticed, suggesting temporal cues to pitch are independent of the place of stimulation.

3.3 The sharpened sawtooth modified-CIS Strategy

The limited spectral resolution of the CIS strategy eliminates the spectral pitch cues that are available to normal hearing listeners. Temporal envelope cues to pitch are, in principle, available through these systems but, up to now, little consideration in the design of speech processing methods has been given to the salience of pitch information. The aim of the signal processing strategies investigated here in the department (Green & al. [10]) is to increase the modulations of the temporal envelope. Indeed, as temporal cues are used by implant users, improvements in pitch perception may be done by emphasizing the way they are coded. The principle is to code the fundamental frequency by modulating the pulses with a special periodic envelope as shown on figure 3.3. Trains of pulses are thus modulated by this specific periodic envelope. We used in the present study the sharpened sawtooth envelope. Such a sharpened envelope should give a clearer information on fundamental frequency than smoother, less modulated, contours. This modulation is done on every channel (8 for the C1 implant of Clarion) in order to give temporal cues to pitch to every channel. The resulting modulated pulses (figure 3.3) will then be modulated by the extracted envelope (rectification and low-pass filtering at 32 Hz rather than 400 Hz) in a same way as in the standard CIS strategy .

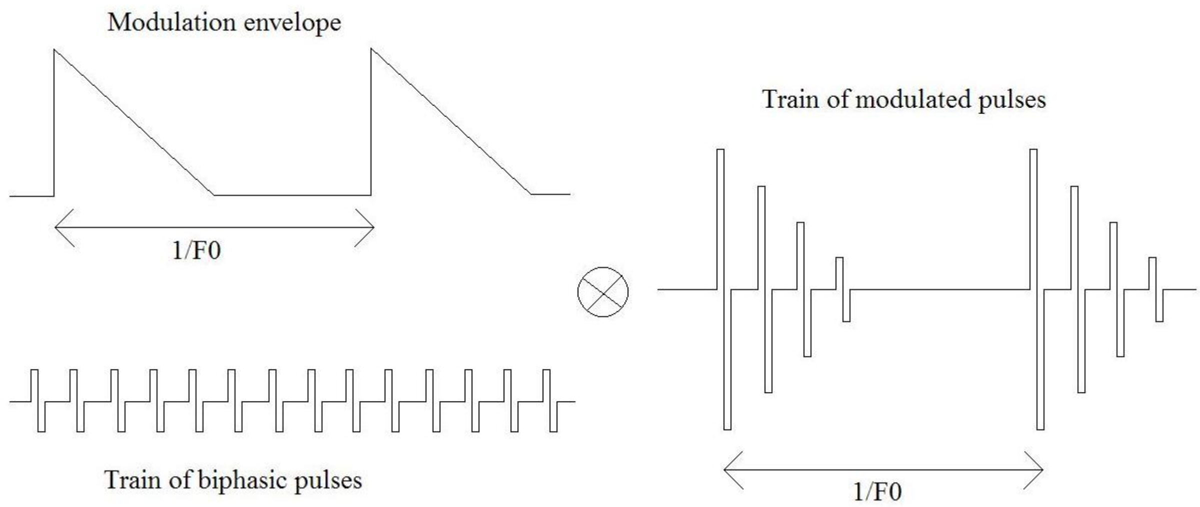

Figure 3.2: Modulation of the pulses by the sharpened sawtooth contour

3.4 Previous Results

This section presents the previous results on pitch perception observed in the department (Green [11]) at UCL. Several modulation envelopes were tested: sine, sawtooth and sharpened sawtooth as shown on figure 3.3.

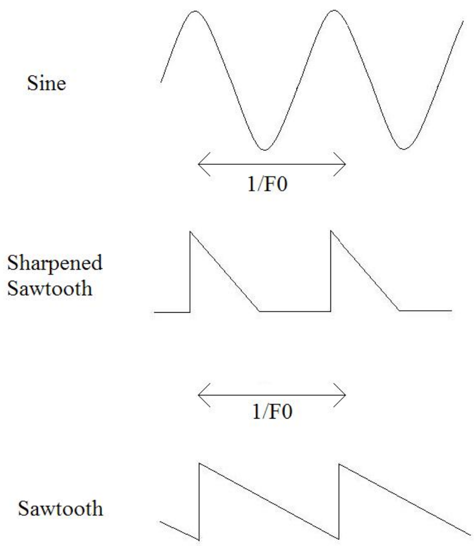

Figure 3.3: Shapes of the three contours previously investigated: sine, sawtooth and sharpened sawtooth

Two series of experiments were made to evaluate these different processing conditions (corresponding to several envelope shapes) and their influence on pitch perception. One was made with cochlear implant patients and another with normally hearing listeners using noise-excited vocoder simulations of CIS implants (as used by Shannon [15]). Glide labelling performance in the noise carrier simulation indicated that intonation information could be derived from temporal envelope cues. A temporally sharpened envelope did enhance the salience of temporal cues. However, as already observed by previous works, the utility of temporal cues declined as F_0 increased towards the higher end of the human voice range.

With cochlear implant patients however, the shape of the modulation envelope didn't seem to have a strong influence. In a similar study, Geurts and Wouters ([7]) found no significant differences between the standard CIS strategy and their modified one (where fluctuations of the temporal envelope were increased). Finally it is important to say that individual differences between implant users were strong and the only noticeable result of their experiment was a decrease of performance in glide labeling with increasing F_0. That observation was thus consistent with the simulations.

One possible reason for finding no differences across various envelope shapes may be the specific procedure of the experiment. Indeed, in every block, different processing conditions were mixed, which can be a hard task for patients, and even harder if they are not used to CIS strategies. On tables 3.1 and 3.2 are some results of an experiment where the blocks corresponding to each processing conditions were separated. Three processing conditions have been tested (standard CIS, sawtooth and sharpened sawtooth). The three patients were asked to label vowel glides as raising or falling. The center frequency was either 113 Hz either 226 Hz. The

Tables show, for each condition, values of slope and intercept obtained by a logistic regression on the data (The higher the slope, the better the performance). The modulation didn't seem to have any effect on the performance of patient LN. Patient PG however showed an improvement with the sharpened sawtooth compared to the standard CIS for the two frequencies. We will come back to these results later as the three patients have taken part in the experiment described in chapter 4.

LN			**PG**			**RH**		
	slope	intercept		slope	intercept		slope	intercept
CIS 113 Hz	14.19	0.00	CIS 113 Hz	10.21	-0.24	CIS 113 Hz	8.80	-0.57
Saw 113 Hz	15.51	-1.09	Saw 113 Hz	9.93	0.16	Saw 113 Hz	4.75	-0.62
Shp 113 Hz	15.55	-0.39	Shp 113 Hz	13.83	0.05	Shp 113 Hz	5.54	-0.16

Table 3.1: Previous Results (glide-labelling experiment: the center-frequency of the glides is 113 Hz. Values of the slope and the intercept after logistic regression on the results.

LN			**PG**			**RH**		
	slope	intercept		slope	intercept		slope	intercept
CIS 226 Hz	8.15	-0.22	CIS 226 Hz	5.86	-0.08	CIS 226 Hz	1.26	-0.80
Saw 226 Hz	9.32	-0.40	Saw 226 Hz	10.85	-0.17	Saw 226 Hz	4.14	-0.44
Shp 226 Hz	8.64	-0.66	Shp 226 Hz	10.70	-0.12	Shp 226 Hz	4.32	-0.82

Table 3.2: Previous Results (glide-labelling experiment: the center-frequency of the glides is 226 Hz. Values of the slope and the intercept after logistic regression on the results.

Chapter 4

Experiments

4.1 Description of the interface

Users of the Clarion cochlear implant C1 have been taking part in experiments to compare two signal processing conditions: standard CIS compared to CIS with modified carriers (using the sharpened sawtooth modulation envelope). Stimuli were generated on a computer and sent to the CLARION Research Interface (CRI). This interface allows the use of a DSP (Digital Signal Processing) board with a Synchronous Serial Interface to directly control stimulus patterns on all 16 Implantable Cochlear Stimulator electrodes simultaneously (8 bipolar electrodes). For the experiments, patients had just to remove the external part of their own device (held to the implanted electronics by a magnet) and use the one linked to the DSP board (as shown on figure 4.1). The system uses a Motorola DSP 56302 evaluation board. A part of my work during these months was to build a Matlab function (MEX-file) in order to control the interface through Matlab. This allows more or less any auditory task, including those involving speech to be run in Matlab. The original software that controls the interface was written in Visual C++ and had been incorporated into a pre-existing C++ program that ran certain forms of psychoacoustic test (2-alternative forced-choice paradigms). That's why speech perception tests (that need more than two choices) couldn't be run. The generated mex-file performs two simple functions: the initialisation of the interface and the sending of a stimulus. These two functions can be called from a Matlab program which provides an easy way to build a graphical user interface to run a speech perception test.

All the stimuli used in the experiments were generated with Matlab and sent to the implant by the interface. The generation of the stimuli was done in several steps. The signal was first bandpassed by eight different filters. The cut-off frequencies of the filters are shown on table 4.1, corresponding to the values of the implant filters. The extraction of the envelopes was done by applying a full-wave rectification (absolute value of the signal) and then low-pass filtering (4^{th} order Butterworth) the signals at 400 Hz for the Standard CIS and 32 Hz for the modified CIS . These envelopes were compressed and mapped between subject's threshold and MCL levels and were then used to modulate pulse trains which had already been modulated by the sawsharp contour (only in the case of the modified CIS strategy). To shape that contour (or envelope of modulation), the fundamental frequency of the signal was necessary, requiring the use of a laryngograph during the recording of the stimuli. For all experiments, we used the standard

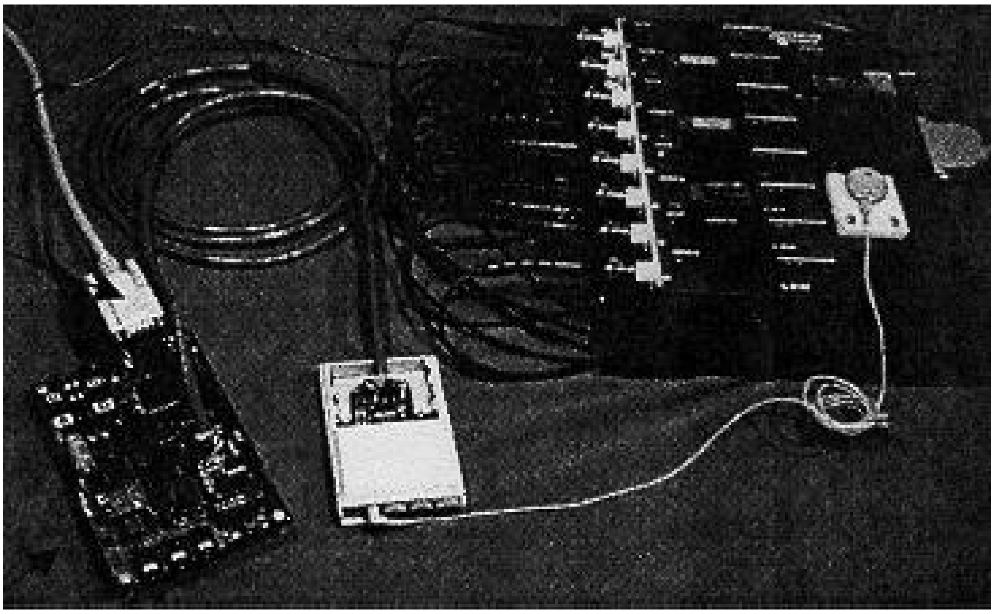

Figure 4.1: The Clarion Research Interface (from [22])

rate of the C1 implant for the CIS strategy (813 pulses per second and per channel) with a pulse duration of 77 μs.

lower freq (Hz)	250	500	730	1015	1450	2000	2600	3800
upper freq (Hz)	500	730	1015	1450	2000	2600	3800	6400

Table 4.1: Values of the cut-off frequencies of the band-pass filters used

4.2 Subjects

Five post-lingually deafened adults using the Clarion C1 cochlear implant device participated in this study. All were native speakers of British English. Three subjects were users of the SAS strategy and two subjects were users of the MPS strategy (the description of these two methods can be found in chapter 2). More details on the patients can be found in tables 4.2 and 4.3.

4.3 Experiment 1: Vowel Identification

4.3.1 Stimuli and procedures

Vowel recognition was measured in a 9 or 5-alternative identification paradigm composed solely of monophtongs. The choice to run a 5 vowel-set or a 9 vowel-set depended on the subject's performances, measured during the first blocks of the experiment. All vowels were presented in a /b/vowel/d/ context. The investigation was made using the words bead, bird, board, bard,

Subject	Age	Years deaf.	Years Exp.	Etiology
PG	72	28	4	Unknown
LN	76	3	4	Unknown
LC	74	5	4	Otosclerosis
RH	72	5	3	Skull fracture
CK	58	40	2	Mastoidectomy; sudden h. loss

Table 4.2: Subject initials, ages, number of years of deafness before surgery, number of years with the implant, Etiology of deafness.

Subject	Strategy	Nb channels	Threshold (mean C.U.)	MCL (mean C.U.)
PG	SAS	8	60	633
LN	MPS	8	91	518
LC	MPS	7	122	369
RH	SAS	8	57	198
CK	SAS	8	64	323

Table 4.3: Subject initials, signal processing strategy, Number of channels used, mean Threshold and mean MCL (Maximum Comfortable Level).

booed, bid, bed, bad, bud for the 9 vowel-set and bead, bird, bard, booed, bad for the 5-vowel set. These 5 vowels were chosen to be most perceptually distinct, according to the values of their formant frequencies (figure 6.1 in appendix 1). Patients LC, RH and LN did the 9-vowel test whereas patients CK and PG did the 5-vowel test.

The stimuli for this test were natural productions from one man and one woman, both native speakers of British English. Two examples of each vowel had been recorded in an anechoic chamber at UCL. Each test block included 36 tokens for the 9 vowel-set (9 vowels / 2 talkers / 2 examplars) and 40 tokens for the 5 vowel-set (5 vowels / 2 talkers / 2 examplars / 1 repetition). A stimulus token was randomly chosen and presented to the subject. Following the presentation, the subject responded by pressing one of the 9 (respectively 5) buttons, each marked with one of the possible responses. The responses buttons were labeled in a /b/vowel/d/ context as can be seen in figure 6.2 in Appendix 2. Feedback was provided by highlighting in yellow the right response button after each response given by the subject. In addition, practice was given before each block: the patients could hear any vowel by clicking on the corresponding button.

The experiment consisted in comparing the standard CIS strategy and the sharpened sawtooth modified-CIS strategy. The patients had 2 blocks to train in a free-field condition: the stimuli were the original recorded words played on a loudspeaker. The patients were thus using their own implant. After that, several blocks of either standard CIS processed vowels or sharpened sawtooth processed vowels were played through the interface. The first two blocks have not been taken into account in the results, because most of the patients were improving their performance progressively, getting used to the processing conditions (which are really different from that they normally use, at least for the SAS users) and to the voices of the speakers.

For the 9-vowel test, 3 blocks (3 times 36 vowels) of each processing condition have been taken into account for LN and 4 blocks for LC and RH. For the 5-vowels test, 2 blocks (2 times 40 vowels) have been taken into account (PG and CK).

4.3.2 Results

The results are presented on figure 4.2. Two patients (LN and RH) show a decrease of performance for the sharpened sawtooth strategy comparing to the standard CIS. The three other subjects get similar results for the two processing conditions. Patients PG and CK, even with the 5-vowel test, show very low performance. There is no significant effect of the sex of the speaker.

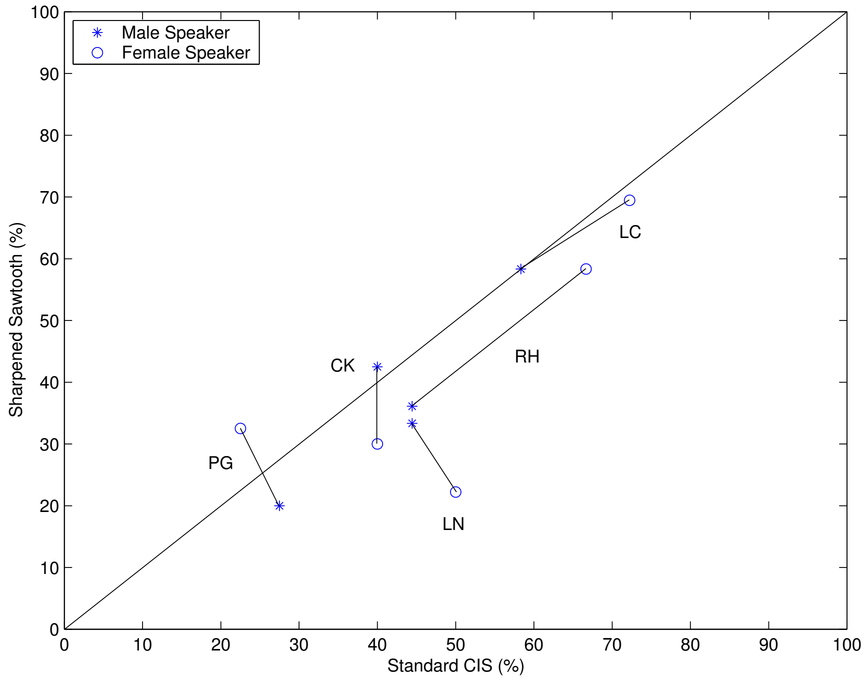

Figure 4.2: Median scores (percent correct) in the vowel identification test for each patient. Plot: results with the sawsharp Vs. results with the standard CIS.

A statistical test of proportions was performed to evaluate the differences observed between the processing conditions ([20]). The test was done for each subject, mixing the results from male and female speakers (table 4.4). Only patient LN had a significant decrease of performance with the sawsharp modulation (level of significance $p < 0.05$).

One has to notice that most of the patients reported a degraded sound of the vowels for the sharpened sawtooth compared to the standard CIS, even if their results didn't show it finally.

Subject	PG	LN	LC	RH	CK
sig.	0.429	0.002	0.401	0.119	0.312

Table 4.4: Statistical significance of the difference of proportions between the two strategies. Vowel Identification Test.

4.4 Experiment 2: Question/Statement Contrast test

4.4.1 Stimuli and procedures

Pitch perception was measured using question/statement contrast tests. The stimuli were processed sentences (c.f. appendix) either spoken like a question (raising pitch at the end) or spoken like a statement (falling pitch). The materials consisted in 240 sentences (30 sentences spoken like a question and 30 like a statement, by two different speakers and in two different processing conditions). The speakers were native speakers of British English but different than those of the vowel identification test. The processing conditions were the standard CIS and the sharpened sawtooth modified CIS.

The test was divided into 3 blocks. Each block contained 10 sentences of the list (either 1 to 10, or 11 to 20, or 21 to 30) as shown in Appendix 3. The stimulus token was chosen randomly among the different speakers, conditions and formulation (question or statement). Each block was presented twice. After each presentation of a stimulus token, the subject had to push one of the two buttons "question" or "statement". No feedback was provided. As for the vowel-test, a practice was given before each block: the subjects could hear examples of the sentences they were about to label.

4.4.2 Results

Figure 4.3 shows a boxplot of the overall results. One can notice a little improvement on the task with the modified-CIS strategy. When looking to the individual results (figure 4.4), this improvement is much more salient. Most of the points of measurement are situated above the diagonal. The patient PG did the task almost perfectly. The better performances of PG and LN compared to other patients may result from their wider dynamic range as shown on table 4.3 (absolute difference between threshold and MCL (Maximum Comfortable Level)). In the previous results (chapter 3), Green & al. ([11]) had observed modulation had no effect on LN's performances in glide-labelling. Here, she seemed to be strongly influenced by the sharpened contour (at least, the most influenced subject). This might be due to variability of some cochlear implant patients or just to the nature of the task (in the glide-labelling experiment, sounds were synthethised vowels whereas here, the stimuli are naturally spoken sentences). The patient PG showed a better performance with the sharpened sawtooth method, as already seen in the previous results.

A statistical test of proportions was performed (same as for the vowel identification test) as shown on table 4.5. Improvements made by PG and LN with the sawsharp modulation are statistically significant. The difference of proportions observed for RH are not relevant because his results are not much above chance.

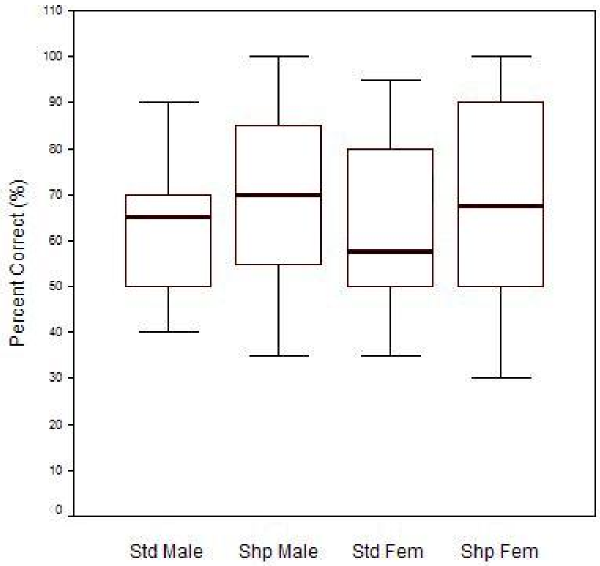

Figure 4.3: Overall results of the Question-Statement test for the 5 cochlear implant patients. Boxes differ by the sex of the speaker and the processing condition. Each box represent 30 values. (6 blocks / 5 patients). Std=Standard CIS, Shp=Sharpened Sawtooth.

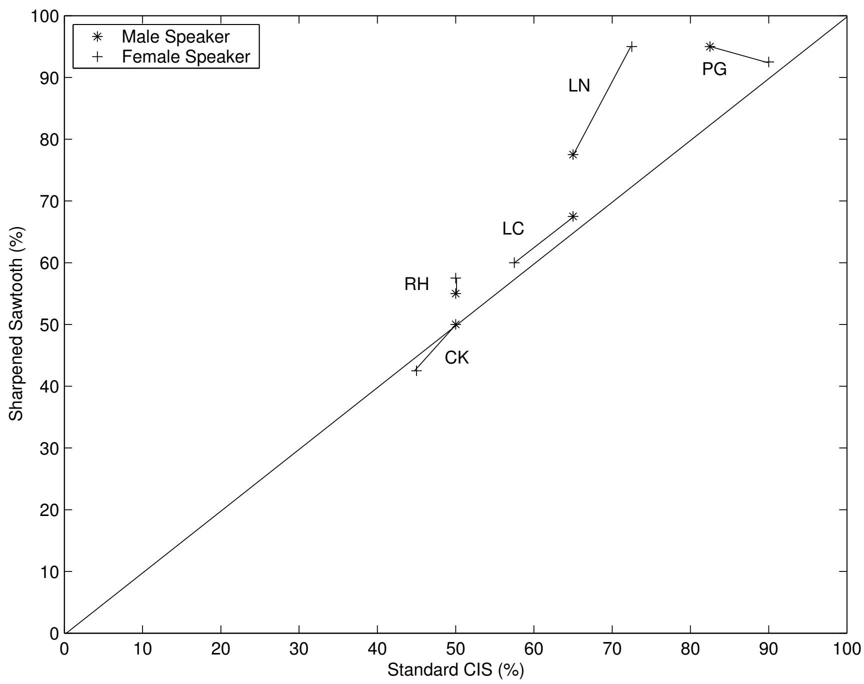

Figure 4.4: Median scores (percent correct) in the Question/Statement test for each cochlear implant patient. Plot: results with the sawsharp Vs. results with the standard CIS.

Subject	PG	LN	LC	RH	CK
sig.	0.029	$< 10^{-5}$	0.255	0.014	0.394

Table 4.5: Statistical significance of the difference of proportions between the two strategies. Question/Statement Contrast Test.

	Standard CIS		Sharpened Sawtooth	
	Day 1	Day 2	Day 1	Day 2
Median Values (%)	55	65	75	67.5

Table 4.6: Comparison of results from two different days for subject LC. (Question/Statement test).

The patient LC came twice to the lab to do this question/statement test. The first time, he just did the test with the stimuli from the male speaker. The second time, he did it with the stimuli from the two speakers. The results shown on figures 4.3 and 4.4 only present the results from the second experiment. It is however interesting to compare them with those from the first test. On table 4.6 can be seen the different median scores from the two tests, indicating a strong variability in the results. Indeed, results from the first test show a huge difference in performance between the two conditions whereas, in the second experiment, there is hardly any difference at all. Therefore, one has to take care of the results observed on one single day not to come to hasty conclusions.

Chapter 5

Noise-vocoding Simulations

5.1 Generation of stimuli

Eight-band noise-excited vocoders were used in acoustic simulations to investigate temporal cues to both pitch and vowel quality for the two processing strategies (standard CIS and sharpened sawtooth). As we have already seen in chapter 3, noise vocoding simulations of cochlear implants have been much used to investigate the cues available in cochlear implants by testing normally hearing listeners. Obviously, the aim is not to reproduce what users of implants are able to hear because there is a strong variability across patients (due to the etiology of their deafness, the number of years of deafness, their available dynamic range, etc). These simulations give however an idea of the cues available in implants.

The generation of stimuli was done in a similar way to that for the tests with patients. The aim of this experiment was to reproduce an 8-channels implant (C1 from Clarion). Thus, 8-band noise-vocoding simulations were used. The initial signal was analysed by 8 different bandpass filters. Envelopes of the signals were then extracted by full-wave rectification and low-pass filtering at a cutoff frequency of 400 Hz (for the standard CIS) and 32 Hz (for the sharpened sawtooth). Eight independent noise-carriers were then modulated by the envelopes after being modulated by the specific waveform for the sharpened sawtooth strategy and the signal generated was band-pass filtered by the same filters as in the first step of the process. The stimulus was finally obtained by adding the eight signals and low-pass filtering the sum at a cutoff frequency of 7 kHz, in order to suppress any possible distortion at high frequencies. The major difference with the electrical stimuli is that we used here noise-carriers rather than pulses.

5.2 Subjects

Five normally hearing native speakers of British English(4 females and 1 male) took part in the experiment. They were aged between 22 and 25. Three of them had already heard such type of simulations and two of them had no experience at all. The two experiments lasted approximately 2 hours and the subjects were paid 10 pounds.

5.3 Experiment 1: Vowel Identification

5.3.1 Stimuli and procedures

The stimuli were from the same recordings as those used in experiments with patients. However, here, vowel recognition was measured in a 12-alternative identification paradigm. Indeed, for the 9-vowel test, all subjects showed very good performance for the two processing strategies (more than 80 %). Thus, we added three more vowels (bared, bode and beard). Each block contained 48 vowels (12 vowels / 2 speakers / 2 signal processing conditions) and performance was measured on the six last blocks of the experiment (3 blocks for the standard CIS and 3 blocks for the sharpened sawtooth modified CIS). We kept the same conditions as for the experiment with patients (practice before each block and feedback provided).

5.3.2 Results

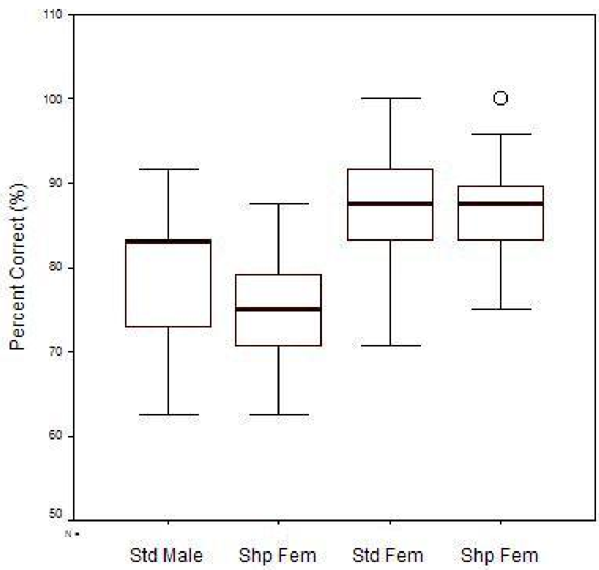

Figure 5.1: Overall results of the Question/Statement test for 5 normally hearing listeners. Boxes differ by the sex of the speaker and the processing condition. Each box represent 15 values. (3 blocks / 5 patients). Std=Standard CIS, Shp=Sharpened Sawtooth.

Global results (figure 5.1) indicate the distribution of results were in a smaller range than results from patients. No substantial effect of the strategy can be found. The individual scores (figure 5.2) show slightly better results for the Standard CIS. Moreover, one can notice results from the male speaker are further from the diagonal than results from the female speaker (at least for subjects JM and LyS). This suggests there is less difference for (higher-pitched) female voice, which is consistent with previous results (chapter 3).

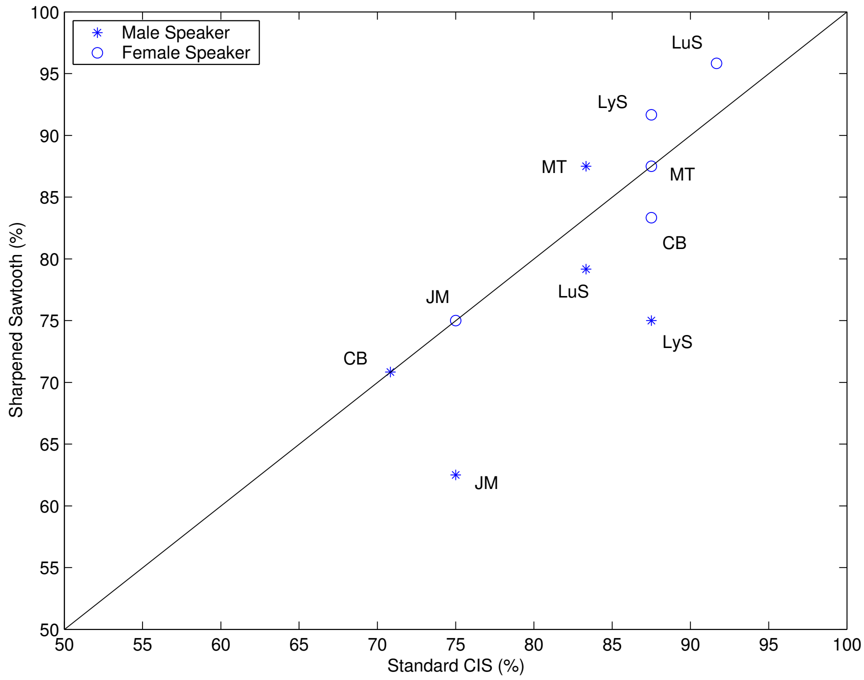

Figure 5.2: Median scores (percent correct) in the Vowel Identification test for each normally hearing listeners. Plot: results with the sawsharp Vs. results with the standard CIS.

source	wave	sex	order	wave*sex	wave*order	sex*order	wave*sex*order
signif.	0.379	0.004	0.299	0.400	0.234	0.941	0.880

Table 5.1: Results of the Repeated-measures analysis of variance for the vowel identification test. "signif." represents the statistical significance of the influence of the following factors: "wave" is the processing condition, "sex" is the sex of the speaker and "order" is the rank of the block. This analysis was performed with SPSS.

A repeated measures analysis of variance was performed. It showed the only statistically significant effect was the sex of the speaker (table 5.1).

5.4 Experiment 2: Question/Statement Contrast test

5.4.1 Stimuli and procedures

The stimuli have been generated from the same recordings as those used in the experiments with patients, and using the same procedure (chapter 4).

5.4.2 Results

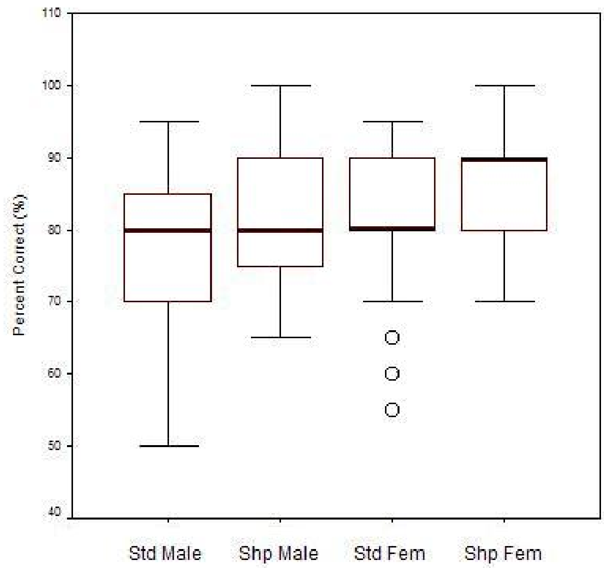

Figure 5.3: Overall results of the Question/Statement test for 5 normally hearing listeners. Boxes differ by the sex of the speaker and the processing condition. Each box represent 30 values. (6 blocks / 5 patients). Std=Standard CIS, Shp=Sharpened Sawtooth.

The overall results (figure 5.3) didn't show strong differences between the processing strategies nor concerning the sex of the speakers. The individual results (figure 5.4) indicated however that subjects performed better with the modified CIS strategy, particularly with the female voice.

A repeated measures analysis of variance run on these data has shown the processing strategy was a statistically significant factor (table 5.2). Moreover, There were no significant interactions involving processing strategy, suggesting the difference between the two processing conditions observed in figure 5.4 is significant. No overall effect of the sex of the speakers was found. The order of the blocks as well as the crossed variable "sex * order" have a significant effect but it

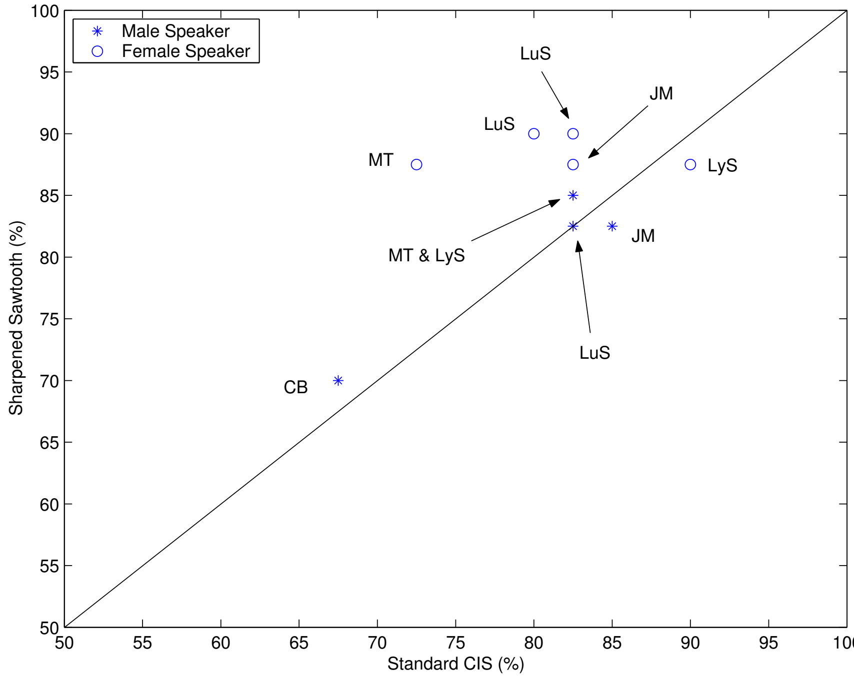

Figure 5.4: Median scores (percent correct) in the Question/Statement test for each normally hearing listeners. Plot: results with the sawsharp Vs. results with the standard CIS.

is not especially associated with improvement of subjects with time. It just shows a variability in the results across the blocks.

source	wave	sex	order	wave*sex	wave*order	sex*order	wave*sex*order
signif.	0.014	0.252	0.01	0.582	0.477	0.012	0.164

Table 5.2: Results of the Repeated-measures analysis of variance for the question/statement tets. Same process than for table 5.1.

Chapter 6

Conclusions

Cochlear implant patients have shown widely varying results. This is probably partly due to the history of their deafness and of their implantation. Yet, considering the results one by one, it appeared that two patients (PG and LN) could improve their perception of question/statement contrasts with the sawsharp modulation. And overall, the three patients who answered above chance for this test have shown better median scores with the modified CIS. This suggests our modulation can convey more salient cues to pitch than the classical CIS strategy. Moreover, this hypothesis was reinforced by the results from simulations where subjects have shown an statistically significant improvement.

In the test of vowel identification, two patients had a decrease of performance with the sawsharp modulation (and only one of them was significant: patient LN). The three others didn't show any differences in their results with the two methods. In the simulations, no significant difference was found neither. This leads us to think the effects of the sharpened modulation were not drastically bad for the perception of vowels.

It is interesting to see that patients reacted really differently to the strategies evaluated. LN showed the biggest differences between the two strategies, and for the two tests. This is all the more strinking since she didn't seem to be strongly influenced by the sawsharp modulation in a previous experiment of glide-labelling. Thus, we have to insist on the effects of variability for each patient. Repeating the measures on several days might help to evaluate the performances of each patient in a more precise way.

Further works and Acknowledgements

More patients are needed to complete this study. It is not possible to say whether the sharpened sawtooth modulation gives any better perception of pitch. Even if the simulations let us think that this modified CIS strategy can improve pitch perception (but not at the expense of speech perception), the results of implant users don't allow a clear conclusion. It would be interesting to continue this study with other patients and also to add other tests such as consonant perception and sentence recognition.

The recent CII cochlear implant system (Clarion) can deliver stimulation rates over 5000 pps to 8 electrodes (whereas the C1 system allows a maximum rate of 813 pps). The future research of the laboratory will continue to investigate such types of processing strategies and more precisely, try to determine the availibility of voice pitch information as a function of pulse rate. This will be possible by testing users of that CII implant. One can imagine a higher pulse rate will allow a better shaping of the trains of pulses by any special contour (such as the sharpened sawtooth we have studied here). Moreover, some works suggest that higher rates may lead to patterns of neural activity that more closely resemble those characteristic of normal hearing (Matsuoka & al. [14]). The future studies will thus attempt to clarify the understanding of the several effects of pulse rate.

These four months spent in the laboratory were very interesting, I've learnt a lot and not only on cochlear implants... partly due to the lectures I've followed and the many different subjects people are working on. I thank Stuart, Tim and Andy for their useful pieces of advice, the discussions, and their kindness. A special thanks to Stuart for the review of this report. And finally, I thank all the people who took part in the experiments presented here.

Appendix

Appendix 1: Phonetics

This appendix shows the differences across vowels, according to their first two formant frequencies. (N.B: The values of these frequencies in a b/V/d context can differ a little from those found in a h/V/d context.)

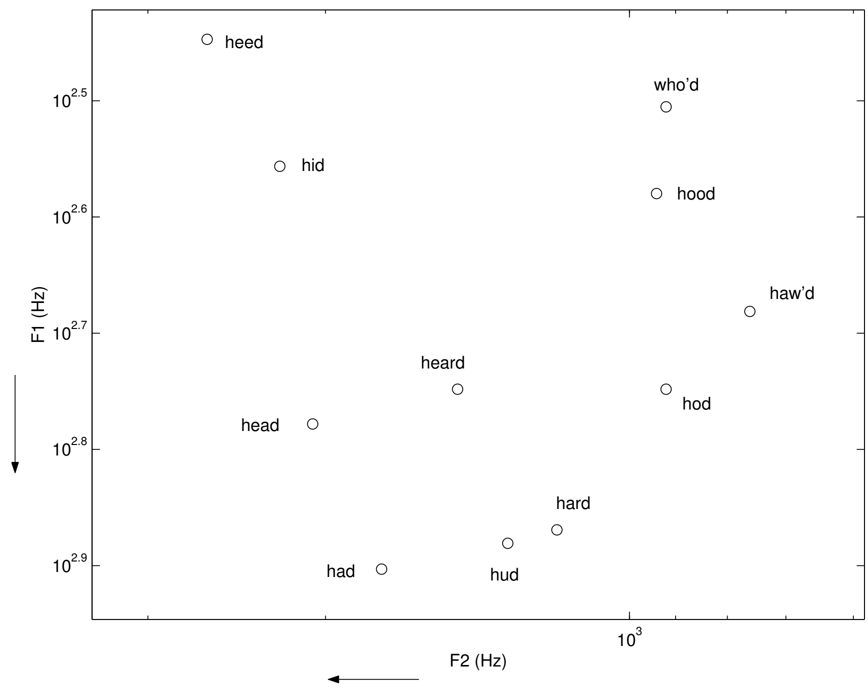

Figure 6.1: Mean of formant frequencies: plot of F2 vs. F1. Adult Male formant frequencies in Hertz collected by J.C. Wells around 1960

had	heed	hid	head	heard	who'd	hard	haw'd	hud	hod	hood
[æ]	[i]	[ɪ]	[ɛ]	[ɜ]	[u]	[ɑ]	[ɔ]	[ʌ]	[ɒ]	[ʊ]

Table 6.1: Phonetics symbols of 11 vowels in a h/V/d context

bad	bead	bid	bed	bird	booed	bard	board	bared	bud	bode	beard
[æ]	[i]	[ɪ]	[e]	[ɜ]	[u]	[ɑ]	[ɔ]	[eə]	[ʌ]	[əʊ]	[ɪə]

Table 6.2: Phonetics symbols of the 12 vowels used in the experiments

Appendix 2: Graphical user interface for the vowel-identification test (9 vowels)

Here is an example of the graphical interface programmed in matlab and used in the experiments.

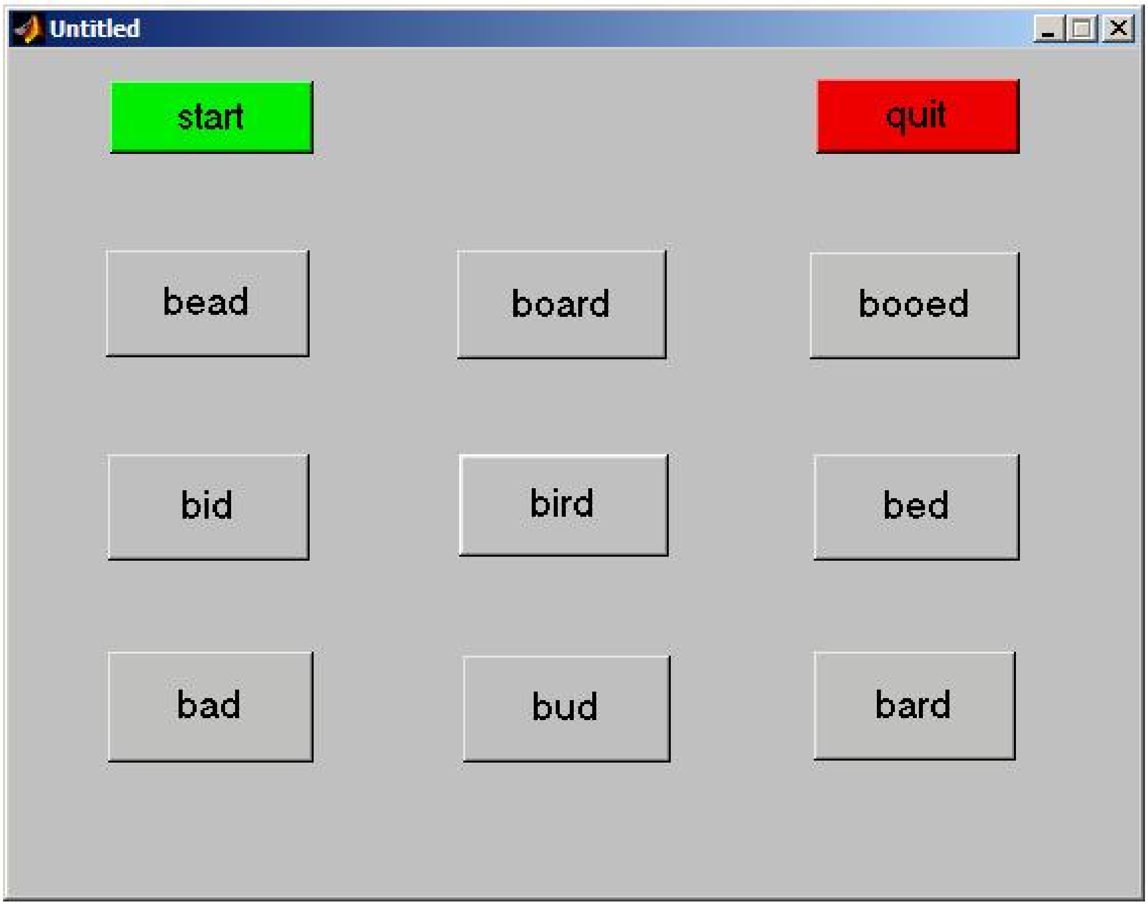

Figure 6.2: Matlab graphical interface

Appendix 3: Question/Statement Sentences

This appendix lists the sentences used for the question/statement test.

From files in the lab (Dept of Phonetics, UCL):

1. They're playing in the garden.
2. It's pouring with rain.
3. She's reading a newspaper.
4. He nearly missed the bus.
5. He was eating a peppermint.
6. He's telephoning now.
7. She didn't want to go.
8. He's learning to drive.
9. He almost crashed the car.

From Matt's project (Lieberman & Michaels (1962), JASA 34, 922-927)

10. They've bought a new car.
11. His friend came by train.
12. You've seen my new house.
13. John found him at the phone.
14. The lamp stood on the desk.
15. They parked near the street light.
16. We talked for a long time.
17. He'll work hard next term.

From MAC test (standard test):

18. They saw the movie.
19. All the way.
20. It's down there.
21. They're not home.
22. It's not enough.
23. At the first house.
24. He can't go.
25. They ate there.
26. Open this one.
27. He'd rather not.
28. That will do it.
29. But not for me.
30. It's his brother.

Bibliography

[1] Philipos C. Loizou. *Mimicking the human hear.* IEEE Signal Processing Magazine, Sept. 1998: 101-130.

[2] Philipos C. Loizou. *Signal Processing techniques for cochlear implants.* IEEE Engineering in Medecine and Biology, May/June 1999: 27-33.

[3] Philipos C. Loizou. *Introduction to cochlear implants.* IEEE Engineering in Medecine and Biology, Jan/Feb 1999: 32-42.

[4] Francis A. Spelman. *The Past, Present, and Future of Cochlear Prostheses.* IEEE Engineering in Medecine and Biology, May/June 1999: 34-46.

[5] B.S. Wilson & al. *Better Speech recognition with cochlear implants.* Nature, 352: 236-238, 1991.

[6] Qian-Jie Fu and Robert V. Shannon. *Effect of stimulation rate on phoneme recognition by Nucleus-22 cochlear implant listeners.* JASA, 107: 589-597, 2000.

[7] Luc Geurts and Jan Wouters. *Coding of the fundamental frequency in continuous interleaved sampling processors for cochlear implants.* JASA, 109(2): 714-726, 2001.

[8] Fan-Gang Zeng. *Temporal pitch in electric hearing.* Hearing Research, 174: 101-106, 2002.

[9] McKay, Mc Dermott & Carlyon. *Place and temporal cues in pitch perception: are they truly independent?.* ARLO, 1(1), Jul 2000.

[10] T. Green, A. Faulkner and S. Rosen. *Spectral and temporal cues to pitch in noise-excited vocoder simulations of continuous-interleaved-sampling cochlear implants.* JASA, 112(5): 2155-2164, 2002.

[11] T. Green, A. Faulkner and S. Rosen. . ARO meeting, February 2003.

[12] B. Townshend & al.. *Pitch perception by cochlear implant subjects.* JASA, 82(1): 106-115, 1987.

[13] Sipke Pijl & Dietrich W.F. Schwarz. *Melody recognition and musical interval perception by deaf subjects stimulated with electrical pulse trains through single cochlear implant electrodes.* JASA, 98(2): 886-895, 1995.

[14] Matsuoka & al.. *The effects of interpulse interval on stochastic properties of electrical stimulation: Models and measurements.* IEEE transactions on Biomedical Engineering, 48: 416-424, 2001.

[15] R.V. Shannon & al.. *Speech recognition with temporal primarily cues.* Science, 270: 303-304, 1995.

[16] Rob Drullman & al.. *Effect of temporal envelope smearing on speech reception.* JASA, 95: 1053-1064, 1994.

[17] Peter Roach. *English Phonetics and Phonology, a practical course.* Cambridge University Press (second edition), 1943.

[18] William A. Yost. *Fundamentals of hearing, an introduction.* Academic Press (fourth edition), 2000.

[19] E. de Boer. *On the Residue and Auditory Pitch Perception.* In Handbook of Sensory Physiology Vol. V Part 3, Springer Verlag, 1976.

[20] H.T. Hayslett & Patrick Murphy. *Statistics made simple.* Doubleday & Company, Inc. 1968.

[21] S. Zimmerman Phillips & C. Murad. *Programming features of the clarion multi-strategy cochlear implant.* Annals of Otology, Rhinology, Laryngology. 108: 17-21, 1999.

[22] Advance Bionics Corporation. *Clarion Research Interface User Manual.*

Contents