THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par : *l'Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)*

Présentée et soutenue le *27 / 04 / 2016* **par :**

EDWARD COLLETT

Strategies of Auditory Categorisation in Cochlear Implant Users and Normal Hearing Listeners.

JURY

PASCAL BELIN	Professeur d'Université	Membre du Jury
PASCAL BARONE	Professeur d'Université	Membre du Jury
OLIVIER DEGUINE	Professeur des universités	Membre du Jury
PASCAL GAILLARD	Maître de conférence	Membre du Jury
GUILLAUME LEMAITRE	Chargé de Recherche	Membre du Jury
EMMANUEL LESCANNE	Professeur des universités	Membre du Jury
ERIC TRUY	Professeur des universités	Membre du Jury

École doctorale et spécialité :
CLESCO : Neurosciences, comportement et cognition
Unité de Recherche :
Centre de Recherches Cerveau et Cognition (CNRS UMR 5549)
Directeur(s) de Thèse :
Pr Pascal Barone et *Pr. Olivier Deguine*
Rapporteurs :
Pascal Belin et *Eric Truy*

ABSTRACT

Auditory categorisation involves the grouping of acoustic events along one or more shared perceptual dimensions which can relate to both semantic and physical attributes. Whilst previous work has established certain categories (for example living and non-living, mechanical, natural and sounds) and identified the importance of perceptions concerning the sound producing event (object and action) the research domain is still to be fully understood.

The process of categorising sounds also involves both high level cognitive processes and low-level perceptual encoding of the acoustical signal, both of which are affected by the use of a cochlear implant (CI), such that auditory abilities are often severely diminished in comparison to normal hearing listeners (NHL).Existing studies have shown that CI users perform better at identifying the auditory category compared to a specific sound, however these studies have used a fixed and limited number of categories. A Free-Sorting-Task (FST) is used in the current work to test categorisation performance in a more real-world manner, whereby participants may categorise sounds however they choose. Subsequent analysis consisted of Multiple Correspondence Analysis (MCA) and HCPC to determine the categories and strategies used by participants.

The FST was used to determine how CI users duration of implantation (from 0 to more than 14 months) would affect the categorisation of everyday sounds. Results showed that experienced listeners performed similarly to NHL, separating vocal, musical and environmental sounds. Experienced CI also showed ability to separate a set of vocal sounds by emotional content.

NHL were also tested with only environmental sounds and demonstrated categorisation based on the sound producing object or action even when sounds were vocoded with only 4 channels. However context information was not strongly perceived and did not aid the identification of sounds.

The presented work adds important research to the study of auditory categorisation and is in agreement with previous works. Results of CI users show the potential of using the FST for further research and contains the possibility of using categorical perception as a useful tool in rehabilitation and assessment techniques.

ACKNOWLEDGMENTS

The completion of this work would not be possible without the significant support and help of many people throughout the past three years.

My thanks go firstly to my supervisors Pascal Barone & Olivier Deguine, for their guidance, patience and refusal to give up on me even in difficult times.

Various researchers and colleagues have also given me considerable amounts of their time and energy throughout my time working on this project. Thank you to everyone at Mirail, especially, Julien Tardieu, Pascal Gaillard and Aurore Berland for the various discussions which greatly helped me with my research.

Also a large thank you to Pr. El Mostafa Qannari of Oniris, Nantes and Brian Gygi for their very help and inspiration with certain aspects of the analysis.

Thank you to the members of the O.R.L at Hopital Purpan for always spending to time to help me with various aspects of patient logistics. Thank you to Pr. Mathieu Marx, ML. Laborde, M. Tartayre, G. Iversenc and all other members of staff who are part of a fantastic team.

To the interns who assisted me, Charles, Yoan, Vivianne and Juliette, who put in much hard work collecting and data and carrying out testing, thank you for your considerable time and effort.

My thanks also to Advanced Bionics for funding my research over the last years and to all the members of staff at AB France and AB Hannover.

Finally thank you to anyone, friends and family, colleagues and acquaintances simply for being there. You may never know just how important you have been.

Edd.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

CA Correspondence Analysis

CCC Cophenetic Correlation Coefficient

CI cochlear implant

CIL Cochlear Implanted Listener

ES	environmental sounds
ESP	environmental sound perception
FST	Free Sorting Task
HCPC	Heirarchical Clustering based on Principal Components
IT	Inferior Temporal Cortex
MCA	MultiCorrespondence Analysis
MDS	Multidimensional Scaling
NHL	normal hearing listeners
NS	Natural sounds
STG	superior temporal gyrus
STS	superior temporal sulcus
EXP	Experienced cochlear implant listeners
INT	Intermediate cochlear implant listeners
NEW	New cochlear implant listeners

1 INTRODUCTION

In everyday life we encounter many different sounds on a daily basis, and often are situated in busy auditory environments, for example a sunday market, that involves many different sound sources. Paying attention to each individual sound, whether it be a nearby person talking, a passing car, coins rattling in your own pocket, would require a large amount of cognitive effort and as a result the auditory system makes use of certain processes to attain to only the most important/relevant information. One process that helps in this manner is auditory categorisation. By identifying or perceiving the category a particular sound belongs to, information about the sound ca be more efficiently extracted from the category knowledge. This also means that the effort required to identify the specific sound is also removed.

This process has been studied in the perception of phonemes, to understand perceptual boundaries, for example when a //ba// sound becomes a //ga// sound and what acoustical properties are important for this change in perception. Work has also been done to show how sounds can be categorises based on the material, size and shape of the sound producing object (references). Regarding the categorysation of more everyday sounds common results have shown categories corresponding to living and non-living objects, for example human and animal vocalisations, human action sounds, mechanical (including too and transport), nature sounds and electronic or synthesised sounds (references). Studies have even shown specific areas of the brain responsible for processing these different types of sounds [78, 42, 79].

The way in which sounds are grouped together to form these categories is based on perceptual similarities. In general it is considered that similarities concern information that is related to the perception of the sound producing event/object in terms of its size, material, action etc. or action, cited as causal similarities by Lemaitre. Another form of similarity is the semantic associations to the sound, for example concepts such as events (parties, sports events) or emotional content (e.g. happy vs sad).

Perceptions of the physical sound signal are also of course important, for example the spectral content, harmonicity, frequency, sound power level, temporal modulations. However these appear to be useful/used in categorisation only when judging the quality or musicality of sounds, and this act of listening to the physical acoustical signal has subsequently been termed musical listening (Gaver) whilst the previous perceptions of causal and semantic information are perceived when a listener enters into everyday listening mode.

It seems that in general categories are based on causal information, which would make sense as often we may hear a sound before seeing it and want to use the information within the sound to inform us about the object/event that produced the sound. This may further be related or include how we may be able to interact with the object.

However the existing research of auditory categorisation is not exhaustive, there is for example no full classification system proposed beyond that of Gaver. And whilst certain broad categories have been demonstrated, there also lies the opportunity to study both sub and super category membership to more fully understand how we perceive auditory information. Finally the relatively small number of studies have also not dealt with how sounds are perceived in the real world, in different environments and under different task requierments.

Auditory categorisation also involves both high level cognitive processes and low-level perceptual encoding of the acoustical signal, both of which are affected by the use of a cochlear implant (CI). Used to rehabilitate people with severe (or complete) hearing impairments, these devices deliver a degraded auditory signal via electric impulses directly to the auditory nerve within the inner ear. Such patients also often have cortical brain reorganisation as a result of their hearing impairment.

It is known that patients with such devices can develop very good perception of speech sounds and the majority of work, both research and rehabilitation is targeted around restoring speech abilities. However, although the perception of environmental sounds is often cited as a benefit to CI users, the accurate perception of these sounds is often lacking. Research is far behind that of speech, one of the main reasons being that environmental sounds encompass a vast range of sound sources all with different acoustical characteristics. This makes it difficult to study exactly which acoustic variables are important for environmental sound perception, and therefore difficult to make specific improvements within the CI device itself. it has also been shown that CI users ability to discriminate vocal from environmental sounds is poor in comparison to NHL (massida).

So how does auditory categorisation function with CI users? Some studies of CI categorical perception have been conducted, and shown that categorical perception is more accurate than the identification of individual sounds. However these studies used fixed categories, therefore leaving the question of how CI users may use categorical perceptions of sounds in the real-world.

The presented work therefore seeks to combine the fields of auditory categorisation with that of CI users auditory perceptions in order to further understand how the two domains of research overlap. It aims to investigate categorical perception in order to advance the understanding of CI users auditory abilities especially in dealing with different kinds of sounds. At the same time adding to the existing research to move forward with the understanding of auditory categorisation in in general. The following chapter introduces the theory and background information associated to the topics presented above in a more detailed manner for the reader to fully comprehend the context of the current work.

1.1 AUDITORY CATEGORISATION

1.1.1 *What is Categorisation?*

Categorisation as a process is one of the ways in which we as humans try to simplify our lives and make things easier. It allows us to deal with the vast amount of sensory input that is encountered on a day by day basis such that we can make quick and efficient judgements about the world around us - the objects we are seeing or the sounds we are hearing. For example by categorising an object as a car we are able to understand that it is likely an object with four wheels, an engine that runs on fuel and something that carries people from position A to position B. We do not need to analyse in detail the specific object to know if all of these things are true. So instead of treating every stimulus, every object as brand new we place it into a category of our own making in order infer knowledge about the category to the object without having to analyse in detail the object or stimulus in question.

Within the study of categorisation there are two commonly held theories of how we as humans go about categorising objects.

Prototype theory: which uses the idea of scaled category membership, for example for the category of birds, how "bird" or bird-esque is a particular bird? There exists a single representative or prototype of a particular stimulus inside your brain that is used to compare new stimuli. Such that if the new stimuli matches, then it is deemed to belong to the category. For example imagine a cow, you may form an image in your mind in which to describe certain qualities, that is your prototype. The prototype is therefore a mix of all the previously encountered stimuli in that category. When studying prototype theory it is very easy to obtain an idea of the strength of category membership or the proximity to the prototype by measuring reaction times of people asked to categorise certain objects. The faster the reaction time the stronger the category membership (or family resemblance as it is called) for example"violin" may be a more typical example of "music" than "bassoon" and therefore categorised more quickly. Prototype theory is often cited as being efficient due to the comparison of new stimuli with a single prototype and does not require the memorising or learning of many different category members - in difference to the exemplar theory.

Exemplar theory: the idea that categorising is based on comparing a new stimulus to previously encountered known examples.So for example the category of "fish" may be created by a person holding the memory of all the fish they have encountered in their life. By comparing a new new object to this category it is then possible to see if it has enough similarities in common with the category exemplars to be classified as a fish as well. Exemplar is useful as it can explain variability amongst category items. For example if the prototype of the category dog includes the requirements of an animal that has four legs and barks - what will be made of a three legged silent dog? Exemplar theory can also (to some extent) explain the creation of abstract (e.g. "beliefs" or "fears") and ad hoc (e.g. "things to take on holiday") categories through the combination of previous experiences, which is not possible with prototype theory.

Overall it is most likely that a combination of the two processes working together is used in daily life as sometimes one works better than another. This allows for the efficiency of prototype theory but also the flexibility of exemplar theory. Most likely this involves using exemplars that are encountered in order to modify or update the currently held prototype of a category.

However although there has been a large amount of research on categorisation it is still not fully understood how the categories we use in daily life come to be created as we are likely influenced by a number of wide ranging aspects. Not only will the frequency of encountering a specific stimuli modify the exemplars and the prototype but the recency of encountering a certain stimuli will also affect the category recognition such that if you were to encounter a bird singing and then immediately afterwards hear an ambiguous yet melodious sounds you might categorise this as birdsong, when in fact it was a flute or recorder being played.

Finally there are a number of things that affect the individuals categories such that one persons categories can be completely different to another persons. This of course will depend on culture, language and of experience - exactly what exemplars has the person witness and learned in their life? This allows categorises amongst people to be different, to be broader or more narrow, to have different subcategories and to have different exemplars and prototypes.The experiences of people can also be influenced by expertise which may give them an ability to create different levels of subcategories and more easily distinguish the borders between two categories. For example a professional musician may have many subcategories of different kinds of instruments and be able to easily distinguish between them all where as a laymen may have broader categories of “brass”, “woodwind” and “percussion” or maybe even simply a category of "musical instruments".

This aspect of familiarity and the experiences that people have had in their lives which govern their use of different categories is a very important aspect.Especially for this manuscript which has tested participants with many years of hearing impairment, such that their familiarity and experiences with auditory stimuli is vastly different to normal hearing listeners (NHL). In addition to this the hearing impaired participants tested are users of cochlear implant (CI) devices which further alters the experience with the auditory environment. Exactly how and to what extent hearing impairment and use of a CI may affect listeners categorisation of sounds is not at all known and will be an important question throughout this work.

1.1.2 *Similarity as a basis for categorisation*

The two proposed theories of categorisation, exemplar and prototype, share one common aspect that they both work on the basis of comparing a certain stimulus to a pre-defined representation of that stimuli in the persons knowledge - be it a single prototype or a collection of exemplars. It would therefore appear that categorisation is based on similarity judgements, but whilst similarity plays a role in categorisation they are two different processes. In general similarity judgements are more strongly based on the perception of a stimulus where as categorisation involves a persons theories, goals and other higher level cognitive processes. The process of judging similarity has also been described as being too flexible, too reliant on perception and does not fit alongside explanations of categories based on theories or objectives. If similarity can be influenced by abstract or alternate theories it must therefore become unconstrained. Or if similarity is restricted to only the direct perceptions of a stimulus then it does not meet the necessary requirements to act as the basis of categorisation.

1.1.3 *Why is auditory categorisation important?*

As mentioned before the process of categorising stimuli allows us as humans to more easily cope with the barrage of sensual information that surrounds us in every day life. This of course applies to the auditory domain where we are nearly constantly aware of the sounds that are going on around us. Being able to categorise these sounds enables us to infer information about the sound and the sound producing object from its category membership. Categories that have been found amongst tests of auditory categorisation have taken the presumption to define four broad categories of sound - being speech, music, environmental (referring to any sounds that are not speech or music) and finally artificial sounds, for example those of a computer or synthesiser [6]. Within specific studies of environmental sounds categories have also been found corresponding to warnings, water, vocalisations (both human and animal) [49, 119, 14, 33, 15]. Further studies have also identified that categories can be created based on the type of action or material that causes a sound for example hitting vs scraping or wood vs metal [57, 77, 79].

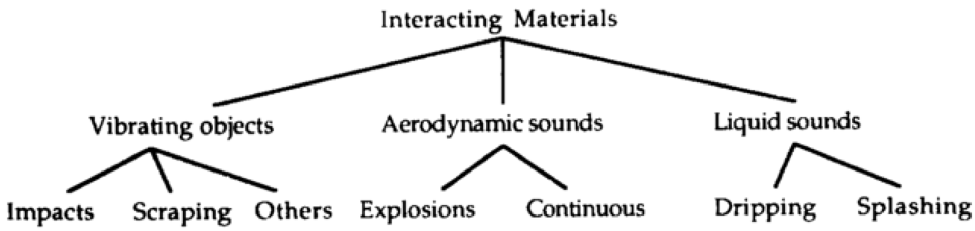


Figure 1.1: Hierarchical description of sound producing events taken from [37]. The first level describes simple interactions of materials, the second level examples of these interactions split into solids, gasses and liquids, and the final level examples of actions within each class of material

Regarding environmental sounds, Gaver [37] initially proposed a mapping related to the interaction of materials, see figure 1.1. At the first level it describes the cateogirsation of sounds in terms of "materials" which are split up into into vibrating objects (solids), liquids and aerodynamic sounds (gasses). The model then describes the interaction of materials at the second level. Thirdly comes the interactions of these materials and basic level interactions within each material set, which for solid sounds include hitting, scraping, for liquids splashing or dripping and for gasses motion caused by pressure changes, for example gusts of wind. More complex interactions are described at a higher level in terms of patterned events - i.e. the repetitive action of basic level events, also compound events which include the combination of basic level events and finally hybrid events that are made up from various materials and actions. These can be seen in figure 1.2 which gives a much more detailed look at Gaver's Hierarchy. This has also been further developed in other studies [57, 77, 91, 75, 79].

Within his ecological approach to auditory perception, Gaver also described two different "modes" of listening. The first, termed *Everyday listening* refers to the perception of the properties of the specific sound generating source, i.e. how big the object may be, how fast it may be moving. In general everyday listening relies on the identification of a sound, be it correct or incorrect and

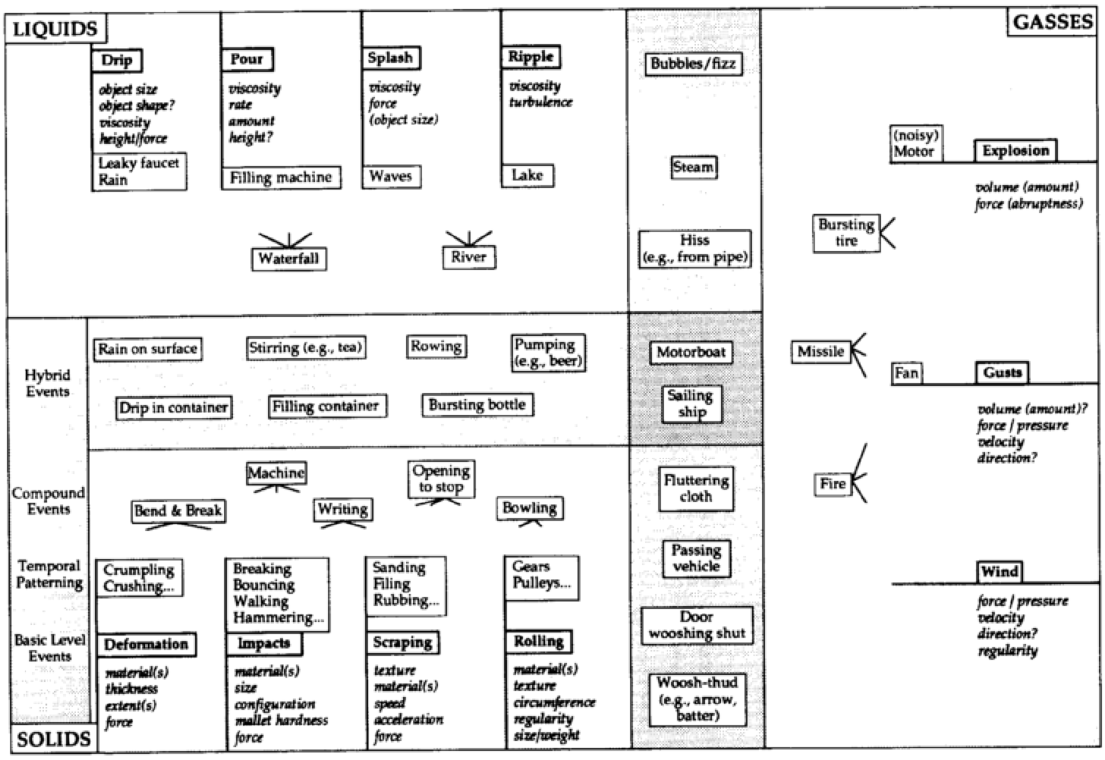


Figure 1.2: Gaver's map of everyday sounds [37]. An more detailed expansion on figure 1.1 whereby interactions of SOLIDS, LIQUIDS and GASSES are firstly divided up into three large sections corresponding the separate materials. Within each section examples of different basic sound producing actions are shown with their underlying attributes, for example how the *size* and *viscosity* of a liquid dripping will affect the sound produced. Examples of sounds that are created via the interaction of basic actions are given, for example a *bursting tyre* and *fire*. Where material sections overlap examples of sounds that are produced via two or three materials are given, for example *rain on a surface*, which involves both liquid and solid actions.

does not appear reliant on acoustic qualities.

The second listening mode, *Musical listening*, refers to the perception of qualitative aspects of the sound itself i.e."what the sound sounds like". This not only refers to acoustic characteristics of the sound, but also the sense of emotion or quality of the sound - e.g. the roughness or smoothness. More concerned with the perception of the raw acoustic signal rather than the sound producing object it has been seen to be encouraged by similarity judgements when directing participants to judge specifically similarity rather than simply categorise a sound. Musical listening and the focus on qualitative information about the sound also becomes important when identification fails, when there is no association of the sound to a specific object. In this sense it could be said that the perception of the listener falls back to the most salient information about the sound.

As a conclusion, from the small amount of studies that have looked at auditory categorisation It would appear that categorisation is predominantly based on the properties of the sound-producing object or event ([49, 52, 76, 6]). Rarely does it appear that the properties of the sound signal are used and when they are it is a result of identification being difficult or failing entirely.

1.2 HUMAN AUDITORY SYSTEM

1.2.1 *Overview*

The human auditory system is responsible for turning the oscillations of air pressure (that constitute sound) into electronic pulses which are then delivered via the auditory nerve to the brain where, depending on the signal different regions of the brain extract different information - for example speech or music and give a a subjective perception of the original sound. Figure 1.3 shows the auditory system which consists of three main areas- outer ear, middle ear and inner ear, each of which are separately described below.

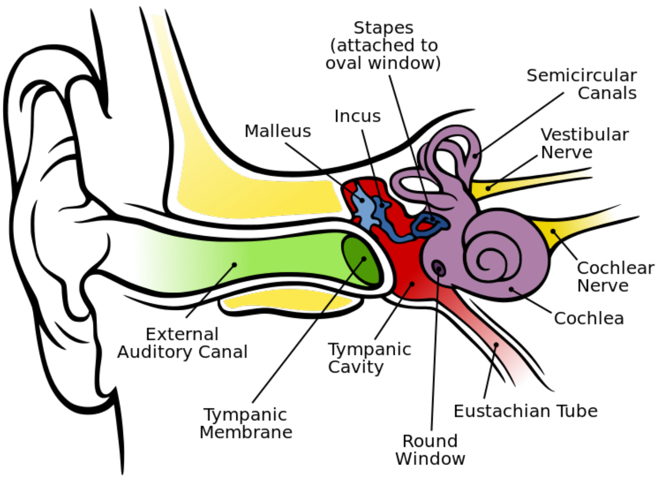


Figure 1.3: Graphic of the human auditory system taken from [26].

Outer ear: the visible part of the auditory system which consists of the pinna, ear canal and ear drum. Vibrations of air pressure are captured in the pinna, which acts like a kind of funnel for the sound, then transmitted along the ear canal before hitting the ear drum (aka tympanic membrane). The ear drum is a membrane stretch across the ear canal, much like a drum skin stretched over the top of the drum. When the vibrations of air pressure hit the ear drum it causes it to physically vibrate and these physical vibrations are then passed on to . . . . the middle ear.

Middle ear: contains the ossicles - three three tiny bones known as the malleus, incus and stapes.These bones form chain of levers connected by the malleus to the ear drum and the stapes to cochlear and their job is to transform the relatively large vibrations at the ear drum into smaller vibrations delivered to the cochlear/inner ear.

Inner ear: apart from the cochlear, a coiled tubular structure that resembles a snails shell, the inner ear also houses the balance organs - the semicircular canals and the vestibule which are responsible for the sense of balance and spatial position experienced by humans. This of course is not essential to the auditory system but performs essential functions. The cochlear itself deserves special attention as it involves a complicated arrangement of anatomy that turns physical vibrations into neural pulses and it is here where the electrode of a cochlear implant (CI) is inserted.

Physically the cochlear divided into three fluid filled sections, as seen in the cross section of figure 1.4. As the upper section, Scala vestibuli is sealed at the oval window (which is connected to the stapes) it is here where vibrations first enter the cochlear and as vibrations travel through the cochlear they are able to traverse the two membranes (Reisnerr and Basilar membranes) that separate the three cochlear sections. The most important of these two membranes is the Basilar membrane.

Basilar Membrane Stretching the entire length of the cochlear the membrane changes in form, at the ape (the beginning) the membrane is narrow and stiff, whilst at its end wider and more flexible. When vibrations pass through the cochlear the basilar membrane is set into motion in the form of a wave - imagine shaking out a bed sheet or the waves along a still lake. Depending on the

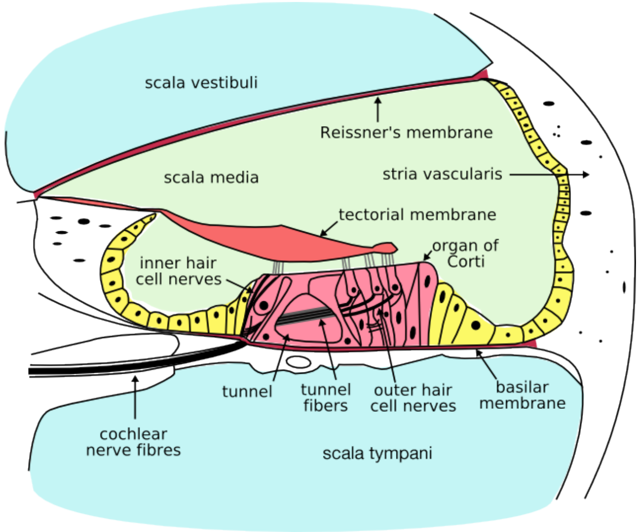


Figure 1.4: Cross section of the cochlear, showing the three fluid channels - *scala vestibuli*, *scala media* and *scala timpani*. The *organ of corti* (in pink) is shown sitting on top of the *basilar membrane*. Inner and outer hair cells are situated between the organ of corti and the *tectorial membrane*. Image taken from https://commons.wikimedia.org/wiki/File:Cochlea-crosssection.png

frequency of the sound/vibrations and due to the physical characteristics of the membrane the peak of this wave occurs at different positions. For example high frequency sounds peak at the beginning of the basilar membrane where it is stiff, and low frequency sounds peak at the end where the membrane is loose and more flexible. This arrangement of responses is described as being "tonotopic" and is illustrated in figure 1.5. The final piece of the auditory system process concerns the creation of electric pulses which happens in the . . . . . . organ of corti.

Organ of Corti: sitting on top of the basilar membrane (see figure 1.4) this is the sensory organ at the heart (excuse the pun) of the auditory system. Within the organ, and again stretching the length of the basilar membrane are tiny hair cells that respond to the motion of the membrane. For example when the motion is large enough, the tiny hairs on the tops of the cells are displaced causing the release of neural impulses that are carried by the auditory nerve to the central auditory system in the brain.

It is here that a very important theoretical point should be made that will have implications on the functionality of cochlear implants and discussion of results in later passages of this manuscript, and that is the theory concerning

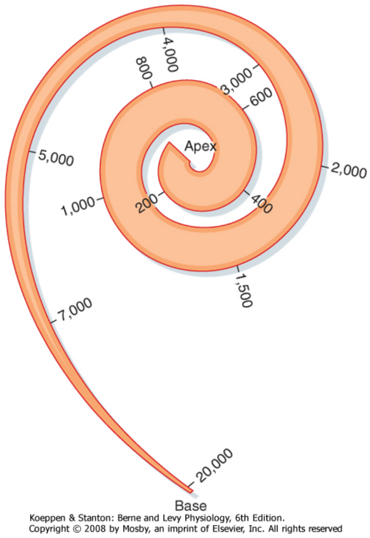


Figure 1.5: Tonotopic arrangement of Basilar membrane. Image shows the basilar memrbanem that is narrow and stiff at its base whilst being wide and flexible at the apex. This infers a tonotopic property to the structure such that high frequencies (limit approx 20, 000 HZ) cause maximum response at the base and lower frequenceis (limit approx 20 Hz) at the apex. Image Taken from Berne and Levy Physiology 6th edition.

how the basilar membrane peaks and subsequent neural impulses are able to code or describe the properties of a sound. Whereas it is relatively to simple to imagine the strength of the neural response coding the intensity of a sound and the time at which a neural response occurs coding the temporal aspects of a sound, when it comes to spectral information there are two different (although complementing) theories that are briefly explained below.

In the coding of spectral information (the frequency content of a sound) there are two different theories that describe how this information is transmitted from the cochlear to the brain. In describing these two theories it is easier to imagine the most simple scenario of using a pure tone (i.e. a single frequency)

PLACE theory: as mentioned before the basilar membrane is tonotopic. Using this idea place theory suggests that frequency information is extracted from the place at which peaks occur on the basilar membrane i.e. the location of the neurons that fire in response to a specific frequency. Complication with this theory is that regarding complex sounds the sensation of pitch (i.e. F_0) may be induced by frequency components.

TIME theory: which suggests the pitch of a sound is related to the "pattern" of neural impulses.These impulses often occur at a particular phase of the sound (known as phase locking) such that the intervals between impulses correspond to the frequency of the sound. The pattern of these intervals can then be used to code the frequency. One problem with this theory is that phase locking is only possible for frequencies up to 5 kHz, although in the real world most everyday sounds (including musical instruments and the human voice) have fundamental frequencies much lower than this, so there that the Time theory is still capable of accounting for many aspects of frequency coding / pitch perception.

Whilst it is not within the scope of this research to pursue these two theories in extravagant detail readers should be aware that recent research often concludes that the auditory system likely uses both of these processes in the coding of frequency information.

1.3 COCHLEAR IMPLANTS

The goal of a cochlear implant (CI) is therefore to transfer the acoustic vibrations of a sound into electric pulses that can be sent to the auditory nerve and subsequently the auditory cortex, enabling what may be called a "synthetic " or electric auditory perception to be realised. In a sense the implant replaces the auditory system shown in figure 1.3, as shown in figure 1.6.

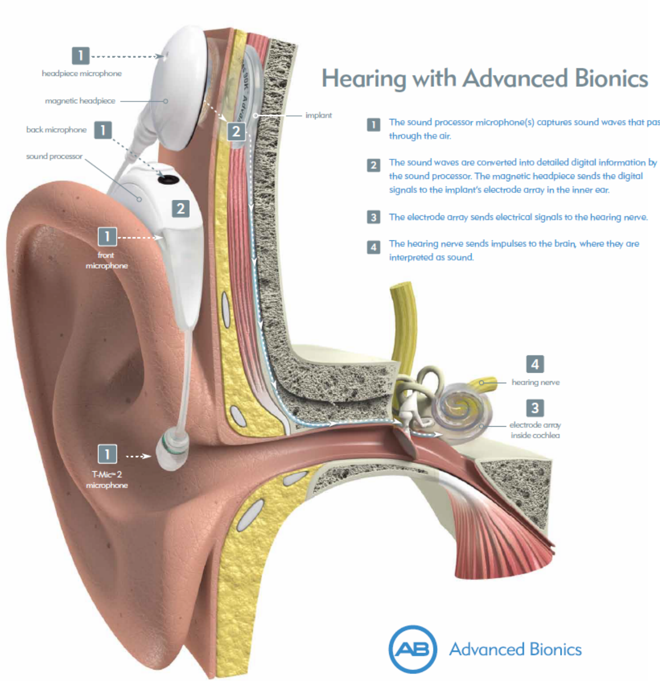


Figure 1.6: Labelled Diagram of a Cochlear Implant. Photographs/Images by courtesy of Advanced Bionics.

The CI itself can be broken down into several different components as indicated in figure 1.6. The function of each element is listed below:

1. Microphone: turns the analogue acoustic signal of a sound into a digital electrical signal fed to the sound processor.
2. Sound Processor: Here different coding strategies (discussed further below) are used to electric signal into a digital stimulus for activating the electrode array.
3. Transmitter / Receiver: The digital stimulus is then passed through from the transmitter to the receiver that is embedded below the scalp of the user.

Figure 1.7: Images of CI transmitter and electrode array for Advanced Bionic CI device. Photographs/Images by courtesy of Advanced Bionics.

At this point the digital signal consists of an arrangement of electrical pulses, shown in the left panel of figure 1.7.

4. Electrode Array: a thin plastic element inserted into the cochlear on which sit a number of electrodes. Electric pulses stimulate directly the auditory nerve based on the coding strategy used by the processor, shown in the right panel of figure 1.7.

1.3.1 *Speech processing with a Cochlear Implant*

The cochlear implant was itself designed specifically to restore speech processing abilities in both pre and post-lingually deaf humans. Speech has also been found to be considerably robust to spectral degradation and vocoding amongst NHL [125]. It is generally considered that most recipients are able to achieve the best improvements which often result in good speech perception in quiet conditions [108] however there are numerous difficulties when listening to competing sound sources, for example multiple talkers, or in complicated/noisy listening environments [35, 100, 129].

The biggest improvement amongst pre-lingually deaf CI users comes from implanting during infancy. It is generally held that earlier implantation leads to the best restoration of speech abilities and that the upper cut-off for this improvement lies between implantation at 4 and 12 years of age [106]. Implantation before 12 months old has also been recommended, for example [29] showed that infants implanted between 12-36 months were delayed by one year in their development of speech in comparison to infants implanted before 12 months old and [56] found that early implantation lead to faster development.

However development is also linked/affected by a range of socio-economic values such that whilst the age at implantation is important it is not the only factor in the restoration of speech abilities [104].

Improvement of auditory abilities are also seen in adult and elderly populations (Dillon et al 2013) and amongst the pre-lingually deaf population most recipients show a benefit from receiving a CI [142]. However even after many years of being implanted these improvements in speech understanding are not as large as compared to those seen in post-lingually deafened populations [145].

Post-lingually deaf adults also benefit greatly from receiving a CI however they require a certain amount of rehabilitation in order to associate the newly perception of sounds to the original representations stored in memory [17]. Within this population the success of CI devices has also been linked to the duration of deafness as a result of the age at onset of deafness. For example Moon et al (2014) found that subjects who became deaf during adolescence had poorer speech performance. They also concluded that while the age at implantation can affect the success of rehabilitation post-implantation, as a result of cognitive abilities, it is not as strong a factor in predicting the success of the implant.

Overall the previous literature shows that CI can be a very successful tool in restoring speech abilities.For pre-lingually deafened CI users the biggest improvements are seen when implantation occurs at the earliest possible stage with benefits drastically reduced if implantation occurs after 12 years old. For post-lingually deafened CI users the age of deafness onset (and subsequent duration of deafness) is seen as a key factor in determining the success of implantation more so than the age of implantation.

Whilst CI may give significant benefits in speech perception there are certain aspects of voice perception that remain difficult for nearly all CI users. Firstly the perception of prosodic information is reduced in CI users, for example intonation [92] and also emotion [89, 99]. [89] showed that CI users correctly recognized emotionally content in half the stimulus presentation whilst NHL performed at 100%. However when reducing the amount of spectral information via vocoding the performance of both NHL and CI users was seen to reduce. The perception of emotion has been linked to the changes in the fundamental frequency of a speaker and also intensity cues [58].

Secondly CI suffer from deficits in recognising the identity and gender of talkers [94, 65, 27, 81]. Gender related information is linked to the anatomy of the human vocal anatomy, for example the size of the vocal tract and vocal folds which on average are larger for males compared to females.This in turn generates differences in the fundamental frequency (F_0) and arrangement of formants in the voice which are used by NHL to accurately discriminate gender [128]. However these cues are often unavailable to CI users whom therefore struggle with gender discrimination, however when the difference in F_0 is large CI users perform significantly better [81, 93].

Considering all of the above it must be noted that there is a considerable amount of variation in CI performance within all the studies undertaken. This variability is a result of many different mitigating factors concerning; the patients existing auditory abilities - for example residual hearing, duration of deafness, age at implant ion, the use of sign language and lip-reading abilities; the CI itself - including insertion depth, the number of active electrodes (often only 4-7 are used by the user); and finally socio-economic status and home life of CI users [104].

1.3.2 *Musical processing*

For many CI users an expectation of receiving an implant is improved perception of music (cited as second most important acoustic stimulus to their lives[41]) , however many CI users actually avoid music due to it being too noisy, and post implantation listening of music is reduced compared to pre-implantation with CI users often stating their enjoyment of music has been reduced (Gfeller, LEa et al 2003, Feldmann and Kumpf , Leek et al.). Mizra et al (2003) reported similar findings and that the enjoyment of music rated from 1 to 10 (where 10 is like very much) also reduced from 8.7 to 2.6, therefore enjoyment is reduced (Tyler et al 2000). It is often cited that the main reason for difficulties with musical sounds and the reduction in enjoyment is linked to problems with the perception of pitch and recognition of timbre, attributes associated to the processing of the spectral fine structure. On the other side of the coin CI users are reported to have temporal processing abilities similar to those of NHL and therefore have little or no problems in the perception of rhythm (REFERENCES).

Pitch

It is generally held that CI users perform much worse than NHL in pitch related tasks. For example in pitch-ranking i.e. placing in order a series of differently pitched tones, CI users have been shown to be unable to rank pitches that are 3 semitones apart and are only able to perform the task at roughly 65% accuracy for pitch intervals of 6 and 12 semitones (Looi 2008), although other data reports pitch ranking thresholds in the wide range of 4-24 semitones [36] or 1 to 12 semitones [41, 102, 60]. Not only are the pitch discrimination abilities of CI variable, but they are worse than NHL whom often performed at 81% in discrimination intervals of only 1 semitone [132]. Pitch perception is not simply the ability to distinguish different pitches and so has been tested in other ways. For example [30] tested the ability of CI users to discriminate between two short melodies, where in one melody the F_0 of a note was altered for half its occurrences.This represents a level of pitch perception within the context of a melody and is therefore more similar to real-world situations of musical pitch perception. The results again showed that CI users performance was around chance.

The CI works by extracting the envelope of a sounds and delivering this signal via electrical stimulation/pulses into the cochlear subsequently the auditory nerve. In doing so most of the fine temporal information is lost such that CI can only perceive repetition rates up to 300 HZ [32] which is much lower than the 2-5kHz limit imposed on NH by the natural auditory system. So without the ability to access the fine temporal information it leaves the perception of pitch in complex tones/sounds difficult.Further complications/hinderences are added when considering the variability in the anatomy of CI users, the survival of auditory nerve, cause of deafness/hearing impairment. Although CI's aim to deliver a large number of frequency information ranging from 16-22 channels, the placement of electrodes and large stimulation given by the electrodes often leads to a greatly reduced number of channels being functional, sometimes as low as 3. This of course has implications to CI users ability to use place coding model of pitch perception, as the spatial resolution of stimulation is much poorer in comparison to a NHL.

Timbre

Timbre is often a difficult quality of musical perception to describe and to quantify. Most simply it can be thought of as the qualities of the sound produced by different musical instrument. More complete it is the combination of the acoustic qualities of a musical or vocal sound that are separate/independent from the frequency and intensity (pitch and loudness) and yet which enable the sound to be perceived/recognised. In measuring the accuracy of timbre perception studies often test musical instrument identification, finding that NHL score close to 100% where as CI users perform at best 54% or less [39, 95, 102]. [86] compared the ability of 15 CI and 15 HA participants in their ability to identify single musical instruments, solo instruments with background accompaniment, and musical ensembles.The mean percent-correct scores for the three respective subtest were: CI group - 61%, 45%, and 43%; HA group - 69%, 52%, and 47%.

Melody
From the findings that pitch and timbre perception are poor in CI it is not surprising that the perception of melodies is less than for NHL, ranging from 17% to 60% for familiar songs in comparison to [41, 64]. This can be improved with the addition of lyrics and rhythmic passages.For example in [39] identification of rhythmic melodies was 20% compared to only 10% for non-rhythmic melodies with identification 90% and 77% for NHL respectively. When sequences are formed of complex tones that contain resolvable harmonics, a pitch sequence advantage is observed, If there are no resolved harmonics, because of bandpass filtering or noise-vocoding , the pitch sequence advantage disappears [31].

Rhythm
In contrast to the processing of spectral aspects of musical perception (pitch and timbre), rhythmic cues are commonly well perceived by CI users whose performance is similar to NHL [32, 88]. [64] showed that CI users also report similar performance in a tempo discrimination task, although when tasked with identify complex rhythms CI users performed 20% lower than NHL. [40] also showed that NHL could detect changes in rhythm at shorter duration of beat (607ms) compared to CI users (1070ms). Therefore in certain "difficult" situations the temporal processing abilities of the implant are not capable of competing with NHL. CI users have also shown to perform similarly to NHL in perceiving temporal gaps (silence) and amplitude modulations. Therefore it appears that the perception of non-complex rhythms does not pose significant issues for CI users . Rhythmic information poses less problems for listeners, mainly because temporal cues are more easily coded by the implant. This could therefore lead to a preference for listening to the rhythm (temporal information)

within music making rhythm more enjoyable for CI users and therefore widen the gap between ease of rhythmic perception and difficulties with pitch and timbre perception.

It would appear that the perception of music is not a trivial task for CI users, with common difficulties in the access to fine spectral information meaning that pitch and timbre are difficult to perceive. This is especially true for complex sounds however there is a range of performance of CI users with some showing significant musical abilities (see cochlear implanted kids playing the piano). Studies have also shown that gross-temporal processing abilities similar to those of NHL especially in the detection of gaps and amplitude modulation leading to often good levels of rhythm perception.

1.3.3 *Environmental sounds*

Studies with CIL

Another reported benefit of a cochlear implant is in the perception of environmental sounds (ES). Often classed as any non-linguistic or non-musical sound, certain es's are amongst the first sounds that CI users report as hearing [137]. In a questionnaire of 22 CI users [146] also reported 77% CI users found the main benefit of receiving their implant to be the perception/awareness of ES.[138]reported similar findings in a questionnaire of 53 experienced CI users. These sounds are important and of benefit for a number of reasons. Firstly certain ES are important for warning or signalling purposes e.g. alarm sounds, telephones, doorbells. Secondly they can just be aesthetically pleasing, e.g. birdsong and waterfalls. Finally and possibly most importantly they contribute the awareness of a listener to give an improved sensation of the surrounding environment as well as aiding the interaction with objects. For example background noises (e.g. air conditioning) are often the first sounds reported by CI users after switching on their implants. environmental sound perception (ESP) is therefore an important part of the quality of life for CI users [146].

However due to the focus of research on understanding and improving speech perception for CI users ESP has only recently, and with limited scope, started to gain attention. The following text outlines these studies in terms of the CI users that were tested, their performance and the testing method used.

In a general study on outcomes of receiving an implant [108] tested 40 patients on the recognition of 40 ES presented at 70 dB(A). The average score of participants at both 9 and 18 months post implantation was 57%. [136] also tested 21 post lingually deaf CI patients at 1,9 1nd 18 months post-implantation. Testing was done using two lists of 18 ES and result de in average scores of 30% correct after 1 month, 35% at 9 months and 37% at 18 months.

[111] tested 11 CI patients (mix of post & pre lingually deaf) who had at least one year of experience with their implant. The test consisted of identifying 40 sounds divide equally amongst 4 different auditory contexts/settings (home, office, kitchen and outside). Testing used a closed-set 10-alternate choice paradigm such that for each context/setting participants were presented with each sound and had to chose from the 10 alternative possibilities, without feedback. Identification was found to be on average 79% (ranging from 87% for office to 78% for kitchen) but with individual participant performance ranging from 45% to 94%

.

[87] looked at ESP with two implant groups, firstly four subjects with 3 months experience (whom they also tested pre-implantation) and secondly 10 experienced CI users with at least 10 months implant experience. They tested 45 sounds (2 tokens for each sound = 90 total) that were subdivided into 9 categories Arriving Home, Bathroom, Household Appliances, Human, Kitchen, Nature, Office, Traffic, Other (included examples of music). Participants were played the sound and asked to identify it from a list of the 45 sounds. Experienced group performed at 59%. Pre-implant with hearing-aid was 40%, post implantation (3 months) 57%.

[59] tested non-linguistic sound perception with 22 post lingually deaf CI users with more than one year of experience. They were tasked with identifying 50 sounds and then categorise it in one of 5 a priori categories (nature, animal/insect, mechanical/alerting, human non-linguistic, musical) there were 10 sounds from each category presented. Participants verbal responses were recorded and later interpreted by testers giving average Identification accuracy was 49% with categorisation accuracy higher at 71% correct.

[122] tested 17 CI users with average experience of 3.2 years whom were mixed post and pre lingually deaf. Participants were presented with 160 envi-

ronmental sounds (40 different sounds of 4 tokens/examples), one at a time and asked to identify it from a list of 60 possible responses listed on a computer screen from which the participant could choose. Results showed identification performance of 45% correct ranging from 16-69%. Using the same test paradigm Shafiro (2015) tested the effects of training on another group of 14 post-industrial deaf CI users with average experience of 5 years implantation. Pretest results were 47% and post-training results were 62%.

The first obvious conclusion from these studies is that the performance of CI users in identifying ES is low, especially in comparison to NH studies where ESP is nearly always close to 100 % . This remains true even for experienced CI users who have been implanted for a long time [119] and for whom it would be expected to improve on ESP in similarity to improvements seen for speech perception with time. There is however evidence that training effects can improve ESP for CI users. [124] reported an improvement of 15.8 % points following 4 training sessions performed by participants at their homes over the period of one week. This is the only study to concretely test the effect of training and showed a positive effect of 46 % points for trained sounds and 50 % points for un-trained examples of the same sounds. However improvements to non-trained stimuli were small and non-significant (only 6% points). So although training effects generalised to non-trained examples of the trained sounds, training effects did not generalise to non-trained sounds.

Whilst identification performance was always seen to be low [111] did show that when presented with four pre-defined categories based on context (auditory environment) categorisation accuracy of the sounds was much higher. Indicating that CI users are able perceive certain information form the sounds and link this to the context in a fairly successful manner even if specific identification of the sound is not always possible.

Within these studies it is also possible to see the wide range of identification score, one of the issues with testing ESP. Highest identification was seen for sounds described as being temporally or spectrally distinct (examples) whilst those with lowest identification are sounds with a uniform envelope. This provides some initial clues as to how future studies could pursue ESP in CI users and how rehabilitation schemes could be best designed. However no study has yet identified specific acoustic variables that account for the identification of ES, most likely again because of the huge variance in the types of sound and the acoustic characteristics that are present when considering ES.

The perception of ES is said to be similar to speech in that both types of sound rely on a common ability of sound recognition [61], and also involve processing of both acoustic and lexical information [69]. [59] recorded a correlation of ES identification and speech performance, whilst [111] also concluded that identification ability was loosely related to monosyllabic word recognition (NU-6 test). However [124] reported only non-significant improvements of speech performance (CNC and SPIN-R tests) during the ES training study. Overall whilst a link between speech and ES performance may make sense these effects are minimal and not conclusive. However the hypothesis has been tested to a greater degree in CI simulation studies and is discussed in more detail below.

There are some issues with the above studies. Firstly and of great significance is the fact that in testing ESP there are almost an infinite number of sounds that fall into this description and it cannot be possible to test them all. [87] also include examples of linguistic sounds, soundscapes and full musical extracts in their study, which do not fall into the definition of ES used by many other studies. Some also only test a small number of sounds which makes it difficult to expand conclusions to all ES. Secondly in testing identification performance there is a variety of methods employed. In testing that used an open-set paradigm, whereby subjects are asked to identify a sound without any options, it is difficult to know what schema was employed by the experimenters for interpreting the descriptions given by subjects. There could be a case that the schema used by one study would give different results when compared to that of another study, and therefore be over or under-estimating the performance of CI users. Other studies use a closed-set paradigm, where a fixed number of options are presented to subjects which mean there is not need to interpret comments, however in doing so this presents subjects with options to choose from and may not allow the assessment of the subjects true perception. Finally the study of [111] provided subjects with categorical information about the sounds which could have helped subjects to identify the sounds.

It is therefore possible to conclude that from the the existing studies that have been conducted on ESP, many have used different paradigms, combined with the variety and amount of ES makes it difficult to make detailed conclusions.

However it does appear that ESP is poor even after training and with experienced CI users. Also although no specific acoustic variables have been identified that can account for ES identification (again influenced by variability of ES) it appears that more easily identified sounds are spectrally or temporally distinct, with sounds that have a uniform envelope structure being more difficult to identify.

CI Simulation studies

Another way of testing for the perception of sounds and also with regards to training effects is to test the more readily available normal-hearing subjects by using vocoder techniques to spectrally degrade sounds and simulate the hearing of a CI. IN similarity to studies concerning speech perception there is still much less known about the perception of environmental sounds in such tests. These studies have been used to assess training effects and also to try and find specific acoustic variables that may be able to explain the perception (both identification and categorised) of ES.

[84] tested training effects and perceptual learning with normal hearing subjects using spectrally degraded sounds that had been passed through an eight-channel sine wave vocoder. 5 groups of NH subjects (155 in total) were trained on different sounds, words (simple & complex), sentences (meaningful & anomalous/nonsense) or environmental sounds and then tested in an open-set identification task. 99 ES were tested covering a range of sounds: vehicles, animals, insects, NLHS, musical instruments, tools, liquids, others. Pre-post gains for all stimuli conditions but generalisation was less uniform, training generalised toward "easier" tasks. For ES sounds generalised to ALL speech materials.

[51] tested the identification of 70 ES under different spectral degradations. Identification was quite robust under High/Low pass filtering (cut-off 300-8000 Hz) especially for high pass where ID never dropped below 70%. Also used octave band filters (fc 212 - 6788 Hz) and noted that performance with higher band cf's was still good 70-80%. Using EMS (vocoder technique) the more channels added the better the ID. Also linked performance in EMN conditions of 1 & 6 channels to the number autocorrelation peaks, burst ratio and cross-correlation. They study also found that the most informative frequency region

for categorising the sounds was between 1200-2400 Hz.

[119] tested the ESP of 60 sounds with 60 NHL under six different spectral-resolution conditions - 2, 4, 6, 8, 16, 24 and 32 channel envelope vocoding. ID improved with increasing number of channels i.e. increasing spectral resolution, but plateaued from 8 channels up, although was still not as good as original sound condition even at 32 channels. Some sounds train whistle and water draining also never reached above 70%. Shafiro 2008 also investigated the perception of ES by NH (seven) in CI simulated conditions (4 channel vocoding) in a pre-post test training paradigm with 5 training sessions. Identification of non-vocoded sounds was 98% correct. In vocoding condition pre-test 33% rising to 63%. The biggest rise was for difficult-trained sounds (61% points)and within this group there was a difference seen for the trained sound tokens (86%) whilst non-trained exemplars rose 36% - showing the effects of specific stimulus training and/or familiarity with sounds.

These studies show that under spectral degradation the identification of ES is reduced for NH subjects and that with an increase in spectral information (the number of frequency channels) identification is seen to increase. Again the variability and range of ESP is demonstrated in similarity to studies of CI users. [51] shows that some sounds were robust to spectral degradation and only required one channel of frequency information (helicopter, galloping horse) whilst others were poorly identified whatever the number of channels / spectral resolution used (flute, electric saw). [119] also reported that sounds could be divided into those which could be correctly identified (at 70%) with 8 channels or less, and those with 16 channels or more. Sounds also be divided into one of these two groups based only on two acoustic variables, the number of bursts in the envelope and the std of the spectral centroid velocity. This suggests that clear temporal events within a sound and the speed of spectral dynamics are factors in deciding the amount of spectral resolution that is required for correct identification for vocoded stimuli and therefore are also possibly important for CI users as well. Finally Shafiro 2004 (doctoral thesis) showed increasing spectral channels only helped ES which were more strongly spectral and those more temporal sounds actually decreased in identification. Along with increased spectral resolution being linked to increased identification performance the frequency region of 1200-24000 Hz was also identified as being important for

ESP [51], and this is also an important region for speech perception.

It is likely that the perception/identification of ES and speech sounds shares some common processing, for example both involve low level processing of acoustic information as well as higher order processing linking the attachment of semantic information to sounds [85, 69]. [61] also identified a specific "familiar sound recognition" ability important for the perception of both speech and ES, especially in the case of degraded sound, whilst there is also evidence that the processing of both sound-types shares cortical areas [78]. [22] reported that training on anomalous speech sentences generalised to both meaningful sentences and ES.However a more recent study by [84] did not find this, and instead found only that training on ES generalised to speech stimuli. They concluded that by training on stimuli with a syntactic structure but no semantic structure reduces the impact of top-down processes and forces the listener to focus on the acoustics of the sound and use bottom-up processes.Training on ES (non-speech) stimuli may therefore improve a listeners access to the spectro-temporal characteristics of stimuli and help when identifying vocoded speech. In similarity to studies on CI users training effects for ES were also reported by [84] and [120] whilst [51] found listeners who had heard sounds once more than naive listeners had improved identification score, with the largest improvements seen for temporally distinct sounds (footsteps, axe chop, scissors).

Overall the existing literature concerning the perception of environmental sounds for CI users (also including CI simulation studies) is informative for a number of things. Training with ES greatly improves the identification ES with large improvements seen for trained sounds and other examples of such sounds. There is a possible link between the perception of ES and speech performance of CI users and that training with ES may also generalise to speech stimuli, however this is not conclusive. Although the identification of ES is reduced by the use of a CI or CI simulations there is a variability amongst ES such that some sounds remain well identified under all forms of spectral degradation whilst others never reach high levels of correct identification. This has regularly been linked to the complexity of the spectral-temporal information of a sound such where spectrally or temporally distinct sounds are often correctly identified where as those with more generic or uniform envelope shapes are poorly identified likely as a result of the confusion between sounds that share these qualities.

However, more detailed conclusions are hard to make due to the difficulty in testing ES. As mentioned there are a large number of different ES with widely ranging acoustic information and the studies, whilst using similar sounds, all use stimuli-sets and test-paradigms that are different. In addition to this none of the studies take into account real-world effects of ESP that may be important for listeners in identifying ES.For example the location of where different sounds are heard can affect its identifiability [52]. Also important is the familiarity of listeners to the sound stimuli ([6]) where one listener may be more familiar with a test stimulus than another listener and therefore have a better chance at identifying it. Finally the manner in which sounds are listened to could be very important. Most ES are part of a sound scape, occurring in the background and not often the focus of listeners attention, which is of course different to the way in which speech is commonly listened to. Finally the use of CI simulations is informative and useful for the progression of research, however whilst this can give clues as to the functioning of ESP in human listeners there is still a large difference on the way NHL and CI users perceiver sounds, both from a low level perceptual manner to a high order cognitive level. For example most CI users are often elderly where as NHL used for studies are often much younger and not often age-matched. [61] identified a number of things important for sound recognition, problem solving strategies to constrain/limit possible response options, the ability to focus attention on spectral and temporal information int eh signal, how efficiently a subject can locate stored information (semantic) about an auditory stimulus. These cognitive processes may be very different between listeners of differing age and add another element into the study of ESP.

1.4 CLUSTERING ANALYSIS

Within this manuscript the main topic of analysis focusses on how different sounds are categorised when they are placed into different groups or clusters. Presented here is a brief introduction to how the data generated by such tasks has been analysed. It should enable the reader to better understand the specific analyses that have been performed throughout the manuscript and which are described in further detail in chapter 2.

Sorting and categorisation tasks have mainly been used in order to study sensory analysis as an efficient way for investigating the similarities between certain kinds of stimuli, for example how a set of perfumes or wines may be grouped based on their odour and taste [23, 1, 109]. Whilst a range of analysis techniques can be employed within the study of categorisation, the present research has used a dual-analysis combining MultiCorrespondence Analysis (MCA) and Heirarchical Clustering based on Principal Components (HCPC) based on the *FAST* model described in [23]. The goals of using these analyses are threefold and described in more detail below:

1. To understand which stimuli (sounds) are being grouped together in order to see what kinds of categories are being created. (HCPC gives us this most easily/simply).
2. To understand HOW the sounds are being grouped together. What strategies (factors) are being used by participants (MCA)
3. The comparison of different groups of participants.

1.4.1 *MultiCorrespondence Analysis (MCA)*

From the results of a sorting/categorisation task MCA is used to uncover the underlying relationships between a data set. This data set is taken either as an indicator/disjunctive matrix (as shown in figure 1.8). For the purposes of the following research the rows (i) represent the sounds to be categorised and the columns (j) the participants. Within each participant column there is a subset array of k element, where K is the total number of categories created by a single participant. This K array is made up of 1's and 0's such that when a sound i

falls into category k it is marked with a 1, therefore indicating which sounds were grouped into which category for each participant.

The total number of stimuli/sounds is given by I, and for a single participant the total number of sounds in a particular category is given by I_k.

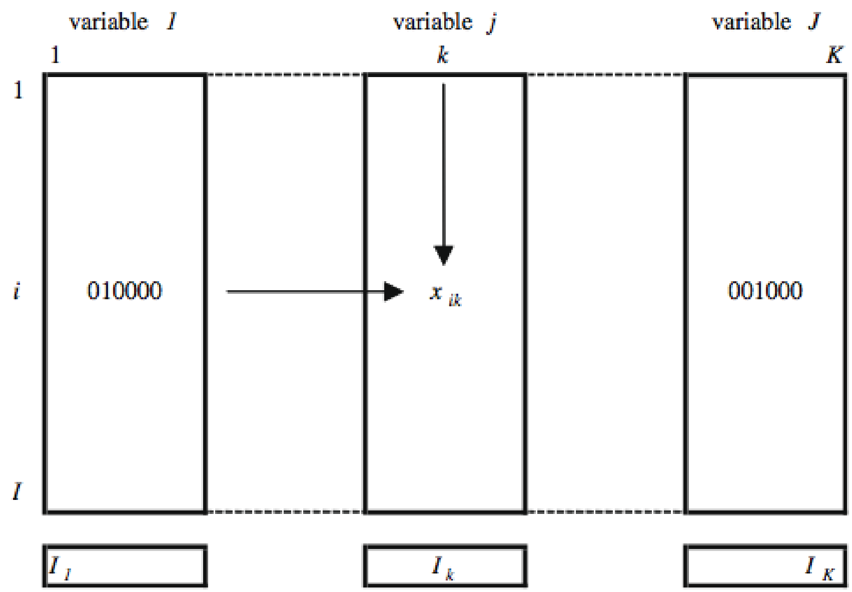


Figure 1.8: Graphic display of an Indicator Matrix. Image taken from [23].

MCA itself uses Correspondence Analysis (CA) in order to analyse the relationships between a number of categorically dependent variables. i.e. the fruitiness of wine, or the saltiness of cheese. Practically speaking the process performs CA on an indicator matrix as shown in figure 1.8. The principal is that MCA seeks to represent these relationships as points within an N-dimensional Euclidean space. For example we can imagine a $3-D$ space with points positioned in groups within this space. Concerning the relationship of stimuli, two stimuli may be superimposed upon each other if all participants placed them in the same category, and the distance between them increases as the number of categorisations decreases.

The dimensions that are created/outputted by the analysis are then ordered in terms of the amount of variance within the data (disjunctive matrix) for which they can account. Whereby the first dimension would accounts for the most variance and reflects a categorisation strategy that is used by the majority of participants.

It is important to note when undertaking a sorting task involving fixed stimuli attributes - for example rating wines on their fruitiness, alcoholic percentage, colour etc - participants are aware of criteria for categorising and the subsequent dimensions reflect either a singular attribute or a mix of. However when conducting a Free Sorting Task (FST) as has been undertaken throughout the present research, the outputted dimensions cannot initially be directly related to any specific aspect of the stimuli, be it physical or cognitive. It has therefore been a large part of the current work to interpret these dimensions using additional information, for example descriptions of the sounds and the categories. This is more fully discussed in the following chapters.

The *FAST* approach (see [23] for details) it is also possible to view an arrangement of the participants (categorical variables). Referred to in the current work as *participant maps* these figures display the distribution of participants across each dimension on a scale of $0 - 1$ and can be used to interpret how strongly each dimension is used by each participant. An example taken from [23] of MCA output is shown in figure 1.9 in order to demonstrate the distribution of stimuli on the *factor map* (panel A) and the participants on the *participant map* (panel B).

1.4.2 *Heirarchical Clustering based on Principal Components (HCPC)*

Following on from theMCA a complementary analysis of Heirarchical Clustering based on Principal Components (HCPC) has also been used throughout the current study. HCPC enables a simplified view of the overall categories as a *dendrogram* or *tree* and an example is given in figure 1.10. The height on the dendrogram displays a measure of similarity between the stimuli, with a greater height representing a greater dissimilarity. The final categories/clusters are given by the HCPC are determined by cutting the dendrogram at a specific height, where this height is calculated based on the change in inertia (variance).

Secondly a Hierarchical Clustering based on Principal Components (HCPC) was performed on the results of the MCA analysis in order to view a simplified version of the categories of sounds in the form of dendrograms. When using this analysis it is not possible to account for all of the variance (inertia) within

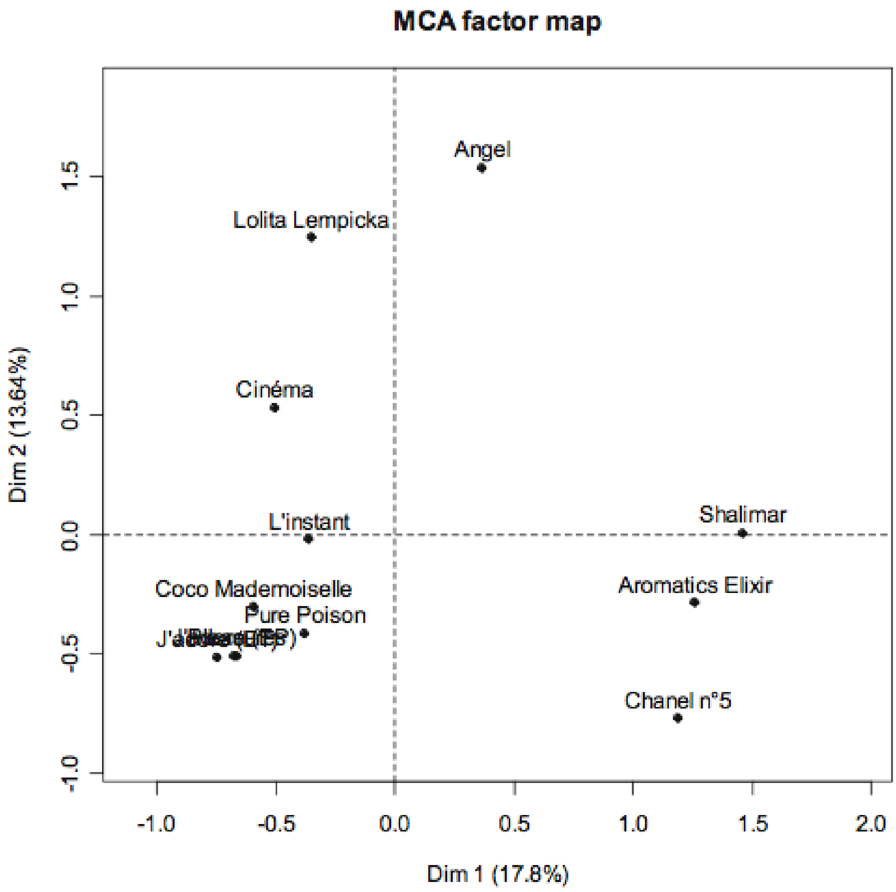


Figure 1.9: Example of the factor map output generated by MCA analysis. The arrangement of a series of perfumes are shown displayed in a 2D euclidean space. Image taken from [23].

the data, i.e. the variability of participant responses, and so a certain amount remains unaccounted for.

By increasing the number of desired categories the inertia can however be reduced and it is using this process that the we can choose a final number of categories: if the number of categories is Q then the optimal number of categories is found when the change in inertia is greater when moving from Q-1 to Q than from Q to Q+1 (François et al., 2010). This can also be defined as the value for Q which minimizes equation 1.1.

$$\frac{Q - Q_{Q+1}}{Q_{-1} - Q} \tag{1.1}$$

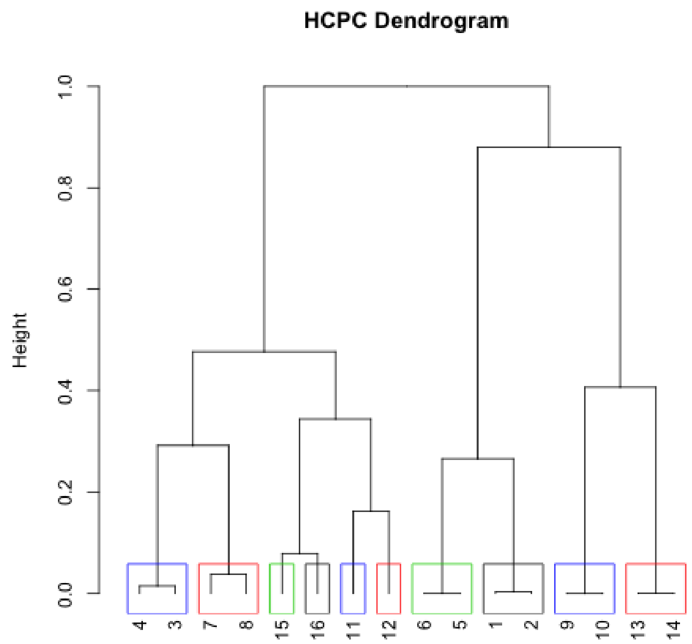


Figure 1.10: Example of a dendrogram / tree as generated by HCPC analysis. Categorised stimuli are represented along the bottom of the tree, with rectangles representing overall categories that were outputted by analysis. The height ordinate represents the degree of similarity between stimuli, where a large height is equivalent to a greater DIS-similarity.

1.4.3 *Comparing Clusterings*

Categorical analysis was concluded by finding the Cophenetic Correlation Coefficient (CCC) [118] which gives a measure of how accurately the distances between items in the raw data are preserved in the dendrogram [25]. CCC has been computed in R' using the functions dist - to calculate the euclidean distance between sounds within the n-dimensional space as computed using MCA; cophenetic - to find the cophenetic distances between the stimuli computed using HCPC; cor - to correlate the two distance matrices and give a final coefficient.

Participants' category descriptions were also evaluated as to whether or not the categorization was based on semantic information concerning the sound source or on qualitative information linked with the acoustic signal. For each participant group the number of descriptions that made reference to either semantic or qualitative acoustic aspects of the sound were totalled across all sounds and all participants and then turned into an overall percentage. Examples of descriptions relating to semantic information were "domestic noise"; "voice, laughter, someone talking"; "musical instruments". Whilst examples of descriptions referring to qualitative acoustic aspects included "melodies"; "noise"; "ringing tones"; "disagreeable sounds". In order to further understand the performance of participants in how they categorized the sounds an estimation of "Categorization Identification" was inferred from the participants' category descriptions. For each sound the associated category descriptions were evaluated as to whether they referenced musical, vocal or environmental sounds. These were then compared to the predefined categories and the percentage of comments in agreement was calculated across all participants for each sound. In some cases participants did not give any specific description and this is reflected in Table 3 by the "% No comment". In addition a value for "Categorisation Accuracy" was found based on the stimuli that were paired together by participants. This was calculated as the percentage of stimuli-pairs present in each participant's category choices compared to the stimuli-pairs contained in the predefined categories, of which there are forty possible pairs, (Table 2).

1.5 AIMS OF THE THESIS

It should be apparent that the theory of auditory categorisation is not firmly established, for example neither a hierarchy of categories or specific categories have been et-out. However agreement is found in that categories are commonly based on the identification of the sound producing object and action, whilst the acoustic information and such similarities become more important when identification fails. This has been described as musical and everyday listening by Gaver, along with his classification of sounds based on material water, gas or liquid (see figure 1.1). Whilst this forms an intuitive base for the classification sounds, these theories probably should be updated somewhat in the face of more recent categorisation studies and as this field of research continues to advance. For example many studies have found similar things, such as categorisation of vocalisations (both animal and human), alarm/warning sounds, transport sounds and action sounds involving different objects.

Categorical perception of sounds also provides an interesting method to test the perception of listeners, whether in how acoustic or semantic properties of the sound (source) are used to categorise stimuli. Rather than identification it offers a broader level of perception, that still shares similar processing mechanism and provides another way of studying how human listeners perceive the auditory world. This may also prove useful when assessing the auditory performance of listeners with hearing impairments, specifically in this case Cochlear Implanted Listener. As stated, some research has been conducted on this topic and has delivered interesting and positive results, however there is still much room to pursue the topic, most notably as a FST has never been tested with a population of CIL, which offers a somewhat more "real-world" opportunity for testing.

In the following chapter we look at how 16 common sounds are categorised amongst NHL. Chapter 1 looks at how the sounds were categorised by NHL and how vocoding the sounds (in order to simulate CI processing) affected the perception and categorisation. Chapter 2 looks spe CIically at CI users, how they categorised the sounds and how the duration of implantation, and therefore the level of experience with a CI, affects the perception and categorisation of the sounds. Finally Chapter 2 conduces by discussing the results of Chapters 1 and 2 to compare the results and give an overall picture. Free-categorisation task of 16 common sounds.The following section describes testing that was undertaken on sixteen common sounds using an Free Sorting Task (FST). The sections below describe the testing that was done involving NHL with natural and vocoded sounds. Testing using sixteen common sounds that were chosen to cover vocal, musical and environmental sounds. NH and vocoded sounds (both 8C and 16C) in order to simulate CI processing. Next is presented the results CI users with duration of implantation greater than 12 months who are hereafter referred to as "experienced" CI users.

Henceforth the following manuscript aims to add further understanding to the topic of auditory categorisation by using a FST and focussing on specific types of environmental and vocal processing to uncover possible sub categories and the kinds of perceptions that are used when categorising these stimuli. The main body of work concerns how CIL perform in a test of free-categorisation concerning environmental, musical and vocal sounds and lays the first step in using this kind of test and/or categorical perception for future research.

2

AUDITORY CATEGORISATION WITH NORMAL HEARING LISTENERS - WITH NATURAL AND CI SIMULATED SOUNDS

The following chapter looks at how 16 everyday sounds were categorised by normal hearing listeners (NHL) in a FST across three different conditions of spectral degradation. Vocoder techniques were applied to the sounds using both 16 and 8 frequency channels in order to simulate CI processing and were compared to a control condition where sounds were presented naturally (NS). As mentioned the 16 sounds were chosen to constitute three overall categories of musical, speech and ES. Results show that in all three stimuli conditions that the participants used similar categorisation strategies which on some level corresponded to the the pre-defined choices and included robust categories of human sounds and machine/motor sounds. Results of 16 and 8 channel vocoded sounds were more similar than compared to the unmodified condition and the increase in vocoding appeared to affect the strength of categorisation of linguistic sounds i.e. speech vs. non-speech sounds.

Chapter Aims

1. Establish how a set of 16 everyday sounds are categorised by NHL in the Natural sounds (NS) condition. The chapter will investigate whether participants categorise the sounds according to the predefined categories of environmental, musical and vocal sounds or by other means. Analysis shall also seek to understand the perceptual similarities that are used to form categorisation strategies. The results and analysis will then form the basis for comparisons with other stimuli conditions in this chapter and then be used again in chapter 4.

2. Acoustic analysis of sounds shall also be used in order to help uncover whether any physical attributes of the sounds are important to the categorisation strategies used by NHL or whether results agree strongly with previous theory that auditory categories are based on semantic information associated to the sound producing event, object or action.

3. To investigate how the categorisation performance and relevant strategies change when reducing the spectral information of sounds. This is done using 16 and 8 channel vocoder techniques in order to simulate CI users auditory abilities. In this way not only is it the goal to investigate auditory categorisation in difficult listening conditions, when identification of sounds is likely reduced, but also provides predictive and comparative results for auditory categorisation performed by CI users (see chapter **??**).

2.1 GENERAL METHOD

The following section describes the experimental method and procedure that was common in all testing and is the same used for chapter 3.

2.1.1 *Stimuli & Procedure*

Sounds were taken from a database owned by the PETRA group at the University of Toulouse Le Mirail (http://petra.univ-tlse2.fr) and were chosen in order to cover as best as possible the broad range of acoustic information that is possible amongst the three predefined categories of sounds - environmental, vocal (both linguistic and non-linguistic extracts featuring male and female voices) and musical (single tones and simple melodies). Sounds were 2-3 seconds in duration and are listed with the predefined categories in Table 2.1 alongside the abbreviated labels used later in the manuscript.. All stimuli were monophonic and recorded in .wav format with a sampling frequency of 44,100 Hz. A limit of sixteen stimuli was used to insure that the test would not be too difficult for CI users to complete. The two participant groups were tested in quiet listening rooms, NHL were tested at the CerCo laboratory and CI users at the Purpan Hospital. Both groups were seated in front of a PC monitor positioned at eye-level with two Roline Digital loudspeakers located on each side at a distance of 1m. Stimuli were presented in stereo at a level of 65 dB SPL (measured with a Sound Level Meter at a distance of 1m) via the loudspeakers in free-field listening conditions. Testing itself was carried out using the open-source TCL-LabX software (http://petra.univ-tlse2.fr/tcllabx/) which acted as the interface for the FST. The sixteen sounds were represented on the computer by sixteen numbered and coloured buttons which were positioned in the same order for all participants. An example of the screen at the beginning of testing is shown in figure 2.1.

The task for participants was to listen to the sixteen sounds and place them into groups i.e. create categories by any means they chose. The exact instruction delivered by the experimenter was:

> " Please group together the sounds that you think belong together.
> You may do this by any means you choose"

Sound ID	Description	Predefined Category
ALRM	alarm clock ringing	Environmental
CAR	car engine starting	
DR	door opening	
FSTP	footsteps	
GLS	glass breaking	
HELI	helicopter flying overhead	
WTR	running water	
BEL	church bells outside	
GTR	arpeggiated notes on an acoustic guitar	Musical
OBOE	single note from an oboe	
VLN	short 7 note melody on a violin	
XLY	single note from a xylophone	
CGH	male voice coughing	Vocal
FEM	female voice speaking	
MALE	male voice speaking	
LGH	female voice laughing	

Table 2.1: 16 common everyday sounds used in the FST. Sounds are divided into three different types of sounds - Environmental, Musical and Vocal sounds. Finally an short identification label (ID) is stated as subsequent figures have been plotted using these labels.

Only minimal feedback was given by the experimenter in order to facilitate the completion of the experiment. Sounds were played by using the PC mouse to double click on the numbered boxes which were positioned numerically in the same order for all participants and groups were created by dragging and positioning boxes together on screen. This was always done by the participants themselves, including the CI users. Once participants had finished positioning the boxes into groups they were asked to provide a brief description for each of the groups they had created.

The experiment was designed in order to test the free-categorization approaches of the participants. There was no limit on the amount of time given to complete the test or on the number of times a specific sound could be listened to. Participants were also given no instruction as to how they should complete the task and were allowed to create as many or as few groups as they wished. Thus a group could contain only a single stimulus or all sixteen. Each participant‘s choice of groups was saved as an indicator matrix in a *.txt* file whilst a separate

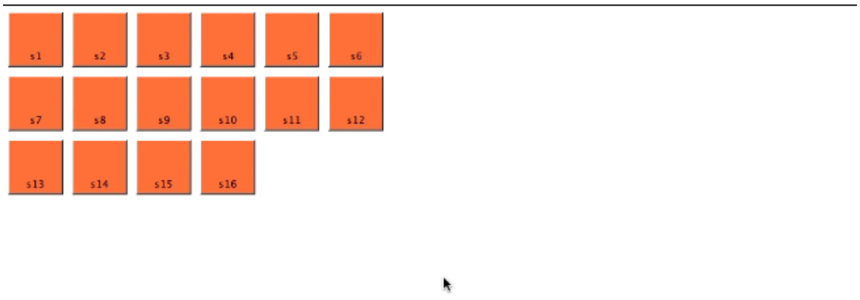


Figure 2.1: Screen-shot of the FST at the beginning of each test. Orange squares represent each of the 16 sounds which could be heard when participants double-clicked on each square with the PC mouse.

.txt file was created for participant's group descriptions and most typical stimuli.

The TCL-LabX software also recorded performance data and statistics for all participants including the number of categories created, the number of times each sound was listened to and the duration of the experiment.

2.1.2 *Analysis*

2.1.2.1 *Categorical Analysis*

To analyse the categories that participants created two different functions were used in R'. Firstly Multiple Correspondence Analysis (MCA) was applied to the indicator matrix outputted by the TCL LabX software. The indicator matrix itself represents the results as an array of categorical variables (participants) as columns and categorical items (sound stimuli) as rows, with each cell containing a number defining the category membership of each sound for each participant. MCA uses Correspondence Analysis (CA) in order to represent each sound as a data point in an n-dimensional Euclidean space based on the categorical values i.e. the categories made by participants. Each dimension is chosen to account for the largest amount of variance possible within the data-set and dimensions are outputted in descending order of variance covered. MCA also performs analysis on the participants (categorical variables) and is able to show how strongly each participant's results can be explained by each dimension This is done using based on co-occurrence matrices of the participants [23].

A total of fifteen dimensions were used in the analysis with the first five selected for use in the HCPC calculations based on the amount of variance being equal to or greater than 8%. The two most significant dimensions (Dim 1 & Dim 2) were also focused on as they account for the most amount of variance in the data and also show the most significant correlations to acoustic variables measured for the sounds (see Table 6). Importantly dimensions are calculated only to account for variability within the data and are not directly related to any perceptual or physical characteristic of the sounds or ecological data of the participants. There is no a-priori knowledge that can be used to automatically make such a relation and so a certain amount of interpretation is used when commenting on what characteristics the dimensions may represent [23].

RV coefficient (RVc) values were also calculated using the function coeffRV. The RVc is a variation on the squared Pearson correlation coefficient that calculates the correlation between two sets of coordinates represented in a matrix [112, 109]. In the case of this study the RVc was used to find the correlation between the coordinate matrices for the first five MCA dimensions of the two participant groups.

Secondly a Hierarchical Clustering based on Principal Components (HCPC) was performed on the results of the MCA analysis in order to view a simplified version of the categories of sounds in the form of dendrograms. When using this analysis it is not possible to account for all of the variance (inertia) within the data, i.e. the variability of participant responses, and so a certain amount remains unaccounted for. By increasing the number of desired categories the inertia can however be reduced and it is using this process that the we can choose a final number of categories: if the number of categories is Q then the optimal number of categories is found when the change in inertia is greater when moving from Q-1 to Q than from Q to Q+1 (François et al., 2010). This can also be defined as the value for Q which minimizes equation 2.1.

$$\frac{Q - Q_{Q+1}}{Q_{-1} - Q} \tag{2.1}$$

2.1.2.2 *Comparing Clusterings*

Categorical analysis was concluded by finding the CCC [118] which gives a measure of how accurately the distances between items in the raw data are preserved in the dendrogram [25]. CCC has been computed in R′ using the functions dist - to calculate the euclidean distance between sounds within the n-dimensional space as computed using MCA; cophenetic - to find the cophenetic distances between the stimuli computed using HCPC; cor - to correlate the two distance matrices and give a final coefficient.

2.1.2.3 *Category Identification*

In order to further understand the performance of participants in how they categorized the sounds an estimation of "Categorization Identification" was inferred from the participants' category descriptions. For each sound the associated category descriptions were evaluated as to whether they referenced musical, vocal or environmental sounds. These were then compared to the predefined categories and the percentage of comments in agreement was calculated across all participants for each sound. In some cases participants did not give any specific description and this is reflected in Table 3 by the "% No comment". In addition a value for "Categorisation Accuracy" was found based on the stimuli that were paired together by participants. This was calculated as the percentage of stimuli-pairs present in each participant's category choices compared to the stimuli-pairs contained in the predefined categories, of which there are forty possible pairs, (Table 2).

2.1.2.4 *Acoustic analysis*

Alongside categorical analysis of participants responses, the sounds themselves were analysed for a range of acoustic variables [51]. To evaluate the functional significance of the MCA dimensions the acoustical values obtained for each sound were correlated using a Pearson correlation to the coordinates of each MCA dimension of the two participant groups. Correlations with MCA dimensions were performed for several acoustic variables in order to reveal the acoustic properties important for categorizing the stimuli as well as any differences in strategy or ability between the participant groups (see section 2.3). Six different acoustical domains were explored and are detailed below.

1. Pitch measures - Including the mean and median frequency, standard deviation of frequency, max frequency, mean pitch salience, max pitch salience. Pitch values are calculated using Slaneys Correlogram model of pitch perception [127] using a temporal windows of 16 ms correlated at different time lags. pitch Salience, which can be assumed to be the perceptual strength of the pitch, is calculated by dividing the maximum value of a sounds correlogram (which occurs at zero time lag) by the estimated pitch taken from the sounds correlogram. Values range from 0 to 1 where 1 indicates a periodic sound with easily perceivable pitch.

2. Spectral measures- Analysis was made on individual sounds to compute the centroid, skew, kurtosis, mean centroid, spectral centroid velocity, spectral centroid uniformity and spectral centroid standard deviation. Centroid, skew and kurtosis are measures of the moments of the spectrum where the centroid is the central mass of the spectrum linked to the brightness of sound. The skew correspond to the asymmetry of the probability distribution of frequencies while the kurtosis is a descriptor of the shape of a probability distribution. The remaining variables related to the centroid values are measures of spectral movements within the sounds.

3. Envelope measures - Including the number of peaks, mean peak, number of bursts, mean burst, total burst duration, duration ratio computed from the wave envelope of each sound. Peaks are defined as fast transient changes in the envelope. Bursts are continuous increases in amplitude of 4dB held for at least 20ms [6], whilst the duration ratio provides an evaluation of the "roughness" of the envelope.

4. Periodicity measures - Values were obtained on the number of autocorrelation peaks, the maximum autocorrelation peak, mean and standard deviation of autocorrelation peaks, range of data. Periodicity is obtained by firstly computing the autocorrelation of each sound (using the function xcorr in MatLab) then calculating the previously stated variables to give

an indication of the frequency, strength and uniformity of periodicities. Periodicity is uncovered by analysing the autocorrelation of each sound. The stated variables then give an indication towards the frequency, strength and uniformity of these periodicities.

5. Cross-Channel correlation - an average value of the correlations between frequency bands of the envelope was calculated for each sound giving an indicator of signal uniformity.
6. RMS values - The RMS power was measured across different frequency bands centred at frequencies of the Bark scale [147].

2.1.2.5 *Statistical analysis on Participant performance*

As previously mentioned the TCL-LabX software recorded performance data and statistics for all participants including the number of categories created, the number of times each sound was listened to and the duration of the experiment. Statistical testing using ANOVA, t-test and kruskall wallis tests were was performed on this data to reveal if participant groups differed in their performance of the FST and also whether this data could add to the interpretation of participants categorisation strategies. For example trying to see if any of the sounds were listened to significantly more or less than others in order to understand more about how the sounds were perceived and whether this was a factor in participants categorisation strategies.

2.2 NHL WITH NATURAL AND VOCODED SOUNDS

2.2.1 *Method - stimuli, subjects and process*

The method and stimuli described above in subsection 2.1 were used to test three groups of NHL. One group was tested with the original natural stimuli (seen in table 2.1) and will here on in be referred to as NHL and will be used as the control group for comparisons both here and in later sections. The other two groups of participants were tested with vocoded versions of the same sounds, using either eight channels or sixteen channels and are referred to as 8C and 16C respectively. Vocoding was done using a noise vocoder in order to simulate the perception of sounds with a CI. A summary of the participants tested is also given in table 2.2.

Stimuli condition ID	Number of Participants	Stimuli description
NS	20	unprocessed natural sound
8C	20	vocoded with 8 channels
16C	21	vocoded with 16 channels

Table 2.2: Table showing the number participants for stimuli condition, NS, 16C and 8C, and a description of how the stimuli were treated in each condition.

2.3 RESULTS

2.3.1 *FST performance*

Performance of the three participant groups in the FST appears similar from figure 2.2. Whilst values for vocoded conditions (8C & 16C) are greater these differences are not significant. Neither for the number of categories created by participants ($kwp = 0.306$), nor the average number of times each participant listened to the sounds ($kwp = 0.5437$). Also the amount of time taken to complete the task seems greater for the 8C condition but is also not significant ($kwp = 0.063$). It should also be noted that one participant in the 8C group took 6405 seconds to complete the task. This result has been regarded a measurement error and has been excluded from the calculations of the average value (shown in figure 2.2) as well as statistical analysis as it greater than the bootstrapped average (425.62s).

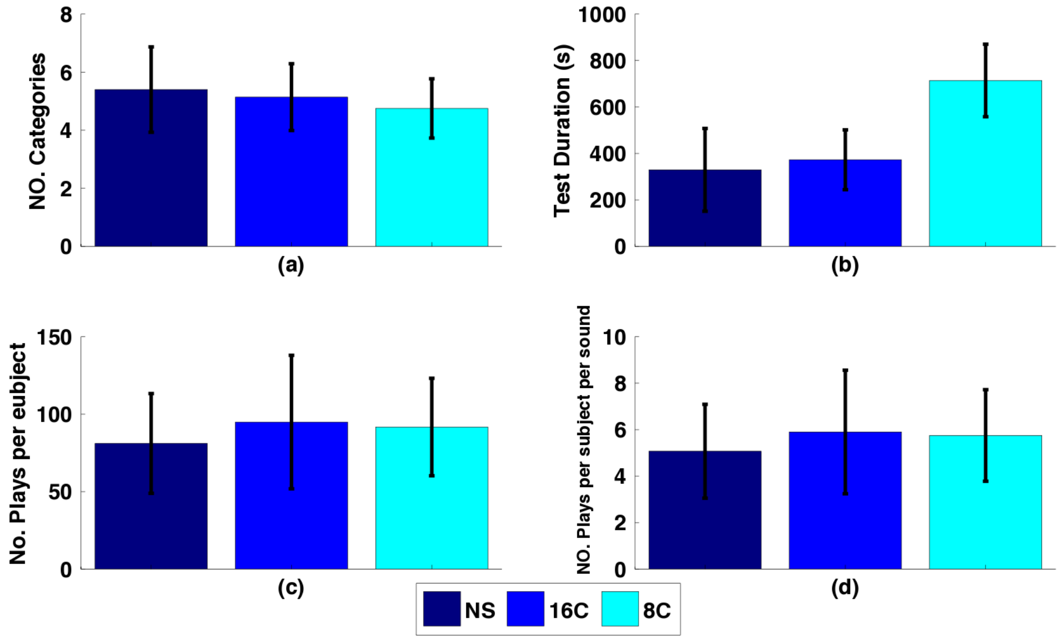


Figure 2.2: Figure showing the FST performance statistics for the three stimuli conditions tested with NHL - NS in dark blue, 16 channel vocoding (16C) in Royal Blue and the 8 channel vocoding condition (8C) in light blue .

Categorical analysis also shows more similarities between the participant groups. From a purely numerical perspective figure 2.3 shows that the amount of variance covered by each of the five dimensions is nearly exactly the same for all three stimuli conditions and values of unaccounted variance are also similar. The percentage of subjects with coordinates greater than 0.8 (see figure 2.10) is also similar although reduced across all dimensions for the 8C condition.

Taken overall figure 2.3 and 2.2 show that within each stimuli condition participants were in agreement and performed the task in a very similar manner to each other showing little variation. Across the three stimuli conditions there also did not appear to be significant differences in performance when vocoding the sounds.

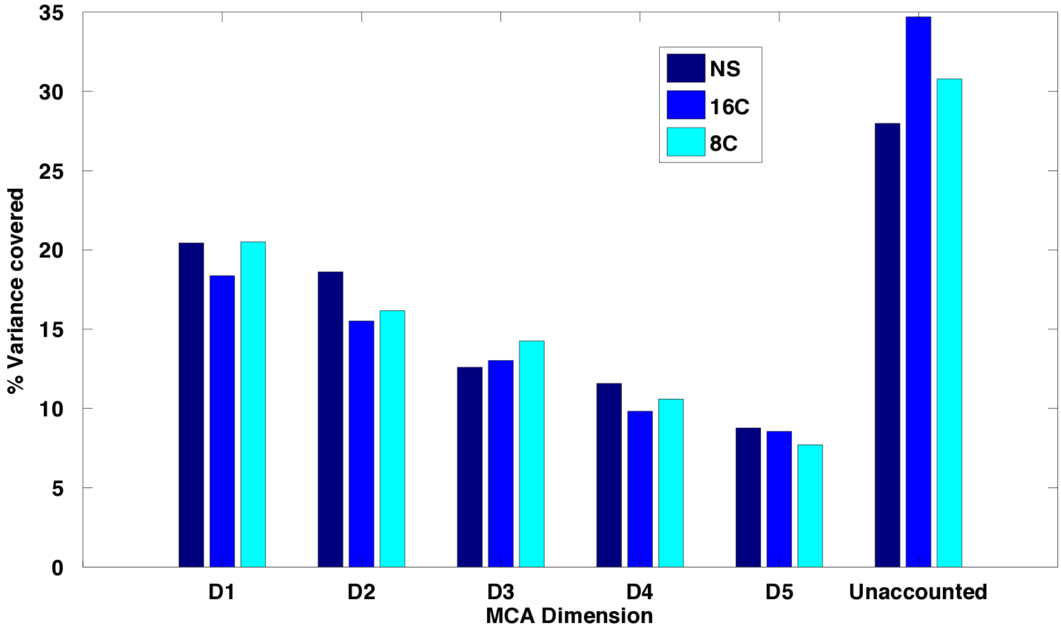


Figure 2.3: Bar plot of the variance covered by each of the five retained MCA dimensions for NHL in NS, 16C or 8C testing conditions. Unaccounted variance is the variance covered by dimensions $6 - 15$ from the analysis.

2.3.2 *HCPC analysis - dendrograms*

Using the retained dimensions a simplified categorisation for each participant group can be seen as dendrograms in figure 2.5. In order to validate these figures Cophenetic correlation coefficients (CpCC) have been calculated for NS = 0.78, 16C = 0.74 and 8C = 0.84. These values show how accurately the dendrograms (figure 2.5) represent the original data. Being close to 1 it means the data are well represented and shows that for each stimuli condition participants performed likely in a uniform way. Interestingly the value for 8C is highest meaning the dendrogram in figure 2.5 is more representative of the 8C participant group than that for 16C and NS.

The ordinate of the dendrograms represents the distance or similarity between stimuli, with a larger distance being indicative of a greater dissimilarity. Average values of the Intra- and inter-category distances are shown in figure 2.4 where the Intra-category distance is significantly different only between NS and 16C (kw, $chisq = 7.72, p < 0.05$) and the inter-category distance is greater for NS compared to both 8C and 16 C(kw, $chisq = 14.48, p < 0.001$). It means that for NS sounds within the same category are perceived as being more similar and that categories themselves are more dissimilar or further separated from each other. It is also interesting that values for the two vocoded stimuli conditions are similar and is further evidence for similar performance in these two conditions. One difference however is seen for the separation of linguistic sounds (*MALE* and *FEM*) which occurs at a height of 0.737 for 8C but only0.614 for 16C. It is therefore likely that in the 8C condition these sounds are more strongly distinguished from the others than in the 16C condition. Finally the category of BEL, WTR, GLS is repeated for both 8C and 16C and does not correspond to any results seen for NH. As NS categorise the BEL sounds amongst musical sounds it is likely that it is not considered as an environmental sounds, such that the categorisation with WTR and GLS is likely due to a change in the perception of this sound when vocoded.

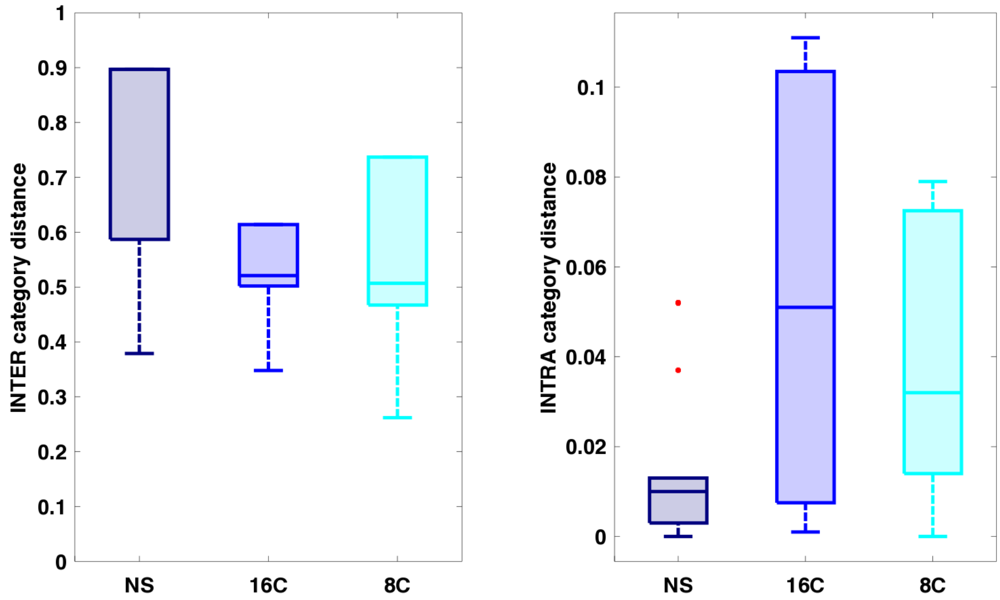


Figure 2.4: Boxplot of Inter & Intra category distances calculated from the heights of dendrograms in figure 2.5.

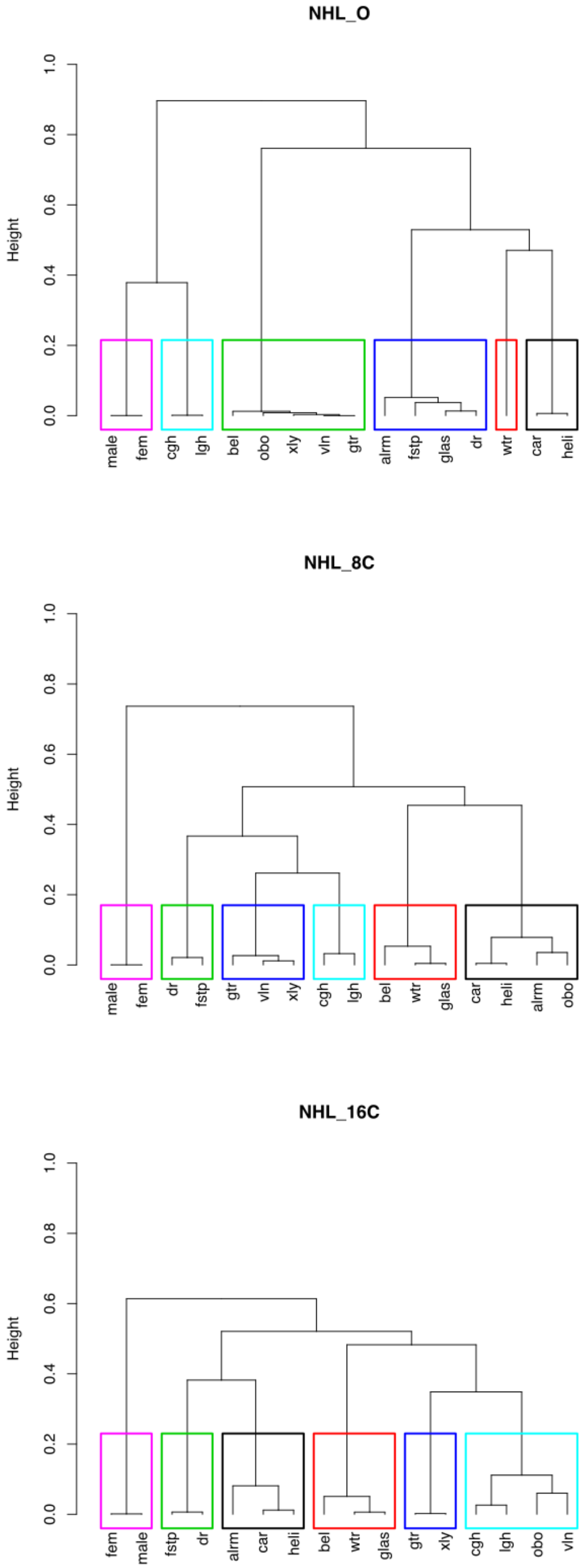


Figure 2.5: HCPC Dendrogram for NHL across three stimuli conditions. Categories outputted by HCPC are indicated by coloured rectangles whilst the upper limit of these rectangles indicates where the tree has been cut (as described in the main text). The height axis gives the perceptual distance between each stimulus Finally stimuli are labelled using the abbreviated sound IDs from Table 2.1.

2.3.3 *Similarity comparison - co-occurrence matrices*

Aside from the similar pairing of linguistic sounds *MALE* and *FEM*, there are also other noteworthy similarities in all stimuli conditions which can be shown in figure 2.6. This figures represent the co-occurrence (similarity) matrices which are a measure of how often a certain sound was paired with another i.e the number of participants that created this pair, where 1 (maroon) would represent a pair created by ALL participants and 0 (blue) a pair never made. In the NS stimuli condition there is stronger agreement amongst participants in pairing together; musical sounds, especially for the sounds GTR and VLN; vocal sounds (MALE with FEM and LGH with CGH) and; between the sounds FSTP, GLS and DR. The pattern of results also shows that participants in the NS condition are more uniform categorisation strategy as there are many sounds either strongly paired together or not at all paired together. Compared to this the results for 8C and 16C demonstrate less agreement between participants such that the figures are more uniform in colour.

Similar pairs do also exist across all three stimuli conditions - MALE-FEM, HELI-CAR, GTR-VLN, GTR-XLY and FSTP-DR.. These pairs likely represent certain kinds of sounds - linguistic, transport/mechanical, musical and finally action sounds. Within the plots for 8C and 16C there is also a shared pair of WTR-GLS that is not strongly represented in the plot for NH. The sound BEL is also linked with these two sounds (as mentioned above) on the plot for 8C, but not for 16C. It is possible that with an decrease in spectral information (moving from sixteen to eight channels) that the sound BEL therefore becomes more perceptually similar to the sounds of WTR and GLS.

2.3.4 *MCA analysis - factor maps*

Figures 2.7, 2.8 and 2.9 show how the sounds are displayed across the various dimensions outputted by MCA analysis. It is clear to see that many of the dimensions can also be interpreted in the same way for all stimuli conditions and they are discussed below:

1. **Dimension 1:** shows a separation of the speech related vocal sounds in comparison to all others. However the non-linguistic vocal sounds (*cgh*

and *lgh*) are only grouped alongside the speech sounds (*male* and *fem*) in the NS condition and not for either of the vocoded conditions.

2. **Dimension 2:** in NS there is a clear separation of musical and environmental sounds but this does not seem to be repeated for 8C and 16C. Instead the two vocoding conditions are more similar to each other in that Dim 2 separates the same two groups of sound (water, glass, bell, heli, car, alarm, oboe) vs. (guitar, xylophone, violin, cough, laugh), which almost, but not quite corresponds to categories of musical vs. environmental sounds. It is interesting to note that when looking at Dim 1 & Dim 2, although the grouping of sounds may not be the same there are many pairs or subgroups of sounds often plotted very close to each other - e.g. ftstp & door, guitar & xylophone, cough & laugh. The interpretation of Dim 2 is likely different for 8C and 16C however, because of the sounds which most strongly contribute to the dimension. For 8C this is car & heli vs guitar, xly & violin whilst for 16C it is water & glass vs. door, guitar, fstp & xly.

3. **Dimension 3 & 4:** figure 2.7 shows that Dimension 3 and 4 are used to distinguish the environmental sounds in the natural sound condition. Vocal and musical sounds are located on/around the zero line eof both dimensions such that they are not heavily important to either dimension and interestingly vocal sounds are also positioned similarly for 8C and 16C. There is a clear separation of the car & heli sounds on dimension 4, as well as the water sound. Figure 2.8 also shows that for the 16 channel vocoded condition the car and helicopter sounds are again distinguished along dimension 4 and for 8C it may be seen on Dim 2. Rather for the 8C condition dimension 3 & 4 separate glass, water & bell (dim 3) and then footstep & door (dim 4).

4. **Dimension 5:** the weakest dimension retained in the analysis is difficult to interpret for 16C, but shows similarities between NS and 8C in the separation of coughing and laughing sounds. Interestingly for NS this is a direct 2-way choice between the linguistic and non-linguistic sounds, where as for 8C it is only the preference of the non-linguistic vocal sounds, suggesting that in the 8C condition these two sounds are perceived as similar, different to the others, but not strongly associated with the other

vocal sounds as is the case in 16C and NS.

This pattern of similarities is also evidenced by calculations of the RV coefficient. Correlations of NS with vocoded conditions show high correlations, with 16C $RV = 0.73; p = 1.8e^{-6}$ and 8C $RV = 0.709; p = 5.57e^{-6}$. A higher correlation is also seen between 8C and 16C, $RV = 0.85; p = 5.95e^{-8}$. These correlations may help show that whilst the results of the vocoded conditions are very similar to those in the NS condition but are more similar to each other, suggesting that the perception of sounds is different with the use of any vocoding independent of the number of channels used.

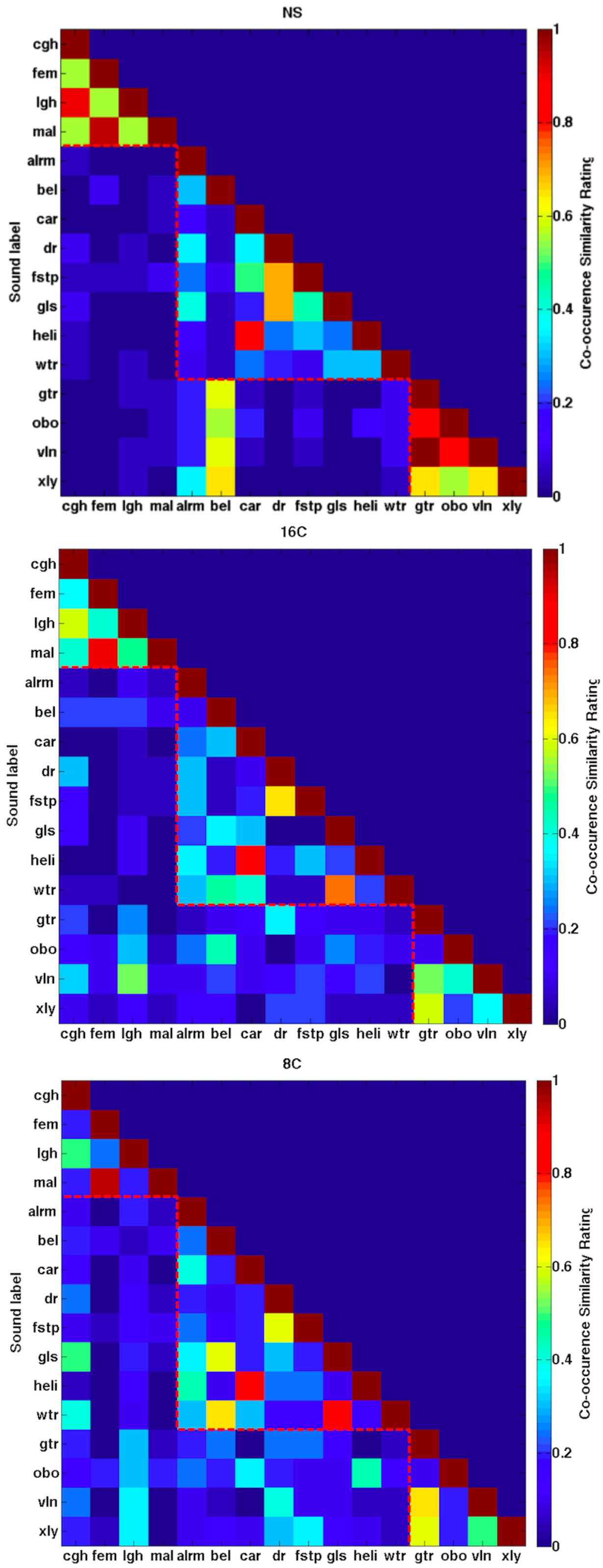


Figure 2.6: Co-occurrence (similarity) matrices for the three testing groups NS, 16C and 8C. Each cell corresponds to the number of participants who paired two stimuli together, with low similarity indicated by blueand high similarity by red. Stimuli are arranged by their predefined category, where the red dotted line separates the three categories of vocal, environmental and musical sounds.

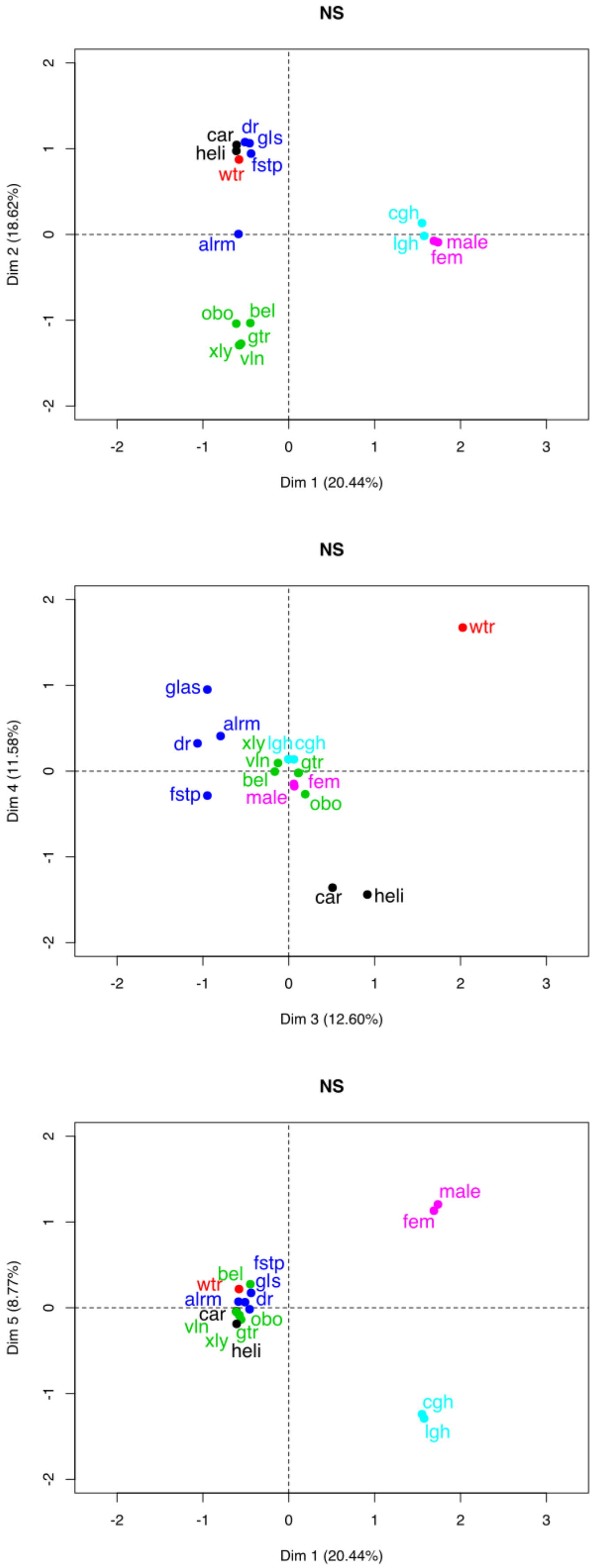


Figure 2.7: MCA Factor maps for NHL in the NS stimuli condition. Dimensions 1-5 are shown across the three panels with percentage of variance covered also shown. Colours correspond to the categories shown in the HCPC dendrograms of figure 2.5 (upper panel).

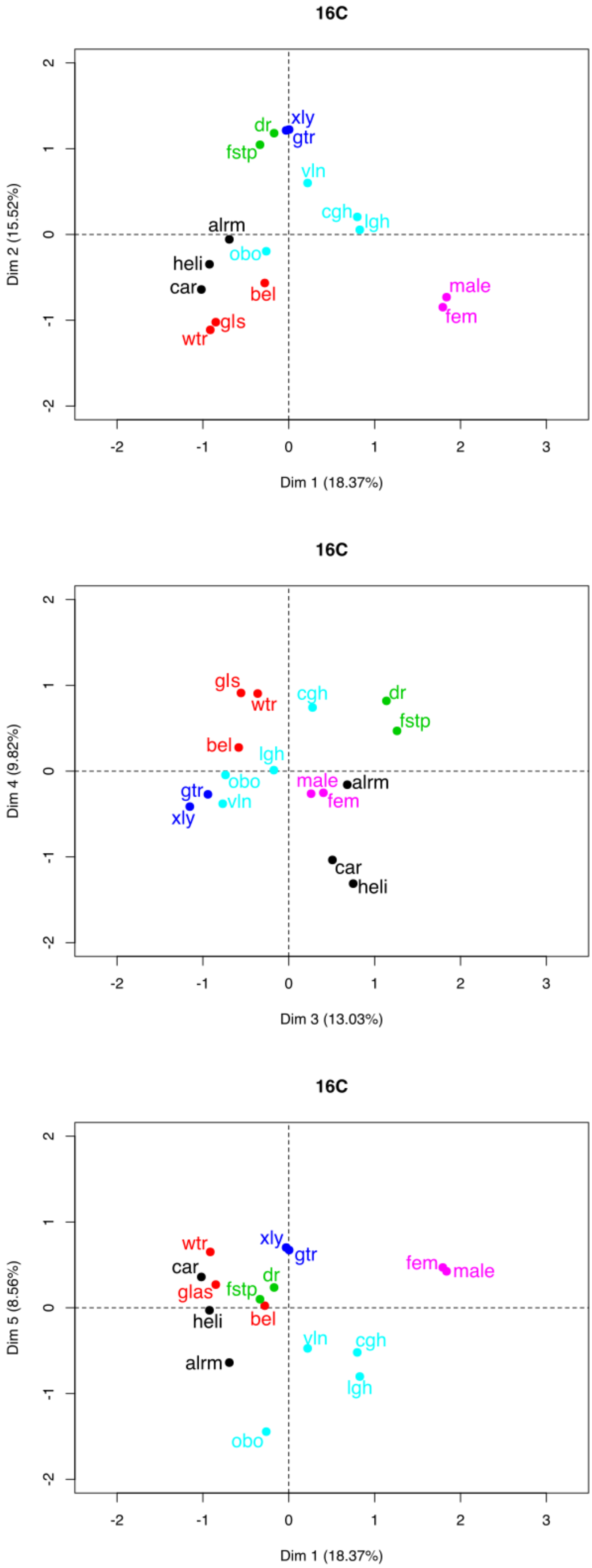


Figure 2.8: MCA Factor maps for NHL in the 16C vocoded stimuli condition. Dimensions 1-5 are shown across the three panels with percentage of variance covered also shown. Colours correspond to the categories shown in the HCPC dendrogram of figure 2.5 (middle panel).

2.3.5 *MCA analysis - subject maps*

The pattern of results displayed by the subject maps in figure 2.10 can also be used to evaluate the level of agreement by participants. Across all three stimuli conditions the coordinates of Dim 1 & Dim 2 show most participants are grouped together toward the top right hand corner of the subject maps indicating strong agreement in the use of these dimensions. For Dim 3-5 (not plotted) the pattern of values is much more dispersed showing less agreement. Although this is a result of the analysis it can be interpreted that the these latter dimensions are only important to certain participants.

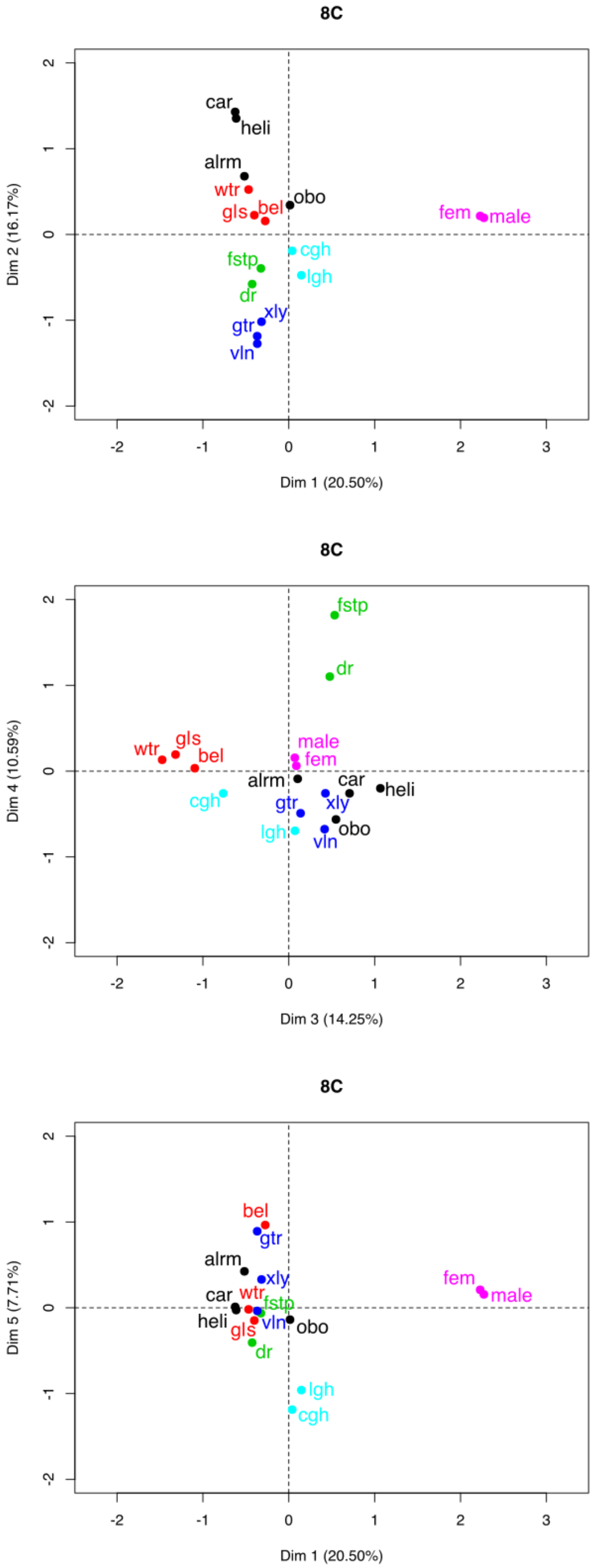


Figure 2.9: MCA Factor maps for NHL in the 8C vocoded stimuli condition. Dimensions 1-5 are shown across the three panels with percentage of variance covered also shown. Colours correspond to the categories shown in the HCPC dendrogram of figure 2.5 (lower panel).

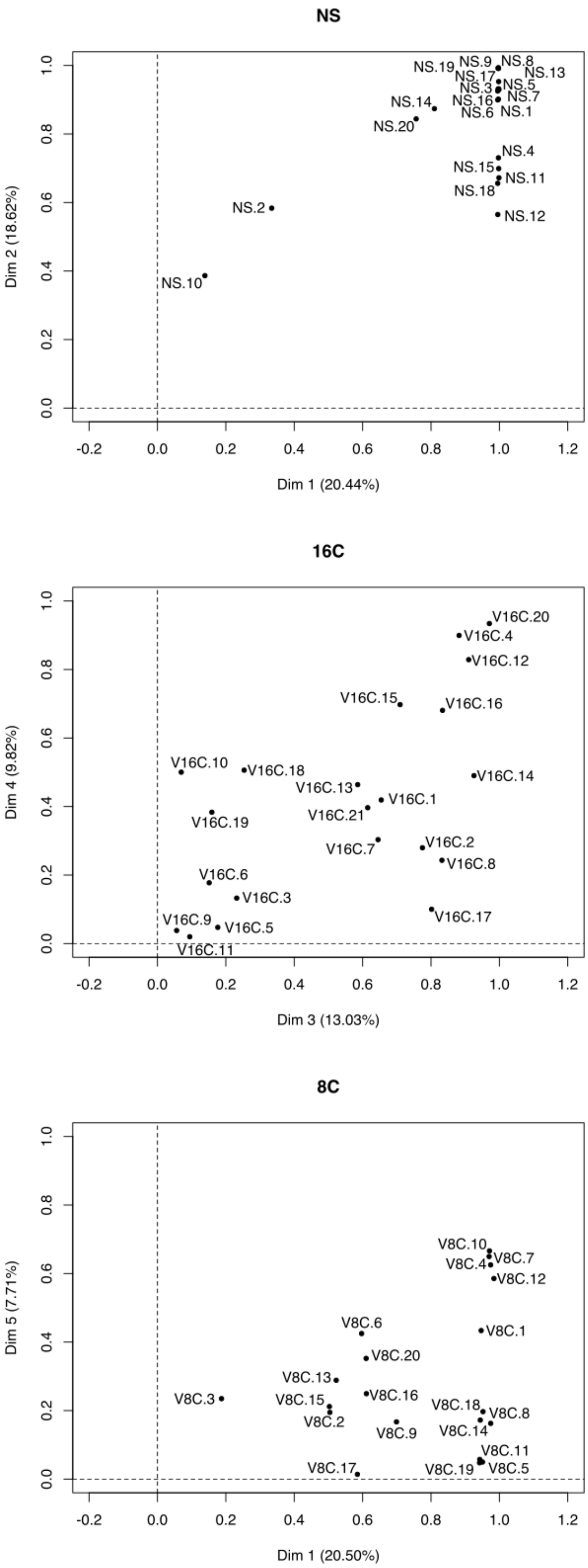


Figure 2.10: Figures showing the distribution of NHL in all 3 stimuli conditions Dim 1 & 2. Results are plotted from $0-1$ and can be interpreted as showing how strongly the results of each participant are matched to each stated dimension, where a value of 1 would show a strong match. The mean value of each dimension is also represented by the Eigenvalues in table 2.3. Finally Dim 1 & 2 are plotted in the top panel, Dim 3 & 4 in the middle panel and Dim 1& Dim 5 in the bottom.

2.3.6 *Categorisation Identification*

Categorisation Identification was also calculated by comparing the participants comments to the originally pre-defined category for each sound. Results are shown in table 2.11 and show that the most correctly categorised sounds were speech sounds and the melodic musical sounds (GTR and VLN). Whist the results for NS are greater there is not a significant difference between the two vocoding conditions.

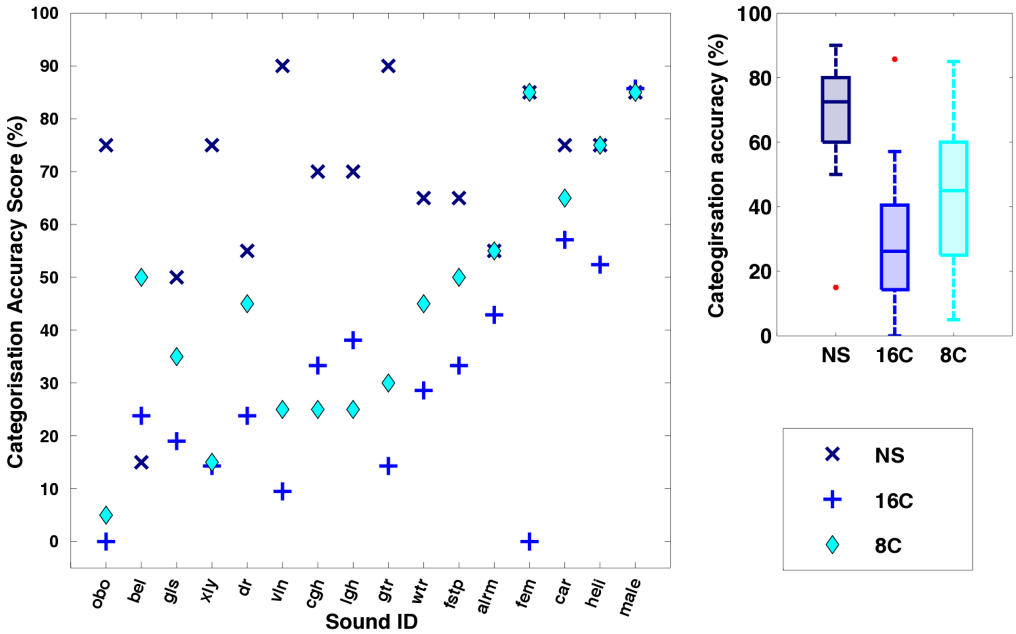


Figure 2.11: Categorisation identification for NHL, where NS is represented by ×, 16C by + and 8C by ◇. Results for individual sounds are plotted in the left hand scatter plot in order of mean category identification across all three stimuli conditions, whilst results for each stimuli condition are plotted as boxplots in the right hand panel.

2.3.7 *Acoustic Analysis*

As mentioned in the methods section acoustic analysis was undertaken on the sounds. These values were then correlated to the coordinates of the MCA dimensions and table 2.3 list the significant correlations that were found in each stimuli condition.

Acoustic Variable	NS	16C	8C	
				Temporally related
Autocorr Mean Peak	D2	D3		
Autocorr std Peak	D2	D3	D2	
Autocorr No Peaks		D2 & D3	D4	
Autocorr Range		D1		
No Bursts	D5			
std Frq			D4	Spectrally related
Mean Salience	D2	D1		
Max Salience	D2	D1	D2	
Spectral kurt	D5		D5	
Spectral STD			D5	
Mean BPF		D1	D2	

Table 2.3: Table of MCA coordinates along each of the five dimensions (listed by D1 ....5) that showed significant correlations ($p < 0.05$) to acoustic variables for each NHL stimuli conditions (NS, 16C and 8C). For example, in the NS condition there was a correlation between MCA Dimension 5 (D5) with measure of the No. Bursts and Spectral Kurtosis of the sound stimuli. Acoustic variables are also divided between those that are related to temporal and spectral measures.

The correlation of the number of peaks in the autocorrelation function to 8C Dim 4 and 16C Dim 3 can be linked to the sounds being periodic and is a result of the fact that the sound footstep (*fstp)* has the highest number of peaks (4). This however results in it having a low average peak in comparison to sounds with only one large peak, resulting in the negative correlation to the mean and std peak. Which in the same way explains the negative correlation for Dim 3 16C - because the sound footstep has low mean autocorr peak value.

Pitch Salience associaated to Dim1 of 16C, cos there is a different pattern of results in comparison to NS and 8C. More gradual dispersion for 16C.

Max salience on Dim 2 for 8C is due to extremities again, with musical sounds (high pitch salience at the lower extremity) and car and heli with low pitch saline at the top. Spectral measures (std and kurtosis) linked with the Dim 5 split of bel and gtr at the extremity, which have low spectral std as they are musical notes, and kurtosis which described the shape of the probability distribution of frequencies such that the frequency distribution for bel and gtr is toward the higher frequencies.

2.4 DISCUSSION

A FST of 16 everyday sounds was performed by three groups of NHL. For each group different levels of vocoding were applied to the sounds, using either 8 channels (8C), 16 channels (16C) or using the natural (unmodified) sounds (NS). Results show strong similarities across all three stimuli conditions in terms of the primary strategies used for categorising the sounds, especially regarding linguistic and musical sounds. Although the exact manner in which the strategies are used may be somewhat different in the face of vocoding which causes altered perceptions. Certain pairs of stimuli are also present in all three stimuli conditions indicating further potentially shared strategies of categorisation. Finally the NS condition shows differences to 8C and 16C with a more distinct separation of categories, a higher agreement between participants and a more uniform performance. Overall results show underlying similarities for the categorisation of sounds by NHL even with vocoded sounds. Differences in vocoding conditions may also may also help to understand how the perception and subsequent categorisation of sounds changes when spectral information is reduced.

For all stimuli conditions vocal sounds seem very important in that they are they first and most distinct category made by participants. Linguistic vocal sounds especially seem robust to the effects of vocoding and the reduction of spectral information, which is somewhat expected knowing that speech sounds can be well identified with as little as four channels in comparison to only 66% for environmental sounds [126, 93]. Non-linguistic vocal sounds, although often paired together were however affected by vocoding in that they were no longer categorised alongside the linguistic sounds, even in the 16C condition.It appears that with decreasing spectral information the similarity between speech and non-speech sounds became much weaker. Inferring from category descriptions some participants from both 8C and 16C conditions did indeed perceive the *coughing* and *laughing* sounds as human vocal sounds. However there is greater variability in perception and more often these sounds are perceived as musical, quite probably due to the strong temporal nature. Certainly for the sound of coughing it carries very little spectral cues that may be used in the identification of the human voice. Why the perceptions differed is not clear although in other studies training effects have played a role in the identification of vocoded sounds [123]. This is somewhat similar to the results of [93] who also found that performance was better for speech vs non-speech sounds when discriminating against environmental sounds.

The robustness of linguistic stimuli (and non-linguistic in the case of natural sounds) is likely related to two factors. Firstly the human vocal system does not vary wildly (especially in comparison to environmental sounds) amongst different people. In this manner it represents a singular source with "fixed" size and smaller variability in the acoustics that it can generate. Compared to non-speech, speech stimuli are also affected differently by vocoding such that certain cues are well preserved in vocoding for example temporal modulations specific to the speech signal and articulation cues [93]. Secondly the addition of linguistic information may aid the listener in providing semantic information i.e. words, phonemes, grammatical and semantic context, which enables top-down processes to aid the discrimination of sounds. This so called "grammar" [8, 51] may help in limiting the choices that participants have when perceiving/identifying a sound.

In contrast to vocal sounds, environmental sounds constitute a huge range of sources which vary wildly in acoustic nature. As such it is incredibly difficult to define a set of acoustic information that can account for environmental sound identification. There is also the additional factor that different environmental sounds are affected in different ways when vocoded. For example in a test of ESP [119] found that identification did not improve significantly beyond 8 channels and even with 32 channels identification accuracy was only at 70%. However some sounds were correctly identified with only 4 channels, thus demonstrating the difference in perception of environmental sounds. using band-limiting the same study also demonstrated the importance of higher frequencies in the perception of environmental sounds in similarity to the finding of [51] who noted the importance of information within the $1.2 - 2.4$ kHz frequency band. Whilst identification of sounds was not tested, results of category identification are significantly lower in both vocoded conditions compared to the natural sounds, showing that this process is reduces the perceptual accuracy of auditory information for naive listeners.

Apart from aiding the understanding of auditory categorisation vocoding tests are highly useful in understanding the perception of sounds by CIL. Testing actual CIL has shown variable results, for example [111] found environmental sound identification of 79% (range $45 - 94\%$) amongst, however in similar non-linguistic sound test (NLST) [59] found scores of $49 \pm 14\%$ for 11 CIL. Notably testing in [111] was done by presenting stimuli from a specific setting/context (home, kitchen, office or outside) in a 10-alternate force choice procedure. whereas [59] used an open-set procedure. The closed-set nature of testing in the first study and the addition of context information may have helped to raise performance levels beyond the norm. Whilst It is clear that performance amongst CIL varies highly in similarity to results of CI simulations, the success of using such simulations is not clear (discussed further in chapter 4).

Commenting specifically on the manner in which NHL categorise auditory information the current study would agree with previous work that suggests auditory categories are based on the perception of semantic attributes of the sound producing object or action rather than qualitative acoustically related perceptions associated to the sound signal [52, 57, 91, 6, 120, 37]. For example categories of human action sounds, mechanical/transport, human voice and musical sounds appear common to all stimuli conditions and therefore also show robustness to vocoding. Musical listening as stated by [37] does not appear to be strongly influencing the results so strongly, which suggests participants are able to identify some or all the sounds to a suitable level to extract semantic meaning from them. For example a similarity shared across all stimuli conditions is the categorising of DOOR and FSTP together. Category descriptions associated to these two sounds are also very similar describing either*everyday noise* or *action* sounds. A similar pattern is also observed regarding the categorisation of CAR and HELI sounds, which are robustly categorised as mechanical sounds in all conditions. This may be another important category as the number of comments referring to *mechanic/motor* sounds increases as the number of channels is reduced. This could also be because the car and heli sounds are more easily identified even when vocoded such that subjects simply place ambiguous sounds alongside them, for example the sounds of ALRM and OBO. Whilst descriptions of the individual sounds were not recorded, the change in category descriptions mirrors what is already known, that vocoding can have very different effects on the perception of different environmental sounds [54, 122].

Importantly it also seems that categorisation is being made on the basis of semantic information concerning the sound producing object even with the lowest amount of spectral information available. Whilst identification performance was not measured it may be that participants are still able to identify sounds to a level where they are able to extract some form of semantic meaning. This definitely seems to be happening with regards to vocal, mechanical/transport and melodic musical sounds.

2.5 CONCLUSIONS

The current chapter tested three groups of NHL with a Free Sorting Task (FST) of 16 everyday sounds that were divided into three predefined categories of environmental, musical and vocal sounds. Two groups of NHL participants listened to sounds that had been passed through either a 16 or 8 channel vocoder in order to simulate a Cochlear Implant (CI) device. Results showed that on a broad level categorisation strategies in each participant group divided stimuli into the three predefined categories with vocal sounds (especially linguistic sounds) strongly separated from others. Across testing conditions similar categories or stimuli pairings were found for mechanical/transport, human action and musical sounds and demonstrated that participants perceptions of sounds were grounded in semantic or causal information associated with the sound producing object and/or action, even when spectral information was reduced the most (8C).

Key Points

- Across all stimuli conditions, Natural sounds (NS) and CI simulated conditions of 16 and 8 channel vocoding the three groups of NHL show very similar performance in categorising the stimuli. On a broad level the sounds are categorised in agreement with the predefined categories of vocal, environmental and musical sounds.
- The most dominant categorisation strategy corresponds to the contrast of vocal sounds with other sounds. This is in keeping with previous research that observes the strong categorisation of vocalisations, as these sounds represent one sound source - the human vocal tract, and are highly familiar to human listeners. However, with increased vocoding this appears to apply only to linguistic vocal sounds as have a speech shaped envelope.
- Musical sounds with a sense of melody also seem to be more strongly categorised together in comparison to sounds of only a single tone, which are more often categorised amongst environmental sounds. This is most likely due to causal uncertainty of such sounds along with the idea that sounds are categorised predominantly based on causal similarities. Rather than being identified as musical simply because a sound contains a strong pitch, unless it is identified as originating from a musical instrument it appears that it will be categorised based on other semantic information, for example originating from a non-musical sound source possibly as a warning sound.
- Human sounds (both vocal and action) as well as machine/transport sounds are robustly categorised even in the 8C condition. This suggests that the perception of these sounds is affected not by the reduction in spectral information and may instead rely more strongly on temporal information. For example a characteristic of engine sounds is their repetitive and broad spectral nature, whilst action sounds often involve short temporal events such as a result of impacts.
- The main impact of vocoding techniques seems to be in altering the perceptions of non-linguistic vocal sounds and single tone musical sounds. Whilst the main categories are preserved, and the first two MCA dimensions are consisted in all stimuli conditions, categorisation performance in later dimensions changes slightly.

3

AUDITORY CATEGORISATION WITH COCHLEAR IMPLANTED LISTENERS - HOW DOES IMPLANTATION DURATION AFFECT PERCEPTION

Three groups of CIL who had different durations of implantation ranging from $0-6$*,* $7-14$ *and* $14+$ *months, were tested with the same FST as previously described in chapter 2. Results show that the more experienced CIL were capable of separating vocal, musical and environmental sounds. This was especially true for linguistic vocal sounds, musical sounds and transport/mechanical sounds and somewhat disagree with previous research that showed CIL struggling in related tasks. The least experienced group of CIL showed some similarities to the other two groups but overall were different in both performance of the task and the categorisation strategies employed. This may indicate that the region of adaptation to a FST lies around 6 months post implantation.*

3.1 INTRODUCTION

The following chapter looks at the performance of auditory categorisation performed by three different groups of CIL who differed in their duration of implantation. They consisted of Experienced cochlear implant listeners (EXP) with more than 14 months of implantation, Intermediate cochlear implant listeners (INT) - from $6-14$ months implantation and New cochlear implant listeners (NEW) with duration less than 6 months. CIL again performed the same FST of 16 common sounds as described in chapter 2 with results analysed using the same MCA and HCPC techniques. This was done primarily in order to find the categorisation strategies in the auditory domain that may be used by CIL and how they compare to the results of NHL in from chapter 2. However this is not discussed in the current chapter but rather explored in greater detail in chapter 4, where the results using vocoded stimuli (8C & 16C) are used for comparisons with CIL results and to evaluate the suitability of vocoding techniques for CI simulations. Chapter **??** will therefore focus on understanding the categorisation strategies of the three CIL groups and how they evolve with increasing duration of implantation. It is known that with increased adaptation to the implant CIL show improved speech abilities that plateau around 6 months post-implantation and reach maximum scores of approximately 80% after 12 months [139, 103, 18]. Speech perception has also been linked with ESP in CIL [121, 79, 111, 51, 59] which suggests that there are similar perceptual and cognitive aspects involved in the perception of both kinds of sounds. However, although some studies have shown improvements with training [121, 85], ESP in CIL is generally considered to be poor. For example in a test of vocal-environmental sound discrimination CIL showed no improvement even after 18 months of implantation [94]. ESP is therefore a complex process for CIL.

Importantly this is the first study to test the performance of CIL in a FST of common sounds, especially mixing vocal, musical and environmental sounds. Whilst closely similar studies do not exist the previous research can help to make predictions. With differences in implantation duration CIL accumulate different amounts of experience with their implants which will therefore likely create different categorisation strategies. These differences are likely to manifest in one of two ways, based on the *everyday* and *musical* listening modes suggested by Gaver [37]. Considering everyday listening, the perception of a sound is related to the semantic information associated with the sound producing event. This relies on identifying the sound source (either correctly or incorrectly) such that categorisation strategies employed by CIL may be based on semantic information, however associated to mis-identified sounds. Alternatively if CIL are completely unable to identify sounds, then musical listening may be employed resulting in categorisation strategies based on the qualitative perception of the sounds that are associated with acoustical characteristics. Of course the easiest way of analysing the effect of implantation duration is to compare the results to those of the NHL tested in the NS condition in chapter 2. Considering this comparison it is likely that the results of EXP will provide a closer match than for INT and even weaker match for NEW - as mentioned this is explored in chapter 4.

Finally it is hoped that by mixing vocal, environmental and musical sounds and posing almost no restrictions on participants that the FST represents a task more similar to "real-world" interactions with the auditory environment for CIL.

Chapter Aims

1. Establish how a set of 16 everyday sounds are categorised by three groups of CIL differing in duration of implantation. The chapter will investigate whether CIL participants categorise the sounds according to the predefined categories of environmental, musical and vocal sounds or by other means. Analysis shall also seek to understand the perceptual similarities that are used to form categorisation strategies. The perception of the different types of sound, especially environmental sounds may also help to further the understanding of CIL auditory abilities within the more real-world FST paradigm.

2. To use a Free Sorting Task (FST) for the first time with CIL and take the first steps to see whether such a test can be used for future research to efficiently test auditory abilities of CIL

3. Within the chapter results will be compared across the three CIL groups in order to understand how categorical perceptions of the sounds change with increased implantation duration and increased listening experience. The results will then be used for comparison with those of NHL from chapter 2.

4. Identification accuracy of auditory categories and individual sounds will be assessed in order to add to the existing research and further understand how the different levels of perception change with implantation duration. Whilst the identification of individual environmental sounds is predicted to be poor the categorical perception of these sounds may be informative to how they can be further studied. For example it will be interesting to see If a link between the identification of individual sounds and categories of environmental sounds can be found.

5. Acoustic analysis of sounds shall also be used in order to help uncover whether any physical attributes of the sounds are important to the categorisation strategies used by CIL or whether results agree strongly with previous theory that auditory categories are based on semantic information associated to the sound producing event, object or action. Acoustic analysis will also be important in uncovering cues that used by CIL for the perception of different auditory categories and individual sounds.

3.2 method & materials

The same method and materials were used here as in chapter 2 however for 3 groups of CIL which are based on duration of implantation and described further below.

3.2.1 *Stimuli & Procedure*

The same stimuli as used in chapter 2 were again used for collecting results with the CIL and are re-stated in table 3.1. The same test procedure was also used as is described in chapter 2 section 2.1. CIL were however tested at the Hospital Purpan

Sound ID	Description	Predefined Category
ALRM	alarm clock ringing	Environmental
CAR	car engine starting	
DR	door opening	
FSTP	footsteps	
GLS	glass breaking	
HELI	helicopter flying overhead	
WTR	running water	
BEL	church bells outside	
GTR	arpeggiated notes on an acoustic guitar	Musical
OBOE	single note from an oboe	
VLN	short 7 note melody on a violin	
XLY	single note from a xylophone	
CGH	male voice coughing	Vocal
FEM	female voice speaking	
MALE	male voice speaking	
LGH	female voice laughing	

Table 3.1: 16 common sounds used in the FST. Sounds are divided into three different types of sounds - Environmental, Musical and Vocal sounds. For each sound an abbreviated identification label (ID) is stated and is used to label stimuli on subsequent figures.

3.2.2 *Analysis of identification performance*

As part of the theory on auditory categorisation it is often stated that categorisation relies (or is strongly based) on identification of a sound-producing object. Therefore understanding how participants perceived the sounds was important and aided the interpretation of the categorisation results and strategies used by participants. For the INT and NEW CI participant groups a measure of identification performance was calculated. This was not however done for the original group EXP and so a representative sample of 5 experienced CIL was tested in order to provide comparative results.

Sound identification is not a trivial thing to analyse and there already exists literature covering this topic with CIL (see [124, 83, 111, 5]. As such whilst an in-depth test of identification was not performed it was important to understand how (accurately) participants were perceiving the sounds and to have a measure that could be used to better understand the link between identification/perception and categorisation, as well as aiding the interpretation of MCA and HCPC analysis. In order to do so participants were asked to identify each sound once the FST had been completed.

Participants were allowed to listen to the sounds again as many times as they chose and were then asked to simply describe what they heard for each sound. Responses that were easily interpreted were immediately scored as correct or incorrect. More complicated responses, including unknown french words, were written down and later evaluated with the aid of a native French speaker. Results were calculated as percentages and correlated with the coordinates of Dimensions outputted by the MCA analysis in order to see if any aspect of the categorisation strategy could be explained by the identification ability of individual subjects or the ease of identification for each sound. Analysis of identification performance can be viewed alongside the categorisation accuracy results in section 3.3.

3.2.3 *CIL participants*

As previously mentioned three groups of CIL were chosen based on the duration of implantation; EXP with duration of implantation greater than 14 months; INT from $6 - 14$ months and NEW less than 6 months. Table 3.2 gives a summary of all etiological and performance data of the three groups with individual data for participants (including the participant ID's and implant manufacturer) in tables 3.3 - 3.5. Duration of implantation was taken as a rounded number from the date of surgery to the date when the FST was performed, as such participants with duration of 6.5 months or more were placed in the INT group. Duration of deafness was taken from medical records as the first date when participants were diagnosed with a "severe hearing loss" The threshold of the non-implanted ear was taken from the participants audio-gram taken prior to implantation and calculated as the average threshold (using headphones) across frequencies $0.25, 0.5, 1, 2, 4$ and 8 kHz. Word recognition and Sentence in noise scores were recorded by the OLR audiologists as part of regular appointments with participants and were recorded either on the same date as the FST or within a few weeks but never exceeded the limits of implantation duration used to define the three CIL groups.

CIL that were tested were not age matched as this was a difficult task to achieve alongside the requirement of having a minimum of 15 participants in each group - as required by the MCA analysis technique. Statistics of table 3.2 however show that there is no significant difference between the three CIL groups in terms of the distribution of age (kruskal-wallis, p =0.0815), nor duration of deafness (kruskal-wallis, p =0.386), nor hearing threshold (of available measures, excluding "deaf" patients) (kruskal-wallis, p =0.2822). Paired comparisons using rank sum tests also showed no significant difference between any two participant groups. This is true even though there exists variability of patient data and differences in the mean values between the three CIL groups. Auditory performance measures also show no difference using kruskal wallis - disyllabic word recog (kruskal-wallis, p =0.144), nor sentence in noise (kruskal-wallis, p =0.388). However there are difference between EXP and NEW for word recognition (rank sum, p =0.0154, corrected value = 0.0462). So whilst NEW have worse mean scores than both INT and EXP, this is only significant for the word recognition in comparison to EXP.

Group ID	N	Defined by duration of implantation	Duration of implantation (months)	Age (years)	Duration of deafness (years)	Threshold (dB)	Word recog. (%)	Phrase in silence (%)
EXP	16	>14	57 ± 34	56 ± 15	4.6 ± 3.7	90 ± 17	80 ± 12	83 ± 22
INT	15	$7 - 14$	10 ± 2	45 ± 22	6.3 ± 7.6	86 ± 12	75 ± 23	86 ± 21
NEW	17	$0 - 6$	3 ± 1	60 ± 14	15.7 ± 17.5	67 ± 42	56 ± 28	63 ± 36

Table 3.2: Summary of etiological data and auditory performance for all CIL. Values given are the mean plus/minus the standard deviation. Groups were chosen based on the duration of implantation, indicated in the third column. The number of participants in each group is given by *N*. Subsequent data is given as the mean and standard deviation for each group of CIL with values for all participants detailed in tables 3.3, 3.4 & 3.5.

Participant ID	Age at testing	Duration of Deafness (years)	Duration of Implantation (months)	Implanted ear	CI model	Coding Strategy	Stimulation speed	Maximum Stimulation	Hearing threshold - non-implanted ear (dB)	Disyllabic word recognition - CI only (% correct)	Sentence in noise - SNR +10dB (% correct)
ALAREN	73	5	62	R	Cochlear CI24RE	ACE	900	11	72.5	80	96
ANDJOS	53	9	53	R	Cochlear CI24RECA	ACE	1200	10	105	80	95
ANSCRI	48	5	35	R	Cochlear CI24CA	ACE	720	9	80	85	99
AUZAND	54	2	34	R	Cochlear Nucleus CI512	ACE	900	10	70	95	73
BERCAT	51	2	62	R	Cochlear Hybrid	ACE	900	8	50	60	64
CHAJEA	72	1	83	L	AB Hires90K CI	Hi-Res S Fidelity 120	NA	NA	93	80	97
DAMCRI	51	2	36	R	MED EL SONATA	FSP	NA	NA	deaf	70	83
ESCPAT	48	3	114	R	Cochlear Nucleus CI24	ACE	720	9	100	80	95
GERJEA	52	8	81	L	Cochlear Nucleus Freedom CI24RECA	ACE 7209	720	9	91	75	85
GIDGEN	63	1	23	R	Cochlear24RECA	ACE	720	10	96	90	89
LAMCEC	63	0.5	17	L	Cochlear C512	ACE	900	10	90	95	95
LERMAR	15	14	126	R	Cochlear CI24RCS	ACE	900	16	-	-	-
MAMSOL	66	5	37	L	AB Hires90K	Hi-Res S	NA	NA	110	75	75
MATARL	76	2	14	R	Cochlear Nucleus C512	ACE	900	9	110	80	81
RAUCLA	70	2	82	L	AB Hires90K	Hi-Res S	NA	NA	80	55	13
SEBMAR	47	8	83	R	Cochlear Nucleus CI24	ACE	900	11	100	100	100

Table 3.3: Individual participant data for Experienced cochlear implant listeners (EXP)who had a duration of implantation greater than 14 months. Participant ID is used on figure 3.4 to identify individual participants. For some participants it was also not possible to obtain certain data, for example the auditory performance of LERMAR. Maximum stimulation is also a variable used for Cochlear branded implants and is therefore not applicable (NA) to the other brands tested.

Participant ID	Age at testing	Duration of Deafness (years)	Duration of Implantation (months)	Implanted ear	CI model	Coding Strategy	Stimulation speed	Maximum Stimulation	Hearing threshold (non-implanted ear (dB)	Disyllabic word recognition - CI only (% correct)	Sentence in noise - SNR +10dB (% correct)
ARNPAT	67	1.5	7.3	R	CI24 RE contour advance	ACE	900	10.0	85	100	100
BOIBEN	35	7	10.6	L	CI422	ACE	900	10.0	82.5	70	93
DESANN	77	5	11.7	R	CI422	ACE	900	10.0	91	100	100
ECLMAT	19	1.5	10.3	R	CI422	ACE	500	12.0	deaf	50	43
ERNMEL	17	3	10.3	L	CI422	ACE	900	9.0	deaf	90	100
FONSEB	36	4	10.1	R	CI24 RE contour advance	ACE	900	10.0	100	75	-
HALVAN	34	10	13.3	L & R		HiRes Optima S	CR3712	NA		84	100
LACHEN	79	1	12.9	L	CI422	ACE	900	10.0	91	80	95
LAIVAL	21	7	7.5	L	CI422	ACE	720	10.0	72	60	87
LOISYL	52	9	8.9	L	CI422	ACE	500	9.0	89	25	56
MILMAR	65	1	8.7	R	CI24 RE contour advance	ACE	900	10.0	76	90	100
PONOLI	32	31	10.3	R	CI422	ACE	720	11.0	91	95	100
SOUPHI	53	3	6.5	R	CI422	ACE	720	11.0	88	85	97
SPECHR	68	9	7.1	R	CI24 RE contour advance	ACE	500	10.0	58	-	-
VANERI	20	1	6.9	R	CI24 RE contour advance	ACE	900	9.0	104	45	51

Table 3.4: Individual participant data for Intermediate cochlear implant listeners (INT) who had a duration of implantation between 6 and 14 months. Participant ID is used on figure 3.4 to identify individual participants. For some participants it was also not possible to obtain certain data, for example the auditory performance of SPECHR. Maximum stimulation is also a variable used for Cochlear branded implants and is therefore not applicable (NA) to the other brands tested.

Participant ID	Age at testing	Duration of Deafness (years)	Duration of Implantation (months)	Implanted ear	CI model	Coding Strategy	Stimulation speed	Maximum Stimulation	Hearing threshold (non-implanted ear (dB)	Disyillabic word recognition - CI only (% correct)	Sentence in noise - SNR +10dB (% correct)
ALBJOE	59	0.42 & 1	4.7	L & R	CI24 RE Contour advance	ACE	900	9	0	90	100
ARQDAN	70	27	2.1	R	CI24 RE Contour advance	ACE	900	9	90	85	100
BAIIVE	69	54	1.9	R	CI24 RE Contour advance	ACE	900	11	83	45	51
BOSPIE	73	4	0.4	R	CI422	ACE	900	11	0	95	-
BUIJEA	87	27	2.1	L	HR90k Advantage / HiFocus 1J	HiRes Optima-S	3712 pps	NA	96	35	6
CAZJOS	66	1	3.2	L	CI24 R (CS) nucleus24 contour	ACE	720.0	10.0	0	60	76
CERPAS	45	9	1.7	L	CI24 RE Contour advance	ACE	250	9	95	40	50
COTJEA	52	8	1.4	L	CI512	ACE	900	9	105	90	97
DASYVE	59	10	2.4	L	CI24RE Contour advance	ACE	900	10	57	80	98
GARSYL	44	42.5	1.9	R	CI24 RE Contour advance	ACE	250	9	101	35	10
GERJAC	62	0.92	1.9	L	CI512		900	11	98	55	87
GIMETI	27	26	3.1	R	CI422	ACE	500	10	76	20	22
HERMAR	62	0.3	3.1	L	CI512	ACE	250	9	0	-	-
MAZSYL	60	0.3	3.6	R	CI422	ACE	720	9	113	70	98
NICCHR	55	2	3.0	L	CI24 RE Contour advance	ACE	900	10	96	50	81
PARNOE	52	2	0.0	L	CI24 RE Contour advance	ACE	720	9	35	40	49
SANNIC	73	30	0.5	L	CI24 RE Contour advance	ACE	500	9	95	0	18

Table 3.5: Individual participant data for New cochlear implant listeners(NEW) who had a duration of implantation less than 6 months. Participant ID is used on figure 3.4 to identify individual participants. For some participants it was also not possible to obtain certain data, for example the auditory performance of HERMAR. Maximum stimulation is also a variable used for Cochlear branded implants and is therefore not applicable (NA) to the other brands tested.

3.3 RESULTS

3.3.1 *Global performance*

The TCL Lab-X software used to carry out the FST also recorded certain data about how the participants completed the task and this is shown for each CIL group in figure 3.1. “Playbacks per participant” is the average number of times that participants listened to all sounds, whilst this number is divided by the number of sounds (16) to calculate the ”Playbacks per participant per sound”. For example, on average each member of EXP participants listened to the sounds 86 times during a test, and each sound was listened to 5.4 times.

Statistical analysis (kruskal-wallis) showed no difference in the number of categories made by each participant group, nor the number of playbacks ($p > 0.05$). There was however a difference in the duration of test, such that NEW were greater than both EXP and INT ($p = 0.0192$). This could be evidence that NEW participants find the task more difficult and which might have been expected knowing that this participant group have the lowest amount of hearing experience with their implants. It also noteworthy that there are no significant differences between the groups of EXP and INT who show similar performance. ANOVA analysis was also performed on the individual participant data and showed that for the participants within each group there was no difference, such that each participant group was uniform in performance.

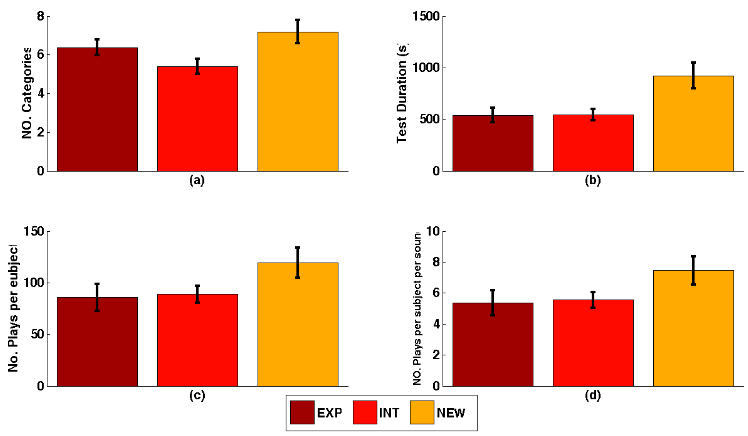


Figure 3.1: FST performance statistics. Statistics for the performance of CIL in performing the FST. Bars represent the average value with error bars showing the standard error. *Playbacks per participant per sound* is taken as an average of the number of playbacks across both stimuli and participants.

3.3.2 *HCPC analysis*

The simplest representation of the categorisation strategies employed by the participant groups is represented by the dendrograms in figure 3.2 where the different coloured rectangles represent the final categories. Importantly this is only a representation based on the categories created by all participants, such that some participants may only weakly follow this representation and some not at all. As a measure of how well the trees reflect the original data (indicator matrix) CCC values are located in table 3.6 and show that the least representative dendrogram is that created for NEW (lower panel of figure 3.2).

The height (or distance) indicated on the dendrograms (figure 3.2) is a representation of how perceptually similar the sounds are deemed to be - with a greater distance indicating that two sounds (or categories) are very dissimilar. The *intra* category distance is calculated as the average distance between sounds within each category (rectangle) and *inter*(or between) category distance the average distance between each category. Distances are similar across all participant groups (kruskal-wallis, $p > 0.05$) and therefore do not show any group to be particular more uniform or varied concerning the categorisation strategies used.

Whilst the same number of categories (six) are seen on the dendrograms for all participant groups, this is somewhat dependent on the number of dimensions retained on the MCA analysis and does not suggest evidence for similarities or differences amongst the participant groups.

Participant Group	INTRA cat. dist.	INTER cat. dist.	CCC
EXP	0.047	0.59	0.80
INT	0.041	0.55	0.83
NEW	0.048	0.56	0.69

Table 3.6: Statistical measures of the dendrogram structures. *INTER cat. dist.* (category distance) calculated as the average height between items located within the same category, whilst *INTRA cat. dist.* is calculated as the mean distance between the specific categories/clusters as noted by the coloured rectangles. Cophenetic Correlation Coefficient (CCC) is used to calculate how accurately the dendrogram represents the original data (the indicator matrix).

The dendrograms show that HCPC analysis produces a category of vocal sounds containing *male*, *fem* & *lgh* for all three participant groups. However whilst this is seen to be the first category for EXP and INT groups this is not the case for NEW and may highlight a difference in categorisation strategy, where the more experienced CIL are able to separate the vocal sounds more distinctly. Another similarity shared by the two experienced CIL groups is that when cutting the dendrograms to produce only three large categories the results correspond to the three predefined categories of musical, vocal and environmental sounds, the only anomalies countering this being:

- for EXP *alrm* alongside musical sounds and *cgh* amongst environmental sounds.
- for INT *gls* and *cgh* amongst the musical sounds.

The results for NEW also correspond to the predefined categories, however only when cutting the dendrogram to create four categories, such that the sounds *dr* and *fstp* are distinct from other environmental sounds.

Overall the global performance statistics and HCPC analysis show that the three CIL groups performed very similarly in the FST, with only the duration of testing and Cophenetic Correlation Coefficient (CCC) being significantly different for NEW. Results are however more similar for EXP and INT whom show a stronger tendency to separate the vocal sounds. Finally the sounds *dr* and *fstp* may be perceived differently by NEW or simply form part of a different categorisation strategy.

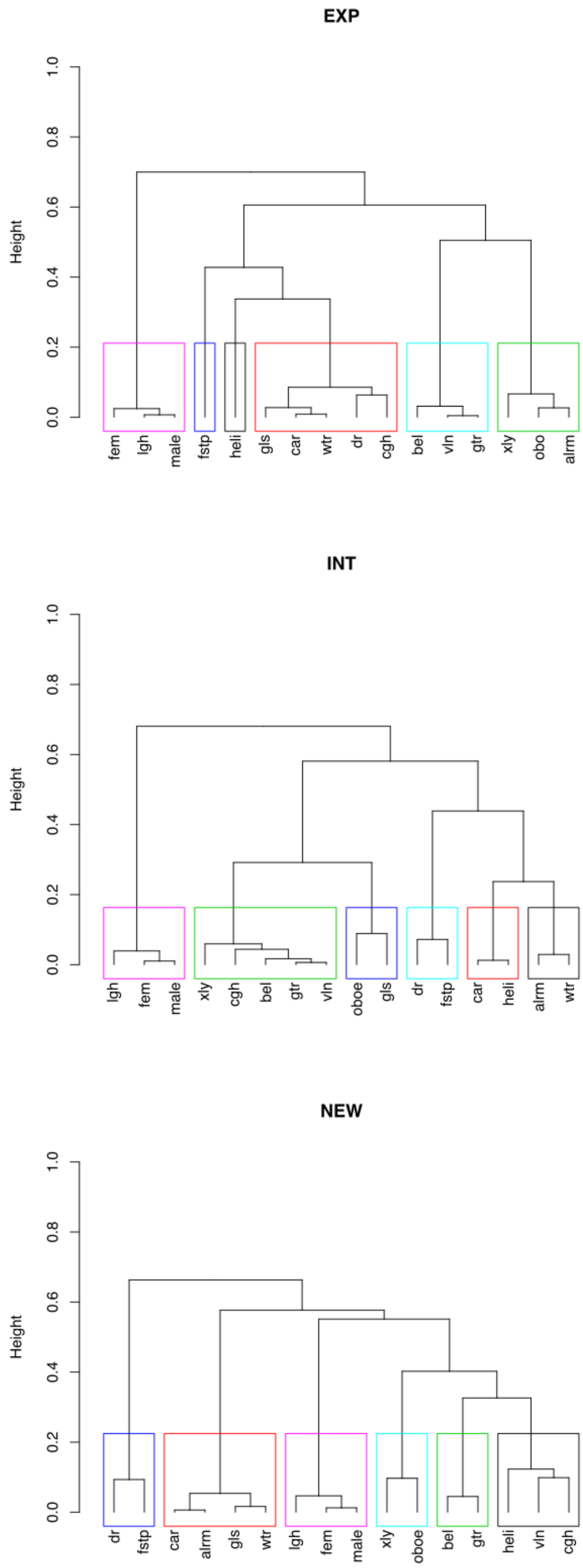


Figure 3.2: Dendrogram for EXP, INT and NEW participants - the overall categories outputted by HCPC are indicated by coloured rectangles whilst the upper limit of these rectangles indicates the point at the dendrogram has been cut (as described in the main text). The height axis gives the perceptual distance between each stimulus whereby a large height indicates that participants deemed those two stimuli to be highly dissimilar and vice versa. Finally stimuli are labelled using the abbreviated sound IDs from Table 3.1.

3.3.3 *MCA analysis*

A more detailed breakdown of the categorisation performance can be given with MCA analysis. In this analysis the original data is represented in an n-dimension euclidean space. Interpreting the dimensions can then give insights into the categorisation strategy of participants, although importantly the generated dimensions are only an output of the analysis and are not directly related to any physical measure of the sounds or the participants.

Whilst the analysis outputs a total of 15 dimensions only those which cover above 8% of the total variance have been retained for interpretation and commentary. This has given 5 dimensions for each participant group, and the variance accounted for by each is shown in panel a. of figure 3.3. Dimensions are ordered in descending order such that D1 covers more variance than than D5. There is no difference between the participant groups for any of the dimensions(kruskal-wallis, $p > 0.05$) nor the unaccounted variance - although this slightly lower for the INT group (40.7 %).

Eigen values shown in panel b. of figure 3.3 are caluculated as the average coordinate from subject maps (figure 3.4) and again show no difference between the participant groups (kruskal-wallis, $p > 0.05$). Alongside the number of participants who have a coordinate greater than 0.8 the eigen values give an idea of how strongly a particular dimension is used. High values would indicate that a large amount of participants are strongly adhering to a particular dimension and strongly using the corresponding categorisation strategy. As can be seen the values are highest for the first two dimension (D1 & D2) whilst the latter dimensions (D3-D5) show lower values, especially in the % participants above 0.8. This is again visualised in figure 3.4 where the majority of participants are located in the top right corner of each panel and demonstrates the general agreement of participants in using D1 & D2. The pattern is similar across all CIL and along with similar values from figure 3.3 suggests that there is a similar amount of variance within the performance of each CIL group.

It should be noted that the dendrograms of figures 3.2, were created using the 5 retained dimensions. Had the full 15 been used more variance would have been included and the resulting dendrograms would be different, with more categories and many sounds individually separated rather than grouped together.

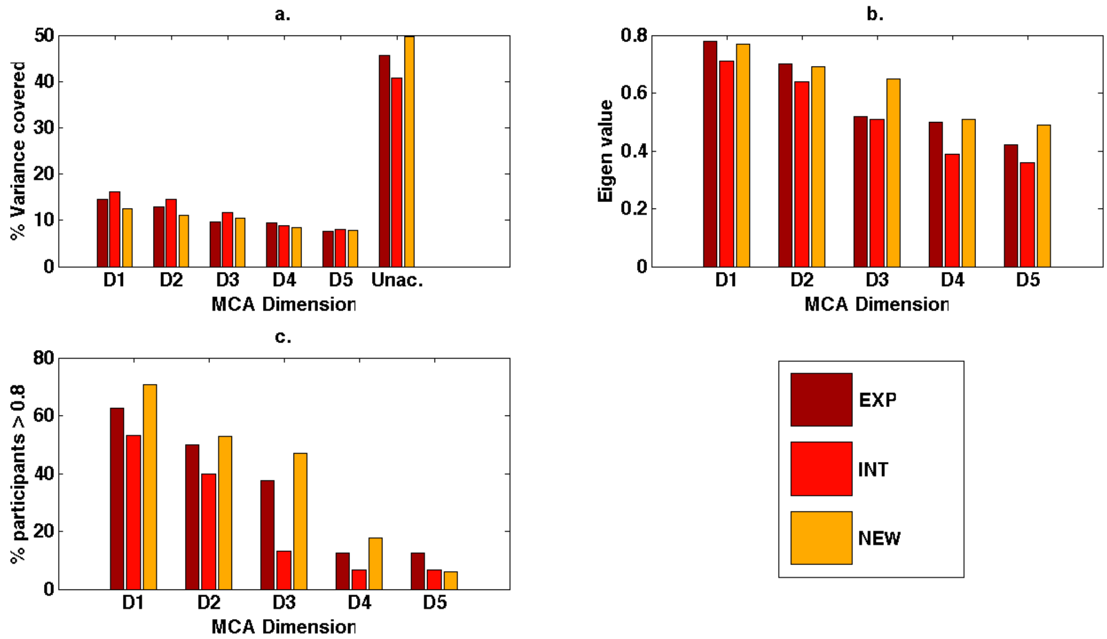


Figure 3.3: The retained dimensions as outputted by MCA analysis for each CIL group. Dimensions covering 8% or more of variance are retained, leading to 5 dimensions for all three CIL groups.The individual and cumulative variance covered for each dimension are given, with the remaining unaccounted variance relating to the non-retained dimensions. *Mean eigen* value is given where eigen values represent the average eigen value of participants for a given dimension whilst % *participants* >0.8 refers to the percentage of participants who had individual eigen values greater than 0.8. Both values which can be seen in figure 3.4.

initial MCA statistics would suggest further similarities amongst the three CIL participant groups.

Overall based on the figures 3.3 & 3.4 there are similarities in the results of the categorisation performance between the three groups of CIL. Certainly, and against expectations, it is NOT the case that the results for EXP show more agreement between participants in comparison to either INT or NEW. Results also suggest that within each participant group there are likely 2-3 aspects of the categorisation strategy that are commonly shared followed by sub-strategies that used by a smaller number of participants. This does not however explain how each dimension may be interpreted, to better understand this, figures 3.5 to 3.8 display the layout of the sounds across the five dimensions.

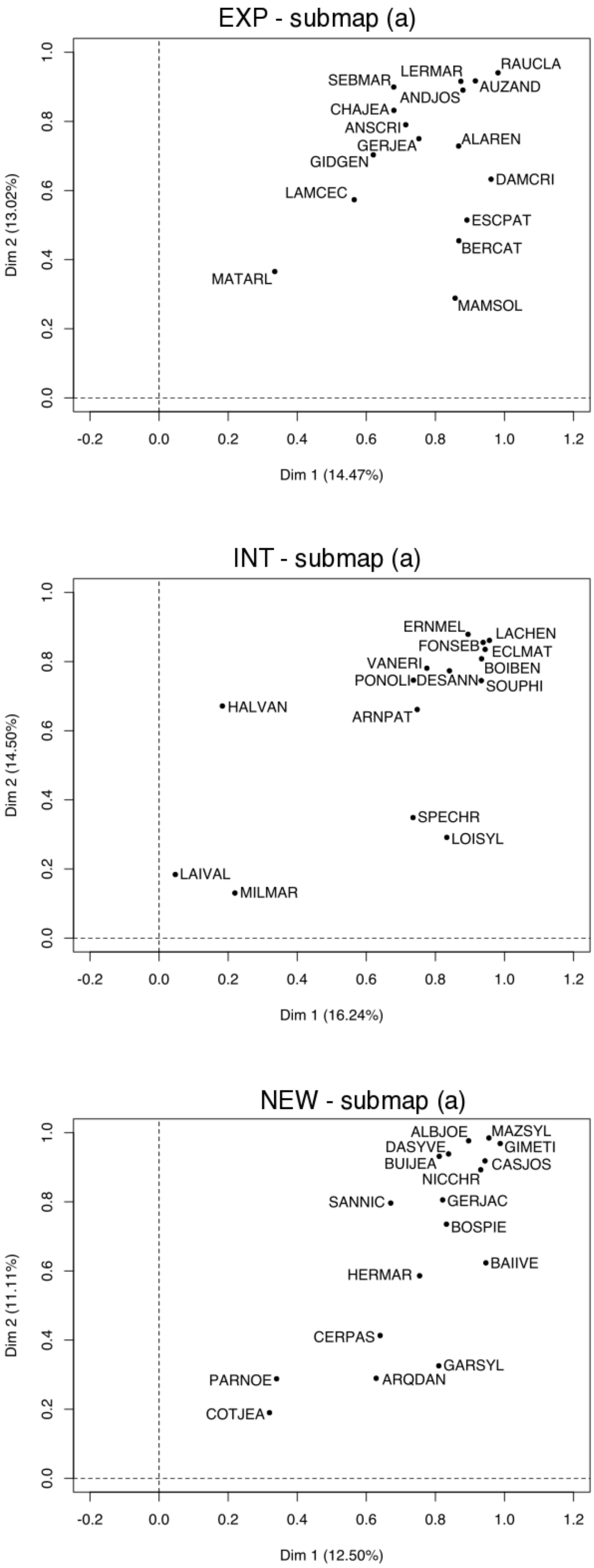


Figure 3.4: Figures showing the distribution of all 3 CIL groups across Dim 1 & 2. Results are plotted from $0-1$ and can be interpreted as showing how strongly the results of each participant are matched to each stated dimension, where a value of 1 would show a strong match. The mean value of each dimension is also represented by the Eigenvalues in table **??**. Finally Dim 1 & 2 are plotted in the top panel, Dim 3 & 4 in the middle panel and Dim 1& Dim 5 in the bottom. Participants are plotted using the ID labels from table 3.3.

Experienced cochlear implant listeners (EXP)

Focussing first on the results for the EXP group (figure 3.5) it would appear that the first two dimensions divide the sounds into three clusters corresponding to vocal, environmental and musical sounds - much like the predefined categories (see table 3.1). Dim 1 would appear to strongly separate the vocal sounds, although interestingly the sound *cgh* is not located with other vocal sounds. Dim 2 is seen to separate the musical sounds from the other groups of vocal and environmental sounds, with the sounds *vln* and *gtr* particularly distinct. This could be a result of these sounds containing multiple frequencies (musical notes) and a sense of melody in comparison to the sounds of *xly* and *obo* which are only single tones. Finally the sound *alrm* is located close to the zero-line of Dim 2, between the musical and environmental clusters which could indicate an ambiguity in the perception of this sound. Participants use of Dim 1 & 2 can be seen in figure 3.4 and shows that whilst there may be only 4 participants who are above 0.8 and therefore strongly using both dimensions, there are only four participants who are outside the limits of 0.6, (*ESCPAT, BERCAT, MAMSOL and MATARL*). Of these, three remain strong users of Dim 1 such that only participant MATARL would seem to be performing, or have a categorisation strategy different from the other participants.

Dim 4 separates the group of *xly*, *obo* and *alrm* from other sounds. These sounds were commonly identified as *telephone ringing*, *door-bell* or simply *ringing sounds* and they are also sounds that constitute sustained single pitches. Therefore the grouping together of these three sound on Dim 4 may be based on similar acoustic information or semantic information - these sounds can be described as "warning" or "'alarming" sounds. However only 2 participants strongly adhere to Dim 4 meaning that it may not be of high importance to the overall categorisation strategy. Finally Dim 3 and Dim 5 are concerned with separating individual sounds (*heli* and *fstp* respectively) from the others.

Overall the results for EXP show that the majority of participants create categories corresponding to the pre-defined categories of vocal, environmental and musical sounds. Categories of human action and warning/alarming sounds may also be important in the overall categorisation strategy used by EXP although not all the participants show this.

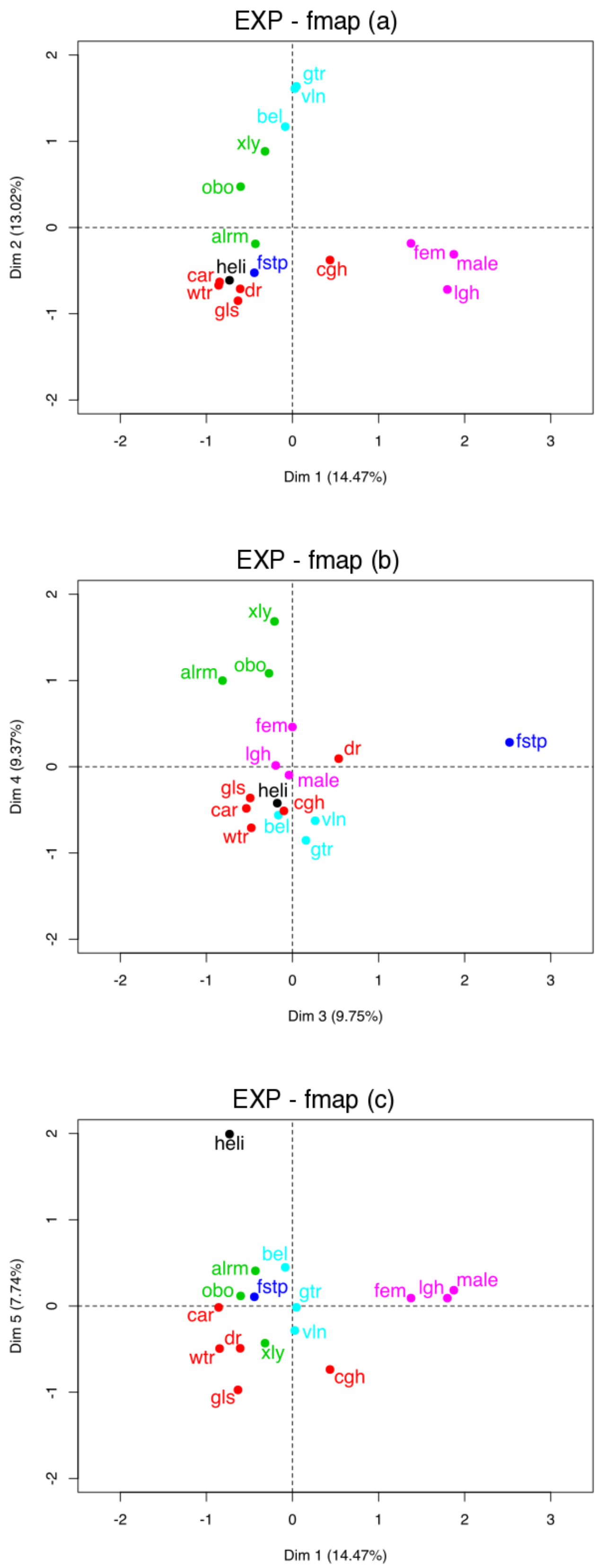


Figure 3.5: Figures showing the distribution of SOUND STIMULI across the 5 retained dimensions, with Dim 1 & 2 in the top panel, Dim 3 & 4 in the middle panel and Dim 1 & Dim 5 in the bottom. Stimuli are plotted using the ID labels from table 3.1 whilst colours are the same as used for the dendrograms in figure 3.2 in order to show the link between the MCA and HCPC analyses.

Intermediate cochlear implant listeners (INT)

The factor maps for the INT participant group (figure 3.6) show similarities to that of EXP. Noticeably the first two dimensions again create three clusters of environmental, musical and vocal sounds. Dim 1 contrasts vocal sounds from other stimuli, although the sound *cgh* appears further separated from vocal sounds, possibly indicative of a difference in perception of the *cgh* sound between the INT and EXP participants.

Again in similarity to EXP results, Dim 2 separates the musical and environmental sounds. However the arrangement is different as the main contributors *fstp* and *dr* rather than *gtr*, *xly* and *vln* as for EXP. Comments in Appendix (???) show that the sounds of *fstp* and *dr* were often described as human action sounds and this is important to interpreting this dimensions and the categorisation strategy. Therefore although Dim 2 is performing the same function for both EXP and INT in separating musical and environmental sounds, it appears that the method of doing this is different.

Interestingly Dim 3 for INT is again similar to EXP as it separates *fstp* from the other sounds, therefore suggesting that human action sounds may be a robust categorisation strategy. However, data from table **??** shows that fewer INT participants are using this strategy. Dim 4 is also comparable with EXP as it it shows the grouping of *heli*and *car* is repeating the categorisation of mechanical/transport sounds. Finally the sounds which contribute to Dim 5 are *wtr*, *xly*, *gtr* and *obo*. This does not correspond to any of the dimensions for EXP and it is not easily interpreted from only the factor map. It could therefore simply be a result of the MCA analysis and not correspond to a "real" categorisation strategy.

Results for INT show strong similarity to those of EXP, especially for Dim 1 & 2 which again separate the stimuli into three rough groups of environmental, musical and vocal sounds. There is again evidence for human action sounds being important as well as a sub-category of transport/mechanical sounds. There is however slightly less subject agreement when comparing the eigenvalues and % of participants above 0.8 which may point towards more variation amongst INT.

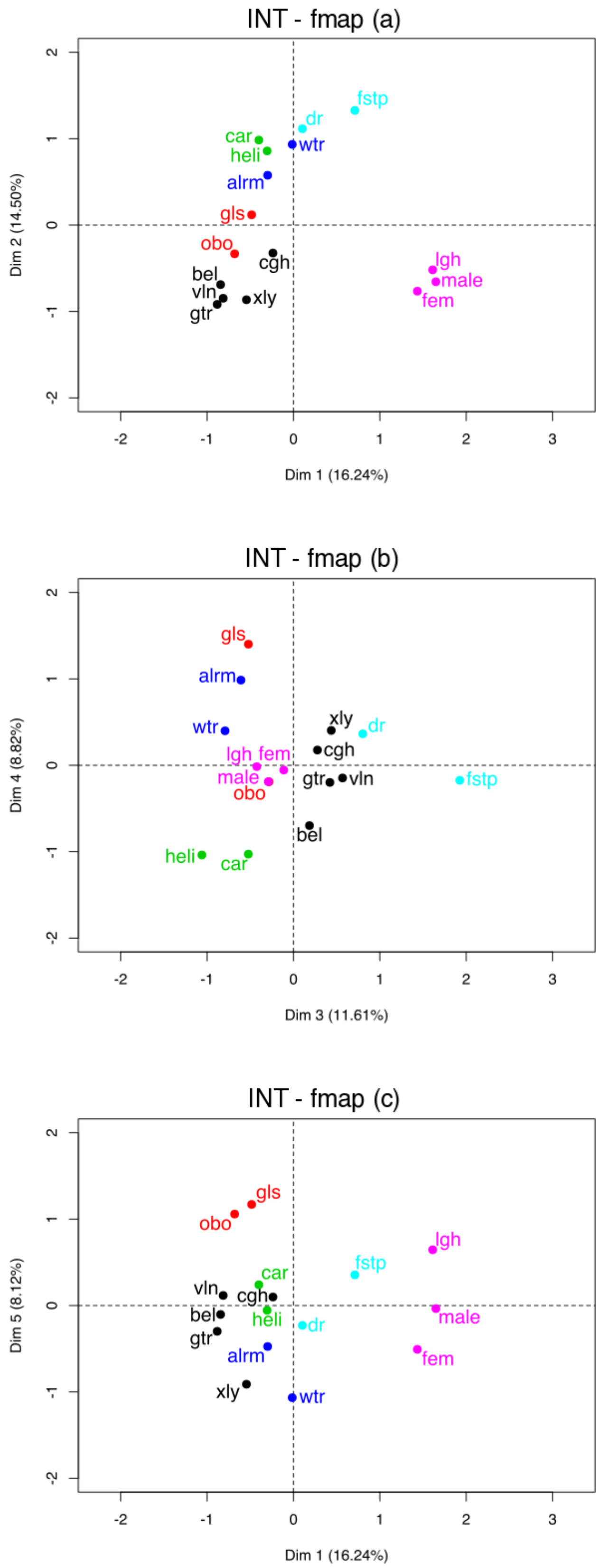


Figure 3.6: Figures showing the distribution of stimuli across the 5 retained dimensions, with Dim 1 & 2 in the top panel, Dim 3 & 4 in the middle panel and Dim 1 & Dim 5 in the bottom. Values of variance located in table **??** are also displayed for each dimension. Stimuli are plotted using the ID labels from table 3.1 whilst colours are the same as used for the dendrograms in figure 3.2 in order to show the link between the MCA and HCPC analyses.

New cochlear implant listeners (NEW)

MCA results for NEW, figure 3.8, immediately stand out as being different to both EXP and INT. Most notably it is not possible to identify three clusters of environmental, musical and vocal sounds using only Dim 1 & Dim 2. Instead Dim 1 separates environmental sounds, to the right, from the musical and vocal sounds and Dim 2 separates the sounds of *fstp* and *dr* from the other environmental sounds. Although *dr* and *fstp* are not grouped together in exactly the same way for EXP and INT there could be a link between the Dim 2 for NEW and the strategy of grouping together human action sounds as seen for EXP and INT.

Dim 3 may also show some similarity to the Dim 1 of EXP and INT in that it involves the distinction of the sound *male* and could therefore relate to a separation of vocal sounds. Although from figure 3.8 panel b the other vocal sounds are not in close proximity to the sound *male* and therefore the categorisation of vocal sounds may be more weakly employed by NEW. Also the *cgh* sound is again not closely grouped with other vocal sounds, suggesting that the perception of this sound amongst NEW is similar. Dim 4 separates only the sound *obo* and again demonstrates latter dimensions involve only individual sounds rather than groups or categories. Finally the extremes of Dim 5 show the sounds *cgh* & *heli* versus *fspt* & *bel*. This is however not particularly informative and Dim 5 may again be simply a result of the MCA analysis.

The results of MCA analysis for NEW are initially quite different compared to EXP and INT. Most striking is that the predominant categorisation strategy involving vocal stimuli is only weakly represented by Dim 3. Rather the dominant categorisation strategies involve environmental sounds and human action sounds.

3.3.4 *Combined Analysis*

Results of all three CIL participant groups were also combined in order to assess in more detail if there effect was of the duration of implantation. Previous analysis had shown that the duration of implantation was correlated with Dim 1 & 2 for NEW, which were interpreted as separating of environmental sounds from other stimuli. A tendency for this correlation was seen for EXP with regards to

Dim 1 - the separation of voice vs. other sounds.

Combining all three CIL participant groups and performing the same HCPC and MCA analysis gives rise to the results shown in figure 3.9. IN the middle panel, the factor map again shows 3 broad groupings of sounds which correspond to the vocal (far left), environmental (to the right) and musical sounds (below). Whilst the exact pattern of results is different to the separate factor maps, it is not surprising that both dimensions can be interpreted as previously. In that Dim 1 corresponds to voice vs. non-voice sounds and Dim 2 the separation of musical vs non-musical sounds. The overall categorisation described by the middle and upper panels of figure 3.9 is however somewhat different to that previously described in separate results and actually acts as a method of concluding the most salient features found in the results so far. I.e. the separation of vocal sounds; the pairing of *fstp* and *dr* which are separated from other environmental sounds; the categorising musical sounds into melodic and single-note stimuli (*xly* & *obo*); the weaker categorisation of *obo* as a musical sound; finally the very weak categorising of *cgh* as a vocal sounds, which is instead grouped alongside musical sounds.

The lower panel of figure 3.9 is very important in understanding the effect of implantation duration (level of experience). An initial observation is that the coordinates for Dim 1 are correlated with those of Dim 2, ($r = 0.76, p < 0.001$) which shows that those participants strongly using the categorisation strategy described by Dim 1 are also using that of Dim 2. Correlations between the the retained MCA dimensions and implantation duration, do not show any significance for the combined results. Taking subsets of participants within the combined results there is a positive correlation between implantation duration and Dim 1 for both EXP ($r = 0.53p < 0.05$) and NEW ($r = 0.63, p < 0.05$), although not for INT participants.

This could therefore suggest that CIL improve agreement to vocal vs environmental sound discrimination (as categories) in a short space of time, 0-6 months for NEW, and also over a longer period of implantation 14 - 126 months for EXP. So for these two groups as duration of implantation increases, so does the adherence to Dim 1 of figure 3.9. This would also fit with research that shows improvements in auditory abilities over first 6 months of implantation, before plateauing between 6-14 months and then further improvements with significant implantation duration and listening experience. It is not clear to which auditory abilities this may be linked, it could be speech processing, environmental sound perception abilities, or also with familiarity and experience of auditory stimuli.

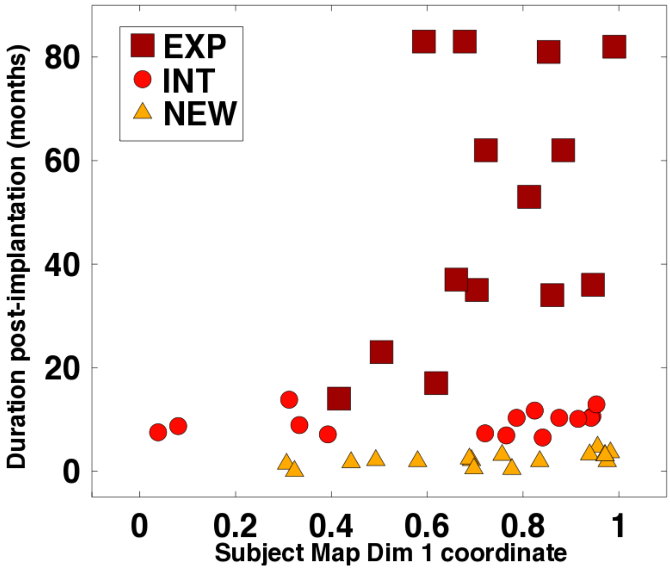


Figure 3.7: Scatter plot of CI user Implantation duration vs. MCA Dim 1 for the combined analysis on ALL CIL participants. Separate groups of CI users are shown by ■ for EXP, INT by • and NEW by ▲. Duration of implantation in months is plotted on the Y-axis and Dim 1 coordinate from the subject map is plotted on the X-axis. Dim 1 coordinates are taken from the lower panel of figure 3.9.

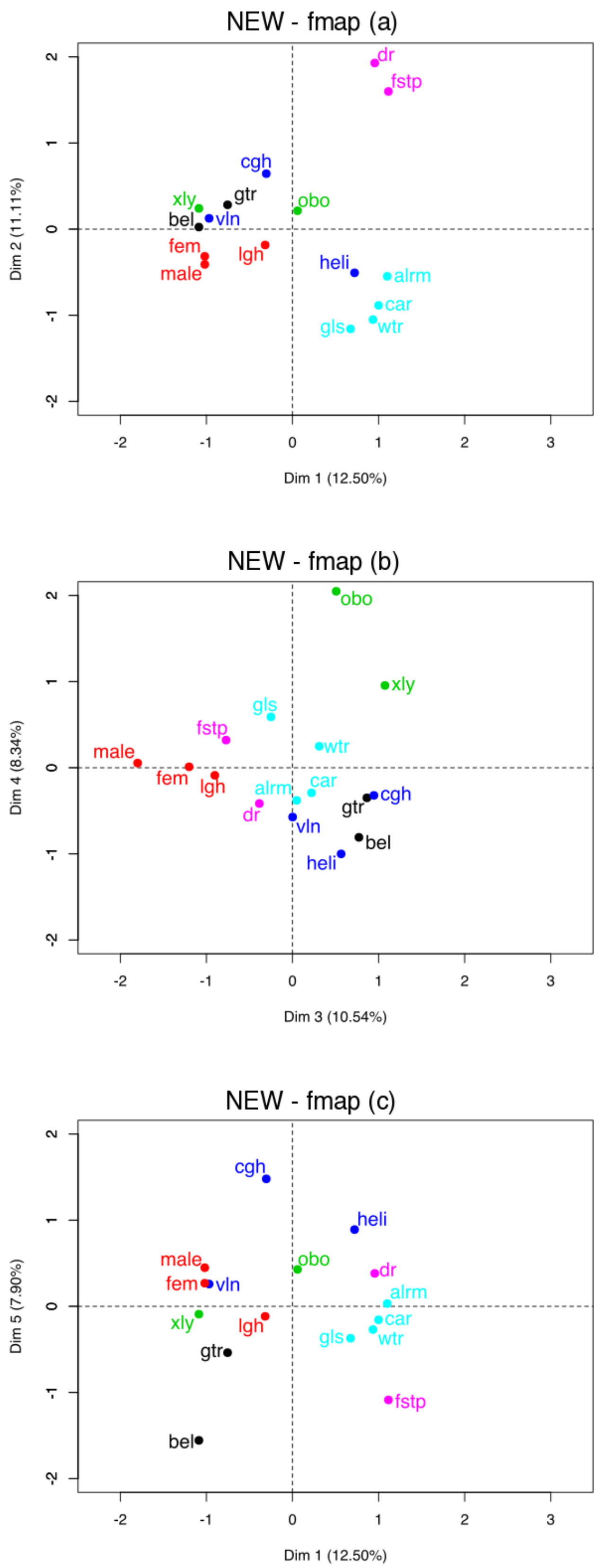


Figure 3.8: Figures showing the distribution of stimuli across the 5 retained dimensions, with Dim 1 & 2 in the top panel, Dim 3 & 4 in the middle panel and Dim 1 & Dim 5 in the bottom. Values of variance located in table **??** are also displayed for each dimension. Stimuli are plotted using the ID labels from table 3.1 whilst colours are the same as used for the dendrograms in figure 3.2 in order to show the link between the MCA and HCPC analyses.

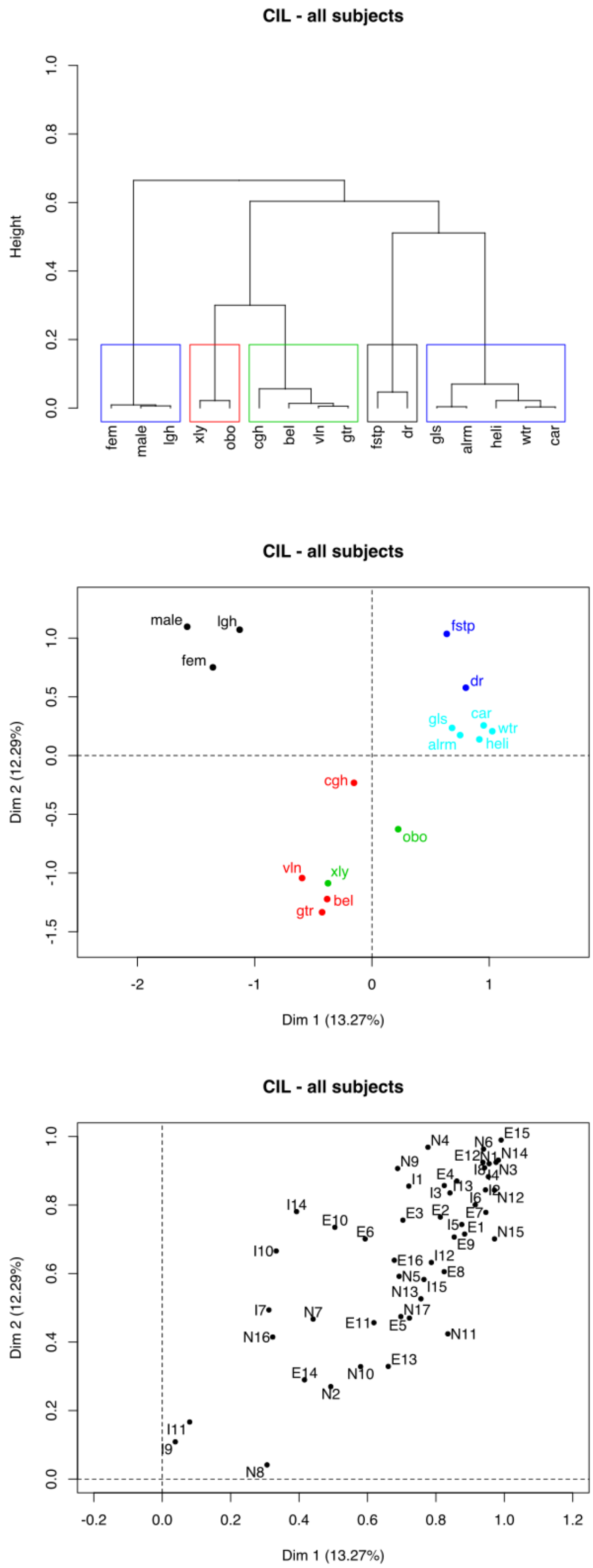


Figure 3.9: Categorisation results for ALL combined CIL. In the participant map, lower panel, E = Experienced cochlear implant listeners, I = INT and N = New cochlear implant listeners. HCPC dendrogram is shown in the upper panel, MCA factor map covering Dimensions 1 & 2 in the middle panel and MCA subject map in the lower panel.

3.3.5 *Acoustic Analysis*

In order to further interpret the dimensions outputted by MCA coordinates of factor maps were correlated, using spearman rank, to acoustic measurements of the sounds. Acoustic variables were calculated in similarity to [49]. The significant correlations ($p < 0.05$) are given in table 3.7 for the three CIL groups. Correlations for EXP show that Dim 2 is positively correlated to the mean and standard deviation values of peaks in the autocorrelation function - which reflects the uniformity of repetitions within the sound. Dim 2 is also correlated to the mean salience - a measure of the strength of pitch within the sound. Figure 3.10 shows the values of pitch salience plotted against mean autocorrelation peak and demonstrates how these two cues may be used in order to categorise the musical sounds together. Along with the sounds BEL and ALRM they have high values of both mean pitch salience and mean autocorrelation peak. The pattern of autocorrelation function is similar for all musical sounds and is displayed for the sound OBO in the right panel of figure 3.10. As can be seen there is only one large peak at the beginning of the sound which could indicate that information at the beginning of the sound is important in the perception of these sounds.

In contrast to the results for EXP the results of INT and NEW show correlations to a wider selection of acoustic cues. Interestingly the two participant groups show similarities with regards to the max and mean waveform peaks (transient increases in amplitude of at least 80% the amplitude range). Both D2 for INT and D1 for NEW are correlated to these variables and both dimensions reflect a separation of environmental sounds vs. vocal and musical sounds (figures 3.6 & 3.8). Figure 3.11 plots the values of mean waveform peak against the MCA coordinates for INT and NEW and shows that in general environmental sounds have low mean peak values. VLN and ALRM sounds have also been plotted in order to visually demonstrate the differences in peak values and also the waveforms of different sounds. It is clear to see that VLN not only has higher peak values but the amplitude varies much more over time in comparison to ALRM. Most environmental sounds surrounding ALRM also share a similar flat waveform pattern. FSTP and DR are however more temporally distinct and similar to the waveform of VLN. This could therefore be a reason why they are better identified (figure 3.14) compared to other environmental sounds. The mean Burst value is also correlated with D3 for INT and further evidences the

Acoustic Variables	EXP	INT	NEW	
Frq max		D2	D1	Spectrally related
Frq mean			D3	
Centroid		D4	D1	
Pitch Salience max		D4		
Pitch Salience mean	D2			
Spectral std		D3	D1	
Spectral skew		D3		
BPF mean	D1	D2	D1	
Autocorr Peak mean	D2			Temporally related
Autocorr Peak std	D2			
Peak max		D2	D1	
Peak mean		D2	D1	
Burst mean		D3		
Duration ratio		D3		
Wav range		D2	D1	

Table 3.7: Table of acoustic correlations between coordinates of MCA dimensions and acoustic measurements of sound stimuli. For each CI user group (EXP, INT & NEW) significant correlations ($p < 0.05$) are shown for the corresponding acoustic variable for the dimensions, which are listed as D1 to D5.

possibility that distinct changes in the amplitude are used by CI users as part of their perception of sounds.

Max frequency values are also correlated to the same two dimensions. This is a result of high frequency values in musical and vocal sounds, and low values in environmental sounds and suggests that the ability to discriminate high and low frequencies is part of the categorisation strategies used by the two CI participant groups.

The only acoustic variable shared by all CI users groups is the mean BPF, which reflects the mean correlation between octave frequency bands of the sound envelope. A high value therefore corresponds to sounds that are more uniform in their envelope response, for example the sounds OBO and HELI. In contrast sounds that have low values are MALE and FEM, as can be seen in figure **??** which also illustrates the envelopes of HELI and MALE across the 6 octave band frequencies used to calculate the mean BPF. Mean BPF is negatively correlated to EXP-Dim 1 and NEW-Dim 1 and positively correlated to INT-Dim 2.

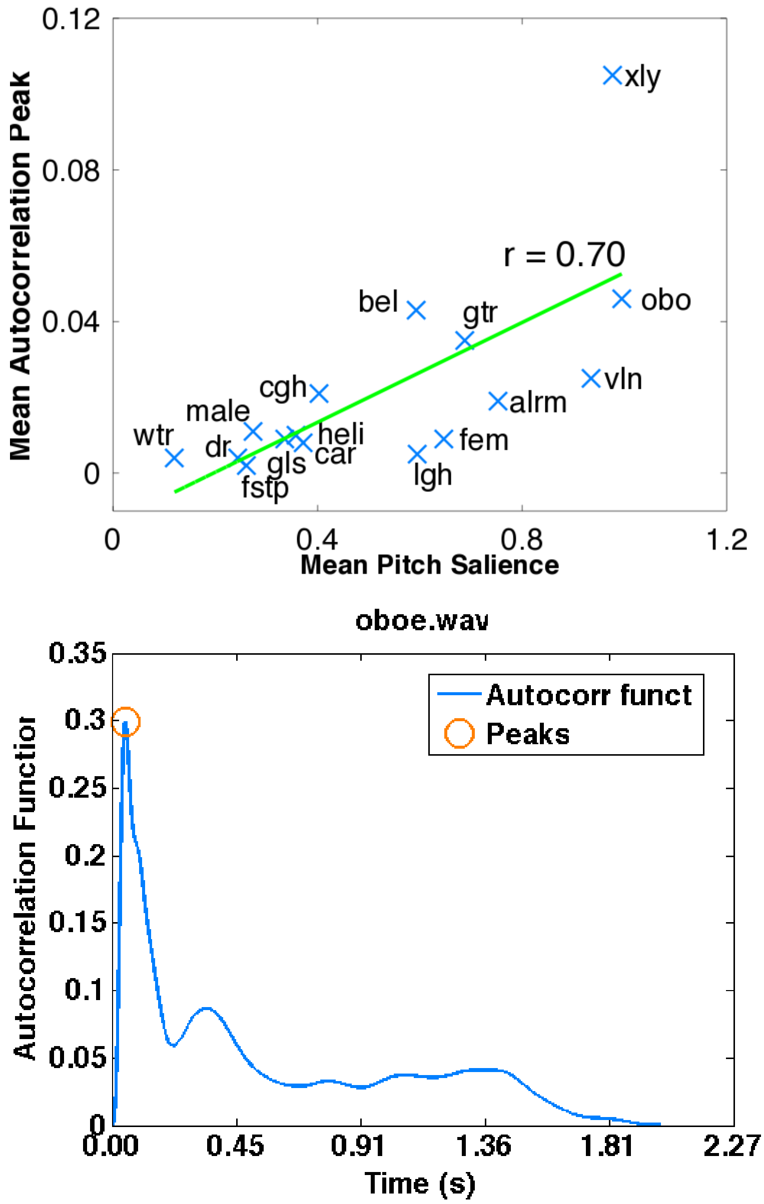


Figure 3.10: Values of Mean Pitch Salience vs. Mean Autocorrelation Peak for the 16 sound stimuli are plotted in the top panel. The plot shows how these two cues may be used to distinguish the musical sounds (obo, vln, gtr, xly) as they have higher values of mean autocorrelation peak and mean pitch salience. A significant correlation (r = 0.7) is also drawn with a line of best fit, in green. The lower panel shows an example autocorrelation function plotted for the sound *oboe*, with peaks highlighted by orange circles.

All three of these dimensions can be seen to contrast environmental from vocal sounds, see figures 3.5-a, 3.6b and 3.8-a.

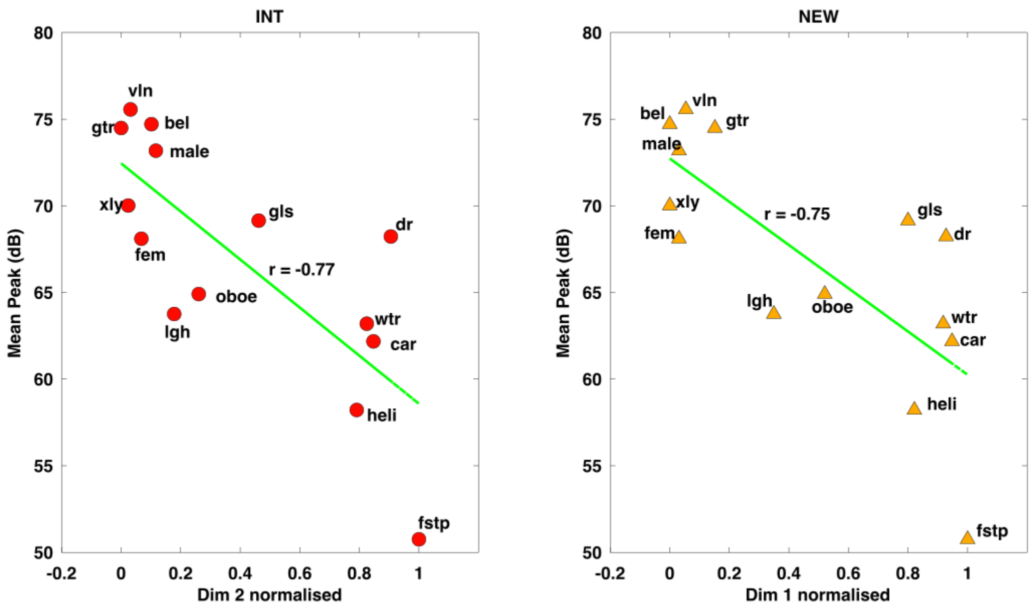


Figure 3.11: Correlations of Peak values in the waveform vs coordinates for MCA dimensions for INT and NEW. Left panel is for INT CI user and plots the correlation between the MCA Dim 2 (x axis) and mean peak (y-axis). Right panel shows the correlation between MCA Dim 1 and mean peak for NEW CI users. It should be noted that MCA dimension coordinates have been normalised from $0-1$ rather than the original data points shown in figures 3.6 & 3.8.

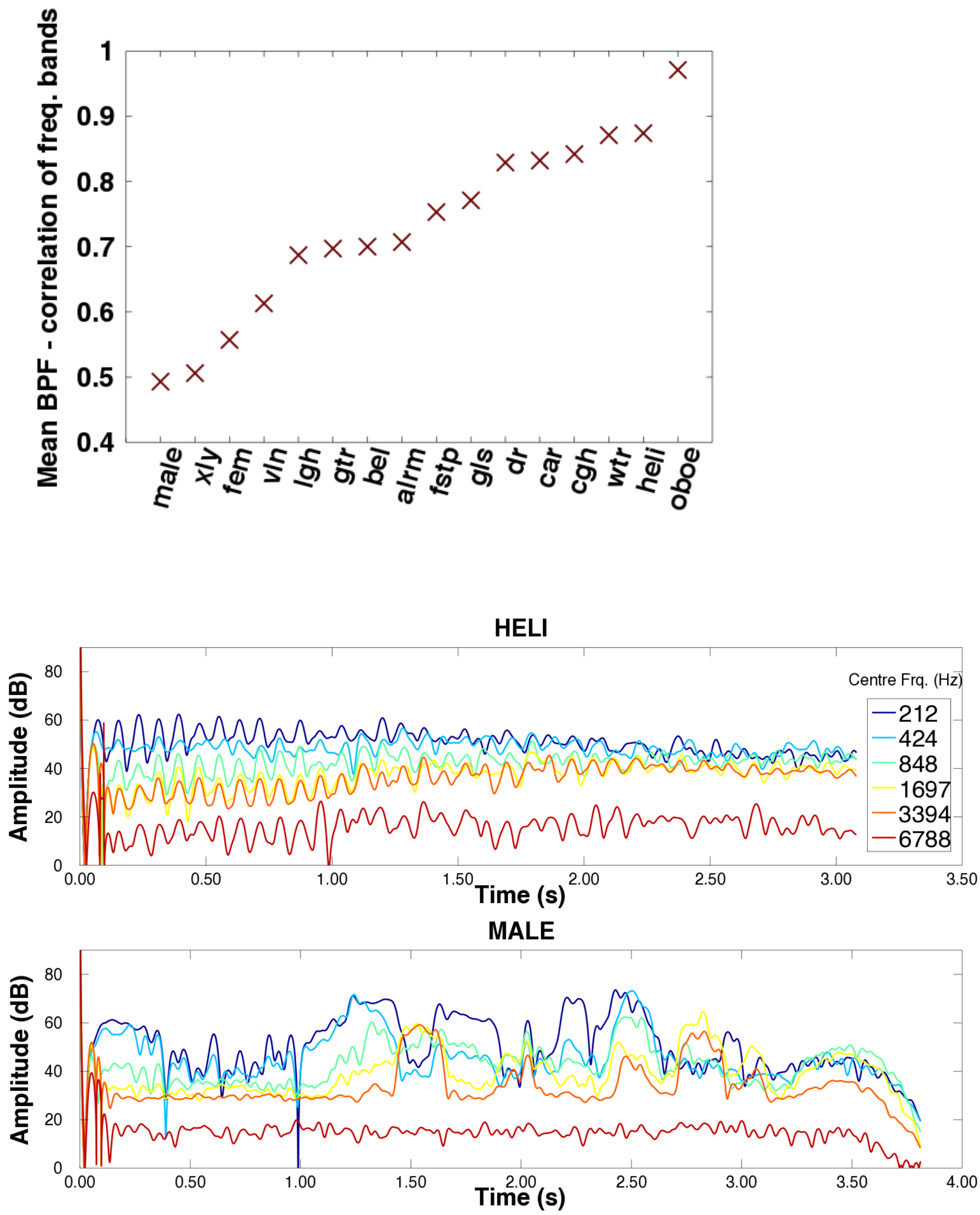


Figure 3.12: Mean BPF values for all sounds, also plotted for individual sounds HELI and MALE to give examples of sounds with high and low mean BPF values, and uniform or non uniform envelope patterns across the six octave frequency bands, which are listed in the legend at 212, 424, 848, 1697, 3394 and 6788 Hz respectively.

3.3.6 *Identification of Categories and Sounds*

In order to aid the interpretation of the outputted MCA dimensions and therefore help to explore the categorisation strategies used by participants a measure of category identification was calculated. Category descriptions given by participants were evaluated as to whether or not they matched the pre-defined categories of environmental, musical and vocal sounds - a full explanation of this process is located in section 3.2.2 and participant category descriptions can be found in appendix APPENDIX B.

Figure 3.13 shows a difference between the three CIL groups concerning category identification, (kruskal-wallis, $p < 0.05$) where the mean value is lower for NEW (42.6%$correct$) compared to INT (55%) and EXP (60%). Sounds are arranged along the x-axis by according to the data of EXP (blue squares). In general the results for INT and NEW follow a similar pattern. At the upper end of the graph are melodious musical sounds (*vln* and *gtr*), human action sounds (*dr* & *fstp*) and vocal sounds (*male* & *lgh*) Interestingly the sound *fem* has a lower identification score. Comments from the sound identification task (see figure 3.14) related to this sound often described it as musical, someone singing or a musical instrument which might explain the lower category identification scores in comparison to *lgh* and *male* sounds.

At the lower end of the scale the poor category identification scores of *obo* and *xly* are a result of these sounds being categorised more often as environmental sounds. This is most likely due to participants perception of these sounds referring to *telephones, ringing, doorbells* combined with CIL difficulties in perceiving the timbre of musical instruments. An additional factor may be that environmental noises such as doorbells and telephones carry important semantic information (warning people about an event) and that they are more familiar to participants rather than the sound of an actual oboe or xylophone.

Category identification is also low for the sound *cgh*, which was categorised as a musical sound and was often described as "drums". This likely reflects the ability of CIL to accurately perceive rhythmic sounds combined with poor identification of timbre. Finally the sound *bel* was categorised as musical rather than environmental. Whilst this sound is highly musical in nature this result highlights the importance of how the pre-defined categories were chosen and this point is explored later in the discussion as part of the methodological considerations.

Results in table 3.8 also show the category identification scores for the predefined categories - environmental (E), musical (M) and vocal sounds (V). The pattern of results is somewhat different across the three participant groups. For EXP the highest category identification is, as might be expected, for vocal sounds (70%), where as for INT it is the musical sounds (62%) and NEW show no difference across the three types of sound. This would suggest better recognition of vocal sounds related that improves with longer duration of implantation.

Category Identification (% correct)				
CIL group		EXP	INT	NEW
Predefined Categories	E	59	53	45
(see table 3.1)	M	55	62	41
	V	70	52	41

Table 3.8: Table of categorisation identification values (as percentages) calculated for individual sounds and predefined categories of environmental *(E)*, musical *(M)* and vocal *(V)* sounds, see table 3.1. Values for the mean and standard deviation (std) are caluculated over individual sounds.

Category identification values were also correlated with the coordinates of the factor map dimensions in figures 3.5 - 3.8 and also patient data from tables 3.3, 3.4 & 3.5. No significant results were found. On first glance this would suggest that the ease of categorising a sound is not part of participants categorisation strategy. Also that there is no aspect of the patient data that can account for the correct/incorrect perception of a sounds category. However the lack of significant correlations could be because category identification was calculated on the basis of only three broad categories where as there may have been more detail/subtlety to the real categories perceived by participants. Perhaps using different categories to reflect this may lead to significant correlations and evidence that the correct perception of a sounds auditory category is linked to the categorisation performance and strategies employed.

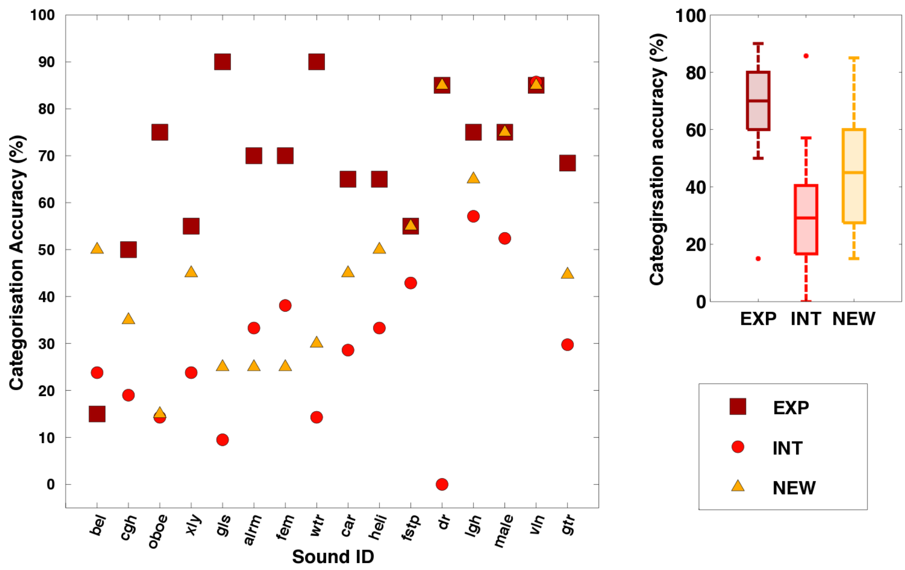


Figure 3.13: Scatter plot of category identification for CIL, where EXP are represented by ■, INT by ● and NEW by ▲.Results for individual sounds are plotted in the left hand scatter plot in order of mean category identification across all three stimuli conditions, whilst results for each stimuli condition are plotted as boxplots in the right hand panel.

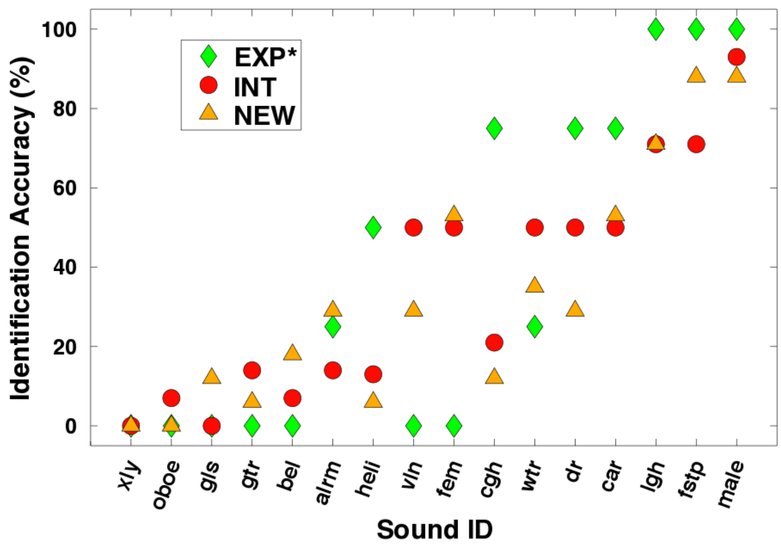


Figure 3.14: Scatter plot of sound identification, given as a percentage of participant comments that were interpreted as correctly identified. INT represented by ● and NEW by ▲. Results are also plotted for a *representative* group of 4 additional EXP participants that were not part of the original testing, these are plotted with ◇

.

Alongside analysing the identification of *categories* INT and NEW were also tested for identifying the individual *sounds*. Unfortunately this data was not recorded for original group of 16 Experienced cochlear implant listeners (EXP) and so a separate group of four Experienced cochlear implant listeners (EXP), referred to as *EXP** were also tested in order to provide representative data. No difference is seen between three sets of data, either for the identification of sounds shown in figure 3.14 kruskal-wallis, $\mathtt{chisq} = 0.04, p = 0.85$, nor for the identification performance of individual participants, $p = 0.581$. Assuming that the EXP* are representative of Experienced cochlear implant listeners (EXP) then it would appear that there is not difference in identification performance across the three groups of CIL nor across the three different kinds of sound. Interestingly for vocal sounds the results appear rather low considering that CIL normally perform around 80% . However it should be noted that the average value for vocal sounds includes that value for *cgh* which is much lower than the other three vocal sounds - *lgh*, *male* and *fem*. When discluding the sound *cgh*, the identification of vocal sounds rise to 71% for INT and NEW whilst remaining similar for EXP* at 67%. The identification of musical sounds is poor ($0 - 18\%$) as participants struggled in identifying the kind of instrument, often simply describing them all as *piano*. This contrasts with the much higher category identification ($41 - 62\%$) as participants were still able to perceive that these sounds were "musical". Identification of environmental sounds is also less than 50% for all CIL groups, which is not a surprise given the theory that even experienced CIL have poor environmental sound perception (ESP) [87, 82, 5, 124].

For the participant groups INT and NEW it was also possible to compare identification of sounds with the identification of categories. Data plotted in figure 3.15 shows that there is a correlation for both participant groups, although this is only significant for INT ($r = 0.58, p < 0.05$) compared to NEW ($r = 0.38, p = 0.14$). This would make sense as there should be a relationship between the correct identification of a sound and its subsequently successful categorisation. In both panels the majority of sounds lie below the dotted line suggesting that identifying the category of these sounds is easier than identifying the actual sound. There are also some sounds which are above or very close to the dotted line and are therefore more easily identified than they are categorise. Interestingly these seem to be the same sounds for both participant groups - *male, fem, fstp, lgh* and *car*.

Finally correlations were made between the *sound identification* data of figure 3.14 the coordinates of the MCA dimensions. Significant correlations were found for the results of INT with respect to Dim 1 (see figure 3.16) and also for NEW with respect to Dim 3 (see figure 3.16). As can be seen from the figures and also from the previously mentioned interpretations of the dimensions, the same four sounds (*male, fem, lgh* & *fstp*) are largely responsible for this correlation. These are of course four of the most correctly identified sounds and at the same time sounds generated by humans (either vocal or human action). Of course the interpretation of the dimensions involved is that they are focussed on grouping together human vocal sounds and the strong correlations presented here may suggest that part of the reason for this is that they are easily identified by CIL.

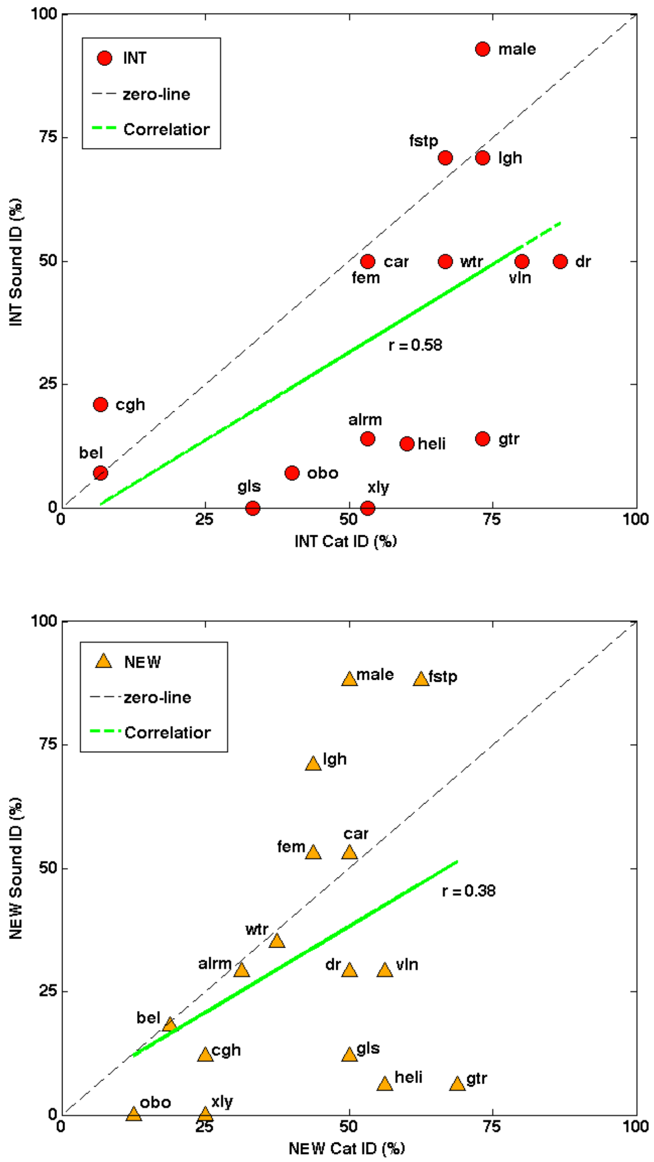


Figure 3.15: Scatter plots for results of Category Identification vs. Sound identification for INT plotted with • in the upper panel and NEWplotted with ▲ in the lower pane. The figures are useful for showing that there is a correlation (as shown by the green line) between CI users ability to identify individual sounds and their ability to identify the predefined categories of musical, environmental and vocal sounds.

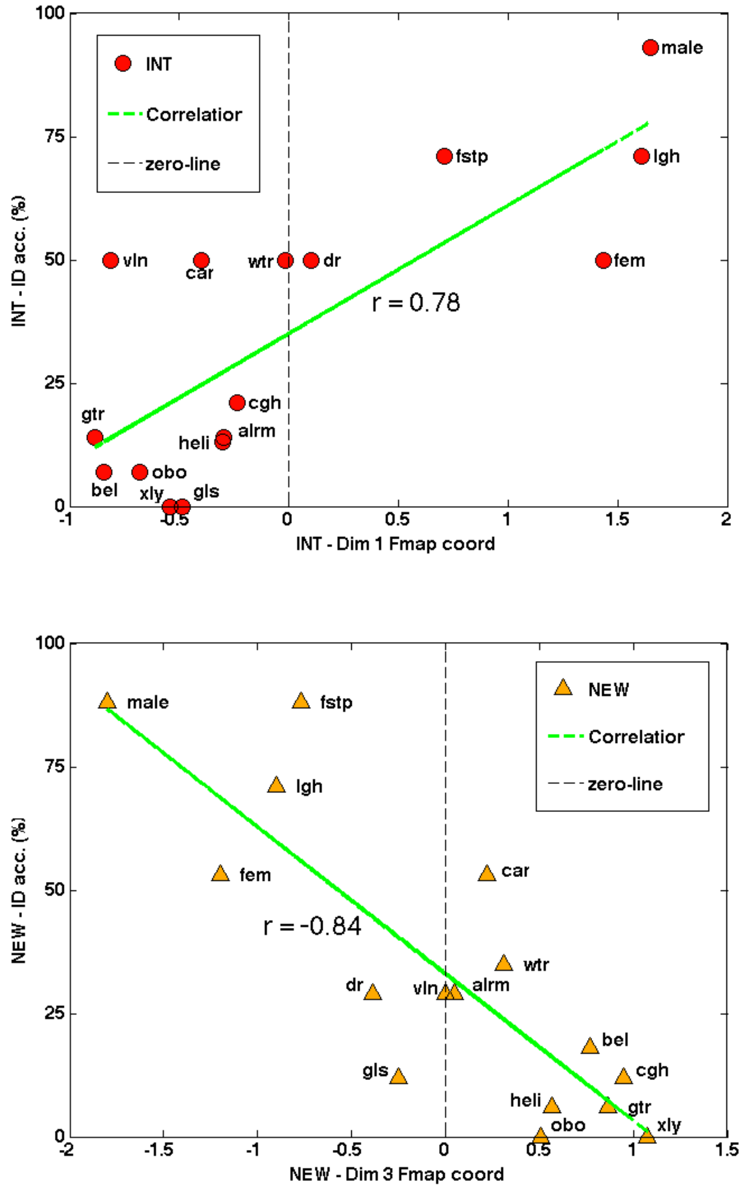


Figure 3.16: Scatter plots for MCA Dimension coordinates vs. Sound identification accuracy (%) (as shown in figure 3.14. Upper panel shows the results for INT CI Users plotted with ● with Dim 1 coordinates on the x-axis an. Lower panel shows for INT, plotted ▲ with Dim 3 coordinates along x-axis. Significant correlations between the data are shown with green lines, where the strength of correlation is also noted.

3.4 DISCUSSION

The same methodology as described in chapter 2 was again used to test the categorisation performance of three groups of CIL differing in the duration of implantation. They consisted of Experienced cochlear implant listeners (EXP) with more than 14 months of implantation., Intermediate cochlear implant listeners (INT) - from 6 — 14 months implantation and New cochlear implant listeners (NEW) with duration less than 6 months. It was predicted that categorisation strategy would change based on the duration of implantation with the perception of sounds improving with greater experience. It was also predicted that due to poor ESP results for all CILwould be different to those for NHL although the relation of chapters 2 3 is explored in chapter 4 and not here. Results showed that the least experienced group (NEW) was different to both of the other groups although similarities were noted across all groups, especially in the categorising of vocal linguistic stimuli and the identification of sounds. There was no difference in the identification performance of NEW, INT and a representative group of EXP participants such that identification of vocal sounds was good but the perception of environmental sounds remained poor even for experienced CIL.

3.4.1 *Human sounds*

Some of the strongest categories formed for CIL are those that involve *human sounds*. This is most clearly seen in the strong categorisation of vocal sounds for the more experienced CIL and concerns those sounds which most closely resemble speech (*male, fem* & *lgh*). The speech-like sounds are also some of the most easily identified sounds for all CIL and this may present a reason as to why they constitute a strong strategy for categorisation. It is however interesting that the vocal category is less dominant in the strategy employed by NEW even though they appear to be successfully identifying the sounds. It is possible that the reason for this difference is because NEW participants found the FST more cognitively demanding than the more experienced CIL. The task required a certain amount of memory when comparing the sounds and so whilst NEW were able to perform adequately at identifying the sounds, the memorisation and comparison of the 16 sounds was more difficult and contributed to the different categorisation strategy shown in the results. In addition scores of word and sentence perception are lower for New cochlear implant listeners (NEW).

Whilst they may be able to successfully identify the vocal sounds as voices, they may find it difficult to perceive other information (such as words) within the sounds and therefore find the task of categorising more difficult compared to EXP and INT. Previous studies in post-lingually deaf CIL have also shown that performance in speech tasks shows large improvements after 3 months implantation and then steady improvement up to 12 months before plateauing [20, 97, 144]. Concerning purely the perception of vocal sounds our results would appear to fit with this pattern such that both INT and EXP groups appear create strong categories of vocal sounds and EXP having slightly higher category identification of vocal sounds compared to NEW. Also whilst all three groups pair the vocal stimuli together, the categorisation strategy is different for NEW suggesting that this group did not distinguish the vocal sounds from non-vocal sounds to the same degree as INT and EXP, results which agree with [93] who showed vocal-environmental sound discrimination improved (slightly) with implantation duration.

Human action sounds also appear to be a robust category shared by the three CIL groups with the *fspt* categorised either separately or together with *dr* and comments related to these sounds referencing the inclusion of a human body in the act of generating the sound i.e. "*somebody* walking" or "a *person* opening a door". Again the sound *fstp* is successfully identified by all CIL and this again likely contributes to why this is a common strategy. Identification of *fstp* may also high due to the repetitive nature of the sound [59]. Regarding the sound *dr* identified as a door other descriptions referred to "human steps" , "hitting" and "grating" such that even though identification was not particular high participants were still able to correctly perceive an action and use this information as part of their categorisation strategy. It remains to be seen how CIL would perform in a paradigm involving ONLY action sounds and but none-the less it is interesting that NEW performed well in this manner.

Of course all of these human sounds share another factor in common in that they are likely to be highly familiar to listeners. Not only will a listener generate these sounds for themselves but they will be regularly heard from other sources in daily life, building up a high familiarity and making them more easily perceived [6].

3.4.2 *Categorisation and Identification*

Whilst the identification of human voice seems relatively adequate for all CIL the identification of other (environmental) sounds remains quite poor. This is not surprising given that even experienced CIL have difficulties with such sounds [93, 122, 111]. However the results of INT bring up an interesting question. If INT and NEW are poor at identifying the sounds how is is that the categorisation performance of INT more closely resembles that of the EXP.

When categorising there are not the same level of requirements for recognition of a stimulus category as for specific stimulus identification, for example the results of [59] show higher categorisation scores compared to identification. This is one of the underlying reasons why categorisation is used as a cognitive process in daily lives. It requires more knowledge, attention, processing power of the fine details to identify the exact source compared to assigning category membership (identifying the category). For example whilst describing the sound of a guitar as "musical" may not correctly identify it, it does correctly categorise the sound and allow the listener to understand certain characteristics.

It is also likely that INT have better access to acoustical information of the sounds that results in different categorisation compared to NEW. This is certainly true for vocal sounds and is evidenced by the known improvement in speech perception by CIL over time. Whist the same improvement is not seen with identification of environmental sounds [87, 108] perhaps we show here evidence that it exists for a broader level of perception i.e. categorisation. As mentioned above the affect of familiarity cannot be ignored not only will CIL access to acoustic information improve as they adapt to using their implant but their familiarity with their auditory environment and to individual sounds will also improve. Neither can the possibility that NEW find the task more cognitively demanding than the more experienced CIL as evidenced by the longer durations taken to complete the experiment and lower word recognition score.

3.4.3 *The effect of Implantation Duration*

The main aim of the research described in this chapter was to test the affect of implantation duration on the categorisation of 16 common sounds. No previous study has tested CIL in such a paradigm before and so there is not any available data for direct comparison. The results do however fit with previous findings that for post-lingually implanted CIL the perception of vocal and musical sounds improves over time and that the perception of environmental sounds remains difficult even after significant amounts of experience.

Most interesting is the difference in categorisation strategy seen for NEW compared to INT and EXP. With increased duration of implantation the category identification is seen to rise and results become more and more similar to those of NHL (see chapter **??** for more detailed commentary on this). However it is not simple to say why this change in performance occurs as it is likely a result of both improved access to acoustic information within the signal as well as increased familiarity with the auditory world post-implantation. Results of task performance also show that increased task difficulty is likely a factor in why NEW perform differently. Whilst the identification of categories would appear to improve with CIL there is no evidence that the identification of individual sounds is changing. Of course the experimental method must be addressed for two particularly reasons that could have affects.

Firstly although category identification values can be used to comment on the improvement across CIL groups the measure has been calculated with regards to the pre-defined categories which are quite broad and may not be fully representative of the real-world categories that human beings use - this point is discussed in more detail in chapter 4. Secondly the manner in which identification of individual sounds has been measured relies on the schema of the experimenter when reading and interpreting the comments that participants used to describe the sounds. There are of course more controlled methods for analysing identification of sounds, however they do not allow the true perception of the sound to be recorded and this was deemed important information for the current study especially in determining how the perception informed participants categorisation strategies. For example whether or not NEW were using "musical" listening more often than the more experienced CIL. Regarding the use of listening modes results would indicate that categorisation strategies for all CIL groups are based on semantic information associated with the sounds, such that participants are still using a more "everyday" listening mode when undertaking the FST. This tells us that independent of implantation duration CIL are still able to extract a certain amount of semantic information from the sound on which categorisation strategies are built. Rather than forming categorisation strategies on more qualitative aspects, such as the emotional content of a sound (whether its pleasant / unpleasant), or the acoustical information (being rough / smooth, high frequency low frequency). Of course that is not to say that these strategies are not used by certain participants. It is likely that very few of the participants follow every one of the strategies detailed above, but rather the results reflect the dominant strategies employed by the overall population and do not detail every specific strategy used by each participant. Instead the results show that is no difference in the amount of variability accounted for by the MCA analysis and this would suggest that there is similar amounts of variation within each CIL group and would discount the possible hypothesis that the more experienced CIL are more uniform in performance. The presence of a large amount of variability amongst each CIL group is also further evidence of the variation in categorisation strategies that are likely used.

3.4.4 *Importance of Acoustic information*

Along with finding that semantic information is responsible for forming categories, the present results suggest that certain acoustical characteristics may also be linked with categorisation strategies. These mainly center around cues related to perceptions of frequency or rhythm / temporal structure of the sound and reflect CIL abilities in perceiving spectral and temporal information. Certainly the inclusion of variables associated to peaks or bursts in the sound envelop are not so surprising given that cochlear implants process temporal cues better than spectral ones [111, 141, 59]. Spectral variables that come out of the results also reflect cues associated with slower spectral dynamics e.g. centroid and pitch salience. Certainly some of the most salient stimuli are ones with clear temporal structure and defined pitch e.g. guitar. There are aspects of the acoustic analysis which show processing of slow spectral dynamics and the possibility that CIL are able to perceive the slower spectral dynamics and then use this is as part of categorisation strategies.

Interestingly the acoustic variables linked to CIL categorisation strategy are different across the three participants groups suggesting an effect of the duration of implantation. It is difficult / impossible to state concretely whether this is a result of improved access to acoustic features or increased familiarity and knowledge of the auditory world, however likely both of these aspects of perception improve with increased duration of implantation. The difference is still present though and purely based on acoustical features shows a change / an evolution from strict acoustical features used by less experienced CIL to features that reflect a perception of the sound for more experienced CIL. For example, where as measures of maximum or mean frequency are direct measures of the spectral content within a sound, the pitch salience reflects the "perceptual strength of pitch/harmonicity". In a similar vein it could be said that measures of "peaks" and "bursts" correspond to direct measure of temporal fluctuations that can be more or less "detected", peaks in the autocorrelation function are related to the periodicity of the sound and reflect a more perceptual quantity.

Finally A common element across all CIL is the contrasting of environmental sounds (notably, car, heli & wtr) versus vocal and musical sounds with links to the the uniformity of the sounds envelope (the mean correlation of band-pass-filtered rms). Even with the influence of semantic information i.e. that car and heli are both transport sounds it shows that characteristically "uniform" sounds may be easily/strongly perceivable for CIL no matter the experience with the implant.

Overall there is likely an association of certain acoustic variables to the perceptions of the CIL participants and the resulting categorisation strategies, for example relating frequency information to the categorisation of musical sounds. The results also reflect previous studies highlighting the importance of temporal structure and slow-spectral dynamics as being important to the categorisation of sounds by CIL. A difference in the associated acoustic cues used by different participants groups could also highlight that more experience CIL use more perceptual based measures as part of their categorisation strategies. However it is vital to remember that the perception of purely acoustic information cannot be easily separated from the perception and use of semantic information. This makes conclusions difficult to arrive at and requires further work to uncover finer details and the specific roles that the different perceptions play in auditory categorisation.

3.5 CONCLUSIONS

The current chapter tested a Free Sorting Task (FST) of common everyday sounds with three groups of CIL with different durations of implantation ranging from 1 - 14 months. Group results showed that the more experienced CIL categorised the sounds in close similarity to the three pre-defined categories of musical, environmental and especially vocal sounds. Whilst the most inexperienced CIL (implantation duration less than 6 months) demonstrated similar strategies by grouping together musical sounds and temporally distinct (and familiar) environmental sounds, results were overall strikingly different. This was shown for example by a poorer category identification ability, even though identification of individual sounds did not differ across the groups of CIL. Acoustic analysis showed that different acoustic cues were linked to the categorisation strategies of different CIL groups possibly highlighting an additional effect of the implantation duration. Overall these results show that the development of auditory abilities post implantation - the ability to access both acoustical and semantic information associated with the perception of auditory objects, is seen be significantly differ for CIL below 6 months implantation. Whilst identification of specific sounds shows no difference with duration of implantation, category identification does and may hint towards future perceptions of categorisation as being important to the study of auditory perception with CIL.

Key Points

- CIL with duration of implantation greater than 6 months (EXP & INT) show categorisation on a broad level of vocal, environmental and musical sounds in agreement with the pre-defined categories. Categorisation strategies as expressed by MCA analysis show strong separation of vocal sounds, not including that of the *coughing* sound which is more often categorised amongst musical sounds due to its rhythmic nature. Clear strategies also exist in the contrasting of musical and environmental sounds, which are in turn centred around melodic sounds and mechanical/transport sounds.

- New cochlear implant listeners (NEW) with less than 6 months implantation differ in their categorisation strategies compared to more experienced CIL (INT & EXP). Categorisation strategies of NEW are predominantly related to human action sounds, which are also temporally distinct and rhythmic. Similarities do exist in the grouping together of vocal sounds, however this is seen as a much weaker categorisation strategy appearing on MCA dimension 3 compared to Dim 1 for more experienced CIL. It also involves mainly only the sound *male* -adult male talking. Therefore suggesting that the perception of vocal sounds is weaker for the most inexperienced CIL and in keeping with the known development of CIL auditory abilities. However whilst category accuracy was similar, identification accuracy of individual vocal sounds did not differ with implantation duration suggesting categorical perception as a way of distinguishing the least experienced CIL.

Key Points

- Interestingly there was no difference in identification performance of environmental sounds for , INT, NEWnor EXP* (representative sample). This is again in keeping with previous research that suggests CIL perception of environmental sounds remains poor even after many months of implantation.
- Acoustic analysis showed correlations of certain cues to the categorisation strategies of CIL. These mainly related to temporal aspects of the sounds, for example the mean value of peak values within the autocorrelation function i.e. a value related to the periodicity of the sound, which was linked to the separation of musical sounds. Correlations also showed that a more limited number of variables were correlated to the results of EXP, which may be evidence for more selective auditory abilities amongst the most experienced CIL.
- When all three groups of CIL were analysed as one single group the duration of implantation was found to be correlated to the categorisation of vocal sounds for subsets of CIL. This was true for EXP and NEW suggesting that the strategy of grouing together vocal sounds and likely the perception of vocal sounds improves from $0-6$ months implantation and then from 14 months onwards. However no correlation was found for the subset of INT suggesting a slower level of improvement.
- Categorisation performance is likely influenced by a combination of adaptation to the implant and familiarity with the auditory world although it is difficult to distinguish the influence of each on the final results and difficult to evaluate the familiarity of individual CIL with the stimuli used.

4

COMPARISON OF AUDITORY CATEGORISATION BY NHL AND CIL

In the previous two chapter the results of a FST *of 16 everyday sounds have been presented for* NHL *and* CIL*. Three sets of* NHL *participants were tested with differing levels of spectral degradation (vocoding) applied to the stimuli in order to simulate* CIL *processing. Three different sets of* CIL *users who had different durations of implantation were also tested. One group of experienced listeners with implantation greater than 14 months (*EXP*), another group of intermediate* CIL *users with duration between 6 and 14 months (*INT*) and finally an newly implanted group with duration less than 6 months (*NEW*). Result from chapter 2 show robustness of certain categorisation strategies when sounds were spectrally degraded and results from chapter* **??** *show that similar categorisation strategies may also be used by the more experienced* CIL*. The following chapter will compare the results of the two previous chapter in order to evaluate the performance of* CIL *users vs.* NHL *and to see how accurately the results of categorisation with vocoded sounds represent* CIL*.*

Chapter Aims

1. The main aim of this chapter is to compare the results of chapters 2 and 3 and the performance of NHL and CIL in categorising 16 everyday sounds. The chapter aims to analyse the similarities in the categories that are created, the strategies used to do so and the perceptions/similarities that these strategies are based on. The performance of the Free Sorting Task (FST) will also be compared to see how the different groups of participants carried out the task.

2. The chapter aims to see how closely, if at all CIL results reflect those of NHL. The three groups of CIL are compared separately as it is predicted that the EXP will show the closest relation to NHL category performance.

3. Results of CI simulated sounds (16C and 8C conditions) are used in comparison to CIL results to evaluate if and how accurately the CI simulated studies mimic those of actual CIL and whether these simulations represent a suitable manner for future testing of auditory categorisation.

4.1 RESULTS

4.1.1 *Global performance - TCL statistics*

Analysis has been performed across all 6 participant groups in order to assess the manner in which the FST was completed and results are summarised in figure 4.1. Concerning the number of playbacks all groups are similar ($5.07 - 5.93$) expect for the NEW group (7.46) which is significantly different, kruskal-wallis $p < 0.001$. NEW also create more categories (7.2) than NHL participant groups ($4.8 - 5.4$), kw, $p < 0.05$, and take the longest to complete the task (871 seconds on average). This shows that the CIL users with the least experience require more time and more attempts at listening to the sounds in order to complete the task. This is not so surprising given this group has had the least amount of experience with their implants. That the results for INT are more related to EXP and NS participant groups suggest categorisation performance likely begins to plateau somewhere between 6-14 months of implant experience.

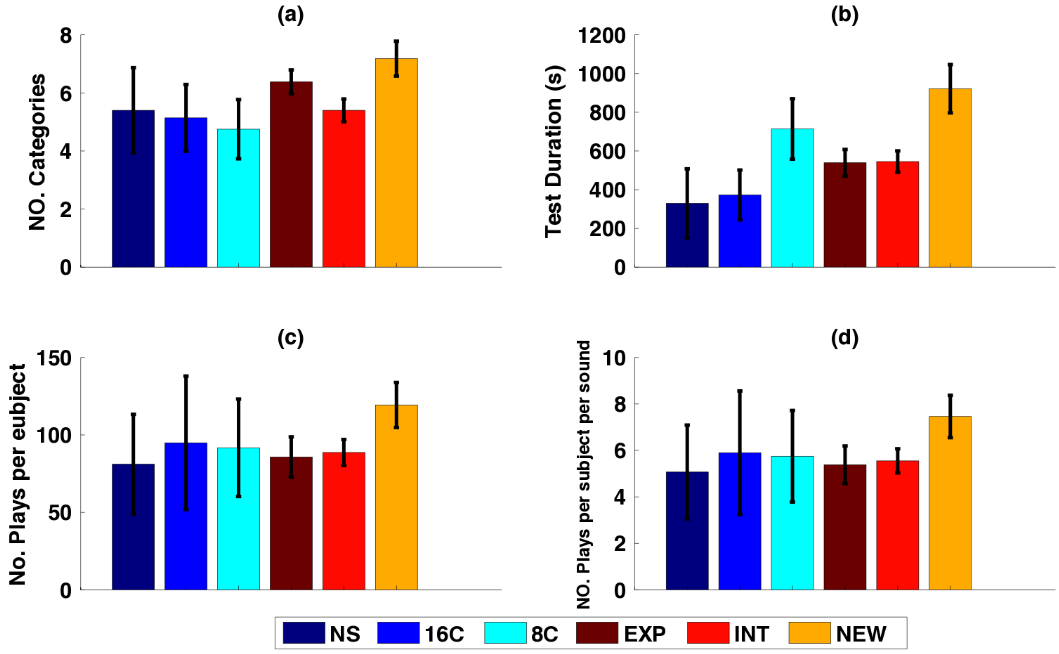


Figure 4.1: Statistics of FST performance for all six participant groups. **Panel (a)** - the average number of categories created by each participant group. **Panel (b)** the average duration taken to complete the FST, in seconds (s). **Panel (c)** - the average number of times that a single participant listened to the sounds. **Panel (d)** - the average number of times a single participant listened to a single sound.

4.1.2 *HCPC analysis (Dendrogram) summary*

Although the dendrograms in the previous chapters are only simplified representations of the final categorisation of stimuli they are useful for comparative purposes. Figure 4.3 displays the dendrograms for all six participant groups for this purpose. Similarities that occur across all groups are the grouping of vocal stimuli together, although this is strongest for *male* & *fem*, sounds which contain speech, and also *lgh*. The last vocal sound *cgh* is only grouped amongst the other vocal sounds by NS participants and based on comments from multiple participants is perceived often as a musical sound (e.g. drums) due to its rhythmic nature. Other pairs/groups of sounds are also common in all dendrograms for example *car* & *heli*, both transport/machine sounds.

It also appears that many musical sounds are also commonly grouped together by most if not all participant groups. There is however a clear difference between those sounds which contain multiple tones and a sense of melody being more often perceived as musical sounds, in comparison to sounds that contain only a single sustained pitch (xly & obo) which were often referred to as household alarm sounds or simply "ringing".

Finally the sounds *dr* & *fstp*, which constitute examples of "action sounds involving a human operator" appear paired together in every dendrogram apart from that of the Experienced cochlear implant listeners (EXP). These sounds also seem especially important in the case of New cochlear implant listeners (NEW), forming the most strongly separated category of sounds.

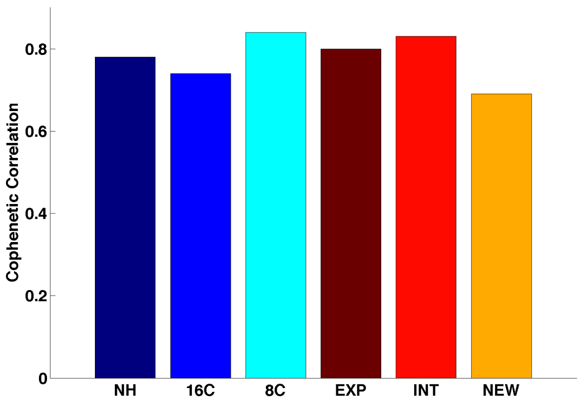


Figure 4.2: Bar chart of Cophenetic Correlation Coefficient (CCC) values calculated for the dendrograms (see figure 4.3 of each participant group

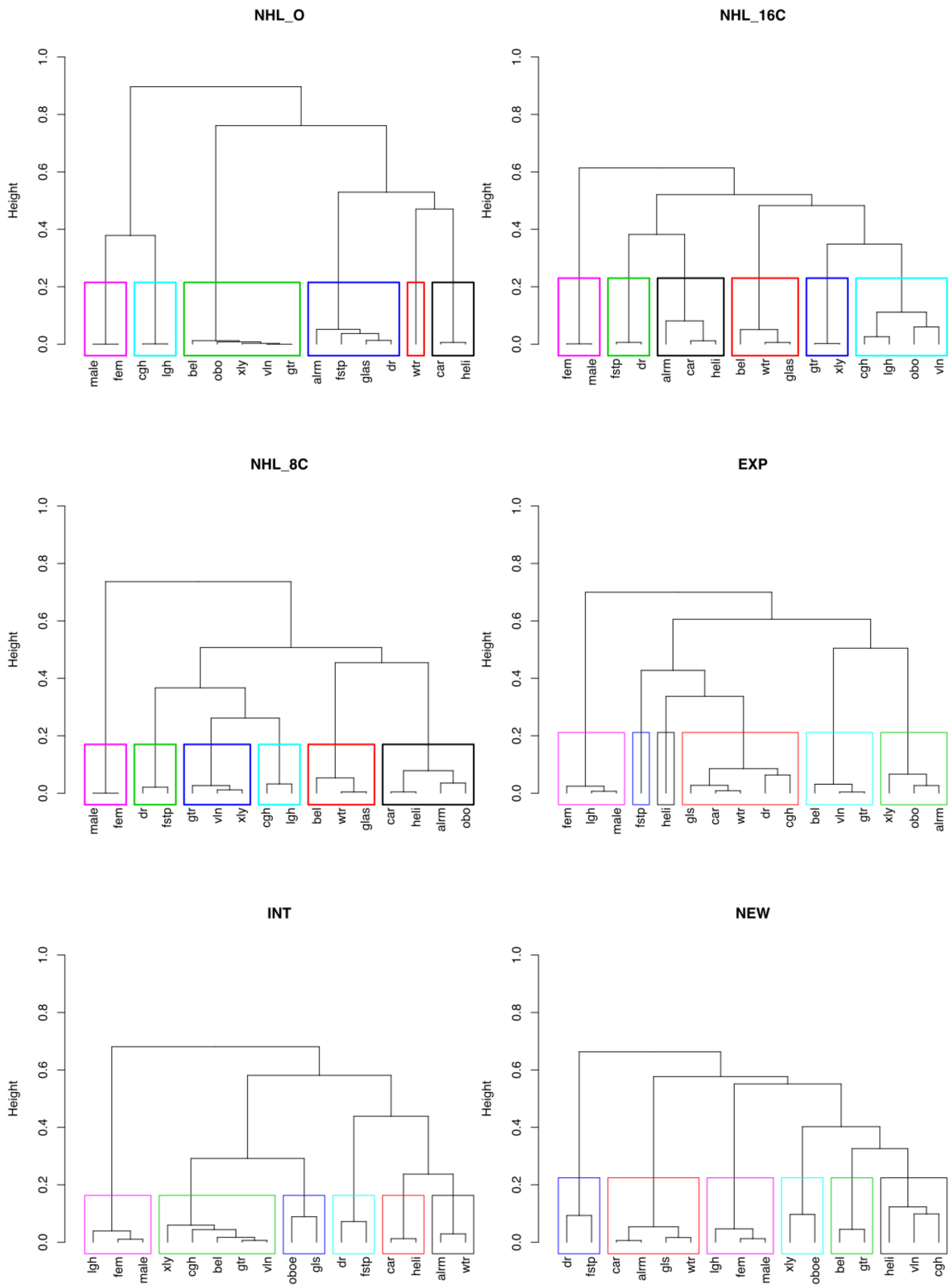


Figure 4.3: Dendrograms showing the final categorisation of sounds by all six participant groups including; NHL participant groups tested with Natural sounds (NS), 16 channel vocoded sounds (16C) and 8 channel vocoded sounds (8C); CI user participant groups - Experienced cochlear implant listeners (EXP), Intermediate cochlear implant listeners (INT) and New cochlear implant listeners (NEW). The height axis gives the perceptual distance between each stimulus whereby a large height indicates that participants deemed those two stimuli to be highly dissimilar and vice versa. Finally stimuli are labelled using the abbreviated sound IDs.

Also CCC values in figure 4.2 give an indication of how representative each dendrogram is. Unsurprisingly in the case of **CIL !** (CIL !)the lowest value is seen for NEW which most likely reflects a higher variability in categorisation strategies amongst these participants that is difficult to represent in the final dendrogram. However in the case of NHL, the condition where perception is most difficult , 8C, has the highest CCC value. This could be because participants found identification of sounds easier in these first two conditions and allows participants more possibilities for categorising sounds rather than in the conditions where identification is more difficult such that certain (more identifiable) sounds act as anchors to other sounds.

The distance within (INTRA) and between (INTER) the categories shown on the dendrograms was also analysed, and these values are summarised in figure 4.4. ANOVA analysis shows that regarding the INTER category distance (left panel), the value for NH (mean = 0.63) is greater than all other participant groups ($p < 0.05$), whilst all other participant groups are of near equal value. For the INTRA category distance (right panel) the lowest value, show the strongest level of categories within categories is seen for NH (mean 0.0162), however this is only significant compared to 16C and EXP ($p < 0.05$). The most concise categories that are also more separate from each other are therefore made by NHL in the natural sound condition (NH). Other participant groups are much more varied

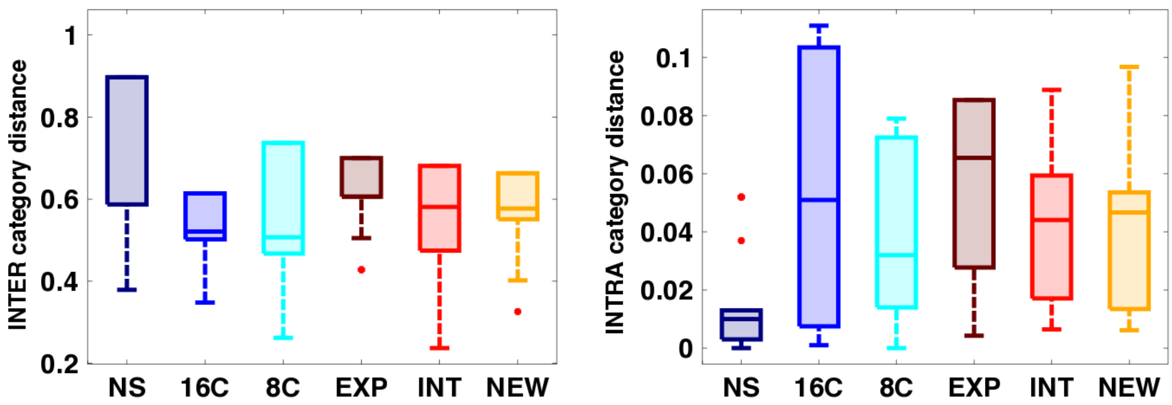


Figure 4.4: INTER (left panel) and INTRA (right panel) category distances calculated from the dendrograms in figure 4.3. Boxplots are plotted to show the overall data for each of the six participant groups.

4.1.3 *Co-occurrence matrices*

Based on the construction of co-occurrence matrices it is possible to make certain analyses to compare the perception of stimuli by the participant groups. Values of RV coefficients in table 4.1 (shown as a colour map in figure 4.5) are all above 0.8 and show that in general the co-occurrence matrices are quite similar. Highest values are seen between the vocoded conditions 16C and 8C and also between the three CIL user conditions (NEW, INT, EXP) indicating these participant groups as highly similar. The lowest value between NS and NEW is somewhat expected as these are the "best" and "worst" performing groups i.e. with most natural and most degraded auditory abilities.

Interestingly figure 4.5 shows that whilst all of the CIL groups (NEW, INT and EXP) are similar, the results for NHL show that the two vocoding conditions are more strongly different to the NS condition (labelled NH). This would suggest that whilst CIL share similarities in perception, the degradation of sound induced by vocoding creates results and perception that are different than the NS condition. Possibly this is a results of participants being un-used to listening to vocoded sounds and results may change if participants were able to become more familiar with vocoded sounds. Previous studies have also found environmental sound identification with vocoded sounds to improve with training with NHL.

RV coefficient	NH	16C	8C	CIL	INT	NEW
NH	-					
16C	0.866	-				
8C	0.838	0.947	-			
CI	0.885	0.895	0.881	-		
INT	0.886	0.900	0.888	0.917	-	
NEW	0.811	0.896	0.873	0.911	0.916	-

Table 4.1: RV coefficients compared between all pairs of results.

Co-occurrence matrices have also been used to calculate the average distance between pairs of stimuli contained in the original pre-defined categories. Table 4.2 shows that all numbers are on the low side of expectations, especially for NS participants. This issue is discussed later on in section 4.2.2.1 which explains why the choice of stimuli and predefined categories may have been mis-judged.

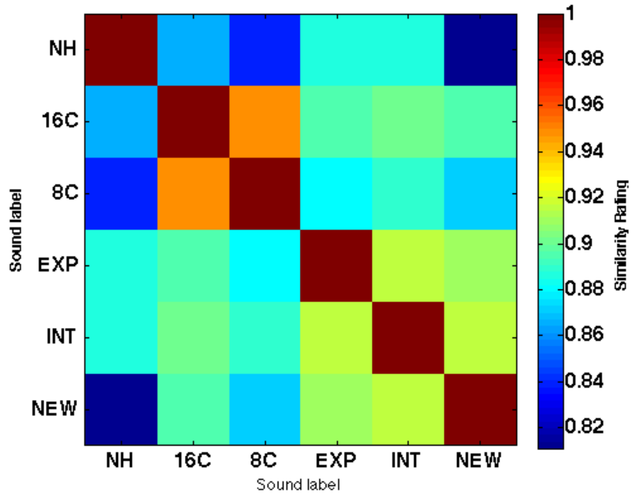


Figure 4.5: Colour map showing the RV coefficients for all six participant groups. RV coefficient calculated as a correlation between two co-occurence matrices, where each cell represents the overall correlation. Values vary from 0.8 (in blue) to the highest correlation 1 (in red).

These results do still however prove useful in showing the highest similarity to the predefined categories being for NS and the lowest similarity for NEW, again as possibly expected based on auditory performance. Values for EXP and INT are the same and may point towards these groups performing similarly and not differing greatly in the ability to complete the task. That 8C has a higher value than 16C would suggest that in the condition of *LESS* spectral information NHL were able to more accurately categorise the sounds. However this may may simply be a result of the choice of predefined categories, such that more by chance than design the stimuli pairs in the 8C condition more strongly resemble those of the predefined categories.

	NS	16C	8C	EXP	INT	NEW
Mean Pairwise Similarity	0.29	0.22	0.25	0.21	0.21	0.16

Table 4.2: Pairwise distance average for all participants groups based on the predefined categories

4.1.4 *MCA analysis summary - interpreting the dimensions*

For the results of each participant group a total of 5 MCA dimensions were retained for further analysis. Dimensions were retained on the criteria that each should cover a minimum of 8% of the total variance of results. Figure 4.6 shows the total variance accounted for by the 5 dimensions for each participant group. Values are highest for the three NHL groups (NH, 16C and 8C), with the highest value as may be expected being in the natural sound condition (NH) at 72%. Lower values for CIL groups suggest that compared to NHL, CIL participants show that there is a greater amount of variance unaccounted for by the retained dimensions and suggests weaker agreement to the categorisation strategies used.

Eigen values shown in figure 4.7 would also give evidence that NHL in the NS condition more strongly use the categorisation strategies described by Dimension 1 and 2. Other participant groups have lower eigen values and thus less participants show agreement. That the values for CIL (especiallyNEW) are high in the latter Dimensions shows that there is a greater number of participants using these latter dimensions and evidences the assumption that CIL are using multiple and varied categorisation strategies compared to NHL, even those in the 16C and 8C vocoded conditions.

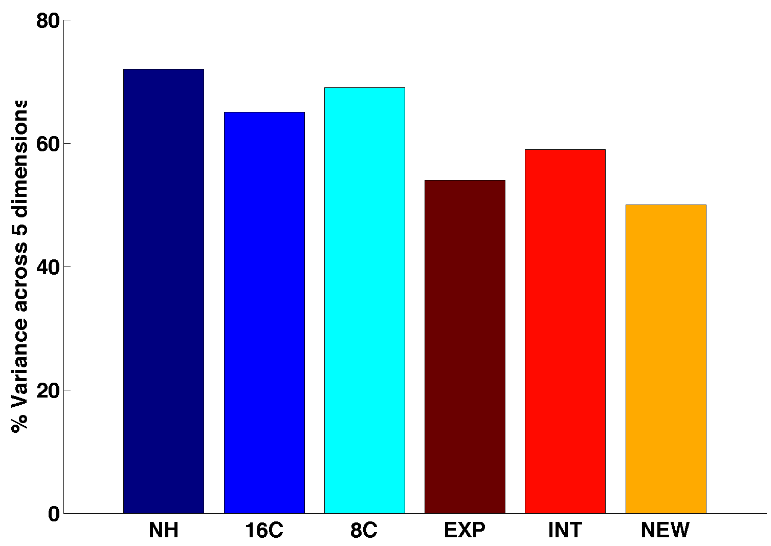


Figure 4.6: % Variance covered by retained MCA dimensions. In each case 5 dimensions retained for each participant group.

A large part of the analysis revolved around the successful interpretations of the dimensions generated by the MCA analysis. Multiple aspects of the data were used to aid interpretations, including the finding the sounds which most

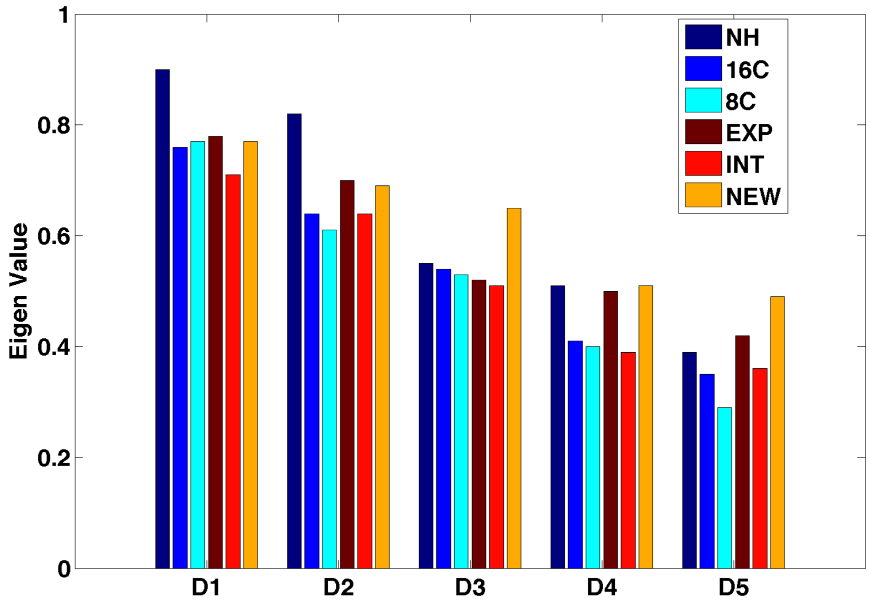


Figure 4.7: Eigen values for retained 5 dimensions across all NHL and CIL participant groups

strongly contributed to each dimension, participants descriptions of categories and sounds (APPENDICES A & B), and also correlations to acoustic characteristics which are summarised in table 4.4. Interpretations are described in fuller detail in the two preceding chapters concerning NHL and CIL separately, however for the purposes of comparison table 4.3 briefly summarises the findings. As can be seen Dim 1 is associated with contrasting the vocal sounds from other stimuli. Specifically this contrast is associated with the speech-related vocal sounds (*fem* & *male*) however NHL-NS also include the sound of laughter and coughing in this strategy. Experienced CIL (INT & EXP) also include the sound of laughter, however as figure 4.3 shows there is confusion of the sound of coughing grouped alongside musical sounds. This Dim1 strategy is not common to NEW however, although it is seen to a weaker degree on Dim 3. Dim 2 is associated with the separation of musical or environmental sounds from the other stimuli. Whilst this manifests as similar results, the exact strategies for doing this likely differ slightly. Strategies that involve the separation of singular sounds are shared in common by all groups except that of NHL-NS. This hints toward the differing categorisation techniques undertaken by CIL and NHL in adverse listening conditions and likely reflects the fact that strong similarities between sounds cannot be identified. Therefore sounds that somehow "stand-out" are categorised on their own. These include the sound of footsteps, the oboe (often perceived as a warning sound) and helicopter. These sounds also related to other shared strategies notably the grouping together of; *human action* sounds, the footstep and door closing; *transport* - sounds of car and helicopter; *musical single tone* - sounds of xylophone, alarm and oboe.

As mentioned above acoustic characteristics were also correlated with the coordinates of MCA dimensions. Perason correlations were conducted using MatLab and those of significant ($p < 0.05$) are summarised in table 4.4.

4.1.5 *Identification of Categories and Sounds*

Comparing categorisation identification scores across all six participant groups also shows similar patterns to the previous analysis. KW and ANOVA analysis show that NS have a significantly higher categorisation identification score (73% correct category identification) compared to 16C (28%), 8C (42%) and NEW (45 %), whilst both EXP (59%) and INT (53%) are higher than 8C, kw $p < 0.001$. These results are plotted as boxplots in figure 4.8.

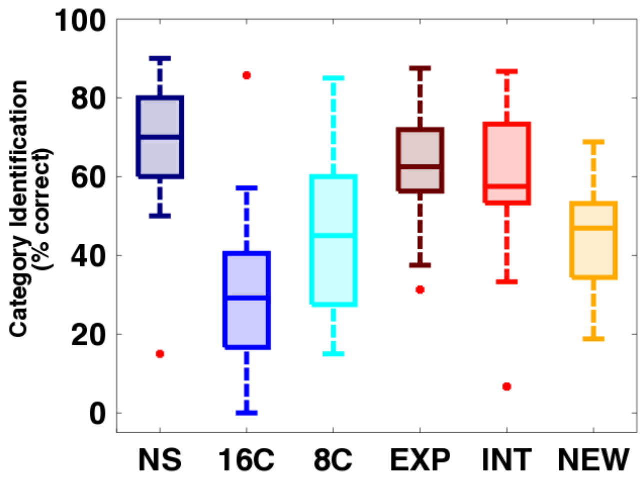


Figure 4.8: Box plot of category identification accuracy (% correct) across all six NHL and CIL participant groups.

Summary table of MCA Dimension Interpretations					
Participant group	**Dim 1**	**Dim 2**	**Dim 3**	**Dim 4**	**Dim 5**
NS	voice / non-voice	ES / musical	nature vs human	transport / nature	speech / non-speech voice sounds
16C	voice / non-voice	(ES+voice) / (musical+human)	human vs musical	Transport	Oboe
8C	voice / non-voice	music / transport	Gls + wtr + bel	human (dr+fstp)	non-speech voice
EXP	voice / non-voice	music / (ES+voice)	Footstep	musical singular note	helicopter
INT	voice / non-voice	ES / (music+voice)	human vs helicopter	Alarm	(glass+oboe) / xylophone
NEW	ES / (music+voice)	human (dr+fstp)	Voice sounds	Oboe	

Table 4.3: Summary of MCA dimension interpretations. Where individual sounds are noted, this refers to categorisation strategies that separate the stated sound from all other stimuli. *Transport* refers to the sounds of *car* and *heli*. ES is used as shorthand for Environmental Sounds.

Summary table of Correlations between MCA Dimensions and Acoustic characteristics of sounds					
Participant group	**Dim 1**	**Dim 2**	**Dim 3**	**Dim 4**	**Dim 5**
NH		Autocorr Mean Peak	Max Salience		No Bursts
		Autocorr std Peak			Spectral kurt
		Mean Salience			
		Max Salience			
16C	Autocorr Range	Autocorr No Peaks	Autocorr No Peaks		
	Mean bpf		Autocorr Mean Peak		
	Mean Salience		Autocorr std Peak		
	Max Salience				
8C	Autocorr Range	Autocorr std Peak		Autocorr No Peaks	Spectral STD
		Mean bpf		Frq std	Spectral kurt
		Max Salience			
EXP		Pitch Salience mean			
		Autocorr Peak mean			
		Autocorr Peak std			
INT		Frq max			
		BPF mean			
		Peak max			
		Env Peak mean			
		Env Peak std			
		Wav range			
NEW	Frq max		Frq mean	Centroid	
	Centroid		Env Peaks NO.	Pitch Salience max	
	Spectral std				
	BPF mean				
	Peak max				
	Peak mean				
	Env Peak mean				
	Env Peak std				
	Wav range				

Table 4.4: Summary of correlated acoustic variables for each participant group.

4.2 DISCUSSION

In the present study a Free-Sorting Task (FST) of every-day sounds was developed and tested with groups of normal hearing listeners (NHL) and CIL. Whilst previous studies have tested auditory categorization amongst similar populations this study is the first to directly compare them as well as the first to test the effect of implantation duration on categorisation performance. Analysis of the categorization has been carried out using MCA and HCPC and is supported by correlations with acoustic analysis and CIL user data to provide more details to the categorization strategies used. Overall the more experienced CIL show comparable levels of discrimination to NHL regarding the three pre-defined categories however with more variation between individual participants. Correlations also showed that certain spectral and temporal information within the raw acoustic signal can be used as predictors of the cues used by CIL users to discriminate between different categories of sounds. This includes information within the lower frequency channels used to categorize vocal sounds, spectral information (pitch saliency) for categorizing musical sounds and temporal information (measures of periodicity) used for categorizing repetitive sounds.

4.2.0.1 *Similarities of categorization processes between NHL and CIL users*

There are obvious similarities when comparing the results of the more experienced CIL participant groups (EXP and INT) and the NHL participants, most notably in the separation of the vocal, environmental and musical sounds. Interpretations of Dimensions 1 & 2 strongly suggest that more experienced CIL are following similar strategies to NHL in contrasting vocal stimuli against all other sounds (Dim 1) and also separating musical and environmental sounds (Dim 2). It could be argued that NEW also show tendencies to use the same strategies, especially regarding vocal stimuli, however overall the categorisation is rather different. Most strikingly Dim 1 for NEW is concerned with contrasting environmental sounds against all others, rather than contrasting vocal sounds as is seen in the cases of all other participant groups. Although NHL and CIL results are similar, the strategies used by the two participant groups to arrive at these results may be different. Firstly the implant delivers a poorer quality stimulus with less spectral information that CIL can use in their perception of sounds. And secondly, due to periods of deafness and adaptation to hearing with an implant, cortical reorganization experienced by CI users could affect their subjective experience of sounds, for example their degree of familiarity with previously encountered sounds and the development of familiarity with new sounds [84, 6].

Previous studies suggest that auditory categories are primarily based on the semantic information of a sound producing object or action, which follows from identification [37, 52, 80, 57]. Less important are the acoustical characteristics of the sound which determine the qualitative perceptions of a listener, for example the pitch, the roughness or the emotional content of a sound. Indeed, the descriptions provided by participants mostly refer to source properties (e.g. vehicle noise, noise in the house) rather than qualitative aspects of the sounds (e.g. "treble tones that crescendo" or"complementing rhythm"). However the *acoustical bias* observed by [43], whereby sounds preceded by non-living sounds were categorized more strongly by acoustical means, suggests that in certain conditions the preference for using semantic information is altered. This could be down to identifiably, as non-living sounds were less well identified, which could therefore have lead to stronger use of the "musical listening mode" and the concentration on acoustical rather than semantic similarities. The current study would support the initial theory that listeners focus predominantly on the semantic source information and then secondly the qualitative acoustic information. In addition the similarity in results for categorization accuracy and category identification suggests that at a category level CI users perceive the sounds in a similar way to NHL. This means that the link between hearing the sound and the semantic representation is still strong for CI users in spite of the reduced auditory input. However, the differences in category descriptions would suggest that NHL have a greater ability to identify the sounds. CI users more often describe the qualitative acoustic characteristics, a process which happens when sound-source identification fails.

4.2.0.2 *Variability of categorization in CIL users*

Compared to NHL, CIL participants show less agreement to the categorisation strategies outputted by MCA and the associated results cover less of the overall variance. Whilst it may be true that some CIL participants use the same strategies slightly differently, the level of agreement is still not as high as for NHL even in the case of vocoded sounds. This suggests that participant variation is not due to the difficulty in perceiving stimuli per se, but more likely due to variability within the perceptions of the CIL participants. There is also consistent variation across all three groups of CIL, suggesting that even participants with greater experience of CI hearing varied in their performance. Regarding the separating vocal sounds, the previous chapter showed that duration of implantation was linked with the adherence to this specific categorisation strategy and also possibly to the strategy separating musical stimuli as well. Therefore the implantation duration can be established as a source of the variance amongst CIL concerning these two aspects of the categorisation. Other factors have also been loosely associated with categorisation results; for example hearing threshold of non-implanted ear and to a lesser extent speech performance scores, which can again explain the variance in agreement towards certain categorisation strategies - discrimination of melodic music sounds and human action sounds. However, the fact that multiple strategies appear to be used by CIL is likely a result of variations in perceptions of the sounds and cannot be explained by any of the participant variables, for example duration of deafness, duration of implantation etc. More likely this variation comes from differences in day-to-day listening habits, the familiarity of participants with the sounds used and other higher cognitive processes, such as listening expertise and musical ability [6, 66, 77, 54, 90]. Of course even with similar levels of implantation duration different participants will have differ in their levels of adaptation to hearing with the CI device and the development / reorganisation of cortical structures and could be cause for much inter-participant variability of CIL results. In comparison, even though when struggling to identifying vocoded sounds, the NHL auditory system is much more stable and less varying than that of CIL.

4.2.0.3 *The importance of vocal sounds (Dimension 1)*

MCA analysis has shown for ALL participant groups, except for NEW, the CIL user group with lowest duration of implantation, that Dim 1 is related to a distinction of vocal vs. non-vocal sounds. The correlation of Dim 1 values for EXP and NS to the RMS of certain vocal related frequency bands and the distinction of vocal sounds supports this idea and is therefore most likely the first and most important distinction that participants are making.

A similar free-sorting approach in NHL [52] found a separation of sounds corresponding to vocal vs. non-vocal sounds as well as clear categories of animal vocalizations and transport/mechanical sounds. In the present study the same vocal non-vocal discrimination can be made for the more experienced CIL as well, which appears to contradict previous studies that showed difficulties in distinguishing these sounds [34, 96, 45, 65, 89]. [93] also showed that when spectral information is reduced both NHL and CIL users present a strong impairment in discriminating vocal from environmental sounds. Although our results show CIL users perform similarly to NHL previous protocols used a two-alternative forced choice paradigm involving much shorter stimuli whereas in the present study sounds were 2-3 seconds long and participants could listen multiple times such that identifying some stimuli was likely easier.

Dim 1 coordinates (for CIL and NS are also correlated to values of RMS within the frequency band centred at 150 Hz, which overlaps the range of human voice fundamental frequencies - 85-180 Hz for males and 165-255 Hz for females [55]. The highest values for RMS at 150 Hz are seen for vocal sounds (*fem* and *male*) and therefore these cues, important to the perception of vocal sounds, give further evidence that Dim 1 is related to vocal perception. Of course the human voice is also special due to its importance for social communication such that most humans hear speech frequently in their everyday lives and familiarity with vocal sounds is very high [12]. In addition speech and other vocal sounds are produced by a single unique source, the human vocal tract, which reduces the variability of spectra compared to environmental and musical sounds. Such particularity might confer to vocal sounds a specific familiarity feature set used to build-up perceptual strategies and leading to more efficient recognition compared to environmental sounds [61]. Finally with regards to CIL users most research is conducted with the aim of improving speech perception such that implants are designed to deal with these vocal sounds better than musical or environmental sounds.

4.2.0.4 *Pitch and Periodicity cues (Dimension 2)*

Dimension 2 contrasts the musical and environmental sounds for NHL (including vocoded conditions) and the more experienced CIL (INT & EXP). Values of coordinates for NS and EXP participant groups are also correlated with the Max Pitch Salience and Autocorrelation peak, suggesting an importance of specific spectral-temporal information in discriminating these sounds. Similar results have also been found by [52] and [59], whilst [110] concluded that the categorization of environmental sounds was linked to the variation in rate of spectral dynamics. High values of pitch salience correspond to slowly changing spectral cues [134] and may therefore be more easily perceived by CI users who normally perform poorly in pitch related tasks [107, 38, 92]. CI users also retain a certain amount of musical perception when presented with simple rhythms of notes [70, 86]. Musical sounds with temporal movement and a sense of melody may therefore have a stronger sense of musicality than singular tones. This may explain why the violin, guitar and bell sounds are grouped together separately from the oboe and xylophone. In addition difficulties in timbre perception [59] and increased causal uncertainty may contribute to the categorization of the oboe and xylophone as environmental sounds. Finally [111]reported a link between temporal and spectral cues used for both speech and environmental sound perception. Although Dim 2 for EXP is linked with certain spectro-temporal variables there is no correlation to the speech perception scores suggesting that different spectro-temporal information may be important for speech and environmental sound perception. This would suggest that although the perceptual differences originate from low level processing abilities, the effect of higher order cognitive processes cannot be dismissed [121].

Although separating the same types of sounds, dimension 2 for INT is likely based on different information. Rather the types of sound associated with this dimension are human action, and the acoustic information is temporal (amplitude changes) rather than spectral. This mix of information is also associated to Dim 1 for NEW. As discussed in chapter **??**it would therefore spear that lesser experienced CIL do not make use of the pitch salience cues in the same way as experienced CIL, an aspect of auditory perception that improves with time (REFERENCES). However they are able to make use of sounds amplitude changes and information of human action sounds in order to separate these kinds of sounds from others.

4.2.1 *Does vocoding predict CIL performance?*

As previously stated experienced CIL show many similar categorisation strategies to NHS in the natural sound condition. Therefore it is not so surprising that in CI-simulated conditions using 16 and 8 channel vocoder techniques, these similar strategies are again found. For example separating vocal sounds, transport sounds and human action sounds. The agreement of participants to the dimensions is similar between CI-simulated conditions and CIL, however CIL are still more varied in their performance. Results also show that the performance of experienced CIL more closely resemble those in the natural sound condition and show that there is still a difference in the perception of natural sounds by CIL compared to the perception of simulated sounds by NHL. This is possibly due to the fact that NHL in these two conditions were naive to hearing vocoded sounds where as experienced CIL are more familiar to the poorer sound quality delivered by the CI device. It has also been shown that when training NHL on vocoded sounds identification scores are seen to increase, therefore with repeated testing or possibly a paradigm that included a period of habituation for NHL to vocoded sounds, the results of the FST would be more similar to NS. So whilst results in vocoding conditions may look similar to CIL, it may not be a fair comparison and this is especially true for the 16C conditions when considering that CIL often only make use of a limited number (4-7) of frequency channels.

Vocoding studies are useful in assessing how the auditory systems of NHL and CIL compare assuming that the signal passed on from the inner ear to the higher levels of the auditory system is the same. In this manner the fact that results for CIL closely resemble those of both NS and 8C (vocoded condition with least amount of spectral information) likely means that there is accurate functioning of higher order processes concerning categorical knowledge semantic association and familiarity with the sound stimuli. This may not be entirely unexpected when considering that the population of CIL consists of post-lingually deaf adults who have most regained built up these higher order processes (to differing degrees) before their hearing impairment manifested. Similarities between CIL simulations and CIL results show that CIL likely maintain higher order processing used in auditory categorisation of everyday sounds. However there are two issues that limit the extent of this conclusion; 1) NHL listening to CI-simulated sounds does not perfectly simulate the perception of natural sounds by CIL and 2) it is not fully understood how adaptation to CI device may change these higher order process. This broad conclusion is still important especially when exploring the why whilst sound identification is poor for CI, the categorisation identification does not seem to be as strongly deteriorated. Of course the categorisation task may be described as being "easier" than that of an identification task but part of the reason may also be be that higher order processes remain to link the perception of a sound to its associated category.

Whether these processes undergo any changes for CIL during deafness and what these changes may be is difficult/impossible to say. Previous studies have reported the time scale for the restoration of a number of auditory processing abilities ([9, 73, 124]) and so it could well be the same for processes concerning the categorisation of sounds that come "on-line" as the brain adapts to receiving auditory signals again. What is more clear is that the development of auditory categorisation abilities would seem to plateau somewhere between 6-12 months implantation, although this is of course based only on our study which is the first to test CIL using a FST and the first to test the effect of implantation duration.

4.2.2 *Influence of Cortical effects as a result of Deafness*

So far the comparison of NHL and CIL users has only concerned the differences in the perceptual abilities of the participants and how CIL users receive a poorer quality stimulus as a result of the implant processing. In addition another important factor to consider is the influence of brain plasticity and the fact that periods of deafness and subsequent adaptation to using a CIL can lead to differences in brain activation and the roles of different cortical structures [130, 72, 73, 21]. For example in post-lingually deaf persons periods of deafness cause areas of auditory processing to be re-purposed to aid visual processing, must often concerning lip-reading and which leads to better abilities for CIL users than NHL in visio-auditory processing concerning speech-reading tasks [116]. And even after long periods of using an implant, in which auditory performance is recovered, CIL patients retain better visio-auditory processing abilities [116]. This cross-modal reorganisation also results in weak activation of the temporal brain area, which is involved in human voice processing [11, 10, 28].

Whilst this reorganisation is helpful in aiding processing of lip-reading for example, it can also be detrimental to the recovery and adaptation to hearing with a CIL [117]. For example [131] show that minimal reorganisation of the auditory superior temporal gyrus (STG)/superior temporal sulcus (STS) is linked to better outcomes of CIL implantation. Inversely [72] showed that with increased durations of deafness, alternative methods for phonological processing are employed by CIL which results in increased activation of the right posterior superior temporal region and is predictive of poor CIL performance.

A final change that is noted in many studies concerns a change in auditory processing abilities, noted as the *dorso-ventral* dissociation related to the duration of deafness [44, 72]. Dorsal regions associated to the planning and control of language production are seen to be more activated in CIL users with recent deafness, decreasing in activity with duration of hearing loss. Alternatively activity in ventral areas involved in semantic processing and memory retrieval was positively correlated to the duration of deafness. It therefore appears that when processing speech the increased deafness the dorsal/phonological route of processing is not maintained and becomes unused, this then results in an almost "fall-back" strategy involving ventral areas and the use of lexical-semantic processing.

Interestingly [71] also found a similar dorso-ventral dissociation concerning an imagery task of non-speech sounds. They conclude that the dorsal route of processing is also used to maintain representations of non-speech sounds, and is seen to reduce in activation with longer periods of hearing loss where cognitive resources are also likely reallocated to speech processing. Applying this to the current projects' results it may be another reason why the processing of speech sounds seems more accurate/dominant than other everyday sounds as they are maintained to a better level of processing. It would therefore be recommended to maintain the dorsal route of processing for non-speech sounds during periods of deafness/hearing loss to help in the restoration of environmental sounds processing post-post-implantation.

Whilst no study has looked at the cortical changes directly involved the categorisation of sounds it is clear that as a result of deafness there are certain cortical reorganisations or changes in processing strategies that affect the perception of the auditory stimuli used in this study. As explored in chapter **??** this is most likely a factor in the different performance of NEW alongside the adaptation to hearing with the implant.

4.2.2.1 *Methodological Considerations*

Using a Free-Sorting Task and based on solid MCA and HCPC analysis, our results provide evidences that CIL and NHS present categorization strategies that are much closer to normal that what would be expected from the technical limitation of a cochlear implant. However, several limitations prevent to make general conclusions on the capacities of CIL to discriminate natural sounds. Firstly our conclusions are based on a choice of 16 sounds that belong to 3 pre-defined categories (environmental, musical and vocal). The limited number of stimuli was chosen such that the FST could be completed in a short time frame and not be fatiguing for the elderly CIL . However this made it somewhat difficult to equally represent all of the environmental, musical and vocal sounds that exist in everyday life and may have made it easier for CIL to produce results that resemble those of NHL. Testing more sounds in a FST would also be possible but would however take a longer amount of time and constitute a higher cognitive load for CIL participants especially. Apart from testing more examples of environmental, musical and vocal sounds another goal of future testing would be also to diversify the categories of sound to be studied. For example no examples of animal vocalization have been used and the results for NHL also show that categories of linguistic, non-linguistic and transport/mechanical sounds likely exist. Such broader investigation of natural sounds would require either using a different data collection method, raising the difficulty of the FST or informing participants of the categories prior to testing, in similarity to [59].

Secondly the FST was chosen over other methods of data collection in order to minimize the complexity of the test protocol for the participants. However in a recent comparative article on different methods of categorical data collection. Giordano (2010) shows that in comparison to hierarchical sorting and similarity ratings the results of a FST have low reliability (repeatability) when applied to different groups of subjects. It is also claimed that FST is inaccurate in representing the raw performance of individuals. Further the use of MDS models may also not cover all of the variance present in the data and may hide some of the similarities and differences between NHL and CIL . However, we present converging results using additional and complementary methods such as HCPC and the more in-depth MCA. Whilst HCPC is able to display the overall categories in a simple manner MCA allows us to more precisely view the possible strategies of categorization as well as the agreement to these strategies amongst the individual participants. Further, we used the raw co-occurrence (similarity) matrices to perform a model-free analysis. Again, this revealed performance values comparable between NHL and CIL , a result which supports the conclusions made using MCA. Finally the results of categorization accuracy and identification have highlighted that the sound BEL was categorized by most participants as a musical rather than environmental sound. In contrary this sound was initially considered by the researchers to belong to the category of environmental sounds as church bells are often heard outside as part of a complex sound environment. This is an important example to remember when considering the possible sounds and pre-defined categories that may be used in future testing.

4.3 CONCLUSION

The current chapter focussed on the comparison of NHL and CIL in the categorisation of sixteen everyday sounds. NHL were divided into three groups who undertook the task either with natural sounds (NS), or CI simulated sounds using 16 and 8 channel vocoder (16C & 8C). CIL were also divided into three groups based on patients implantation duration, either New cochlear implant listeners (NEW) $0-6$ months, INT $6-14$ months and Experienced cochlear implant listeners (EXP) 14+ months.

The more experienced CIL (INT and EXP) showed similar performance to NHL-NS in the strong separation of vocal stimuli from other sounds, however EXP were the only group of CIL who made use of spectral cues (pitch salience) when categorising melodic musical sounds together,and temporal cues (mean autocorrelation function peak) when categorising environmental sounds together, again in similarity to NHL-NS. The least experienced CIL group, NEW, showed weaker ability when separating vocal sounds and more prominent was the focus on human action sounds that contained fast changes in amplitude (peaks). All results suggest that categorisation is based predominantly on the semantic information associated to the sound producing event, object or action. It appears that CIL also rely on semantic perceptions, however certain acoustic variables (mentioned above) shed light on certain types of sounds that may be more easily perceivable for CIL. NHL in CI-simulations however maintained categorisation based on semantic information even at the poorest quality signal using 8 channels. In comparison to CI simulations, CIL performance was more varied suggesting the use of more categorisation strategies. However the dominant categorisation strategies were closely related indicated that in the more experienced groups of pre-lingually deafened CIL's, the higher cognitive functions associated with auditory categorisation were maintained. Overall results show that experienced CIL are categorise sounds in the same manner as NHL, however with more inter-participant variation that appears linked to the duration of implantation.

Key Points

- Similar strategies shared by ALL NHL groups and more experienced CIL user groups with implantation greater than 6 months. These include dominant strategies based on the separation of vocal sounds for all cases bar that of New cochlear implant listeners (NEW). However the importance of vocal stimuli is still present for these participants.
- Categorisation strategies are based on semantic or causal perceptions of sounds for all participants. It does not appear that any participant group is making categories based on perceptions of the physical characteristics of the sounds, even though these characteristics may be important for the perception of sounds. For example certain temporal (autocorrelation function peak, and amplitude peaks) and spectral (pitch salience) are of importance for sounds that are more easily perceived by CIL
- There is much more inter-participant variability in the results of all CIL groups and is independent of implantation duration. In contrast NHL, NHL tested with CI simulated sounds, show stronger inter-participant agreement to the use of categorisation strategies. This is likely a result of the varying auditory abilities of CIL participants and possibly the greater variation in age found with CIL.
- Similarities of results between CI-simulations and CIL show that higher order cognitive functions associated with auditory categorisation are maintained in experienced post-lingually deafened CIL.

5

CATEGORISATION OF ENVIRONMENTAL SOUNDS AND THE EFFECTS OF CONTEXT

Previous chapters looked at the categorisation of a set of everyday sounds that included vocal, musical and environmental sounds. Whilst significant literature exists on the study of musical and vocal sounds, especially regarding the abilities of CIL, environmental sounds have been focussed upon less. These sounds, which are important in the daily lives of listeners, are also normally heard in specific locations or contexts and this aspect of environmental sound perception (ESP) has also not received the attention it probably deserves. The following chapter therefore seeks to add to the existing research on the perception of ES by using a FST of 20 sounds which were chosen to reflect different contexts - Bathroom, Kitchen, Exterior and Office. Not only is ESP assessed, but the hypothesis of whether categories can be based on purely the context is also tested. Additional testing steps using CI simulated sounds (4-channel vocoder) also look at whether the absence/addition of context information can aid the identification performance of the same 20 sounds. Results will show that categories appear to be based on some contexts, however this may actually reflect the type of activity associated with the context. Other strategies used in categorisation reflect previous theory and are based on water sounds, machine/transport sounds as well as the perception of the sound producing action. Analysis of identification performance shows that in the natural sound condition the addition of context information improves identification performance. In the CI simulated condition this effect was not seen, although this may be heavily influenced by the high task difficulty of perceiving the sounds with only 4 frequency channels.

5.1 INTRODUCTION

Whilst the previous chapter looked at a mix of vocal, musical and environmental sounds, an additional goal of the current work was to aid the understanding of the perception of only environmental sounds (ES). As discussed in chapter 1 the study of environmental sound perception (ESP) has fallen behind that of speech and music sounds, for example auditory categorisation has mainly focussed on speech processing and phoneme categorisation. In addition ESP studies have often looked at single details of perception, for example the size, shape or material of simple objects [68, 62, 46, 43] or the identification of simple sound-producing actions [143].

As previously stated in chapter 1 identification and categorisation are two processes very closely linked when it comes to auditory stimuli. Identification of a sound source leads to the ability to associate an auditory stimulus with higher forms of conceptual knowledge often used in categorisation. Understanding the categorisation of ES also suffers from lack of research however some studies have used tasks involving categorisation/sorting and similarity in order to study ESP using different sounds.

In a free-classification test of 60 environmental sounds, [57] found that participants classified the sounds very differently without a clear consensus. Lexical analysis did however reveal four broad categories of solids, liquids, gasses and machines, which the authors concluded corresponded to Gavers original hierarchy [37], although differed to other models [33]. Sounds were also classified based on the sound source first and secondly the sound-producing action.

Another free-classification task [91] of 120 sounds included environmental, human (vocal and other) and musical sounds. Judges evaluated participants responses and arrived at a total of 23 similar categories which overall corresponded to sources (e.g. transport, animals, humans) locations (household, kitchen) or abstract themes (warning/alarm, hygiene). Source categories also included some based on the material of the source (paper, water). There were also clear categories of musical sounds and human sounds, which included such sub-categories as sleeping and sickness sounds.

[113] performed two categorisation tasks looking at the specific difference of task instruction. Using the same xxx sounds subjects were first asked to group them based on *"sounds that are together in the environment"* then secondly, a different set of participants were asked to categorise the sounds based on *"sounds that were acoustically similar"*. Categories that appeared in the first test showed corresponded to sounds inside a house, transportation and animal sounds. Based on the acoustic experiment categories were based on similarities in rhythm, pitch and amplitude modulations. In similarity, [140] found that participants subjects categorised 20 environmental sounds based on similarities of the sound-producing event or of acoustic similarity. Finally when categoris-

ing soundscapes (not individual sounds) [48] found that they were categorised based on the activity present - cafés, markets and parks.

Clearly these studies show that there are multiple categories that can exist within the encompassing title of environmental sounds. Whilst participants may use many different descriptions for naming these categories there exist overlapping similarities in the kinds of categories that are formed. These include machines, transport, kitchen/house, animal sounds, human sounds. It seems therefore that above these category titles important perceptions concern the sound-producing *object*, sound-producing *action*, location (and attached activity) as well as emotional content (i.e. warning sounds).

Therefore these studies show that the categorisation of ES can be based on different information and different perceptions. Gaver eluded to this when he described two different listening modes which he defined as musical and environmental listening. [77] offered a slightly different way of thinking that may be more applicable, in that three different kinds of similarities are used when grouping environmental sounds together:

1. **Acoustical:** relating to the perception of physical acoustical attributes of a sound, for example the pitch, temporal patterning or loudness.
2. **Causal:** similarities associated with the sound producing event i.e. the object or action that generates the sound, for example the kind of impact, the material (water, metal, wood).
3. **Semantic:** similarities of associated meaning with the sound. This could involve similarities in the associated location, event or activity, emotion or even other more abstract information.

Apart from helping to inform the grouping of sounds, the location or context within which sounds are heard or associated can also affect the identification of sounds [75, 8, 91, 50, 101]. Most studies of ESP do not take this into consideration, which is odd when considering that may ES's only occur in specific contexts (locations or events), or simply that auditory events in the real world always take place within an auditory environment that can affect perception, for example via masking (Gygi 2007).

The addition of context information has been seen to have differing effects. In a test where context was created by a sequence of sounds, target stimuli that were congruent with the sequence were not identified any greater than in an isolated condition [8]. [50] however showed a 5% increase in detection for sounds incongruent with an auditory scene, suggesting that incongruent contexts can help listeners perceive sounds. The authors described this as a "pop-out" effect such that incgonruent sounds appeared to pop-out or stand out from the simultaneous auditory context, where as congruent sounds became part of the surrounding context and therefore more difficult to detect. The addition of context has not however been tested with CI simulations in order to clarify whether or not it could be of help for Cochlear Implanted Listener (CIL), for whom it is known suffer from deficits in ESP. Therefore (as listed in the box below) one aim of the current chapter is to see if the addition of context information could aid the identification of environmental sounds under CI simulation using a 4 channel vocoder. In similarity to categories chosen by [111] four categories corresponding to Bathroom, Kitchen, Exterior and Office were chosen for testing.

Information concerning the auditory context (location or environment) can also be used to form categories of sounds, as mentioned above. However it is not known if different contexts form categories more strongly than others. The use of the four different contexts as mentioned above, therefore provided an opportunity to test whether participants would choose to categorise a set of environmental sounds based on the context information whilst providing additional information in the study of environmental sound perception. A summary of the chapter aims is written below in the shaded box.

Chapter Aims

1. To further the understanding of auditory categorisation concerning ONLY environmental sounds. Whilst research has been carried out on this topic, the field is still not vast and could benefit from additional studies. For example how strongly are auditory categorisation strategies based on the environmental context in which sounds are heard? Sounds are therefore taken from four specific contexts (Bathroom, Kitchen, Office and Exterior) in order to see whether this information is used by listeners as part of their categorisation strategies.

2. To test the categorisation performance of NHL with CI simulated environmental sounds to investigate I) how NHL perform in adverse listening conditions i.e. what kinds of perceptions are possible and how do these perceptions then form the basis for forming categories II) how CI users may perform in such a task. To study the categorical perceptions of NHL to understand whether or not they might prove useful in improving the perception of environmental sounds by CI users.

3. To test whether or not the addition of environmental context information can help listeners to more accurately identify environmental sounds. In the real-world sounds are always heard within a specific context, which can affect the ability of a listener to detect or identify sounds, for example in aiding the detection of incongruent sounds [53]. Most studies of environmental sound perception do not take into account the influence of the environmental context and so it is important to try and make steps to understanding this. This is especially important for understanding whether context information can help the identification of CI simulated sounds and therefore be used to aid CI users environmental sound perception in future research.

5.2 METHOD & MATERIALS

In order to test the above hypotheses, four separate testing steps were established and are described in section 5.2.2. In each Step a separate group of normal hearing participants was recruited, summarised in table 5.1 and the same 20 environmental sounds were used throughout all testing and are described in table 5.2. The four testing steps are also listed directly below such that the reader is informed before reading ahead.

- **Step1-Typ-Typ** - typicality ratings of the stimuli as to whether they were typical of the designated context.
- **Step2-FST-NS** - Free Sorting Task (FST) of stimuli in Natural sounds (NS) condition to test categorisation strategies of environmental sounds.
- **Step3-FST-SIM** - again a FST however sounds were passed through a 4 channel vocoder in order to simulate listening with a CI
- **Step4-FC-SIM** - Forced categorisation of CI simulated sounds as used in Step3-FST-SIM

5.2.1 *Participants*

For each Testing step a different group of participants were tested so that there was no overlap between testing. Table **??** details the the four different groups of participants:

5.2.2 *Stimuli & Procedure*

For all four testing steps the same 20 everyday/environmental sounds were used and are listed in table 5.2, with abbreviated ID used throughout the results and discussion, a small description of the sound and also the corresponding Environmental Context that each sound was chosen to represent. Stimuli were taken from the database of sounds owned by PETRA at Université of Toulouse III - (Mirail). Stimuli were selected to represent four different contexts that would be common to participants daily lives with the choice also inspired from

Testing stage	Task	N (n males)	Age (mean ± std)	No. French native
Step-1-Typ-NS	Typicality Rating, NS	9 (4)	29 ± 5	7
Step-2-FST-NS	FST with NS	16 (9)	27 ± 4	13
Step-3-FST-SIM	FST with CI-simulated sounds	18 (8)	30 ± 6	12
Step-4-FC-SIM	Forced Categorisation with CI-simulated sounds	17 (3)	28 ± 7	12

Table 5.1: Participant information for Chapter 5 showing the number of participants, with the number of males shown in brackets, age (mean and standard deviation) and the number of french natives.

[111] who used four categories of General Home, Kitchen, Office and Outside. Based on the stimuli available it was decided to alter the General Home to Bathroom. This also eliminated any ambiguity that could arise from Kitchen sounds also being associated to General Home sounds. For each testing step a a slightly altered procedure was used, and they are subsequently described below:

Step1-Typ-NS: Typicality ratings

In order to judge if the chosen sounds were typical of the corresponding contexts (see table 5.2) 9 NHL rated the typicality of each sound. Participants were asked to judge whether each stimuli sounded as if it belonged to the corresponding context on a scale of 1-10, where 1 was rated as *not at all typical* and 10 *highly typical*. Sounds were listed by context as in table 5.2 such that participants knew directly the which sounds came from which context.

Step2-FST-NS: FST in Natural Sound (NS) condition

Using the same FST as described above participants were asked to group sounds together on the basis of their personal preferences. With the only instruction that they should group together those sounds which sounded as if they belonged together. Following completion of the categorisation, participants were asked to verify choices by listening to each sound once again. Once this had been accomplished and any final changes made, participants were asked to describe each of their categories before finally being asked to identify each sound. Whilst undertaking this they were permitted to listen to the sounds multiple times again.

Step3-FST-SIM: FST in CI simulated condition

Using exactly the same procedure as described for Step2-FST-NS, however this time the 20 sounds were passed through a 4-channel vocoder in order to simulate CIL perceptions. Only 4 channels were used for vocoding based on [93] - where vocal-environmental sound discrimination was found to be comparable between experienced Cochlear Implanted Listener (CIL) and NHL listening to sounds that had been vocoded with 4 frequency channels. Following completion of the categorisation task participants were asked to describe both their categories and sounds as in Step2-FST-NS.

Step4-FC-SIM: Forced categorisation in CI simulated condition

Instead of using the FST paradigm, participants were presented with a written list of the four contexts and asked to categorise the sounds into these four categories i.e. Bathroom, Kitchen, Exterior and Office. Following the categorisation task they were asked to identify each category and provide a short description of the sounds as in the previous two testing steps.

5.3 RESULTS

Results and analysis of the four testing steps are detailed below. Typicality measures are briefly covered and show variance across the stimuli, which reflects real world perceptions of sounds and means that categorisation strategies are not strongly biased in favour of the context information. MultiCorrespondence Analysis (MCA) and Heirarchical Clustering based on Principal Components (HCPC) analyses are again used as in previous chapters to assess the categorisation strategies used by participants in Step2-FST-NS and Step3-FST-SIM. Results show similarities to previous work regarding categories that reflect water sounds, machine/transport sounds and categories related to the perception of the sound-producing action. Perceptions of the sound producing action also remain robust in the CI simulated condition. Identification performance is assessed across all conditions as the percentage of participants that correctly identified each sound, and this is analysed for vocoded vs natural sounds as well as the presence or absence of context information. Finally correlations are made between the results of MCA with - acoustic measurements of sounds, iden-

ID	Sound description	Environmental Context
shv	Shaving	Bathroom
shwr	Shower	
crt	Shower curtain	
snk	Sink	
tth	Toothbrush	
tub	Bathtub	
cer	Cereal	Kitchen
brd	Bread	
egg	Egg beating	
ovn	Oven	
plt	Plate	
bskt	Basketball	Exterior
rain	Rain	
tctr	Tractor	
svl	Shovel	
lght	Lightning	
ppr	Paper folding	Office
cpy	Photocopier	
scsr	Scissors	
tpe	Tape	

Table 5.2: Description of 20 environmental sounds used to test the effect of context in Chapter 5. Stimuli are arranged into four auditory contexts. ID tags are also provided and are used throughout the following chapter to reference the stimuli.

tification performance, participant information (age) and typicality measures.

5.3.1 *Typicality*

The results of Step-1-Typ-NS are shown in figure 5.1 and were recorded with 9 participants. Error bars display the Standard Deviation and show that some sounds were not perceived as being very typical for their corresponding context - these sounds, *shv, crt* and *ovn* all have mean values lower than 5.

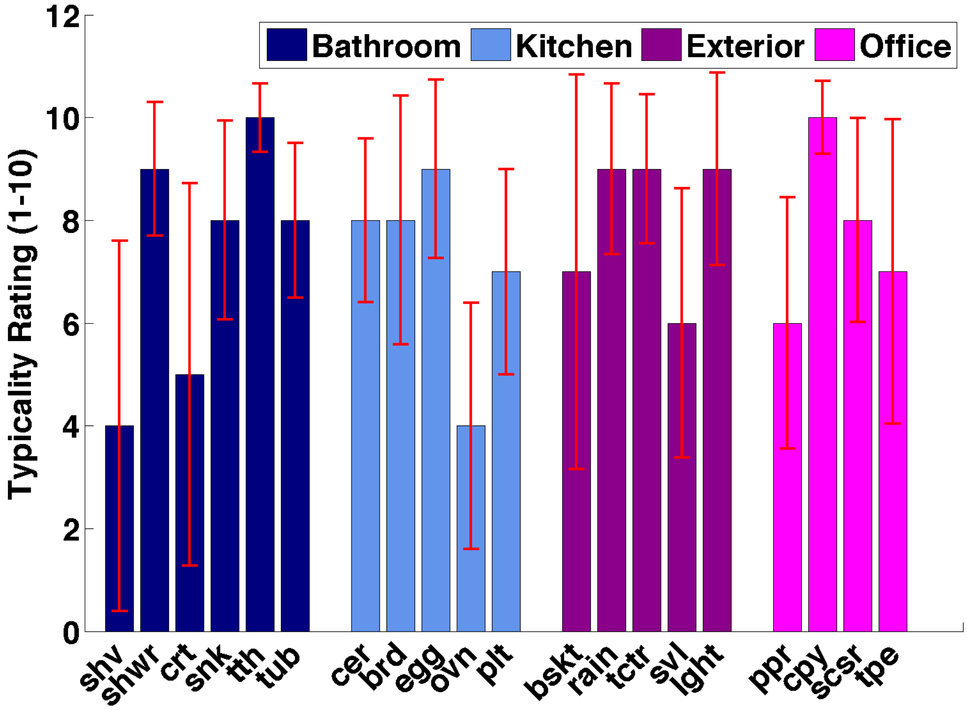


Figure 5.1: Typicality measurements for the 20 environmental sounds (detailed in table 5.2 for the 4 different contexts; Bathroom, Kitchen,Exterior,Office. Standard error is shown by red error bars.

5.3.2 *Step2-FST-NS - Free-Sorting-Task of Natural sounds*

In order to analyse the categorisation performance the same analyses conducted in the previous two chapters was again carried out. Beginning with the aim of establishing how the 20 environmental sounds were categorised figure 5.2 shows the dendrogram output of HCPC for Step2-FST-NS. As can be seen participants put together all kitchen sounds apart from the oven, *rain* and *lght* are grouped with the bathroom sounds apart from *crt*, the remaining exterior and office sounds are mixed into two categories whilst *bskt* is separated individually. This initially suggests that categories may be created based on the contexts of kitchen and bathroom, but not for the exterior or office. That these latter two contexts are not clearly represented may arise because; unfamiliarity of participants with office sounds; the likelihood that the exterior context incorporates many more sounds and is not as narrowly defined as other contexts.

MultiCorrespondence Analysis (MCA) was also used to analyse the potential categorisation strategies in more detail. For Step2-FST-NS, a total of 5 dimen-

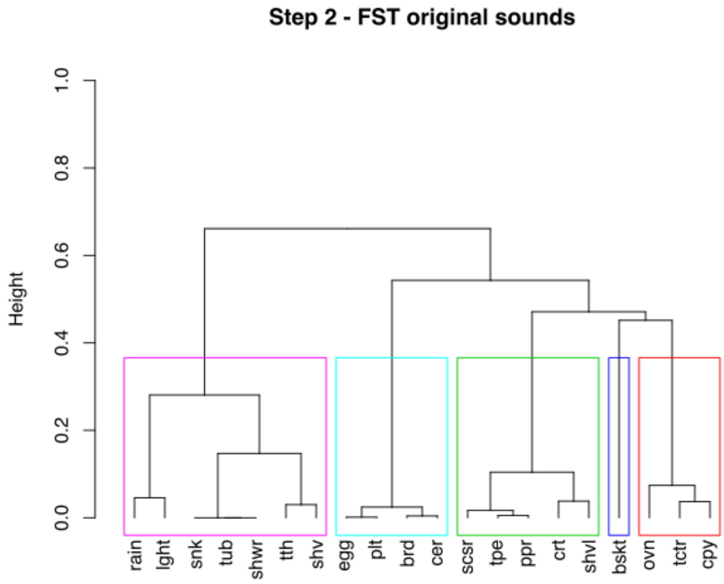


Figure 5.2: Dendrogram for **Step2-FST-NS**, the overall categories outputted by HCPC are indicated by coloured rectangles whilst the upper limit of these rectangles indicates the point at the dendrogram has been cut (as described in the main text). The height axis gives the perceptual distance between each stimulus whereby a large height indicates that participants deemed those two stimuli to be highly dissimilar and vice versa. Finally stimuli are labelled using the abbreviated sound IDs from Table 5.2.

sions were retained in the MCA following the criteria that dimensions should cover 8% or more of the total variance. Figure 5.3 displays these 5 dimensions and from interpretations each can be explained as the following:

- **Dim 1:** would appear to contrast the sounds belonging to the bathroom context with all other sounds. However the stimuli at the extreme far left all involve *water* as a sound producing medium, which may be more important to the perception of participants.
- **Dim 2:** Contrasts sounds belonging to the kitchen context, with the sounds of the photocopier (*cpy*) and tractor (*trct*).
- **Dim 3:** Separates specifically basketball sound from all others, suggesting something different/important with regards to this sound.
- **Dim 4:** Contrasts *trct* and *cpy* with sounds involving the manipulation of an object.

- **Dim 5:** Contrasts the lighting sound (*lght*), which is very isolated, with sounds generated by manipulating an object, for example the toothbrush (*tth*)and basketball (*bskt*) sounds.

MCA results therefore agree with those of the dendrogram in figure 5.2 in that they suggest participants may indeed categorise some sounds on the basis of the context in which they are often heard, which appears true for the Bathroom and Kitchen sounds. However this is not conclusive and is discussed in more detail in section 5.4. The sounds *cpy* and *trct* seem strongly grouped together along with *ovn* (shown in fig 5.2). These sounds all come from different context suggesting a different strategy of categorisation. As they are also the only mechanical sounds in the stimuli set this may therefore be a more likely reason why they are grouped together.

The pattern of results on figure 5.4 show that 75 % of participants have coordinate greater than 0.8 for Dim 1, and 41% for Dim 2. A large majority of participants are therefore following the strategy described by Dim 1, such that either Bathroom or Water sounds are quite easily grouped together.

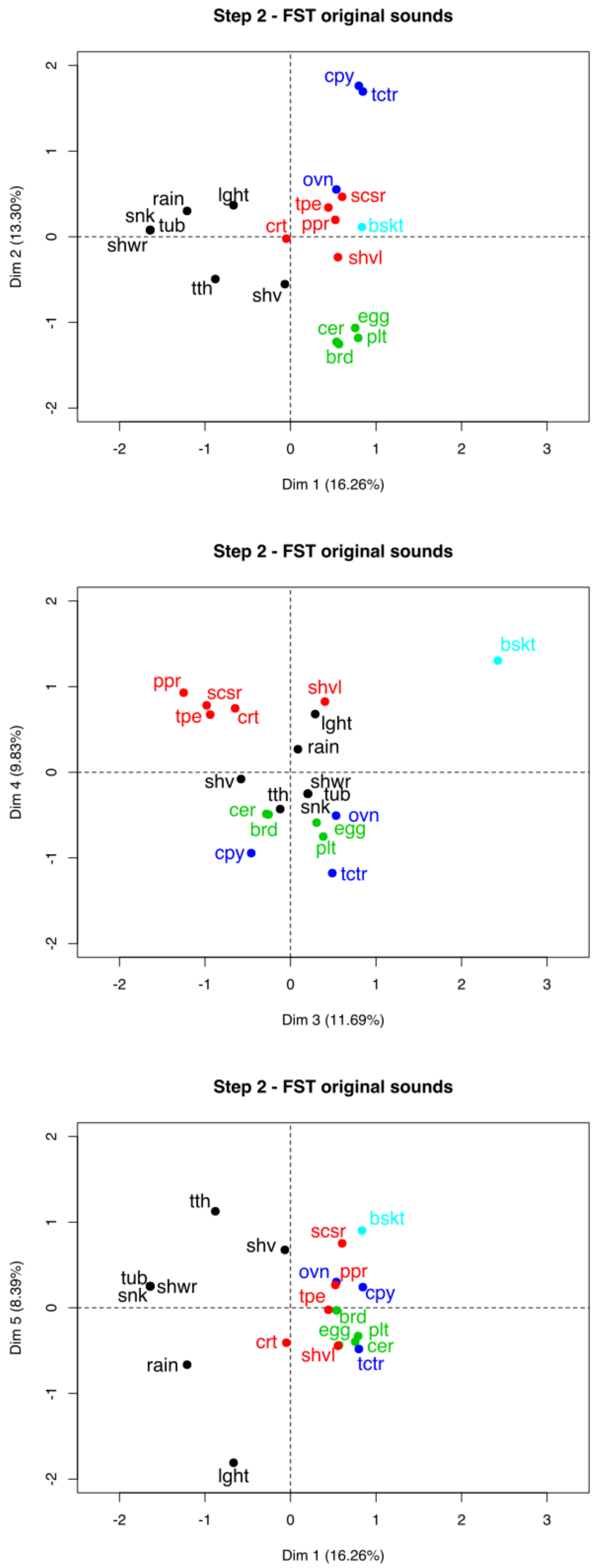
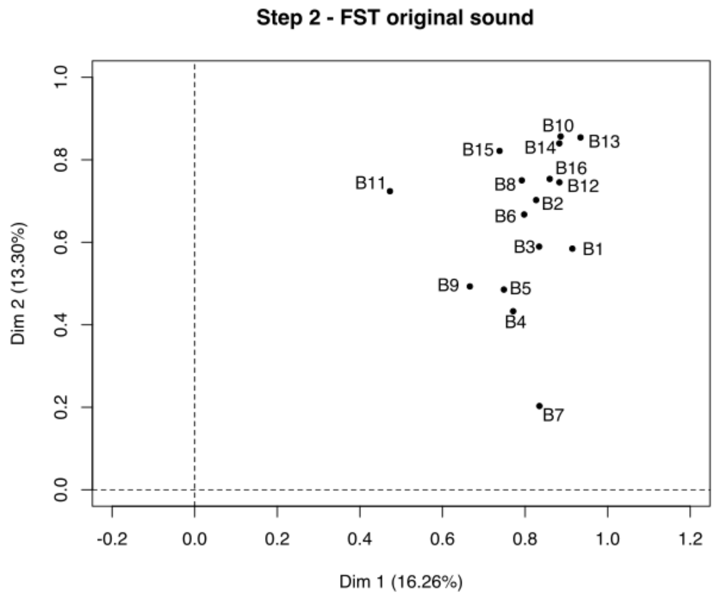


Figure 5.3: Distribution of sounds across the 5 retained dimensions, with Dim 1 & 2 in the top panel, Dim 3 & 4 in the middle panel and Dim 1 & Dim 5 in the bottom. Stimuli are plotted using the ID labels from table 5.2 whilst colours are the same as used for the dendrograms in figure 5.2 in order to show the link between the MCA and HCPC analyses.



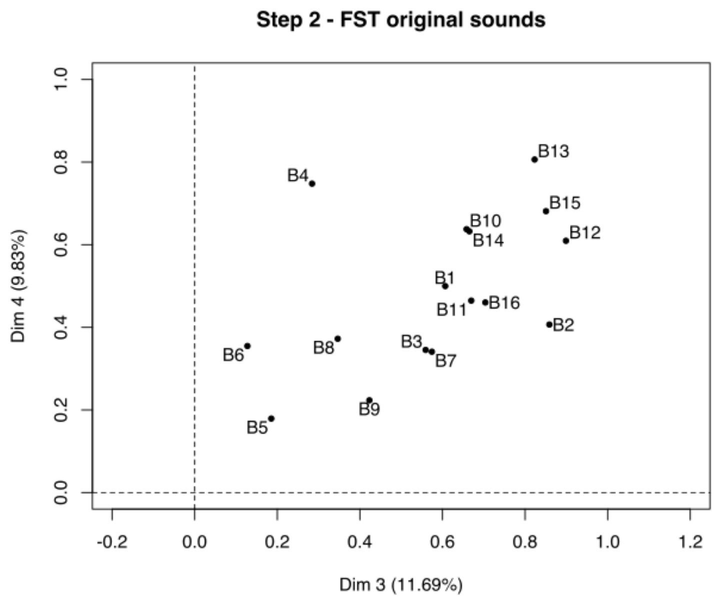


Figure 5.4: Distribution of participants across the retained dimensions for **Step2-FST-NS**, with Dim 1 & 2 in the top panel, Dim 3 & 4 in the lower panel.

5.3.3 *Step3-FST-SIM - Free Sorting Task of CI simulated (Vocoded) sounds*

Step3-FST-SIM involved the same process as Step2-FST-NS, however using sounds that had been vocoded with 4 frequency channels. The dendrogram for Step3-FST-SIM (figure 5.5) does not show categories that resemble any of the the four contexts (Bathroom, Kitchen, Exterior or Office), which is in contrast to the results of Step2-FST-NS, figure 5.2. However the dendrogram does show water-related sounds strongly separated from other sounds at a height of 0.712 on the ordinate. A similar category of water sounds is also seen in figure 5.2 however this separates at a lower height of 0.661 on the ordinate. Within these categories of water-related sounds the INTRA category distance is also smaller in Step-3-FST-SIM (0.0036) than in Step2 (0.18). This means that sounds perceived as water related are more similar in the vocoded condition. This is most likely because in the natural sound condition participants were able to more easily distinguish the different kinds of sounds. However it is still very interesting that water sounds remain a robust category even after vocoding.

The remaining sounds seem to be randomly categorised together and are note easily interpreted from only the dendrogram, although the grouping of *cpy* and *trct* together is again evident and may again represent a category of mechanical sounds in similarity to Step-2-FST-NS results. Figure **??** also shows a large majority (66%) of participants are strongly using Dim 1, where as following dimensions are much lower, for example only 17% for Dim 2. This would suggest that whilst Dim 1 is commonly shared categorisation strategy, following this there are no other strategies held by a large number of participants, who instead likely use multiple strategies based on the difficulty and poor agreement in perceiving the vocoded sounds.

Results of MCA analysis for Step-3-FST-SIM are shown in figure 5.6

- **Dim 1:** separates water based sounds from others, especially the group of sounds that involve human actions. For example *brd, tth* and *plt* all described as rubbing/scratching/cutting.
- **Dim 2:** separates the mechanical sounds from all others. Also included here, but to a lesser degree are *snk* and *ovn* which were described as a door, something opening. Whilst the perception is not 100% mechanical, it gives an idea that there was a level of perception similar to trct and cpy.

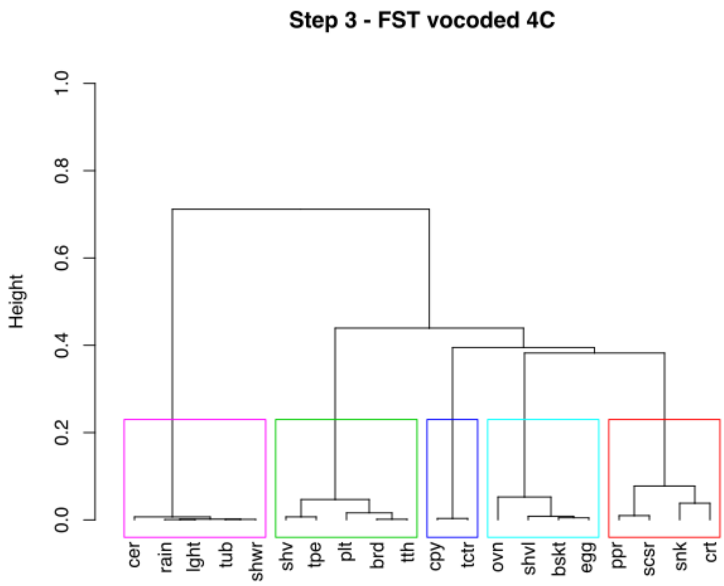


Figure 5.5: Dendrogram for **Step3-FST-SIM**, the overall categories outputted by HCPC are indicated by coloured rectangles whilst the upper limit of these rectangles indicates the point at the dendrogram has been cut (as described in the main text). The height axis gives the perceptual distance between each stimulus whereby a large height indicates that participants deemed those two stimuli to be highly dissimilar and vice versa. Finally stimuli are labelled using the abbreviated sound IDs from Table 5.2.

- **Dim 3:** contrasting *shv & tpe* vs. *bskt* & egg. From comments may be contrasting two different types of action sound - hitting/cutting vs. scratching/rubbing.
- **Dim 4:** separating *crt*, *scsr & ppr* from the rest. Again comments indicate that these sounds referred to shared properties such as *paper, branches, turning a page/curtain*

Whilst initial interpretations are difficult only using figures 5.5 & 5.6, the use of participants comments helps greatly. They can give an insight into what participants were perceiving, which perceptions are shared between sounds and ultimately what shared information might be being used to form categories.

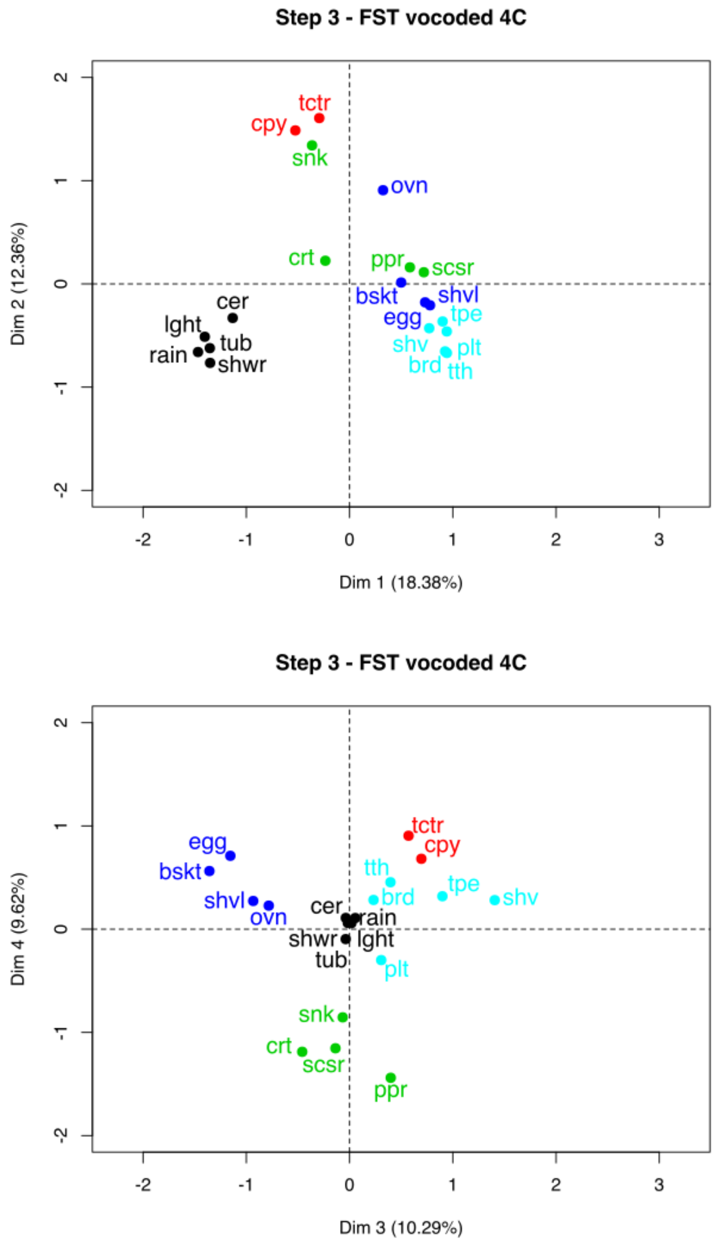


Figure 5.6: Distribution of sounds across the 5 retained dimensions, with Dim 1 & 2 in the top panel, Dim 3 & 4 in the middle panel and Dim 1 & Dim 5 in the bottom. Stimuli are plotted using the ID labels from table 5.2 whilst colours are the same as used for the dendrograms in figure 5.2 in order to show the link between the MCA and HCPC analyses.

5.3.4 *Step4-FC-SIM - Fixed Categorisation of CI simulated (Vocoded) sounds*

Results for Step-4-FC-SIM must be handled differently due to the difference in testing methodology that was used, whereby participants were asked to categorise the sounds into 4 categorises titled Bathroom, Kitchen, Exterior and Office. Hence the results show 4 categories in both the dendrogram and factor maps of figures **??** & 5.7. Also the latter MCA dimension are insignificant to the analysis as Dim 1 & Dim 2 suitably divide the stimuli into the 4 categories. The interesting point concerning these results is to analyse how participants choices correspond to the original categories and how the perception of the vocoded sounds was changed from Step-3-FST-SIM to Step-4-FC-SIM.

The following text summarises the results of HCPC and MCA analysis in terms of the original four contexts of Bathroom, Kitchen, Office and Exterior to determine how accurately each is represented.

Bathroom: consisting of the sounds *rain* and *shwr*. Interestingly only two sounds have been grouped together as Bathroom sounds, both of which were categorised by 67% of participants as such. Obviously both of these sounds reflect very similar events - droplets of water falling and impacting on a hard surface. Also interesting is that the second most used category for *shwr* was Kitchen, whilst for *rain* it was Exterior - so even in the case where sounds are severely degraded by the vocoding there is still a difference in perception (for some participants) relating to whether or not the sounds take place inside or outside.

Regarding the categorisation of the other original Bathroom sounds, the lower panel of figure 5.8 shows that that correct categorisation for all the Bathroom sound was only 32% and only the sound *shwr* was correctly categorised above 50%.

Kitchen: the created Kitchen category contains the sounds *cer, brd, tth, egg, shv* and *plt*. The original Kitchen sounds are all located in this group, except for *ovn* - which was correctly categorised by 0% of participants and was instead split roughly $50-50$ between Bathroom and Office. Overall the original Kitchen sounds were on average correctly categorised by 51% participants. Concerning the created category of Kitchen sounds, the perception of *cer, brd & egg* are quite stable, citing "cutting, pouring grains, bread, egg beating". However *egg* was also perceived numerous times as the sound of a keyboard being tapped and therefore had lower category identification of 50% compared to over 80% for the other two sounds. The results for *tth & shv* are more varied with category identification of Kitchen only at 33 & 28% and perceptions of these sounds are also very varied (see Appendix C). Although perceptions linked to Kitchen include"peeling, cut, grains, bread, seeds".

Office: the sounds *ppr, scsr* and *tpe* constitute the created category of Office sounds and all have high category identification (67, 83 & 83% respectively). This also leads to the original Office sounds being correctly categorised by a large percentage of participants, 65% from figure 5.8. Only the original Office sound of *cpy* is poorly categorised, most commonly as an Exterior sound (39%) which is likely due to its perception as a mechanical/car sound. The perceptions of *ppr & scsr* are associated with paper or the manipulation thereof, which makes sense as to why they should be categorised as Office rather than any of the other three categories. The sound*tpe* is also often identified as "scotch" with other perceptions detailing the sound producing action e.g. "sharpen" or "pull".

Exterior: this category holds the most sounds - *tub, ovn, crt, cpy, bst, lght, trct, snk* and *shvl*. A likely reason for this is that the category of Exterior is much broader than the other three and encompasses many more sounds. Bathroom, Kitchen and Office might generate very specific and limited options of sounds, such that when encountered with ambiguous sounds it is more simple for participants to dismiss these three categorise and instead place sounds into a group of Exterior sounds. Figure 5.8 also shows that the original Exterior sounds were correctly categorised on average by 66% participants. With the sounds *bskt, lght* and *trct* all being correctly categorised by above 80% of participants. Interestingly for these three sounds, the perceptions recorded by participant descriptions remain quite stable over the three testing conditions (see Appendix ???). Figure 5.8 also shows that the average category identification is greatest for Exterior sounds - 65%.

The discussion of results above is however true only for the results of the entire group of 17 participants. There is in fact no one single participant who follows exactly the categorisation detailed by the dendrogram in figure 5.7. With regards to the original categories detailed in table 5.2, the lower panel of figure **??** details the percentage of sounds that were correctly categorised by each of the 18 participants was on average $50 \pm 12\%$, with maximum and minimum performance values of 75 and 30% respectively. In comparison to the categories described by figure 5.7 the agreement of participants to this are roughly the same - $56 \pm 9\%$.

The subject map of figure 5.7 (lower panel) also highlights the fact that these categorisation strategies are a result of the category choices of all participants. Only one participant has coordinate above 0.8 for Dim-1, and there are none for Dim -2. This would suggest that categorising these sounds, bearing in mind they are vocoded, on the basis of context is not easily intuitive for the participants in Step-4-FC-SIM. Rather some sounds.

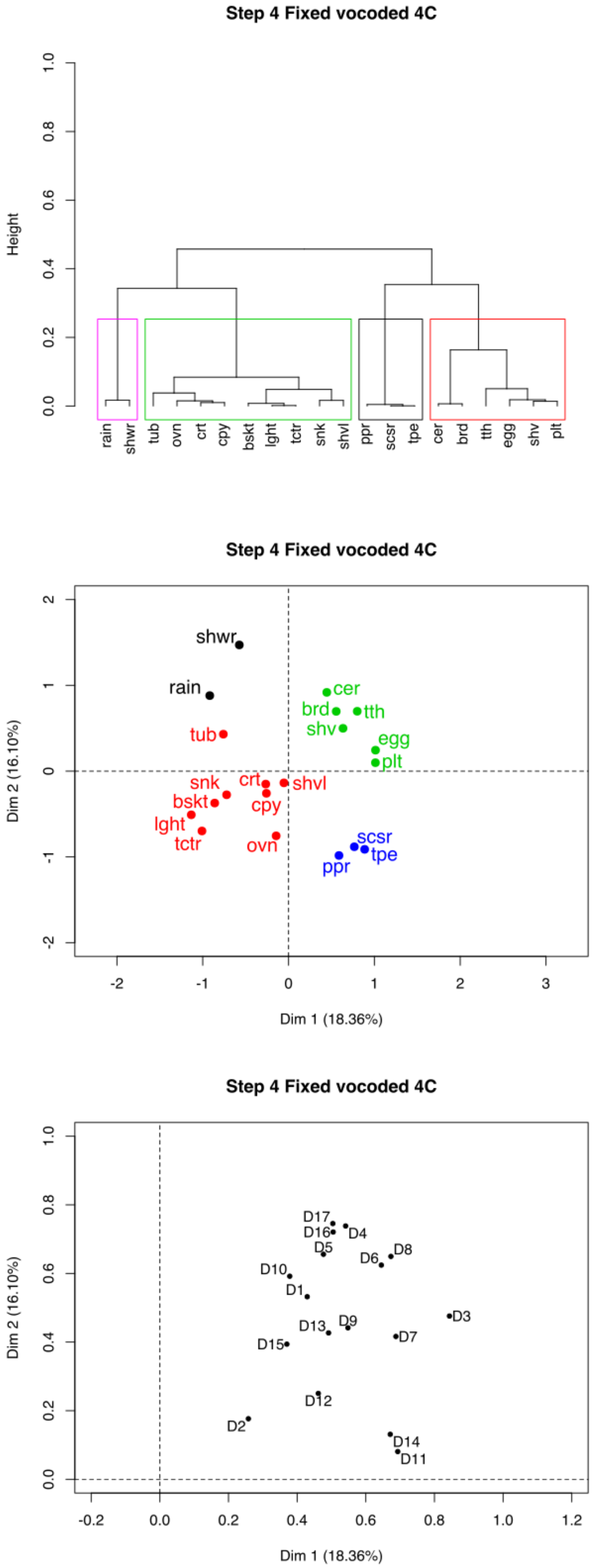


Figure 5.7: MCA and HCPC analysis for **Step-4-FC-SIM**: upper panel shows the HCPC dendrogram output of results, with coloured rectangles representing the final categories; middle panel shows the MCA factor map output covering Dimensions 1 & 2; lower panel MCA subject map across Dimensions 1 & 2.

5.3.5 *Identification of Categories and Sounds*

The second aim of this chapter 5 was to investigate whether the addition of the context information could aid the identification of the 20 environmental sounds. At each testing step participants were asked to give a description of each sound that was later evaluated by the experimenter as to whether or not it correctly identified the sound. Results are shown in figure 5.9 in three different panels.

In Step-1-Typ-NS, rating the typicality of sounds, participants were aware of the association of each sound to the 4 contexts. In this way the results represent a baseline performance when context information is present, and it is of no surprise therefore that the average identification performance is highest for Step-1-Typ-NS, 72% correct , in the lower panel of figure 5.9. ANOVA testing shows this difference to be significant in comparison to all other testing Steps ($p < 0.05$) whilst the results of Step-2-FST-NS (mean $51\% correct$) are also significantly greater ($p < 0.05$) than those of Step-3-FST-SIM (16% correct) and Step-4-FC-SIM (20% correct). This is of course not totally surprising that the sounds in natural condition were better identified than when vocoded using 4 channels. The results of upper panel of figure 5.9 therefore shows that the addition of context information helps the identification of the sounds, that is except for the sounds *shwr, lght* and *ppr* where performance is greater in Step-2-FST-NS. The biggest difference is for the sound *shwr* which in Step-1-Typ-NS was also alternatively described as “water from a tap” or “running water”, both of which were deemed incorrect due to not being specific enough. It may simply be that causal uncertainty of this sound is the reason for the difference.

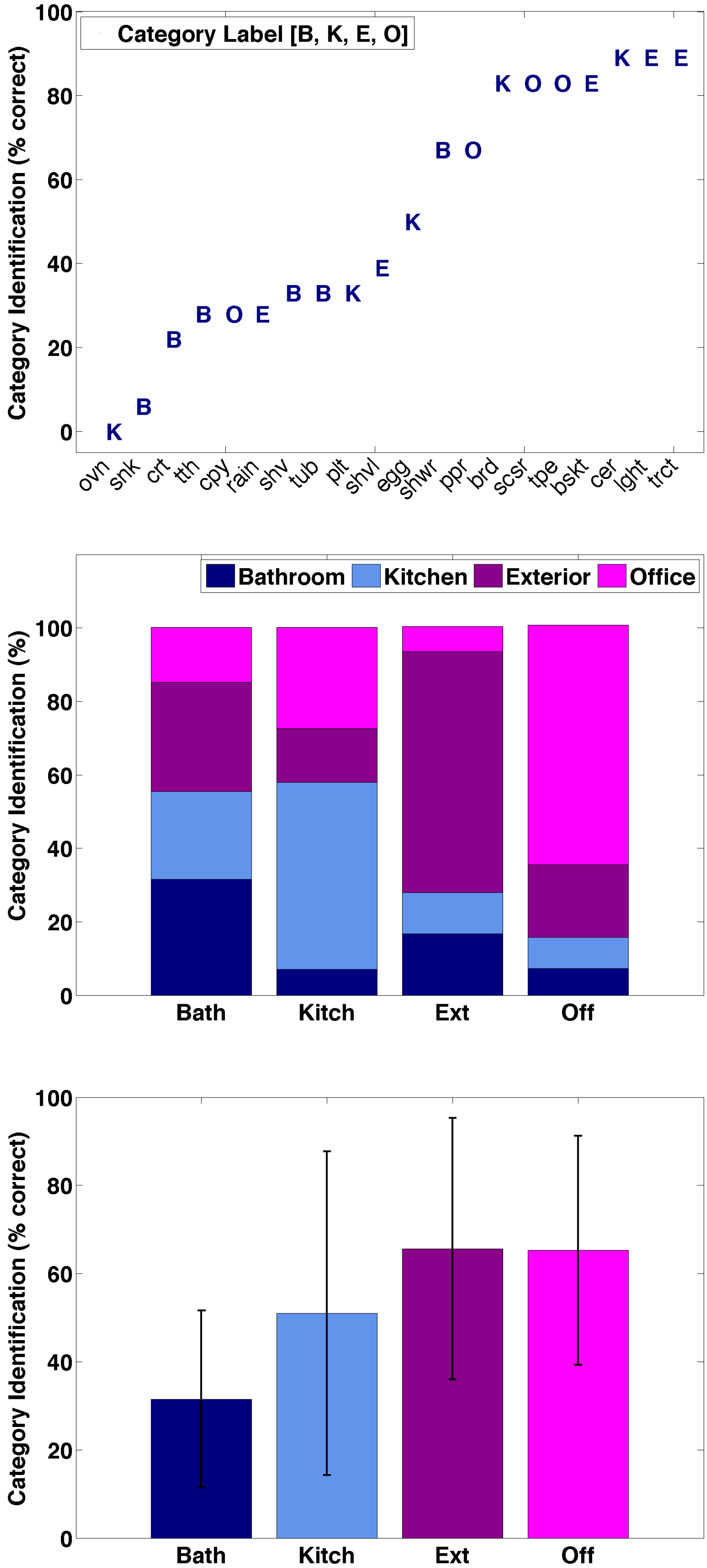


Figure 5.8: Category Identification accuracy for **Step-4-FC-SIM**: *Upper panel* shows the correct category identification of each individual sound, where the corresponding category is represented by **B**-Bathroom, **K**-Kitchen, **E**-Exterior and **O**-Outside. *Middle panel* shows the category identification of each environmental context Bathroom, Kitchen,Exterior,Office. *Lower panel* shows the correct category identification of each overall environmental context i.e. the correct responses from the middle panel.

Middle panel of figure 5.8 shows the effect of adding context information to the vocoded sounds. On average there is an increase in identification performance of 4 ± 19 percentage points. A total of 11 of the 20 sounds also show an increase in identification whilst 6 show a decrease. However multiple ANOVA testing shows that this small increase is not significant ($p > 0.05$), meaning that the addition of the four contexts does not appear to help participants with their identification. In the case of Natural sounds, comparing Step1-Typ & Step2-FST-NS (top panel of figure 5.9, the difference is significant such that the addition of context information in Step1-Typ increases the average identification of sounds by 21 ± 26 percentage points.

The results would therefore suggest that in the case of *natural sounds* the addition of context information helps improve the identification of sounds. However this effect is not seen for vocoded sounds. Three possible ideas on why this might be are listed below:

1. Step-1-Typ-NS gives participants the direct relation of context to sound, where as in Step-4-FC-SIM the sounds were given separately to a sheet of paper with the four contexts written down. So in Step-4-FC-SIM, participants still had to make to the correct relation from sound to category and maybe this induced variability and error to their perceptions.
2. The addition of context could make identification MORE difficult if it limits participants ability to identify the sound.
3. Vocoding simply makes the identification of the sounds too difficult. The resulting difference in Step-3-FST-SIM to Step-4-FC-SIM is simply a results of variability in performance.

Whilst the identification results do not show a difference, and it is not possible to compare the category identification due to the differences in methodology between the different steps, co-occurrence (similarity) matrices can be analysed to compare the participants performance. Figure 5.11 shows co-occurrence matrices for Step-2-FST-NS to Step-4-FC-SIM where each cell corresponds to the percentage of participants that paired each stimulus together with every other sound. Cells have been arranged to reflect the predefined categories in table 5.2, and are divided by the dotted red line, such that the upper triangle of results outlined by the red line represents the similarity of Bathroom sounds, and the lower triangle the Office sounds. By comparing the average results of cells that correspond to the pre-defined categories it is possible to get an idea of the agreement between the categorisation results of participants in each test-step and categorisation based on contexts. The results are summarised in table 5.3. All values are low, below 0.5, which shows that participants are not strongly categorising sounds based on the context, however value of Step-3-FST-SIM is lower than for the other two test conditions. That the value of pair-wise comparison for Step-4-FC-SIM is higher than Step-3-FST-SIM suggests that the addition of context has an effect on the perception of the sounds such that participants are able to more accurately pair stimuli from the same context together.

Test step	Step-2-FST-NS-FST-NS	Step3-FST-SIM	Step4-FC-SIM
PW value	0.389	0.225	0.367

Table 5.3: Pairwise comparison of similarity matrices for each testing step to the pre-defined categories described in table 5.2. Similarity matrices or co-occurrence matrices are displayed for Step3-FST-SIM and Step4-FC-SIM in figure 5.11.

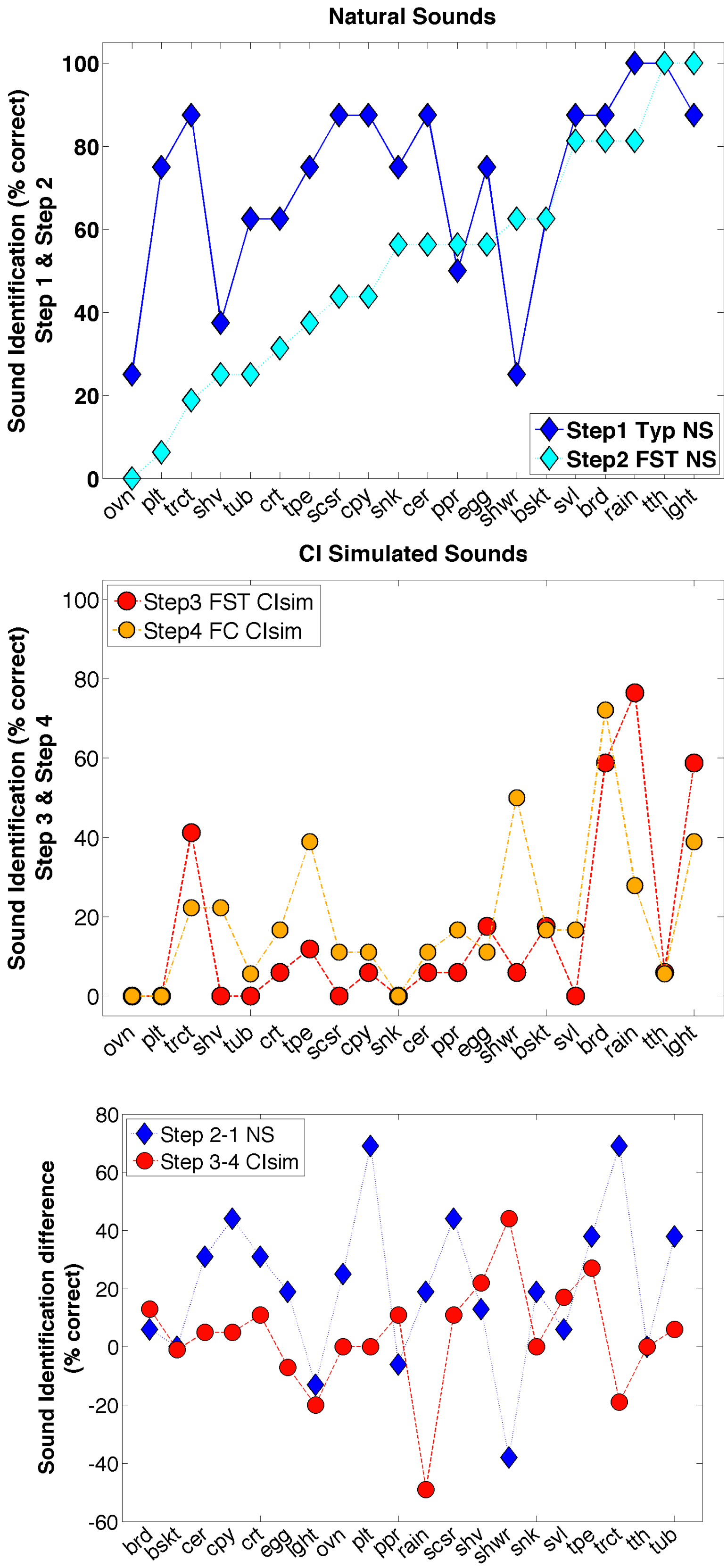


Figure 5.9: Identification performance across the 4 testing steps. *Upper panel* shows identification of sounds in the **Natural sound** condition for the typicality ratings (**Step1-Typ-NS**) and free-sorting-task (**Step2-FS-NS**). *Middle panel* shows sound identification in the CI simulated (4 channel vocoded) conditions for free-sorting (**Step3-FS-SIM**) and fixed categorisation (**Step4-FC-SIM**).*Lower panel* shows the effect of informing participants of the four environmental contexts, i.e. the difference between Step-2-FST-NS to Step1 and from Step-3-FST-SIM to Step-4-FC-SIM.

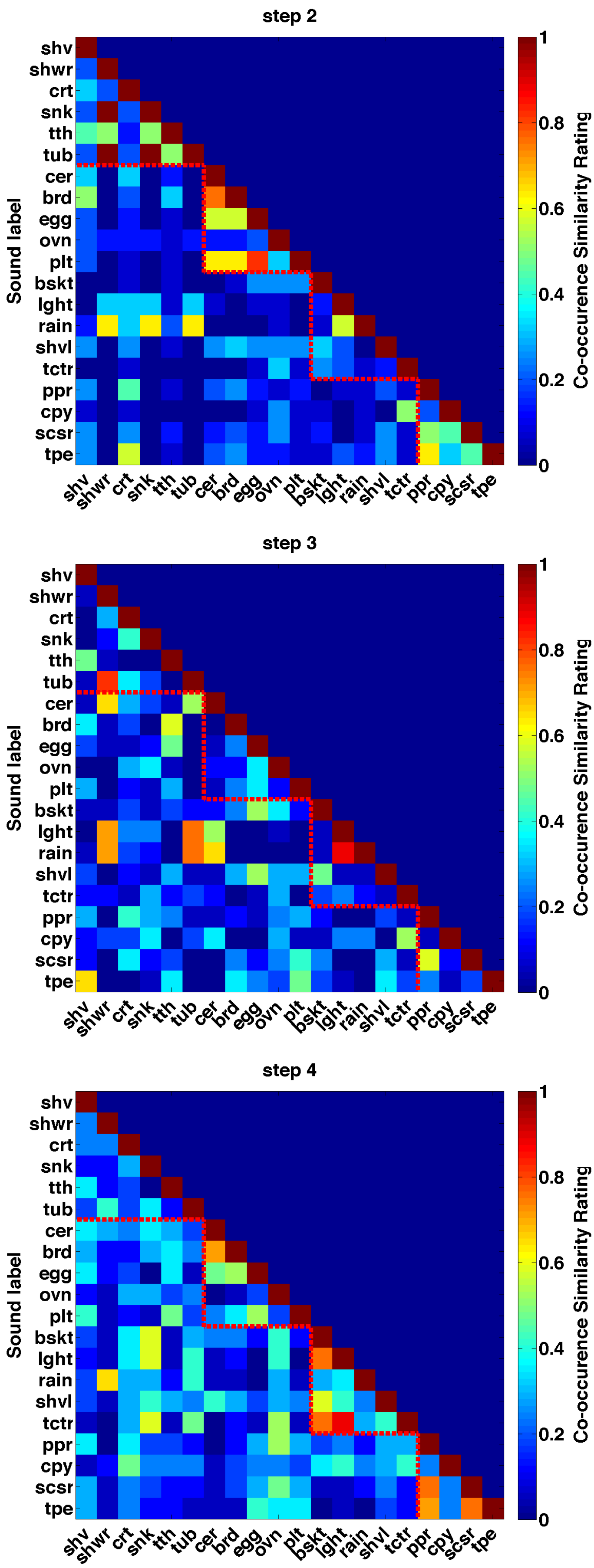


Figure 5.10: Co-occurrence (similarity) matrices calculated for **Step2-FST-NS**, **Step3-FST-SIM** & **Step4-FC-SIM**. Each cell corresponds to the number of participants who paired two stimuli together, with low similarity indicated by blueand high similarity by red. Stimuli are arranged by the four environmental contexts, which are separated by the red dotted line.

5.3.6 *Acoustic Analysis*

In keeping with the acoustical analysis of previous chapters the same measurements were carried out for both the natural and CI simulated sounds. Acoustic measurements were subsequently correlated, using Pearson correlation, with the results of MCA (coordinates of factor maps) as well as the sound-identification scores of all four testing steps to try and understand if any acoustic variables were of importance to the perception of the sounds. Whilst no significant correlations ($p > 0.05$) were found with the identification scores, the MCA results of Step3-FST-SIM were correlated ($p < 0.05$) with the RMS in multiple frequency bands. Specifically the coordinates of Dim2 (as shown in figure 5.6) were *positively* correlated with the RMS in lower frequency bands from $350 - 1850$ Hz AND also *negatively* correlated with the RMS in higher frequency bands from $2150 - 4000$ Hz. Figure **??** displays the sounds plotted against the average RMS in each range of frequency bands, RMS in lower frequency bands along the X-axis, and higher frequency bands along the Y-axis. By plotting the figure in this way it is possible to see how sounds are displayed / separated and when comparing to the MCA results (upper panel of figure 5.6) it the pattern shows similarities, for example the group of sounds *trct, sink* & *cpy* are separated along Dim2 of the factor map and also have high RMS in the lower frequency bands compared to other sounds. That the sound *bskt* also shares similar RMS values, but is not grouped with the same sounds on the factor map indicates that the RMS is not only way sounds are categorised.

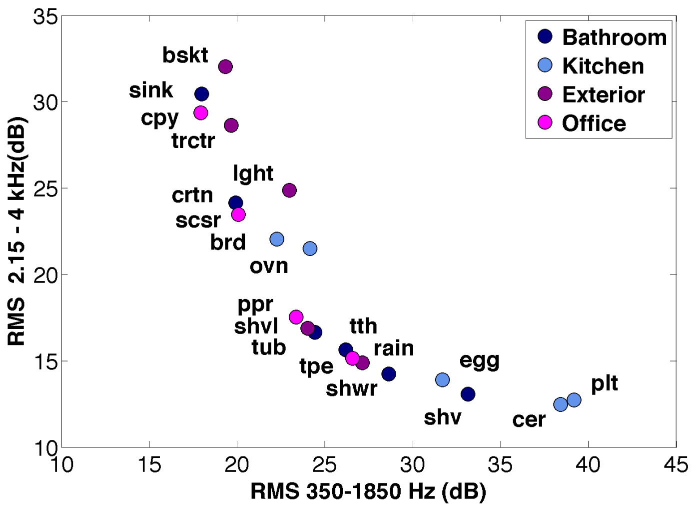


Figure 5.11: RMS in high frequency ($2.15 - 4$kHz) plotted along the the y-axis and low frequency ($350 - 1850$ Hz) bands, plotted on the x-axis for CI simulated sounds (4 channel vocoding). The 4 environmental contexts are again indicated by colours: Bathroom, Kitchen, Exterior, Office. Whilst individual sound labels (see table 5.2) indicate the different sounds used.

5.4 DISCUSSION

The current chapter aimed to firstly test how a collection of environmental sounds were categorised in a Free Sorting Task (FST) and whether or not the information that related to shared context would be used as part of participants categorisation strategies. Results showed that context information was used to create a category of Kitchen sounds alongside other categories of water, paper, machine/mechanical and action sounds. Thus suggesting that for the tested stimuli the context/location was not the most salient or important information that participants used. When testing the categorisation of CI simulated sounds with a 4-channel vocoder identification of the sounds was poor, however perceptions relating to the material and action of the sound-producing event remained robust and central to categorisation strategies. In the case of a fixed-categorisation task, the association/link between the context and perceived material or action was used as the basis to categorise sounds. The addition of context information was seen to help participants identify sounds in the Natural sounds (NS) condition, however not in the case of CI simulated sounds. However the task difficulty may have been too high for participants to make use of the additional information when listing to CI simulated sounds with only 4-frequency channels.

Secondly, consecutive testing steps (described below) were used to see if the addition of auditory context could aid the identification of the same 20 sounds following vocoding. Vocoding was done with 4-channels in order to simulate CIL ESP as identified by [93]. Results did not show any significant effect of additional context information in aiding the identification of sounds.

- ***Step1-Typ-Typ:*** *Rating of typicality to establish how suitable the sounds were in relation to the chosen context*
- ***Step2-FST-NS:*** *Categorisation of the natural sounds in a FST*
- ***Step3-FST-SIM:*** *Categorisation of vocoded sounds (using 4 channels) in a FST*
- ***Step4-FC-SIM:*** *Fixed categorisation of vocoded sounds (using 4 channels) into 4 categories relating to Bathroom, Kitchen, Exterior and Office.*

5.4.1 *Categorisation of (Natural) Environmental sounds*

The results of the current chapter demonstrate that auditory categories can be formed using different information. In the case of natural un-processed sounds

(Step-2-FST-NS) categories mainly reflect attributes of the sound-producing event as specific objects (bouncing ball, machine/transport) or as a certain material (water, paper) [75]. Categories also seemed based on perception of the sound producing action (manipulations of paper sounds, cutting, chopping) [76, 91, 57]. However, it does not appear that categorisation strategies are based upon the identification of specific contexts. There is some evidence to suggest that a category of Kitchen sounds is perceived as in similarity with [91], however these sounds also share a common activity - cooking. [48] found that the level and type of activity within a soundscape was responsible for how different soundscapes were categorised. It is possible therefore that categories reflecting a context (or location) are actually formed based on the activity that takes places within a certain context, rather than simply the idea of a context as simply a location. This argument may also explain the category of Office sounds, which are grouped together based on the activity of "sorting papers" rather than specifically because they belong to the category of "office". [91] for example found a category referencing "paper" sounds but no specific mention of "office". This is also in similarity to the category of water sounds, which may be presumed to represent the context of bathroom, however perceptions are associated to many different objects and include both bathroom and exterior sounds such that the common perception is that of the material - water, [57].

5.4.2 *Categorisation of Vocoded sounds*

The results of CI simulations using a 4 channel vocoder should be first discussion on the basis of what they show about the categorisation of sounds. Specific discussion on how this relates to the auditory perceptions and abilities of CIL is discussed in the following section and not here. Whilst existing studies have looked the identification of vocoded, or spectrally degraded environmental sounds [51] there does not appear to be any significant insights into the categorisation of these sounds. The results hereby constitute the first attempt at doing so and are therefore of interest in aiding the understanding of auditory categorisation.

Identification performance of the CI simulated sounds is, as expected, much less than in the case of natural sounds. Theory of Gaver would suggest that in such cases, where identification is difficult, that musical listening would be employed by participants to compare sounds based on their acoustical similarities of the sound signal. However the majority of descriptions used by participants refer to semantic qualities related to the sound-producing event, object or action. This is more in keeping with a continued used/reliance of "everyday listening" mode [37], or the "Causal" and "Semantic" listening modes described by [76] and suggests that the way environmental sounds are perceived requires further understanding.

As mentioned above, they study of [51] also tested the identification of spectrally degraded sounds. Results found that the most easily identifiable sounds were those with *salient temporal patterning* and certain sounds within the current study reflect this - e.g. *cutting bread* and *tearing tape*. The frequency region of $1200 - 2400$ Hz was also identified as important to the identification of environmental sounds by [51] and whilst identification performance of the current study showed no correlation to the RMS in any frequency region, the correlation with Dim 2 shows that the current stimuli may be categorised based on the RMS in ($2150 - 4000$ Hz) and lower ($350 - 1850$ Hz) frequency regions.

There are similarities to the categorisation of natural environmental sounds (Step-2-FST-NS) in the categorisation of water, machine/transport and action sounds together. This is very surprising that even with the difficulty identifying sounds, the perception of these categories appears quite robust. As explained above this is based on the perception of the semantic properties associated with the sound producing event. Understanding that information related to the sound-producing object and action is still perceived when identification is poor is important as it shows that participants are still processing sounds in a more everyday listening way. This result could also provide a step towards understanding if a possible hierarchy of categories exists.

Categorisation and identification can be thought of in the same way in that they are both processes that seek to attach specific information to the perception of an auditory stimulus, however they differ in that identification is more specific or detailed [105, 91, 76]. At a level of category perception the current results suggest a perceptual space that involves categories of water, machine and action sounds. This is important for two reason, firstly it shows the kinds of information listeners are able to perceive when spectral information is severely degraded, also secondly the categories that may exist at a broader level of perception above those categories found with natural sounds. Of course there is the possibility that other categories exist at this level of perception, as the current study only used 20 sounds. It should also be noted that in the discussion of action sounds, these are specifically sounds that hint towards the involvement of a human carrying out the action - for example bouncing a basketball or chopping bread, rather than a non-human action sound such as helicopter blades, a photocopier, waves breaking or a tree falling.

The addition of context information does not seem to aid the direct identification of sounds and would therefore seem to disprove the initial hypothesis. However, in forcing participants to categorise the sounds based on only the 4 contexts, perceptions of the sounds appear to change. Also using only 4 frequency channels may have increased the task to a point where participants were unable to utilise the context information. For example results in the natural sounds condition show an effect of context agreeing with [8]. The task difficulty may also have been raised beyond that representative of CIL environmental sound perception. Studies have shown that ESP for CIL is in the range of XXX - XXX, which is higher than the results of Step3-FST-SIM & Step4-FC-SIM, which over both conditions is only $18 \pm 21\%$ correct. With an increased number of channels, identification of the sounds would likely increase, as evidenced in chapter 2. In this case it would be more likely to note an effect of the context on identification as noted in natural sound condition (Step1-Typ and Step2-FST-NS). In other words, in order for the context to make be of use to participants, they must be able to identify sounds accurately enough to enable them to establish/create a link between the sounds and the context.

Previous works have also highlighted the acoustic complexity and speed of spectral dynamics as being important to the categorisation of environmental sounds [110, 111]. The present study does not demonstrate any results to counter or agree with this. One reason may be that in comparison to examples of simple and complex sounds used in [110], the stimuli used in the current work would all fall into the category of *simple* sounds. These simple sounds are defined as by slowly varying spectral dynamics and had value of **SSV!** (SSV!) less than 1. Complex sounds included examples of vocalisation (both animal and human) and musically related sounds, which were absent from the stimuli tested in this chapter. Acoustic properties of the stimuli also beomce more important to categorisation strategy when identification fails and when listeners adopt the "musical-listening" mode as proposed by [37]. In the case acse of vocoded sounds, identification is poor and so it would therefore be presumed that participants strategies would be more strongly based on the acoustical characteristics of the sounds. However participants comments related to identification still refer almost exclusively to qualitative aspects of the sound producing event (material, object or action) and comments describing participants categories rarely refer to acoustical characteristics, such as similarities in frequency, smoothness/roughness, complexity or loudness. This would suggest that Gavers theory is not so distinct as initially laid out and that ideas proposed by [75] pose a good direction in which to continue researching auditory categorisation i.e. that categories are based on causal, semantic and also acoustic perceptions of the sound. This also reflects the sentiment shared in [7] that "the recognition of environmental sounds is directed to produce semantic interpretations of the sound" rather than reflect the perceptions of acoustical information concerning the sound signal.

5.4.3 *Methodological Considerations*

Similar considerations as described in previous chapters must also be mentioned again here. Firstly, the categories that are outputted from the HCPC and MCA analysis represent the "overall" performance of all participants. It is also possible that very few participants completely follow the categorisation strategies described above. And participants may actually categorise the same stimuli in different ways as found in previous research [57, 91]. Secondly, the results and categories produced by participants are dependent on the stimuli that are used in testing. As an example [57] found categories of solids, liquids, gasses and machine sounds, however when testing only solid sounds results showed a difference between impact and continuous sounds not found in the first experiment. In the current study a limited number of 20 sounds were used in order that testing could be completed in a short time frame. However this of course again limited the possible categories that could be created and made it more possible for certain sounds to drive/dominate the categorisation in comparison to a test of many stimuli [57].

Regarding the four different contexts that were chosen, Bathroom, Kitchen, Office and Exterior, typicality ratings of individual sounds are not equal. The most typical sounds are likely to be more familiar to listeners and could therefore affect the categorisation strategies used [6]. For example, within the Kitchen sounds, the sound *ovn* has the lowest typicality rating and is categorised apart from the other Kitchen sounds. On the other hand all stimuli are common to the given contexts and variability in typicality and familiarity represents a real world scenario whereby sounds do not share the same level of familiarity amongst different listeners. The method of testing may also have influenced the results, whereby had participants been given written labels for the sounds they would have recorded much higher typicality ratings for all stimuli. However if all sounds were of equal and high typicality this may have actually biassed the results to drive a categorisation strategy more strongly based on context, although it is the opinion of the experimenter that this is unlikely due to the hypothesis that categorisation based on context is actually driven by the activities linked to certain contexts. Therefore to drive a categorisation strategy of context it would be more easily done with sounds that shared related activities common to specific contexts.

5.5 CONCLUSION

The current chapter delved into auditory categorisation looking at specifically environmental sounds divided between four specific contexts/locations. It was hypothesised that these contexts may firstly form the basis of auditory categories, however findings suggest that information related to the sound-producing object and action are more important/salient to auditory categorisation processes. CI simulated sounds also showed that perceptions of the sound-producing object (potential the material) and sound-producing action remain robust and useful for categorisation even when identification of sounds is reduced by degraded sound quality. This may provide a step towards understanding a possible hierarchy of environmental sound categorisation which may extend to all the categorisation of all auditory stimuli. It is important for understanding what auditory categories exist or that we create to deal with the world, but also gives big clues on how we listen to sounds and the information that is perceived or used in higher order cognitive processes. However this requires much further study. Although the understanding of auditory categorisation is starting to be formed, it still seems quite variable, participant & test dependent and also no classification system has yet been established beyond what has been laid out by [37, 77].

Key Points

- The categorisation of a set of 20 environmental sounds is based predominantly on information concerning the sound producing object. With regards to sounds in the Natural sounds (NS) condition this results in categories found in previous research - machine/transport and water/nature sounds.
- The same 20 sounds were vocoded with 4 frequency channels to simulate CI users discrimination of vocal vs environmental sounds. Even in this condition where spectral information was severely degraded, and participants struggled to identify sounds, categorisation strategies were based on the perception of the sound producing object and sound producing action. This again resulted in perceptions associated to machine/transport and water sounds as well as perceptions of the material e.g. water, metal and wood; and the sound producing action e.g. cutting, chopping, scratching. This is an important observation that even in difficult listening conditions, NHL are still aiming to identify information associated to the sound producing object/event, whether it be *causal* or *semantic* [57], rather than focussing on the *acoustic* characteristics of the physical sound signal.
- However, perceptions of the sound producing action may also be related to the temporal structure of the sounds. Whereby the rhythmic nature of some sounds (e.g. *basketball* and *shovel*) helps participants to perceive actions like chopping/scraping in the CI simulated condition.

Key Points

- Categorisation strategies do not appear to be strongly based on the environmental context. From the initial choice of Bathroom, Kitchen, Office and Exterior sounds, the only apparent categories seen in the results was for Kitchen. However this may be based on the *activity* of preparing food rather than the perception of the kitchen as a specific location. In this way it reflects the findings of [74] who found that soundscapes were categorised based on the events, or activities that were perceived in each soundscape. Certain Office sounds were also categorised together, however this is again more likely based on the shared perception of manipulating paper, rather than the perception of the Office as a semantic category.

- The addition of environmental context appeared to help identification of sounds in the Natural sounds (NS) condition, with an average improvement of 20% points. This was seen between from Step2-FST-NS to Step-1-Typ-NS. However in Step1-Typ-NS participants were given the direct association between the sounds and their corresponding context. Context information did not however aid the identification of sounds in the CI simulated condition i.e. the different in results between Step3-FST-SIM and Step4-FC-SIM. In this case it may that using a 4 channel vocoder makes it too difficult to make use of the context information, or that in the FST paradigm participants could not make use of the additional environmental context information because the direct link between the sounds and corresponding context was NOT given.

- Results therefore suggest that the environmental context may not hugely important to participants perception of environmental sounds. It may make more sense to focus on the kinds of information that are perceived by listeners in adverse listening conditions as a way of enhancing the understanding of CI users auditory abilities, especially concerning the perception of environmental sounds.

6 CATEGORISATION OF HUMAN VOICE - THE IMPORTANCE OF AGE, GENDER AND EMOTION

6.1 introduction

Much work on the perception and categorisation of vocal stimuli has concentrated on the gender of voices, prosodic processing, vocal discrimination/recognition [133, 4, 63, 135, 47] and concluded that the categorisation of voices is determined by the age and gender of a speaker [47]. This would indicate that the age and gender are important features of a speakers voice. [47] used a Multidimensional Scaling (MDS) analysis to show that in a test of categorising syllables stimuli were categorised first on the basis of age and then gender (see figure **??**).

Of course there are other features of voice stimuli that may be used to identify and categorise individuals, for example sexual orientation, ethnicity, accent, personality [133]. Emotion is also an incredibly important feature of vocal sounds vital to human communication. As a first step to further understanding how human voices are categorised it was therefore decided to test vocal stimuli that varied across three different vocal characteristics the talker, these were Age, Gender and Emotion. This also enabled the chance to see how results would differ from those of [48] and give possible clues as to the importance (or hierarchy) of the the three characteristics stated. In order to test for a possible hierarchy it was decided to limit construct each vocal characteristic with only two options. Of course for gender this is already decided between male and female. Regarding age, voices of adults and children were chosen whereby adult voices existed in the age range of approximately 25-50 years old and child voices between the ages of 5-10. This was done to reduce the chances of voices being perceived as generated by babies, adolescents, or elderly voices. For Emotion, happy and sad emotions were chosen to be clearly different and this was achieved using the emotional actions of laughter and crying sounds.

This testing was also deemed relevant to the study of vocal perception by Cochlear Implanted Listener (CIL), who have difficulties wither speaker discrimination and gender discrimination in comparison to NHL [94]. Studies of emotional perception with CIL have also found that CIL do not perceive the emotional content of voice stimuli as well as NHL [98]. The most easily perceived emotions have also been found to be anger and sadness due to specific acoustic cues which aid CIL's perceptions [98] (Banse & Scherer, 1996; Most, Wiesel, & Zaychik, 1993; Pereira, 2000).

The following chapter therefore looks at the NHL categorisation of voices based on age, gender, emotion to see if the same or different hierarchy exists as previously hinted at. Preliminary (but not conclusive) results are also presented for a group of 11 CIL in order to see how performance compares to NHL and if CIL are capable of perceiving and utilising all forms information within the vocal stimuli.

Chapter Aims

1. To test the categorisation performance of a set of 16 vocal sounds that differ in age (child or adult), gender (male or female) and emotional content (laugher or crying). Stimuli have been created so that each variable (age, gender, emotion) is clearly distinct and identifiable. In this way it is hoped to find which aspect of the stimuli is most important to categorisation.
2. It is known that CIL have difficulties identifying speakers and discriminating voices. This has been seen with regards to gender and also age, it has also been seen that CIL users perceive different emotions with more ease/difficulty. However which of these vocal characteristics is more or less important and how they are distinguished when presented together has not been tested. It is hoped to better understand this by testing the categorisation performance of CILand to find a relative hierarchy that may indicate the saliency of each stimulus characteristics for CIL

6.2 METHOD & MATERIALS

6.2.1 *Participants*

Two groups of participants were recruited for testing the categorisation of human voices in this chapter. The first group consisted of 25 normal hearing listeners (NHL) who were native French speaking adults (16 males age 29 ± 9 years). The second group were 11 Cochlear Implanted Listener (CIL), also native French speaking adults (6 males, age 46 ± 19 years). Full details of the CIL are listed in table 6.1.

6.2.2 *Stimuli*

As mentioned above the goal was to test the how voices would be categorised as a novel way of testing the perception of voice and associated attributes. As this testing constituted a first step towards fully understanding the categorisation of voices it was decided to test stimuli with a limited number of three salient vocal characteristics. These were chosen to be the age, gender and emotion of the voice and for each characteristic two extremes were chosen so that participants would, hypothetically, be able to easily distinguish the stimuli. Thus, for age child and adult voices were chosen, for gender male and female and for emotional voicing laughter and crying. This was done so that each vocal characteristic was of equal perceptual salience and therefore would not create bias in the goal to evaluate the categorisation strategies of participants. The construction of stimuli based on these characteristics is shown in table 6.2 and for each stimuli construction there are two examples present.

Stimuli in table 6.2 were chosen from a selection of 48 stimuli that were evaluated for their typicality to each of the three vocal characteristics. Typicality were carried out by with 30 NHL, divided into three groups and where each group listened to and evaluated 16 of the 48 overall sounds. The 16 sounds with the highest overall typicality, i.e. that best represented each of the age, gender and emotion characteristics were then used for the categorisation testing and the results for these sounds are shown in figure 6.1.

Importantly to note, throughout the chapter the following ID tags have been used to identify stimuli, whereby **B = boy** (i.e. male child), **G = girl** (female

Patient ID	Age at testing	Duration of Deafness	Duration of Implantation (months)	Implanted ear	CI model	Dysialabic word Word recognition	Sentence in Noise (SNR +10dB)
PILCAT	50	47	27	R	HiRes optima S	62	74
ASEVER	50	43	16	L	HiRes optima S (clearvoice off)	76	88
HALVAN	34	22	14	L & R		68	89
POTMAX	22	birth	24	R	Hires S-w Fidelity 120	80	99
POTTHO	20	birth	25	R	HiRes optima S	52	79
PELMAR	65		18	L	COCHLEAR CI422CP920	75	92
NOVCHR	59	10	98	R	COCHLEAR 24RECACP910	50	64
CAPJUL	27	birth	152	R	COCHLEAR CI 512CP910		
BUIJEA			12	R			
ESTMAR	67	9	86	R	COCHLEAR CIU 24 Contoure	70	95
BRIDAV	48		36	R	COCHLEAR 24 RECACP910	92.5	+5dB (92L, 100R)

Table 6.1: Individual participant data for CI users tested in chapter 6. Participants in the top half of the table were tested at CHU Lyon, whilst participants in the bottom half were tested at Hopital Purpan, Toulouse.

child), **M = Man** (adult male) and **W = woman** (adult female).

ID	Age	Gender	Emotion
BC-1			
BC-2		Male	Crying
BL-1			Laughing
BL-2	Child		
GC-1			Crying
GC-2		Female	
GL-1			Laughing
GL-2			
MC-1			Crying
MC-2		Male	
ML-1			Laughing
ML-2	Adult		
WC-1			Crying
WC-2		Female	
WL-1			Laughing
WL-2			

Table 6.2: Stimuli description for Voice-categorisation. Stimuli ID are listed in the left hand column where *M* = man, *W* = woman, *B* = boy, *G* = girl, *L* = laughing, *C* = crying. Subsequent columns detail how stimuli were divided between AGE (child/adult), Gender (female/male) and Emotion (laughing/crying). As described in the text, two example of each stimuli type were used for testing.

6.2.3 *Procedure*

The same procedure of using a Free Sorting Task (FST) was employed here in similarity to the that used in previous chapters. The FST was used to test participants categorisation of a number of different sounds. The current chapter differs in that the stimuli were all vocal sounds as described in table 6.2. Stimuli were also specifically chosen to vary along the three dimensions of age (child/adult), gender (male/female) and emotion (laugh/cry).

Analysis was again carried out using Heirarchical Clustering based on Principal Components (HCPC) and MCA with results being presented in the form of dendrograms, factor maps and subject maps in the following section.

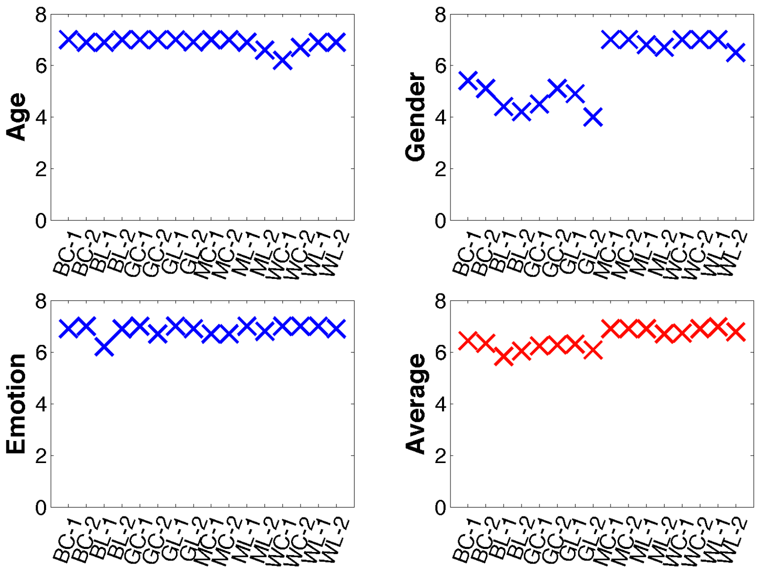


Figure 6.1: Figure of typicality ratings for Voice-stimuli, as listed in table 6.2. Stimuli were rated on their typicality for Age (child-adult), Gender (male-female) and Emotion (laughing-crying) on a scale of $1 - 7$. The Average results of the three ratings is also shown in the bottom right panel.

6.3 RESULTS

The following section is divided into two sections. The first covers the categorisation performance of participants using MultiCorrespondence Analysis (MCA) and Heirarchical Clustering based on Principal Components (HCPC) analyses. This will set out what kinds of categories were created , what aspects of the stimuli were used for categorisation and overall describe the categorisation strategies used by participants in the three testing conditions. A second section will present results on how accurately stimuli were perceived in the three different conditions using participants verbal descriptions when asked to identify the stimuli. This information is also important in understanding the categorisation strategies.

6.3.1 *Categorisation Performance*

Figure 6.3 shows the MCA analysis for the voice-categorisation in the Natural sounds (NS) condition. Dimensions can be quite easily interpreted due to the simplistic nature of the stimuli, in that they varied only in three salient features of age, gender and emotion. Dim 1 shows crying sounds strongly contrasted with laughter sounds, such that the first dimension and the strongest manner in which participants are grouping/separating the stimuli, is based on the *emotion*. There is also no different in terms on age or gender, such that laughter and crying stimuli have almost the exact same value (respectively) on Dim 1. Figure 6.4 shows the map of participants across the MCA dimensions. As can be seen in the upper panel participant agreement to Dim 1 is very high, 96%, only one participant does not appear to be using emotion as the first manner of categorising the stimuli. In fact participant *ns03* categorised stimuli into two groups, described as *adults* or *children*, highlighting the variability in categorisation strategy that exists.

Dim 2 is also quite clear in that it contrasts the adult and child stimuli, reflecting an association to the *age* of the voice. There is a small difference associated to the emotion of the voice however, in that the laughter stimuli seem more strongly separated by age, than the crying stimuli. This could reflect a perceptual difference in that the age of the voice is more easily perceived for the laughter sounds. This is further evidenced when noting that Dim 3 also contrasts the stimuli based on age, but more strongly for the crying sounds, with adults (W & M) to the right of the axis and child voices (B & G) to the left. In fact the coordinates of participants on Dim 2 and Dim 3 are significantly correlated ($r = 0.87, p < 0.05$) showing that those participants who strongly use Dim 2 are the same whom use Dim 3, a total of 13 participants. Of course this makes sense that those participants who would categorise the laughter sounds via age, would also categorise crying sounds in the same manner.

It is not until Dim 4 that the gender of the stimuli appears as part of the categorisation strategy and noticeably this is exclusively linked to the adult voices. Whilst this is not unexpected, based on gender typicality values in figure 6.1, the fact that gender should appear so 'late' in the categorisation strategy is surprising based on previous studies that focus more on gender discrimination / categorisation with regards to vocal stimuli. Interesting split of those participants that do or do not use Dim 1,2,3,4. There does not appear to be a graded agreement for any dimension. This is again likely influenced by the distinct characteristics of the stimuli.

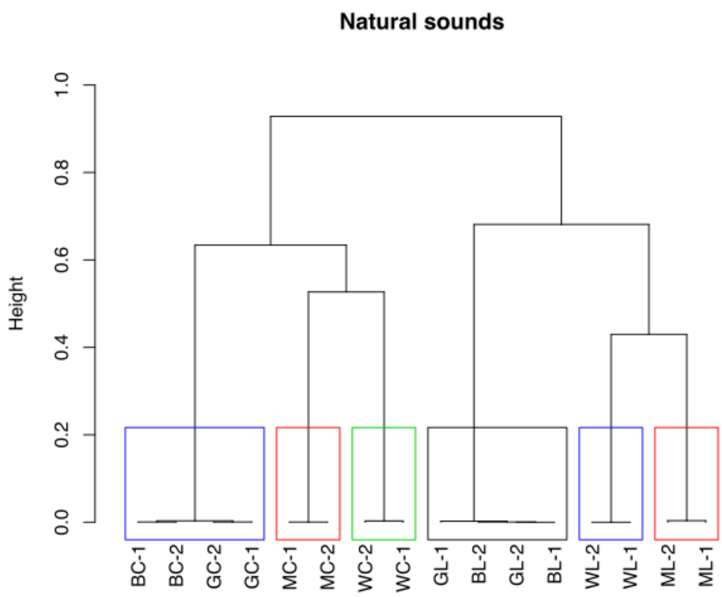


Figure 6.2: Dendrogram output of the HCPC analysis for Natural sounds. The height axis indicates the level of similarity between stimuli, such that a greater height corresponds to a greater DIS-similarity.

Figure 6.2 shows the results of the categorisation in a simplified form and it is again easy to see that the first node corresponds to Dim 1, dividing the stimuli between laughter and crying, the second and third nodes to Dim 2 & 3 dividing stimuli between age. Finally the fifth and sixth nodes show a of gender, but ONLY for the adult stimuli, where as the child stimuli are grouped all together both boys and girls.

Overall these results suggest that the most salient feature within the voice stimuli is that of the emotion (laughing/crying), followed by the age (child/adult) and finally the gender (male/female). The contrast of gender regarding the child voices was only observed on Dim 6, which covered only 4.2% of the variance and is therefore not shown. Only 2 participants strongly used this dimension and there is also no division of gender for child stimuli in figure 6.2. This therefore shows that the ambiguity of child gender is not a strong salient feature used in the categorisation strategy.

Results taken with 11 CIL, all of whom had more than 12 months experience with their implants were tested with the same stimuli and paradigm. Results in this case must be treated as preliminary because of the nature of the analysis which requires 15 participants to be valid. However due to the simplistic nature of the stimuli the results presented below give a very strong indication to results that would be achieved with more participants and therefore provide valid comparisons for discussion and conclusions.

Results of MCA and HCPC analysis are shown in figure 6.5. The factor map in the upper panel is again easily interpreted, with Dimension 1 showing a contrast of the laughter and crying sounds, and Dim 2 contrasting the crying sounds with regards to age, child vs adult. Further dimensions were unable to be interpreted in terms of differences in age, gender or emotion, instead only showing the separation of singular sounds from other stimuli, for this reason they are not shown.

The dendrogram, lower panel of figure 6.5 also shows that the first distinction of the voice stimuli is regarding the emotion, with laughter sounds all grouped together on the let hand side. Crying sides, separated in the right hand branch are subsequently divided by age, which is not seen for the laughing sounds. Analysis of how sounds were perceived shows that 'age' was perceived to a similar level of accuracy for both laughing and crying sounds (figures 6.6 & 6.7), so it is not a matter of age being more easily perceived for different kinds of sounds. More likely is the possibility that adult-crying sounds and child-crying sounds somehow perceived in different ways and this is discussed further in section 6.4. Noticeably one sound, *BL-2*, is"mis-categorised" amongst crying sounds, when it is in fact the sound of laughing.

Middle panel of figure 6.5 helps understand how participants use the strategies described above. 9 of the 11 participants have a coordinate along Dim 1 greater than 0.8 and therefore share the same strategy of categorising the sounds based on emotion. Only two participants (X106 & X105) do not have a coordinate above 0.8. Participant X106, the furthest away from using the commonly held strategy related to emotion, categorised sounds on the basis of loudness, describing categories as "normal", "louder" or "less loud but the same word". Participant 105, somewhat followed the strategy of emotions, describing two categories as "baby *crying*" and "woman *laughing*", however they also used a category of "coughing" sounds differentiating their personal categorisations strategy from that of other participants. With regards to Dim 2, there is more variation amongst CIL participants, such that only 4 of them strongly use Dim 2. However, other participants still use the perception of age (for example participant X27 and X103) to categorise different sounds. This reflects the variability in perceiving the age of the voice stimuli, a result of varying auditory abilities amongst CIL.

Overall results show strong similarities to those of NHL in the use of emotion and age for categorising stimuli and there appears no use or perception of the gender of voices. Variability in CIL auditory abilities is also explains variability in results.

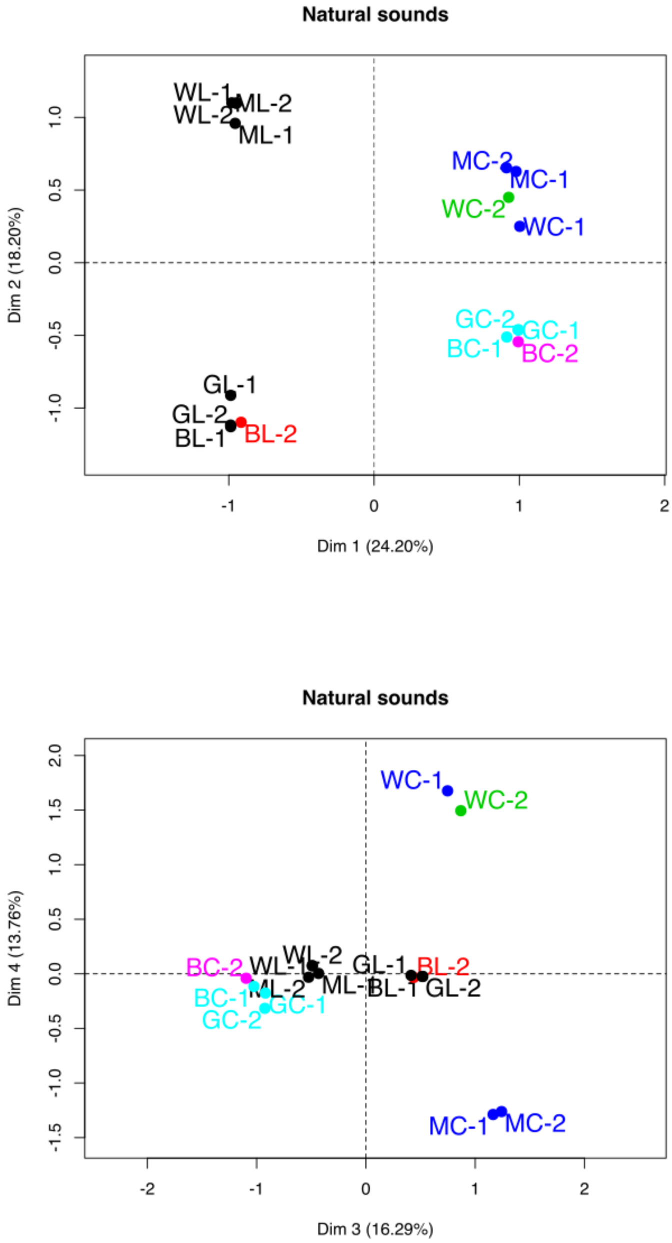


Figure 6.3: MCA factor maps for Natural sounds NS condition, tested with xxxx participants. Upper panel shows dimensions 1&2 with dimensions 3&4 in the lower panel. Amount of variance covered by each dimension is also shown on the corresponding axes.

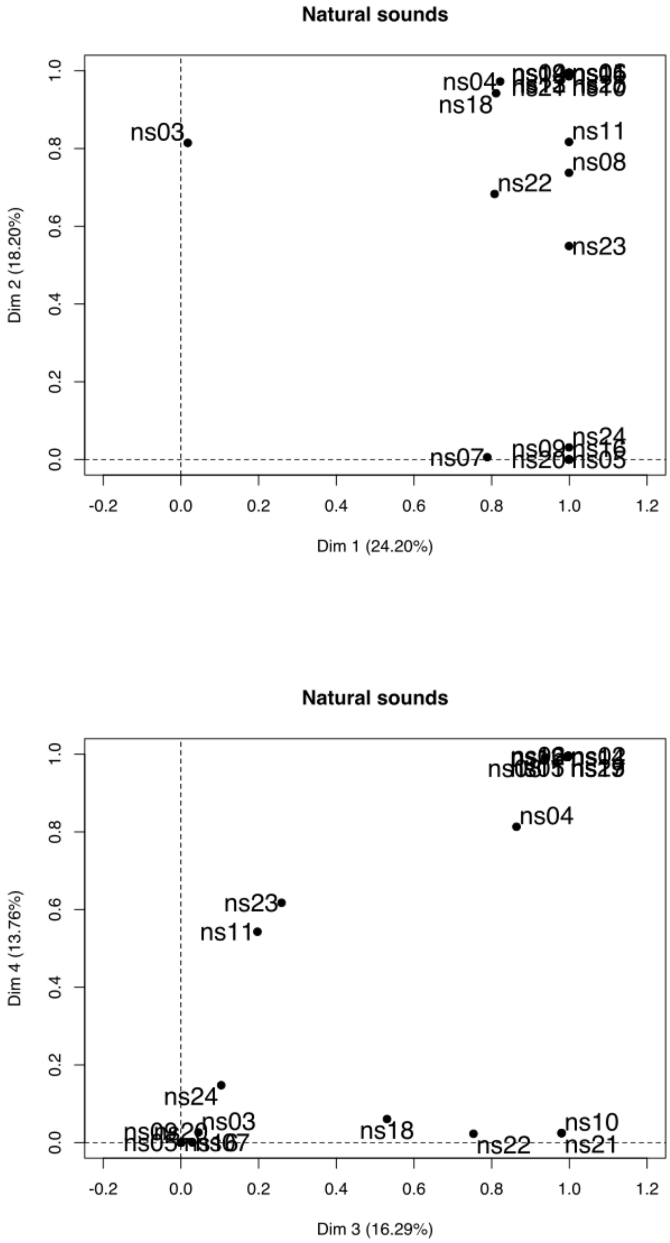


Figure 6.4: MCA factor maps for Natural sounds NS condition, tested with xxxx participants. Upper panel shows dimensions 1&2 with dimensions 3&4 in the lower panel. Amount of variance covered by each dimension is also shown on the corresponding axes.

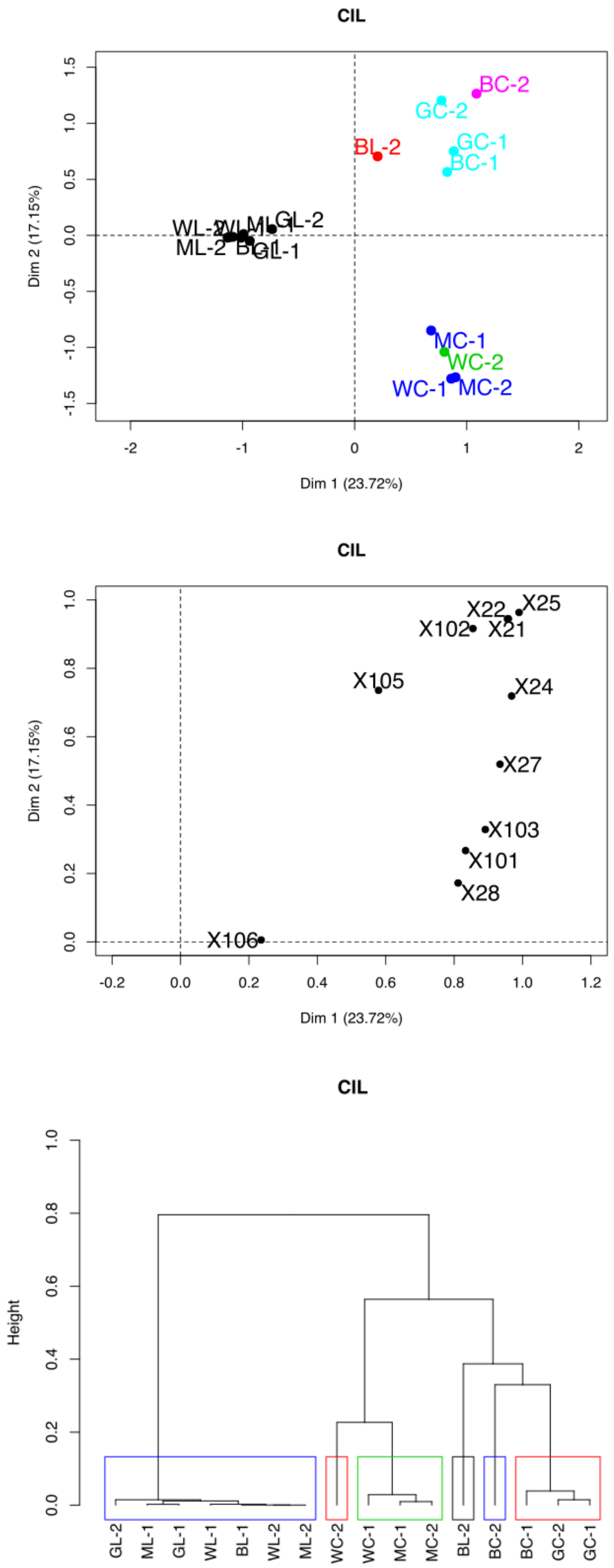


Figure 6.5: Results of categorisation performance of 11 CIL. Upper panel shows the factor map output of MCA analysis, with Dimensions 1 & 2 displayed. Middle panel also shows the participant map for the same two dimensions whilst the lower panel shows dendrogram output of HCPC.

6.3.2 *Stimuli Perception*

In order to aid the interpretation of categorisation strategies described in the preceding section and to further understand the perceptions of voice-stimuli by the two groups of participants, NHL and CIL. Perceptual accuracy of each stimuli characteristics (Age, Gender & Emotion) was analysed by evaluating the descriptions of sounds given by each participant. Figure 6.6 displays the results split between the three stimuli characteristics and shows firstly that the perceptions of CIL participants are lower than NHL. This is true for all, Age, Emotion and Average results (kruskal-wallism $p < 0.001$) however not in the case of Gender perception, where both NHL and CIL struggle to perceive the gender of the child voices. This is further highlighted in figure 6.7, and shows that NHL perceive gender quite well for the case of adult voices whilst CIL do not accurately perceive gender in any condition. The participant groups are also different in that the perception of age is much stronger in NHL than in CIL. This figure also helps show that Emotion is accurately perceived in all cases for both NHL and CIL, although CIL appear to perceive laughter sounds more strongly than crying sounds, as previously mentioned.

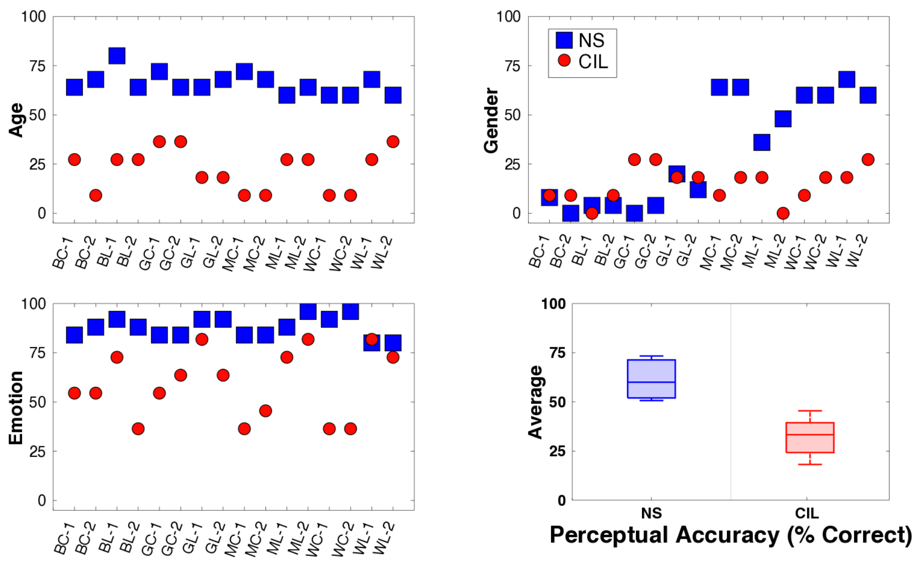


Figure 6.6: Perception accuracy of voice stimuli as rated for each stimuli. The three scatter plots show how accurately the Age, Gender and Emotion were perceived for each stimulus. Results are plotted for normal hearing listeners (NHL) with ■ and for Cochlear Implanted Listener (CIL) with •. A boxplot of the overall perceptual accuracy (the average across all three scatter plots) is also shown.

The analysis of perceptual accuracy show that the stimuli characteristics appear to differ in their salience (or accuracy of perception). Where Emotion is the most strongly salient, well perceived by both participant groups,; followed by Age, again used by both groups, and finally the Gender, which is only well-perceived in the case of NHL listening to adult sounds. This pattern corresponds to the interpretation of the MCA analysis for both groups in that Emotion, then age then gender are used to categorise the stimuli such that the most salient information is used as the basis for categorisation strategies. As mentioned above, age seems to be used as a manner of separating the crying sounds by CIL, however the results of figures 6.6 & 6.7 suggest that age is poorly perceived by CIL in all stimuli conditions. Therefore the perceptual difference between adult-crying and child-crying sounds may be due to another perception of the stimuli.

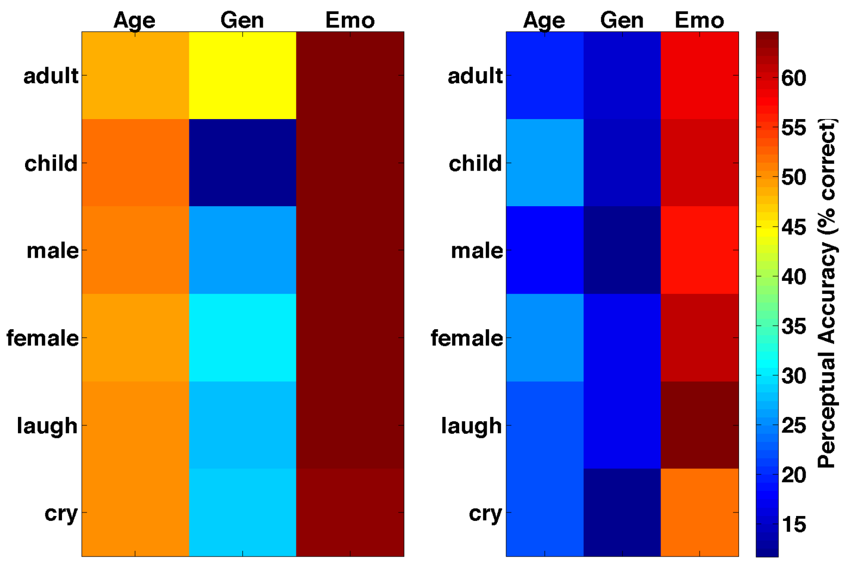


Figure 6.7: Colour map of the average Perception accuracy scores for each stimuli variable - Age, Gender and Emotion. Scores are subsequently subdivided between Adult/Child, Male/Female and Laughing/Crying in order to demonstrate the perception of each variable within each sub-variable. For example in the right hand panel for CIL, the perception of Emotion is stronger/more accurate for laughing sounds as compared to crying sounds.

Overall it can be observed that both participant groups easily perceive the emotional information of the stimuli and subsequently use this as the first part of their categorisation strategies. Age may also be used by both participant groups, certainly by NHL, although there is further discussion required (below) regarding the perception of adult and child crying sounds by CIL. Finally the gender of voice does not constitute a significant part of the categorisation strategies especially for CIL where it is very poorly perceived. This is however slightly clouded by the ambiguity of gender within the child voices. The results would also suggest a change to the importance of information contained within the identification of voice, where the emotional content is the most important, followed by age and gender, although this conclusion is of course made based on the stimuli used in the study and may change with different or more varied stimuli.

6.3.3 *Acoustic Analysis*

In similarity to analysis conducted in previous chapters, the coordinates of MCA dimensions were correlated with acoustic measurements of the stimuli, for both the NHL and CI users tested. Only two significant correlations (Spearman rank $p < 0.05$) were found for the CI user data and are summarised thusly:

For the CI users, Dim 1 (as depicted in figure 6.5) was correlated to RMS at 250 Hz ($r = 0.55$) and 350Hz ($r = 0.51$). Dim 1 is associated to the separation of crying and laughing sounds and it therefore appears that crying sounds have a higher RMS in this frequency range (overall $200 - 400$ Hz).

Dimension 3 is also correlated to the RMS at 150Hz ($r = 0.81$). This correlation is however due to the sound **BL-2** having very high RMS in this frequency range (59 dB) compared to the other sounds. Although Dimension 3 is not shown in the figures, it is not easily interpreted in terms of age, gender or emotional content and therefore this correlation does not appear to aid in the categorical perception of the stimuli by the CIL tested.

6.4 DISCUSSION

A Free Sorting Task (FST) of 16 vocal sounds was tested amongst groups of 25 normal hearing listeners (NHL) and 11CIL where stimuli varied across three vocal characteristics of Age (child vs. adult), Gender (male, vs. female) and Emotion (laughter vs. crying).

Categorisation strategies for both participant groups were based firstly on emotion, secondly on age and thirdly (but only for select stimuli) on gender. This suggests a possible hierarchy of information for NHL that was in part reflected by CIL. Of important note was also the perception of the emotional content by CIL which was fairly unanimously used for categorisation strategies of all the CIL participants.

The results are similar to those of [47] in demonstrating groups/categories based on the age and gender of voice and as mentioned above the results demonstrate a possible hierarchy of vocal information for use in categorising vocal stimuli. The importance that the emotion seems to carry is interesting as it may contravene previously held theory on auditory categorisation that suggests categories are based on successful source identification. Perceiving the emotional content of the stimulus does not rely on accurate source identification and in this way categorisation theory may need to take into account emotional attributes more strongly, for example the categorisation of warning or alarm sounds which has been commonly found [57, 91]. This may also be why the emotion is perceived more accurately than the other characteristics of age and gender. However, one small note is that emotion was elicited using laughter and crying, which themselves could also constitute physical actions more strongly related to source identification. Previous research has used sentences spoken in different emotions, if similar results were found using such stimuli it may remove the perception of a physical emotional action and help to show how strongly auditory (vocal) categorisation is based on emotion vs. source information. This should also be done of course with many different emotions to see how this influences the categorisation strategy and if the hierarchy observed her remains robust.

Whilst CIL clearly discriminate stimuli based on emotion the conclusions that age is also used may be somewhat blurred. CIL show no difference in the perception of age regarding the laugher vs. crying sounds and so it is somewhat unclear why categorisation strategy should be different. There must be a difference in the perception of adult-crying and child-crying sounds. Possibly it is due to a difference in emotional perception, whilst sounds identified as "crying" a baby/child crying is evident of very different needs and actual emotion than an adult crying, whereby a child more likely cries as a result of physical injury and *pain* whilst an adult is more likely to cry as a result of an emotional injury and "sadness". Therefore perception may differ due to the emotion or the emotional intensity within the adult and child crying sounds. If this is the true reason it shows a significant amount of detailed emotional perception by the CIL group that was not originally predicted.

One reason why gender might not be used so strongly as part of NHL's categorisation strategies is the ambiguous nature of the child stimuli. The results of NHL show that gender can and is used by listeners to discriminate between the adult sounds, however due to the age and emotion being more typical and more easily perceived for ALL stimuli the gender is simply less useful in the case of the stimuli described here. For CIL the perception of gender is also difficult with adult voices and so therefore does not play a part in their categorisation. In order to remove ambiguity of child gender subsequent testing (not shown here) with audio-visual and solely visual stimuli (i.e. pictures corresponding to the auditory stimuli) have also been tested in a FST. With the addition of visual information the ambiguity was removed, gender typicality measures were no different between child and adult stimuli. Significantly the results of both tests showed emotion to be used as the first manner of categorising the stimuli in similarity with the the results of only the audio condition presented here. Therefore even when the gender is non-ambiguous it remains lower down in the possible hierarchy of vocal information used by listeners when discriminating different voices.

Evidently CIL are capable of perceiving the emotional actions of laughter and crying. This is important for furthering the understanding of emotional perception with this group of listeners who have previously shown difficulties in this area of perception [98]. Of course further research would require testing a wider variety of emotions and emotional actions and this would be possible using the interesting and novel approach of categorical perception and the FST.

6.5 conclusion

Both NHL and CIL showed similar abilities in categorising a set of vocal sounds that were divided between age (child/adult), gender (male vs. female) and emotion (laughing vs. crying). Interestingly the most salient of these features was the emotion, whilst gender did not constitute a major part of the categorisation strategies employed by either group of participants. This may place greater importance on the emotional content and perception associated with vocal recognition and discrimination for all humans. Regarding CIL their performance at perceiving the various attributes of stimuli was poorer than NHL. Although the results present useful steps towards understanding their abilities in the discrimination of voices and how emotional stimuli may a play a part in this, maybe even aiding it.

Key Points

- Typicality ratings of stimuli (conducted with NHL) show that age and emotion are highly typical and therefore easily perceived for all stimuli. Gender is also highly typical for *adult* voices, however there is ambiguity of gender for *child* voices most likely due to the similarity in fundamental frequency.
- NHL categorise stimuli firstly in terms of *emotion* (laughter vs crying), secondly in terms of age (child vs adult) and finally in terms of the gender of adult voices. The gender of voice is not strongly used as part of categorisation strategies, but maybe influenced by the ambiguous gender of child voices.
- The emotional content of the vocal stimuli is the most salient vocal characteristic in comparison to age and gender. The difference between *laughter* and *crying* sounds is used by both NHL and CIL to discriminate the vocal stimuli and this discrimination represents the first aspect of categorisation strategies for both participant groups.
- CIL also use the emotional characteristics of the stimuli for the purpose of categorisation and results show that the emotion is more accurately perceived than the age and gender. CIL are also poorer at perceiving age, gender and emotion in comparison to NHL. These results are most likely due to the reduced ability of CIL to process spectral information in comparison to temporal information. Noticeably the laughter stimuli used for testing are rhythmic and temporally distinct in nature.

7 OVERALL DISCUSSION

The work presented covers the topic of auditory categorisation - How the perception of everyday sounds is used to group different sounds together. Using a FST, looked at how humans perceive a collection of everyday sounds and how similarities in these perceptions are used to group or categorise sounds together.

Common categories were found in similarity to existing studies - human vocal sounds, musical sounds, mechanical, natural and specific kinds of action sounds. Spectrally degraded sounds showed that categorisation by NHL was robustly based on perceptions of the sound producing event (material / action) and very rarely were acoustic characteristics used as part of categorisation strategies, suggesting that categorisation theories should be further developed to understand the role of potential listening modes as suggested by Gaver (musical vs everyday listening) and LeMaitre (acoustical vs causal vs semantic).

Categorisation of auditory stimuli was also found to be associated to the auditory environment/context, however more likely based on activities that are commonly related to specific places. It was hoped that the addition of context information could also aid the successful perception of spectrally degraded sounds, however results proved inconclusive possibly biased by the task difficulty, or showing that the context information was not useful for participants at all.

The work also present the first results of testing CIL in a FST. Comparison of CIL and NHL show that most experienced CIL follow very similar results to NHL in the separation of human, musical and environmental sounds, with lesser experienced CIL showing the use of similar strategies regarding vocal and human action stimuli. Duration of implantation was seen to be important for the segregation of vocal vs environmental sounds, although overall CIL presented more variation and less participant agreement to the uncovered strategies. Categorisation identification was also found to b improved for more experienced CIL and suggests that categorical perception should be further investigated as manner of improving CIL perception of real-world sounds and possibly incorporated into rehabilitation schemes.

7.1 how are sounds categorised?

The present work would seem to echo auditory categories that have been found in previous work, including those of human vocalisations, machine/transport sounds, human action sounds and nature sounds [52, 76, 37, 48, 79]. Categories are predominantly based upon perceptions associated with the sound producing event (object / action) in accordance with the theory of everyday listening (Gaver) and causal similarities (LeMaitre). The influence of context, whilst not forming specific categories simply based on a location, may be used to create categories based on the activities associated with said location, in a way modulating the categorisation of actions (that may otherwise be categorised based on only the specific action) [48]. Finally the importance of emotion, or emotional reaction, cannot be ignored as another factor that modulates categorisation, as shown un previous studies on visual and auditory categorisation [13, 19]. CIL participants for example described stimuli as those that were "liked" and "not liked" as well as common perceptions of warning sounds ("doorbell" and "telephone ringing"). Therefore influence of emotional reaction influence seems to be an increasing factor in the absence of spectral information and increased difficulty in identifying sounds, as noted by categorisation strategies of CIL and NHL in CI simulations. This is also evidenced by the categorisation of the emotion or emotional reaction within vocal stimuli that furthers previous research into the perception of voices [47] in that it over-rides the perception of age and gender. This is likely due to the importance of emotions in communication and the evolution of these process in humans [19].

Concerning the manner in which stimuli are perceived it appears that the theory of Gaver is not entirely clear-cut. When sounds are not easily identifiable it is not the case that perception is concerned "only" with factors associated with the sound signal. Rather listeners still seek to extract information about the sound producing event, for example the action, and make the sound relevant to their needs. However the grouping of musical sounds still suggests that these sounds, or when sounds are perceived as musical, can be categorised differently and are based the perception of the sound signal. Overall an approach based on that suggested by [77] regarding the use of acoustical, causal or semantic perceptions/similarities seems a likely basis for understanding auditory categorisation.

Extracting information about an object via the sound it produces is of course understandable as an evolutionary mechanism [79], and is evidenced by the finding of specific brain areas associated to different kinds of action sounds - human, animal, natural and mechanical. [79] also developed a categorisation model based on the concreteness, the size and the effect ability of the sounds, further showing the importance of the relationship between the listener and the sound producing object.

Of course there are possible further aspects to this relationship that could be explored and added to the model. This is especially true as the model is only concerned with *action* sounds and does not involve musical sounds nor human vocal sounds, however it could well be the case that vocal and human-action sounds are subsets of *human* sounds, much in similarity to the categorisation of faces and bodies concerning visual stimuli [19].

The results do not specifically outline a hierarchy of auditory classification, although they do agree on certain categories related predominantly to the perception of the sound producing event which in turn relates to how a listener may interact with the object. Certain aspects of a hierarchy may be apparent for example the robustness of the sound producing action being a low-level form of categorisation. Living and non-living sounds also appear to be part of the hierarchical structure whilst musical, environmental and human sounds may be broad level categories below this along with vocal and human action being subsets of human sounds.The influence of other information may be used to group lower level perceptions, for example emotion or the type of activity to which certain objects/actions are associated. Alongside this form of perception that follows on from theory by Gaver and LeMaitre, the perception of musical sounds suggests that these sounds are considerable different in their and a form of "musical listening" indeed exists further work should be undertaken to understand the exact conditions of when and why this is employed.

7.2 AUDITORY CATEGORISATION WITH CIL

The process of categorisation is an essential part of understanding how the world is perceived. Category knowledge can help guide the identification of a specific object and help discriminate between different sounds. In effect identification and categorisation are one in the same process which seeks to extract meaning from a stimulus via the identification of certain properties. Identification is simply a more refined/accurate level in this process [19, 115].

The study here has shown that at a level of categorical perception that experienced CIL are able to use this knowledge to discriminate between different kinds of sounds much in the same way as NHL. There is an effect of the duration of implantation that shows more experienced CIL (implantation greater than 14 months) are able to more accurately distinguish categories of musical environmental and especially human vocal sounds. CIL have also been shown to demonstrate better ability to identify auditory categories rather than identify individual sounds, in similarity to previous studies [111, 59]. However in previous studies CIL were made aware of possible categories (e.g. home, animal, office, mechanical). The current study therefore represents the first time CIL have been tested with a Free Sorting Task (FST), a more "real-world" approach to testing auditory perception/categorisation. Experienced CIL also show significant ability to categorise voices based on emotional action (laughing/crying) however struggle to perceive age or gender information [94, 65, 81, 34].

The performance of CIL is very similar to NHL in regards that categories are based on the perception (correct or in-correct) of semantic information associated to the sound producing event. This would suggest that "top-down" processes involved in categorisation are still in tact and still driving the strategies used by post-lingually deafened CIL. A strong factor in these processes is the the idea of how the listener may interact with the object [80, 76]. However the accurate processing of the acoustic information is clearly essential in allowing the listener (NHL or CIL) to create an association between the sound signal and the access to such semantic information. It appears that the functioning of the CI device delivers an adequate representation of the acoustic signal for semantic information to be access at the categorical level quite well. This seems especially true for vocal sounds and is also likely aided by the fact that CI devices and subsequent rehabilitation programmes are focussed towards restoring speech perception. However when it comes to identification of specific information, for example identifying certain objects or identifying the gender or age of a speaker, there is insufficient detail to the signal delivered by the CI device for a listener to make a link to previously held concepts, or causal information that they had previously learned during periods of functional hearing. Of course an additional aspect in this is that during periods of deafness there are cortical reorganisation that may disrupt or alter the sources of conceptual and semantic knowledge associated with auditory stimuli. The current work echoes that of others in that temporal are more easily perceived by CIL, especially those with a clear rhythmic structure [111, 122, 59, 5]. Certain spectral information also appear to be used for categorising purposes when they are highly salient or easily extracted from the signal, for pitch salience when categorising melodic musical sounds. These could be likely areas to focus on with regards to the kind of details in the acoustical signal that can be accurately delivered by the implant, perceived by the device user and further investigated for improving CIL auditory performance both auditory processing and categorisation, discrimination and possibly identification. They may represent the level of detail that may be delivered by the CI device an passed up as bottom-up processing.

As mentioned above these top-down processes seem to be used by CIL even after periods of deafness. Whether there is cortical reorganisation of such areas as identified by Lewis is yet to be understood and poses a good opportunity to further understand the development of auditory cognition in CIL in a similar way to that already done with speech and voice related stimuli [16, 79, 115, 114, 24, 67]. It is also stated that environmental sound processing shares similar networks as speech processing [59, 111]. Understanding how the whole auditory system of CIL undergoes change may play an important role to improving rehabilitation techniques for the future. It is likely that some form of cognitive reorganisation or adaptation takes place concerning these sounds and the process of categorisation, because results with CI simulations differ certain aspects, which shows the way spectrally degraded stimuli are dealt with by NHL and CIL auditory systems is different (exuding the effects of short-term training effects).

If categorical perception / categorisation processes are sufficiently retained in CIL it shows that not only are higher order processes active, but that the signal delivered to theInferior Temporal Cortex (IT) is of sufficient detail to represent the auditory category. This is possibly where the difference between categorisation and identification of a sound stimuli differ. The CI device delivers much more crude information, especially spectrally. and this reduced level of acoustic information may not allow the listener to identify individual stimuli.

It may therefore be possible to use this improved ability of categorical perception and robust top-down processing as part of rehabilitation schemes, especially in helping CIL discriminate between different kinds of sounds. By learning to differentiate categories of sounds this could help CIL to use the semantic categorical knowledge to improve the discrimination of different kinds of sounds. This may enhance the link between detection of the sound signal and the cortical representation of the auditory category. Once certain categories have been learned it may be possible to train CIL with sub-categories whereby the acoustic information would be more difficult to differentiate. In this way training would resemble a reverse hierarchy of learning [3, 2]. Whilst it may not be possible to arrive at identification performance equivalent of NHL the learning of categories could allow more accurate identification of certain sounds. For example environmental sounds pose difficulty in that they constitute a widely varying acoustics, can be generated by many sources, occur in many different locations. By dividing stimuli into categories that reflect the way in which auditory stimuli are perceived, for example the kind of source or action or activity, it may limit the variability of acoustic characteristics and allow sounds be more readily learned. It would also present a feasible manner in which to present a multitude of stimuli to a listener in which the task, to categorise stimuli was less taxing than specifically identifying individual stimuli, and would allow listeners to increase their familiarity with different sounds, an aspect important to auditory perception [6].

In order to make such a scheme a reality it is therefore essential to understand the mechanisms by which sounds are categorised. The current work suggests that context or location may not be the most ideal form of auditory category, possibly more strongly related to the activity which takes place in a particular location instead [48]. Visual aids may act differently to aid such categorisation. Rather it has been shown and in other works that the way in which a listener interacts with a sound forms the basis of many categories. Different forms of perceptions as suggested by [75, 52] would also point towards the fact that the same sound can be listened to in different ways, an evolution of the listening modes proposed by [37]. This is also a vital part of understanding that should be investigated and clarified as to how it affects auditory categorisation and how listeners relate to different sounds. The model suggested by [80] (mentioned previously) should also be considered as vital when considering the categorisation of action sounds and how the model can be applied to CIL in order to improve the perception of such sounds. It is likely that the same cognitive networks exist for post-lingually deaf CIL, however investigating whether modification due to deafness arise would be an important step to furthering understanding of action sound perception.

8

CONCLUSION

The current work tested the categorisation a variety of different sounds with groups of NHL and CIL. Testing showed that predefined categories of environmental, musical and vocal were broadly represented in the categorisation strategies of participants, however such large encompassing categories are not accurately representative of the way in which auditory categories are perceived in the real-world. The work poses another step towards better understanding of auditory perception, by focussing at the level of category perception. This is important in understanding how NHL deal with auditory environments/stimuli and how this knowledge could be applied to the understanding of auditory perception in CIL.

Testing found similar categories and categorisation strategies to previous work. Based on causal or semantic information, of which some perceptions remain robust even in the face of sever spectral degradation. It would also appear that context/location is not a strong part of these categorisation strategies. However whilst agreement exists the domain of auditory categorisation is still not firmly established and requires some more rigorous/exact studies in order to generate a fuller picture of how perceptions and potential listening modes combine.

The study also represents the first to test CIL in a FST of mixed everyday sounds. Shows that experienced CIL are able to follow the same strategies as NHL and show the same level of category identification. There is also an effect of implantation duration (implant experience) such that more experienced CIL are more similar to NHL, with more accurate category identification and a greater ability at discriminating vocal from environmental sounds.

To conclude the study a proposal is made to use categorical perception as a useful tool in rehabilitating CIL, especially in learning how to differentiate different"kinds of sound" and especially for dealing with real world environments when many different kinds of sounds could be going on at the same time. Although in order to successfully do this the theories, understandings, classification models proposed for auditory stimuli should be more clearly established. For example specific categories that exist, the perceptions that are used when judging category membership (and similarity) for example acoustic, causal and semantic, and finally the potential listening modes (e.g. musical and everyday from Gaver) that may modulate perceptions and classifications.

9

APPENDIX A: CATEGORY IDENTIFICATION BY NHL

The following appendix contains descriptions of the categories created by the three groups of NHL as described in chapter 2. As explained, participants completed a FST of sixteen common sounds and were then tasked with describing the categories that they created. Each participant group listened to stimuli in one of three conditions, for the first group sounds were natural and unmodified (NS), for the second group sounds were passed through a 16-channel vocoder (**16C!** (16C!)) and the third group listened to sounds passed through a 8-channel vocoder (**8C!** (8C!)). Vocoding methods were employed in order to simulate CI perceptions of the sounds.

As mentioned above the comments in this appendix refer to the categories that were identified by participants and NOT to individual sounds. In the following tables, the category descriptions of participants are shown relative to each of the sixteen stimuli. Therefore allowing the reader to see how each individual sound was categorised and shedding light on how it was perceived.

The descriptions were used to evaluate the category identification accuracy of each participant relevant to the three predefined categories of environmental, musical and vocal sounds. This was done by evaluating the description associated with each sound as to whether or not they corresponded to the pre-defined categories.

9.1 APPENDIX - A1 - CATEGORY DESCRIPTIONS FOR NS, 16C & 8C

Subject ID	DR	GTR	CGH	LGH	CAR	BEL	FEM	STP	WTR	XLY	VLN	HELI	ALRM	GLASS	OBOE	MALE
NS-1	Mouvements	Musicalités	Vocalises humaines	Vocalises humaines	Mouvements	Musicalités	Vocalises humaines	Mouvements	Mouvements	Musicalités	Musicalités	Mouvements	Musicalités	Mouvements	Musicalités	Vocalises humaines
NS-2	depart	joie	destruction	joie	depart	couple	couple	depart	joie	bruit	joie	destruction	bruit	destruction	bruit	couple
NS-3	vie quotidi-enne	Musique	Bruits personnels	Bruits personnels	Pollution auditive	Musique	Paroles	vie quotidi-enne	Désordre, panique	Reveil	Musique	Pollution auditive	Reveil	Désordre, panique	Musique	Paroles
NS-4	claquement	musique	voix	voix	moyen de locomotion	musique	voix	claquement	bruit	musique	musique	moyen de locomotion	bruit	bruit	moyen de locomotion	voix
NS-5	Bruits de la maison	Instruments de Musique	Bruits d'êtres humains	Bruits d'êtres humains	Moyens de transport	Instruments de Musique	Bruits d'êtres humains	Bruits de la maison	Moyens de transport	Instruments de Musique	Instruments de Musique	Moyens de transport	Bruits de la maison	Bruits de la maison	Instruments de Musique	Bruits d'êtres humains
NS-6	son produit à partir d'une action humaine sur des objets	son produit à partir d'un onstrument de musique	son produit à partir d'organe vocal humain	son produit à partir d'organe vocal humain	son produit à partir d'un vehucile à moteur	son produit à partir d'un onstrument de musique	son produit à partir d'organe vocal humain	Autres	son naturel	son produit à partir d'un onstrument de musique	son produit à partir d'un onstrument de musique	son produit à partir d'un vehucile à moteur	son produit à partir d'un onstrument de musique	son produit à partir d'une action humaine sur des objets	son produit à partir d'un onstrument de musique	son produit à partir d'organe vocal humain
NS-7	bruits do-mestiques	instruments de musique	rires	rires	transports: helicoptère, voiture, radeau	instruments de musique	paroles	bruits do-mestiques	transports: helicoptère, voiture, radeau	instruments de musique	instruments de musique	transports: helicoptère, voiture, radeau	bruits do-mestiques	bruits do-mestiques	instruments de musique	paroles
NS-8	Vie quotidi-enne	Musique	sons qui viennent de la gorge	sons qui viennent de la gorge	Vie quotidi-enne	Musique	gens qui parlent	Vie quotidi-enne	eau qui coule	sonnette	Musique	Vie quotidi-enne	Telephone	vaisselle qui tombe	Musique	gens qui parlent
NS-9	BRUITAGE	MUSIQUE	EXPRIME	EXPRIME	BRUITAGE	CLOCHE	PAROLE	BRUITAGE	BRUITAGE	CLOCHE	MUSIQUE	BRUITAGE	SONNERIE	BRUITAGE	MUSIQUE	PAROLE
NS-10	l e matin	depart de la journee	l e matin	telephone copine	depart de la journee	enterrement	enterrement	depart de la journee	l e matin	telephone copine	depart de la journee	accident	l e matin	l e matin	depart de la journee	depart de la journee

Table 9.1: Appendix A-1: category descriptions of sixteen common everyday sounds by NHL in NATURAL SOUND condition (NS).

Subject ID	DR	GTR	CGH	LGH	CAR	BEL	FEM	STP	WTR	XLY	VLN	HELI	ALRM	GLASS	OBOE	MALE
NS-11	Bruits de notre environnement	Sons provennant d'instrument de musqiue	Voix humaines	Voix humaines	Sons violents, agressifs, forts.	Bruits de notre environnement	Voix humaines	Bruits de notre environnement	Son de la nature, relaxant	Sons provennant d'instrument de musqiue	Sons provennant d'instrument de musqiue	Sons violents, agressifs, forts.	Sons violents, agressifs, forts.	Bruits de notre environnement	Sons provennant d'instrument de musqiue	Voix humaines
NS-12	bruits domestiques	Sons instruments	Voix	Voix	Bruits exterieurs	Bruits exterieurs	Voix	Bruits exterieurs	Bruits naturels	Sons instruments	Sons instruments	Bruits exterieurs	bruits domestiques	bruits domestiques	Bruits exterieurs	Voix
NS-13	bruits	notes de musique	bruits de bouche	bruits de bouche	quotidien	notes de musique	bruits de bouche	quotidien	riviere	notes de musique	notes de musique	helicptere	sonnerie de telephone	bruits	notes de musique	bruits de bouche
NS-14	répétition d'une sonorité dans le fond	musiques	répétition d'une sonorité dans le fond	répétition d'une sonorité dans le fond	moyens de lomotion	sonneries	personnes	bruit de pas	eau	sonneries	musiques	moyens de lomotion	sonneries	verre qui casse, pourrait être associé à la personne	musiques	personnes
NS-15	bruits secs	bruits plus musicaux, sonnerie	bruits de gorge bruits de raclement	bruits de gorge bruits de raclement	bruits de moyens de transport, eau qui coule de meme intensitŽ	bruits plus musicaux, sonnerie	bruits de parole	bruits secs	bruits de moyens de transport, eau qui coule de meme intensitŽ	bruits plus musicaux, sonnerie	bruits plus musicaux, sonnerie	bruits de moyens de transport, eau qui coule de meme intensitŽ	bruits plus musicaux, sonnerie	bruits secs	bruits de moyens de transport, eau qui coule de meme intensitŽ	bruits de parole
NS-16	Bruits objets quotidiens	Mélodie	Gens	Gens	Bruits objets quotidiens	Mélodie	Gens	Bruits objets quotidiens	Nature (eau...)	Mélodie	Mélodie	Bruits objets quotidiens	Bruits objets quotidiens	Bruits objets quotidiens	Mélodie	Gens
NS-17	Interactions humaines	musiques	Communication humaine	Communication humaine	véhicules motorisés	Melodies simples	humaine Communication	Interactions humaines	nature	Melodies simples	musiques	véhicules motorisés	Interactions humaines	Interactions humaines	musiques	Communication humaine
NS-18	bruit s'objet de la vie quotidienne	son melodieux, musique Most	voix	voix	bruit s'objet de la vie quotidienne	son melodieux, musique Most	voix	bruit s'objet de la vie quotidienne	son melodieux, musique Most	son melodieux, musique Most	son melodieux, musique Most	bruit s'objet de la vie quotidienne	bruit s'objet de la vie quotidienne	bruit s'objet de la vie quotidienne	son melodieux, musique Most	voix
NS-19	deplacement	musique	emotion	emotion	deplacement	temps	voix	deplacement	environment	musique	musique	environment	temps	ecat	musique	voix
NS-20	Autres	Instuments de musique	Tout ce qui provient de l'homme	Tout ce qui provient de l'homme	Véhicules	Instuments de musique	Tout ce qui provient de l'homme	Tout ce qui provient de l'homme	Autres	Instuments de musique	Instuments de musique	Véhicules	Instuments de musique	Autres	Instuments de musique	Tout ce qui provient de l'homme

Table 9.2: Appendix A-1: category descriptions of sixteen common everyday sounds by NHL in NATURAL SOUND condition (NS).

Subject ID	dr	gtr	cgh	lgh	car	bel	fem	fstp	wtr	xly	vln	heli	alrm	gls	obo	male
V16C-1	Mechanics with a wooden resonance	Bruits mÈtalliques, percussions	Mechanics with a wooden resonance	Bruits mÈcaniques dans lair	Mechanical noise in water	Mechanical noise in water	Voix de monstres	Mechanics with a wooden resonance	Mechanical noise in water	Bruits mÈtalliques, percussions	Bruits mÈtalliques, percussions	Bruits mÈcaniques dans lair	Bruits mÈcaniques dans lair	Mechanical noise in water	Bruits mÈcaniques dans lair	Voix de monstres
V16C-2	briut de bois	rythme	rythme	rythme	moteur	son brouillon	paroles	briut de bois	moteur	son brouillon	rythme	moteur	moteur	rythme	son brouillon	paroles
V16C-3	sons de lenvironnement non naturels	sons de lenvironnement non naturels	Voix humaines	Voix humaines	sons de lenvironnement non naturels	Sons de lenvironnement naturels	Voix humaines	sons de lenvironnement non naturels	Sons de lenvironnement naturels	Sons de lenvironnement naturels	sons de lenvironnement non naturels	sons de lenvironnement non naturels	Sons de lenvironnement naturels	Sons de lenvironnement naturels	Sons de lenvironnement naturels	Voix humaines
V16C-4	Porte et grrr	Bruits intenses	Porte et grrr	Bruits intenses	Véhicules:Bru voiture et hélicoptaire	Bruit de pluie et dorage	Paroles	Bruit interprétables: bruit de bobine et de pas	Bruit de pluie et dorage	Bruits continus sourds	Bruits intenses	Véhicules:Bru voiture et hélicoptaire	Bruit interprétables: bruit de bobine et de pas	Bruit de pluie et dorage	Bruits continus sourds	Paroles
V16C-5	Ces bruits évoquent des activités humaines (sur lenvironnement)	Ces bruits évoquent des bruits de lenvironnement	bruits langagiers	bruits langagiers	Ces bruits évoquent des bruits de lenvironnement	Bruits animaliers	Bruits animaliers	Ces bruits évoquent des activités humaines (sur lenvironnement)	Ces bruits évoquent des bruits de lenvironnement	Ces bruits évoquent des activités humaines (sur lenvironnement)	Ces bruits évoquent des activités humaines (sur lenvironnement)	Ces bruits évoquent des activités humaines (sur lenvironnement)	Ces bruits évoquent des activités humaines (sur lenvironnement)	Ces bruits évoquent des bruits de lenvironnement	Bruits animaliers	bruits langagiers
V16C-6	bruits de pas et de porte susceptibles dêtre entendus dans une maison	bruits de pas et de porte susceptibles dêtre entendus dans une maison	bruits de pas et de porte susceptibles dêtre entendus dans une maison	film dhorreur	eau qui coule	eau qui coule	film dhorreur	bruits de pas et de porte susceptibles dêtre entendus dans une maison	eau qui coule	film dhorreur	inclassables	inclassables	eau qui coule	inclassables	inclassables	film dhorreur
V16C-7	son amplifié	raisonnance	son amplifié	film dhorreur	écoulement	écoulement	Voies déformées	clic cloc	écoulement	raisonnance	raisonnance	écoulement	film dhorreur	écoulement	film dhorreur	Voies déformées
V16C-8	bruit fait par un humain en interagissant avec un objet	bruits artificiels	voix humaine	voix humaine	bruit fait par un humain en interagissant avec un objet	bruits deau	voix humaine	bruit fait par un humain en interagissant avec un objet	bruits deau	bruits artificiels	bruits artificiels	bruit fait par un humain en interagissant avec un objet	entre bruit artificiel et bruit crée par un humain	entre bruit artificiel et bruit crée par un humain	bruits artificiels	voix humaine
V16C-9	Savais pas trop ou les classer	Savais pas trop ou les classer	Voix déformées	Voix déformées	Aquatique	Voix déformées	Voix déformées	Bruits "simples"	Aquatique	Bruits "simples"	Voix déformées	Aquatique	Aquatique	Aquatique	Voix déformées	Voix déformées
V16C-10	bruits insolites	bruits insolites	Voix humaines	Voix humaines	Véhicules	Véhicules	Voix humaines	Horloge	Eau	Horloge	Voix humaines	Véhicules	Nature (vent, insectes	Eau	Nature (vent, insectes	Voix humaines
V16C-11	bruits dobjects metaliques ou en bois qui se cognent	bruits dobjects metaliques ou en bois qui se cognent	bruits de voix (ou ressemblance de voix)	bruits de voix (ou ressemblance de voix)	bruits en rapport avec leau	bruits dobjects metaliques ou en bois qui se cognent	bruits de voix (ou ressemblance de voix)	Inclassables de par leur bizzarerie	bruits en rapport avec leau	bruits dobjects metaliques ou en bois qui se cognent	Inclassables de par leur bizzarerie	bruits en rapport avec leau	Inclassables de par leur bizzarerie	bruits en rapport avec leau	Inclassables de par leur bizzarerie	bruits de voix (ou ressemblance de voix)

Table 9.3: Appendix A-2a: category descriptions of sixteen common everyday sounds by NHL in VOCODED 16 channels condition (V16C).

Subject ID	dr	gtr	cgh	lgh	car	bel	fem	fstp	wtr	xly	vln	heli	alrm	gls	obo	male
V16C-12	Une personne qui exécute une action	Percussions	Eau	Voix	Engins	Eau	Voix	Une personne qui exécute une action	Eau	Percussions	Percussions	Engins	Engins	Eau	Soufle	Voix
V16C-13	bruits déplacements	bruits mécaniques	voix	bruits mécaniques	bruits mécaniques	bruits mécaniques	voix	bruits déplacements	bruits deau	? différent des autres	bruits mécaniques	bruits mécaniques	bruits deau	bruits deau	bruits mécaniques	voix
V16C-14	bruit intérieur	bruit de quelque chose qui tombe	son humain	son humain	son véhicule en général	bruit deau	son humain	son véhicule en général	bruit deau	bruit de quelque chose qui tombe	indéterminé	son véhicule en général	bruit intérieur	bruit deau	indéterminé	son humain
V16C-15	son réel	ressemble tous les 2 au même instrument	son amplifié, de même type	son amplifié, de même type	2 véhicules	son amplifié, de même type	voix déformées	son réel	son réel	ressemble tous les 2 au même instrument	son amplifié, de même type	2 véhicules	son réel	son amplifié, de même type	son amplifié, de même type	voix déformées
V16C-16	bruits quotidiens à la maison: porte, réveil, pas, personne qui parle	trucs qui rebondissent dans escalier, ou boomerang, ou truc qui raisonnent	bruit dans la nature: vent, eau, trucs qui rebondies dans environnement qui raisonne	trucs qui rebondissent dans escalier, ou boomerang, ou truc qui raisonnent	** bruits de véhicules (voiture, hélico) **	nature: bruits danimaux et orage	nature: bruits danimaux et orage	bruits quotidiens à la maison: porte, réveil, pas, personne qui parle	nature: bruits danimaux et orage	trucs qui rebondissent dans escalier, ou boomerang, ou truc qui raisonnent	trucs qui rebondissent dans escalier, ou boomerang, ou truc qui raisonnent	** bruits de véhicules (voiture, hélico) **	bruits quotidiens à la maison: porte, réveil, pas, personne qui parle	même catégorie que S3	même catégorie que S3	bruits quotidiens à la maison: porte, réveil, pas, personne qui parle
V16C-17	bruit clair et distinct	bruit métallique avec écho	bruits bizarres avec échos, à priori il ya un rire et des bruits de toux	bruits bizarres avec échos, à priori il ya un rire et des bruits de toux	dans le lot il y a un moteur qui démarre mais ça me fait penser à de leau qui coule	bruits bizarres avec échos, à priori il ya un rire et des bruits de toux	paroles incompréhensibles	bruit clair et distinct	dans le lot il y a un moteur qui démarre mais ça me fait penser à de leau qui coule	bruit métallique avec écho	bruits bizarres avec échos, à priori il ya un rire et des bruits de toux	bruit clair et distinct	bruit clair et distinct	dans le lot il y a un moteur qui démarre mais ça me fait penser à de leau qui coule	bruits bizarres avec échos, à priori il ya un rire et des bruits de toux	paroles incompréhensibles
V16C-18	Bruits qui contiennent des "impulsions"	Bruits qui contiennent des "impulsions"	Bruits qui contiennent des "impulsions"	Bruits qui contiennent des "impulsions"	bruits continus	parole	parole	Bruits qui contiennent des "impulsions"	Bruits "deau"	Bruits qui contiennent des "impulsions"	Bruits qui contiennent des "impulsions"	bruits continus	bruits continus	Bruits "deau"	Bruits "deau"	parole
V16C-19	bruits autres	bruits autres	bruits autres	bruits danimaux	moyens de locomotions	moyens de locomotions	voix humaines	moyens de locomotions	eau	bruits autres	bruits danimaux	moyens de locomotions	bruits autres	eau	moyens de locomotions	voix humaines
V16C-20	Deplacement dans un batiment	Musique concrete	Bruits humains	Bruits humains	Vehicules	Souffle	Paroles	Deplacement dans un batiment	Eau	Musique concrete	Musique concrete	Vehicules	Reveil	Eau	Souffle	Paroles
V16C-21	sons saccadés	bruit dimpact sur objet ou dans lair	sons saccadés	sons saccadés	bruit en rapport avec de leau	bruit dimpact sur objet ou dans lair	voix trafiquée	sons saccadés	bruit en rapport avec de leau	bruit dimpact sur objet ou dans lair	voix trafiquée	son dhélice	son dhélice	bruit en rapport avec de leau	bruit en rapport avec de leau	voix trafiquée

Table 9.4: Appendix A-2b: category descriptions of sixteen common everyday sounds by NHL in VOCODED 16 channels condition (V16C).

Subject ID	dr	gtr	cgh	lgh	car	bel	fem	fstp	wtr	xly	vln	heli	alrm	gls	obo	male
V8C-1	Humain	Drums	Humain	Humain	Humain	bruits aquatiques	Voix	Humain	bruits aquatiques	Drums	Drums	Humain	bruits aquatiques	bruits aquatiques	Humain	Voix
V8C-2	Electronic sounds (pinball)	Electronic sounds (pinball)	Weird other sounds	Weird other sounds	Water sounds	Artificial Voice	Artificial Voice	Regular sounds	Water sounds	Weird other sounds	Electronic sounds (pinball)	Regular sounds	Weird other sounds	Weird other sounds	Artificial Voice	Artificial Voice
V8C-3	bruit de la vie active (talon, douche..)	bruit de la vie active (talon, douche..)	indeterminé	sports (tir, balle, frisbie)	moyen deplacement, transport (hélico, bateau, voiture ...)	moyen deplacement, transport (hélico, bateau, voiture ...)	sports (tir, balle, frisbie)	bruit de la vie active (talon, douche..)	moyen deplacement, transport (hélico, bateau, voiture ...)	sports (tir, balle, frisbie)	indeterminé	moyen deplacement, transport (hélico, bateau, voiture ...)	bruit de la vie active (talon, douche..)	bruit de la vie active (talon, douche..)	moyen deplacement, transport (hélico, bateau, voiture ...)	bruit de la vie active (talon, douche..)
V8C-4	Sons aigus enchainés vite	Répétitions fréquences graves	Sons aigus enchainés vite	Sons aigus enchainés vite	Bruits de moteur	Répétitions fréquences graves	Voix moidifiée	Bruits de pas	Sons aigus enchainés vite	Répétitions fréquences graves	Sons aigus enchainés vite	Bruits de moteur	Bruits de moteur	Sons aigus enchainés vite	Chasse deau	Voix moidifiée
V8C-5	Sons qui me font penser à des bruits darmes ou de passages dans les jeux dordinateurs	Sons qui me font penser à des intépéries météorologiques un orage qui gronde, de la grèle qui tombe sur de la tole, une pierre ...	Sons qui me font penser à des intépéries météorologiques un orage qui gronde, de la grèle qui tombe sur de la tole, une pierre ...	Catégorie de son qui me font penser à des véhicules: une voiture qui démarre, un hélicoptère, un camion qui roule, un vieille vo...	Catégorie de son qui me font penser à des véhicules: une voiture qui démarre, un hélicoptère, un camion qui roule, un vieille vo...	Sons qui me font penser à des intépéries météorologiques un orage qui gronde, de la grèle qui tombe sur de la tole, une pierre ...	Les sons sont comme si ce sont des personnages électroniques qui parlent (robots ou jeux vidéo)	Son qui me fait penser à qulquun qui tape un pclou dans une planche de bois avec un marteau	Sons qui me font penser à des intépéries météo: un orage qui gronde, de la grèle qui tombe sur de la tole, une pierre ...	Sons qui me font penser à des bruits darmes ou de passages dans les jeux dordinateurs	Sons qui me font penser à des bruits darmes ou de passages dans les jeux dordinateurs	Catégorie de son qui me font penser à des véhicules: une voiture qui démarre, un hélicoptère, un camion qui roule, un vieille vo...	Catégorie de son qui me font penser à des véhicules: une voiture qui démarre, un hélicoptère, un camion qui roule, un vieille vo...	Sons qui me font penser à des intépéries météo: un orage qui gronde, de la grèle qui tombe sur de la tole, une pierre ...	Catégorie de son qui me font penser à des véhicules: une voiture qui démarre, un hélicoptère, un camion qui roule, un vieille vo...	Les sons sont comme si ce sont des personnages électroniques qui parlent (robots ou jeux vidéo)
V8C-6	bruit mat	bruit resonnant	parole, voix	parole, voix	bruit mat	bruit resonnant	parole, voix	bruit mat	bruit mat	bruit resonnant	bruit resonnant	bruit mat	bruit mat	bruit mat	bruit mat	parole, voix
V8C-7	portes qui souvrent et se ferment	choses qui tombent	pluie-eau qui tombe	divers	bruits de machines (voiture, hélico...)	choses qui tombent	jeux vidéos	portes qui souvrent et se ferment	pluie-eau qui tombe	portes qui souvrent et se ferment	divers	bruits de machines (voiture, hélico...)	bruits de machines (voiture, hélico...)	pluie-eau qui tombe	divers	jeux vidéos
V8C-8	entrée dans une pièce	rythme	eau	bruit de machine	eau	eau	voix informatique	entrée dans une pièce	eau	rythme	rythme	bruit de machine	bruit de machine	eau	bruit de machine	voix informatique
V8C-9	Ce groupe pur moi correpond plus à des omatopées. On peut les entendre comme bruit de fond pour un film	Ce groupe pur moi correpond plus à des omatopées. On peut les entendre comme bruit de fond pour un film	Ce groupe pur moi correpond plus à des omatopées. On peut les entendre comme bruit de fond pour un film	Ce groupe pur moi correpond plus à des omatopées. On peut les entendre comme bruit de fond pour un film	Ce groupe est le groupe des moyens de transport la voiture et lhellicoptere	ce groupe correspond au bruit que lon peut entendre dans la nature	Ce dernier gropue correpond au differentes "voies" que lon peut entendre cris animaux ou homme	ce groupe correspond au bruit que lon peut entendre dans la nature	ce groupe correspond au bruit que lon peut entendre dans la nature	Ce groupe pur moi correpond plus à des omatopées. On peut les entendre comme bruit de fond pour un film	Ce groupe pur moi correpond plus à des omatopées. On peut les entendre comme bruit de fond pour un film	Ce groupe est le groupe des moyens de transport la voiture et lhellicoptere	ce groupe correspond au bruit que lon peut entendre dans la nature	Ce groupe pur moi correpond plus à des omatopées. On peut les entendre comme bruit de fond pour un film	Ce dernier gropue correpond au differentes "voies" que lon peut entendre cris animaux ou homme	Ce dernier gropue correpond au differentes "voies" que lon peut entendre cris animaux ou homme

Table 9.5: Appendix A-3a: category descriptions of sixteen common everyday sounds by NHL in VOCODED 8 channel condition (V8C).

Subject ID	dr	gtr	cgh	lgh	car	bel	fem	fstp	wtr	xly	vln	heli	alrm	gls	obo	male
V8C-10	bruits "de la maison" ** s8 "pas sur le parquet" ** s1 "porte qui souvre	de lunivers du train, du métallique	sons avec effet "dans leau" ** s9 "chaine qui sort de leau" ** s3 "coups de feu dans leau	les "boomrangs	bruit de moteur ou dengins motorisé... ** s13 "moteur qui démarre pas" ** s5 "moteur qui démarre	de lunivers du train, du métallique	"voix modifiées"...	bruits "de la maison" ** s8 "pas sur le parquet" ** s1 "porte qui souvre	sons avec effet "dans leau" ** s9 "chaine qui sort de leau" ** s3 "coups de feu dans leau	pas très liés, mais sons "unique et long"	les "boomrangs	bruit de moteur ou dengins motorisé... ** s13 "moteur qui démarre pas" ** s5 "moteur qui démarre	bruit de moteur ou dengins motorisé... ** s13 "moteur qui démarre pas" ** s5 "moteur qui démarre	sons avec effet "dans leau" ** s9 "chaine qui sort de leau" ** s3 "coups de feu dans leau	pas très liés, mais sons "unique et long"	"voix modifiées"...
V8C-11	le reste... plutôt machine, percussif, en extérieur	percussifs avec echo, vaste intérieur	sons avec composante liquide importante	percussifs avec echo, vaste intérieur	sons avec composante liquide importante	sons avec composante liquide importante	composantes vocales	percussifs avec echo, vaste intérieur	sons avec composante liquide importante	percussifs avec echo, vaste intérieur	le reste... plutôt machine, percussif, en extérieur	le reste... plutôt machine, percussif, en extérieur	percussifs avec echo, vaste intérieur	sons avec composante liquide importante	le reste... plutôt machine, percussif, en extérieur	composantes vocales
V8C-12	rythme, tons	rythme, tons	raclement de gorge	raclement de gorge	démarrage, moteur	temps (orage, pluie)	voix déformés	rythme, tons	temps (orage, pluie)	rythme, tons	rythme, tons	démarrage, moteur	temps (orage, pluie)	temps (orage, pluie)	temps (orage, pluie)	voix déformés
V8C-13	bruits de lenvironnement	instruments, musique	voix humaines	voix humaines	bruits de lenvironnement	bruits de lenvironnement	voix humaines	bruits de lenvironnement	bruits de lenvironnement	bruits de lenvironnement	instruments, musique	bruits de lenvironnement	instruments, musique	bruits de lenvironnement	voix humaines	voix humaines
V8C-14	bruit humain dans un contexte manufacturé (maison, route...)	bruit humain dans un contexte manufacturé (maison, route...)	bruit humain dans un contexte manufacturé (maison, route...)	bruit humain dans un contexte manufacturé (maison, route...)	vehicules	bruits naturels (on secroirait en pleine nature je veux dire)	voix mofiées (en tout cas jai limpression)	bruit humain dans un contexte manufacturé (maison, route...)	bruits naturels (on secroirait en pleine nature je veux dire)	bruit humain dans un contexte manufacturé (maison, route...)	bruit humain dans un contexte manufacturé (maison, route...)	vehicules	bruits naturels (on secroirait en pleine nature je veux dire)	bruits naturels (on secroirait en pleine nature je veux dire)	bruit humain dans un contexte manufacturé (maison, route...)	voix mofiées (en tout cas jai limpression)
V8C-15	bruits quotidients	DJ	eau	DJ	bruits quotidients	voix trafiquée	voix trafiquée	bruits quotidients	eau	DJ	DJ	bruits quotidients	bruits quotidients	eau	voix trafiquée	voix trafiquée
V8C-16	Bruits de notre environnement	Sons provenant dinstruments de musqiue	Voix humaines	Voix humaines	Sons violents, agressifs, forts.	Bruits de notre environnement	Voix humaines	Bruits de notre environnement	Son de la nature, relaxant	Sons provenant dinstruments de musqiue	Sons provenant dinstruments de musqiue	Sons violents, agressifs, forts.	Sons violents, agressifs, forts.	Bruits de notre environnement	Sons provenant dinstruments de musqiue	Voix humaines
V8C-17	sons "musicaux"	sons "musicaux"	sons "musicaux"	nature	bruits "moteurs"	nature	bruits "humain	bruits "humain	nature	bruits "moteurs"	sons "musicaux"	bruits "moteurs"	sons "musicaux"	nature	bruits "moteurs"	bruits "humain
V8C-18	sons du quotidien	sons de tambours avec fréquence	sons du quotidien	sons de tambours avec fréquence	moteurs	évenements climatiques	voix	sons du quotidien	évenements climatiques	sons du quotidien	sons de tambours avec fréquence	moteurs	son non déterminé	évenements climatiques	moteurs	voix
V8C-19	son composé	son simple	bruit deau	son simple	bruit mecanique	bruit deau	parole	son simple	bruit deau	son simple	son composé	bruit mecanique	bruit mecanique	bruit deau	bruit mecanique	parole
V8C-20	autres indeterminés	rythmique	monstres, horreur, voix deformées	monstres, horreur, voix deformées	voitue, hélico (transports)	eau, orage, mauvais temps	monstres, horreur, voix deformées	rythmique	eau, orage, mauvais temps	rythmique	rythmique	voitue, hélico (transports)	eau, orage, mauvais temps	eau, orage, mauvais temps	autres indeterminés	monstres, horreur, voix deformées

Table 9.6: Appendix A-3b: category descriptions of sixteen common everyday sounds by NHL in VOCODED 8 channel condition (V8C).

10

APPENDIX B: CATEGORY AND SOUND IDENTIFICATION BY CIL

The following appendix contains descriptions of the categories created by the three groups of CIL, Experienced cochlear implant listeners (EXP), Intermediate cochlear implant listeners (INT) and New cochlear implant listeners (NEW) as described in chapter 3. As explained, participants completed a FST of sixteen common sounds and were then tasked with describing the categories that they created.

As mentioned above the comments in this appendix refer to the categories that were identified by participants and NOT to individual sounds. In the following tables, the category descriptions of participants are shown relative to each of the sixteen stimuli. Therefore allowing the reader to see how each individual sound was categorised and shedding light on how it was perceived.

The descriptions were used to evaluate the category identification accuracy of each participant relevant to the three predefined categories of environmental, musical and vocal sounds. This was done by evaluating the description associated with each sound as to whether or not they corresponded to the pre-defined categories.

Subject ID	DR	GTR	CGH	LGH	CAR	BEL	FEM	STP	WTR	XLY	VLN	HELI	ALRM	GLASS	OBOE	MALE
EXP-1	No comment	deux mêmes bruits	No comment	deux voix	bruits de véhicule	deux mêmes bruits	sons se ressemblent	No comment	bruits de véhicule	cloche	sons se ressemblent	motos	motos	bruits de véhicule	bruits de véhicule	deux voix
EXP-2	BALAYAGE DANS UNE COUR AVEC DE L EAU	melodies	VOIX HUMAINES	VOIX HUMAINES	ENGINS A MOTEUR	SONNERIES	VOIX HUMAINES	MARTEAU, PAS D'AUTRE ASSOCIATION	BALAYAGE DANS UNE COUR AVEC DE L EAU	SONNERIES	melodies	ENGINS A MOTEUR	SONNERIES	BALAYAGE DANS UNE COUR AVEC DE L EAU	SONNERIES	VOIX HUMAINES
EXP-3	déplacement humain	sonneries musicales	sons saccadés	sons saccadés	sons désagréables	sonneries musicales	voix	déplacement humain	sons désagréables	sonneries musicales	sonneries musicales	sons saccadés	sonneries musicales	sons désagréables	sons désagréable, mais continu et plus aigu	voix
EXP-4	bruit de moteur	instruments de musique	voix humaines	voix humaines	bruit de moteur	instruments de musique	voix humaines	briut de pas	No comment	sonnette d'entrée	bruits de trompette	bruit de moteur	reveil	No comment	bruits de trompette	voix humaines
EXP-5	bruits domestiques	Instruments de musique	bruits domestiques	Des Sons émis par des gens	bruits d'extérieur (	bruits d'extérieur (	Des Sons émis par des gens	Instruments de musique	bruits d'extérieur (	Instruments de musique	Instruments de musique	bruits d'extérieur (	bruits domestiques	bruits domestiques	Instruments de musique	Des Sons émis par des gens
EXP-6	Sonnerie de la porte, ouverture et pas	musique	bruits de démarrage et de moteur qui tourne	rire	bruits de démarrage et de moteur qui tourne	vache et klaxon	Sonnerie de la porte, ouverture et pas	Sonnerie de la porte, ouverture et pas	bruits de démarrage et de moteur qui tourne	Sonnerie de la porte, ouverture et pas	musique	bruit non reconnu	Sonnerie de la porte, ouverture et pas	bruits de démarrage et de moteur qui tourne	vache et klaxon	bruit non reconnu
EXP-7	bruits de la rue	mélodies	mélodies	voix, rires, personnages qui parlent	bruits de la rue	bruits de la rue	voix, rires, personnages qui parlent	bruits de la rue	bruits	son de cloche	mélodies	bruits de la rue	bruits	bruits de la rue	bruits de la rue	voix, rires, personnages qui parlent
EXP-8	bruits non identifiés	mélodies	mélodies	voix humaines	bruits de la vie courante	musiques	musiques	bruit non identifié (sabot de cheval ?)	bruits non identifiés	bruits de la vie courante	mélodie	bruits de véhicules (klaxons, etc...)	bruits de la vie courante	bruits de la vie courante	bruits de véhicules (klaxons, etc...)	voix humaines
EXP-9	bruit	musique	bruit	rire	bruit	musique	parole	pas	bruit	sons de piano	parole	bruit	bruit	bruit	sons de piano	parole

Table 10.1: Appendix B-1: category descriptions of sixteen common everyday sounds by Experienced CIL (EXP)

Subject ID	DR	GTR	CGH	LGH	CAR	BEL	FEM	STP	WTR	XLY	VLN	HELI	ALRM	GLASS	OBOE	MALE
EXP-10	Bruits de la maison	musique	Bruits de la maison	une personne, un être humain	bruits de la rue ** S9 perçu comme un marteau piqueur	bruits de la rue ** S9 perçu comme un marteau piqueur	une personne, un être humain	une personne, un être humain	bruits de la rue ** S9 perçu comme un marteau piqueur	SONNERIES	musique	bruits de la rue ** S9 perçu comme un marteau piqueur	SONNERIES	Bruits de la maison	SONNERIES	une personne, un être humain
EXP-11	groupe "voitures et motos" : porte, pas, klaxxn, moteur, bruits de verre, de moteur	histoire : quelq'un s'éclaircit la voix pour chanter puis joue de la musique,	histoire : quelq'un s'éclaircit la voix pour chanter puis joue de la musique,	Histoire : réveil qui sonne puis femme qui dit m'enfin et qui rit.	groupe "voitures et motos" : porte, pas, klaxxn, moteur, bruits de verre, de moteur	histoire : quelq'un s'éclaircit la voix pour chanter puis joue de la musique,	Histoire : réveil qui sonne puis femme qui dit m'enfin et qui rit.	groupe "voitures et motos" : porte, pas, klaxxn, moteur, bruits de verre, de moteur	groupe "voitures et motos" : porte, pas, klaxxn, moteur, bruits de verre, de moteur	début d'un générique d'émission	début d'un générique d'émission	n'allait dans aucun autre groupe	Histoire : réveil qui sonne puis femme qui dit m'enfin et qui rit.	groupe "voitures et motos" : porte, pas, klaxxn, moteur, bruits de verre, de moteur	groupe "voitures et motos" : porte, pas, klaxxn, moteur, bruits de verre, de moteur	histoire : quelq'un s'éclaircit la voix pour chanter puis joue de la musique,
EXP-12	BRUITS DE LA VIE COURANTE IDENTIFIES	INSTRUMENTS A CORDES FRAPPEES	INSTRUMENTS DE MUSIQUE A PERCUSSIONS	VOIX HUMAINES	BRUITS NON IDENTIFIES	INSTRUMENTS A VENT	VOIX HUMAINES	BRUITS DE LA VIE COURANTE IDENTIFIES	MOINS CERTAIN DE LA NATURE EXACTE DU BRUIT MAIS A RECONNU LE BRUIT D EAU QUI COULE	INSTRUMENTS A CORDES FRAPPEES	INSTRUMENTS A CORDES FROTTEES	BRUITS DE LA VIE COURANTE IDENTIFIES	BRUITS NON IDENTIFIES	INSTRUMENTS DE MUSIQUE A PERCUSSIONS	INSTRUMENTS A VENT	VOIX HUMAINES
EXP-13	ENTEND 2 SONS	SONS DE MUSIQUE	LES GENS RIENT	LES GENS RIENT	ENTEND BEAUCOUP DE SONS	SONS DE MUSIQUE	PAROLE	SONS DE MUSIQUE	SONS DE MUSIQUE	ENTEND 2 SONS	SONS DE MUSIQUE	SONS DE MUSIQUE	N A PAS ENTENDU LE SON	ENTEND BEAUCOUP DE SONS	ENTEND 2 SONS	PAROLE
EXP-14	rythme se completent.	tonalités se rapprochant	rythme se completent.	bruits aigus	Plus aigues tonalités crescendo	tonalités se rapprochant	rythme se completent.	bruit de pas	Plus aigues tonalités crescendo	rythme se completent.	Plus aigues tonalités crescendo	moteurs	moteurs	bruits aigus	Plus aigues tonalités crescendo	tonalités se rapprochant
EXP-15	PORTE QUI S OUVRE	MELODIES	TOUX A REGROUPER AVEC LE GROUPE "PAROLES"	VOIX PAROLES	ENVIRONNI	MELODIES	VOIX PAROLES	DES PAS	ENVIRONNI	MELODIES	MUSIQUE A REGROUPER AVEC GROUPE "MELODIES"	HELICOPTE A REGROUPER AVEC ENVIRONNEMENT	SONNERIES	SONNERIES	KLAXON	VOIX PAROLES
EXP-16	bruits du quotidien	mélodies	voix humaines	voix humaines	bruits du quotidien	mélodies	sonneries	bruits du quotidien	bruits du quotidien	sonneries	mélodies	bruits du quotidien	sonneries	bruits du quotidien	sonneries	voix humaines

Table 10.2: Appendix B-1: category descriptions of sixteen common everyday sounds by Experienced CIL (EXP)

Subject ID	DR	GTR	CGH	LGH	CAR	BEL	FEM	STP	WTR	XLY	VLN	HELI	ALRM	GLASS	OBOE	MALE
INT-1	vie quotidienne	vie quotidienne	vie quotidienne	No comment	vie quotidienne	music	vie quotidienne	music	music	music	voie humaine	music	vie quotidienne	vie quotidienne	voie humaine	voie humaine
INT-2	bruits de la maison	bruits de la maison	voiture, demarrage de vehicule	la vie courant	savez pas	la musique, lorchestre, achords des instruements	la musique, lorchestre, achords des instruements	la musique, lorchestre, achords des instruements	la musique, lorchestre, achords des instruements	la vie courant	la vie courant	la musique, lorchestre, achords des instruements	voiture, demarrage de vehicule	revei matin	conversation	conversation
INT-3	sounds that were liked	sounds that were liked	sounds that were liked	Sounds that were not-liked	sounds that were liked	sounds that were liked	sounds that were liked	sounds that were liked	sounds that were liked	sounds that were liked	sounds that were liked	sounds that were liked	sounds that were liked	sounds that were liked	sounds that were liked	sounds that were liked
INT-4	bricolage	travaux, electromenager	?	bricolage	Musiaue et instruments	Musiaue et instruments	Musiaue et instruments	Cloches, sonneries	Musiaue et instruments	Musiaue et instruments	Voix humaine	?	travaux, electromenager	travaux, electromenager	Cloches, sonneries	Voix humaine
INT-5	petit bruit, quoticien	gravs, roulemet	bruit de la rue	bruit de la rue	petit bruit, quoticien	sonnerie, jingle	vache, monclemnet	plus agu	vache, monclemnet	petit bruit, quoticien	voie	sonnerie, jingle	gravs, roulemet	sonnerie, jingle	plus agu	voie
INT-6	bruits domestique	bruits domestique	bruit mechanique, voiture avec klaxin	son naturel, pas et rigole denfant	bruits domestique	guitare classique	bruit mechanique, voiture avec klaxin	guitare classique	violin bass classique ensemble	spielberg, , john belucci	son naturel, pas et rigole denfant	violin bass classique ensemble	bruit mechanique, voiture avec klaxin	bruits domestique	spielberg, , john belucci	spielberg, , john belucci
INT-7	je ne saia pas de quel son il s agit	deux buits qui se ressemblent s2 et s9	deux bruits que je ne sais pas definir	je ne saia pas de quel son il s agit	des bruits musicaux que je reconnais bien	des bruits musicaux que je reconnais bien	des bruits musicaux que je reconnais bien	deux buits qui se ressemblent s2 et s9	des bruits musicaux que je reconnais bien	deux buits qui se ressemblent s2 et s9	rire enfant	des bruits musicaux que je reconnais bien	deux bruits que je ne sais pas definir	des bruits musicaux que je reconnais bien	je ne saia pas de quel son il s agit	je ne saia pas de quel son il s agit
INT-8	quelqun qui marche	pluit qui tombe	pluit qui tombe	quelqun qui marche	je sais pas	music	police siren	music	music	quelqun qui marche	personnes qui parle	music	pluit qui tombe	sonniere	personnes qui parle	personnes qui parle
INT-9	le pas	les chose qui coule	les chose qui coule	le pas	les chose qui coule	la musique	moteurs, klaxon, vehicule	la musique	la musique	la musique	voix, parole	la musique	moteurs, klaxon, vehicule	les chose qui coule	voix, parole	voix, parole
INT-10	la vie courante	la vie courante	la vie courante	la vie courante	la musique	la musique	la musique	la musique	la musique	la musique	la parole vocale	la musique	la vie courante	la vie courante	la parole vocale	la parole vocale
INT-11	bruits du cuisine	moto	sonnerie	bruits du cuisine	fete	sonnerie	music	fete	bruits du cuisine	personne	music	personne	personne	music	moto	personne
INT-12	sons du vie quotidien	sons du vie quotidien	sons danimaux	sons du vie quotidien	sons rhytmique	sons musique	sons danimaux	sons musique	sons musique	sons rhytmique	sons emit pars les humaines	sons danimaux	sons rhytmique	sons du vie quotidien	sons emit pars les humaines	sons emit pars les humaines
INT-13	porte	eclat du voix	qqc qui trainer	parole	qqc qui trainer	siren	sons musical	porte	sons musical	sons musical	eclat du voix	siren	eclat du voix	qqc qui trainer	parole	parole
INT-14	automobile	coulement deau, ne mattre pas avec es autres	automobile	teppement, ne mettre pas avec les autres	sonneries differents	sons musical, orchestre	sonneries differents	sonneries	sons musical, orchestre	sons musical, orchestre	sons humain, parole	sons musical, orchestre	automobile	sonneries	sons humain, parole	sons humain, parole
INT-15	bruits quotidien	bruits quotidien	des klaxonx, revels sonerie	bruits quotidien	bruits quotidien	sons musical, instruemets musique	des klaxonx, revels sonerie	sons musical, instruemets musique	sons musical, instruemets musique	de vie quotidien	voix humaine	sons musical, instruemets musique	de vie quotidien	des klaxonx, revels sonerie	bruits quotidien	voix humaine

Table 10.3: Appendix B-2: category descriptions of sixteen common everyday sounds by Intermediate CIL (INT)

Subject ID	DR	GTR	CGH	LGH	CAR	BEL	FEM	STP	WTR	XLY	VLN	HELI	ALRM	GLASS	OBOE	MALE
NEW-1	porete qui freme	leau qui coule	les pas	bruit metalique	musique, bruit instrument	musique, bruit instrument	musique, bruit instrument	musique, bruit instrument	musique, bruit instrument	sons emit de les personne	banoit du vide, leau a vide	bruit du motor	bruit du motor	alert	sons emit de les personne	sons emit de les personne
NEW-2	etaps, vhaval qui marche	demarerv du motor	etaps, vhaval qui marche	demarerv du motor	notes de musique	notes de musique	une seul note	notes de musique	engroue, la gorge	demarerv du motor	notes de musique	bruit des motor	demarerv du motor	bruit des motor	quelqun qui parle	quelqun qui parle
NEW-3	sons du pas	sons du voiture	sons du pas	sons humaines	musique	sons humaines	cloche, blooaaam	rire, cattarie	cloche, blooaaam	rire, cattarie	musique	sonnerie	sons du voiture	sonnerie	voix	voix
NEW-4	bruits des personnes	bruit des moteurs, les bruits different	bruits des personnes	bruit des moteurs, les bruits different	bruits musical, des instruements	bruits des personnes	bruits musical, des instruements	bruits musical, des instruements	bruits musical, des instruements	bruits des personnes	bruits musical, des instruements	bruit des moteurs, les bruits different	bruit des moteurs, les bruits different	bruits des personnes	bruits musical, des instruements	bruits des personnes
NEW-5	port grince	sounds of things falling	talon	sounds of things falling	sound of a door sonnerie	unidentified sounds	sound of a door sonnerie	musique	unidentified sounds	laughter	not sure	unidentified sounds	sounds of things falling	sounds of things falling	sons de parole	sons de parole
NEW-6	remarrer du voiture/-moteur	bruits forts	tapping	bruits forts	piano, musique	sonnerie	cloche	piano, musique	piano, musique	parole, voix	je sias pas	sonnerie du telephone	remarrer du voiture/-moteur	sonnerie du telephone	parole, voix	parole, voix
NEW-7	Lange parle, communication	Bruit mechanique - motos	Bruits dalerte	Bruits de la maison	Bruits dalerte	Bruit mechanique - motos	Bruits de la maison	Bruits de la maison	Lange parle, communication	Bruits de la maison	Bruits de la maison	Lange parle, communication	Bruits dalerte	Bruits dalerte	Bruits dalerte	Bruits de la maison
NEW-8	bruits nouveau	Music	voix, music/parole	bruits denvironments	Music	bruits denvironments	peut-etre rie, mais je ne sais pas	bruits nouveau	bruits nouveau	bruits denvironments	Music	bruits nouveau	Music	voix, music/parole	bruits denvironments	bruits denvironments

Table 10.4: Appendix B-3: category descriptions of sixteen common everyday sounds by Newly implanted CIL (NEW)

Subject ID	DR	GTR	CGH	LGH	CAR	BEL	FEM	STP	WTR	XLY	VLN	HELI	ALRM	GLASS	OBOE	MALE
NEW-9	plus difficile, presence humaines	objets, quand on a deplace	objets, quand on a deplace	objets, quand on a deplace	plus difficile, presence humaines	bruit dappereils, des utils	bruit dappereils, des utils	melody, instruemtn musique	objets, quand on a deplace	objets, quand on a deplace	melody, instruemtn musique	objets, quand on a deplace	objets, quand on a deplace	objets, quand on a deplace	melody, instruemtn musique	melody, instruemtn musique
NEW-10	bruits des pas	Souflement	bruits des pas	siflement	Souflement	Souflement	murmure, en bas, tres doucement	murmure, en bas, tres doucement	Souflement	murmure, en bas, tres doucement	murmure, en bas, tres doucement	murmure, en bas, tres doucement	siflement	siflement	la parole, quelquen qui parle, plus fort que lautre categorie	la parole, quelquen qui parle, plus fort que lautre categorie
NEW-11	pas (steps) du personnse	bruit dexterior, matapiquer, pluie	pas (steps) du personnse	bruit dexterior, matapiquer, pluie	piano, music	klaxon	klaxon	musique, violin, rhétique	pas (steps) du personnse	ricalement	piano, music	voiture relatif, bruit de moteur	voiture relatif, bruit de moteur	bruit dexterior, matapiquer, pluie	chanter	chanter
NEW-12	tourne un cle dun porte	trrrrrrr	pas	fourchette tombe	musique	sais problem	sonnete de porte	musique	musique	humaines	grave	matopiquer	voiture demarre	perceuse	humaines	humaines
NEW-13	toc toc, tappe	capte bien, ni grave, ni haute	toc toc, tappe	glisse ordi	musique	capte bien, ni grave, ni haute	leglise, son ddu cloche	parole	toc toc, tappe	parole	telephone distant	motor	capte bien, ni grave, ni haute	capte bien, ni grave, ni haute	leglise, son ddu cloche	parole
NEW-14	aigu - moyen	aigu	aigu	aigu	grav	aigu - moyen	grav	moyen - grave	grav	aigu	grav	je sais pas - (pas les characteristiques dentendre)	aigu	aigu	grav	moyen - grave
NEW-15	DEUX CLAQUEMENT	IL Y A DU BRUIT GENRE APPLAUDISSEMENT ET MITRAILETTE, LES RYTHMES ET LES SONS SONT PROCHES	UN BRUIT DE PAS	UN BRUIT DUN OBJET QUE LON POSE	UN BRUIT DUN PIANO	RAS	UN BRUIT DUN PIANO	UNE FEMME QUI PARLE	UN BRUIT DE TAMBOUR	UN HOMME QUI PARLE EN RIGOLANT	UN BRUIT DUN PIANO	IL Y A DU BRUIT GENRE APPLAUDISSEMENT ET MITRAILETTE, LES RYTHMES ET LES SONS SONT PROCHES	IL Y A DU BRUIT GENRE APPLAUDISSEMENT ET MITRAILETTE, LES RYTHMES ET LES SONS SONT PROCHES	IL Y A DU BRUIT GENRE APPLAUDISSEMENT ET MITRAILETTE, LES RYTHMES ET LES SONS SONT PROCHES	UNE FEMME QUI PARLE	UN HOMME QUI PARLE
NEW-16	nothing to say	** patient thought they sounded similar. ** the idea of continuous sounds was present	easily identified as "pas"	interpreted sounds: ** again only comments were relating to it sounding similar, in terms of temporal pattern	musical sounds	interpreted comments: ** subject described the temporal structure again	described as a hit on a cloche	musical sounds	interpreted comment: ** subject easily identified the rhythm and had a vague idea it was related to the voice	musical sounds	musical sounds	** patient thought they sounded similar. ** the idea of continuous sounds was present	** patient thought they sounded similar. ** the idea of continuous sounds was present	** patient thought they sounded similar. ** the idea of continuous sounds was present	interpreted sounds: ** again only comments were relating to it sounding similar, in terms of temporal pattern	interpreted sounds: ** again only comments were relating to it sounding similar, in terms of temporal pattern
NEW-17	[]	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Table 10.5: Appendix B-3: category descriptions of sixteen common everyday sounds by Newly implanted CIL (NEW)

11

APPENDIX: C

The following appendix contains data relevant to the testing of the effect of context (location) on the identification and categorisation of 20 environmental sounds that is described in chapter **??**. Sounds were chosen to represent one of four different contexts - Bathroom, Kitchen, Exterior and Office and are listed in table **??**. The typicality of each sound to its associated context was measured by asking 9 participants to rate the typicality on a scale of 1-10, where 10 was highly typical and 1 not very typical. Results of this test, referred to in chapter throughout the the manuscript as *testing Step-1*, are shown in table **??** with a mean value also calculated.

The current appendix also contains figures which display the descriptions given by participants when identifying each of the 20 sounds across three different test conditions. The three test conditions, which are fully detailed in chapter **??** are again briefly described below. These figures were created in order to show the possible changes in perception that could have occurred across the three different test conditions.

- **Step1-Typ-Typ** - typicality ratings of the stimuli as to whether they were typical of the designated context.
- **Step2-FST-NS** - Free Sorting Task (FST) of stimuli in Natural sounds (NS) condition to test categorisation strategies of environmental sounds.
- **Step3-FST-SIM** - again a FST however sounds were passed through a 4 channel vocoder in order to simulate listening with a CI
- **Step4-FC-SIM** - Forced categorisation of CI simulated sounds as used in Step3-FST-SIM

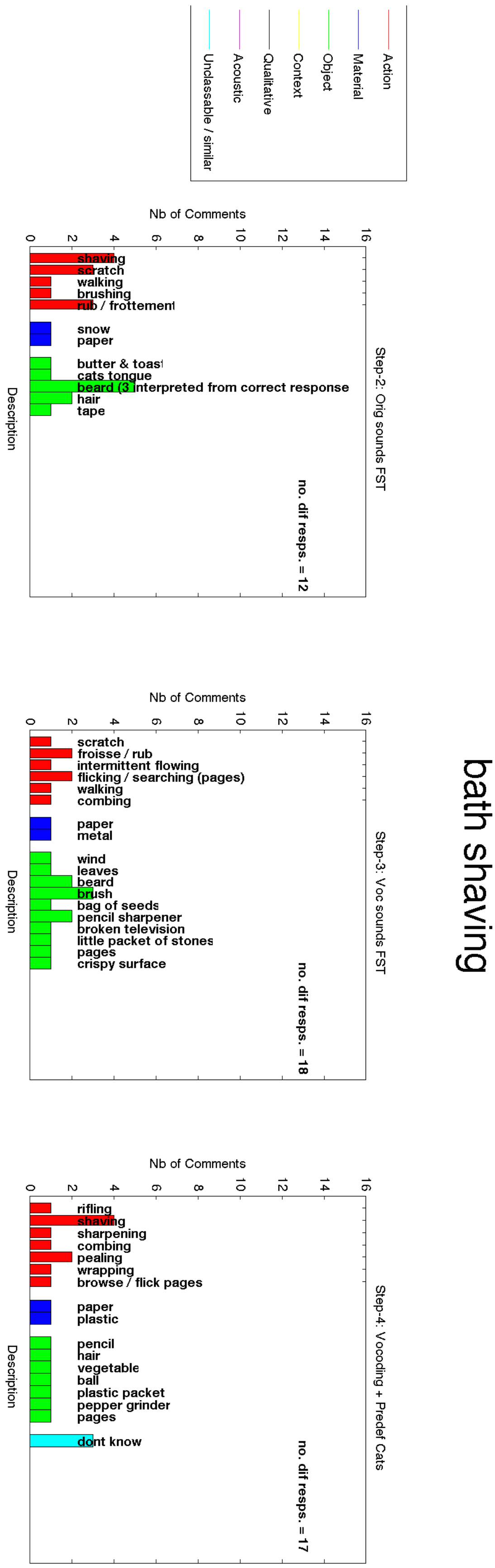
In creating the comment-figures participants descriptions were evaluated as to whether the perception described referenced different attributes of the either the sound producing event, or the sound signal itself. These attributes are listed in the legend on every figure and are also described below:

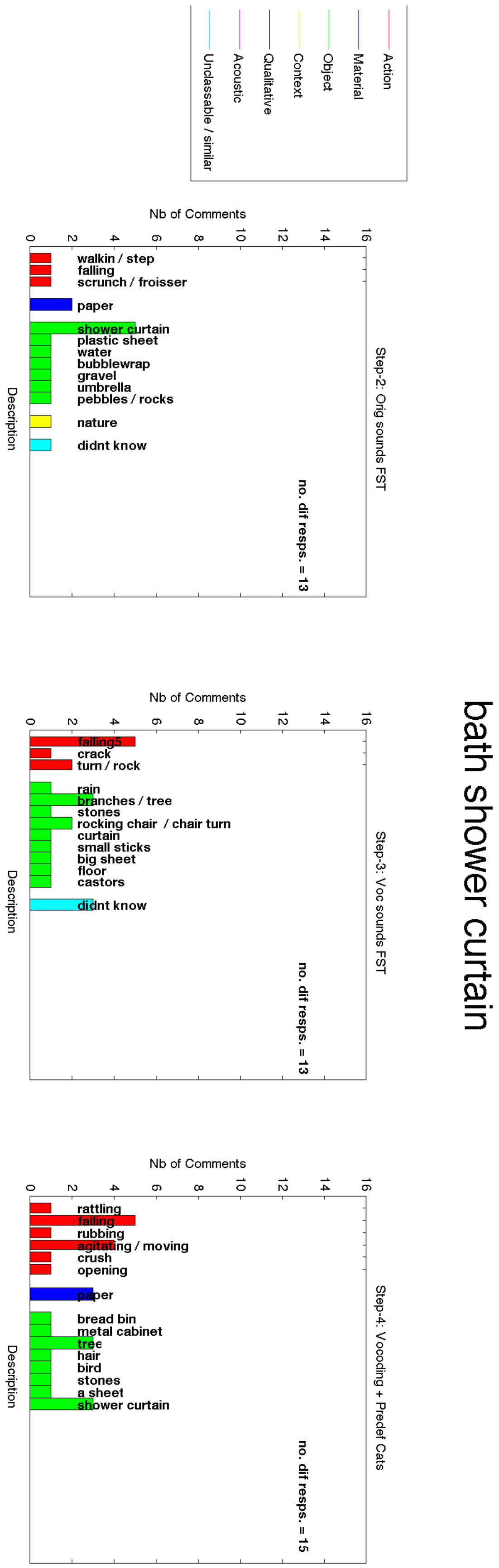
- **Action:** referencing the action involved in producing the sound, e.g. hitting, smashing, scraping, scratching.
- **Material:** referencing the material of an object involved in producing the sound e.g. water, glass, wood, metal.
- **Object:** specific referencing of a real physical object e.g. car, telephone, tree, human being.
- **Context:** mention of a context or location associated to the sound e.g. bathroom, park, train station.
- **Qualitative:** perceptions that pertain to the sound signal itself, but however using such vocabulary as "soft/hard" to describe the texture. Also referencing more emotional reactions to sounds for example happy/sad, busy/calm. Qualitative and Acoustic may be read as quite similar.
- **Acoustic:** again referencing the sound signal however in a more analytical manner, such as describing specific aspects of the frequency content - e.g. high/low pitched, describing the sound as a kind of "noise". Qualitative and Acoustic may be read as quite similar.
- **Unclassifiable / "Similar":** Descriptions that were not able to be interpreted as any of the above. This also included conditions where participants were only able to give responses such as "i don't know" or where participants were only able to state that sounds were "similar" without giving further details on this similarity.

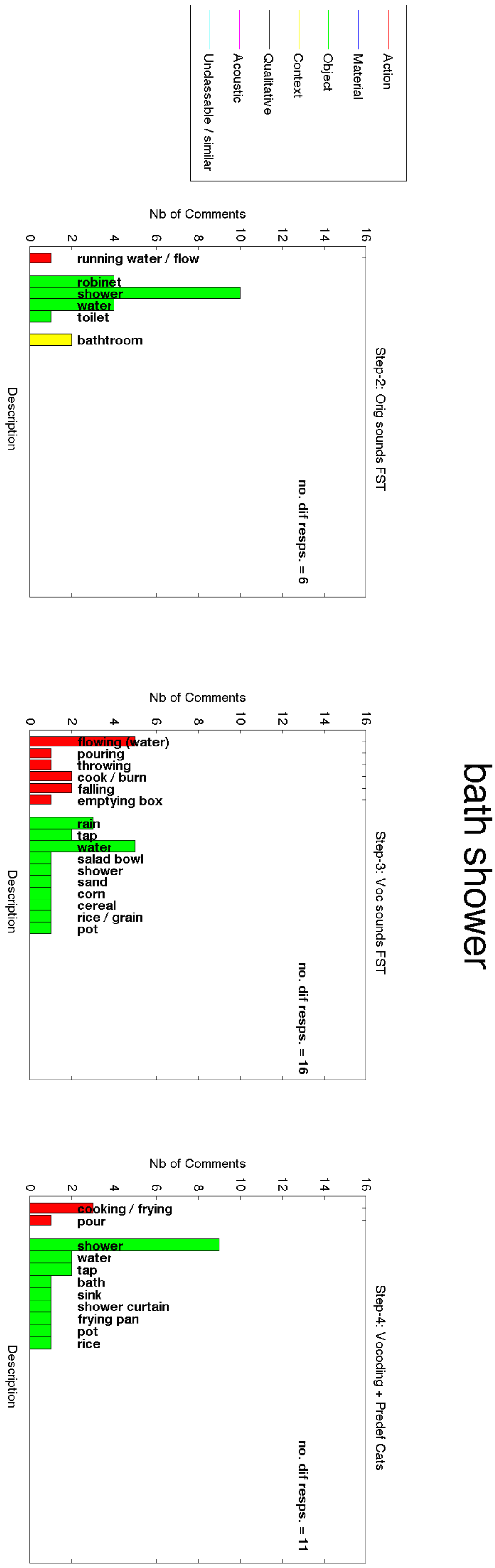
Participants responses were given in both French and English and

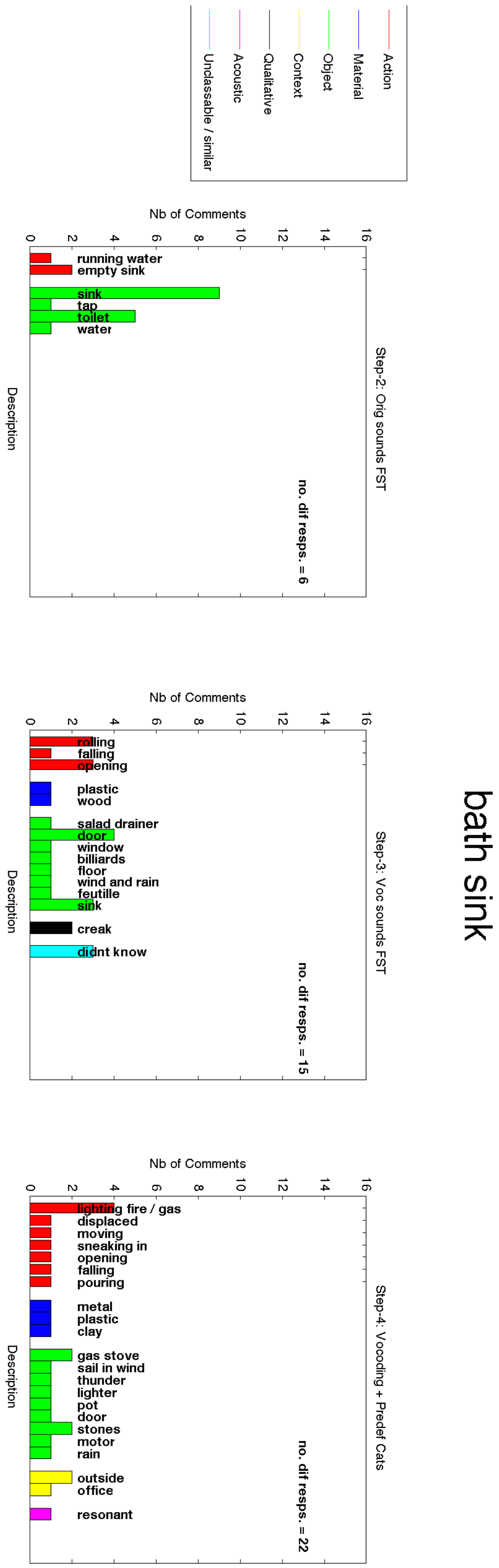
		Participant Number									
Context	Sound ID	1	2	3	4	5	6	7	8	9	Mean
Bathroom	Shaving	1	2	2	1	1	3	10	8	8	4
	Shower	10	7	10	10	10	10	10	7	9	9
	Shower curtain	5	1	0	7	1	10	10	5	5	5
	Sink	5	8	10	9	10	8	10	6	6	8
	Toothbrush	10	10	10	10	8	10	10	10	10	10
	Bathtub	7	9	6	8	10	10	10	7	9	8
	Water	5	9	7	7	9	10	10	3	10	8
Kitchen	Cereal	10	8	10	8	7	10	10	7	6	8
	Bread	10	9	10	4	4	7	10	9	8	8
	Egg beating	5	10	10	8	10	10	10	8	10	9
	Oven	1	1	1	5	5	5	4	3	8	4
	Plate	10	5	5	7	5	8	10	6	7	7
Exterior	Basketball	10	9	10	10	2	6	10	0	9	7
	Rain	10	10	10	9	5	10	10	10	10	9
	Tractor	10	8	10	6	8	10	10	10	10	9
	Shovel	7	7	1	3	6	7	10	5	7	6
	Lightning	10	10	10	7	5	9	10	10	7	9
Office	Paper folding	2	6	9	4		7	10	5	6	6
	Photocopier	10	10	10	9	10	10	10	10	8	10
	Scissors	10	7	9		5	6	10	6	7	8
	Tape	1	9	4	5	7	6	10	10	7	7

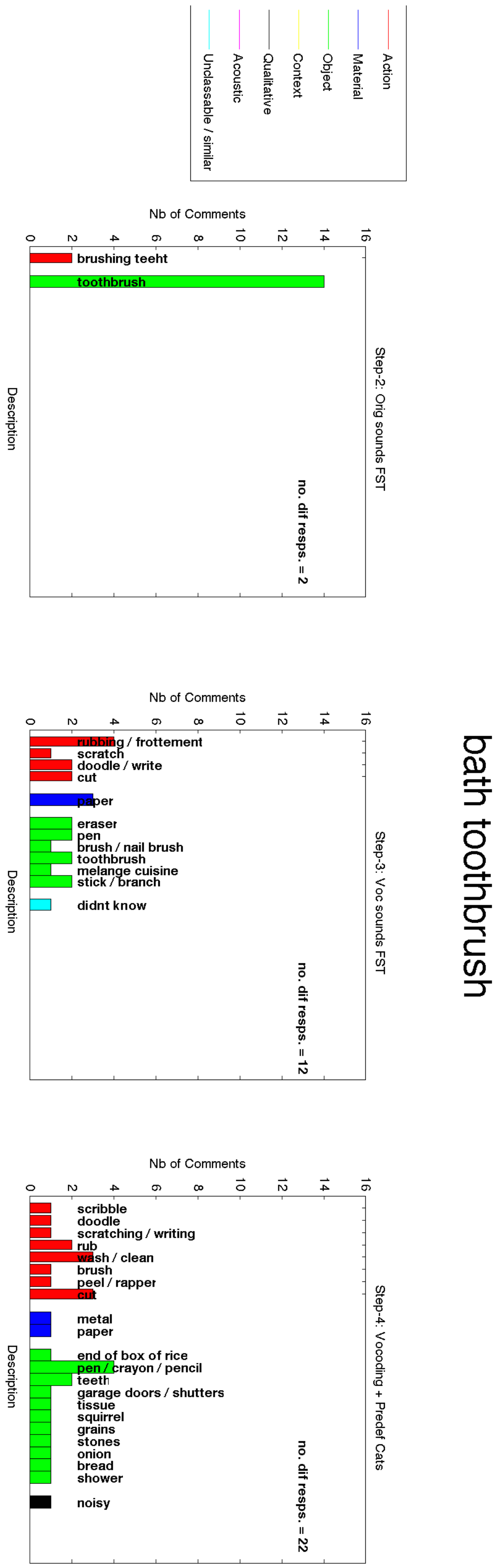
Table 11.1: Typicality ratings of 20 environmental sounds used for testing the effect of context in chapter 5.

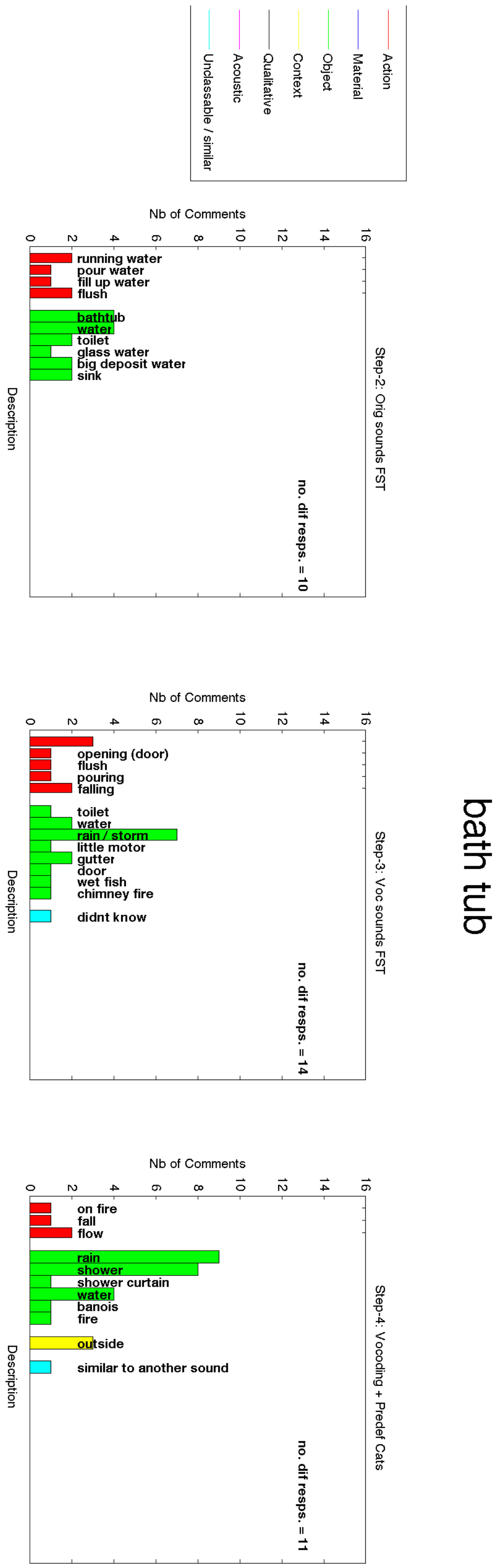

bath shaving
Action
Material
Object
Context
Qualitative
Acoustic
Unclassable / similar
Step-2: Orig sounds FST
Nb of Comments
Description
shaving
scratch
walking
brushing
rub / frottement
snow
paper
butter & toast
cats tongue
beard (3 interpreted from correct response
hair
tape
no. dif resps. = 12
Step-3: Voc sounds FST
scratch
froisse / rub
intermittent flowing
flicking / searching (pages)
walking
combing
paper
metal
wind
leaves
beard
brush
bag of seeds
pencil sharpener
broken television
little packet of stones
pages
crispy surface
no. dif resps. = 18
Step-4: Vocoding + Predef Cats
rifling
shaving
sharpening
combing
pealing
wrapping
browse / flick pages
paper
plastic
pencil
hair
vegetable
ball
plastic packet
pepper grinder
pages
dont know
no. dif resps. = 17

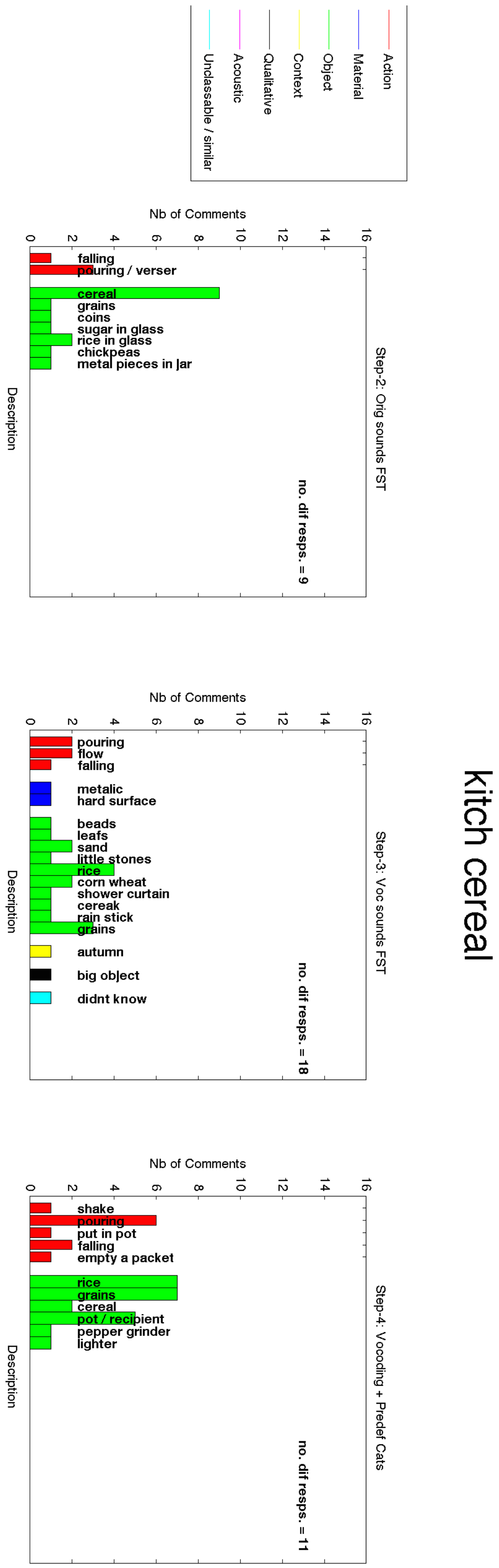

bath shower curtain
Action
Material
Object
Context
Qualitative
Acoustic
Unclassable / similar
Step-2: Orig sounds FST
Nb of Comments
Description
walkin / step
falling
scrunch / froisser
paper
shower curtain
plastic sheet
water
bubblewrap
gravel
umbrella
pebbles / rocks
nature
didnt know
no. dif resps. = 13
Step-3: Voc sounds FST
falling5
crack
turn / rock
rain
branches / tree
stones
rocking chair / chair turn
curtain
small sticks
big sheet
floor
castors
didnt know
no. dif resps. = 13
Step-4: Vocoding + Predef Cats
rattling
falling
rubbing
agitating / moving
crush
opening
paper
bread bin
metal cabinet
tree
hair
bird
stones
a sheet
shower curtain
no. dif resps. = 15

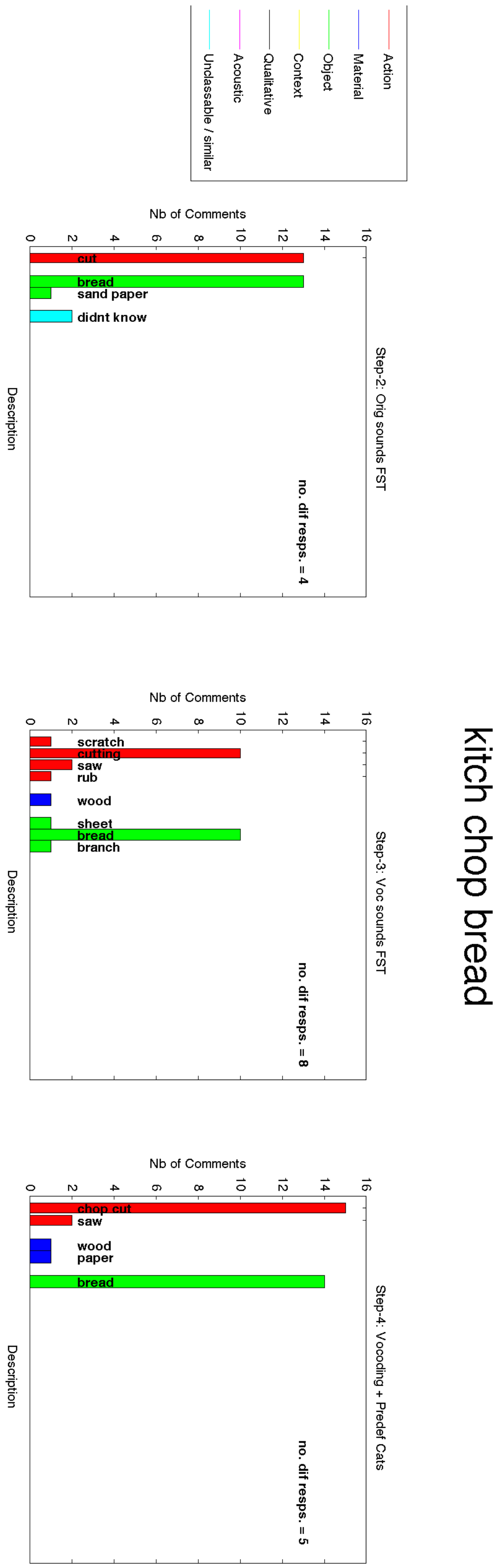

Action
Material
Object
Context
Qualitative
Acoustic
Unclassable / similar
bath shower
Step-2: Orig sounds FST
Nb of Comments
Description
running water / flow
robinet
shower
water
toilet
bathroom
no. dif resps. = 6
Step-3: Voc sounds FST
Nb of Comments
Description
flowing (water)
pouring
throwing
cook / burn
falling
emptying box
rain
tap
water
salad bowl
shower
sand
corn
cereal
rice / grain
pot
no. dif resps. = 16
Step-4: Vocoding + Predef Cats
Nb of Comments
Description
cooking / frying
pour
shower
water
tap
bath
sink
shower curtain
frying pan
pot
rice
no. dif resps. = 11

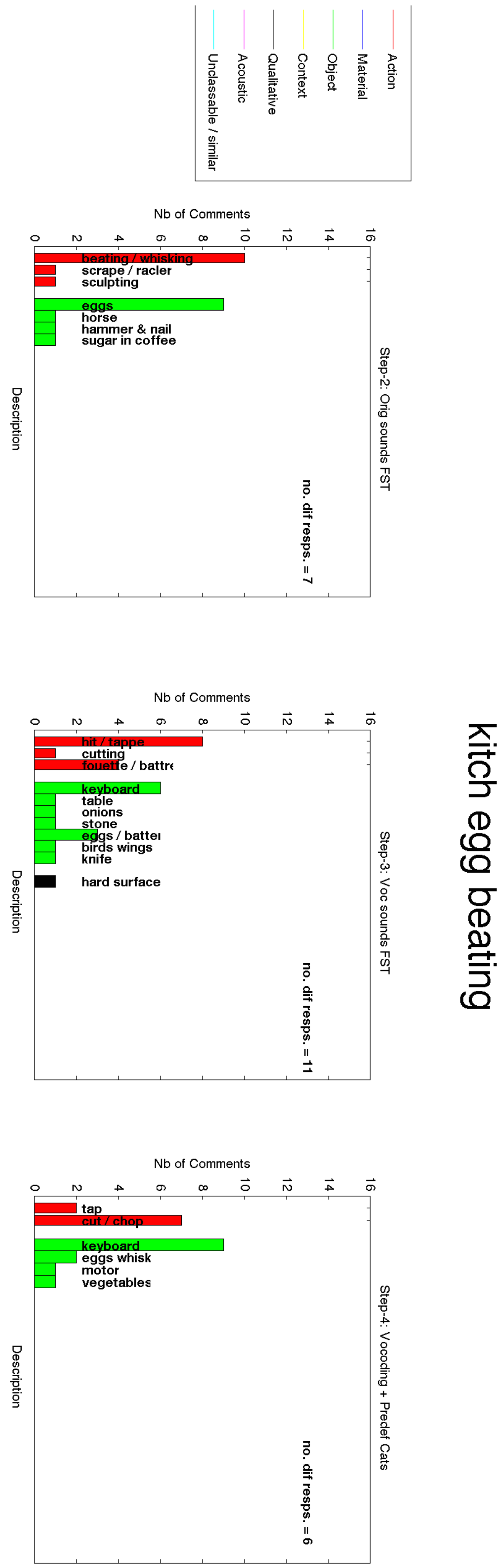

bath sink
Action
Material
Object
Context
Qualitative
Acoustic
Unclassable / similar
Step-2: Orig sounds FST
Nb of Comments
Description
no. dif resps. = 6
running water
empty sink
sink
tap
toilet
water
Step-3: Voc sounds FST
no. dif resps. = 15
rolling
falling
opening
plastic
wood
salad drainer
door
window
billiards
floor
wind and rain
feutille
sink
creak
didnt know
Step-4: Vocoding + Predef Cats
no. dif resps. = 22
lighting fire / gas
displaced
moving
sneaking in
opening
falling
pouring
metal
plastic
clay
gas stove
sail in wind
thunder
lighter
pot
door
stones
motor
rain
outside
office
resonant

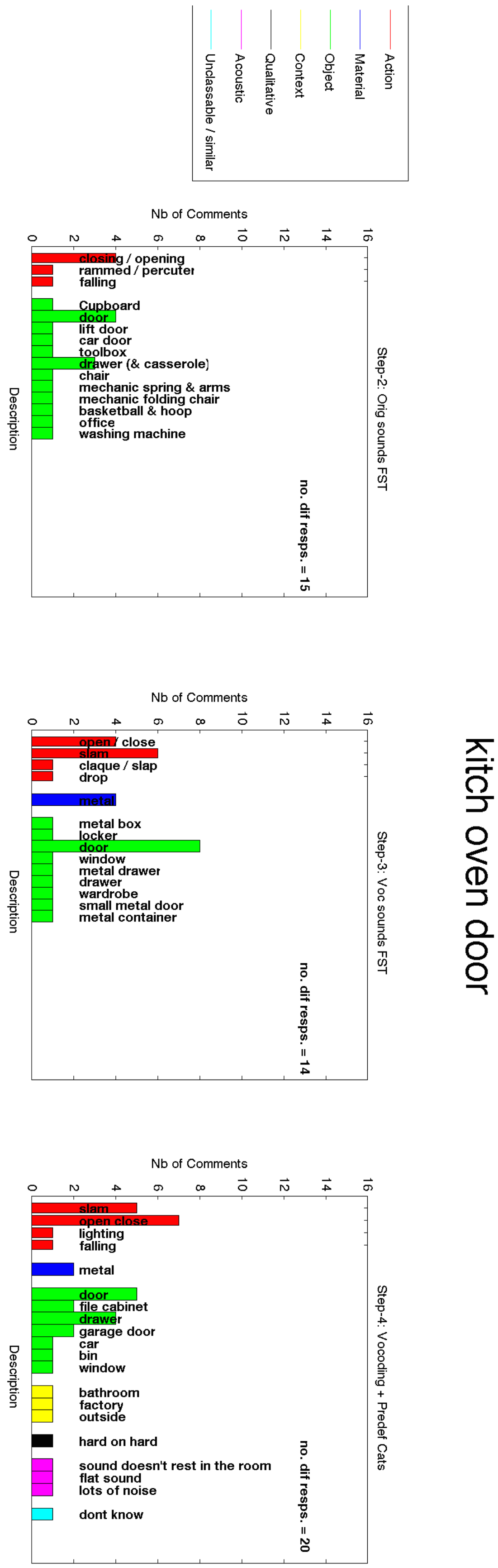

bath toothbrush
Action
Material
Object
Context
Qualitative
Acoustic
Unclassable / similar
Step-2: Orig sounds FST
Nb of Comments
Description
brushing teeht
toothbrush
no. dif resps. = 2
Step-3: Voc sounds FST
rubbing / frottement
scratch
doodle / write
cut
paper
eraser
pen
brush / nail brush
toothbrush
melange cuisine
stick / branch
didnt know
no. dif resps. = 12
Step-4: Vocoding + Predef Cats
scribble
doodle
scratching / writing
rub
wash / clean
brush
peel / rapper
cut
metal
paper
end of box of rice
pen / crayon / pencil
teeth
garage doors / shutters
tissue
squirrel
grains
stones
onion
bread
shower
noisy
no. dif resps. = 22

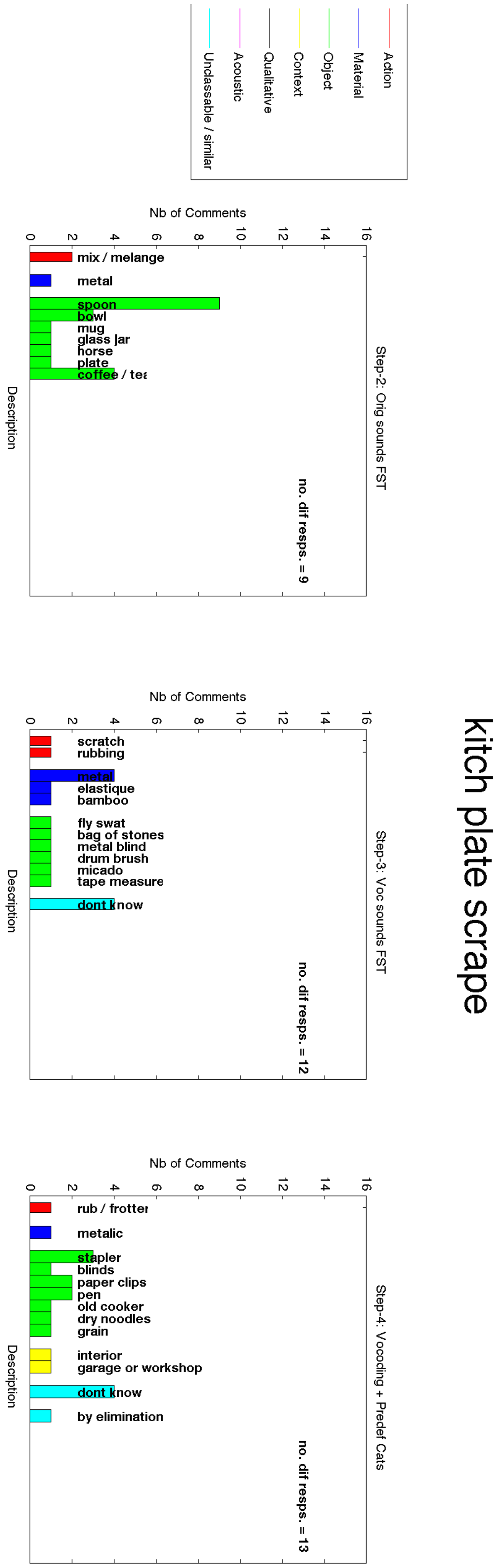

Action
Material
Object
Context
Qualitative
Acoustic
Unclassable / similar
bath tub
Step-2: Orig sounds FST
Nb of Comments
Description
no. dif resps. = 10
running water
pour water
fill up water
flush
bathtub
water
toilet
glass water
big deposit water
sink
Step-3: Voc sounds FST
Nb of Comments
Description
no. dif resps. = 14
opening (door)
flush
pouring
falling
toilet
water
rain / storm
little motor
gutter
door
wet fish
chimney fire
didnt know
Step-4: Vocoding + Predef Cats
Nb of Comments
Description
no. dif resps. = 11
on fire
fall
flow
rain
shower
shower curtain
water
banois
fire
outside
similar to another sound
0
2
4
6
8
10
12
14
16

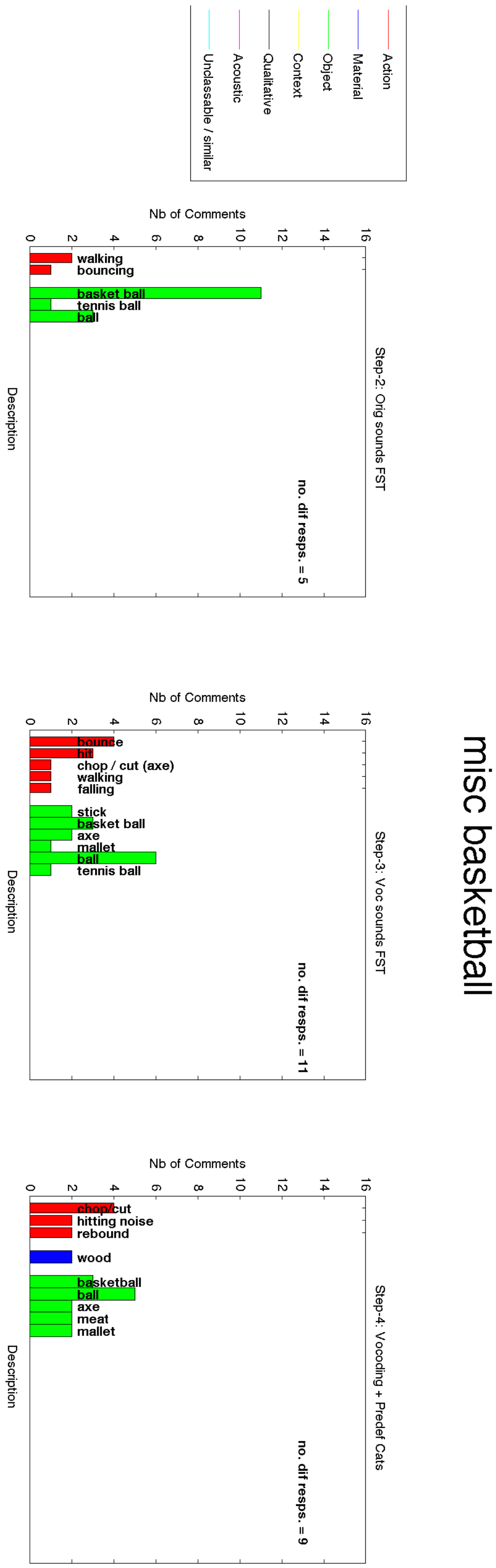

kitch cereal
Action
Material
Object
Context
Qualitative
Acoustic
Unclassable / similar
Step-2: Orig sounds FST
Nb of Comments
Description
falling
pouring / verser
cereal
grains
coins
sugar in glass
rice in glass
chickpeas
metal pieces in jar
no. dif resps. = 9
Step-3: Voc sounds FST
pouring
flow
falling
metalic
hard surface
beads
leafs
sand
little stones
rice
corn wheat
shower curtain
cereak
rain stick
grains
autumn
big object
didnt know
no. dif resps. = 18
Step-4: Vocoding + Predef Cats
shake
pouring
put in pot
falling
empty a packet
rice
grains
cereal
pot / recipient
pepper grinder
lighter
no. dif resps. = 11

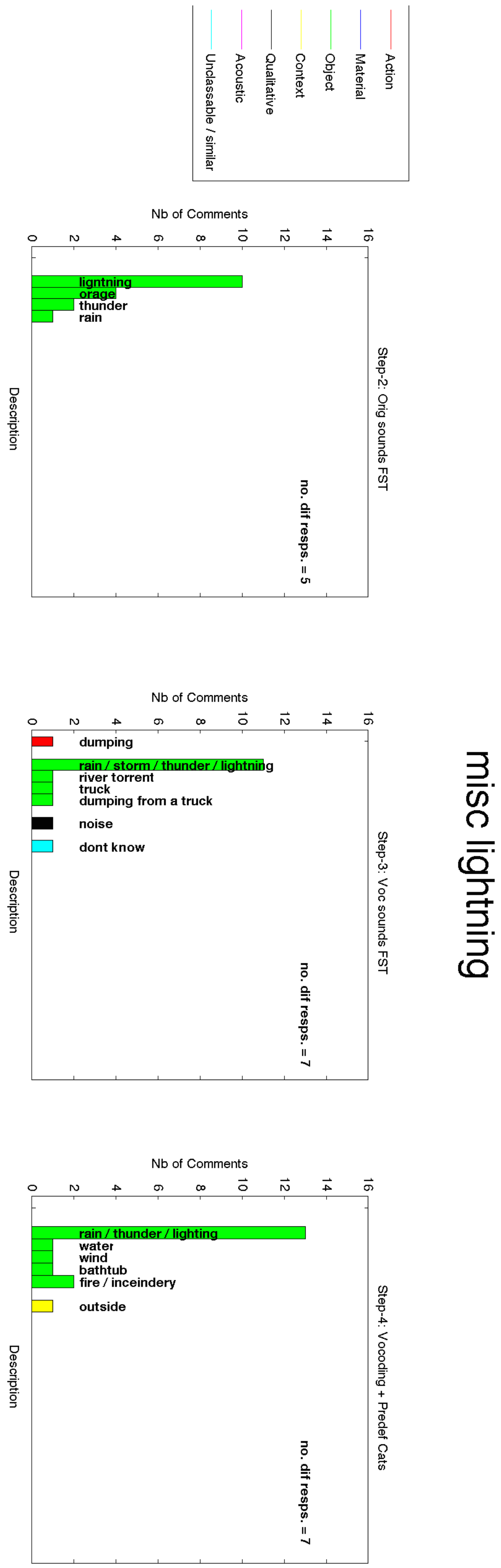

kitch chop bread
Action
Material
Object
Context
Qualitative
Acoustic
Unclassable / similar
Step-2: Orig sounds FST
no. dif resps. = 4
cut
bread
sand paper
didnt know
Nb of Comments
Description
Step-3: Voc sounds FST
no. dif resps. = 8
scratch
cutting
saw
rub
wood
sheet
bread
branch
Nb of Comments
Description
Step-4: Vocoding + Predef Cats
no. dif resps. = 5
chop cut
saw
wood
paper
bread
Nb of Comments
Description

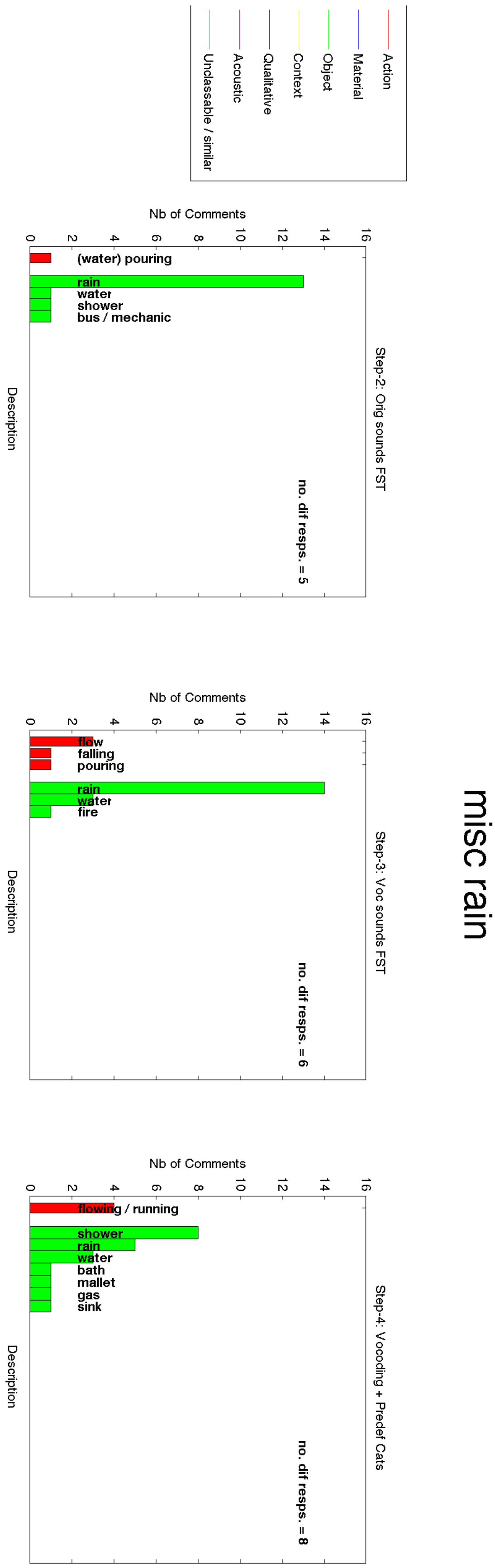

kitch egg beating
Action
Material
Object
Context
Qualitative
Acoustic
Unclassable / similar
Step-2: Orig sounds FST
Nb of Comments
0
2
4
6
8
10
12
14
16
beating / whisking
scrape / racler
sculpting
eggs
horse
hammer & nail
sugar in coffee
no. dif resps. = 7
Description
Step-3: Voc sounds FST
Nb of Comments
hit / tappe
cutting
fouette / battre
keyboard
table
onions
stone
eggs / batter
birds wings
knife
hard surface
no. dif resps. = 11
Description
Step-4: Vocoding + Predef Cats
Nb of Comments
tap
cut / chop
keyboard
eggs whisk
motor
vegetables
no. dif resps. = 6
Description

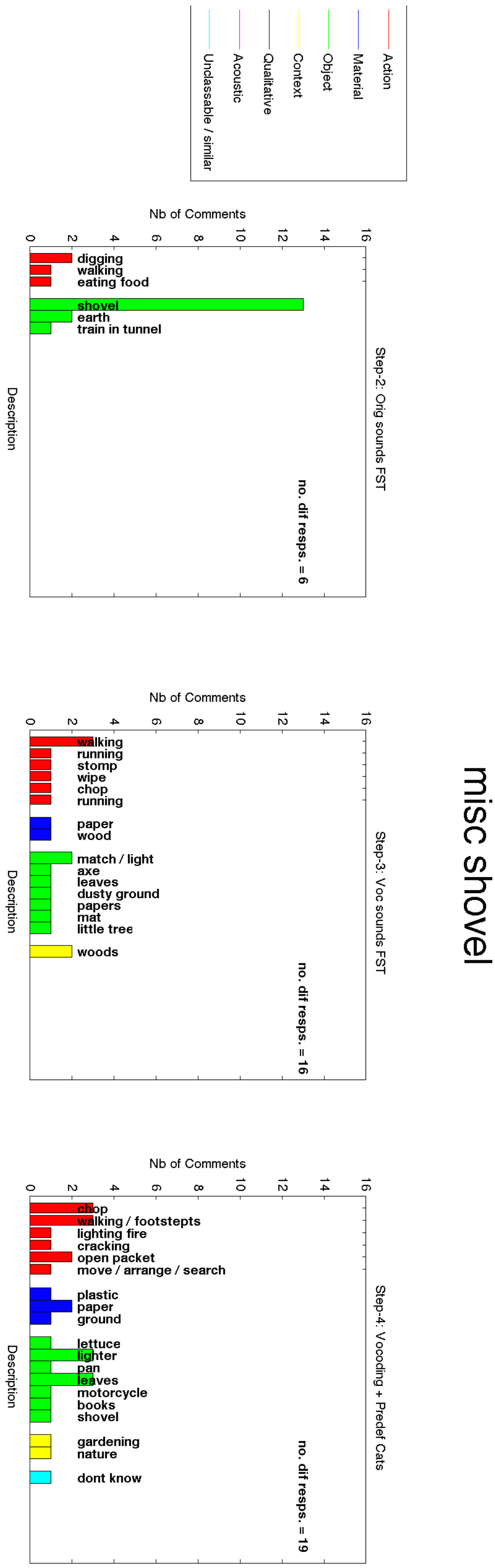

kitch oven door
Action
Material
Object
Context
Qualitative
Acoustic
Unclassable / similar
Step-2: Orig sounds FST
Nb of Comments
Description
closing / opening
rammed / percuter
falling
Cupboard
door
lift door
car door
toolbox
drawer (& casserole)
chair
mechanic spring & arms
mechanic folding chair
basketball & hoop
office
washing machine
no. dif resps. = 15
Step-3: Voc sounds FST
open / close
slam
claque / slap
drop
metal
metal box
locker
door
window
metal drawer
drawer
wardrobe
small metal door
metal container
no. dif resps. = 14
Step-4: Vocoding + Predef Cats
slam
open close
lighting
falling
metal
door
file cabinet
drawer
garage door
car
bin
window
bathroom
factory
outside
hard on hard
sound doesn't rest in the room
flat sound
lots of noise
dont know
no. dif resps. = 20

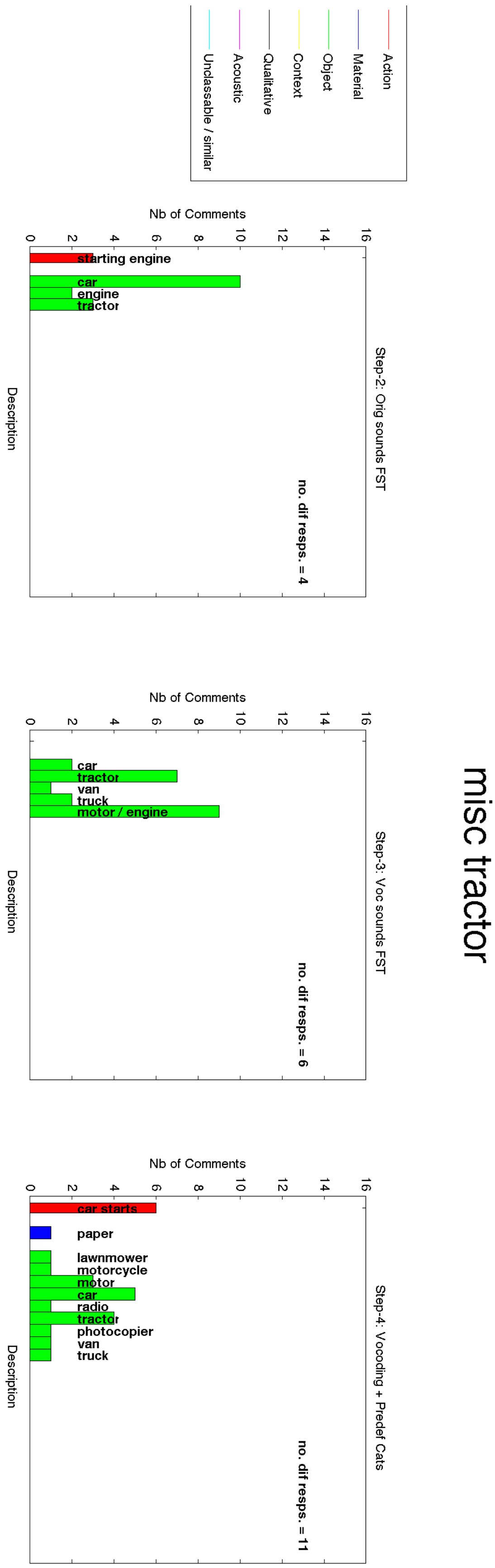

kitch plate scrape
Action
Material
Object
Context
Qualitative
Acoustic
Unclassable / similar
Step-2: Orig sounds FST
Nb of Comments
Description
no. dif resps. = 9
mix / melange
metal
spoon
bowl
mug
glass jar
horse
plate
coffee / tea
Step-3: Voc sounds FST
Nb of Comments
Description
no. dif resps. = 12
scratch
rubbing
metal
elastique
bamboo
fly swat
bag of stones
metal blind
drum brush
micado
tape measure
dont know
Step-4: Vocoding + Predef Cats
Nb of Comments
Description
no. dif resps. = 13
rub / frotter
metalic
stapler
blinds
paper clips
pen
old cooker
dry noodles
grain
interior
garage or workshop
dont know
by elimination

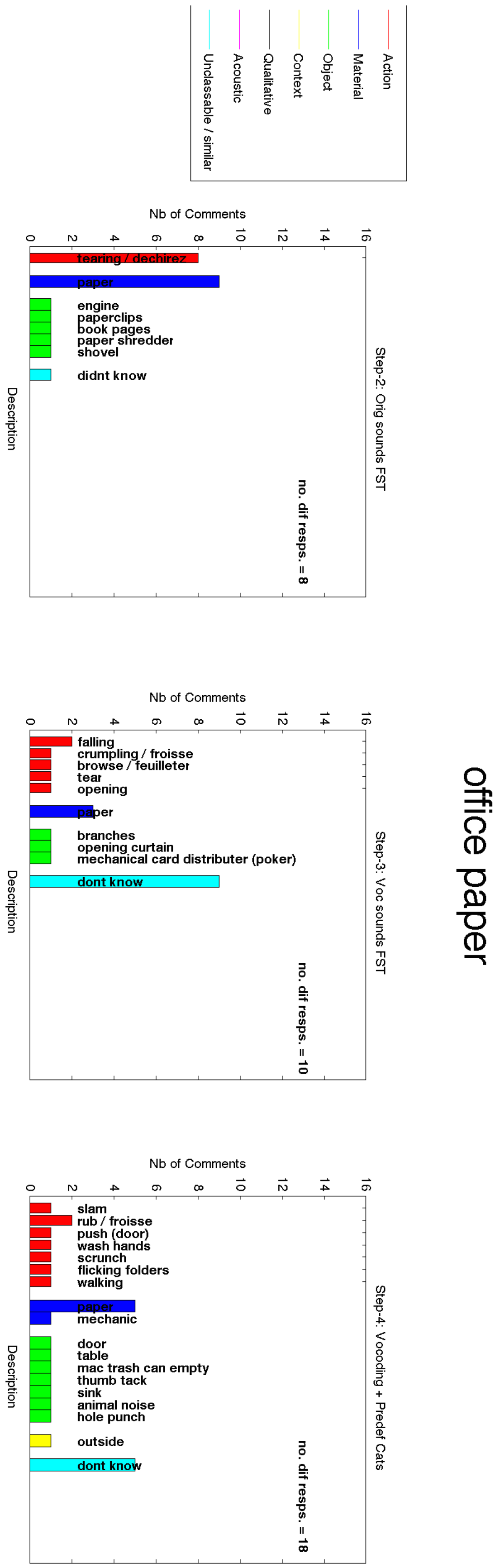

misc basketball
Action
Material
Object
Context
Qualitative
Acoustic
Unclassable / similar
Step-2: Orig sounds FST
Nb of Comments
Description
walking
bouncing
basket ball
tennis ball
ball
no. dif resps. = 5
Step-3: Voc sounds FST
bounce
hit
chop / cut (axe)
walking
falling
stick
basket ball
axe
mallet
ball
tennis ball
no. dif resps. = 11
Step-4: Vocoding + Predef Cats
chop/cut
hitting noise
rebound
wood
basketball
ball
axe
meat
mallet
no. dif resps. = 9


misc lightning
Action
Material
Object
Context
Qualitative
Acoustic
Unclassable / similar
Step-2: Orig sounds FST
Nb of Comments
Description
ligntning
orage
thunder
rain
no. dif resps. = 5
Step-3: Voc sounds FST
dumping
rain / storm / thunder / lightning
river torrent
truck
dumping from a truck
noise
dont know
no. dif resps. = 7
Step-4: Vocoding + Predef Cats
rain / thunder / lighting
water
wind
bathtub
fire / inceindery
outside
no. dif resps. = 7

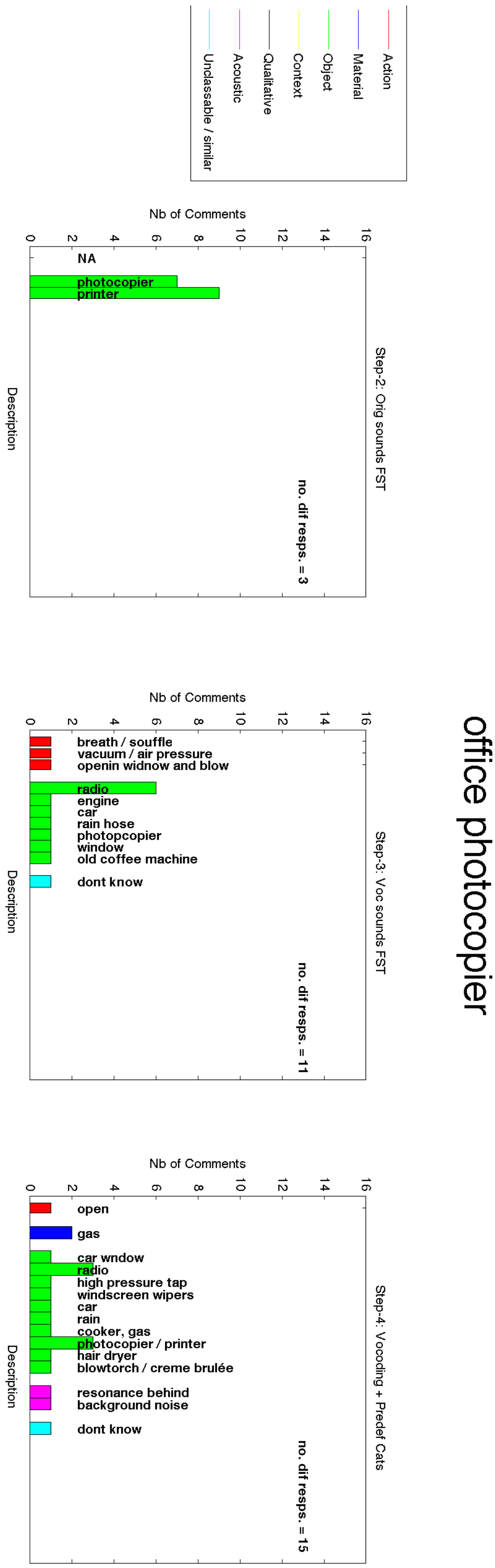

misc rain
Action
Material
Object
Context
Qualitative
Acoustic
Unclassable / similar
Step-2: Orig sounds FST
Nb of Comments
Description
(water) pouring
rain
water
shower
bus / mechanic
no. dif resps. = 5
Step-3: Voc sounds FST
flow
falling
pouring
rain
water
fire
no. dif resps. = 6
Step-4: Vocoding + Predef Cats
flowing / running
shower
rain
water
bath
mallet
gas
sink
no. dif resps. = 8
0
2
4
6
8
10
12
14
16

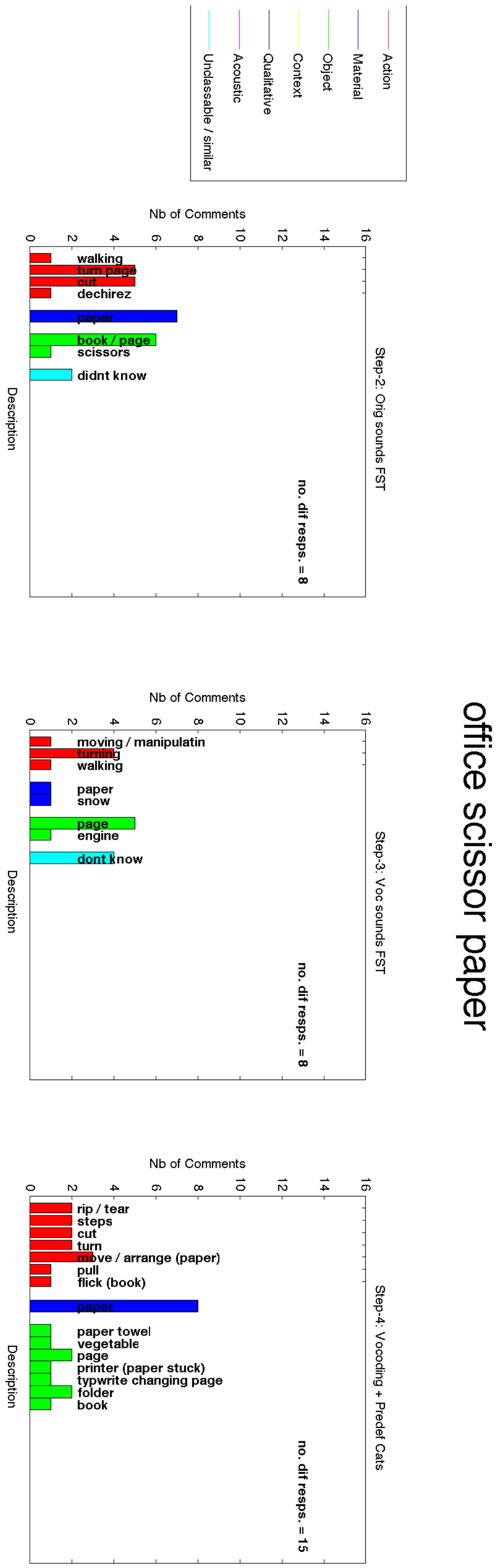

misc shovel
Action
Material
Object
Context
Qualitative
Acoustic
Unclassable / similar
Step-2: Orig sounds FST
Nb of Comments
Description
digging
walking
eating food
shovel
earth
train in tunnel
no. dif resps. = 6
Step-3: Voc sounds FST
walking
running
stomp
wipe
chop
running
paper
wood
match / light
axe
leaves
dusty ground
papers
mat
little tree
woods
no. dif resps. = 16
Step-4: Vocoding + Predef Cats
chop
walking / footstepts
lighting fire
cracking
open packet
move / arrange / search
plastic
paper
ground
lettuce
lighter
pan
leaves
motorcycle
books
shovel
gardening
nature
dont know
no. dif resps. = 19

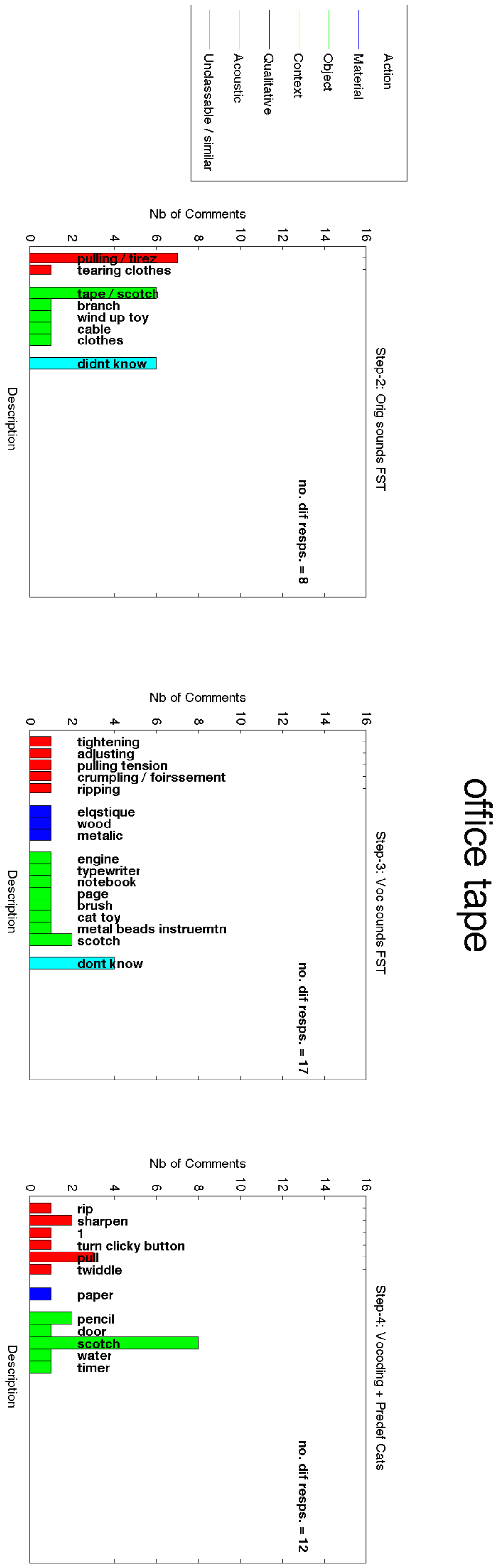

misc tractor
Action
Material
Object
Context
Qualitative
Acoustic
Unclassable / similar
Step-2: Orig sounds FST
no. dif resps. = 4
starting engine
car
engine
tractor
Nb of Comments
Description
Step-3: Voc sounds FST
no. dif resps. = 6
car
tractor
van
truck
motor / engine
Nb of Comments
Description
Step-4: Vocoding + Predef Cats
no. dif resps. = 11
car starts
paper
lawnmower
motorcycle
motor
car
radio
tractor
photocopier
van
truck
Nb of Comments
Description



office paper
Action
Material
Object
Context
Qualitative
Acoustic
Unclassable / similar
Step-2: Orig sounds FST
Nb of Comments
Description
tearing / dechirez
paper
engine
paperclips
book pages
paper shredder
shovel
didnt know
no. dif resps. = 8
Step-3: Voc sounds FST
Nb of Comments
Description
falling
crumpling / froisse
browse / feuilleter
tear
opening
paper
branches
opening curtain
mechanical card distributer (poker)
dont know
no. dif resps. = 10
Step-4: Vocoding + Predef Cats
Nb of Comments
Description
slam
rub / froisse
push (door)
wash hands
scrunch
flicking folders
walking
paper
mechanic
door
table
mac trash can empty
thumb tack
sink
animal noise
hole punch
outside
dont know
no. dif resps. = 18


office photocopier
Action
Material
Object
Context
Qualitative
Acoustic
Unclassable / similar
Step-2: Orig sounds FST
Nb of Comments
0 2 4 6 8 10 12 14 16
Description
NA
photocopier
printer
no. dif resps. = 3
Step-3: Voc sounds FST
Nb of Comments
0 2 4 6 8 10 12 14 16
Description
breath / souffle
vacuum / air pressure
openin widnow and blow
radio
engine
car
rain hose
photopcopier
window
old coffee machine
dont know
no. dif resps. = 11
Step-4: Vocoding + Predef Cats
Nb of Comments
0 2 4 6 8 10 12 14 16
Description
open
gas
car wndow
radio
high pressure tap
windscreen wipers
car
rain
cooker, gas
photocopier / printer
hair dryer
blowtorch / creme brulée
resonance behind
background noise
dont know
no. dif resps. = 15


office scissor paper
Action
Material
Object
Context
Qualitative
Acoustic
Unclassable / similar
Step-2: Orig sounds FST
Nb of Comments
Description
walking
turn page
cut
dechirez
paper
book / page
scissors
didnt know
no. dif resps. = 8
Step-3: Voc sounds FST
moving / manipulatin
turning
walking
paper
snow
page
engine
dont know
no. dif resps. = 8
Step-4: Vocoding + Predef Cats
rip / tear
steps
cut
turn
move / arrange (paper)
pull
flick (book)
paper
paper towel
vegetable
page
printer (paper stuck)
typwrite changing page
folder
book
no. dif resps. = 15


office tape
Action
Material
Object
Context
Qualitative
Acoustic
Unclassable / similar
Step-2: Orig sounds FST
no. dif resps. = 8
pulling / tirez
tearing clothes
tape / scotch
branch
wind up toy
cable
clothes
didnt know
Step-3: Voc sounds FST
no. dif resps. = 17
tightening
adjusting
pulling tension
crumpling / foirssement
ripping
elqstique
wood
metalic
engine
typewriter
notebook
page
brush
cat toy
metal beads instruemtn
scotch
dont know
Step-4: Vocoding + Predef Cats
no. dif resps. = 12
rip
sharpen
1
turn clicky button
pull
twiddle
paper
pencil
door
scotch
water
timer
Nb of Comments
0
2
4
6
8
10
12
14
16
Description

BIBLIOGRAPHY

[1] H Abdi and D Valentin. Encyclopedia of Measurement and Statistics, 2007.

[2] M Ahissar, M Nahum, I Nelken, and S Hochstein. Reverse hierarchies and sensory learning. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 364(1515):285–299, February 2009.

[3] Merav Ahissar and Shaul Hochstein. The reverse hierarchy theory of visual perceptual learning. *Trends in cognitive sciences*, 8(10):457–464, October 2004.

[4] Merle-Marie Ahrens, Bashar Awwad Shiekh Hasan, Bruno L Giordano, and Pascal Belin. Gender differences in the temporal voice areas. *Frontiers in Neuroscience*, 8:711–9, July 2014.

[5] Janna Maree Arnephy. Environmental Sound Perception for Cochlear Implant Users. 2008.

[6] J a Ballas. Common factors in the identification of an assortment of brief everyday sounds. *Journal of experimental psychology. Human perception and performance*, 19(2):250–267, 1993.

[7] James A Ballas and Jr James H Howard. Interpreting the Language of Environmental Sounds. *Environment and behavior*, 19(1):91–114, January 1987.

[8] James A Ballas and Timothy Mullins. Effects of Context on the Identification of Everyday Sounds. *Human performance*, 4(3):199–219, September 1991.

[9] Pascal Barone and Olivier Deguine. Multisensory Processing in Cochlear Implant Listeners. In *Auditory Prostheses*, pages 365–381. Springer New York, New York, NY, July 2011.

[10] P Belin, R J Zatorre, and P Ahad. Human temporal-lobe response to vocal sounds. *Cognitive Brain Research*, 2002.

[11] P Belin, R J Zatorre, P Lafaille, P Ahad, and B Pike. Voice-selective areas in human auditory cortex. *Nature*, 2000.

[12] Pascal Belin, Shirley Fecteau, and Catherine Bédard. Thinking the voice: neural correlates of voice perception. *Trends in cognitive sciences*, 8(3):129–135, March 2004.

[13] Penny Bergman, Anders Sk o ld, Daniel V a stfj a ll, and Niklas Fransson. Perceptual and emotional categorization of sound. *The Journal of the Acoustical Society of America*, 126(6):3156–3167, 2009.

[14] Penny Bergman, Anders Sköld, Daniel Västfjäll, and Niklas Fransson. Perceptual and emotional categorization of sound. *The Journal of the Acoustical Society of America*, 126(6):3156, 2009.

[15] Terri L Bonebright. PERCEPTUAL STRUCTURE OF EVERYDAY SOUNDS : A MULTIDIMENSIONAL SCALING APPROACH Department of Psychology. pages 73–78, 2001.

[16] Lewis Bott, Aaron B Hoffman, and Gregory L Murphy. Blocking in category learning. *Journal of Experimental Psychology: General*, 136(4):685–699, 2007.

[17] B Boudia, O Koenig, N Bedoin, and L Collet. Phonological representations in postlingual deaf subjects using a multichannel cochlear implant. *International Journal of Pediatric Otorhinolaryngology*, 47(2):157–164, 1999.

[18] J Bradley, P Bird, P Monteath, and J E Wells. Improved speech discrimination after cochlear implantation in the Southern Cochlear Implant Adult Programme - New Zealand Medical Journal. *NZ Med J*, 2010.

[19] Tobias Brosch, Gilles Pourtois, and David Sander. The perception and categorisation of emotional stimuli: A review. *Cognition & Emotion*, 24(3):377–400, April 2010.

[20] Laísa Flávia Soares Fernandes Peixoto Buarque, Joseli Soares Brazorotto, Hannalice Gottschalck Cavalcanti, Luiz Rodolpho Penna Lima Júnior, Danielle do Vale Silva Penna Lima, and Maria Ângela Fernandes Ferreira. Auditory performance, during a period of time in cochlear implant users with post lingual hearing loss. *Audiology - Communication Research*, 18(2):120–125, June 2013.

[21] Kristi A Buckley and Emily A Tobey. Cross-Modal Plasticity and Speech Perception in Pre- and Postlingually Deaf Cochlear Implant Users. *Ear and hearing*, page 1, September 2010.

[22] R A Burkholder. *Perceptual learning of speech processed through an acoustic simulation of a cochlear implant*. Unpublished Doctoral Dissertation, 2005.

[23] Marine Cadoret, S e bastien L e, and J e r o me Pag e s. A Factorial Approach for Sorting Task data (FAST). *Food Quality and Preference*, 20(6):410–417, 2009.

[24] T Carlson, D A Tovar, A Alink, and N Kriegeskorte. Representational dynamics of object vision: The first 1000 ms. *Journal of Vision*, 13(10):1–1, August 2013.

[25] D B Carr, C J Young, R C Aster, and X Zhang. Cluster analysis for CTBT seismic event monitoring. 1999.

[26] Lars Chittka and Axel Brockmann. Perception space–the final frontier. *PLoS biology*, 3(4):e137, April 2005.

[27] M Cleary and D B Pisoni. Talker discrimination by prelingually deaf children with cochlear implants: Preliminary results. *Annals of Otology Rhinology and Laryngology*, 111(5):113–118, May 2002.

[28] A Coez, M Zilbovicius, E Ferrary, D Bouccara, I Mosnier, E Ambert-Dahan, E Bizaguet, A Syrota, Y Samson, and O Sterkers. Cochlear Implant Benefits in Deafness Rehabilitation: PET Study of Temporal Voice Activations. *Journal of Nuclear Medicine*, 49(1):60–67, December 2007.

[29] V Colletti, M Carner, V Miorelli, M Guida, L Colletti, and F G Fiorino. Cochlear implantation at under 12 months: Report on 10 patients. *The Laryngoscope*, 115(3):445–449, March 2005.

[30] W B Cooper, E Tobey, and P C Loizou. Music perception by cochlear implant and normal hearing listeners as measured by the Montreal Battery for Evaluation of Amusia. *Ear and hearing*, 2008.

[31] Marion Cousineau, Laurent Demany, Bernard Meyer, and Daniel Pressnitzer. What breaks a melody: Perceiving F0 and intensity sequences with a cochlear implant. *Hearing Research*, 269(1-2):34–41, October 2010.

[32] W R Drennan and J T Rubinstein. Music perception in cochlear implant users and its relationship with psychophysical capabilities. *Journal of rehabilitation research . . .* , 2008.

[33] D Dubois. Categories as acts of meaning: The case of categories in olfaction and audition. *Cognitive Science Quarterly*, 2000.

[34] Qian-Jie Fu, Sherol Chinchilla, and John J Galvin. The Role of Spectral and Temporal Cues in Voice Gender Discrimination by Normal-Hearing Listeners and Cochlear Implant Users. *JARO*, 5(3):253–260, May 2004.

[35] Qian-Jie Fu, Robert V Shannon, and Xiaosong Wang. Effects of noise and spectral resolution on vowel and consonant recognition: Acoustic and electric hearing. *Journal of the Acoustical Society of America*, 104(6):3586–3596, December 1998.

[36] S Fujita and J Ito. Ability of nucleus cochlear implantees to recognize music. *Annals of Otology Rhinology and Laryngology*, 108(7):634–640, July 1999.

[37] W W Gaver. What in the World Do We Hear?: An Ecological Approach to Auditory Event Perception. *Ecological Psychology*, 5(1):1–29, 1993.

[38] Luc Geurts and Jan Wouters. Coding of the fundamental frequency in continuous interleaved sampling processors for cochlear implants. *Journal of the Acoustical Society of America*, 109(2):713–726, February 2001.

[39] K Gfeller and C R Lansing. Melodic, rhythmic, and timbral perception of adult cochlear implant users. *Journal of Speech and Hearing Research*, 34(4):916–920, July 1991.

[40] K Gfeller, C Olszewski, M Rychener, K Sena, J F Knutson, S Witt, and B Macpherson. Recognition of "real-world" musical excerpts by cochlear implant recipients and normal-hearing adults. *Ear and hearing*, 26(3):237–250, June 2005.

[41] K Gfeller, S Witt, and G Woodworth. Effects of frequency, instrumental family, and cochlear implant type on timbre recognition and appraisal. *ANNALS OF ...*, 2002.

[42] B L Giordano, S McAdams, R J Zatorre, N Kriegeskorte, and P Belin. Abstract Encoding of Auditory Objects in Cortical Activity Patterns. *Cerebral Cortex*, 23(9):2025–2037, July 2013.

[43] Bruno L Giordano, John McDonnell, and Stephen McAdams. Brain and Cognition. *Brain and Cognition*, 73(1):7–19, June 2010.

[44] A L Giraud and H J Lee. Predicting cochlear implant outcome from brain organisation in the deaf. *Restorative neurology and neuroscience*, 2007.

[45] Julio Gonzalez and Juan C Oliver. Gender and speaker identification as a function of the number of channels in spectrally reduced speech. *The Journal of the Acoustical Society of America*, 118(1):461–10, 2005.

[46] M Grassi. Do we hear size or sound? Balls dropped on plates. *Perception & psychophysics*, 67(2):274–284, February 2005.

[47] C Guastavino and P Belin. VOICE CATEGORIZATION USING FREE SORTING TASKS. *sea-acustica.es*.

[48] Catherine Guastavino. Categorization of environmental sounds. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 61(1):54–63, 2007.

[49] B Gygi. Factors in the identification of environmental sounds. 2001.

[50] B Gygi and V Shafiro. The effects of auditory context on the identification of environmental sounds. 2007.

[51] Brian Gygi, Gary R Kidd, and Charles S Watson. Spectral-temporal factors in the identification of environmental sounds. *Journal of the Acoustical Society of America*, 115(3):1252–1265, February 2004.

[52] Brian Gygi, Gary R Kidd, and Charles S Watson. Similarity and categorization of environmental sounds. *Perception & psychophysics*, 69(6):839–855, 2007.

[53] Brian Gygi and Valeriy Shafiro. The incongruency advantage for environmental sounds presented in natural auditory scenes. *Journal of experimental psychology. Human perception and performance*, 37(2):551–565, March 2011.

[54] Brian Gygi and Valeriy Shafiro. Auditory and cognitive effects of aging on perception of environmental sounds in natural auditory scenes. *Journal of Speech, Language, and Hearing Research*, 56(5):1373–1388, September 2013.

[55] James D Harnsberger, Rahul Shrivastav, W S Brown Jr, Howard Rothman, and Harry Hollien. Speaking Rate and Fundamental Frequency as Speech Cues to Perceived Age. *Journal of Voice*, 22(1):58–69, January 2008.

[56] Rachael Frush Holt and Mario A Svirsky. An exploratory look at pediatric cochlear implantation: Is earliest always best? *Ear and hearing*, 29(4):492–511, August 2008.

[57] Olivier Houix, Guillaume Lemaitre, Nicolas Misdariis, Patrick Susini, and Isabel Urdapilleta. A lexical analysis of environmental sound categories. *Journal of Experimental Psychology: Applied*, 18(1):52–80, 2012.

[58] David House. Perception and production of mood in speech by cochlear implant users. *ICSLP 1994*, 1994.

[59] Y Inverso and C J Limb. Cochlear implant-mediated perception of nonlinguistic sounds. *Ear and hearing*, 2010.

[60] Robert Kang, Grace Liu Nimmons, Ward Drennan, Jeff Longnion, Chad Ruffin, Kaibao Nie, Jong Ho Won, Tina Worman, Bevan Yueh, and Jay Rubinstein. Development and Validation of the University of Washington Clinical Assessment of Music Perception Test. *Ear and hearing*, 30(4):411–418, August 2009.

[61] Gary R Kidd, Charles S Watson, and Brian Gygi. Individual differences in auditory abilities. *The Journal of the Acoustical Society of America*, 122(1):418, 2007.

[62] Roberta L Klatzky, Dinesh K Pai, and Eric Krotkov. Perception of Material from Contact Sounds. *Presence*, 9(4):399–410, 2000.

[63] Thomas R Knösche, Sonja Lattner, Burkhard Maess, Michael Schauer, and Angela D Friederici. Early Parallel Processing of Auditory Word and Voice Information. *NeuroImage*, 17(3):1493–1503, November 2002.

[64] Ying-Yee Kong, Rachel Cruz, J Ackland Jones, and Fan-Gang Zeng. Music perception with temporal cues in acoustic and electric hearing. *Ear and hearing*, 25(2):173–185, April 2004.

[65] Damir Kovačić and Evan Balaban. Voice gender perception by cochlear implantees. *The Journal of the Acoustical Society of America*, 126(2):762–14, 2009.

[66] Jody Kreiman, Bruce R Gerratt, and Kristin Precoda. Listener Experience and Perception of Voice Quality. *Journal of Speech, Language, and Hearing Research*, 33(1):103–115, March 1990.

[67] Nikolaus Kriegeskorte, Marieke Mur, Douglas A Ruff, Roozbeh Kiani, Jerzy Bodurka, Hossein Esteky, Keiji Tanaka, and Peter A Bandettini. Matching Categorical Object Representations in Inferior Temporal Cortex of Man and Monkey. *Neuron*, 60(6):1126–1141, December 2008.

[68] Andrew J Kunkler-Peck and M T Turvey. Hearing shape. *Journal of Experimental Psychology: Human Perception and Performance*, 26(1):279–294, February 2000.

[69] L Lachs, K McMichael, and D B Pisoni. Speech perception and implicit memory: Evidence for detailed episodic encoding of phonetic events. *Rethinking implicit memory*, 2003.

[70] L Lassaletta, A Castro, and M Bastarrica. Musical perception and enjoyment in post-lingual patients with cochlear implants. *Acta . . .*, 2008.

[71] D S Lazard, a L Giraud, E Truy, and H J Lee. Evolution of non-speech sound memory in postlingual deafness: implications for cochlear implant rehabilitation. *Neuropsychologia*, 49(9):2475–2482, 2011.

[72] D S Lazard, H J Lee, M Gaebler, C A Kell, E Truy, and A L Giraud. Phonological processing in post-lingual deafness and cochlear implant outcome. *NeuroImage*, 49(4):3443–3451, February 2010.

[73] Diane S Lazard, Hamish Innes-Brown, and Pascal Barone. Adaptation of the communicative brain to post-lingual deafness. Evidence from functional imaging. *Hearing research*, 307(c):136–143, January 2014.

[74] Dubois Lcpe and Paris E-mail. PERCEPTION , REPRESENTATION AND KNOWLEDGE : ACOUSTIC PHENOMENA BETWEEN NOISE AND SOUNDS. 2003.

[75] Guillaume Lemaitre and Laurie M Heller. Auditory perception of material is fragile while action is strikingly robust. *The Journal of the Acoustical Society of America*, 131(2):1337, 2012.

[76] Guillaume Lemaitre and Laurie M Heller. Evidence for a basic level in a taxonomy of everyday action sounds. *Experimental brain research. Experimentelle Hirnforschung. Expérimentation cérébrale*, 226(2):253–264, 2013.

[77] Guillaume Lemaitre, Olivier Houix, Nicolas Misdariis, and Patrick Susini. Listener expertise and sound identification influence the categorization of environmental sounds. *Journal of Experimental Psychology: Applied*, 16(1):16–32, March 2010.

[78] J W Lewis. Distinct Cortical Pathways for Processing Tool versus Animal Sounds. *Journal of Neuroscience*, 25(21):5148–5158, May 2005.

[79] James W Lewis, William J Talkington, Aina Puce, Lauren R Engel, and Chris Frum. Cortical networks representing object categories and high-level attributes of familiar real-world action sounds. *Journal of cognitive neuroscience*, 23(8):2079–2101, August 2011.

[80] James W Lewis, William J Talkington, Katherine C Tallaksen, and Chris a Frum. Auditory object salience: human cortical processing of non-biological action sounds and their acoustic signal attributes. *Frontiers in systems neuroscience*, 6(May):27, 2012.

[81] Tianhao Li and Qian-Jie Fu. Voice gender discrimination provides a measure of more than pitch-related perception in cochlear implant users. *International journal of audiology*, 50(8):498–502, July 2011.

[82] S Y Liu, T C Liu, Y L Teng, L A Lee, T J Lai, and C M Wu. Environmental Sounds Recognition in Children with Cochlear Implants. *PloS one*, 2013.

[83] J L Loebach and D B Pisoni. Perceptual learning under a cochlear implant simulation. *... on Spoken Language Processing Progress Report ...*, 2007.

[84] Jeremy L Loebach and David B Pisoni. Perceptual learning of spectrally degraded speech and environmental sounds. *The Journal of the Acoustical Society of America*, 123(2):1126, 2008.

[85] Jeremy L Loebach, David B Pisoni, and Mario A Svirsky. Effects of semantic context and feedback on perceptual learning of speech processed through an acoustic simulation of a cochlear implant. *Journal of Experimental Psychology: Human Perception and Performance*, 36(1):224–234, 2010.

[86] V Looi, H McDermott, C McKay, and L Hickson. Music perception of cochlear implant users compared with that of hearing aid users. *Ear and hearing*, 2008.

[87] Valerie Looi and Janna Arnephy. Environmental sound perception of cochlear implant users. *Cochlear implants international*, pages n/a–n/a, January 2009.

[88] Valerie Looi, Kate Gfeller, and Virginia Driscoll. MUSIC APPRECIATION AND TRAINING FOR COCHLEAR IMPLANT RECIPIENTS: A REVIEW. *Seminars in hearing*, 33(4):307–334, November 2012.

[89] Xin Luo, Qian-Jie Fu, and John J Galvin. Vocal Emotion Recognition by Normal-Hearing Listeners and Cochlear Implant Users. *Trends in amplification*, 11(4):301–315, December 2007.

[90] Olivier Macherey and Alexia Delpierre. Perception of musical timbre by cochlear implant listeners: a multidimensional scaling study. *Ear and hearing*, 34(4):426–436, June 2013.

[91] M M Marcell, D Borella, M Greene, and E Kerr. Confrontation naming of environmental sounds. *Journal of clinical and ...*, 22(6):830–864, 2000.

[92] M Marx, C James, J Foxton, and A Capber. Speech Prosody Perception in Cochlear Implant Users With and Without Residual Hearing. *Ear and ...*, 2014.

[93] Z Massida, P Belin, C James, J Rouger, B Fraysse, P Barone, and O Deguine. Voice discrimination in cochlear-implanted deaf subjects. *Hearing research*, 275(1-2):120–129, May 2011.

[94] Z Massida, M Marx, P Belin, C James, B Fraysse, P Barone, and O Deguine. Gender Categorization in Cochlear Implant Users. *Journal of Speech, Language, and Hearing Research*, 56(5):1389–1401, October 2013.

[95] Hugh J McDermott. Music Perception with Cochlear Implants: A Review. *Trends in amplification*, 8(2):49–82, June 2004.

[96] E Molin, A Leijon, and H Wallsten. Spectro-temporal discrimination in cochlear implant users. In *Acoustics, Speech, and Signal Processing, 2005. Proceedings. (ICASSP '05). IEEE International Conference on*, pages –iii/28 Vol. 3. IEEE, 2005.

[97] Il Joon Moon, Eun Yeon Kim, Jin Ok Jeong, Won-Ho Chung, Yang-Sun Cho, and Sung Hwa Hong. The influence of various factors on the performance of repetition tests in adults with cochlear implants. *European Archives of Oto-Rhino-Laryngology*, 269(3):739–745, 2012.

[98] Tova Most and Chen Aviner. Auditory, visual, and auditory-visual perception of emotions by individuals with cochlear implants, hearing AIDS, and normal hearing. *Journal of Deaf Studies and Deaf Education*, 14(4):449–464, December 2008.

[99] Takayuki Nakata, Sandra E Trehub, and Yukihiko Kanda. Effect of cochlear implants on children's perception and production of speech prosody. *The Journal of the Acoustical Society of America*, 131(2):1307–1314, 2012.

[100] Peggy B Nelson, Su-Hyun Jin, Arlene Earley Carney, and David A Nelson. Understanding speech in modulated interference: Cochlear implant users and normal-hearing listeners. *Acoustical Society of America Journal*, 113(2):961–968, February 2003.

[101] M E Niessen and L Van Maanen. Disambiguating sound through context. *International Journal of ...*, 02(03):327–341, 2008.

[102] Grace L Nimmons, Robert S Kang, Ward R Drennan, Jeff Longnion, Chad Ruffin, Tina Worman, Bevan Yueh, and Jay T Rubinstein. Clinical assessment of music perception in cochlear implant listeners. *Otology & Neurotology*, 29(2):149–155, February 2008.

[103] Seung-Ha Oh, Chong-Sun Kim, Eun Joo Kang, Dong Soo Lee, Hyo-Jeong Lee, Sun O Chang, Soon-hyun Ahn, Chan Ho Hwang, Hong Ju Park, and Ja Won Koo. Speech Perception after Cochlear Implantation over a 4-Year Time Period. *Acta Oto-laryngologica*, 123(2):148–153, July 2009.

[104] M J Osberger and L Fisher. Preoperative predictors of postoperative implant performance in children. *Annals of Otology Rhinology and Laryngology*, 109(12):44–46, December 2000.

[105] Thomas J Palmeri and Isabel Gauthier. Visual object understanding. *Nature reviews. Neuroscience*, 5(4):291–303, April 2004.

[106] Nathaniel R Peterson, David B Pisoni, and Richard T Miyamoto. Cochlear implants and spoken language processing abilities: Review and assessment of the literature. *Restorative neurology and neuroscience*, 28(2):237–250, 2010.

[107] D Pressnitzer. Music to Electric Ears: Pitch and Timbre Perception by Cochlear Implant Patients. *Annals of the New York Academy of Sciences*, 1060(1):343–345, December 2005.

[108] D W Proops, I Donaldson, H R Cooper, J Thomas, S P Burrell, R L Stoddart, A Moore, and I M Cheshire. Outcomes from adult implantation, the first 100 patients. *Journal of Laryngology and Otology*, 113(24):5–13, August 1999.

[109] El Mostafa Qannari, Philippe Courcoux, and Pauline Faye. Food Quality and Preference. *FOOD QUALITY AND PREFERENCE*, 32(PA):93–97, March 2014.

[110] R K Reddy, V Ramachandra, N Kumar, and Nandini Chatterjee Singh. Categorization of environmental sounds. *Biological cybernetics*, 100(4):299–306, 2009.

[111] C M Reed and L A Delhorne. Reception of environmental sounds through cochlear implants. *Ear and hearing*, 26(1):48–61, February 2005.

[112] P Robert and Y Escoufier. A Unifying Tool for Linear Multivariate Statistical Methods: The RV- Coefficient. *Applied statistics*, 25(3):257, 1976.

[113] Gerard Roma, Jordi Janer, Stefan Kersten, Mattia Schirosa, Perfecto Herrera, and Xavier Serra. Ecological Acoustics Perspective for Content-Based Retrieval of Environmental Sounds. *EURASIP Journal on Audio, Speech, and Music Processing*, 2010(2):1–11, 2010.

[114] E Rosch. Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, 1975.

[115] E Rosch and B B Lloyd. Cognition and Categorisation, Hilsdale, 1978.

[116] J Rouger, S Lagleyre, B Fraysse, S Deneve, O Deguine, and P Barone. Evidence that cochlear-implanted deaf patients are better multisensory integrators. *Proceedings of the National Academy of Sciences of the United States of America*, 104(17):7295–7300, April 2007.

[117] Julien Rouger, Sébastien Lagleyre, Jean-François Démonet, Bernard Fraysse, Olivier Deguine, and Pascal Barone. Evolution of crossmodal reorganization of the voice area in cochlear-implanted deaf patients. *Human Brain Mapping*, 33(8):1929–1940, May 2011.

[118] Sinan Saraçli, Nurhan Doğan, and İsmet Doğan. Comparison of hierarchical cluster analysis methods by cophenetic correlation. *Journal of Inequalities and Applications*, 2013(1):1–8, 2013.

[119] Valeriy Shafiro. Development of a large-item environmental sound test and the effects of short-term training with spectrally-degraded stimuli. *Ear and hearing*, 29(5):775–790, 2008.

[120] Valeriy Shafiro. Identification of Environmental Sounds With Varying Spectral Resolution. *Ear and hearing*, 29(3):401–420, June 2008.

[121] Valeriy Shafiro, Brian Gygi, Min-Yu Cheng, Jay Vachhani, and Megan Mulvey. Perception of Environmental Sounds by Experienced Cochlear Implant Patients. *Ear and hearing*, 32(4):511–523, July 2011.

[122] Valeriy Shafiro, Brian Gygi, Min-yu Cheng, Jay Vachhani, and Megan Mulvey. Factors in the Perception of Environmental Sounds by Patients with Cochlear Implants. pages 1–10, 2013.

[123] Valeriy Shafiro, Stanley Sheft, Brian Gygi, and Kim Thien N Ho. The Influence of Environmental Sound Training on the Perception of Spectrally Degraded Speech and Environmental Sounds. *Trends in amplification*, 16(2):1084713812454225–101, August 2012.

[124] Valeriy Shafiro, Stanley Sheft, Sejal Kuvadia, and Brian Gygi. Environmental Sound Training in Cochlear Implant Users. *Journal of Speech, Language, and Hearing Research*, 58(2):509–519, April 2015.

[125] Robert V Shannon, Qian-Jie Fu, John Galvin, and Lendra Friesen. Speech Perception with Cochlear Implants. In *Cochlear Implants: Auditory Prostheses and Electric Hearing*, pages 334–376. Springer New York, New York, NY, 2004.

[126] Robert V Shannon, Fan-Gang Zeng, Vivek Kamath, John Wygonski, and Michael Ekelid. Speech Recognition with Primarily Temporal Cues. *Science*, 270(5234):303–304, October 1995.

[127] M Slaney and R F Lyon. A perceptual pitch detector. In *Acoustics, Speech, and Signal Processing, 1990. ICASSP-90., 1990 International Conference on*, pages 357–360. IEEE, 1990.

[128] J L Smith, C L Morgan, and P H White. Investigating a measure of computer technology domain identification: A tool for understanding gender differences and stereotypes. *Educational and Psychological Measurement*, 65(2):336–355, April 2005.

[129] Ginger S Stickney, Fan-Gang Zeng, Ruth Litovsky, and Peter Assmann. Cochlear implant speech recognition with speech maskers. *The Journal of the Acoustical Society of America*, 116(2):1081–1091, August 2004.

[130] K Strelnikov, J Rouger, J F Demonet, S Lagleyre, B Fraysse, O Deguine, and P Barone. Does Brain Activity at Rest Reflect Adaptive Strategies? Evidence from Speech Processing after Cochlear Implantation. *Cerebral Cortex*, 20(5):1217–1222, April 2010.

[131] K Strelnikov, J Rouger, J F Demonet, S Lagleyre, B Fraysse, O Deguine, and P Barone. Visual activity predicts auditory recovery from deafness after adult cochlear implantation. *Brain*, 136(12):3682–3695, December 2013.

[132] Catherine M Sucher and Hugh J McDermott. Pitch ranking of complex tones by normally hearing subjects and cochlear implant users. *Hearing research*, 230(1-2):80–87, August 2007.

[133] Simone Sulpizio, Fabio Fasoli, Anne Maass, Maria Paola Paladino, Francesco Vespignani, Friederike Eyssel, and Dominik Bentler. The Sound of Voice: Voice-Based Categorization of Speakers' Sexual Orientation within and across Languages. *PLoS ONE*, 10(7), 2015.

[134] Ernst Terhardt, Gerhard Stoll, and Manfred Seewann. Algorithm for extraction of pitch and pitch salience from complex tonal signals. *Journal of the Acoustical Society of America*, 71(3):679–688, February 1982.

[135] N Titova and R Näätänen. Preattentive voice discrimination by the human brain as indexed by the mismatch negativity. *Neuroscience letters*, 308(1):63–65, 2001.

[136] N TYEMURRAY, R S Tyler, G G Woodworth, and B J GANTZ. Performance Over Time with a Nucleus or Ineraid Cochlear Implant. *Ear and hearing*, 13(3):200–209, June 1992.

[137] R S Tyler, A J Parkinson, and G G Woodworth. Performance over time of adult patients using the Ineraid or Nucleus cochlear implant. *The Journal of the . . .* , 102(1):508, 1997.

[138] Richard S Tyler and Danielle Kelsay. ADVANTAGES AND DISADVANTAGES REPORTED BY SOME OF THE BETTER COCHLEAR-1MPLANT PATIENTS. *Otology & Neurotology*, 11(4):282, July 1990.

[139] Taina T Välimaa, Martti J Sorri, and Heikki J Löppönen. Speech perception and functional benefit after multichannel cochlear implantation. *Scandinavian Audiology*, 30(1):45–47, October 2009.

[140] Nancy Jean VanDerveer. Ecological Acoustics, 1979.

[141] vhamacgygib. 2007 Gygi. pages 1–6, May 2007.

[142] Susan B Waltzman, J Thomas Roland, and Noel L Cohen. Delayed implantation in congenitally deaf children and adults. *Otology & Neurotology*, 23(3):333–340, May 2002.

[143] W H Warren and R R Verbrugge. Auditory perception of breaking and bouncing events: a case study in ecological acoustics. *Journal of experimental psychology. Human perception and performance*, 10(5):704–712, October 1984.

[144] Robert A Williamson, Kristen Pytynia, John S Oghalai, and Jeffrey T Vrabec. Auditory Performance After Cochlear Implantation in Late Septuagenarians and Octogenarians. *Otology & neurotology : official publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology*, 30(7):916–920, October 2009.

[145] Su Wooi Teoh, David B Pisoni, and Richard T Miyamoto. Cochlear Implantation in Adults with Prelingual Deafness. Part I. Clinical Results. *The Laryngoscope*, 114(9):1536–1540, September 2004.

[146] F Zhao, S D G STEPHENS, S W Sim, and R MEREDITH. The use of qualitative questionnaires in patients having and being considered for cochlear implants. *Clinical Otolaryngology and Allied Sciences*, 22(3):254–259, June 1997.

[147] E Zwicker. Subdivision of the Audible Frequency Range into Critical Bands (Frequenzgruppen). *The Journal of the Acoustical Society of America*, 33(2):248–248, 1961.