Model-based optimisation for enhanced training of individuals based on abilities, learning styles and preferences

V. Govender

A thesis submitted to the
School of Computational and Applied Mathematics,
Faculty of Science,
University of the Witwatersrand,
in fulfilment of the requirements for the degree of
Masters in Science.

07 September, 2012.

DECLARATION

I declare that the contents of this thesis are original except where due references have been made. It is been submitted for the degree of Master of Science. It has not been submitted before for any degree or examination to any other institution.

Govender V.

ACKNOWLEDGEMENTS

Firstly, I would like to thank the all Mighty. His grace and guidance has aided me throughout life.

This thesis would not have been possible without the support and dedication of several people.

To my loving parents, Pragesh and Daisy Govender, and my brother Julian, thank you for all the support and motivation throughout my MSc.

I would like to express my sincere gratitude to my mentors: Prof. M. Ali, Dr. Louis Coetzee and Dr. Quentin Williams for their continuous support during my MSc, for their patience, motivation, enthusiasm, and immense knowledge. Their guidance helped me achieve my goals in this research project. I am entirely grateful for their availability after office hours.

I am especially grateful to my girlfriend Seshni, for helping me with the review of my thesis and for overall support.

My sincere thanks also goes out to the CSIR for offering me an MSc Studentship.

I thank my fellow colleagues: Dario, Asha, Franklin, Morgan, Charles, Obakeng, Tumelo, Eric, Yohana, Simasa, Innocent, Terry and Tanya for the discussions and advice throughout my MSc.

Last but not least, I would like to thank Dr. R. Narain, for the advice he has given me throughout my MSc.

ABSTRACT

Computer based training of individuals is becoming more common. Computer based systems increasingly are filled with devices and appliances that enhance the user's interaction with the computer. These new devices and appliances present new modalities of interaction with the user. This opens new possibilities for computer based training. However, not much is known about mapping these modalities to the user for enhanced learning. This thesis presents an artificial learning model for on-line training of individuals. The model supplied is a multi-modal system in that it links multiple input and output modalities to a user profile. The model contains a non-linear mapping between the user profile and the modalities. The non-linear mapping has been achieved through the use of an Artificial Neural Network. The learning model has been extended to include time dependencies of the suggested modalities via a feedback mechanism within the Artificial Neural Network. The presented results indicate the complexity in choosing the most appropriate mapping for an individual. Results are presented showing the robustness of the learning model. By taking cognisance of the user profile and context (e.g. the user is *bored* or *tired*) appropriate modalities are suggested which facilitate learning.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Ever since computers were introduced, humans have been developing new mechanisms for interacting with the computer allowing new possibilities for user training based on their needs [3]. The integration of the touch screen, stylus and text-to-speech devices to the computer system makes user training even more adaptable and meaningful. Training of a user takes place through the use of modalities based on the user profile, for example, the user's abilities and learning style. The core of the problem lies in effectively choosing the correct modalities for a particular user.

The HCI (Human Computer Interaction) training system based on only one modality is called a uni-modal system of learning (e.g. the use of one input modality and receive output via an output modality, such as using a mouse to click a document and receive text on screen). On the other hand, a multi-modal system of learning, maps a combination of two or more modalities (e.g. using the keyboard and mouse to play a game and receive instructions on screen and via audio). A uni-modal system of learning is rudimentary when compared to a multi-modal system [4]. According to Blattner and Glinert [5], the strengths and weaknesses of each modality are understood clearly in a uni-modal system, but when these modalities are integrated into a multi-modal system they are not understood at the same level, effectively making the use of the multi-modal system an open research area.

There are currently a number of research papers that focus on HCI systems for training of individuals [1, 6, 7, 8]. Gellevij [6], for example, compared a uni-modal system with a multi-modal system using text-only instructions for users. It was found that a multi-modal system was time efficient and had an increased performance when compared to the uni-modal system.

Oviatt [7] investigated the benefits of multi-modal interactions between a computer and a user. The investigation also included non-specialist users (the user is not a specialist in a specific field, e.g. a mathematician studying biology). Depending on the user's ability, this multi-modal system provides the user with a choice of switching modalities. Oviatt [7], however, did not consider users with different abilities (e.g. able-bodied versus a person with disabilities).

Kawai et al. [9], on the other hand, developed a multi-modal training mechanism for both the able-bodied users and users with disabilities. A toolkit of basic mechanisms was presented for selecting the user's preferred modalities.

Coetzee et al. [1] presented a multi-modal system to identify appropriate modalities for a given user profile under the assumption of a linear mapping between the user profile and modalities. The mapping allows the multi-modal system to adapt according to the user profile. A functional approach is assumed, making this multi-modal system different from what is available in current literature. A disadvantage of this approach is that, in real life, a linear relationship may not always be appropriate.

This thesis is an extension of the model suggested by Coetzee et al. [1] by incorporating a non-linear mapping between the user profile and modalities. Artificial Neural Networks (ANN) [10, 48, 54] are introduced to mimic the non-linear mapping. For a given user, a set of modalities is suggested by the non-linear mapping. The ANN based model is extended to include the user's change of status. For example, during the training process if the user gets *bored* or *tired* due to exhaustion and/or noise then this information is fed back to the ANN. Through the use of the feedback ANN model a new set of modalities can be suggested by the non-linear model for the user making the system dynamically adapt to changing circumstances. To date, a utilization of a feedback system has not been covered in literature.

1.1 Background to the problem

The Ability Based Training Intervention (AbTi) is a broader research project undertaken by a number of researchers at the Council for Scientific and Industrial Research (C.S.I.R)[1]. The aim of this project is to develop a framework for multi-modal training via an HCI system, in essence to train individuals with varied profiles and thus increase efficiency of the training. The research carried out in this thesis is a component of the AbTi project.

AbTi attempts to facilitate enhanced information dissemination, which will then be used to provide people with computerised training [11, 12]. Figure 1.1 depicts the various components in the AbTi project. Three architectural layers are identified (as shown in Figure 1.1): the back-end (which will be responsible for the delivery of optimised content in the form of training lessons), the training intervention framework (which will be responsible for dialogue management, interfacing with the back-end, optimised mapping, assessment and other control functionalities) and the multi-modal interaction framework (which will be responsible for managing the interaction with the user through the various modalities) [12]. Each of these layers consists of various sub-modules responsible for specific tasks (such as the integration of inputs or the determination of the optimised choices of modalities for a user based on the preferences, learning styles and abilities). This thesis deals with the third component of AbTi, i.e. the multi-modal interaction framework. In particular, the research involved in this thesis addresses the optimised choices of modalities based on user profiles (otherwise referred to as the enhanced ability based training for an individual).

[1]The C.S.I.R in South Africa performs multidisciplinary research and technological innovation with the aim of contributing to the quality of life of people of this country and to industrial development, and is increasing its scope on the wider continent.

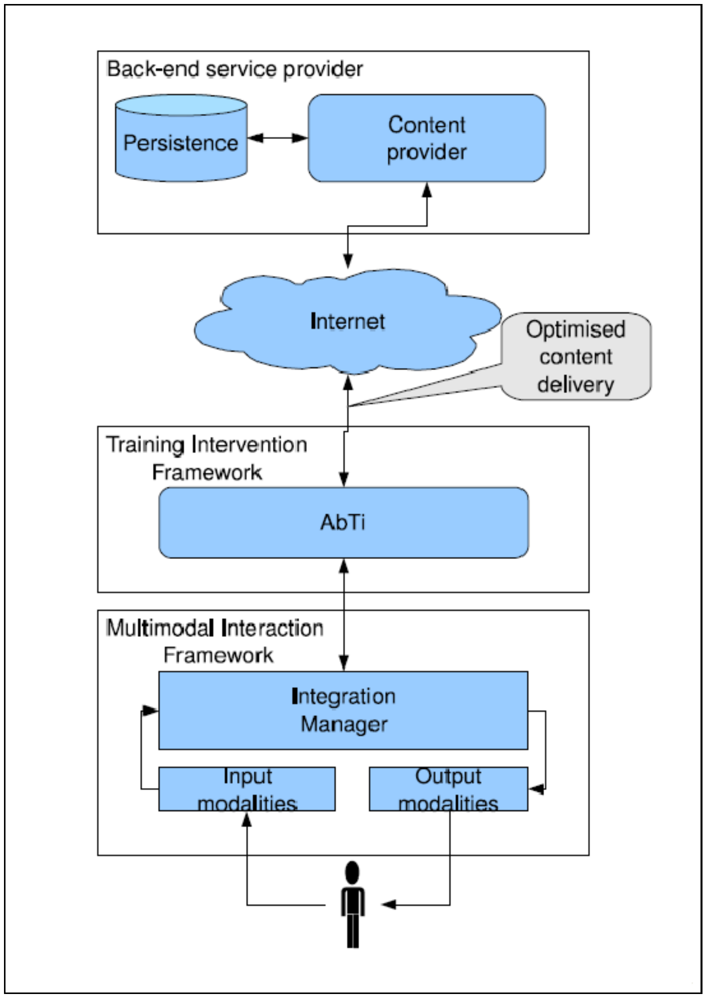

Figure 1.1: Components in the AbTi project [1]

1.2 The problem description

The objective of AbTi is to train individuals with particular content associated with a topic. AbTi will take an individual's profile (trainee's profile) as input, choose the optimal input/output modality combination and interact with the individual via input and output modalities. Some inputs are straightforward while some are found by the interaction with the user. Examples of straightforward inputs are language and literacy level, while inputs like the trainees learning style and preferences can be found through interaction. The learning styles are influenced by a trainee's perceptual and sensory learning preferences. Henceforth, learning styles will refer to both learning styles and preferences.

Thus the problem at hand involves optimally mapping the individual's profile (trainee's abilities, literacy level, learning styles and language), and input and output modalities (instruments or computer devices), assistive technologies and standard computer application and devices. In doing this, the modalities used to facilitate HCI are optimised to allow the user to be trained according to the methods and mechanisms that increase information intake.

1.3 Outline of the thesis

Chapter 2 presents a review of the literature. A number of recent papers on the multi-modal learning system are reviewed. These are methodological papers with applications using test cases. This is then followed by research on the literature referred to as E-learning. E-learning (on-line learning) is also a multi-modal based technology, developed to assist people with training of individuals. The missing elements required for enhanced ability based training for an individual are discussed.

Chapter 3 describes a potential user profile and possible modalities. The user profile consists of the user's abilities, literacy level, learning styles and language. The chapter discusses which aspects of the user's abilities and learning styles are taken into account for the research carried out in this thesis. A set of modalities are presented and an explanation of their usefulness is discussed. The research presented in Coetzee et al. [1] is presented in detail, to provide the foundation for the research carried out in this thesis.

Chapter 4 presents the feed-forward ANN [2]. Elements of the ANN such as sigmoidal units, backward propagation and Levenberg-Marquardt Optimization are discussed. The methodology of data generation and the training of the ANN are addressed.

Chapter 5 presents the numerical results. Results were obtained by implementing both the non-linear learning model and the non-linear learning model with feedback. Numerical results on various experiments are analysed, compared and remarks are made.

Chapter 6 presents concluding remarks. In particular, the chapter presents the strengths and weaknesses of the models presented in the thesis. Recommendations for future research directions are suggested.

1.4 Conclusion

This thesis shows that HCI can be dynamically adapted by changing the input and output modalities used by the computing device to match the user, based on his language, literacy level, learning styles and 'state' at a particular moment. This allows for more efficient user training.

Chapter 2

Literature Study

This chapter presents a review of topics relevant to the research carried out in this thesis. There are three components to the review. Firstly, a number of methodological papers on HCI-based multi-modal learning systems are presented. Secondly, a number of recent practical papers on the multi-modal learning system, where the emphasis was put on on-line learning, for example, E-learning, are presented. Thirdly, a research paper on the on-line system where the learning styles of a trainee have been predicted by the use of ANN, is reviewed.

2.1 Human computer interaction

In current literature, there are a number of research papers that use HCI [13, 14, 15] in training. Carrol [16] gives a detailed history of HCI by reviewing the progression of HCI toward a science of design. Oviatt [7] investigated and described the benefits and use of multi-modal interactions with users. These interactions resulted in the development of multi-modal interfaces that promoted new forms of computing, aiming at non-specialist individuals. Two important case studies were conducted by Oviatt [7, 17] to analyse the use of multiple modalities.

In her first study Oviatt [17] showed a 41% reduction in total error for spoken language within a multi-modal architecture. This indicated that a multi-modal system can be designed to enhance interaction in comparison to a uni-modal system.

A second study [17] was conducted to determine if the multi-modal performance improvements obtained with speakers with an accent were specific to a population. The findings showed a decrease in the total error rate indicating that a multi-modal system can be designed to enhance interaction better than a uni-modal system. Oviatt [17] also indicated the scope of a possible research in multi-modal interfaces whereby user profiles can be adapted.

Kawai et al. [9], on the other hand, developed a multi-modal training mechanism for both the able-bodied users and the users with disabilities. A toolkit of basic mechanisms is presented for selecting the users' preferred modality.

It is reported in [4] that a uni-modal system of mapping is rudimentary when compared to a multi-modal system, in addition, Blattner and Glinert [5], compare the strengths and weaknesses of multi-modal systems over uni-modal systems. Blattner and Glinert suggest a generic platform to support multi-modal interaction.

A user interface, called *Cumulate*, suggested by Brusilovsky [18] for adaptive learning for students [19, 20] exists. The data (students' characteristics) were collected by multiple servers with which the students had interacted with. A database of the students' characteristics was then built, which then formed the basis for individual adaptation. Brusilovsky et al. [21] further tested the interaction of *Cumulate* [18] with an ontology server. This server stores the ontological structures of data and provided the basis for the exchange of higher-level information between different servers (user model that collects data).

Kobsa [22] presents an overview on the development of user-modeling systems. One such system is the *Doppelganger* [23] which gets information about the user via software and hardware sensors. Unsupervised clustering is then used to cluster information of a particular user stereotype, in turn fitting modalities to the user. This stereotyping is a probabilistic rather than deterministic method. Users can also inspect and edit their user models.

Previous literature has shown that adaptive multi-modal systems enhance HCI and user training. However, little has been done in regards to training of users with disabilities and real time feedback mechanisms.

2.2 E-learning

E-learning is computer-based learning, developed mainly for teaching purposes. E-learning facilitates learning via the Internet (internal participation) for both locally-based and international users. Hence it serves as a web-based distance education.

Panjawaranonda and Srivihok [24] proposed an E-learning system for web-based distance education. The design of this E-learning system is based on the assumption that all teachers/users being trained are at the same educational level, otherwise referred to as a 'one-size-fits-all' model.

Blochl et al. [25] proposed an E-learning system to eliminate the 'one-size-fits-all' model. This is a user-centric system focussing on the user's skills, learning styles and learning strategies. In this system, user learning activities are monitored and combined into a user profile. The E-learning system is adjusted according to the dynamic user profile. Blochl et al. [25], however, did not accommodate trainees with disabilities. Hence, no adaptive or assistive technologies were introduced.

Panjawaranonda and Srivihok [24] also presented a user-centric E-learning system. This was developed for teaching and aiding mathematics at primary education level where active participation from both the trainee and the teacher was required.

Panjawaranonda and Srivihok [24] did not consider the use of adaptive/assistive technologies. Sonwalkar [26] thoroughly motivated the use of adaptive/assistive technologies (e.g screen-reader, Braille, etc.) for the E-learning system. A reference to the feedback mechanism in E-learning systems is also made by Sonwalkar [26]. Sonwalkar [26], however, did not present any methodological suggestions on how to use the adaptive technologies in the E-learning system.

2.3 Artificial Neural Network

Villaverde et al. [27] presented a feed-forward ANN based methodology that identified the learning preferences of the individual user within the E-learning system. When a trainee uses a number of actions (such as exam revision) in the e-learning environment as input to the ANN, a number of learning styles are predicted as the output. In essence the ANN is used here to model the learning style. This was not a holistic approach as no modalities, no assistive technologies and no trainee with disabilities were considered.

2.4 Non-linear model

The non-linear multi-model learning system in this thesis relates to the work presented by Coetzee et al. [1] in terms of mapping between input/output modalities and user profile.

The research presented makes use of assistive technologies, as suggested by Sonwalkar [26], in developing the non-linear learning system.

The non-linear model also relates to the work of Villaverde et al. [27] in that a feed forward based ANN has been used to identify the non-linear mapping between modalities and the user profile.

Chapter 3 presents a detailed problem statement.

Chapter 3

Detailed Statement of Problem

This chapter explains learning styles, abilities, literacy level and language of users (user profile). Influences of the user profile on the HCI are also presented, followed by a detailed description of the modalities, in particular input and output modalities used. The chapter also briefly presents the details of the linear model presented by Coetzee et al. [1].

3.1 User profile

The user profile consists of learning styles, abilities, literacy level and language. The type of trainee is determined by the user's learning style, for example, a user that learns by listening to facts (content presented) or audio devices is called an *aural* user, while a *visual* trainee prefers a screen device. It is important to know the type of user, so one can acquire the characteristic modalities based on the different learning styles.

A user profile determines the modalities that a user interacts with, for example, a user who *cannot See* requires an audio device for comprehension. The literacy level places the user in categories of knowledge and learning strengths.

The language aspect of a user aids the user from different linguistic backgrounds, for example, different languages aid different people who may wish to learn in a language other than the English. However, language is an independent factor in the learning of a user.

The main components discussed above classify the user according to learning styles, abilities and literacy. These components are discussed in further detail below.

3.1.1 User's learning style

The learning style of a trainee is defined by Dunn et al. [28] as the way each trainee begins to concentrate on, process and retain new and difficult information. There are currently a number of different schools and models that define learning styles and present assumptions, which are often conflicting [29, 30]. Some schools of learning styles are defined by Felder et al. [31], Honey et al. [32], Dunn and Dunn [33] and Kolb [34] to name a few. It is noted that these schools define different learning styles to aid the content (the materials used to train the trainee) to enable effective learning based on the needs of the individual.

The research carried out in this thesis uses the learning styles as defined by Fleming et al. [35] and Kolb [34, 36] (Kolb's learning style cycle). The reason behind the use of these learning styles is that AbTi has adopted these schools of thought.

Honey et al. [32] made some adaptations to Kolb's learning style model [34, 36]. The learning styles considered in this thesis are taken from Honey et al. [32] and Flemming et al. [35]. Four categories of learning styles suggested by Honey et al. [32], together with their meanings are presented first:

1. *Reflectors* : stand back, gather data, ponder and analyse, delay reaching conclusions, listen before speaking, thoughtful.

2. *Theorists* : think things through in logical steps, assimilate disparate facts into coherent theories, rationally objective, reject subjectivity and flippancy.

3. *Pragmatists* : seek and try out new ideas, practical, down-to-earth, enjoy problem solving and decision-making quickly, *bored* with long discussions.

4. *Activists* : ‘here and now’, seek challenge and immediate experience, open-minded.

In addition to the model from Honey et al. [32], Flemming’s VARK model has aspects which are important. Next, a brief overview of each category of learning style in the VARK (*visual*, *aural*, *read/write*, *kineasthetic*) model suggested by Flemming et al. [35] is presented below:

1. *Visual* trainees use their eyes to learn most efficiently. They prefer to see how to do things or how they work rather than talking about it.

2. *Aural* trainees prefer spoken works and they like to read out loud, they enjoy music and will notice sound effects in movies.

3. The *read/write* trainees like information and material to be displayed in words. Their preferences are text-based display.

4. *Kineasthetic* trainees learn by doing things and through experiments. These trainees are usually good at sport [37, 38].

Trainees differ in their abilities. The next section describes factors associated with a user’s ability.

3.1.2 User’s abilities

Each individual has different abilities which impact on the way the user uses a computer through the modalities. When considering an individual user, cognisance of what the user can do (his abilities) rather than disabilities (what he cannot do) must be taken into account. For example, consider a person who *cannot See*, with assistance from an appropriate modality (through an assistive technology), he can interact with the computer using an automatic speech recognition system (for simplicity sake we use a microphone device to achieve this). Thus the way the learning system is configured (modalities in use) influences the ability of the user to access the information and to enhance his interaction.

The abilities of the user determine the requirements on the configuration of the system. Table 3.1 presents a list of abilities that are involved in creating the output from a user. Specific assumptions in Table 3.1 with regard to a user's ability are made, for example, a user must be able to see (to have the ability) to understand *Sign Language*.

Table 3.1: List of abilities linked with output for the linear and non-linear model

Ability associated with output	**Assumptions**
can See	None
can Hear	None
can Read	User can see
can Understand South African *Sign Language*	User can see
can Feel	None
can Understand Braille	User can feel

Table 3.2 presents a list of abilities that are linked with input into a computer with assumptions, for example, a user needs to feel to use the keyboard.

Table 3.2: List of abilities linked with input for the linear and non-linear model

Ability associated with input	**Assumptions**
can Talk	Voice combined with speech recognition
can Click	Input switch
can Move Pointer	Standard mouse or eyetracker
can Utilise Keyboard	User can feel
can Make Physical Movements	Through sensors

3.1.3 User's literacy level and language

The literacy level is an independent component of the user's profile. As mentioned earlier, South Africa has eleven official languages therefore language needs to be considered when training individuals. This is because some of the individuals, who are going to be trained, will not have English as their mother tongue. Thus, an individual may be literate in a specific language but not in another. Whether an individual has no high school education, has a basic education or has a tertiary qualification, needs to be considered before training takes place. Literacy levels [39, 40] may also be influenced by a disability, for example, a deaf individual is literate in *Sign Language*, but not necessarily literate in a spoken language.

3.2 Input and output modalities

Input modalities are the devices that can be used for input into a computer when the user is using the computer. Such devices are typically the keyboard, mouse, touch screen, microphone, and any device that can be used to communicate to the computer and relay the commands. In the ability based training scenario, the user gives the inputs to the computer, based on what he is asked, then the learning model will compose an individual profile for that user, based on his abilities. For example, AbTi will compile the user's profile. The user gets questions and answers relayed back to him, by the output modalities, such as speakers, a screen, any other output device. Thus the learning model is communicating back to the user. The profile determines the appropriate output modalities with which the user will be trained. A list of the modalities is given in Tables 3.3 and 3.4.

Table 3.3 presents a list of output modalities. These are the devices that the user will use to receive information.

Table 3.4 lists the input modalities. These type of devices are used to input information to the computer. To gain a deeper understanding of the interaction between the user profile and the modalities some scenarios are presented next.

Table 3.3: List of output modalities for the non-linear model

Output devices	**Description**
Screen	An image or representation of an object
Video	A recorded video file, the *visual* component
Sign Language	Audio or text presented by a *Sign Language* interpreter
Symbols	Picture that represents something else by association
Text-To-Speech	Text played as audible output
Audio	Audible sound component
Text	Printed works
Braille	Text output onto a Braille device
Motion	Events presented through force feedback
Vibration	Vibration alerts
Heat	Heat signals for alerts
ZoomUi	Users can change the scale of the viewed area

3.2.1 Interaction of user profile and modalities

Individuals have different needs as based on their abilities, learning styles as well as literacy levels and language [41, 42]. These needs impact on the effectiveness of computer based training. The profile of an individual is unique. To highlight the differences in user profiles a number of examples are presented below.

For example, analyse an individual who *cannot See*, speaks Zulu (a national language in South Africa), is highly educated and is a *read/write* trainee. If it is discovered that it is easier for the individual to be trained in English, rather than his mother tongue, it will have an impact on how the individual is trained. This implies that the individual is a novice in reading and writing Zulu, but an expert in English.

As stated before, individuals have different needs based on their abilities. Most individuals have their five senses [43]: sight, touch, smell, hear and taste. In the case of

Table 3.4: List of input devices for the non-linear model

Input Devices	**Description**
Microphone	Automatic speech recognizer to create character string
Touch and feel	Joystick, sends pointer events
Eye Tracker	Requires helper application to send pointer events
Camera	Requires helper application to create pointer events
Mouse	Sends pointer events
Puff and sip	Sends pointer events
Touch Screen	Sends pointer events
Keyboard	Sends character string
Stylus	Sends pointer events
Switches	Sends pointer events

an individual who *cannot Hear*, there would be no need to stream for audio and have it played. Thus the learning system needs to de-activate audio output, but keep the screen on for *visual* interaction. The same idea is applied to the remaining senses. When an individual has more than one disability, then the problem at hand becomes more complex. This reduces the number of modalities that can be used. For instance, consider a person who *cannot See*. He can use an operating system like Windows, because he can use assistive technology such as a Braille display, screen-reader and audio device. With an individual who *cannot See* and also has a physical disability in his hands, he is prevented from using a Braille display. In this case he has to rely on a text-based text-to-speech application to complete his tasks on a computer. This therefore limits his options and it becomes the task of the learning system to suggest alternative modalities [37].

In the case of a user who is a *visual* trainee but *cannot See*, the user would normally use the visually based modalities (screen device). Given that the user *cannot See*, what other alternate suggestions can be given to this user? The non-linear learning model presented in Chapter 5 gives suggestions of modalities that may be suited for a *visual* trainee that *cannot See*.

3.3 The linear multi-modal system

Coetzee et al. [1] presented a multi-modal system to identify appropriate modalities for a given user profile under certain assumptions. An important feature of the approach is the assumption of the linear relationship between user profile and modalities. This model has been referred to as the linear multi-modal system. Research conducted in this thesis introduces a non-linear relationship between the user profile and modalities. Before the linear model is presented, it is useful to present five tables. Tables 3.1 and 3.2 present a list of abilities that are involved in creating the output and input. Table 3.5 presents and describes the list of preferences. Table 3.6 lists the input modalities used in the linear model. These type of devices, input modalities, are used to input information to the computer. Finally, Table 3.7 presents a list of output modalities used in the linear model. These devices are used to output information to the computer. For completeness, the linear multi-modal system of Coetzee et al. [1] is presented below.

Table 3.5: List of preferences for the linear and non-linear model

Preference	**Description**
visual	User prefers pictures, graphs, diagrams.
aural	User prefers spoken words
read/write	User prefers reading and writing texts
kineasthetic	User prefers by doing (working with hands, moving)

Let $A = \{screen, video, audio, text, ...\}$, see Tables 3.6 and 3.7, and $B = \{can\ See, can\ Hear, can\ Feel, ...\}$, see Tables 3.1 and 3.2, be the set of input and output modalities, and the set of user abilities, respectively. Similarly, let $C =$

Table 3.6: List of input devices for the linear model

Input Devices	**Description**
Microphone	Automatic speech recognizer to create character string
Joystick	Sends pointer events
Eye Tracker	Requires helper application to send pointer events
Camera	Requires helper application to create pointer events
Mouse	Sends pointer events
Head Pointer	Requires helper application to send pointer events
Touch Screen	Sends pointer events
Keyboard	Sends character string
Stylus	Sends pointer events
Switches	Sends pointer events

$\{visual, aural, read/write, kineastethic\}$, see Table 3.5, be the set of perceptual preferences and $D = \{blind, deaf, fully-able, paralysed, ...\}$ be the set of users.

Let E_i be a column vector of size n associated with the user, where $i \in B \cup D$. Elements in E_i take values in $[0, 1]$. These values represent the measure of abilities associated with the user's ability i. Similarly, let F_j be a matrix of $n \times n$. The elements of F_j also take values in $[0, 1]$, $j \in C$. These values represent correlated measures of learning preferences of the user with learning style j (see Table 3.5).

The following relationship is used to calculate the adjusted values for the user profile:

$$p_k = F_j \; E_i, \tag{3.1}$$

where the index represents the adjusted user profile. For example, if i = fully-able, $j = visual$ then k = fully-able with learning preference *visual*. Hence $k \in C \cup D$.

Let G be the dominance matrix of size $m \times n$ (where m is the number of available input and output modalities and n is the number of user abilities) where entries in the

Table 3.7: List of output modalities for the linear model

Output devices	**Description**
Image	An image or representation of an object
Video	A recorded video file, the *visual* component
Animation	Simulation of motion by presenting a series of pictures
Sign Language	Audio or text presented by a *Sign Language* interpreter
Symbols	Picture that represents something else by association
Icons	A small or abstract representation of an object or event
Text-To-Speech	Text played as audible output
Audio	Audible sound component
Music	Audible music sounds
Sound	Audible sounds
Earcons	Audible abstract sounds
Text	Printed works
Simple Text	Printed text converted to a simplified version
Captions	Printed text captions
Braille	Text output onto a Braille device
Tactile	Events presented through force feedback
Vibration	Vibration alerts
Sound Vibrations	Sound converted to vibrations
Heat	Heat signals for alerts

first column are members of A and the entries in the first row are members of B.

The following relationship is then used to calculate the values of the suggested modalities:

$$a_l = G\ p_k, \tag{3.2}$$

where the index l represents values of modalities based on the adjusted user profile. For example, $l =$ fully-able with *visual* preference value. Thus $l \in C \cup D$. It is clear that the linear relationship provides the estimated values for the input and output modalities.

A simple example [1] is used to illustrate the above process. Let E_i be a column vector of size 3, $i \in D$ and F_j be a matrix of size 3×3, $j \in C$. Consider

$$E_i = [0.3, 0.3, 0.3]^T, \tag{3.3}$$

$i \in D$, e.g. $i =$ fully-able. The matrix is then given by

$$F_j = \left(\begin{array}{cccc} & Can\ See & Can\ Hear & Can\ Read \\ Can\ See & 0.6 & 0.0 & 0.0 \\ Can\ Hear & 0.0 & 0.2 & 0.0 \\ Can\ Read & 0.0 & 0.0 & 0.2 \end{array} \right), \tag{3.4}$$

where $j \in C$, e.g. $j = visual$.

The adjusted user profile for a fully-able user with *visual* preference, can be calculated using equation (3.1).

$$p_k = [0.18, 0.06, 0.06]^T, \tag{3.5}$$

$k \in C \cup D$, e.g. $k =$ fully-able with *visual* preference.

A possible dominance matrix G representing values of abilities against modalities (for example, Text, Audio, Image and Video) is given as:

$$G = \left(\begin{array}{cccc} & Can\ See & Can\ Hear & Can\ Read \\ Text & 0.0 & 0.0 & 100 \\ Audio & 0.0 & 100 & 0.0 \\ Image & 100 & 0.0 & 0.0 \\ Video & 100 & 0.0 & 0.0 \end{array} \right). \tag{3.6}$$

By utilizing equation (3.2) and applying equations (3.5) and (3.6), one gets

$$a_l = [6.0, 6.0, 18.0, 18.0]^T. \tag{3.7}$$

The first component in a_l represents the importance of text to the user, the second cell represents the importance of audio, the third cell represents importance of image and the final cell represents importance of video. Thus equation (3.7) indicates the user prefers image and video, which indicates the user has *visual* preference.

Next a different example is considered, i.e. a user who can only hear. The E_i is given by

$$E_i = [0.0, 1.0, 0.0]^T, \tag{3.8}$$

$i \in D$, e.g. $i =$ can only hear. The matrix F_j is given by

$$F_j = \begin{array}{cc} & \begin{array}{ccc} \mathit{Can\ See} & \mathit{Can\ Hear} & \mathit{Can\ Read} \end{array} \\ \begin{array}{c} \mathit{Can\ See} \\ \mathit{Can\ Hear} \\ \mathit{Can\ Read} \end{array} & \begin{pmatrix} 0.2 & 0.0 & 0.0 \\ 0.0 & 0.6 & 0.0 \\ 0.0 & 0.0 & 0.2 \end{pmatrix} \end{array}, \tag{3.9}$$

where $j \in C$, e.g. $j =$ *aural*.

By using equations (3.8), (3.9) and (3.6) and applying equation (3.2) one has

$$a_l = [0.0, 60, 0.0, 0.0]^T. \tag{3.10}$$

Equation (3.10) indicates that the user prefers only audio, which is true for a user who can only hear.

3.4 Linear to non-linear

Coetzee et al. [1] presented a linear multi-modal system which mapped the user profile to modalities. A non-linear system's output is not directly proportional to its inputs. Hence, there may be a combination of inputs that Coetzee et al. [1] model cannot solve. The non-linear model in this thesis aims to develop a model that is equivalent to Coetzee et al.'s model or improve their model.

The core of the non-linear learning model is briefly described. The model requires input and output modalities, and other inputs such as user profile (abilities, learning styles, literacy level and language). A non-linear mapping is created through the use of ANNs, that are built through training data whose creation is presented in Chapter 4. Given all modalities and the user profile, the non-linear model suggests an appropriate combination of input and output modalities to be used for a particular user [11]. The model will monitor the state of the user and continue to optimize and generate a combination of modalities and suggest these to AbTi.

The following chapter provides a theoretical view of ANNs.

Chapter 4

The Artificial Neural Network based Non-linear Model

4.1 Artificial Neural Networks

This chapter presents a brief description of the ANN. A general framework for the multi-layer perceptron ANN is given. The framework also includes the different types of activation functions [44, 45]. A brief overview of backward propagation is given [2]. Finally, the details of how the ANN was modeled for the problem described in Chapter 3 is presented. In particular, how a ANN model was developed to map the user's profile with modalities is discussed.

An ANN is a model that is inspired by the biological nervous system, such as the brain that processes information [46]. The first artificial neural model was developed by McCulloch et al. [10]. Since then, numerous research papers have been published in the field [47, 48, 49, 50]. A very basic description of how the ANN works is presented in Figure 4.1. There are parameters in the ANN which can be adjusted to achieve certain goals. In Figure 4.1, such a goal is considered to be an approximation of an unknown function [2].

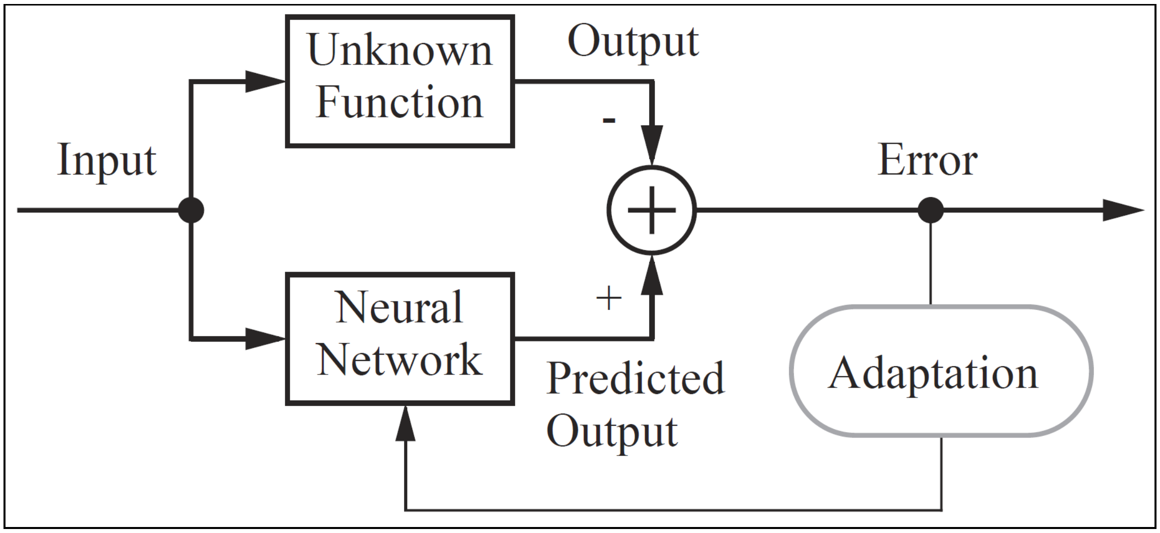

Figure 4.1: Neural Network as a function approximator [2]

4.2 Multi-layer neural network

The multi-layer perceptron consists of layers of neurons. The neurons are stacked together to produce a layer. When the layers are cascaded together, they form the multi-layer neural network. Figure 4.2 depicts a three layer feedforward perceptron ANN. These layers are the input layer, hidden layer and output layer. The input layer is associated with the input vector $(p_1, p_2, \ldots, p_R)^T$, and the output layer is associated with the output vector $(a_1, a_2, \ldots, a_s)^T$. The number of layers in between these two layers are referred to as the hidden layers. Figure 4.2 has only one hidden layer. A more detailed view of a ANN is presented in a number of figures that follow.

4.2.1 Single input neuron

Figure 4.3 shows a single input neuron. The scalar input p is multiplied by the weight w. The unit, 1, is multiplied by bias b. The terms wp and b are then sent to the summation function. The net input (this is the output from the summation), $n = wp + b$, goes into the transfer function f. The scalar neuron output, a, is then calculated as:

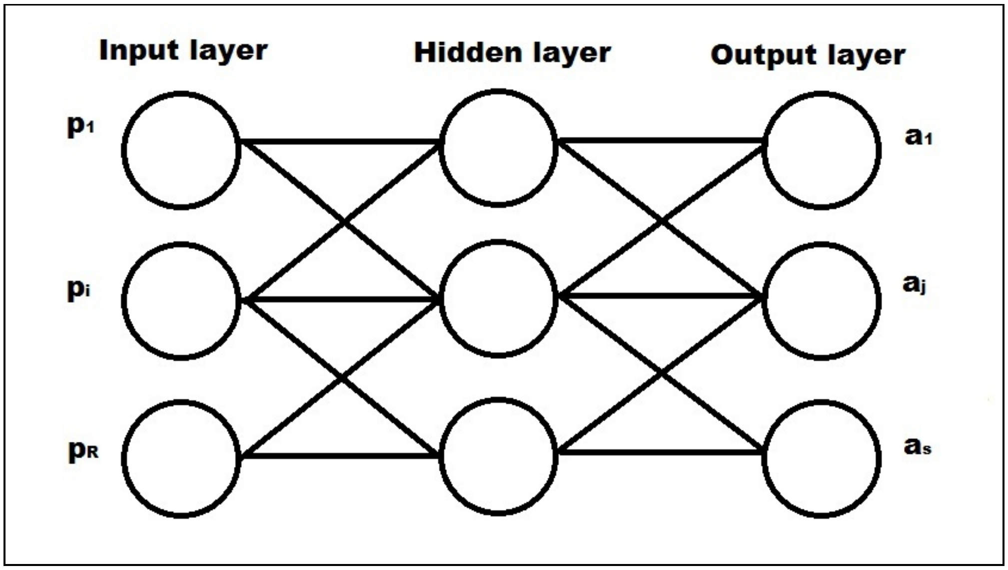

Figure 4.2: General feedforward neural network

$$a = f(wp + b) \tag{4.1}$$

Note that b and w are both adjustable network parameters. The transfer function, f, in Figure 4.3 can be linear or non-linear in nature. Many choices exist for the linear or non-linear transfer functions [51] in a multi-layer network. The choices of transfer functions for the hidden layer neurons may often be different from that of the output layer. This is because neurons at the output layer and neurons at the hidden layer perform different roles. One of the best known transfer functions is the logistic sigmoid transfer function, as shown in Figure 4.4. The function's outputs lie in the range (0,1). It is given by:

$$f(a) = \frac{1}{1 + e^{-a}} \tag{4.2}$$

Another transfer function that can be used is the $tanh$ function and is given by

$$f(a) = tanh(a) = \frac{e^{a} - e^{-a}}{e^{a} + e^{-a}} \tag{4.3}$$

A heaviside, or stepwise, activation function of the form

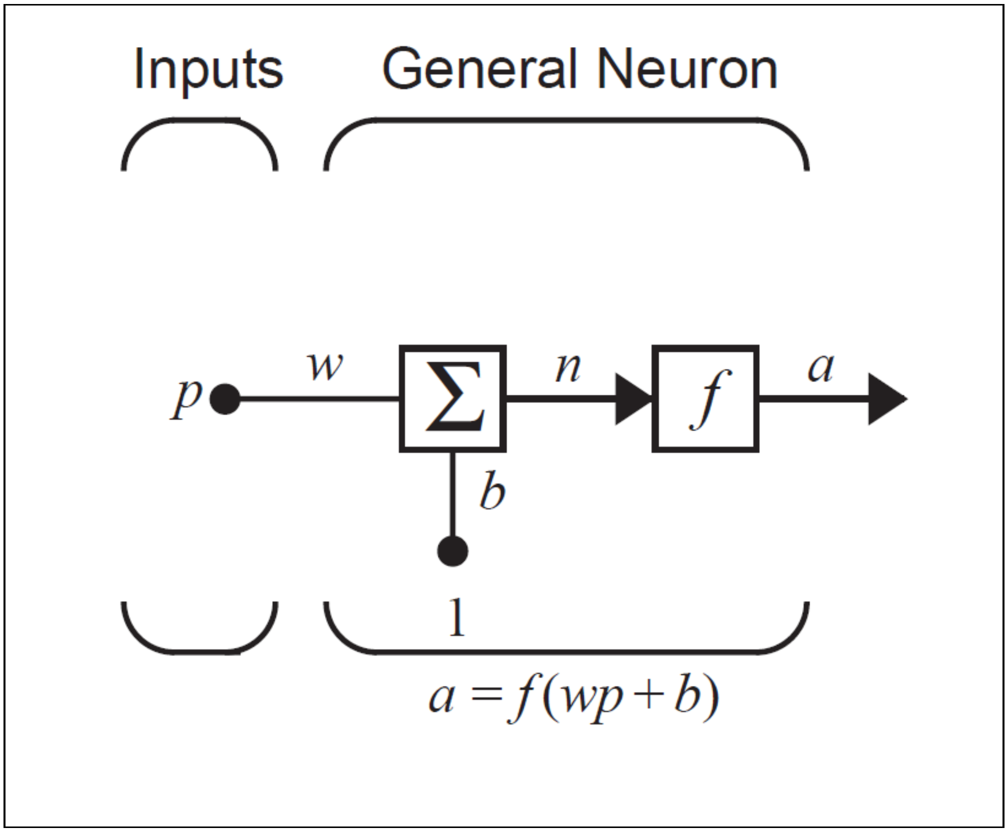

Figure 4.3: Single input neuron [2]

$$f(a) = \begin{cases} 0 & if\ a < 0 \\ 1 & otherwise \end{cases} \tag{4.4}$$

can also be used.

4.2.2 Multiple input neuron

Generally, a neuron has more than one input. Figure 4.5 shows a neuron with R inputs. Similar to the single input neuron, the individual inputs, $p_1, p_2, \ldots, p_R$ are multiplied by weights $w_{j,1}, w_{j,2}, \ldots, w_{j,R}$, where j is the number of layers in the Neural Network, b is the bias and n refers to the net input. Once again, a bias $\vec{b_j}$ is sent to the summation with the weighted inputs to form net input $\vec{n_j}$,

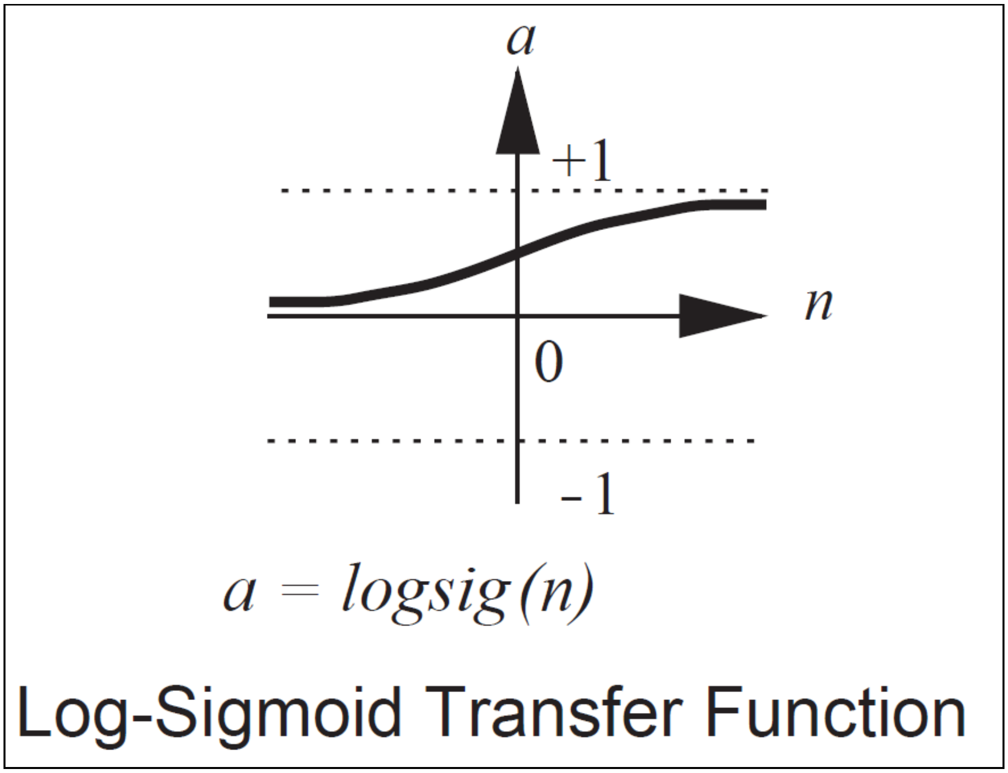

Figure 4.4: Log-Sigmoid transfer function [2]

$$n_j = w_{j,1}p_1 + w_{j,2}p_2 + \ldots + w_{j,R}p_R + b_j \tag{4.5}$$

This can be written as:

$$\vec{n} = W\vec{p} + \vec{b} \tag{4.6}$$

where $\vec{b}$ is a vector and the matrix W is the weight matrix with elements $w_{j,1}, w_{j,2}, \ldots, w_{j,R}$ at the jth row. For the single neuron, W has only one row and in this case n_1 is written as

$$n_1 = w_{1,1}p_1 + w_{1,2}p_2 + \ldots + w_{1,R}p_R + b_1 \tag{4.7}$$

The neuron output can be written as:

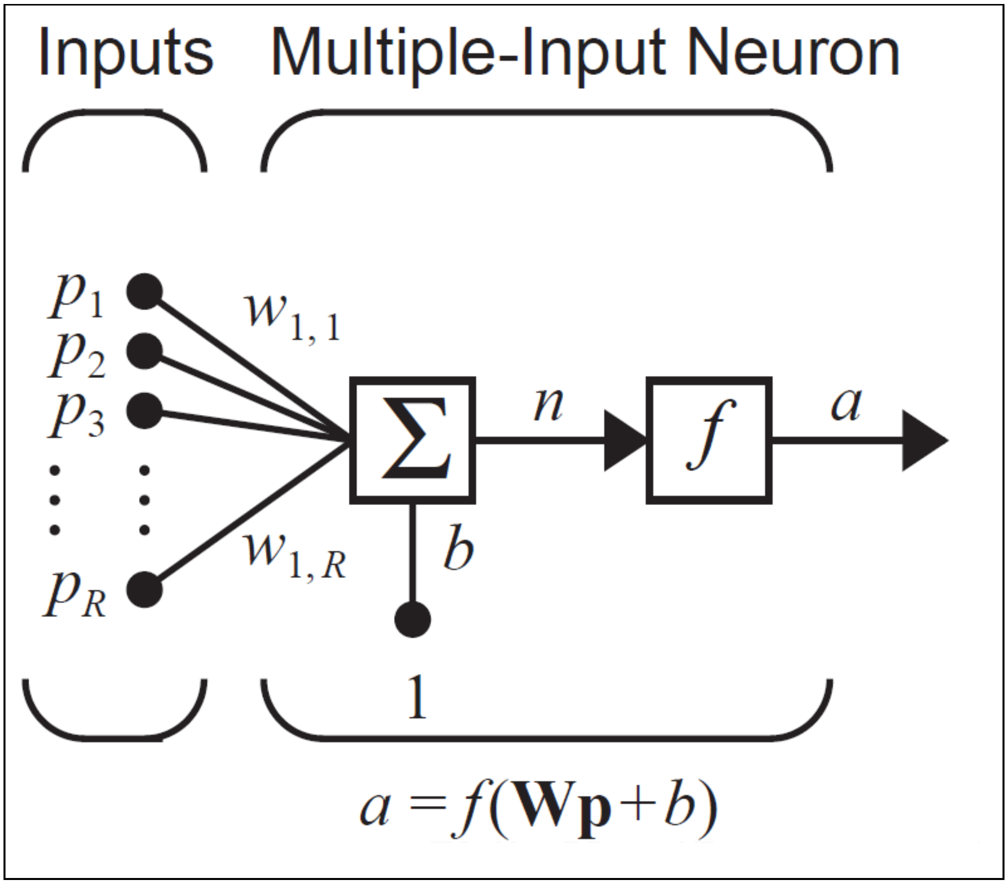

Figure 4.5: Multiple input neuron [2]

$$a = f(Wp + b) \tag{4.8}$$

Figure 4.6 shows a multiple input neuron and is an abbreviated representation of Figure 4.5. As before, $\vec{p}$ is multiplied with the matrix W to form the product $W\vec{p}$. Bias $\vec{b}$ and $W\vec{p}$ enter into the transfer function f. The neuron output a is a scalar in Figure 4.6. If more than one neuron was used, the network output would have been a vector.

Generally, one neuron with many inputs may not be sufficient. One might need a number of neurons working in parallel. This is often referred to as a layer of neurons as depicted in Figure 4.7.

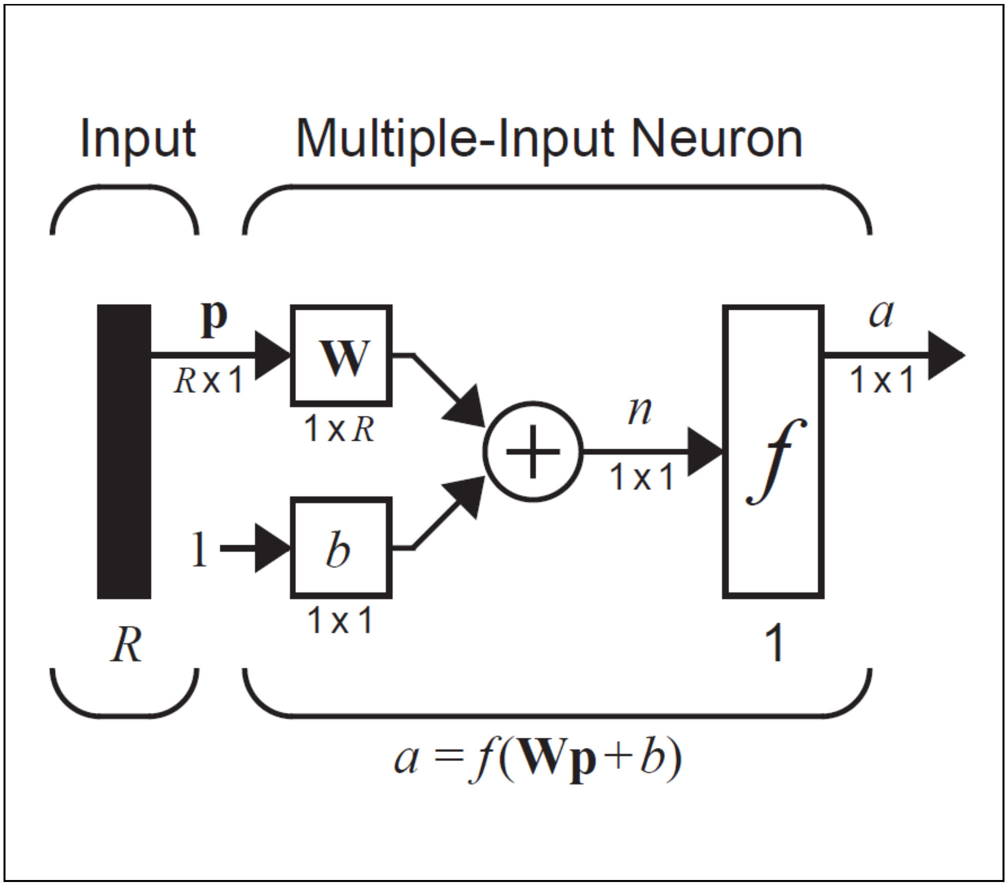

Figure 4.6: Neuron with R inputs [2]

4.2.3 Layer of neurons

Figure 4.7 shows a single-layer network of S neurons. Note that each of the R inputs is connected to each of the neurons and that the weight matrix now has S rows. An abbreviated form of a single layer of S neurons is given in Figure 4.8. The layer includes the weight matrix, the summation, the bias vector $\vec{b}$, the transfer function boxes and the output vector $\vec{a}$. Each element of the input vector $\vec{p}$ is connected to each neuron through the weight matrix W. Each neuron has a bias $\vec{b_j}$, a summation, a transfer function f and an output vector $\vec{a_j}$.

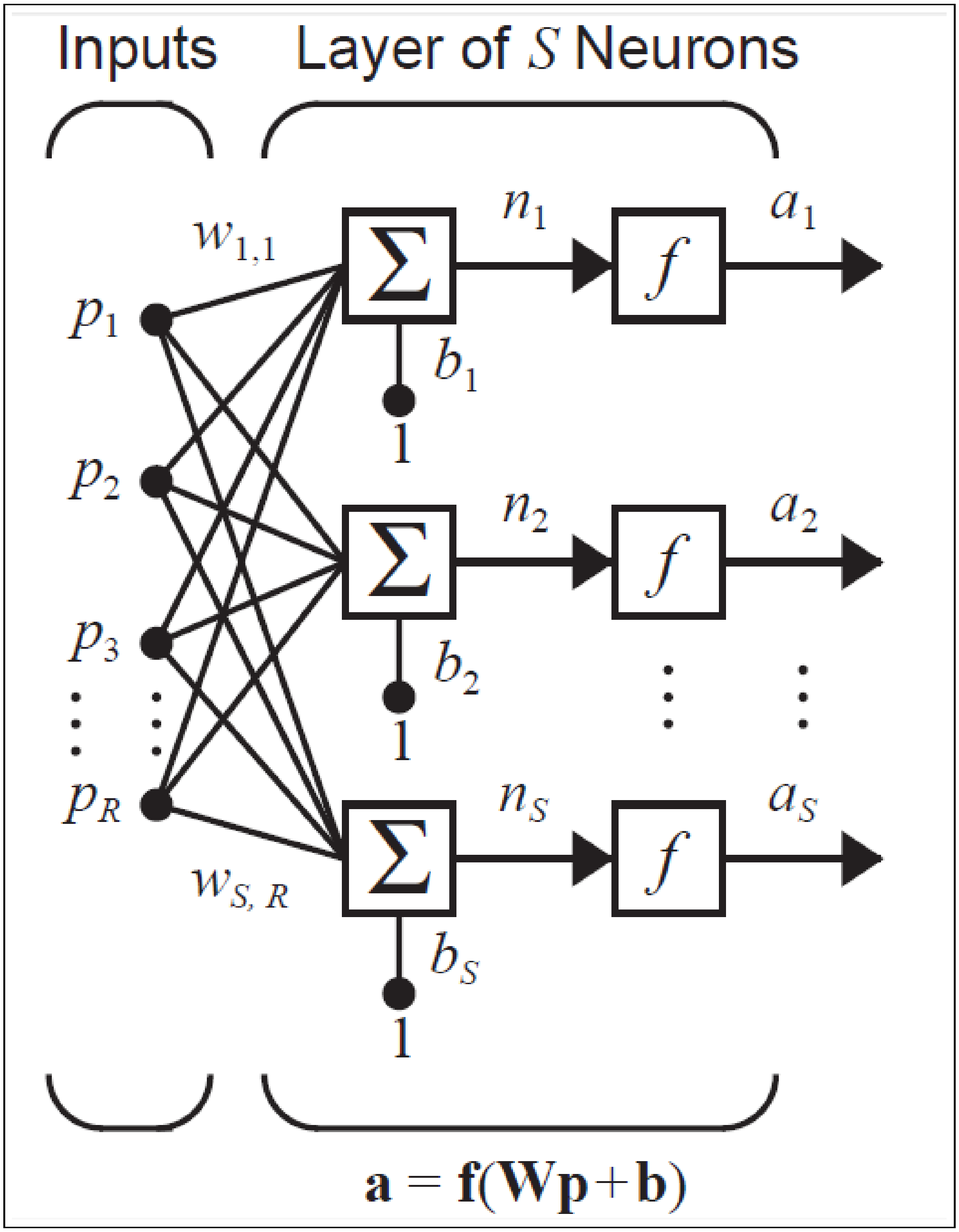

Figure 4.7: A layer of S neurons [2]

4.2.4 Multiple layers of neurons

Consider a network with several layers. Each layer has its own vector bias $\vec{b}$, its own weight matrix, W, net input vector n and an output vector $\vec{a}$. Such a network is presented in

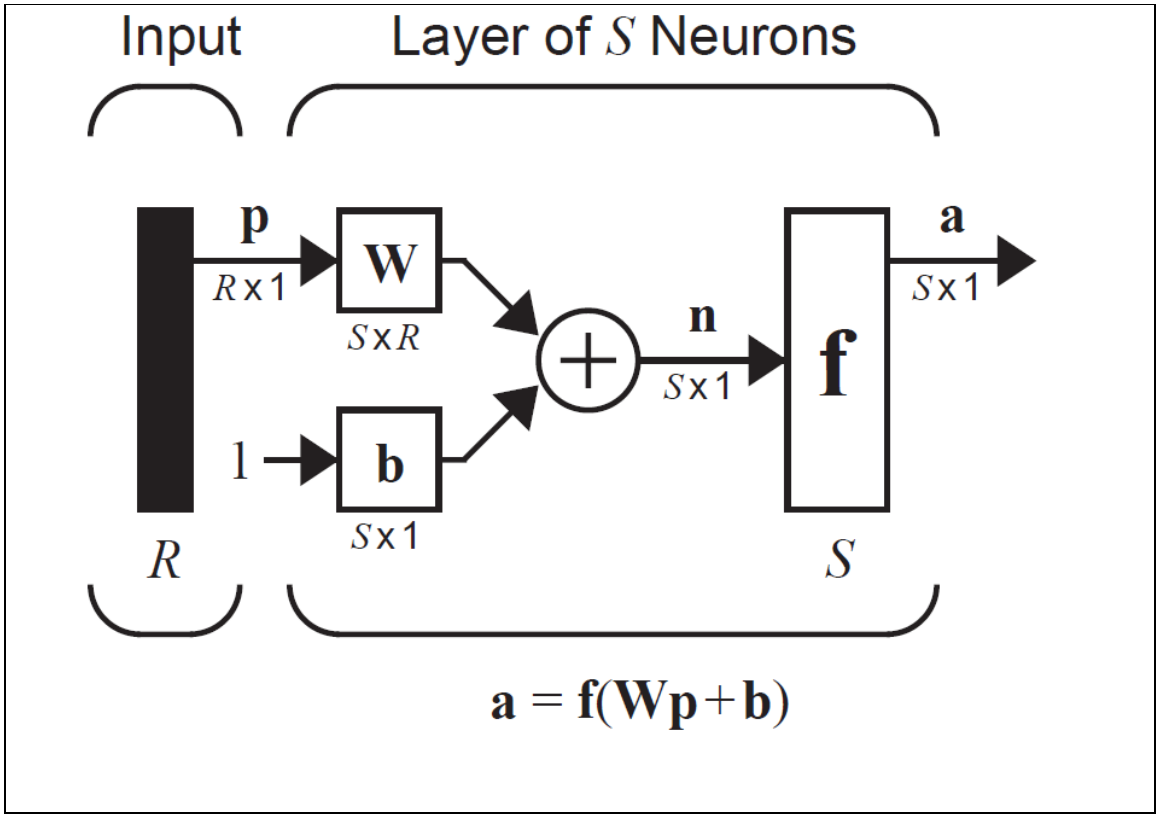

Figure 4.8: Layer of S neurons in an abbreviated form [2]

Figure 4.9, where the weights W^1 and W^2 refer to the first weight layer and second weight layer, respectively. Figure 4.9 is a three layer network.

In Figure 4.9, there are R inputs, and S^1, S^2 and S^3 neurons in the first, second, third layers respectively. A layer with output a is called the output layer of the ANN, e.g. the last layer. The other layers are referred to as the hidden layers of the ANN. Thus Figure 4.9 has an output layer and two hidden layers.

4.3 Backward propagation

Backpropagation [52, 2] involves minimizing the total squared error of the output computed by the net. The training of a network by backpropagation involves three stages: the

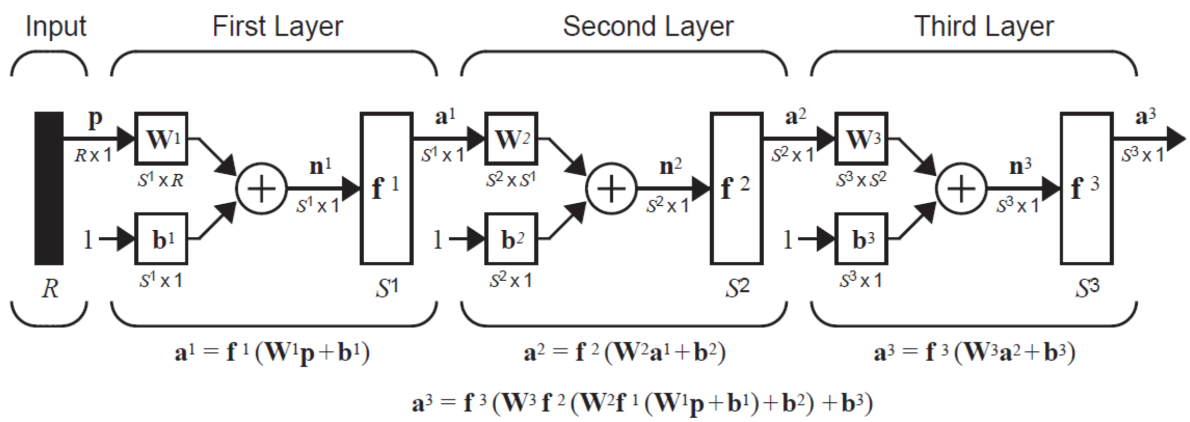

Figure 4.9: Three layer network [2]

feedforward of the input pattern, the calculation and backpropagation of the associated error, and the adjustments of the weights.

There are different methods that can be used in backpropagation, for example, quasi-Newton and conjugate gradient are more efficient than steepest decent algorithms [53]. Another method is the Levenberg-Marquardt algorithm [54] which was used in this study. This algorithm is very efficient for training small to medium-size networks. The network is trained using the Levenberg-Marquardt algorithm [2, 52] with the use of a sigmoid activation function.

4.4 The Levenberg-Marquardt algorithm

The Levenberg-Marquardt algorithm is an iterative technique that locates the minimum of a function that is expressed as the sum of squares of non-linear real valued functions [44].

It is considered to be a combination of steepest decent and the Guass-Newton method. If the current solution is far from the optimal solution, the algorithm behaves similar to the steepest decent method, i.e. slow but is guaranteed to converge. It becomes a Guass-Newton method when the current solution is close to the optimal solution. To show that the Levenberg-Marquardt algorithm is considered a combination of the methods discussed above, a short description of the Levenberg-Marquardt learning cycle [54] follows.

The change (Δ) in the weights ($\vec{w}$) is obtained by solving

$$\alpha\Delta = -\frac{1}{2}\nabla E \tag{4.9}$$

where E is the mean squared network error

$$E = \frac{1}{N}\sum_{k=1}^{N}(\vec{y}(x_k) - \vec{d_k})^2, \tag{4.10}$$

N is the number of profiles (data entries), $\vec{y}(x_k)$ is the network output corresponding to the profile, x_k, and $\vec{d_k}$ is the desired output for that particular profile.

The elements of the α matrix are given by

$$\alpha_{ij} = (1 + \lambda\delta_{ij})\sum_{r=1}^{p}\sum_{k=1}^{N}[\frac{\partial y_r x_k}{\partial w_i}\frac{\partial y_r x_k}{\partial w_j}] \tag{4.11}$$

where p is the number of outputs of the network and

$$\delta_{ij} = \begin{cases} 1 & if\ i = j \\ 0 & otherwise \end{cases} \tag{4.12}$$

Beginning with random weights, α and ∇E are evaluated, and by solving (4.9), the new weights are obtained ($\vec{w'} = \vec{w} + \Delta$).

4.5 Modeling the user's profile and modalities with neural networks

In our work the ANN architecture has been used to find a mapping between the user's profile and a set of modalities as depicted in the prototype ANN in Figure 4.10. To achieve this goal, the inputs of the network need to be identified, with the network's corresponding outputs. It is also necessary to determine other architectural parameters, such as the number of hidden layers to be used, the number of processing units (neurons) in each of the hidden layers, and the activation function to be used in the processing units. Each of these issues is analysed in the following subsections.

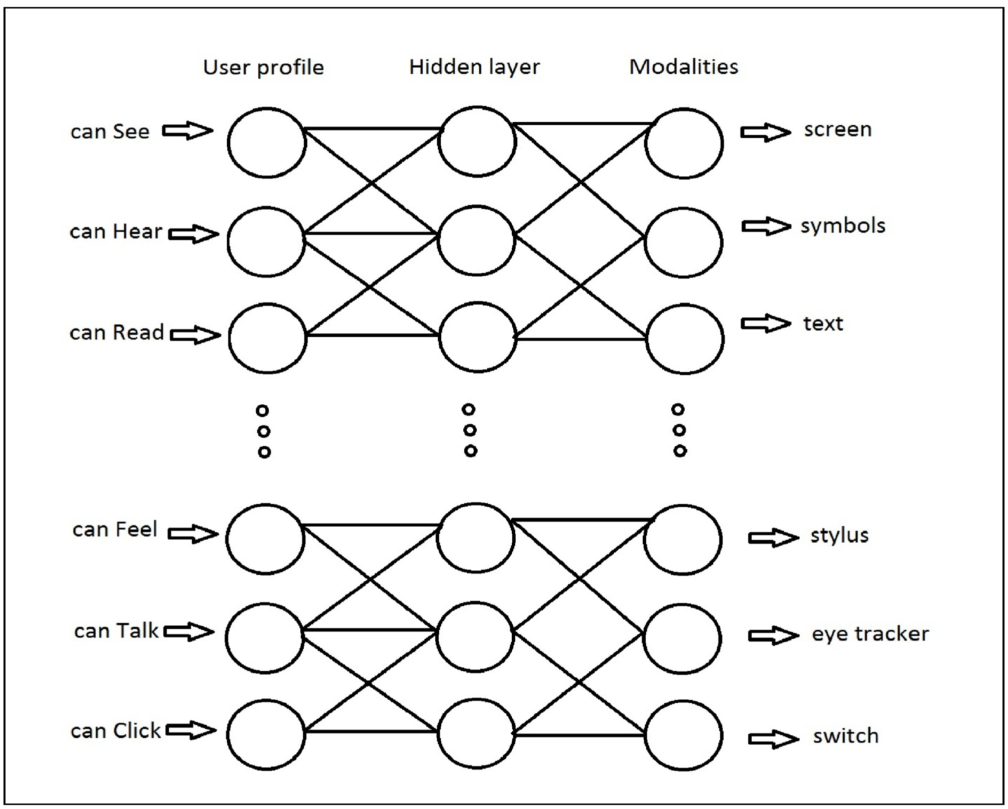

Figure 4.10: Feedforward neural network with user profile and modalities

4.5.1 Data generation

For training the ANN, input and output data sets are required. The input data set of 11 user abilities,

$$\begin{aligned}&\{can\ See, can\ Hear, can\ Read, can\ understand\ South\ African\ Sign\ Language,\\&can\ Feel, can\ Talk, can\ Click, can\ Move\ Pointer, can\ Utilise\ Keyboard,\\&can\ make\ Physical\ Movements, can\ understand\ Braille\}\end{aligned} \tag{4.13}$$

was used, as originally defined in Tables 3.1 and 3.2. Each ability in this set takes on either 0 or 1 values. This means that there is a set of 2^{11} vectors, E, of dimension 11×1 (see E in Chapter 3.3) which can be considered. However, for the numerical experiment a set of only 250 randomly generated vector values were considered. Data were generated from a uniform distribution, $Unif$. For example, $E(i) = E(can\ Hear)\ \sim Unif(\{0, 1\})$ and $E(i) = E(can\ See) \sim Unif(\{0, 1\})$.

For each vector, E, four learning styles (*visual*, *aural*, *read/write*, *kineasthetic*) were considered and four different diagonal matrices of size 11×11 (see F in Chapter 3.3) were created (i.e. one for each learning style).

These diagonal matrices are associated with the learning styles and are created using a rule based technique. For example, the diagonal matrix associated with the learning style *aural* was created as follows.

For an *aural* trainee[1] the diagonal element $F(i, i) = F(can\ Hear, can\ Hear) = 0.4$ and $F(j, j) = F(can\ Talk, can\ Talk) = 0.4$. The remaining diagonal elements will have low values, which have been randomly generated from the set

[1]For a *visual* learning style, the element $F(can\ See, can\ See) = 0.6$ was used, the remaining values were used from the set $\{0.01, 0.02, \ldots, 0.22\}$. For a *read/write* learning style, the element $F(can\ Read, can\ Read) = 0.4$, and $F(can\ utilise\ keyboard, can\ utilise\ keyboard) = 0.4$ were used, the remaining values were used from the set in Equation (4.14). For a *kineastethic* learning style, the element $F(can\ Feel, can\ Feel) = 0.4$, and $F(make\ physical\ movements, make\ physical\ movements) = 0.4$ were used, the remaining values were used from the set in Equation (4.14).

$$\{0.01, 0.02, \ldots, 0.04\}, \tag{4.14}$$

due to the fact that these abilities are not correlated with *aural*.

The vector E and the diagonal matrix F are then multiplied to create the vector p of size 11×11 (see Equation (3.1), Chapter 3.3), i.e. $p = FE$. Hence there are four p vectors due to the four learning styles.

Elements $p_1, p_2, \ldots, p_{15}$ of each of these p vectors were used as the inputs to the ANN, see Figure 4.7. The elements $p_1, p_2, \ldots, p_{11}$ are the values of p, and the elements, e.g. $p_{12}, p_{13}, p_{14}, p_{15}$ are the values of the learning style, for example *aural*. Only one of these will have the value 1 and the rest will have the value 0, e.g. $aural = 1$.

The set of 22 elements,

{*screen, symbols, text, video, audio, Braille, Sign Language, text-to-speech, vibration, touch screen, heat, motion, zoomUI, mouse, touch and feel, microphone, keyboard, camera, stylus, eye tracker, switch, puff and sip*} (4.15)

together with the set in equation (4.13) were used to create the dominance matrix, D, of size 11×22. The elements of this dominance matrix are again found by the rule based method as follows.

For a user who *cannot See*, the *screen, video, symbols* and *Sign Language* are deactivated, while for a user who *cannot Hear*, *audio* are deactivated, i.e. $D(audio, can\ See) = 0$. The dominance matrix is defined as

$$D = \begin{pmatrix} & can\ See & can\ Hear & \ldots & can\ Talk \\ screen & d_{11} & d_{12} & \ldots & d_{1m} \\ audio & d_{21} & d_{22} & \ldots & d_{2m} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ switch & d_{n1} & d_{n2} & \ldots & d_{nm} \end{pmatrix}. \tag{4.16}$$

One can let $d_{11} = D(screen, can\ See) = 100$, because one needs the ability 'can See' to view the *screen*. A user needs the ability 'can See' to read *text* hence $D(text, can\ See) = 50$. The first value 100 is due to the direct relationship between *screen* and 'can See', while the second value 50 is the result of the indirect relationship between *text* and 'can See'. Similarly, $D(audio, can\ See) = 0$, because the user does not need the ability 'can See' to listen to *audio*. The values of the indirect relationship are chosen from the set {0, 22, 50}.

The vector p is now multiplied with the dominance matrix to give the importance vector a (see equation (3.2), Chapter 3.3) of dimension 22×1. Hence, for each vector E there are four a vectors. The elements of each of these a vectors are now used as the outputs $a_1, a_2, \ldots, a_{22}$ in the ANN, see Figure 4.7.

For each vector E there are four sets of inputs, $p_1, p_2, \ldots, p_{15}$, and four sets of corresponding outputs, $a_1, a_2, \ldots, a_{22}$. Hence the total number of data consists of 1000 (250×4) inputs and 1000 outputs. The non-linear model is then created by training the ANN using 1000 input and output data sets.

As in the linear model, the non-linear model with the change in state was created using the ANN. Four different states, (*energetic*, *bored*, *tired*, *lazy*), were considered. The objective was to develop a time dependent learning model. If a trainee moves from one state to another (for example, from *energetic* to *tired*), his suggested modalities are affected. To create the learning model to consider such a change of state, the ANN is again trained. The previous set of p values were used, but for each vector $p = (p_1, p_2, \ldots, p_{15})^T$ a new variable H is introduced. Corresponding to the four different states the value of H falls into four sub-intervals [0,25], (25,50], (50,75], (75,100] respectively for *lazy*, *tired*, *bored*, *energetic*. The vector p is extended with H to give rise to dimension 16, where $p_{16} = H$. For each vector p of dimension 16, four values of H, one from each sub-interval, are generated randomly. This results in a total of 4000 p vectors. The 4000 dominance matrices are again created using a rule based method.

Consider the dominance matrix, D of dimension 22×16, given in Equation (4.16), as the dominance matrix for the state *energetic*[2]. The following concepts are used to obtain the matrix D for the remaining states. The non-linear model without feedback is first used to identify a set of preferred modalities for each type of learning style (*visual*, *aural*, *read/write*, *kineasthetic*). An example for the *aural* trainee is presented. Consider the change of state *energetic* to *bored*. The dominance matrix, D, for the *energetic* state is used as a base-line matrix.

For the *aural* trainee the values of 80% of the preferred modalities are scaled down by 25%. For example, if there are five preferred modalities, all with values 100 in the matrix, then four of these (chosen at random) values are reduced. That is, in the new matrix $D(can\ Hear, microphone) = 100$, and $D(can\ Hear, audio) = 75$, since *audio* is one of the four randomly chosen preferred modalities. This can occur because, for example, when the trainee's state is changed from *energetic* to *bored*, the *bored* trainee might prefer to talk rather than to listen. The remaining values of the dominance matrix, which are between 0 to 50, are also scaled down to 10% to 30%. This percentage is chosen at random. Therefore for any change of state from one state to another, the dominance matrix for the *energetic* state is used as the base line matrix and the above procedure is used to find the corresponding dominance matrix for the trainee. As before the p vector is used as the input and the a ($Dp = a$) vector is used as output. Hence the total number of data consists of 4000 (1000×4) inputs and 4000 outputs. The non-linear model with feedback is then created by training the ANN using these parameters.

4.5.2 Hidden Layer

There is no theory about how many hidden layer units are needed to approximate any given function [27]. There are some rules [55] for determining the number of neurons in the hidden layer, see Figure 4.9, of a multi-layer feed-forward ANN for example. Ultimately, the number of hidden layers are determined via trial-and-error experimentation. A total of 15 units was found to be appropriate for the study carried out here. This ensured that

[2]This matrix now contains H values from the sub-interval (75,100].

the training process was not very complex and time costly.

4.5.3 Training of the neural network

The proposed model for modality recognition is a three-layered feed-forward ANN with Levenberg-Marquardt learning algorithm. In this model, the first layer contains a total of 15 input units (16 for the non-linear model with feedback); the hidden layer contains 15 processing units, which were found by a trial and error process, that are connected to 22 output units in the third layer.

The data has been generated and the components input, hidden and output layers of the ANN have been decided. The programming of the ANN was carried out in Matlab by Viren Govender with the aid of the built in functions found in Matlab. For the training of the ANN, one has to randomly split the data into two categories,

Training data : these are presented to the network during training, and the network is adjusted according to its error. Out of 1000 (4000 for the non-linear model with feedback), 800 (respectively 3200 for the non-linear model with feedback) data sets were used in training the ANN.

Test data : the remaining 20% of the data was used to test the ANN. These have no effect on training and so provide an independent measure of network performance.

Utilizing the ANN approach with data generated as described, a number of experiments was conducted to measure the effectiveness of the non-linear approach. These results are presented in Chapter 5.

Chapter 5

Results and Findings

This chapter compares the results of the non-linear model as presented in Chapter 4 with those obtained by the linear model from Coetzee et al. [1]. A comparison of both models using a baseline user profile is presented. Results produced by the non-linear model for different profiles are then presented. This is then followed by an analysis of profiles against the change in trainee states. Note that the puff and sip devices have been used to indicate 'yes' and 'no' respectively, while the zoomUi device used to zoom in or zoom out on the screen.

5.1 Comparison of the linear and non-linear model

The non-linear model and the linear model presented by Coetzee et al. [1] will be compared using a sample problem. Results[1] obtained by Coetzee et al. [1] for a *Fully-able* trainee are summarised in Figure 5.1.

In Figure 5.1, the vertical axis presents the cost value of the suggested modalities against four different learning styles (*visual*, *aural*, *kineasthetic* and *read/write*) in a scale of 0 to 10, while the horizontal axis presents the output content elements, i.e. suggested

[1]The first author Coetzee in Coetzee et al. [1] was kind enough to provide us with experimental results associated with Figures 5.1 and 5.3.

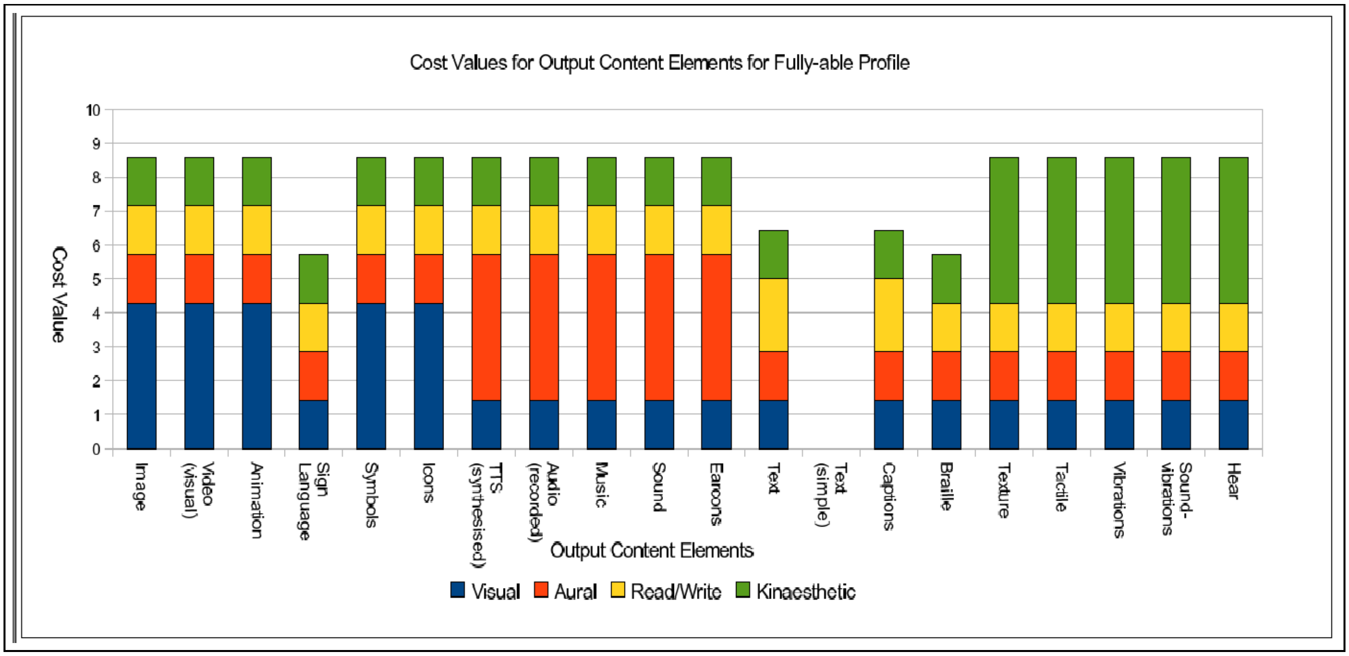

Figure 5.1: Cost values for a *Fully-able* trainee based on output modalities [1]

modalities. The cost value allows for the identification of the appropriate HCI components for a specific user profile. Experimental results of the non-linear model obtained using the same trainee profile, are summarised in Figure 5.2. In Figure 5.2, the vertical axis presents suggested modalities, while the horizontal axis presents the values (importance values) of the suggested modalities against four different learning styles on a scale of 0 to 10. The importance values are a measure of one modality against another. For example, if a trainee has a cost value of 10 for *audio* and 9 for *microphone*, then this results in the trainee prefering *audio*, followed by *microphone*. Two different output modalities, *video* and *text*, against the four learning styles are analysed in the comparison of the linear and non-linear models.

In Figure 5.1, the cost values of the modality *video* are 4.5, 1.5, 1.5 and 1.5 for *visual*, *aural*, *kineasthetic* and *read/write* preferences, respectively. In Figure 5.2, the suggested modality values for the same modality are 5, 1.7, 1.7 and 1.7 for *visual*, *aural*, *kineasthetic* and *read/write*, respectively. The comparison clearly shows that the non-linear model

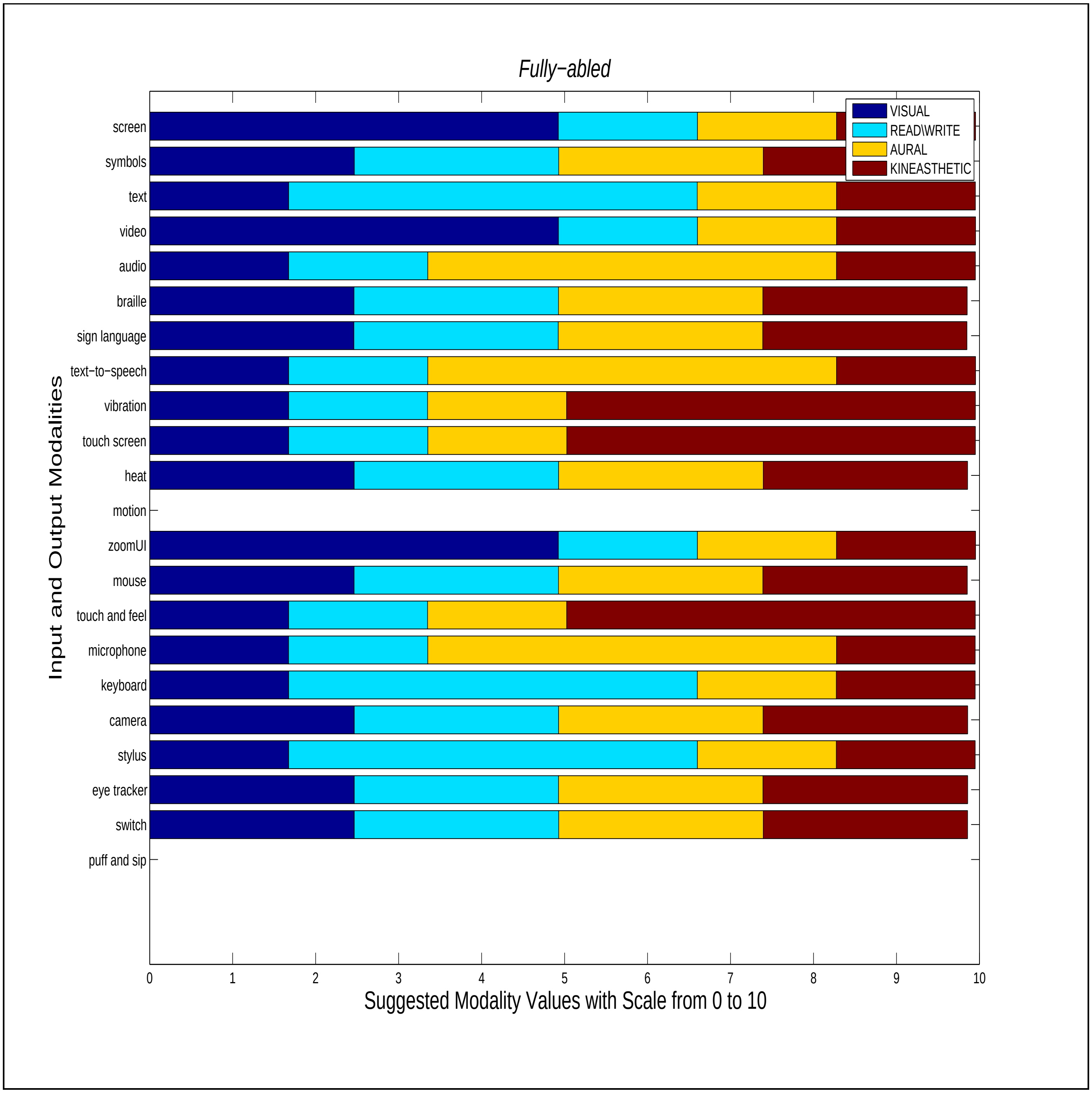

Figure 5.2: Modality values for a *Fully-able* trainee

predicts higher values for *video* against all learning styles.

In Figure 5.1, the cost values of the modality *text* are 1.5, 1.5, 1.5 and 2 for *visual*, *aural*, *kineasthetic* and *read/write*, respectively. In Figure 5.2, the suggested modality values for the same modality are 1.7, 1.7, 1.7 and 5 for *visual*, *aural*, *kineasthetic* and *read/write* preferences, respectively. For a *fully-able* profile, results between the linear and non-linear models are comparable. This indicates that the non-linear model has the capacity to effectively represent, model and provide reasonable suggestions for the problem as described in Chapter 3.

In addition to the above comparison, Figure 5.1 shows that the importance value of *text* is 2 for the *read/write* trainee and the value of *video* for the *read/write* trainee is 1.5. These two values are similar. Intuitively, for a *read/write* trainee the value of *text* should have been higher.

This kind of discrepancy did not occur in the non-linear model as the predicted values of 5 and 1.7 are far apart for *text* and *video* respectively for a *read/write* trainee.

The output modalities have been described above. The comparison is extended by analysing the input modalities. In Figure 5.3, the horizontal axis presents the input elements, i.e. suggested input modalities, while the vertical axis presents values (importance values) of the suggested modalities against four different learning styles (*visual*, *aural*, *kineasthetic* and *read/write*) in the scale of 0 to 20. Two different modalities, *microphone* and *stylus*, against the four learning styles are compared to Figure 5.2.

In Figure 5.3, the cost values of the modality *microphone* are 2, 10, 1 and 1 for *visual*, *aural*, *kineasthetic* and *read/write*, respectively. In Figure 5.2, the suggested modality values (scaled up values on the scale of $0 - 20$) for the same modality are 3.3, 10, 3.3 and 3.3 for *visual*, *aural*, *kineasthetic* and *read/write*, respectively. The non-linear model has, once again, predicted comparable values to the linear model as suggested by the input modality *microphone*.

In Figure 5.3, the cost values of the modality *stylus* are 2, 1, 10 and 1 for *visual*, *aural*, *kineasthetic* and *read/write*, respectively. In Figure 5.2, the suggested modality values (scaled up) for the same modality are 3.3, 3.3, 3.3 and 10 for *visual*, *aural*, *kineasthetic* and *read/write*, respectively. In terms of the high importance value, it is noted that

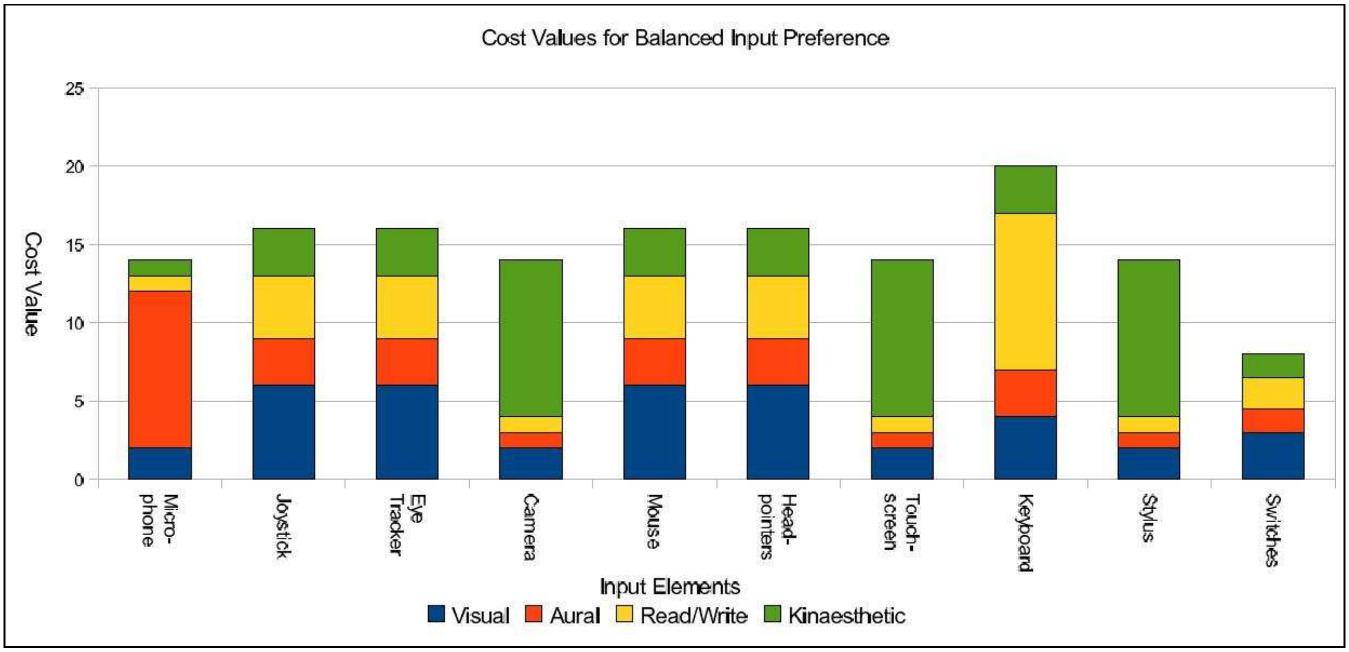

Figure 5.3: Importance values for a *Fully-able* trainee based on input modalities [1]

the linear model suggests *stylus* for the *kineasthetic* trainee, while the non-linear model suggests the *stylus* for the *read/write* trainee. Hence the prediction by the non-linear model is more intuitive, because the *stylus* device should have a higher correlation with the *read/write* trainee than with the *kineasthetic* trainee. Based on the analysis presented in this section, the non-linear model presents comparable results to those of the linear model. It has highlighted some discrepancies in the linear model. The results also suggest that the non-linear model will be effective for more complex profiles.

The following sections present results of the non-linear model for a number of scenarios.

5.2 Numerical results for the non-linear learning model

The non-linear model was tested and the results were analysed using various profiles. The following were taken into consideration:

- A trainee who *cannot See*,

- A trainee who *cannot Hear*,
- A trainee who is *Fully-able* and can understand *Sign Language*,
- A trainee who has recently (has not learnt to adapt to the disability) acquired disabilities, namely he cannot *Hear* and he cannot understand *Sign Language*.

First the non-linear learning model was tested using two disabilities (cannot *See* and cannot *Hear*) and the suggested modality values were recorded on four learning styles of the trainee. Next, the suggested model was tested using the *Fully-able* profile. This profile has been tested to see how the results obtained compare to those of the trainees with disabilities. Finally, a trainee who has recently acquired a disability was considered. This is considered to verify if the set of modalities predicted by the model are realistic.

5.2.1 A trainee who cannot *See*

The results obtained by the model for a trainee who *cannot See* are summarised in Figure 5.4. In Figure 5.4, the horizontal axis presents the suggested percentage of importance of modality values, ranging from 0 to 100. The vertical axis presents input and output modalities. The figure depicts the predicted percentage of modality values for four different learning styles, e.g. *visual*, *aural*, *kineasthetic* and *read/write*. A discussion on four different modalities and their predicted values against four learning styles are presented below.

The first modality chosen was the *screen* for a blind trainee. Figure 5.4 shows that the *screen* has a value of 0, for all the four different learning styles. This is realistic because the trainee has no utility for a screen.

The next modality considered was the *keyboard*. Figure 5.4 shows that the *keyboard* has a value of 68 for the *visual* trainee. A reason for this is that this trainee can visualise the *keyboard* and use this device. The modality *keyboard* has a value of 33 for both the *aural* and the *kineasthetic* learning styles. This means that the *kineasthetic* and *aural* trainees prefer other devices (e.g. *audio* for the *aural* trainee and *vibration* for *kineasthetic*

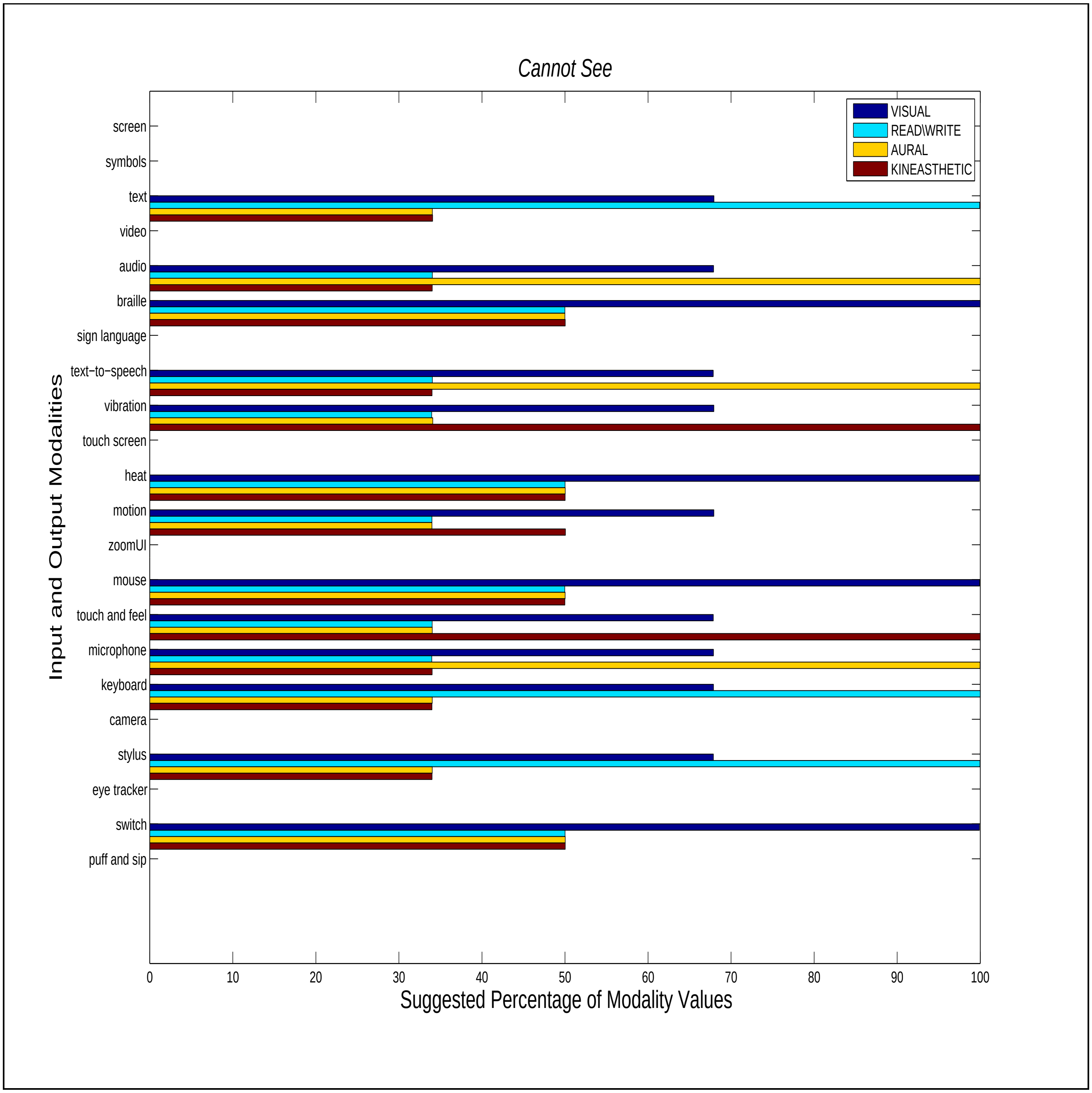

Figure 5.4: Importance values for the trainee who cannot *See*

trainee, see Figure 5.4) that are better suited to them instead of the *keyboard*. Finally, the *keyboard* has a value of 100 for the *read/write* trainee. This suggested modality value is a realistic one for the *read/write* trainee because the *keyboard* is a device that is generally preferred by a *read/write* trainee [1].

The third modality considered was the *audio* device. This modality has a value of 68 for the *visual* trainee. Although the profile of this trainee is *blind*, the trainee can have a *visual* preference. This can be due to the fact that the content is being described in *visual* terms i.e. using descriptive adjectives (the trainee is making use of his imagination). On the other hand, *audio* has an importance value of 33 for both the *read/write* and *kineasthetic* trainees. There are other modalities that are better suited for these trainees, e.g. *vibration* for the *kineasthetic* trainee and *keyboard* for the *read/write* trainee. The *audio* device has the value of 100 for the *aural* trainee. An *aural* trainee prefers using *audio* because naturally the trainee likes to listen and to speak.

The final modality considered was the *vibration* device. Figure 5.4 shows that the *vibration* device has an importance value of 68 against the *visual* trainee. This is a high value because, although the trainee is *blind*, he can feel the *vibration* and imagine accordingly. The *vibration* device has a value of 33 against both the *read/write* and *aural* preferences. These trainees prefer other modalities. The *read/write* trainee prefers *keyboard* and the *aural* trainee prefers *audio*. The *vibration* device has a value of 100 against the *kineasthetic* profile. A kineasthetic trainee prefers a device that requires movement. Hence, the values predicted by the non-linear model for different learning styles against *vibration* are realistic.

To summarize, Figure 5.4 shows that *screen*, *symbols*, *video*, *Sign Language*, *touch screen*, *eyetracker* and the *puff and sip* device have importance values of zero for the trainee who cannot *See*. This is due to the fact that the trainee cannot *See* the *screen*, *symbols*, *video*, *Sign Language* and he cannot use the *eyetracker*. It is noted that *text*, *audio*, *Braille*, *text-to-speech*, *vibration*, *heat*, *motion*, *mouse*, *touch and feel* device, *microphone*, *keyboard* and *stylus* are important modalities for a trainee who cannot *See*. These modalities make intuitive sense. For instance, a trainee who cannot *See* can use *text* that will feed into the *text-to-speech* device. *Audio* and *braille* are also important modalities as they have good values for all four learning styles.

The predicted modalities with the highest importance values corresponding to the various learning styles are presented below.

The *Braille* device is of high importance for the learning style *visual*. The *visual* trainee prefers a device that he can visualise. Similarly, *text*, *keyboard* and *stylus* are of importance for the learning style *read/write*. The *read/write* trainee prefers the *keyboard* for writing (typing). The *audio* device, *text-to-speech* and *microphone* are important for the learning style *aural*. The *aural* trainee prefers *audio*. The *vibration*, and *touch and feel* device are of high importance for the learning style *kineasthetic*. The *kineasthetic* trainee prefers doing things, thus the *touch and feel* device is preferred. Clearly, learning styles affect the importance values of various modalities. The modalities values suggested by the non-linear model for *blind* trainee are however realistic.

5.2.2 A trainee who cannot *Hear*

The non-linear model has been tested using this ability and the results of this study are summarised in Figure 5.5. Once again four modalities have been chosen and analysed using four learning styles (*visual*, *aural*, *read/write*, *kineasthetic*).

In this instance the first modality chosen was the *audio* device. Figure 5.5 shows that *audio* has a value of zero against all learning styles. This is realistic as the *audio* is replaced with other modalities.

The second modality considered was the *screen*. This modality has a value of 100 against the learning style *visual*. This is because a *visual* trainee prefers to see things and visualise accordingly. The modality *screen* has an importance value of 33 against the remaining three learning styles. This indicates that *screen* is preferred less by these trainees. A *read/write* trainee would still use this device (*screen*), for example, view *text* on the screen. An *aural* trainee can still have this device on to look at *Sign Language*. Similarly, a *kineasthetic* trainee can have this device on to look at *Sign Language*. Hence, the predicted values for the learning styles are realistic.

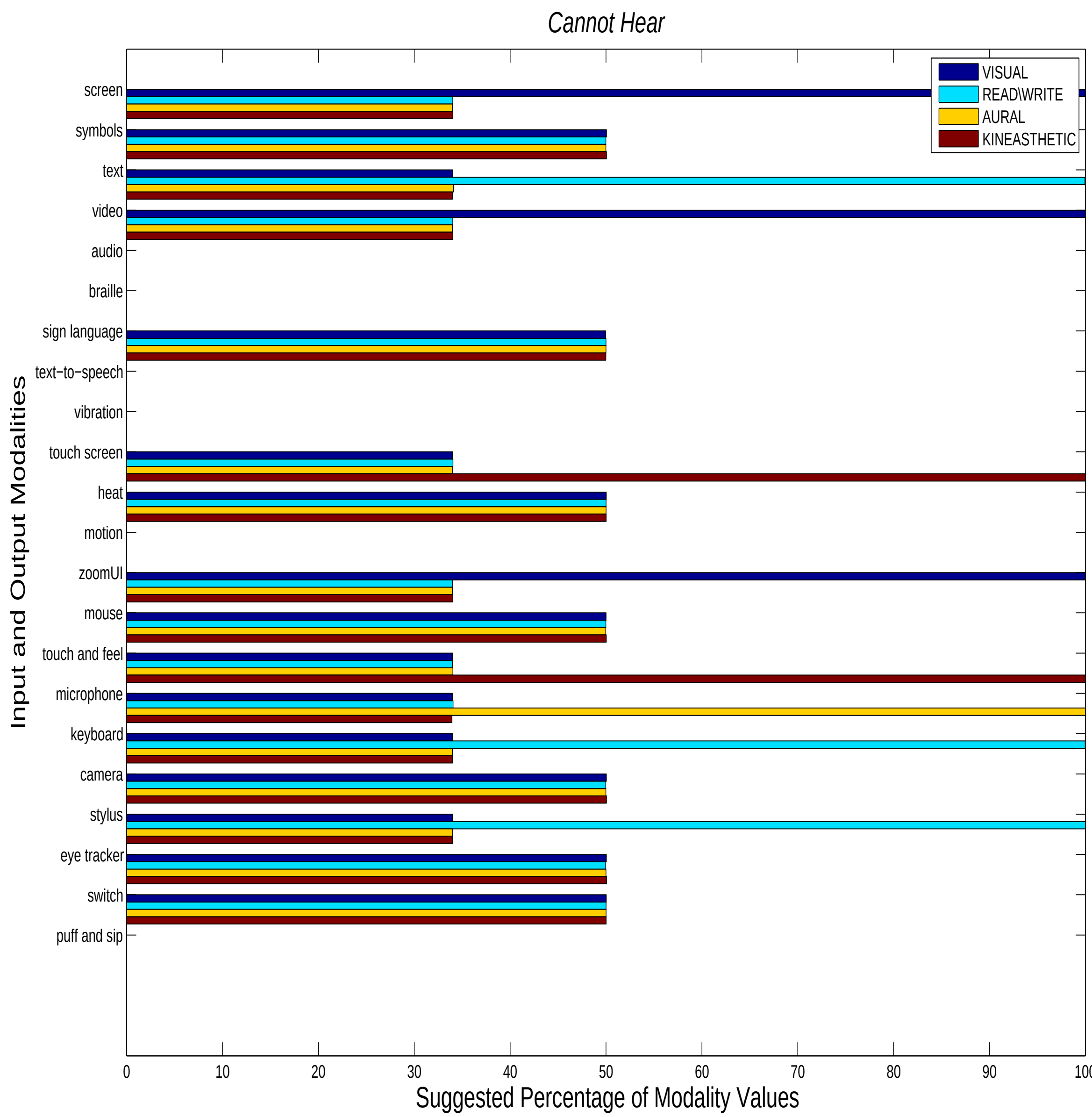

Figure 5.5: Importance values for the trainee who cannot *Hear*

The third modality chosen was the *touch screen* device. This modality has an importance value of 100 against the *kineasthetic* trainee. As mentioned earlier, a *kineasthetic* trainee likes movement and touching. The modality *touch screen* device has a value of 33 against the remaining three learning styles. This shows that *touch screen* is of less importance to these trainees. For example, a *visual* trainee, who has the *screen* device switched on, may like to use the *touch screen*. The *read/write* trainee may like a virtual keyboard to be presented by the *touch screen*. Finally, the *aural* trainee may consider the *touch screen* as a response mechanism to the input modality *microphone*. For example, the trainee can talk with the *microphone* and use the modality *touch screen* to communicate with the system. The predicted modalities are therefore acceptable.

The final modality considered was the *microphone*. This modality has a value of 100 for the *aural* trainee. This is quite logical because this trainee likes to talk although he *cannot Hear*. The modality *microphone* has a value of 33 against the remaining three learning styles. For example, the *visual* trainee can use this input modality to speak and get a reply on the output modality *screen*. Similarly, the *read/write* trainee can use this device and get a reply via *text*. A *kineasthetic* trainee can also use the device and get a reply via the *touch screen*. Hence all predicted values are realistic.

To summarize, Figure 5.5 shows that *audio*, *Braille*, *text to speech*, *vibration*s, *motion* and the *puff and sip* device all have the importance value of zero. This is due to the fact that the trainee *cannot Hear* the *audio* and he *cannot Hear* the information from the *text to speech* device. *Braille* is of no use, because this trainee, who is not blind, has never learnt to use the *braille* device. It can be seen that the *screen*, *symbols*, *text*, *Sign Language*, *heat*, *mouse*, *touch and feel* device, *microphone*, *keyboard*, *camera*, *stylus*, *eyetracker* and *switch* are important modalities for a *hearing* impaired trainee profile. These modalities make intuitive sense. A *hearing* impaired trainee will use the *screen* with the *keyboard* and *mouse*. The trainee can also use the *microphone* to communicate with and receive replies with the aid of *video* and/or *Sign Language*.

The predicted modalities with the highest importance values corresponding to the various learning styles are presented next.

The modalities *screen*, *video* and *zoomUi* have the highest importance value for the learning style *visual*. A *visual* trainee likes to see devices like the screen. Similarly, *text*, *keyboard* and *stylus* have the highest importance for the learning style *read/write*. This trainee likes to write on the *stylus*. The *microphone* has the highest importance value for the learning style *aural*. Since the trainee is *aural*, he likes to talk on the *microphone*. The *vibration* and *touch screen* have the highest importance value for the learning style *kineasthetic*. The trainee likes interaction, thus *vibration* is a device preferred by the *kineasthetic* trainee.

The modalities suggested by our non-linear model for a trainee who *cannot Hear* are realistic justifying the robustness of the non-linear model.

5.2.3 A trainee who is *fully-able*

Next, the non-linear model is tested with a *Fully-able* trainee profile where the trainee understands *Sign Language*.

The results of this study are summarised in Figure 5.6. Once again the relationship of four modalities with four learning styles (*visual*, *aural*, *read/write*, *kineasthetic*) is described.

The first modality considered was the *Braille* device. This device has a value of 0 for all the different learning styles, as shown in Figure 5.6. A reason for this is that the trainee can *See* and does not require the *braille* device. Another reason could be that the trainee does not know how to read the *braille* device.

The second modality, *video*, has an importance value of 100 for the *visual* trainee. Again a *visual* trainee likes to see the content. The modality *video* has an importance value of 33 for the remaining learning styles. Even though this value is low, these trainees can use *video*. The *read/write* trainee can use video where the lecture (content) is displayed in video. An *aural* trainee may like to listen to the *audio* of the *video*. Finally, the *kineasthetic* trainee will prefer watching interactive content in video.

The third modality considered was the *mouse* device. This modality has an importance value of 50 for all the different learning styles. With regard to a *visual* trainee, the *mouse*

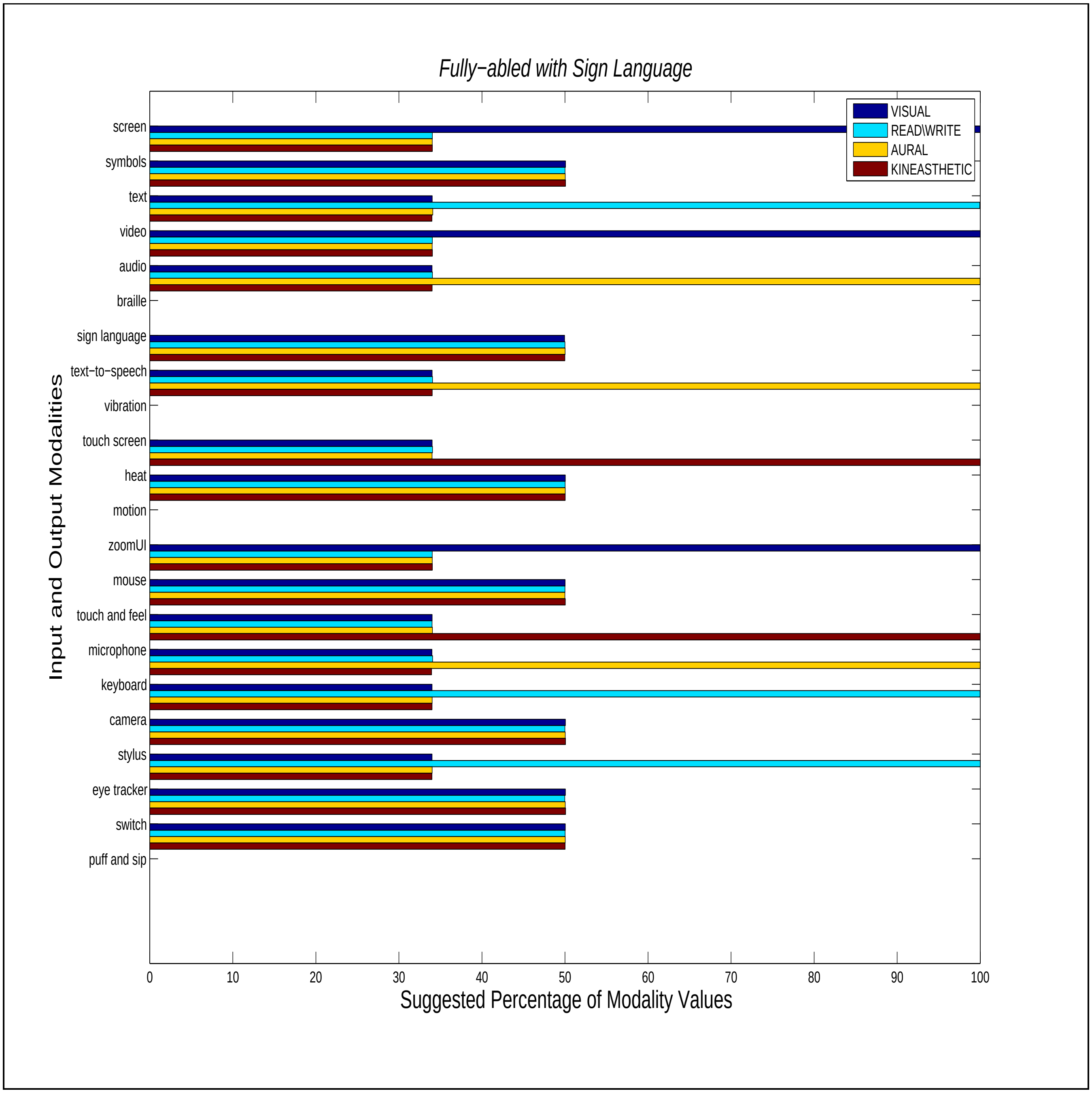

Figure 5.6: Importance values for a *Fully-able* trainee

can be used in combination with the *screen* by clicking an icon. Meanwhile, a *read/write* trainee may use this device to highlight *text*, while an *aural* trainee may use this device to click on a music icon. Finally a *kineasthetic* trainee may use this device to scroll between the content presented. Therefore, the predicted importance values for this modality are reasonable.

The final modality considered for a *Fully-able* trainee was the *keyboard*. This modality has an importance value of 33 for the *visual* trainee. This trainee can use the *keyboard* to enter commands, to view content on the *screen*. The modality *keyboard* has a value of 100 for a *read/write* trainee. This is realistic because a *keyboard* is used for typing. The value of the *keyboard* is 33 for *aural* and *kineasthetic* trainees. The *aural* trainee can use the *keyboard* to highlight *text* and feed it into the *text-to-speech* device. A *kineasthetic* trainee can use the *keyboard* as an interactive device.

To summarize, Figure 5.6 shows that the *Braille*, *vibration*, *motion* and *puff and sip* devices all have the importance value of zero. This is due to the fact that the trainee does not need these devices, being a *Fully-able* trainee. It is noted that the trainee cannot use *braille* as he has no knowledge of it. This trainee can use all other devices. The predicted modalities with the highest importance values corresponding to the various learning styles are now presented.

Screen, *video* and *zoomUi* have the highest importance value for the learning style *visual*. This makes sense because a *visual* trainee prefers seeing things. Similarly, *text*, *keyboard* and the *stylus* device all have the highest importance value for the learning style *read/write*. This is true because a *read/write* trainee prefers reading and writing. *Audio*, *text to speech* and the *microphone* have the highest importance value for the learning style *aural*. An *aural* trainee prefers to listen and speak. Finally, the *touch screen* and *touch and feel* devices have the highest importance value with regard to the learning style *kineasthetic*.

The above analysis using the summarised results from Figure 5.6 shows that the non-linear model is realistic as it never predicted unrealistic configuration values.

5.2.4 A trainee who cannot *Hear* but cannot understand *Sign Language*

In the previous subsections the non-linear model was tested on a number of trainee profiles. Here we test the model with a trainee who has recently become disabled, i.e. a trainee who cannot *Hear* and cannot understand *Sign Language*. The results of this study are summarised in Figure 5.7.

Here four different modalities and their predicted values are analyzed against the four learning styles (*visual*, *aural*, *read/write*, *kineasthetic*).

The first modality considered was the *audio* device. This modality has an importance value of 0 for all different learning styles. This is realistic because the trainee *cannot Hear*.

The second modality considered was *Sign Language*. This modality, *Sign Language* has a value of 0 for all the different learning styles, as shown in Figure 5.7. This is a realistic view point because the trainee cannot understand *Sign Language*.

The third modality considered was *text*. This modality has a value of 100 for the *read/write* trainee. This is because this trainee likes to read. *Text* has a value of 33 for the remaining learning styles. A *visual* trainee likes to read the content that is presented in a *visual* manner, while an *aural* trainee may like to read *text* if no other modalities are available. Finally a *kineasthetic* trainee may like to read *text* interactively (e.g. material that offers the reader different options).

Finally, the fourth modality considered was the *screen*. This modality has an importance value of 100 for a *visual* trainee. This is because a *visual* trainee prefers to see things and visualise this accordingly. The *screen* device has a value of 33 for the remaining three learning styles. This indicates that *screen* is preferred less by this trainee. A *read/write* trainee would still have this device on to view *text* for example. An *aural* trainee can still have this device on to look at *Sign Language*. A *kineasthetic* trainee can have this device on so that he can make use of the *touch screen* device.

Notice that the results presented in Figure 5.5, Section 5.2, and Figure 5.7 are gener-

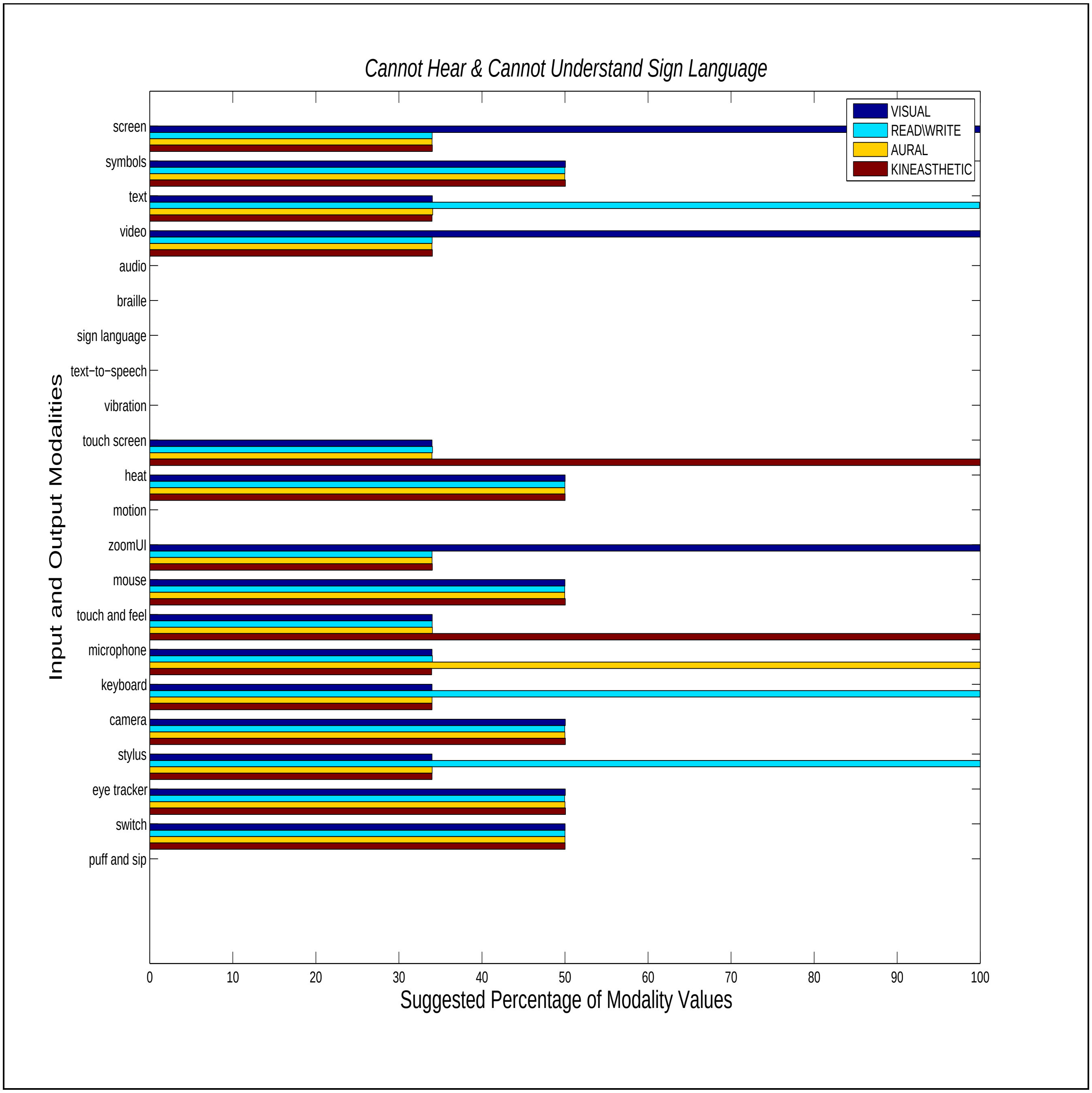

Figure 5.7: Importance values for a trainee who cannot *Hear* and cannot understand *Sign Language*

ated, respectively, for the profile cannot *Hear*, and for the profile cannot *Hear* and cannot understand *Sign Language*.

To summarize, Figure 5.7 shows that *screen*, *symbols*, *text*, *heat*, *mouse*, *touch and feel* device, *microphone*, *keyboard*, *camera*, *stylus*, *eyetracker* and *switch* are important modalities for a *hearing* impaired trainee profile. *Sign Language* has zero importance, this make intuitive sense, due to the fact that the trainee cannot understand *Sign Language*. Hence, a comparison between Figure 5.5 and Figure 5.7 shows that except *Sign Language* (which has the importance value of zero in Figure 5.7) the other predicted modalities and their values are the same. The model has correctly suggested modalities, thus the model is realistic.

The results presented in this section indicate the value of the non-linear model for various trainee profiles in a state context (see the states energetic, bored, lazy, tired in the next section). The next section presents results for the non-linear model where the trainee profile does not remain static, i.e. feedback from the trainee is simulated.

5.3 Numerical results for the non-linear learning model with feedback

In the previous section, reliable predictions by the non-linear model have been presented. However, time dependency in the preference of the modalities have not been considered. During the trainee's training period, the suggested modalities may become influenced by time passing. This characteristic can be due to the fact that the trainee gets *bored*, becomes *lazy* or gets *tired*. This can be modeled by providing feedback to the learning model. Figure 5.8 depicts this non-linear feedback mechanism. When a user is suggested a set of modalities, the user may suggest an H value. Thus let H be the happiness of a particular user. The H value is then sent to the trigger mechanism, which evaluates the users performance and places him in a particular state. Four different states are introduced. These are:

- *energetic*,
- *bored*,
- *lazy*, and
- *tired*.

Each state has a threshold, i.e. (0,25], (25,50], (50,75] and (75,100] for *tired*, *lazy*, *bored* and *energetic* respectively. For example, if $H = 70$, then the state *bored* is fed into the neural network. This state brings about a change in the modalities suggested to the user.

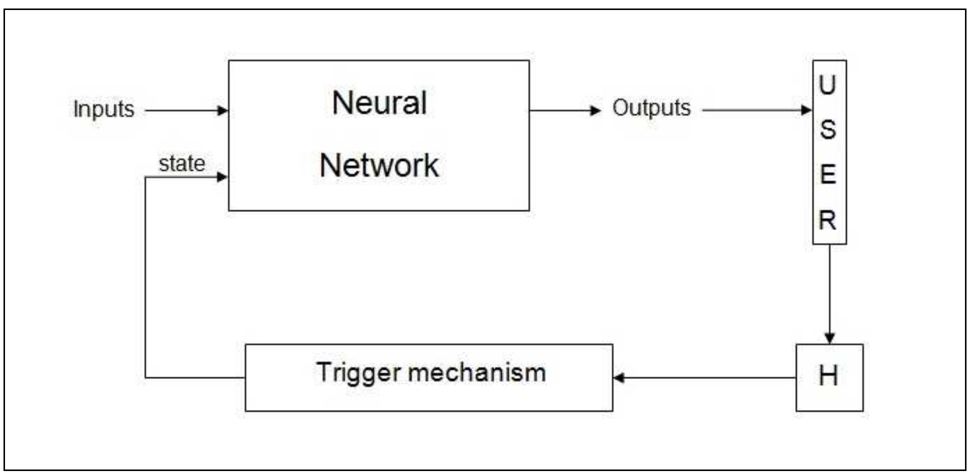

Figure 5.8: Non-linear model with feedback

The results summarised in Figures $5.4-5.7$ were based on the *energetic* trainee. Next numerical studies are carried out to see the effects of the remaining states (*bored*, *lazy* and *tired*) on the results presented in Figures $5.9 - 5.20$.

5.3.1 The *bored* trainee

A *bored* trainee who cannot *See*

The non-linear model with feedback mechanism has been tested and the results obtained for a *visually* impaired *bored* trainee are summarised in Figure 5.9. In Figure 5.9, the horizontal axis presents the suggested importance of modalities values, while the vertical axis presents input and output modalities, similar to Section 5.1. Discussions on three different modalities and their predicted values against four learning styles (*visual*, *aural*, *read/write*, *kineasthetic*) for the *bored* trainee who *cannot See* are presented below.

The first modality considered was the *screen* device. It is important to note that Figure 5.4, see page 47, contains the summarised results obtained by the non-linear model for the trainee 'who *cannot See*'. On the other hand, Figure 5.9 presents the summarised results obtained by the non-linear model with feedback mechanism for the same trainee. The only difference is that the state *bored* is now activated within the non-linear model. As in Figure 5.4, this modality has a value of 0 for all the different learning styles, see Figure 5.9. This is again due to the fact that the trainee cannot *See*.

The second modality considered was the *audio* device. Figure 5.9 shows that *audio* has an importance value of 74 for an *aural* trainee. The modality *audio* has an importance value of 0 for the three remaining learning styles. A comparison with Figure 5.4 shows that these importance values are different in Figure 5.9. In Figure 5.4 the modality *audio* has an importance value of 100. Thus different states bring about a change in the modalities suggested to the trainee, indicating the clear effects of the state *bored*.

Finally the modality *Braille* was considered. This modality has a value of 100 for a *visual* trainee. This is exactly the same value predicted by the non-linear model without feedback in Figure 5.4. Clearly, the integration of the state *bored* in the model does not affect the trainee's choice of *braille*. This is simply because the trainee *cannot See*. The modality *braille* has a value of 28 for a *read/write* trainee. The *braille* device has an importance value of 27 for an *aural* trainee. These values are however different in

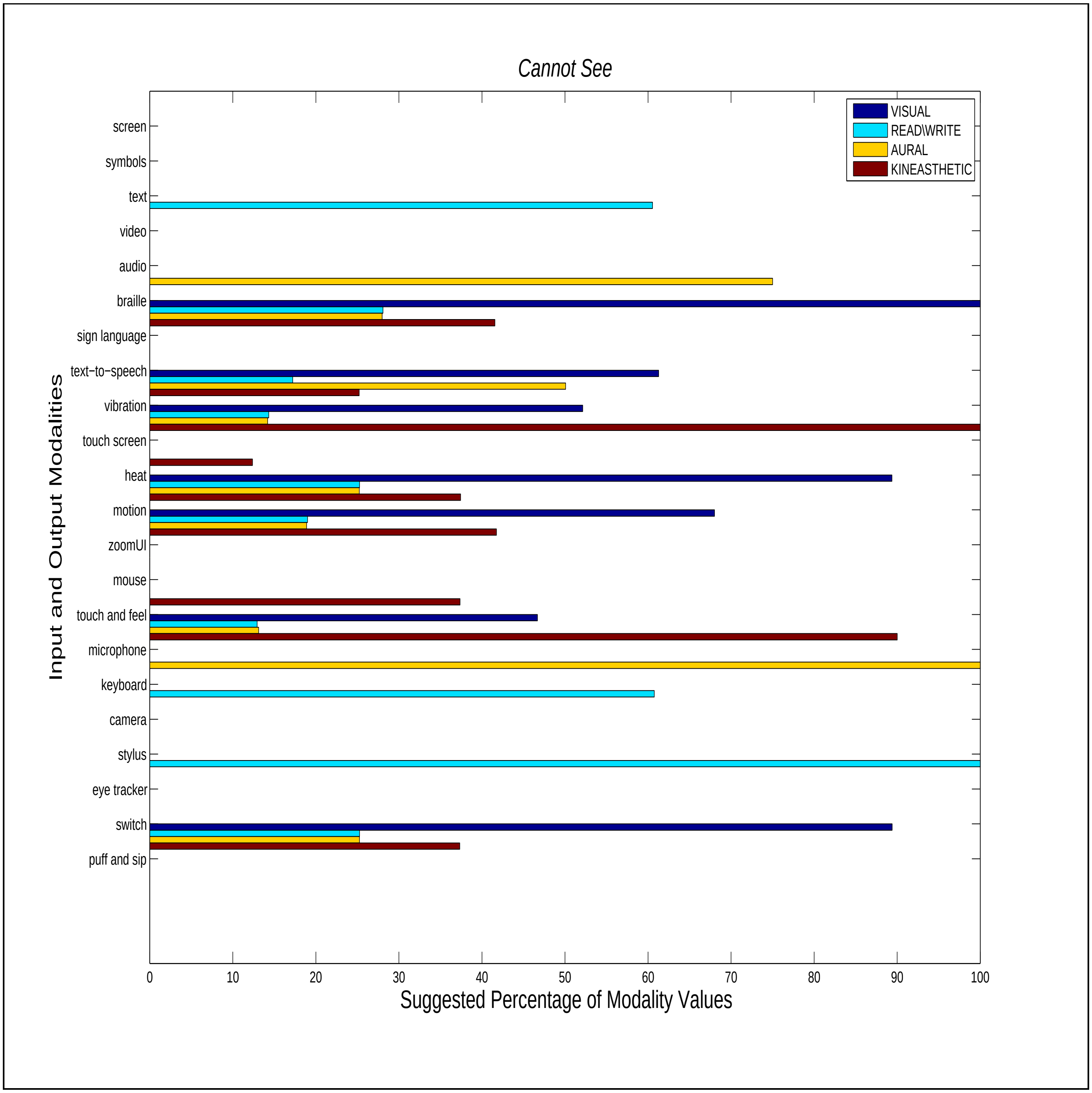

Figure 5.9: Importance values for a *bored* trainee who cannot *See*

Figure 5.4, indicating the effect of the change of trainees state.

To summarize, Figure 5.9 shows that *screen*, *symbols*, *video*, *Sign Language*, *touch* *screen*, *eyetracker* and the *puff and sip* device have zero importance value for the *visually* impaired *bored* trainee. The values of these modalities are exactly as predicted by the non-linear model without feedback, as presented in Figure 5.4.

The values of *microphone* in Figure 5.9 are zeroes for learning styles *visual*, *read/write* and *kineasthetic*. However, these values are non-zero in Figure 5.4. This is due to the trainee's change of state, e.g. from *energetic* to *bored*. When a trainee becomes *bored* the model then considers the trainee's learning style to suggest the modalities. For example, an *aural* trainee prefers the *microphone*, while a trainee with the remaining learning styles does not prefer this modality.

Upon further analysis it is noted that *text*, *audio*, *Braille*, *text-to-speech*, *vibration*, *heat*, *motion*, *mouse*, *touch and feel* device, *microphone*, *keyboard* and *stylus* are important modalities for a *visually* impaired *bored* trainee profile. These modalities make intuitive sense. The values of these modalities however different in Figure 5.4 and Figure 5.9. This again, clearly shows the effect of the trainee's change of state.

A *bored* trainee who cannot *Hear*

The non-linear model with feedback mechanism was tested using this ability and the results of this study are summarised in Figure 5.10. Once again, three different modalities and their predicted values are analyzed against the four learning styles (*visual*, *aural*, *read/write*, *kineasthetic*).

The first modality chosen was the *audio* device. Results presented in Figure 5.5, see page 50, are obtained by implementing the non-linear model without feedback. The only difference between the results in Figure 5.5 and Figure 5.10 is that the trainee state *bored* is now activated in producing the results in Figure 5.10. Figure 5.10 shows that the *audio* device has an importance value of 0 for all the four different learning styles. This is because the trainee *cannot Hear*.

The second modality, the *screen* device, has an importance value of 100 against a *visual* trainee. The *screen* device has an importance value of 0 for all the remaining learning

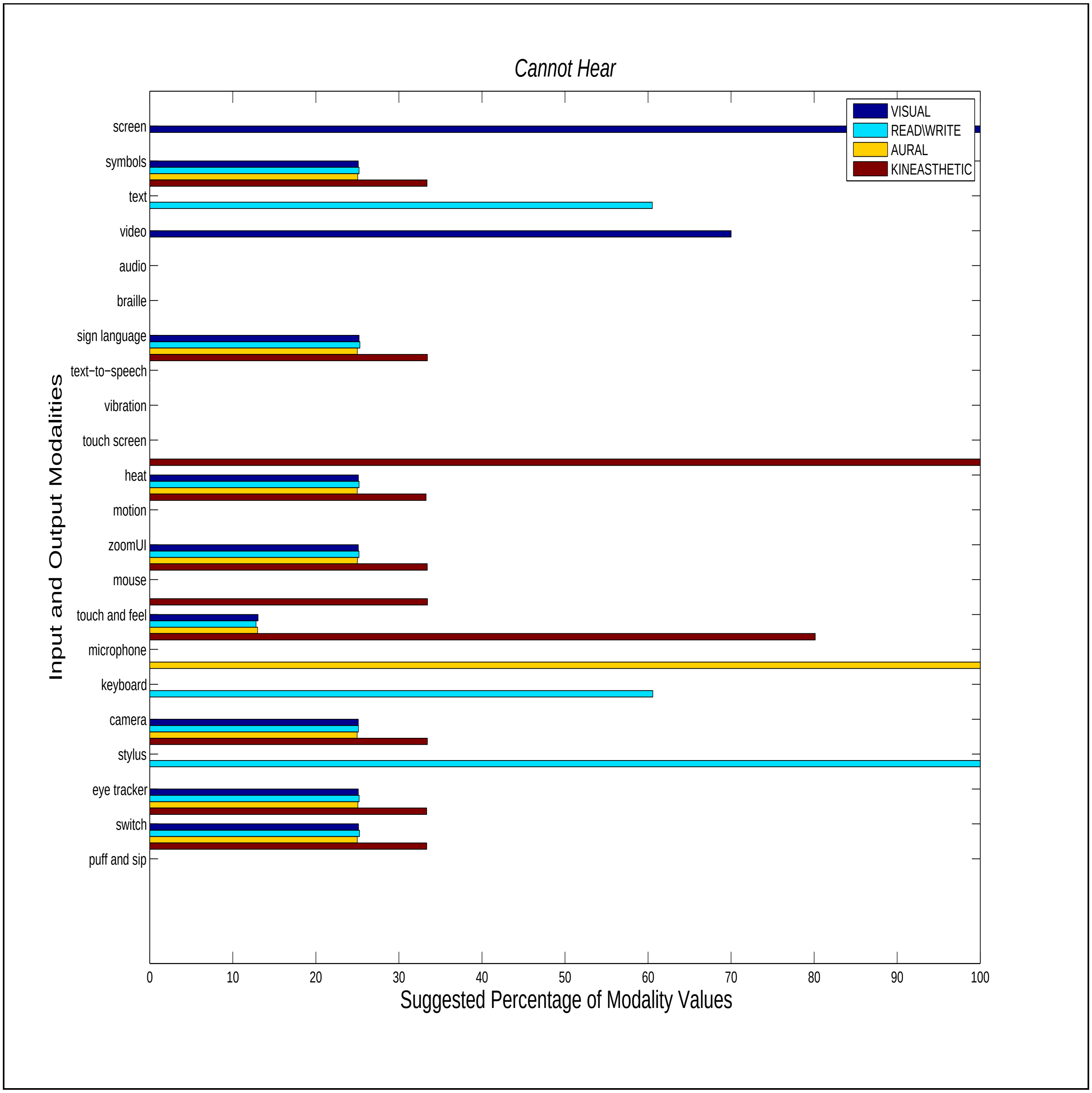

Figure 5.10: Importance values for a *bored* trainee who cannot *Hear*

styles. A comparison with Figure 5.5 shows that these importance values are different in Figure 5.10, indicating the clear effects of the state *bored*.

Finally, the modality *Sign Language* was considered. This modality has a value of 36 against the *kineasthetic* trainee. This is exactly the same value predicted by the non-linear model without feedback in Figure 5.5. Clearly, the integration of the state *bored* in the model does not affect the trainee's choice of *Sign Language*. It is simply because the trainee *cannot Hear*. *Sign Language* has the importance value of 26 for the other learning styles.

To summarize, Figure 5.10 shows that *audio*, *Braille*, *text to speech*, *vibration*s, *motion* and the *puff and sip* device have zero importance value for the *bored* trainee. The values of these modalities are exactly as predicted by the non-linear model without feedback, see Figure 5.5. The values of *screen* in Figure 5.10 are zeroes for learning styles *aural*, *read/write* and *kineasthetic*. However, these values are non-zero in Figure 5.5. This is due to the fact that different states bring about different suggestions of modalities. In Figure 5.5, *screen* and *video* were both suggested to the *visual* trainee. In Figure 5.10 *screen* is suggested first to the trainee. This is due to the trainee's state *bored*, where the trainee may prefer the *screen* compared to video. Therefore, the non-linear learning model has suggested a more accurate set of modalities in Figure 5.10.

It is noted that the *screen*, *symbols*, *text*, *video*, *Sign Language*, *heat*, *mouse*, *touch and feel* device, *microphone*, *keyboard*, *camera*, *stylus*, *eyetracker* and *switch* are important modalities for a *hearing* impaired trainee profile. These modalities make intuitive sense. The values of these modalities however different in Figure 5.5 and Figure 5.10. This clearly shows the effect of the trainee's change of state.

A *bored* trainee who is *Fully-able*

The effect of our model on a *bored* trainee who is *Fully-able* and understands *Sign Language* is now considered. The results of this study are summarised in Figure 5.11. Once again, the relationship of three different modalities with four learning styles (*visual*, *aural*, *read/write*, *kineasthetic*) has been described.

The first modality chosen was the *Braille* device. It is important to note that Fig-

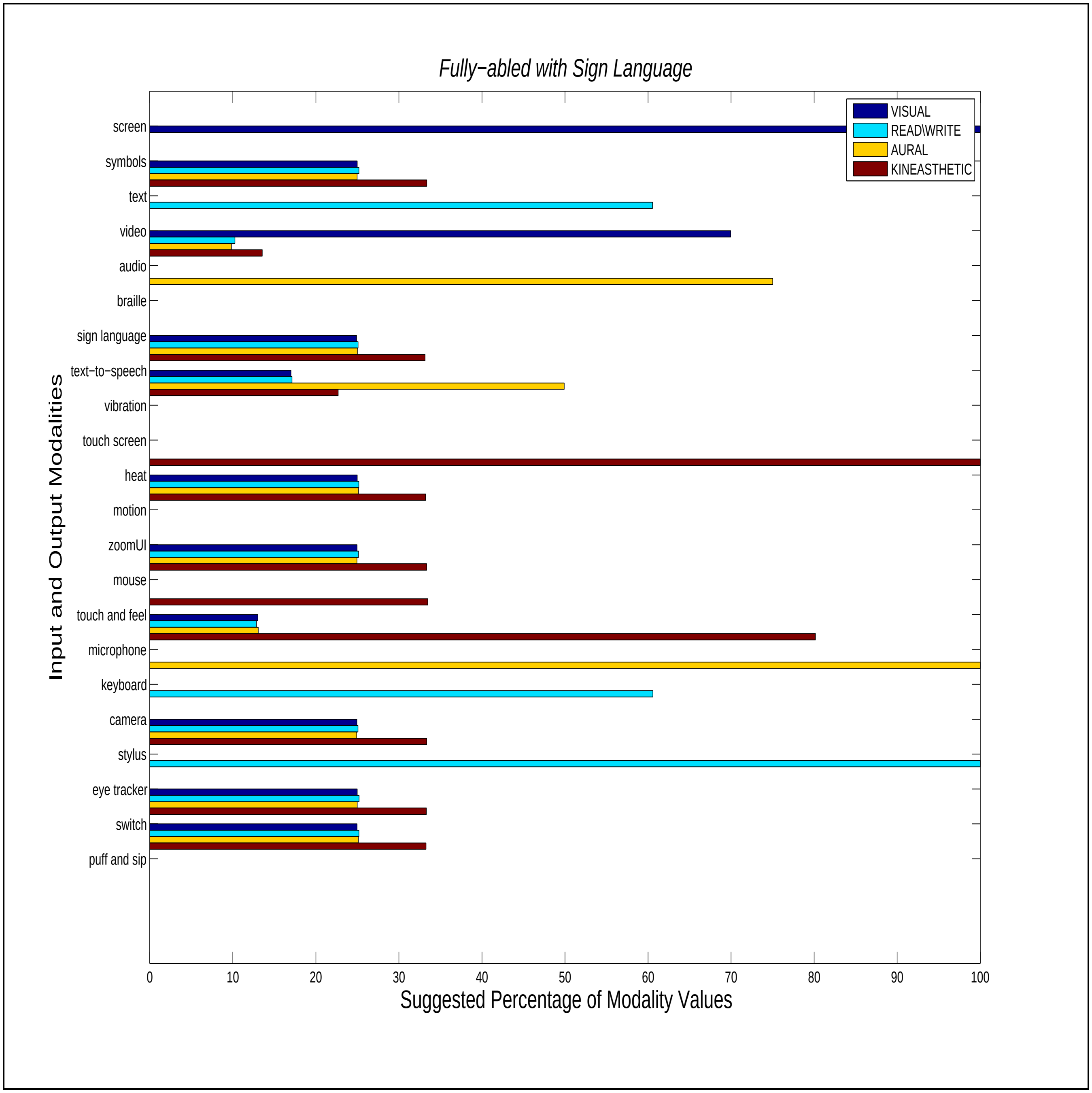

Figure 5.11: Importance values for a *bored Fully-able* trainee

ure 5.6, see page 53, contains the summarised results obtained by the non-linear model for the trainee who cannot use the *braille* device. On the other hand, Figure 5.11 presents the summarised results obtained by the non-linear model with feedback mechanism for the same trainee. The only difference is that the state *bored* is now activated within the non-linear model with feedback. Figure 5.6 and Figure 5.11 predicted the same importance value of zero for *braille* for the *Fully-able* trainee. As before, this is due to the fact that the trainee cannot understand *braille*.

Next, the modality *text-to-speech* is chosen. This modality has a value of 18 for both the *visual* and *read/write* trainee. The *test-to-speech* device has a value of 52 for the *aural* trainee. The modality has a value of 23 for the *kineasthetic* trainee. A comparison with Figure 5.6 shows that these importance values are different, see Figure 5.11, indicating the clear effects of the state *bored*.

To summarize, Figure 5.11 shows that the trainee cannot use *Braille* as he has no knowledge of it. The trainee does not need *vibration*, *motion* and the *puff and sip* device, thus these values are zero. The values of these modalities are exactly as predicted by the non-linear model without feedback, see Figure 5.6. The values of *screen* in Figure 5.11 are zero for learning styles *aural*, *read/write* and *kineasthetic*. However, these values are non-zero in Figure 5.6. In Figure 5.6 *keyboard* and *stylus* were both suggested to the *read/write* trainee. In Figure 5.11, the *stylus* is suggested first to the trainee. When a trainee is *bored*, he may prefer to use the *stylus* to write on, instead of typing on a *keyboard*. Therefore, the non-linear learning model has suggested a more accurate set of modalities in Figure 5.11.

A *bored* trainee who cannot *Hear* but cannot understand *Sign Language*

The effect of the non-linear model on a *bored* trainee who cannot *Hear* and cannot understand *Sign Language* is now considered. The results of this study are summarised in Figure 5.12.

When comparing Figure 5.10, see page 62, and Figure 5.12, we note that the only difference lies in that this trainee cannot use *Sign Language*. The learning styles are identical to Figure 5.10, except for the fact that *Sign Language* has no importance to the trainee. A comparison between Figure 5.10 and Figure 5.12 shows the exclusion of *Sign Language* (which has the importance value of zero in Figure 5.12) the other predicted modalities and their values are the same (see Figure 5.10 and Figure 5.12).

Clearly, from the results presented on the *bored* trainee, one can infer that the model is realistic.

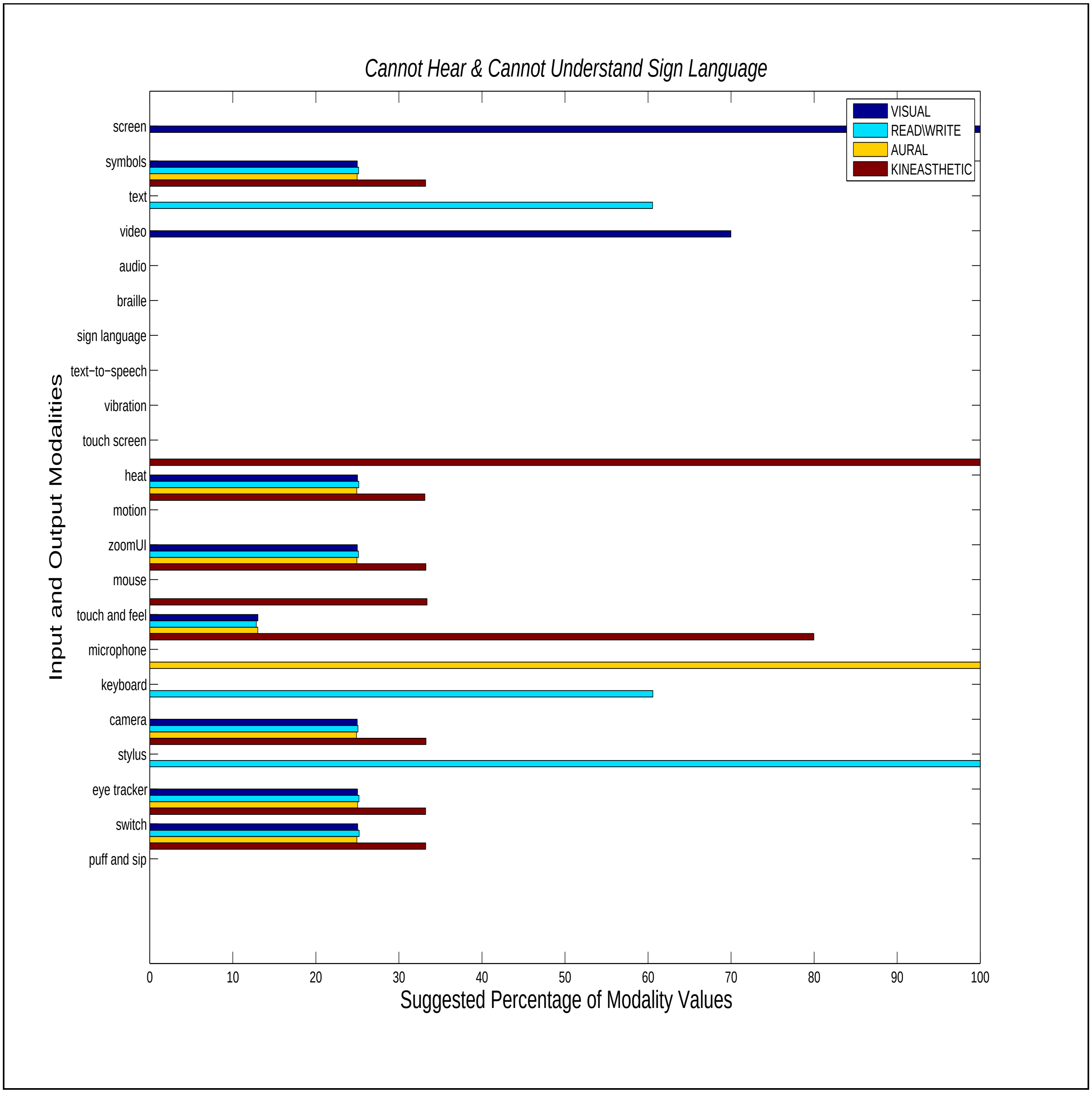

Figure 5.12: Importance values for a *bored* trainee that cannot *Hear* and cannot understand *Sign Language*

5.3.2 The *lazy* trainee

A *lazy* trainee who cannot *See*

The non-linear model with feedback mechanism has been tested and the results obtained for a *visually* impaired *lazy* trainee are summarised in Figure 5.13. A discussion about the three different modalities and their predicted values against four learning styles (*visual*, *aural*, *read/write*, *kineasthetic*) is presented.

The first modality chosen was the *screen* device. It is important to note that Figure 5.4, see page 47, contains the summarised results obtained by the non-linear model for the trainee 'who *cannot See*'. On the other hand, Figure 5.13 presents the summarised results obtained by the non-linear model with feedback mechanism for the same trainee but to include the state *lazy*. As in Figure 5.4, the *screen* has a value of 0 for all of the four different learning styles, shown in Figure 5.13. This is again due to the fact that the trainee cannot *See*.

Second, the modality *microphone* was chosen. Figure 5.13 shows that this device has a value of 100 for an *aural* trainee. The *microphone* device is at 0 for the remaining three learning styles. A comparison with Figure 5.4 shows that these importance values are different. In this case, Figure 5.4 also has a value of 100 for the *aural* trainee, but has non-zero values for the remaining learning styles, thus indicating the clear effects of the state *lazy*.

Finally, the modality *keyboard* was considered. This modality has a value of 100 for the *read/write* trainee. This is exactly the same value predicted by the non-linear model without feedback in Figure 5.4. Clearly, the integration of the state *lazy* in the model does not affect the trainee's choice of *keyboard*. The *keyboard* has a value of 0 for the remaining learning styles, as in Figure 5.13.

To summarize, Figure 5.13 shows that *screen*, *symbols*, *video*, the *Sign Language*, *touch screen*, *eyetracker* and the *puff and sip* device have zero importance values for the trainee who cannot *See*. These values are exactly as predicted by the non-linear model with

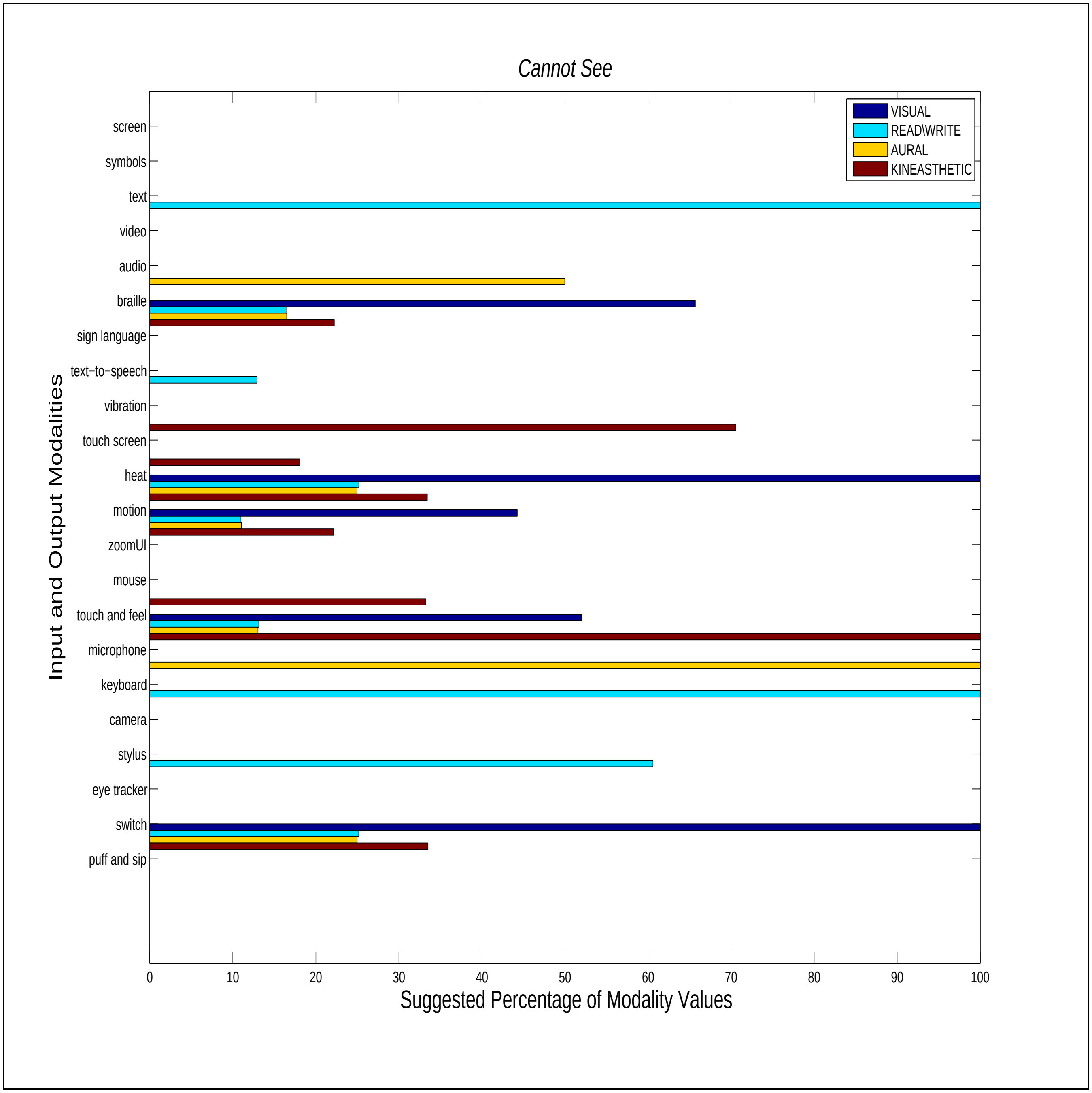

Figure 5.13: Importance values for a *lazy* trainee who cannot *See*

feedback, see Figure 5.4.

The values of *microphone* in Figure 5.13 are zeroes for learning styles *visual*, *read/write* and *kineasthetic*. However, these values are non-zero in Figure 5.4. This is due to the fact that different states bring about different suggestions of modalities. In Figure 5.4 *audio* and *microphone* were both suggested to the *aural* trainee. In Figure 5.13 *microphone* is suggested first to the trainee. This can happen when the trainee becomes *lazy* and may want to talk rather than listen. The non-linear learning model has suggested an accurate set of modalities in Figure 5.13.

Upon further analysis, it is noted from Figure 5.13 that *text*, *audio*, *Braille*, *text-to-speech*, *vibration*, *heat*, *motion*, *mouse*, *touch and feel* device, *microphone*, *keyboard* and *stylus* are important modalities for a *visually* impaired *lazy* trainee profile. These modalities make intuitive sense. The values of these modalities are however different in Figure 5.4 and Figure 5.13. This clearly shows the effect of the trainee's change of state.

A *lazy* trainee who cannot *Hear*

The non-linear model was tested using this profile and the results of this study are summarised in Figure 5.14. A discussion on three different modalities and their predicted values against four learning styles (*visual*, *aural*, *read/write*, *kineasthetic*) follow.

The first modality considered was the *audio* device. It is important to note that Figure 5.5, see page 50, contains the summarised results obtained by the non-linear model for the trainee 'who *cannot Hear*'. On the other hand, Figure 5.14 presents the summarised results obtained by the non-linear model with feedback mechanism (due to the state *lazy*) for the same trainee. This modality has a value of 0 for all the different learning styles. The same value was also suggested by the non-linear model, see Figure 5.5. This is true because the trainee *cannot Hear*.

Next, the modality *video* was considered. This modality has a value of 100 for the *visual* trainee. The modality *video* has a value of 0 for the remaining learning styles. A comparison with Figure 5.5 shows that these importance values are different in Figure 5.14. In Figure 5.5 the modality *video* has a value of 100 for the *visual* trainee, but has importance values of 33 for the remaining learning styles. This indicates the clear

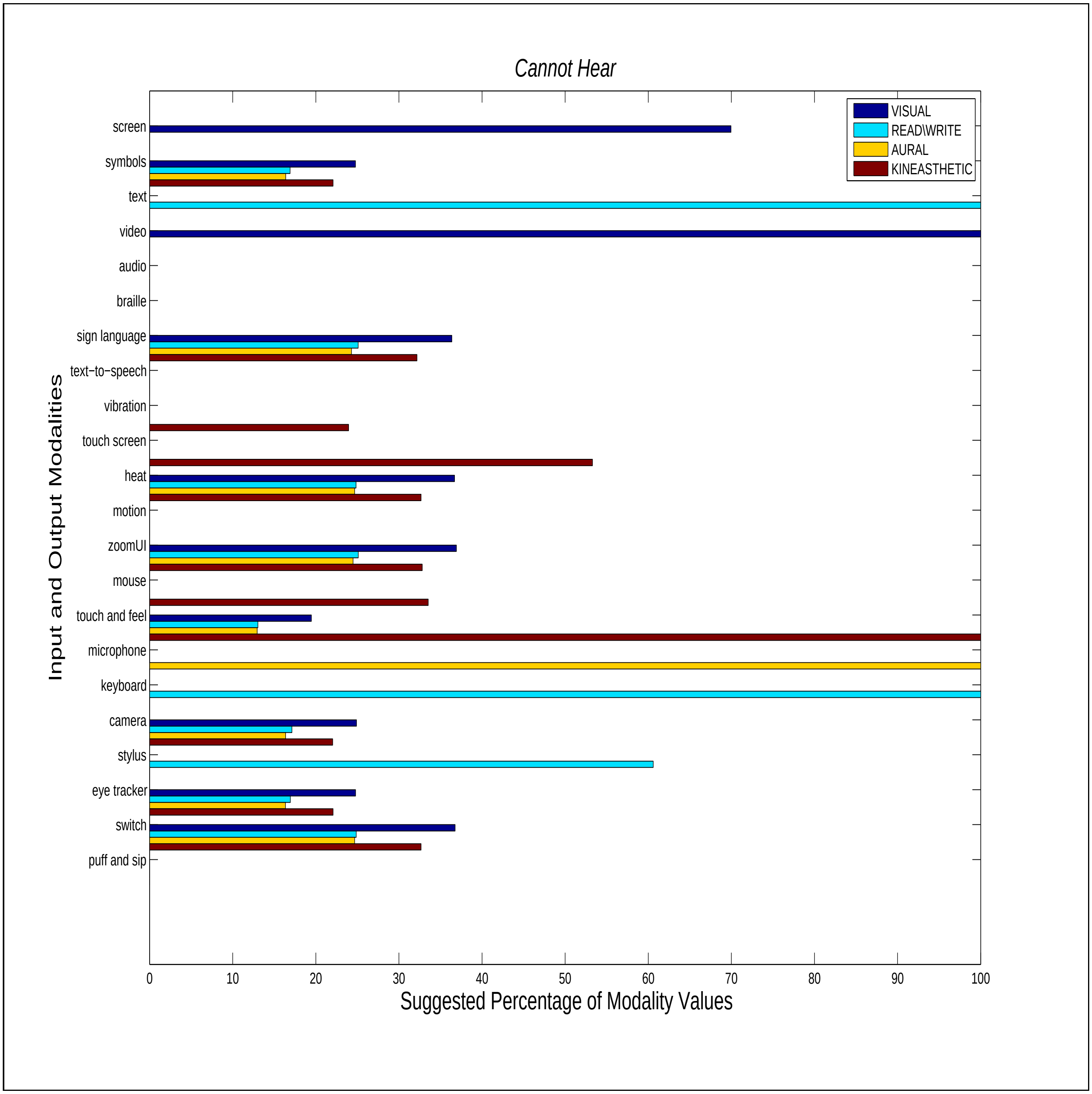

Figure 5.14: Importance values for a *lazy* trainee who cannot *Hear*

effects of the state *lazy*.

Finally, the modality *touch* and *feel* device was considered. This device has a value of 18 for the *visual* trainee, while this device has a value of 14 for the *read/write* and *aural* trainees. Finally, for the *touch* and *feel* device, has a value of 100 against the learning style *kineasthetic*. The values of these modalities are different to the values predicted by the non-linear model without feedback, see Figure 5.5, except for the learning style *kineasthetic* in which both have a value of 100. In Figure 5.5, it is noted that the remaining learning styles have values of 33. This shows a clinical difference.

To summarize, Figure 5.14 shows that *audio*, *Braille*, *text to speech*, *vibrations*, *motion* and the *puff and sip* device have zero importance value for the *lazy* trainee. The values of these modalities are however different in Figure 5.5 and Figure 5.14. This clearly shows the effect of the trainee's change of state. The values of *video* in Figure 5.14 are zero for learning styles *aural*, *read/write* and *kineasthetic*. However, these values are non-zero in Figure 5.5. This is due to the fact that different states bring about different suggestions of modalities. In Figure 5.5 *screen* and *video* were both suggested to the *visual* trainee. In Figure 5.14 *video* is suggested first to the trainee. When the trainee is feeling *lazy*, the video may be a preferred method of learning. The non-linear learning model has suggested a more accurate set of modalities in Figure 5.14.

It is clear from Figure 5.14 that the *screen*, *symbols*, *text*, *video*, the *Sign Language*, *heat*, *mouse*, *touch and feel* device, *microphone*, *keyboard*, *camera*, *stylus*, *eyetracker* and *switch* are important modalities for a *hearing* impaired trainee profile. These modalities make intuitive sense.

A *lazy* trainee who is *Fully-able*

The effect of our model on a *lazy* trainee who is *Fully-able* and can understand *Sign Language* is considered. The results of this study are summarised in Figure 5.15. The relationship of three different modalities with four learning styles (*visual*, *aural*, *read/write*, *kineasthetic*) has been described.

The first modality considered was the *screen* device. Figure 5.15 shows that the *screen* has a value of 70 for the *visual* trainee. These values are however different in Figure 5.6, see page 53, and Figure 5.15. This clearly shows the effect of the trainee's change of state. While this device has a predicted value of 7 for the *read/write* and *aural* trainees, it has a value of 9 predicted for a *kineasthetic* trainee.

Next, the modality *text* was considered. This modality has a value of 100 against the *read/write* trainee. This is exactly the same value predicted by the non-linear model without feedback in Figure 5.6. Clearly, the integration of the state *lazy* in the model does not affect the trainee's choice. The modality *text* has a value of 0 for the remaining learning styles. These values are however different in Figure 5.6 and Figure 5.15. This clearly shows the effect of the trainee's change of state.

Finally, the modality *microphone* was considered. Figure 5.15 shows that this modality has a value of 100 for an *aural* trainee. The *microphone* has a value of 0 for the remaining three learning styles. The value is the same for the *aural* learning style when comparing Figure 5.6 and Figure 5.15 but the values differ for the remaining learning styles. This clearly shows the effect of the trainee's change of state.

To summarize, Figure 5.15 shows that the trainee cannot use *Braille* as he has no knowledge of it. The modalities *vibration*, *motion* and the *puff and sip* device have zero importance value to the trainee. This is exactly the same value predicted by the non-linear model without feedback in Figure 5.6. Clearly, the integration of the state *lazy* in the model does not affect the trainee's choice of modalities. The values of *video* in Figure 5.15 are zeroes for learning styles *aural*, *read/write* and *kineasthetic*. However, these values are non-zero in Figure 5.6. This is due to the fact that different states bring about different suggestions of modalities. A different set of modalities are suggested to the trainee depending on the state the trainee is in. Thus each state gives a new suggestion that may aid the trainee. In Figure 5.6, *screen*, *video* and *zoomUi* were suggested to the *visual* trainee. On the other hand, in Figure 5.15, *video* is suggested first to the trainee. This is due to the trainee's state *lazy*, where the trainee may prefer to watch *video*.

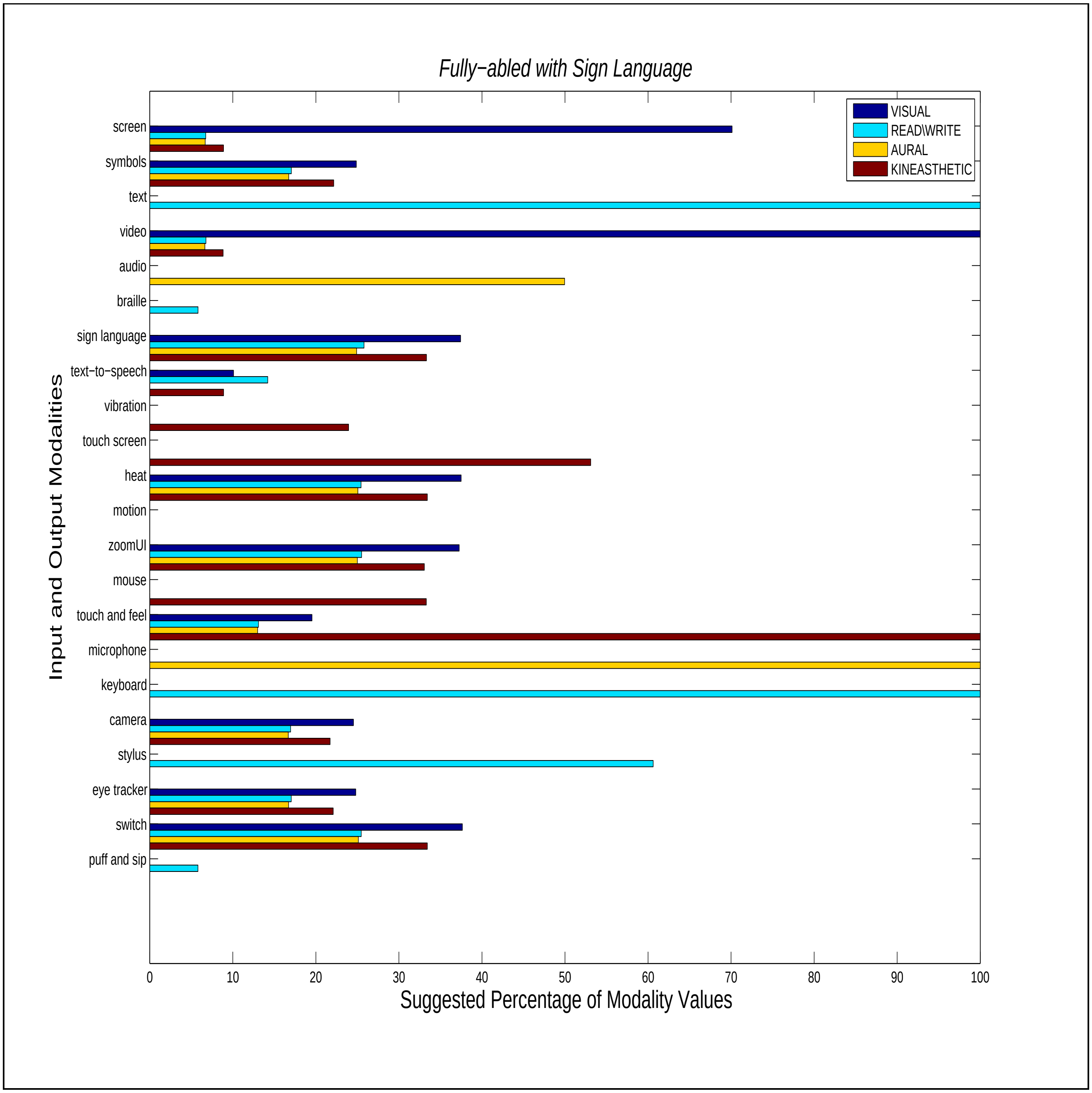

Figure 5.15: Importance values for a *lazy Fully-able* trainee

Therefore, the non-linear learning model has suggested a more accurate set of modalities in Figure 5.15. This clearly shows the effects of the state *lazy*.

A *lazy* trainee who cannot *Hear* but cannot understand *Sign Language*

The effect of the model on a *lazy* trainee who cannot *Hear* and cannot understand sign is now presented. The results of this study are summarised in Figure 5.16.

By inspecting Figure 5.16 it is noted that it is similar to Figure 5.14, see page 71, as mentioned earlier (except *Sign Language* is excluded in Figure 5.16). The modality *Sign Language* is disabled (zero) in Figure 5.16, because the trainee cannot understand *Sign Language*.

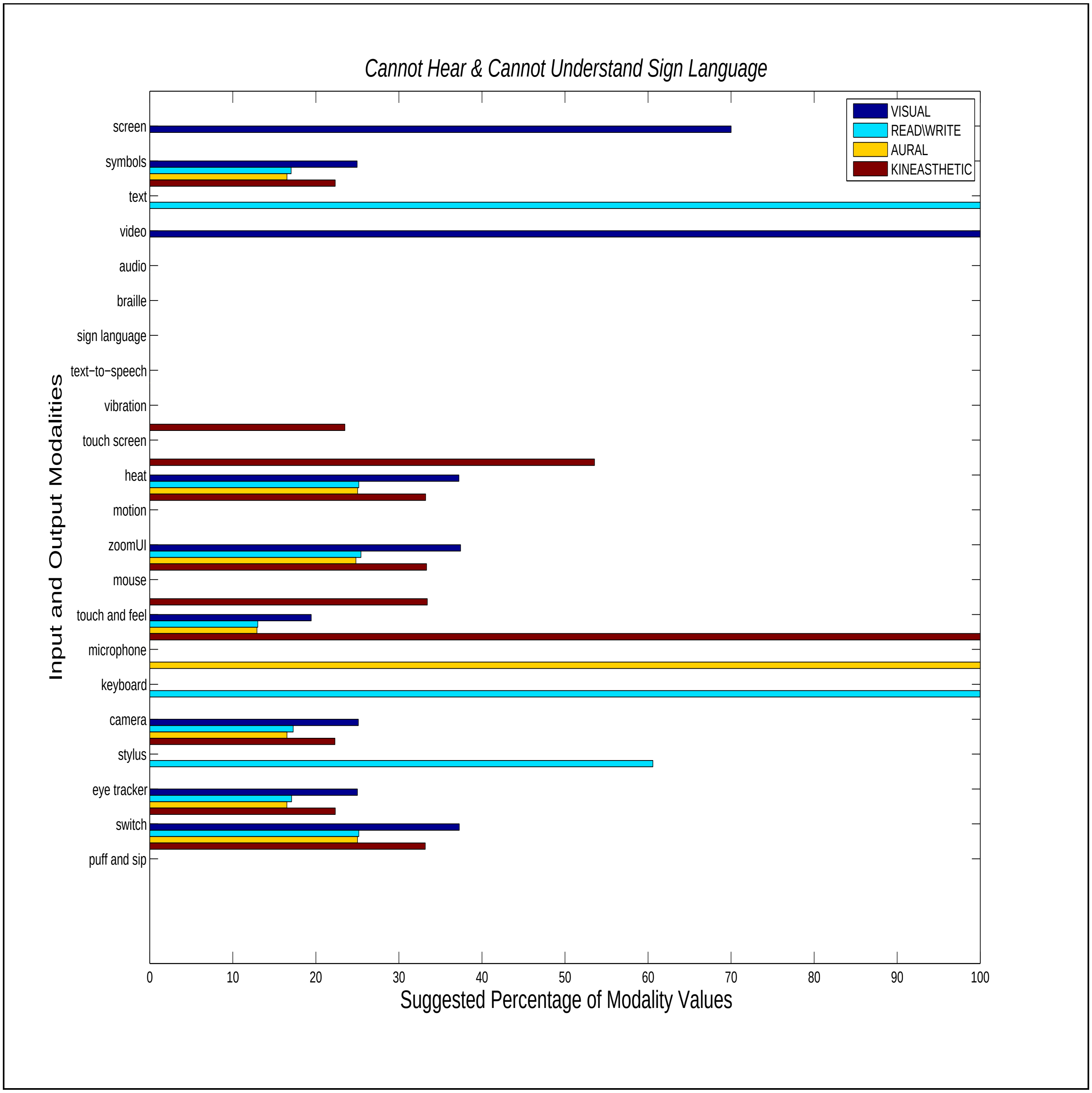

Figure 5.16: Importance values for a *lazy* trainee who cannot *Hear* and cannot understand *Sign Language*

5.3.3 The *tired* trainee

A *tired* trainee who cannot *See*

The non-linear model with feedback mechanism has been tested and the results obtained for a *visually* impaired *tired* trainee are summarised in Figure 5.17. A discussion on two different modalities and their predicted values against four learning styles (*visual*, *aural*, *read/write*, *kineasthetic*) are presented below.

The first modality considered was the *screen* device. It is important to note that Figure 5.4, see page 47, contains the summarised results obtained by the non-linear model for the trainee 'who *cannot See*'. On the other hand, Figure 5.17 presents the summarised results obtained by the non-linear model with feedback mechanism for the same trainee. The only difference is that the state *tired* is now activated within the non-linear model with feedback. As in other trainee who cannot *See*, Figure 5.17 shows that the screen has a predicted value of 0 for all the learning styles. The trainee cannot *See* and thus does not need the device.

Finally, the modality *text-to-speech* was considered. This modality has a value of 67 against the *visual* trainee, who can listen to the content in a visual manner. The modality has a value of 33 for a *read/write* trainee. The device has a value of 71 against an *aural* trainee. Finally, this device has a value of 19 against a *kineasthetic* trainee. Once again, the values of these modalities differ in Figure 5.4 and Figure 5.17. This clearly shows the effect of the trainee's change of state.

To summarize, Figure 5.17 shows that *screen*, *symbols*, *video*, the *Sign Language*, *touch screen*, *eyetracker* and the *puff and sip* device have zero importance for the trainee who cannot *See*. The values of *vibration* in Figure 5.17 are zero for learning styles *visual*, *aural* and *read/write*. However, these values are non-zero in Figure 5.4. This is due to the fact that the trainee's state is altered in Figure 5.17, e.g. from *energetic* to the *tired* state. However, both Figure 5.4 and Figure 5.17 have predicted the value of *vibration* for the *kineasthetic* learning style.

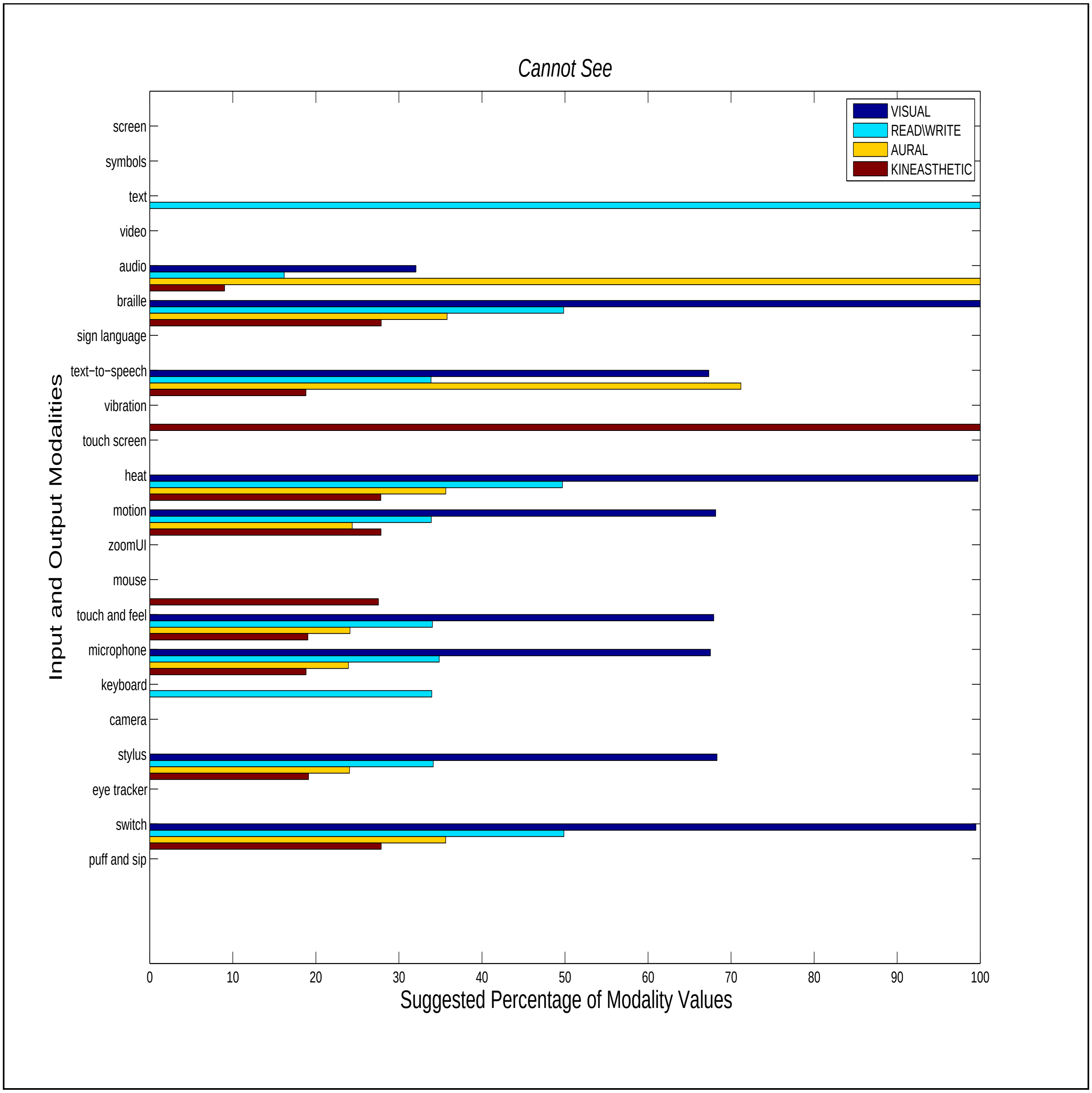

Figure 5.17: Importance values for a *tired* trainee who cannot *See*

It is noted from Figure 5.17 that *text*, Braille, *text-to-speech*, *vibration*, *heat*, *motion*, *mouse*, *touch and feel* device, *microphone*, *keyboard* and *stylus* are important modalities for a trainee who cannot *See* profile. These modalities make intuitive sense. These values are however different in Figure 5.4 and Figure 5.17. This clearly shows the effect of the trainee's change of state.

A *tired* trainee who cannot *Hear*

The non-linear model has been tested using the profile of a *tired* trainee who cannot *Hear*. The results of this study are summarised in Figure 5.18. We now discuss two different modalities and their predicted values against four learning styles (*visual*, *aural*, *read/write*, *kineasthetic*).

The first modality *audio* was considered. The difference between Figure 5.5, see page 50, and Figure 5.18 is that the state *tired* is now considered within the non-linear model with feedback. This modality has a value of 0 against all the different learning styles. This is sensible as a *deaf* trainee *cannot Hear*.

The second modality considered was *zoomUi*. This modality has a value of 100 against the *visual* trainee. The *zoomUi* modality has a value of 0 against the remaining learning styles. The results of this modality are the same in Figure 5.5 and Figure 5.18 for the learning style *visual*. The results are different for the remaining learning styles, when Figure 5.5 and Figure 5.18 are compared. The values in Figure 5.5 are 33 for the modality zoomui, while they are 0 in Figure 5.18. This clearly shows the effect of the trainee's change of state.

To summarize, Figure 5.18 shows that *audio*, *Braille*, *text to speech*, *vibration*s, *motion* and the *puff and sip* device have zero importance value for the *tired* trainee. The values of these modalities differ in Figure 5.5 and Figure 5.18. This clearly shows the effect of the trainee's change of state. The values of *text* in Figure 5.18 are zero for learning styles *visual*, *aural* and *kineasthetic*. However, these values are non-zero in Figure 5.5. It is noted that different states bring about different suggestions of modalities to the trainee. In this instance, the trainee may prefer *text* to the *keyboard*. Since the user is *tired*, he may prefer to read, rather than type. The non-linear learning model has suggested a more

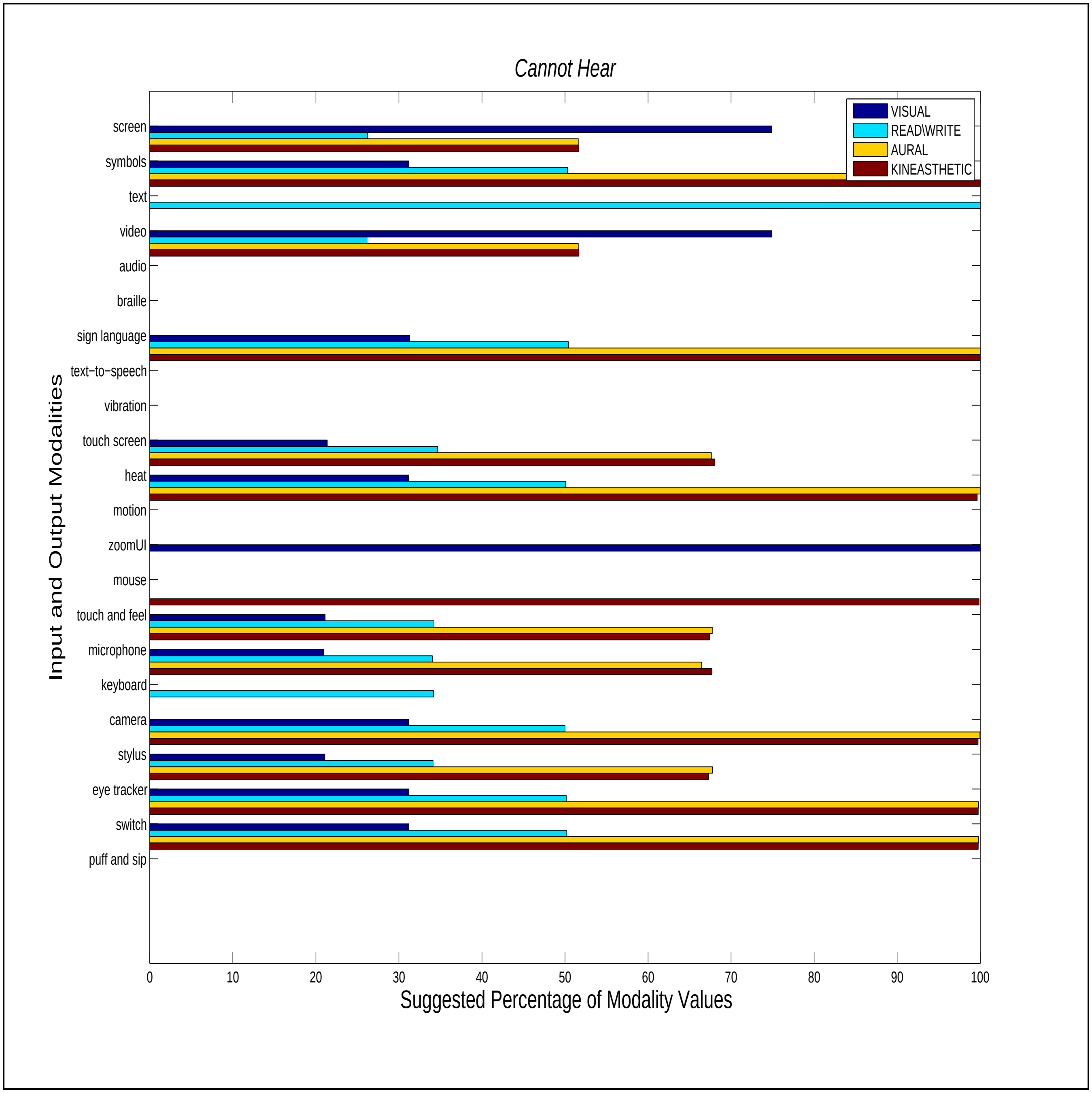

Figure 5.18: Importance values for a *tired* trainee who cannot *Hear*

accurate set of modalities in Figure 5.18 for the *tired* trainee.

It is noted from Figure 5.18 that the screen, *symbols*, *text*, *video*, the *Sign Language*, *heat*, *mouse*, *touch and feel* device, *microphone*, *keyboard*, *camera*, *stylus*, *eyetracker* and *switch* are important modalities for a *hearing* impaired trainee profile. These modalities make intuitive sense. The values of these modalities are different in Figure 5.5 and Figure 5.18. This clearly shows the effect of the trainee's change of state.

A *tired* trainee who is *Fully-able*

The effect of the model on a *tired* trainee who is *Fully-able* and can understand *Sign Language* is considered. The results of this study are summarised in Figure 5.19. Two different modalities and their predicted values against four learning styles (*visual*, *aural*, *read/write*, *kineasthetic*) are presented below.

The first modality considered was the *Braille* device. This modality has a value of 0 for all the different learning styles. When comparing Figure 5.6, see page 53, and Figure 5.19, the values of these modalities are exactly as predicted by the non-linear model without feedback, see Figure 5.6.

Finally, the *eyetracker* device was considered. This modality has a value of 31 against the *visual* trainee. The *eyetracker* has a value of 10 against the *read/write* trainee. The *eyetracker* device has a value of 36 against the *aural* trainee. Finally, the *eyetracker* device has a value of 100 against the *kineasthetic* trainee. By inspecting Figure 5.6 and Figure 5.19, one can see that the values of the modality *eyetracker* are different. In Figure 5.6 one finds that eyetracker has a value of 50 for the four learning styles. This clearly shows the effect of the trainee's change of state.

To summarize, Figure 5.19 shows that the trainee cannot use *Braille* as he has no knowledge of it. The modalities *vibration*, *motion* and the *puff and sip* device are of zero importance. This is exactly the same value predicted by the non-linear model without feedback in Figure 5.6. Clearly, the integration of the state *tired* in the model does not affect the trainee's choice of modalities. The values of *zoomUi* in Figure 5.19 are zero for learning styles *aural*, *read/write* and *kineasthetic*. However, these values are non-zero in Figure 5.6. One can clearly note that different states bring about different suggestions of

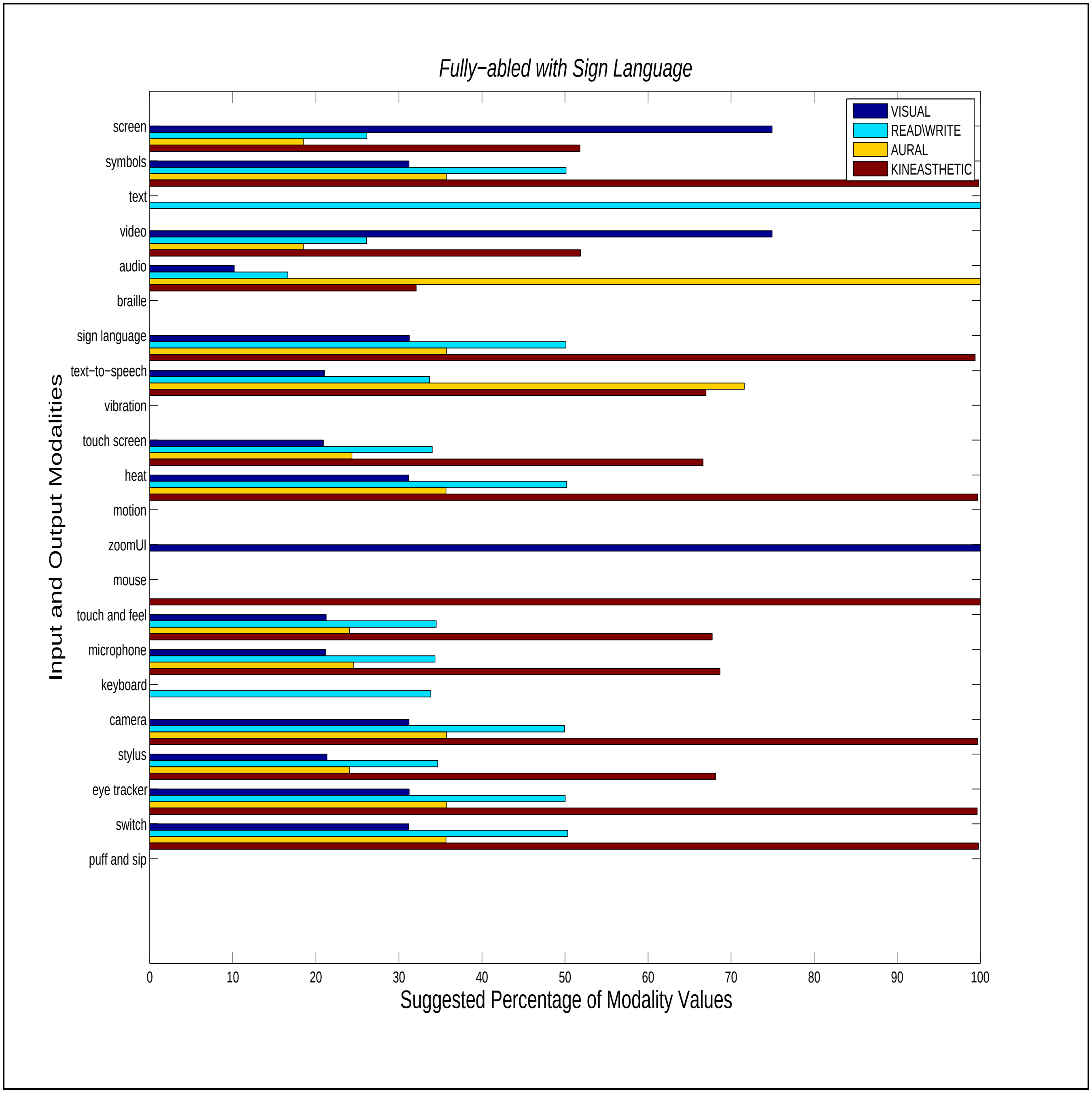

Figure 5.19: Importance values for a *tired Fully-able* trainee

modalities. In Figure 5.6, *screen*, *video* and *zoomUi* were suggested to the *visual* trainee. In Figure 5.19, *zoomUi* is suggested first to the trainee. This is due to the trainee's state *tired*, where the trainee may prefer the *zoomUi*. Therefore, the non-linear learning model has suggested a more accurate set of modalities in Figure 5.19. This clearly shows the effects of the state *tired*.

A *tired* trainee who cannot *Hear* but cannot understand *Sign Language*

The effect of our model on a *tired* trainee who cannot *Hear* and cannot understand sign is considered. The results of this study are summarised in Figure 5.20.

Once again, when analyzing Figure 5.20 it is noted that it is similar to Figure 5.18, see page 80, as mentioned earlier (a similar argument was described about *Sign Language* in Figure 5.12, see page 67). Thus in Figure 5.20, *Sign Language* is zero, otherwise the remaining results are identical to the results in Figure 5.18.

5.4 Trainee profile and change of state

This section studies the effect of different states against the modalities suggested for a given trainee profile. Thus the effect of how the different states affect a trainee is shown. Only the modalities with the highest importance value are considered. The results are taken from the preceding section.

5.4.1 List of abbreviations

For simplicity sake, the following abbreviations are applied to Tables 5.1 – 5.4.

- IV = Importance Value
- SL = *Sign Language*
- TTS = *text-to-speech*
- TF = *touch and feel*
- TS = *touch screen*

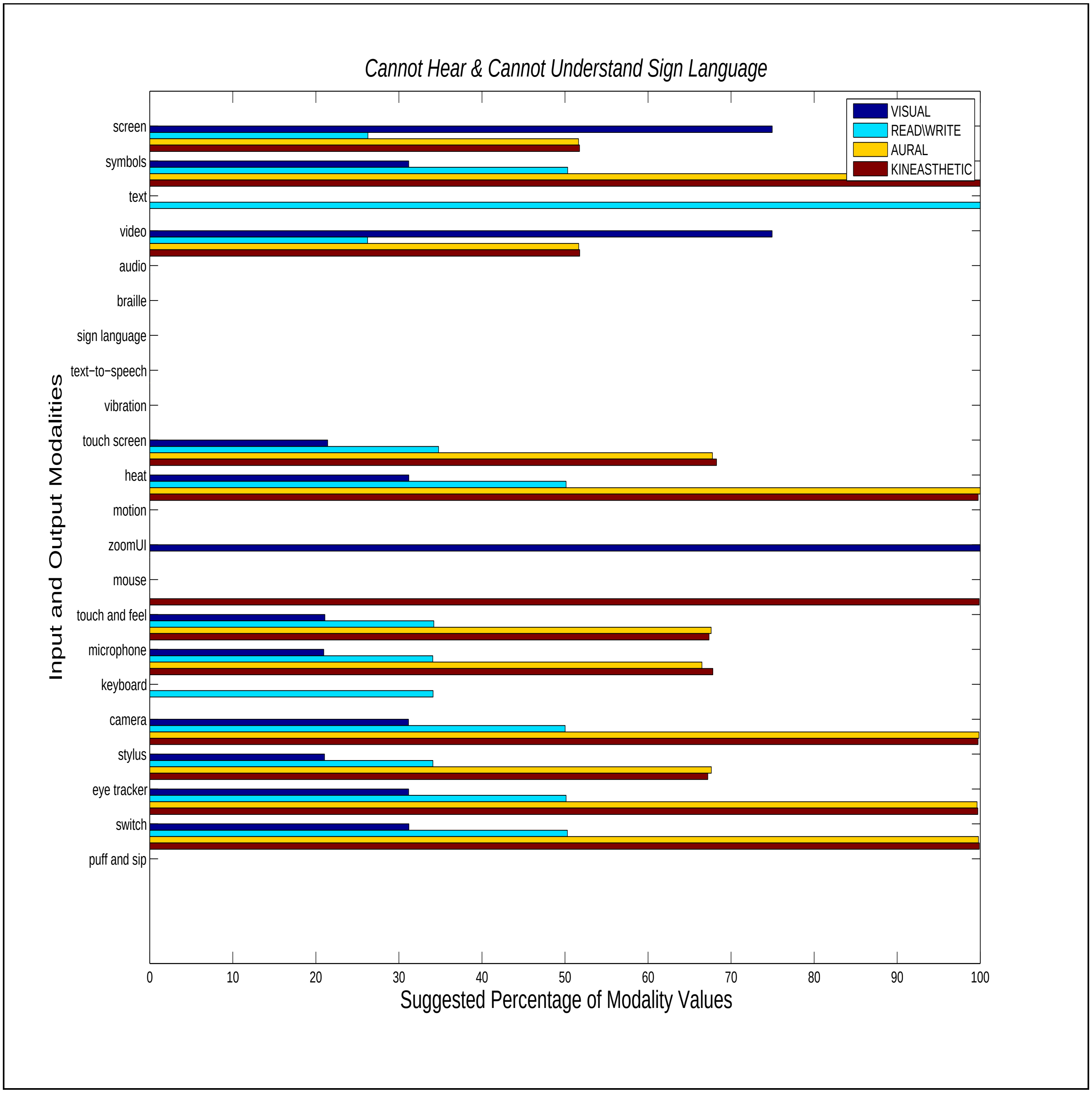

Figure 5.20: Importance values for a *tired* trainee who cannot *Hear* and cannot understand *Sign Language*

5.4.2 The *visual* learning style

Table 5.1 summarizes the modalities for various states for the learning style *visual*. In Table 5.1 the first column presents the trainee's state and the second column presents the trainee profile. The third column presents the modalities and the fourth column presents the importance value of the modalities, respectively. Table 5.1 shows that for the learning style *visual*, the suggested modalities are different for different states. For example, consider the trainee profile 'cannot *Hear* '. When the trainee is *energetic*, the modalities screen, video and zoomUi are suggested to the trainee. When the trainee becomes *bored*, the model now suggests the modality screen only. If the trainee falls into the state of *laziness*, the model suggests the modality video. When the trainee's state is *tired*, the modality zoomUi is suggested. One can see that a change of states brings about the change in modalities. One can intuitively note that when the change from *bored* to *lazy* occurred, the modality screen is replaced with video. A reason for this can be that a trainee prefers watching video when the trainee is *lazy*.

5.4.3 The *read/write* learning style

Table 5.2 summarizes the various modalities for the different states for the *read/write* learning style.

Consider the fully-able trainee profile. In the *energetic* state, the modalities text, *keyboard* and *stylus* are suggested. During the *bored* state the suggested modality is the *stylus*, while in the *lazy* state, the modalities *text* and *keyboard* are suggested. If the trainee becomes *tired*, *text* is suggested. If the change of state from *lazy* to *tired* is observed, *text* and *keyboard* becomes *text*. The intuitive reasoning behind this, is that the trainee now prefers to read *text* rather than to type on a *keyboard*.

Table 5.1: List of modalities for the *visual* learning style

State	**Trainee Profile**	**Modalities**	**IV**
energetic	*cannot See*	Braille	100
	cannot Hear	screen, video, zoomUi	100
	Fully-able with SL	screen, video, zoomUi	100
	cannot, Hear and use SL	screen, video, zoomUi	100
bored	*cannot See*	Braille	100
	cannot Hear	screen	100
	Fully-able with SL	screen	100
	cannot, Hear and use SL	screen	100
lazy	*cannot See*	heat	100
	cannot Hear	video	100
	Fully-able with SL	video	100
	cannot, Hear and use SL	video	100
tired	*cannot See*	Braille, heat, switch	100
	cannot Hear	zoomUi	100
	Fully-able with SL	zoomUi	100
	cannot, Hear and use SL	zoomUi	100

5.4.4 The *aural* learning style

Table 5.3 summarizes the various modalities for the different states for the *aural* learning style. Table 5.3 shows that for the learning style *aural*, the modalities suggested are different for various states.

For example, consider a trainee who *cannot See*. When the trainee is *energetic*, the modalities *audio, text-to-speech* and *microphone* are suggested. If the trainee's state changes to *bored*, the modality *microphone* is suggested. When the state becomes *lazy*, the *microphone* is still the modality suggested, while if the trainee becomes *tired*, the modality *audio* is suggested. Consider the state change from *bored* to *lazy*; the modality

Table 5.2: List of modalities for the *read/write* learning style

State	Trainee Profile	Modalities	IV
energetic	*cannot See*	text, keyboard, stylus	100
	cannot Hear	text, keyboard, stylus	100
	Fully-able with SL	text, keyboard, stylus	100
	cannot, Hear and use SL	text, keyboard, stylus	100
bored	*cannot See*	stylus	100
	cannot Hear	stylus	100
	Fully-able with SL	stylus	100
	cannot, Hear and use SL	stylus	100
lazy	*cannot See*	text, keyboard	100
	cannot Hear	text, keyboard	100
	Fully-able with SL	text, keyboard	100
	cannot, Hear and use SL	text, keyboard	100
tired	*cannot See*	text	100
	cannot Hear	text	100
	Fully-able with SL	text	100
	cannot, Hear and use SL	text	100

microphone is replaced with the modality *audio*. This could occur because the trainee prefers *audio* when *tired* compared to talking using the *microphone*.

5.4.5 The *kineasthetic* learning style

Table 5.4 summarizes the various modalities for the different states for the *kineasthetic* learning style and analyses the modalities suggested for these states. Consider the trainee profile of a trainee that *cannot, Hear* and use *Sign Language*. During the *energetic* state, the trainee's suggested modalities are *vibration* and the *touch screen* device. In the *bored* state, the suggested modality is the *touch screen*, while in the *lazy* state, the suggested

Table 5.3: List of modalities for the *aural* learning style

State	**Trainee Profile**	**Modalities**	**IV**
energetic	*cannot See*	audio, TTS, microphone	100
	cannot Hear	microphone	100
	Fully-able with SL	audio, TTS, microphone	100
	cannot, Hear and use SL	microphone	100
bored	*cannot See*	microphone	100
	cannot Hear	microphone	100
	Fully-able with SL	microphone	100
	cannot, Hear and use SL	microphone	100
lazy	*cannot See*	microphone	100
	cannot Hear	microphone	100
	Fully-able with SL	microphone	100
	cannot, Hear and use SL	microphone	100
tired	*cannot See*	audio	100
	cannot Hear	heat, camera, SL	100
	Fully-able with SL	audio	100
	cannot, Hear and use SL	heat, camera	100

modality is the *touch and feel* device. If the trainee's state falls under the category *tired*, the suggested modalities are *symbols, heat* and the *mouse*. When the trainee's change of state occurs, for example, from *energetic* to *lazy*, the suggested modalities change from the *touch screen* and *vibration* to the *touch and feel* device. A reason for this is that the *kineasthetic* trainee may prefer a *touch and feel* device (joystick) in the *lazy* state as compared to touching a screen.

Table 5.4: List of modalities for the *kineasthetic* learning style

State	Trainee Profile	Modalities	IV
energetic	*cannot See*	vibration and TF	100
	cannot Hear	vibration and TS	100
	Fully-able	TS and TF	100
	cannot, Hear and use SL	vibration and TS	100
bored	*cannot See*	vibration	100
	cannot Hear	TS	100
	Fully-able	TS	100
	cannot, Hear and use SL	TS	100
lazy	*cannot See*	TF	100
	cannot Hear	TF	100
	Fully-able	TF	100
	cannot, Hear and use SL	TF	100
tired	*cannot See*	vibration	100
	cannot Hear	symbols, SL, heat, camera	100
	Fully-able	symbols, SL, heat, mouse	100
	cannot, Hear and use SL	symbols, heat, mouse	100

5.4.6 Conclusion

In this chapter, the linear model was compared to the non-linear model. Both models were shown to be compatible in certain aspects, i.e. both models suggested the same modality most of the time. The non-linear model identified discrepancies in the linear model, for example the linear model suggested the *stylus* for the *kineasthetic* trainee, while the non-linear model suggested the *stylus* for the *read/write* trainee. Thus the non-linear model corrected these discrepancies by identifying the accurate modalities.

Over time, a trainee may require a different set of modalities due to him being *bored*, *tired* or *lazy*. A feedback mechanism is necessary to tackle this change in state. The non-linear model with feedback found that at different states, different modalities are suggested, an indication that this model can be applied to more complex scenarios making it an alternative model. Chapter 6 presents a conclusion to the model-based optimisation for enhanced training of individuals based on abilities, learning styles and preferences.

Chapter 6

Conclusion

The aim of the research carried out in this thesis was to enhance and broaden the scope of individual training based on abilities, learning styles and the trainee's preferences. To achieve this a set of modalities and learning preferences was used. A relationship between these sets with the trainee profile was identified. This relationship was established through the use of ANN. The model obtained was non-linear and thus an extension of the previously suggested linear model.

Results obtained by both the linear model and non-linear model were compared. For a base-line trainee profile, both the systems predicted comparable results. However, for a number of test cases, the linear model suggested inaccurate results whereas the non-linear model suggested more adequate results. For example, the linear model suggested the *stylus* as a modality for the *kineasthetic* trainee. However, the non-linear model correctly suggested the *stylus* modality for the *read/write* trainee. Hence the usefulness of the non-linear model was identified and confirmed. The model presented in thesis can be used as an alternative to the model in the literature.

An important aspect of this research was that the non-linear model is further extended to include time dependency in suggesting the modalities. When an individual gets training, there may be a change of his preferred modalities over time. This scenario has to date not been addressed in the literature. The model suggested can thus handle this time dependency. This model was also tested using a number of profiles, for example, '*cannot See*', '*cannot Hear*', '*fully-able*' and '*cannot Hear* and cannot use *Sign Language*'. The results obtained were realistic. This shows that the model can be implemented practically.

The research presented in this thesis adds value to the AbTi research project. In addition it has provided a novel way of mapping a user's profile to modalities. In this context, the approach followed and the subsequent results obtained have successfully addressed the challenge as laid out in the problem statement.

Future research can be carried out in a number of directions. For example, the model can be augmented using different theories from the field of psychology. The use of clustering within the model can be looked into for efficient implementation. The non-linear model was developed using ANN through the generation of artificial data. This model can be trained using real test cases, i.e., by linking the user profile with modalities. This is a first attempt to use a non-linear model (with feedback) for training of individuals. Thus the model gives results that can be interpreted in more than one way. This thesis has investigated and presented the problem and solution in one manner. Thus future work should investigate these results more thoroughly.

This thesis showed that HCI and user training can be improved utilising a non-linear mapping between computing modalities and a user profile to dynamically and in an ongoing fashion adapt interaction mechanisms.

Bibliography

[1] L. Coetzee, I. Viviers & E. Barnard. Model based estimation for multi-modal user interface component selection. In *PRASA09*, pages 11-16, 2009.

[2] M. T. Hagan & H. B. Demuth. Neural networks for control. In *Proceedings of the 1999 American Control Conference*, pages 1642-1656, 1999.

[3] F. Karray, M. Alemzadeh, J. A. Saleh & M. N. Arab. Human-computer interaction: Overview on state of the art. *SMART SENSING AND INTELLIGENT SYSTEMS*, 1:137–159, 2008.

[4] J. G. Merrienboer. Research on cognitive load theory and its design implications for e-learning. *ETR&D*, 53:5–13, 2005.

[5] M. Blattner & E. Glinert. Multimodal integration. *IEEE Multimedia*, 3:14-24, 1996.

[6] M. Gellevij, H. Meij & J. Pieters. Multimodal versus unimodal instruction in a complex learning context. *Experimental Education*, 70(3):215-239, 2002.

[7] S. Oviatt. Designing robust multimodal systems for universal access. Technical report, Oregon Graduate Institute of Science & Technology, 2001.

[8] J. Tomlinson, M. Russell & N. M. Brooke. Integrating audio and visual information to provide highly robust speech recognition. *IEEE ICASSP*, pages 821–824, 1996.

[9] S. Kawai, H. Aida & T. Saito. Designing interface toolkit with dynamic selectable modality. Technical report, The University of Tokyo, 1996.

[10] W. McCulloch & W. Pitts. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5:115–133, 1943.

[11] I. Viviers & L. Coetzee. Introduction to computer modalities. Technical report, C.S.I.R, South Africa, 2008.

[12] L. Coetzee & I. Viviers. Architectural components. Technical report, C.S.I.R, South Africa, 2008.

[13] S. Oviatt, P. Cohen, L. Wu, J. Vergo, E. Duncan, B. Suhm, J. Bers, T. Holzman, T. Winograd, J. Landay, J. Larson & D. Ferro. Designing the user interface for multimodal speech and gesture applications: State-of-the-art systems and research directions. Technical report, Oregon Graduate Institute of Science and Technology, 2000.

[14] S. Oviatt. Mutual disambignation of recognition errors in a multimodal architecture. Technical report, Oregon Graduate Institute of Science & Technology, 1999.

[15] A. Adjoudani & C. Benoit. Audio-visual speech recognition compared across two architectures. Technical report, Proceedings of the Eurospeech Conference, Madrid, Spain, 1995.

[16] J. M. Carroll. Human-computer interaction: Psychology as a science of design. *Annual Review of Psychology*, 48:61–83, 1997.

[17] S. Oviatt. Integration and synchronization of input modes during multimodal human-computer interaction. Technical report, Oregon Graduate Institute of Science & Technology, 1997.

[18] P. Brusilovsky. A distributed architecture for adaptive e-learning. In *Thirteenth International World Wide Web Conference*, pages 104-113, 2004.

[19] P. Brusilovsky & E. Millian. User models for adaptive hypermedia and adaptive educational systems. *The Adaptive Web: Methods and Strategies of Web Personalization*, 2007.

[20] N. Henze & P. Brusilovsky. Open corpus adaptive educational hypermedia. *The Adaptive Web: Methods and Strategies of Web Personalization*, 2007.

[21] P. Brusilovsky, S. Sosnovsky & M. Yudelson. Ontology-based framework for user model interoperability in distributed learning environments. In *World Conference on E-Learning, E-Learn 2005*, 2005.

[22] A. Kobsa. Generic user modeling systems. *User Modeling and User-Adapted Interaction*, 11:46-63, 2001.

[23] J. Orwant. Heterogenous learning in the Doppelganger user modeling system. User modeling and user-adapted interaction. *The Journal of Personalization Research 4*, pages 107–130, 1995.

[24] A. Panjawaranonda & A. Srivihok. Individualized e-learning system: Agent for learning mathematics in primary education in Thailand. Technical report, Kasetsart University, Bangkok, 2005.

[25] M. Blochl, H. Rumetshofer & W. Wob. Individualized e-learning systems enabled by a semantically determined adaptation of learning fragments. *Proceedings of 14th International Workshop on Database and Expert Systems Applications (DEXA03). IEEE Computer Society*, 2003.

[26] N. Sonwalkar. Adaptive learning technologies: From one-size-fits-all to individualization. Technical report, Educause Center for Applied Research, Colorado, USA, 2005.

[27] J. E. Villaverde, D. Godoy & A. Amandi. Learning styles' recognition in the e-learning environments with feed-forward neural networks. *Computer Assisted Learning*, 22:197–206, 2006.

[28] R. Dunn, K. Dunn & J. Perrin. *Teaching young children through their individual learning style*. Biston, MA, 1994.

[29] N. Dlodlo. Ability based technology interventions. research into technology interventions based on the abilities, learning styles and literacy levels of participants. Technical report, C.S.I.R, South Africa, 2008.

[30] L. Bostrom & L. M. Lassen. Unraveling learning, learning styles, learning strategies and meta-cognition. *Education and Training*, 48:178-189, 2006.

[31] R. M. Felder & L. K. Silvermann. Learning and teaching styles in engineering education. *Journal of Engineering Education*, 78:674-681, 1988.

[32] P. Honey & A. Mumford. The manual of learning styles. *Peter Honey*, 1986.

[33] R. S. Dunn & K. J. Dunn. Learning styles/teaching styles: Should they... can they... be matched. *Educational leadership*, 36:238–244, 1979.

[34] D.A. Kolb. Experiential learning: experiences as the source of learning and development. *Prentice-Hall*, 1984.

[35] N. D. Fleming & C. Mills. VARK a guide to learning styles. Last accessed 30 March 2009.

[36] D. A. Kolb. *Learning Styles and Disciplinary Differences*. Jossey-Bass Inc., 1981.

[37] R. Alberts, M. Pistorius & M. Selepe. Training needs per ability. Technical report, C.S.I.R, South Africa, 2008.

[38] R. Alberts & M. Selepe. Training needs per perceptual and sensory learning preference. Technical report, C.S.I.R, South Africa, 2008.

[39] G. Olivrin. Adjusting literacy with language transformations and alternative representations. Technical report, C.S.I.R, South Africa, 2008.

[40] L. Coetzee & I. Viviers. Documentation of content formats. Technical report, C.S.I.R, South Africa, 2008.

[41] R. Alberts & M.Pistorius. Scope document. Technical report, C.S.I.R, South Africa, 2007.

[42] M. Greeff & E. V. Macagnano. Intelligent Environments for Independent Living. Technical report, C.S.I.R, South Africa, 2008.

[43] F. A. Geldard. *The human senses.* Oxford, 1953.

[44] C. M. Bishop. *Neural Networks for Pattern Recognition.* Oxford University Press Inc., New York, 1995.

[45] K. Hornik, M. Stinchcombe & H. White. Multilayer feedforward networks are universal approximators. *Neural Networks 2*, pages 359–366, 1989.

[46] C. Stergiou & D. Siganos. Neural networks. Accessed via website:. http://www.doc.ic.ac.uk/ nd/surprise_96/journal/vol4/cs11/report.html in August 2009.

[47] Moody & Darken. Fast learning in networks of locally-tuned processing unites. *Neural Computation*, 1:281–294, 1989.

[48] S. Haykin. *Neural Networks: A Comprehensive Foundation.* NJ: Prentice-Hall, 1999.

[49] M. Agarwal. A systematic classification of neural-network-based control. *IEEE Contr. Syst. Mag.*, 17:75–93, 1997.

[50] J. B. D. Cabrera & K.S. Narendra. Issues in the application of neural networks for tracking based on inverse control. *IEEE Trans. Automat. Contr.*, 44:2007–2027, 1999.

[51] A. N. Kabundi, K. Hornik, D. J. Marais & L. Greyling. Predicting nonlinear models: Neural networks versus linear regression analysis, Rand Afrikaans University, South Africa, 2002.

[52] K. Gurney. *Introduction to Neural Networks*. UCL Press, 1997.

[53] J. Nocedal and S. J. Wright. Numerical optimization. *Springer*, 1999.

[54] G. Lera & M. Pinzolas. Neighborhood Based Levenberg-Marquardt Algorithm for Neural Network Training. *IEEE Transactions on NN*, 13:1200-1203, 2002.

[55] E. B. Baum & D. Haussler. What size net gives valid generalization? *Neural Computation*, 1:151–160, 1988.