A Secure Radio-Frequency Assistive Listening Device for Hard of Hearing People

by

Paul Sui-King Chan

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF APPLIED SCIENCE

in

THE FACULTY OF GRADUATE STUDIES
(DEPARTMENT OF ELECTRICAL ENGINEERING)

We accept this thesis as conforming to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA
September, 1993.
© Paul Sui-King Chan, 1993
In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

(Signature)

Department of Electrical Engineering

The University of British Columbia
Vancouver, Canada

Date Oct 6 1993
ABSTRACT

Assistive listening devices allow hearing impaired listeners to achieve higher comprehension than with hearing aids alone in difficult listening conditions which occurs during business meetings, lectures, and presentations. These devices minimize the acoustic pathway between the speaker and listener by using some form of electromagnetic transmission. A major problem with existing assistive listening devices is that they provide little security for the users, as the signal may be picked up by unauthorized listeners outside the confines of the meeting room or hall.

The design, implementation, and testing of a secure radio-frequency assistive listening device are described in this thesis. The device combines digital voice encoding technology with the direct sequence spread spectrum technique to provide a secure digital voice communication channel for hearing impaired people.

It is shown that the CCITT G.722 and the μ-law digital voice encoding algorithms are acceptable for hearing impaired listeners. A transmitter and receiver based on the Arlan 650TM wireless network card is implemented. Digitized speech samples are grouped into packets and transmitted using the TCP/IP protocol. The packet error rates of the device measured under some typical environment are presented.

The degradation of intelligibility under packet loss conditions is also presented. It is shown that speech intelligibility decreases as the number of lost speech segments increases and the lost speech segments are replaced by two simple packet loss replacement methods.
# Table of Contents

ABSTRACT ........................................................................................................... ii

List of Tables ...................................................................................................... vi

List of Figures ..................................................................................................... vii

ACKNOWLEDGMENTS ......................................................................................... xi

Chapter 1  Introduction ....................................................................................... 1

Section 1.1  Motivation ....................................................................................... 1

Section 1.2  Assistive Listening Devices ............................................................. 4

Section 1.3  Objective & Overall Design .............................................................. 6

Section 1.4  Spread Spectrum techniques ............................................................ 8

Section 1.5  Outline of the Thesis ....................................................................... 10

Chapter 2  Selection of a speech encoding method .............................................. 12

Section 2.1  Speech Characteristics and intelligibility ........................................ 12

Section 2.2  Subjective Measurement of different speech encoding techniques ..................... 16

Section 2.3  Selecting a speech encoding technique for hearing impaired listeners ................................................................. 23

2.3.1 Procedures of the SPIN test .................................................................. 26

2.3.2 Results in Quiet Conditions .................................................................... 30

2.3.3 Results in Noisy Conditions .................................................................... 32

2.3.4 Speech encoding techniques for the hearing impaired listeners .................... 37
Chapter 3  Hardware Design & Implementation ................. 38
Section 3.1  Overall design ........................................ 38
Section 3.2  Spread Spectrum modem .......................... 39
Section 3.3  Arlan 650™ characteristics ..................... 41
Section 3.4  Codec Interface ...................................... 45
Chapter 4  Software Design & Implementation .............. 53
Section 4.1  Overall design ........................................ 53
Section 4.2  TCP/IP protocol ..................................... 54
Section 4.3  Voice packet transmission in the radio-frequency ALD system ........................................... 61
Section 4.4  Voice packet size in the radio-frequency ALD .......... 72
Section 4.5  Lost packet replacement strategies ................ 76
Chapter 5  Evaluation of the radio-frequency ALD ............ 79
Section 5.1  Packet error rates of the radio-frequency ALD ........ 79
Section 5.2  Miller and Nicely audiometric test ................ 87
  5.2.1  Miller and Nicely Test procedures .................... 92
  5.2.2  Results of the Miller and Nicely Test ................ 95
    Results of the pilot subject ............................... 95
    Results using normal hearing subjects .................. 96
    Results using hearing impaired subjects ............... 97
  5.2.4  Effect of packet lost on speech intelligibility ....... 99
<table>
<thead>
<tr>
<th>Chapter 6</th>
<th>Conclusions</th>
<th>106</th>
</tr>
</thead>
<tbody>
<tr>
<td>Section 6.1</td>
<td>Summary</td>
<td>106</td>
</tr>
<tr>
<td>Section 6.2</td>
<td>Suggestions for further work</td>
<td>107</td>
</tr>
<tr>
<td>Appendix A</td>
<td>Example of a Revised SPIN Test Form</td>
<td>108</td>
</tr>
<tr>
<td>Bibliography</td>
<td></td>
<td>110</td>
</tr>
</tbody>
</table>
## List of Tables

| Table 1.1 | Total population and persons who report impaired hearing in Canada. | 1 |
| Table 2.1 | Frequency bands making equal (5 percent) contributions to articulation index when all bands are at their optimum levels. Composite data for men’s and women’s voices | 15 |
| Table 2.2 | Five-point scales for quality and impairment, and associated number scores. | 17 |
| Table 2.3 | Revised SPIN Test Score (Mean) | 37 |
| Table 3.1 | Known commercially available spread spectrum modems. | 40 |
| Table 3.2 | Radio frequency and data rate of channels in Arlan 650™ | 42 |
| Table 3.3 | Signal description of the TMS320C30 serial ports | 46 |
| Table 3.4 | Signal description of the MC14402 | 48 |
| Table 4.1 | Categorization of packet loss distortions. | 73 |
| Table 5.1 | Mean packet error rate in measurement # 1–7 | 84 |
| Table 5.2 | Classification of consonants used to analyze confusion | 89 |
| Table 5.3 | Overall Scores of the pilot subject | 95 |
| Table 5.4 | Miller and Nicely test results for the normal hearing subjects. | 96 |
| Table 5.5 | Miller and Nicely test results for the hearing impaired subjects. | 97 |
List of Figures

Figure 1.1  Distribution of persons with impaired hearing aged 15 and over residing in households, Canada. .......... 3

Figure 1.2  Proposed Assistive Listening Device. ................. 7

Figure 2.1  Idealized speech spectrum measured at one meter from lips. ............................................. 12

Figure 2.2  Articulation index versus cut-off frequency. All bands are at their optimum levels. Curve is based on about equal number of men's and women's voices. ............. 14

Figure 2.3  Subjective speech quality of different speech encoding techniques versus encoding bit rate. ........ 18

Figure 2.4  Relation between AI and various measures of speech intelligibility. ................................. 22

Figure 2.5  Scoring Nomograph for the Revised SPIN test. .... 25

Figure 2.6  Setups for applying speech encoding techniques to recordings of Revised SPIN test forms. .......... 27

Figure 2.7  Classification of hearing impairment in relation to handicap for speech recognition. ............. 29

Figure 2.8  Test setup in quiet condition ......................... 30

Figure 2.9  Audiograms of subjects taking Revised SPIN test in quiet condition. ............................... 31

Figure 2.10 Revised SPIN test in quiet conditions. ............ 32
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.11</td>
<td>Test setup in noisy conditions, babble is presented at 8 dBHL below speech signal.</td>
<td>33</td>
</tr>
<tr>
<td>2.12</td>
<td>Frequency response of the Realistic FM wireless microphone system (Model # 32-1221).</td>
<td>34</td>
</tr>
<tr>
<td>2.13</td>
<td>Audiograms of subjects taking Revised SPIN test under noisy conditions.</td>
<td>35</td>
</tr>
<tr>
<td>2.14</td>
<td>Revised SPIN test results for noisy conditions.</td>
<td>36</td>
</tr>
<tr>
<td>3.1</td>
<td>Overall design of the radio-frequency ALD system.</td>
<td>38</td>
</tr>
<tr>
<td>3.2</td>
<td>Radio frequency spectrum of Arlan 650™ Channel 7 to 9.</td>
<td>44</td>
</tr>
<tr>
<td>3.3</td>
<td>Block diagram of the Codec Interface.</td>
<td>45</td>
</tr>
<tr>
<td>3.4</td>
<td>TMS320C30 Serial Port Timing</td>
<td>47</td>
</tr>
<tr>
<td>3.5</td>
<td>Transmit and receive timing diagram for MC14402.</td>
<td>49</td>
</tr>
<tr>
<td>3.6</td>
<td>TMS320C30 Processor Board and MC14402 Interface signal timing.</td>
<td>50</td>
</tr>
<tr>
<td>3.7</td>
<td>Clock generation circuit in the Codec Interface.</td>
<td>51</td>
</tr>
<tr>
<td>3.8</td>
<td>MC14402 to TMS320C30 Processor Board interface circuit.</td>
<td>52</td>
</tr>
<tr>
<td>4.1</td>
<td>Overall software design of the radio-frequency ALD system.</td>
<td>53</td>
</tr>
<tr>
<td>4.2</td>
<td>The three level of service provided by TCP/IP.</td>
<td>55</td>
</tr>
<tr>
<td>4.3</td>
<td>TCP/IP reference model.</td>
<td>57</td>
</tr>
<tr>
<td>4.4</td>
<td>Communication process in the TCP/IP protocols.</td>
<td>60</td>
</tr>
<tr>
<td>Figure 4.5</td>
<td>Communication processes in the radio-frequency ALD system.</td>
<td>62</td>
</tr>
<tr>
<td>Figure 4.6</td>
<td>Structure of an IP datagram.</td>
<td>64</td>
</tr>
<tr>
<td>Figure 4.7</td>
<td>The five subfields that comprise the Type of service field.</td>
<td>65</td>
</tr>
<tr>
<td>Figure 4.8</td>
<td>Fragmentation of an IP datagram</td>
<td>66</td>
</tr>
<tr>
<td>Figure 4.9</td>
<td>IP address classes.</td>
<td>68</td>
</tr>
<tr>
<td>Figure 4.10</td>
<td>Structure of an Ethernet frame in radio-frequency ALD</td>
<td>70</td>
</tr>
<tr>
<td>Figure 4.11</td>
<td>Overhead accompanying each voice packet.</td>
<td>74</td>
</tr>
<tr>
<td>Figure 4.12</td>
<td>Data transmission rate versus packet size.</td>
<td>75</td>
</tr>
<tr>
<td>Figure 4.13</td>
<td>Packet replacement techniques for the radio-frequency ALD.</td>
<td>77</td>
</tr>
<tr>
<td>Figure 5.1</td>
<td>Location of radio-frequency ALD transmitter and receiver in Measurement #1, 2, 3, 6, 7, 8, 9 and 10.</td>
<td>81</td>
</tr>
<tr>
<td>Figure 5.2</td>
<td>Location of radio-frequency ALD transmitter and receiver in Measurement #4 and 5.</td>
<td>82</td>
</tr>
<tr>
<td>Figure 5.3</td>
<td>Packet error measurement of the radio-frequency ALD.</td>
<td>83</td>
</tr>
<tr>
<td>Figure 5.4</td>
<td>Packet error measurements under interference</td>
<td>85</td>
</tr>
<tr>
<td>Figure 5.5</td>
<td>Packet error measurement under interference.</td>
<td>86</td>
</tr>
<tr>
<td>Figure 5.6</td>
<td>A confusion matrix.</td>
<td>90</td>
</tr>
<tr>
<td>Figure 5.7</td>
<td>A confusion matrix grouped by voiced and voiceless consonants</td>
<td>91</td>
</tr>
<tr>
<td>Figure 5.8</td>
<td>Confusion matrix for voiced and voiceless consonant</td>
<td>92</td>
</tr>
<tr>
<td>Figure 5.9</td>
<td>Simulated packet loss in the Miller and Nicely syllables</td>
<td>93</td>
</tr>
<tr>
<td>Figure 5.10</td>
<td>Set up for the Miller and Nicely test</td>
<td>94</td>
</tr>
<tr>
<td>-------------------</td>
<td>---------------------------------------------</td>
<td>----</td>
</tr>
<tr>
<td>Figure 5.11</td>
<td>Audiogram of the hearing impaired subjects</td>
<td>98</td>
</tr>
<tr>
<td>Figure 5.12</td>
<td>Results of the Miller and Nicely Tests</td>
<td>99</td>
</tr>
<tr>
<td>Figure 5.13</td>
<td>Overall scores for the hearing impaired subjects</td>
<td>100</td>
</tr>
<tr>
<td>Figure 5.14</td>
<td>Results of different articulatory scores in silence substitution for normal hearing subjects</td>
<td>101</td>
</tr>
<tr>
<td>Figure 5.15</td>
<td>Results of different articulatory scores in packet repetition for normal hearing subjects</td>
<td>102</td>
</tr>
<tr>
<td>Figure 5.16</td>
<td>Results of different articulatory scores in silence substitution for hearing impaired subjects</td>
<td>103</td>
</tr>
<tr>
<td>Figure 5.17</td>
<td>Scores of the 16 syllables under different packet lost conditions</td>
<td>104</td>
</tr>
</tbody>
</table>
ACKNOWLEDGMENTS

I would like to thank Dr. C.A. Laszlo and Dr. C.S.K. Leung for their patience, suggestions and help during the project.

Special thanks are extended to Dr. M.K. Pichora-Fuller in the Department of Audiology and Speech Sciences for her advice and use of equipment to carry out the intelligibility tests. I would also like to thank M. Frauendorf, J. Nicol, S. Rosenberg, O.G. Gilbert for help in getting the software for intelligibility listening test to work.

In addition, I would like to thank all volunteers who participated in the intelligibility listening tests.

Financial support from the B.C. Science Council and NSERC grant OGP001731 is gratefully acknowledged.
Chapter 1 Introduction

1.1 Motivation

According to a Statistics Canada 1992 survey [15], of 25,061,270 Canadians, 1,022,220 report they have difficulty in hearing. This means that more than 1 of every 25 Canadian adults (4.1%) report that their hearing is impaired to some degree. More detailed statistics are given in Table 1.1. The reported rates of impaired hearing increase markedly with age, from slightly less than 1% for persons under 25 years of age residing in households, to almost half (47.5%), for persons 85 years of age and older. For persons with impaired hearing aged 15 and over residing in households, aging is the primary cause of impaired hearing (29.14%). Disease or stroke is the second most frequent cause (20.98%). Almost 1 out of 5 persons (18.11%) indicates that the impairment resulted from something associated with work. Accidents and injuries account for 8.59% and 6.86% have impaired hearing present at birth.

<table>
<thead>
<tr>
<th>Residence</th>
<th>Total Population</th>
<th>Hearing Impaired</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Number</td>
</tr>
<tr>
<td>Total</td>
<td>25,061,270</td>
<td>1,022,220</td>
</tr>
<tr>
<td>Under 15</td>
<td>5,325,185</td>
<td>48,390</td>
</tr>
<tr>
<td>15 years and over</td>
<td>19,736,085</td>
<td>973,830</td>
</tr>
<tr>
<td>Residing in ...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Household</td>
<td>24,806,180</td>
<td>908,825</td>
</tr>
<tr>
<td>Under 15</td>
<td>5,322,315</td>
<td>47,970</td>
</tr>
<tr>
<td>15 years and over</td>
<td>19,483,865</td>
<td>860,855</td>
</tr>
<tr>
<td>Health-related Institutions</td>
<td>255,090</td>
<td>113,395</td>
</tr>
<tr>
<td>Under 15</td>
<td>2,870</td>
<td>420</td>
</tr>
<tr>
<td>15 years and over</td>
<td>252,220</td>
<td>112,975</td>
</tr>
</tbody>
</table>
Chapter 1 Introduction

The hearing impaired population consists of people of all ages, with varied occupations, interests, hearing loss and needs. Within the hearing impaired population there are two groups, the deaf and the hard of hearing. The needs and characteristics of these two groups are very different. The primary mode of communication for the deaf is sign language, while the primary mode of communication for the hard of hearing is speech. The overall impact of hearing loss is significant, causing retardation of educational progress and the socialization process for children. For those who acquire hearing loss later in life, the consequences include loss of self-esteem, tensions in inter-personal relationships, difficulties in the work place, and varied psychosocial problems.

Impaired hearing may refer to minor difficulties or to the complete inability to use hearing for conversation. Using the responses to questions about the ability to carry on a conversation, three categories of impaired hearing have been defined in the survey. Fig. 1.1 shows the distribution of the three categories of hearing impaired aged 15 and over residing in households in Canada.

Category II, the middle range of hearing difficulty, has the largest number of persons, with 587,065 (68.2%), while 211,930 (24.6%) are in Category I, and 45,575 (5.3%) are in Category III. The fact that fewer cases of impaired hearing are found in Category I than would be anticipated is probably due to persons with mild hearing impairments not reporting them as often as those with more severe impairments, because they are unaware of them or do not regard them as limiting. The fact that the survey defines hearing impairment as one that limits the individual in daily activities reinforces this explanation.
Chapter 1 Introduction

Category I - persons who say they have no difficulty hearing one person but have at least partial difficulty hearing in groups;
Category II - persons who say they have partial difficulty hearing one person and have at least partial difficulty hearing in groups;
Category III - those who are completely unable to hear in one person conversations;
IND. - refers to those persons who have impaired hearing but whose degree of impairment cannot be determined because key answers are missing.

Fig. 1.1 Distribution of persons with impaired hearing aged 15 and over residing in households, Canada. [15]

Dramatic changes have occurred in hearing health care during the last decade due to the development of new surgical interventions and the application of technology. Some communication problems of the hard of hearing population can now be solved effectively due to advances in hearing aid technology. Unfortunately, hearing aids are able to provide only a partial solution. There are many situations in which the hearing aid does not provide adequate intelligibility of speech. Usually such problems occur
in noisy environments with low signal-to-noise ratio or in poor acoustic circumstances where reverberation reduces speech intelligibility. Some problem areas for the hard of hearing population are:

1. auditoriums or meeting rooms where speaker-to-listener distance, reverberation and noise reduce speech intelligibility.

2. television, telephone, and radio listening where poor fidelity and interference from room noise affects intelligibility, and

3. person-to-person communication in such noisy environments as restaurants, automobiles and parties.

Hearing aids, which amplify sounds regardless of their origin, become almost useless when the acoustical signal-to-noise ratio is low. While persons with normal hearing understand speech when the signal-to-noise ratio is 6 dB to 12 dB [31], many hard of hearing people require a signal-to-noise ratio of 15 dB to 20 dB [28] to function adequately.

1.2 Assistive Listening Devices

Assistive listening devices offer hearing impaired listeners better communications than hearing aids alone can provide in difficult listening conditions. These devices minimize the acoustical pathway between the speaker and listener using some form of electromagnetic transmission. When used properly, these systems provide the hard of hearing listener with a high level acoustic signal of good quality. These systems are having a great impact on the ability of hard of hearing individuals to participate in professional and business meetings, and educational and leisure activities. There are four
Chapter 1 Introduction

types of systems available:

1. Hard-wire system consisting of a microphone, an amplifier and a number of earphones connected by a wire cord.

2. Magnetic induction loop system consisting of a microphone, an amplifier and a coil of wire placed around the room. The amplified electrical speech signal is fed to the coil of wire, producing a modulated magnetic field in the room. Hearing aids with built-in induction coils can pick up the speech signal in the magnetic field and convert it into sound.

3. Infrared system consisting of one or more microphones wired into an amplifier/driver that powers one or more infrared emission panels. The modulated infrared light travels to individual receivers which convert the light back into an electrical signal, and provide acoustical, magnetic or direct input to the hearing aid of the user.

4. FM system consisting of a microphone with a transmitter worn by the speaker and a radio receiver worn by a listener. Current radio-frequency systems are allocated the 72–76 MHz radio band by the FCC in the US and DOC in Canada. Within this band 32 separate channels are available.

Each of these systems have specific advantages and operational limitations. The hard-wire system is the simplest but the mobility of the user is limited by the length of the cord. Systems based on magnetic induction loops require either permanent installation or placement of wires around the room each time it is used. In addition, the strength of the magnetic field emitted from the coil varies as a function of distance and shows other unpredictable variations within the room. The physical orientation of the induction coil in the hearing aid will also markedly affect the signal received by the user. Most
systems based on infrared technologies also require extensive installation, governed by the need to provide even illumination by placing the infrared emission panels in appropriate locations in the room. In addition, infrared systems cannot be used in environments where the infrared transmission may be swamped by sunlight or other hot light sources. The FM system requires minimum expertise to set up and offers maximum mobility for the user.

A major problem with the loop and FM systems is that they provide minimum security for the users, as the signal may be picked up by unauthorized listeners outside the confines of the room or meeting hall. In some applications, this is not a problem, but in many business meetings, lectures, and presentations the possibility of being overheard is not acceptable.

1.3 Objective & Overall Design

The objective of this study is to develop a specialized radio-frequency assistive listening device (ALD) for the hearing impaired population that combines the best features of existing methodologies. The design criteria for this device include transmission security, transmission quality to satisfy the needs of the hard of hearing user, and the ease of use of the system.

Fig. 1.2 shows the basic arrangement for the proposed ALD system. Listeners receive transmitted speech via a light-weight headphone, or couple the ALD to their personal hearing aid directly or via a personal loop if special amplification or signal processing is required to enhance their listening ability. This flexible approach will make the technology not only useful to the hearing impaired population, but it will also be attractive to users of other communication devices, such as walkie-talkies and cordless...
Chapter 1 Introduction

telephones, who want security in their systems.

Fig. 1.2 Proposed Assistive Listening Device.

In this thesis an investigation on the use of spread spectrum techniques to secure the voice signal transmitted between the speaker and hearing impaired listeners is presented. The process requires digitizing voice into bits of data. Each bit of data is encoded by a pseudorandom sequence for transmission. At the receiver end, decoding is accomplished by correlating the receiving signal with a synchronized replica of the same pseudorandom sequence which was used to transmit the data. The decoded digital signal is then converted back into voice signals. The idea behind this scheme is that while each authorized receiver could tune in to different transmitters by using their assigned pseudorandom code sequences, unauthorized receivers would be “locked out”. Given
the large number of possible code sequences, it would be difficult for an unauthorized user to recover the digitized voice sent by the transmitter. In this fashion, authorized user would be able to freely use their communication system, but illegal eavesdropping could be made quite difficult.

1.4 Spread Spectrum techniques

Spread spectrum systems were first developed in the mid-1950s. The initial applications have been in anti-jamming of military tactical communications, in missile guidance systems, and in anti-multipath systems. The spread-spectrum technique is a means of transmission in which the signal occupies a bandwidth in excess of the minimum necessary to send the information. The band spread is accomplished by means of a pseudorandom code which is independent of the data. Synchronized reception with the pseudorandom code at the receiver is used for de-spreading and subsequent data recovery. Although the use of the spread-spectrum technique means that each transmission requires the use of a wide band of the spectrum, many of the requirements demanded by our application can be satisfied simultaneously. These include:

• Low probability of intercept;
• Anti-jamming;
• Anti-interference;
• Low-density power spectra for hiding the signal;
• Message screening from eavesdroppers; and
• Selective addressing capability.
Chapter 1 Introduction

There are several means by which the spectrum of a signal can be spread.

1. Modulation of a carrier by a digital pseudorandom code sequence whose bit rate is much higher than the information bit rate. Such a technique is known as “direct sequence” modulation.

2. Carrier frequency shifting in discrete increments in a pattern dictated by a pseudorandom code sequence. The transmitter jumps from frequency to frequency within some predetermined set. This is called “frequency hopping”.

3. Conceptually similar to the “frequency hopping” technique is the “time hopping” technique in which the bursts of data signals are transmitted at pseudorandom times.

4. Hybrid combinations of the above techniques are also frequently used.

Direct sequence (DS) methods are the most common in spread spectrum systems. This is because of their relative simplicity and efficiency of the technique. For example, DS methods do not require a high speed frequency synthesizer as in the frequency hopping systems. Compared to DS systems, simple time-hopping systems offer little in the way of interference rejection since a continuous carrier at the signal center frequency can block communications effectively. Because of this relative vulnerability to interference, time hopping techniques are usually combined with other spread spectrum techniques. The advantage of combining two spread spectrum techniques in hybrid systems is to obtain characteristics which are not available in using a single spread spectrum technique. The specific construction of hybrid systems usually depends on its application, and its implementation is usually more complicated than systems using a single spread spectrum technique. Each of the spread spectrum techniques discussed can be used to achieve the desired spectrum spreading effect.
Chapter 1 Introduction

Each technique is important in the sense that it has useful applications. The historical tendency has been to use each method mainly in a particular field of applications. [26]

1.5 Outline of the Thesis

Chapter 2 examines speech encoding techniques suitable for this application. A discussion of the characteristics of speech signals and factors affecting speech intelligibility appears in Section 2.1. Section 2.2 is a review of some commonly used speech encoding techniques, with emphasis on their relative quality and bit rates. Section 2.3 presents the results of using the Speech Perception in Noise (SPIN) test to select a speech digitizing technique for the radio frequency assistive listening device.

Chapter 3 deals with the hardware design of the radio frequency assistive listening device. Section 3.1 discusses the overall design. A review of known spread-spectrum modems is presented in Section 3.2. Section 3.3 presents the features and characteristics of the Arlan 650™ spread-spectrum modem. Section 3.4 discusses the hardware interface used to send the digitized voice data to the modem.

Chapter 4 deals with the software implementation aspects of the radio frequency assistive listening device. Section 4.1 presents the overall software design. Transmission Control Protocol/Internet Protocol (TCP/IP), the protocols used to transport the digitized voice signal, is discussed in section 4.2. Section 4.3 deals with the transmission of the digitized voice data in the TCP/IP hierarchy. Section 4.4 is an analysis of packet size with respect to the bit transmission rate. Section 4.5 explains how missing packets can be replaced.

Chapter 5 contains the results of tests carried out with the modem. Section 5.1
Chapter 1 Introduction

describes the packet error rates of the radio-frequency assistive listening device in different environments. Section 5.2 deals with applying the Miller and Nicely audiometric test to determine the effect of missing packets.

Chapter 6 contains the conclusions and suggestions for future work.
Chapter 2 Selection of a speech encoding method

2.1 Speech Characteristics and intelligibility

Speech consists of a succession of sounds varying rapidly from instant to instant both in intensity and in frequency. The sounds of speech contain energy between at least 100 and 8000 Hz. Spectral analyses of spoken English have been studied extensively in the past [17], [20], [8], and was found that the spectra of individual voices differ considerably. For comparison and testing purposes an idealized speech spectrum (Fig. 2.1), based on average measurement of a group of speakers, has been developed by French and Steinberg [20].

Fig. 2.1 Idealized speech spectrum measured at one meter from lips [20].
This spectrum is measured at a distance of one meter from the lips in a sound field free from reflections. The intensity of this spectrum, integrated over the entire frequency range, amounts to 65 dB relative to $10^{-16}$ watt/cm². Audibility in the entire frequency range is not required for good intelligibility of speech. Fletcher [12] states that substantially complete fidelity for the transmission of speech is obtained by a system having a frequency range from 100 to 7000 cycles per second and a range of 40 decibels in amplitude.

The contribution of an individual speech sound to the comprehension of an utterance is a very complex matter. Frequency, intensity, temporal characteristics, speech context and speaker characteristics interact and contribute to the intelligibility of speech. The effective proportion of the speech signal available to a listener depends upon the intensity of the various sound components in their ears and the intensity of unwanted sounds that may be present.

The Articulation Index (AI) [1], [17], [20] has been used as a quantitative measure of the intelligibility of speech transmitted over communication systems. It is based on the concept that any narrow band of speech frequency of a given intensity carries an independent contribution to the total index and that the total contribution of all the bands is the sum of the contributions of separate bands. The magnitude of this index is taken to vary between zero and unity, the former applying when the received speech is completely unintelligible, the latter to the condition of best intelligibility.

French and Steinberg [20] investigated the relationships between articulation index A, speech intensity, and frequency response of communication systems using a group of listeners and talkers. They derived the following equation: 13
Chapter 2 Selection of a speech encoding method

\[ A = \sum W_n \cdot (\Delta A)_{\text{max}} \]

where \( A \) = Total articulation index

\((\Delta A)_{\text{max}}\) = maximum contribution of any one band

\(W_n\) = percent of maximum contribution contributed by each band

The relationships between \(W_n\), and the respective levels of speech and of acoustical noise in the ear is presented in [20]. Methods for the calculation of articulation index is given in [1].

![Articulation index vs Cut-off frequency](image)

**Fig. 2.2** Articulation index versus cut-off frequency. All bands are at their optimum levels. Curve is based on about equal number of men's and women's voices [20].
Chapter 2 Selection of a speech encoding method

From their experimental data, French and Steinberg [20] derived a curve of articulation index versus cut-off frequency of low pass filters under the special condition of optimal loudness at the ear and in negligibly low noise level (Fig 2.2). Two important points can be concluded from their studies:

1. Extending the frequency range of a communication system below 250 Hz or above 7000 Hz contributes almost nothing to the intelligibility of speech.

2. Each of the following frequency bands in Table 2.1 makes a 5 percent contribution to the articulation index, provided that all bands are at their optimal levels.

Table 2.1 Frequency bands making equal (5 percent) contributions to articulation index when all bands are at their optimum levels. Composite data for men’s and women’s voices [20].

<table>
<thead>
<tr>
<th>Band</th>
<th>Frequencies</th>
<th>Band</th>
<th>Frequencies</th>
<th>Band</th>
<th>Frequencies</th>
<th>Band</th>
<th>Frequencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>250-375</td>
<td>6</td>
<td>955-1130</td>
<td>11</td>
<td>1930-2140</td>
<td>16</td>
<td>3255-3680</td>
</tr>
<tr>
<td>2</td>
<td>375-505</td>
<td>7</td>
<td>1130-1315</td>
<td>12</td>
<td>2140-2355</td>
<td>17</td>
<td>3680-4200</td>
</tr>
<tr>
<td>3</td>
<td>505-645</td>
<td>8</td>
<td>1315-1515</td>
<td>13</td>
<td>2355-2600</td>
<td>18</td>
<td>4200-4860</td>
</tr>
<tr>
<td>4</td>
<td>645-795</td>
<td>9</td>
<td>1515-1720</td>
<td>14</td>
<td>2600-2900</td>
<td>19</td>
<td>4860-5720</td>
</tr>
<tr>
<td>5</td>
<td>795-955</td>
<td>10</td>
<td>1720-1930</td>
<td>15</td>
<td>2900-3255</td>
<td>20</td>
<td>5720-7000</td>
</tr>
</tbody>
</table>

It should be noted that the perceived naturalness of the speech is considered as a separate item here. It is known that although low frequency enhancement (50 Hz to 200 Hz) does not contribute to intelligibility, it will increase the naturalness of speech [21].

The required performance level of a given communication system can only be evaluated by the users of the system. Hirsh et al. [13] suggests that for low pass filtered speech, intelligibility dropped only slightly when frequencies above 1600 Hz were removed. There is no single value of the articulation index which can be specified.
Chapter 2 Selection of a speech encoding method

as a criterion for “acceptable” communication. Present-day commercial communication systems are usually designed for operation under conditions that provide articulation indexes in excess of 0.5 [1].

2.2 Subjective Measurement of different speech encoding techniques

With the advent of digital communications, considerable interest has been focused on the efficient encoding of speech. Over the past decade a large number of digital coding algorithms have been investigated for a wide variety of applications [24], [19], [14], [27]. Numerous signal processing techniques taking advantage of speech production and perception properties have been proposed and studied for the purpose of reducing the required transmission or storage rate for digitized speech. These techniques range from low to high complexity in design, and offer a corresponding trade-off between performance and complexity.

The performances of different speech encoding techniques are usually evaluated by quality or intelligibility tests. The testing procedure will depend on whether the issue is the quality or the intelligibility of the digitized voice. In the high quality range, intelligibility is good and the perceived naturalness of the digitized speech is usually assessed by qualitative measurements. In the low quality range, quality is low anyway and intelligibility criteria will determine whether the coding technique is acceptable.

Opinion rating on a subjective five-point scale is commonly used to assess the degree of speech quality or speech impairment (Table 2.2).
Table 2.2 Five-point scales for quality and impairment, and associated number scores.

<table>
<thead>
<tr>
<th>Score</th>
<th>Quality Scale</th>
<th>Impairment Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Excellent</td>
<td>Imperceptible</td>
</tr>
<tr>
<td>4</td>
<td>Good</td>
<td>(Just) Perceptible but not Annoying</td>
</tr>
<tr>
<td>3</td>
<td>Fair</td>
<td>(Perceptible) Slightly Annoying</td>
</tr>
<tr>
<td>2</td>
<td>Poor</td>
<td>Annoying (but not Objectionable)</td>
</tr>
<tr>
<td>1</td>
<td>Unsatisfactory (Bad)</td>
<td>Very Annoying (Objectionable)</td>
</tr>
</tbody>
</table>

The final result from these tests, in the simplest form, is the pooled average judgement called the Mean Opinion Score (MOS) for a group of listeners. A MOS of 5.0 implies perfect quality, but this is hardly ever attained, even by undigitized speech. Unimpaired and extremely high quality speech tends to get an MOS rating between 4 and 5. This is due to the fact that subjects may sometimes award a score such as 4 to a speech sample that ideally deserves a 5; or they may occasionally rank a slightly impaired stimulus higher than the original. A MOS approaching 4.5 signifies high-quality or near-transparent coding. A score of 3.5 on the MOS scale indicates that there is some detectable distortion but very little degradation of intelligibility. Lowest in the hierarchy of speech coding is synthetic speech with a MOS not exceeding 3.0, this quality is characterized by high intelligibility but has an inadequate level of naturalness and speaker recognizability.
Fig. 2.3 Subjective speech quality of different speech encoding techniques versus encoding bit rate.

Fig 2.3 is a quantitative description of speech quality with various commonly used speech encoding techniques as a function of their bit rates in kilobits per seconds (Kbps) [33], [24], [21]. The subjective quality of digitized speech tends to increase as the bit rate is increased. The complexity of different encoding techniques is also illustrated in Fig. 2.3. The level of complexity is quantified by implementation criteria such as the number of multiply/add operations involved per waveform sample. The high complexity encoding techniques are also characterized by the highest levels of encoding delay, making the more complex techniques less useful than others in real time applications. However, advances in digital technology tend to make complex digitization techniques more practical.

A brief description of the digital encoding techniques shown in Fig. 2.3 follows.
Chapter 2 Selection of a speech encoding method

Pulse Code Modulation (PCM) used in telecommunications usually samples at a rate of 8 KHz and quantizes the amplitude of each sample by rounding off each sample value to a set of discrete values. Amplitude compression is typically used following either the so called $\mu$-law or A-law standards. They are characterized by fine quantizing steps for the very frequently occurring low amplitude speech segments; coarser quantizing steps are used for the occasional large amplitude segments. [21]

The Adaptive Differential Pulse Code Modulation (ADPCM) exploits the high amplitude correlations in adjacent speech samples to reduce the bit rate. It is based on the notion of adaptively quantizing the difference in amplitude of adjacent samples. [21]

Sub-band coding (SBC) decomposes the 0–4 KHz frequency band using band pass filters into the sub-bands 0–500, 500–1000, 1000–2000, 2000–3000 and 3000–4000 Hz. An ADPCM coder with fixed first-order prediction is employed for each band. [33]

The CCITT G.722 algorithm is a form of SBC aimed to provide a 7 KHz audio bandwidth at 64 Kbps. Two sets of filters are used to divide the audio signal sampled at 16 KHz into a high band and a low band. Two ADPCM coders are used to quantize the high and low band components to be transmitted. [24]

In Adaptive Transform Coding (ATC), the input speech is blocked into frames of data and transformed by a symmetric discrete Fourier transform [33]. Each frame is represented by a set of transform coefficients, which are separately quantized and transmitted. At the receiver, the quantized coefficients are inversely transformed to produce a replica of the original input frame. [14]

The Code-Excited Linear Prediction (CELP) is one of a class of coders known as analysis-by-synthesis coders. In the actual encoding process, the encoder first buffers an
input speech frame of about 20 ms or so, and then performs linear prediction analysis on the buffered speech. During the analysis stage, it attempts to find the best parameter values so that the error between the input speech frame and its synthesized output frame at the decoder is minimized. These values are encoded and sent to the decoder. The decoder decodes and reproduces speech frame-by-frame. [16]

The Adaptive Predictive Coder (APC) is an ADPCM coder which uses two stages of prediction, a high-order short-term predictor based on the spectral envelope, and a long-term predictor based on the pitch or periodicity information. The prediction results in a small prediction error which is quantized by an adaptive step-size quantizer. The coefficients of the long term predictor are adapted in every 10 ms, where the coefficients of the short-term predictor are adapted every frame or every alternate frame. [33], [14]

When low bit rate encoding techniques are used, especially under imperfect communication conditions (e.g. errors in transmitting digitized voice, listeners with hearing impairments), intelligibility can be a serious issue. In these situations, the intent will not be to measure speech quality which will be quite low anyway, but rather to measure features that preserve information contrast (e.g. consonant in speech). Various speech intelligibility tests have been developed by audiologists to estimate a person's ability to understand conversational speech. Three commonly used intelligibility tests are phonetically balanced (PB) word tests, syllable tests, and sentence tests.

The phonetically balanced word tests consist of word lists in which the phonetic composition of all words in the lists are all equivalent and representative of everyday English speech. These tests utilize an open-set response format, i.e. the listener is not presented with a closed set of several alternatives of monosyllabic words for each test.
item. The set of possible responses to a test item is open and is limited only by the listener's vocabulary.

The syllable tests have a closed-set format and are offered as an alternative to the conventional open-set PB word tests. The advantages of the closed-set format include the elimination of examiner bias, ease of administration and simplified scoring techniques. Another advantage of syllable tests is the possibility to obtain a somewhat detailed picture of the type of errors made by the listener and not just an indication of the total number of errors made.

In an attempt to approximate everyday speech more closely, several intelligibility tests have been developed using sentences as the basic items. An advantage of speech tests using sentences over other material is that sentences approximate the spectral and contextual characteristics of connected discourse.
Chapter 2 Selection of a speech encoding method

Fig. 2.4 Relation between Al and various measures of speech intelligibility.

The relationships between articulation index and intelligibility score for a given group of talkers and listeners is presented in Fig. 2.4 [1]. These curves show that the intelligibility score, in percent correct, is highly dependent on the constraints placed upon the message being communicated. The greater the constraints, the higher the percent intelligibility score for a given articulation index. The constraints here refer to grammatical structure, contextual information found in sentences, or limitations in vocabulary size and syllabic length of words.
Chapter 2 Selection of a speech encoding method

2.3 Selecting a speech encoding technique for hearing impaired listeners

Based on the Mean Opinion Scores in Fig. 2.3, two relatively simple speech encoding techniques with high MOS, the $\mu$-law PCM and the CCITT G.722 algorithm, were selected as candidates for the proposed radio-frequency ALD. The $\mu$-law PCM coder has a bandwidth of 3.4 KHz and the CCITT G.722 algorithm has a bandwidth of 7 KHz. Under ideal listening conditions, according to Fig. 2.2 the bandwidth of the $\mu$-law and the CCITT G.722 algorithm has articulation indexes of 0.8 and 1.0 respectively. These algorithms should thus produce a high intelligibility score for most of the intelligibility tests in Fig. 2.4.

For the hearing impaired population, intelligibility rather than quality is the important factor in deciding which speech coder is acceptable. A sentence-based intelligibility test, the Revised Speech Perception in Noise (SPIN) test, is used to determine whether the two encoding techniques are acceptable for the hard of hearing population.

The SPIN test was originally developed by Kalikow et. al. [5] with the purpose of making a speech test which better reflects everyday listening conditions than the tests which use isolated words as stimuli. Each test item in the SPIN test is a sentence of five to eight words in length. There are ten forms in the SPIN test, and each form contains 50 sentences. For each of the 50 sentences in a form, half are high predictability sentences, which provide the listener with linguistic clues about the final word (e.g. The watchdog gave a warning GROWL). The other 25 sentences are low predictability sentences, which provide little or no information about the final word (e.g. I had not thought about the GROWL). Recordings of the test sentences on magnetic tape are provided by the
developer of the test. The listener is asked to repeat the last word of each sentence. The last word of the sentence is always a monosyllabic noun. The nouns used have word frequencies of from 5 to 150 per million words in the Thorndike-Lorge lists [9]. Each of these nouns is used twice, once in a high predictability sentence and once in the low predictability sentence. The background noise in the test is a babble of 12 voices on a second sound track. Kalikow et. al. generated the babble track by adding together tape recordings of 12 talkers reading aloud from continuous text. Each recorded test form is preceded by a calibration tone (1,000 Hz) to control the sound level output.

The SPIN test was revised by Bilger [25] in order to standardize its use for hard of hearing subjects. The test was administered to 128 hard of hearing subjects, with the purpose of equalizing the difficulty (i.e. score) for each form. From the original ten forms, eight equivalent forms known as the Revised SPIN test were constructed (Appendix A).

The score of the Revised SPIN test is known as the “percent hearing for speech” and is determined by the nomograph shown in Fig. 2.5. The construction of the nomograph was based on statistical analysis done on the test results of the 128 hard of hearing subjects [25]. Once the numbers of correctly identified high and low predictability items are determined, the nomograph can be used to determine the “percent hearing for speech” for that person (e.g. a high predictability score of 25 and a low predictability score of 22 will give a percent hearing of speech of 97.0). This percentage should fall in the acceptance region. If it does not, then the score probably underestimates the subject’s hearing for speech.
The acceptance region is the region where the high and low predictability item scores of a Revised SPIN test are acceptable. In general, scores on the diagonal (high predictability item score = low predictability item score) or below the diagonal on the nomograph are unacceptable; the exceptions are for pairs of high-low scores such as...
Chapter 2 Selection of a speech encoding method

25–24, 24–24, 1–1 or 0–1 indicating the subject has extremely good or extremely poor hearing. Based on the high-low pair distributions of the 128 hard of hearing subjects, regions above the diagonal that falls far away from the distribution are rejected. The acceptance region is the region between the lines drawn through the matrix. Percentages not rejected, but outside the acceptance region are left to the discretion of the audiologist.

The difference between a person’s scores on the high and low predictability sentences can be used as an indicator of how well that individual makes use of the context of a sentence. The scores on the low predictability sentences should reflect how well the peripheral auditory system (outer ear, middle ear and inner ear) processes speech, while the score for the high predictability sentences should reflect how well the encoded information about speech is used by the individual.

2.3.1 Procedures of the SPIN test

The purpose of the experiment is to determine whether the intelligibility of a hearing impaired listener will be reduced when the speech material is processed by digital encoding methods under quiet and noisy conditions. A magnetic recording of the eight Revised SPIN test forms was obtained from the U.B.C. School of Audiology and Speech Sciences. Magnetic recordings of the Revised SPIN Tests processed by the μ-law and the CCITT G.722 algorithms were made using the setup shown in Fig. 2.6.
Only the speech channel is processed by the speech coding methods. It is assumed that only a small amount of ambient noise will get into the ALD system in a realistic application. Ambient noise reaching the listener’s ear has a bandwidth which is not limited by the bandwidth of the ALD system.

Ten hearing impaired listeners participated in the experiment. The ten subjects were divided into two groups of five; one group was tested under quiet conditions and the other group under noisy conditions. In each group, unprocessed and processed sentences are randomly selected and given to each subject in a random order. The unprocessed sentences are from the original Revised SPIN Test and the processed sentences are from the tapes processed by the $\mu$-law and the CCITT G.722 algorithms. All the participating subjects had bilateral sensorineural hearing loss (hearing loss caused by damage to the inner ear or the auditory nerve). Prior to the Revised SPIN tests, each subject underwent a basic hearing screening test to determine their degree of hearing loss.

The basic tool for the assessment of the degree of hearing is the audiogram. An
audiogram is a chart used to record graphically the hearing threshold of an individual at different frequencies. The hearing threshold is typically defined as the lowest (softest) sound level needed for a person to detect the presence of a signal approximately 50% of the time. The audiogram plots the signal frequency against the hearing level in decibel hearing level (dBHL). The horizontal line at 0 dB represent normal hearing sensitivity for the average young adult. Results plotted on the audiogram can be used to classify the extent of hearing loss. Classification schemes using the pure tone audiogram are based on the fact that there is a strong relationship between the threshold for those frequencies known to be important for hearing speech (500, 1000, 2000 Hz) and the lowest level at which speech can be recognized accurately 50% of the time. Given the pure tone thresholds at 500, 1000, and 2000 Hz, the effect of the impairment can be estimated. This is accomplished by calculating the average (mean) loss for these three frequencies. A typical classification scheme is shown in Fig. 2.7 [10]. This scheme reflects the different classifications of hearing loss as well as the likely effects of the hearing loss on an individual’s ability to hear speech.
Each subject was tested individually while seated in a sound proof room (Fig. 2.8). Sound was administered to the subject through Grason-Stadler GSI-16 Speakers. Sound level control was accomplished using a Grason-Stadler GSI-16 audiometer and the calibration tone presented at the beginning of each Revised SPIN test form recording. Sound was presented at the subject’s most comfortable loudness (MCL) and the subjects were allowed to use their personal hearing aid if they wished. In administering the Revised SPIN test to these subjects, they were asked to respond to every item in the Revised SPIN Test. The 25–item Revised SPIN Practice Tape was used to accustom
Chapter 2 Selection of a speech encoding method

subjects to the procedure and to determine the subject’s MCL. Whenever a subject failed to respond promptly, the tape was stopped and the subject was asked to guess. Also, whenever the subject gave an uncertain response, the tape was stopped and the subject was asked to repeat the response or to spell the word. Each test session was taped, so that the test items, the subjects responses, and any interchange between the test administrator and the subject appeared on the tape. A second test administrator then scored all the tests.

2.3.2 Results in Quiet Conditions

In the quiet condition the babble track was not presented to the subjects. Three different recordings, one from the original Revised SPIN test and the other two from the tapes processed by the μ-law and the CCITT G.722 algorithm were randomly selected and given to each subject in a random order.

![Fig. 2.8 Test setup in quiet condition](image)

The subjects’ audiograms and MCLs are shown in Fig. 2.9. Their hearing impairment ranged from mild hearing loss to severe hearing loss according to Fig. 2.7. Except for subject 1, all subjects used their hearing aids in all of the three tests.
Chapter 2 Selection of a speech encoding method

Fig. 2.9 Audiograms of subjects taking Revised SPIN test in quiet condition.
Chapter 2 Selection of a speech encoding method

The results of the SPIN test in the quiet environment is shown in Fig. 2.10.

No perfect aggregate score was achieved by any subject, but most of them came close to 100% on high predictability items. As expected, scores on low predictability items were lower. Scores for the original Revised SPIN test recording and the two recordings processed by the two digital encoding methods varied for individual subjects. The means of the scores of the original Revised SPIN test recording and the two recordings processed by the two digital encoding methods are within ± 0.8% of each other (Table 2.3 on p. 37).

2.3.3 Results in Noisy Conditions

The noisy conditions were simulated by mixing the babble track with the speech track. The babble was set at 8 dBHL below the speech sound level, since this is the
Chapter 2 Selection of a speech encoding method

median speech to babble ratio encountered in a wide range of real life situations [25]. The set up shown in the Fig. 2.11 was used. Voice output from the tape recorder was passed through an FM wireless microphone system (Model Realistic # 32-1221) before being input to the audiometer. All other experimental procedures were the same as for the experiment in quiet condition.

The purpose of introducing the FM wireless microphone system in the experiment was in order to collect data for the evaluation of speech scrambling methods. As described in the Introduction, the speech scrambling aspect involves secure transmission of scrambled speech signals using the FM wireless microphone system.

Fig. 2.11 Test setup in noisy conditions, babble is presented at 8 dBHL below speech signal.

Ideally, the setup of Fig. 2.8 should have been used for the evaluation of noisy conditions. However, we have encountered difficulties in obtaining the agreement of hard of hearing subjects to participate in repeated experiments. Realizing that experiments will
Chapter 2 Selection of a speech encoding method

be needed to evaluate the performance of the speech coding scheme and the FM wireless microphone system, we have decided to combine the two experiments. The setup of Fig. 2.11 was used. Compared to the setup of Fig. 2.8, Fig. 2.11 shows the FM wireless microphone system.

The measured frequency response of the FM wireless microphone system has a bandwidth of approximately 7 KHz as shown in Fig. 2.12.

![Bandwidth of FM Transceiver](image)

**Fig. 2.12 Frequency response of the Realistic FM wireless microphone system (Model # 32-1221)**

The introduction of the FM wireless microphone system in the experiment limited all speech signal bandwidths to 7 KHz. This should not affect intelligibility test results as frequencies of a speech signal below 250 Hz or above 7000 Hz contributes very little to the speech intelligibility [12]. In addition, experiments in quiet conditions show that for hard of hearing people, limiting the speech material to 3 KHz by the $\mu$-law coder does not affect the intelligibility scores significantly compared to the original Revised SPIN test recording. Any change in the results of the Revised SPIN tests under noisy conditions is mainly due to the introduction of the babble noise.
Chapter 2 Selection of a speech encoding method

Fig. 2.13 Audiograms of subjects taking Revised SPIN test under noisy conditions.
Chapter 2 Selection of a speech encoding method

The subjects' audiograms and MCLs are shown in Fig. 2.13. Their hearing impairment ranged from mild hearing loss to severe hearing loss according to Fig. 2.7. Except subject 10, all subjects used their hearing aids in all of the three tests. The results of the Revised SPIN test in noisy conditions are shown in Fig. 2.14.

As expected, scores for low predictability items are lower than those for the high predictability items. Test scores of subject 6's CCITT G.722 recording (High score 19, Low score 18) and subject 7's original Revised SPIN recording (High score 22, Low score 17) fall outside the acceptance region of the Revised SPIN test nomograph [25], indicating that the percent-hearing-for-speech scores for these subjects probably underestimate their ability to hear speech. These two scores were not taken into account when the mean
percent-hearing-for-speech was calculated. Scores of the original Revised SPIN test and for the tests processed by the two digital encoding methods varied for individual subjects, but the means are within ± 2.3% of each other (Table 2.3).

2.3.4 Speech encoding techniques for the hearing impaired listeners

The results in Table 2.3 show that applying either the $\mu$-law or CCITT G.722 speech encoding techniques to the Revised SPIN Test forms does not reduce the intelligibility score significantly for hearing impaired persons.

<table>
<thead>
<tr>
<th></th>
<th>Original SPIN</th>
<th>CCITT G.722</th>
<th>$\mu$-law</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quiet Condition</td>
<td>96.3</td>
<td>95.9</td>
<td>96.7</td>
</tr>
<tr>
<td>Noisy Condition</td>
<td>89.1</td>
<td>89.5</td>
<td>91.4</td>
</tr>
</tbody>
</table>

Both encoding techniques produce a data stream at 64 Kbps. The logical choice would be the CCITT G.722 encoding technique, as it would give a larger bandwidth which could provide extra audio clues for hearing impaired listeners. However, a CCITT G.722 device in integrated circuit (IC) form is not currently available. In 1988, Philips offered the PCB 2322 chip that implements the CCITT G.722 algorithm. Unfortunately, this chip is no longer in production. Implementation of the G.722 algorithm is possible using digital signal processors, but this approach would result in a larger physical size and higher cost. Since our Revised SPIN test results show no significant difference in terms of intelligibility for hard of hearing persons between the two encoding techniques, and a $\mu$-law codec in IC form is readily available, the $\mu$-law codec was chosen for the radio-frequency ALD.
Chapter 3  Hardware Design & Implementation

3.1 Overall design

The development of the radio-frequency ALD is carried out with the aid of personal computers (PCs) as shown in Fig. 3.1. The Codec Interface performs voice digitization and recovery, as well as provides the timing signals for the PC to read and write to the Codec Interface.

In the radio-frequency ALD transmitter, the voice input is sampled and digitized by the Codec Interface. The PC reads the digitized voice data from the Codec Interface, and performs error detection/correction to ensure reliable data transmission. The PC then coordinates the flow of encoded data to the wireless spread spectrum modem which transmits the data using a spreading code known to authorized receivers.
In the radio-frequency ALD receiver, the received spread spectrum signal is demodulated. The PC performs error detection/correction on the demodulated data and coordinates the flow of decoded data to the Codec Interface. The Codec Interface performs voice recovery and reproduces the analog voice signal. The use of more than one radio-frequency ALD receiver enables voice broadcast from the radio-frequency ALD transmitter to be received by a group of users with knowledge of the de-spreading codes. If there are several transmitters in operation, authorized receivers can select the appropriate de-spreading codes to receive transmissions from the desired transmitter only.

In this part of the thesis the examination of the feasibility of transmitting voice using spread spectrum techniques is presented.

### 3.2 Spread Spectrum modem

The spread spectrum modem must be capable of transmitting the data from the Codec Interface of the transmitter to the receiver. Since the μ-law codec converts analog voice into digital data stream at 64 Kbps, this implies the spread spectrum modem must have a data rate of no less than 64 Kbps.

As data are transmitted over the radio-frequency link, errors will inevitably occur due to interference or noise present in the transmission environment. These errors affect the quality of the reproduced analog voice signal by the radio-frequency ALD receiver. Error correction or detection capability can be implemented in the radio-frequency ALD system to minimize the effect of data lost. The degree of error detection or correction capability selected will depend on the channel error rate of the spread spectrum modem. An analysis of the error correction or detection capability required for the radio-frequency
Chapter 3 Hardware Design & Implementation

ALD system and the required transmission rate of the radio-frequency spread spectrum modem will be discussed in Chapter 4.

There are chip sets available for spreading and de-spreading direct sequence data (e.g. STEL-1032, STEL-3310 from Stanford Telecom, OCI 100011-1 from O’Neill Communications Inc.). While relatively simple circuits can be used to digitally spread and de-spread the data, developing the radio frequency (RF) sections required for transmission at 902–928 MHz is considerably more involved. The 902–928 MHz band has been allocated for industrial, scientific and medical (ISM) applications in Canada and USA where license-free spread spectrum transmitters are allowed to operate. [4]

It was decided to investigate the feasibility of integrating commercially available spread spectrum modems in the design of the radio-frequency ALD system. In order to transmit the voice data output by the codec, the spread spectrum modem must be capable of transmitting and receiving data at a rate of at least 64 Kbps.

Table 3.1 Known commercially available spread spectrum modems.

<table>
<thead>
<tr>
<th>Manufacturer</th>
<th>model</th>
<th>spread spectrum technique</th>
<th>data rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>NCR</td>
<td>WaveLAN</td>
<td>Direct Sequence</td>
<td>2 Mbps</td>
</tr>
<tr>
<td>Intermec</td>
<td>Model 9181</td>
<td>Direct Sequence</td>
<td>256 Kps</td>
</tr>
<tr>
<td>Solid State Electronics Corp.</td>
<td>Model 5093</td>
<td>Direct Sequence</td>
<td>64 Kbps</td>
</tr>
<tr>
<td>Prism</td>
<td>RangeLAN</td>
<td>Direct Sequence</td>
<td>242 Kbps</td>
</tr>
<tr>
<td>O’Neil Communication Inc.</td>
<td>LAWN</td>
<td>Direct Sequence</td>
<td>19.2 Kbps</td>
</tr>
<tr>
<td>Telesystems</td>
<td>Arlan 650</td>
<td>Direct Sequence</td>
<td>1.35Mbps</td>
</tr>
</tbody>
</table>
A number of commercially available spread spectrum modems were reviewed, as shown in Table 3.1. None of them allow the user to change the spreading code directly, and some of them do not have the required data transmission rate. The Arlan 650\textsuperscript{TM}, manufactured by Telesystems, was selected for use in the radio-frequency ALD design because of its data rate and software compatibility with the TCP/IP protocols.

### 3.3 Arlan 650\textsuperscript{TM} characteristics

The Arlan 650\textsuperscript{TM} Wireless Network Card from Telesystems SLW Inc. installs directly into the PC-bus of any IBM PC/AT\textsuperscript{TM} or compatible. It contains an on-board spread spectrum radio transceiver. A TNC-type connector at the rear panel of the Arlan 650\textsuperscript{TM} Wireless Network card connects to an attached 8-inch half wave dipole antenna.

In Canada, Arlan 650\textsuperscript{TM} operates under DOC regulations allowing license-free use of spread spectrum transmitters in the 902–928 MHz band. The Arlan 650\textsuperscript{TM} has a maximum RF power output of 1 Watt. It can be set up to operate in one of the 13 channels and the data rates vary from 215 Kbps to 946 Kbps, as shown in Table 3.2. [30]
Table 3.2 Radio frequency and data rate of channels in Arlan 650™ [30]

<table>
<thead>
<tr>
<th>Channel</th>
<th>Rate kbps</th>
<th>Centre Frequency MHz</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>215</td>
<td>908</td>
</tr>
<tr>
<td>1</td>
<td>215</td>
<td>910</td>
</tr>
<tr>
<td>2</td>
<td>215</td>
<td>913</td>
</tr>
<tr>
<td>3</td>
<td>215</td>
<td>915</td>
</tr>
<tr>
<td>4</td>
<td>215</td>
<td>917</td>
</tr>
<tr>
<td>5</td>
<td>215</td>
<td>920</td>
</tr>
<tr>
<td>6</td>
<td>215</td>
<td>922</td>
</tr>
<tr>
<td>7</td>
<td>344</td>
<td>911</td>
</tr>
<tr>
<td>8</td>
<td>344</td>
<td>915</td>
</tr>
<tr>
<td>9</td>
<td>344</td>
<td>*919</td>
</tr>
<tr>
<td>10</td>
<td>630</td>
<td>915</td>
</tr>
<tr>
<td>11</td>
<td>860</td>
<td>915</td>
</tr>
<tr>
<td>12</td>
<td>946</td>
<td>915</td>
</tr>
<tr>
<td>13*</td>
<td>1050</td>
<td>915</td>
</tr>
<tr>
<td>14*</td>
<td>1350</td>
<td>915</td>
</tr>
</tbody>
</table>

* Current Canadian DOC regulations prohibit the use of Radio Channels 13 and 14. [30]

Fig. 3.2 shows the RF spectrum of Arlan 650™ transmitting data using Radio Channels 7 to 9. The range of the spectrum shown is from 900 MHz to 930 MHz, with a horizontal scale of 3 MHz per division. The bandwidth of each channel overlaps and interferes with adjacent channels. The direct sequence spread spectrum technique described in Section 1.4 is used to minimize interference between adjacent channels. Each channel uses a different spreading code to spread the bandwidth of data to be transmitted. By using different radio channels, the spreading code of the system can be changed indirectly.

In addition to the radio channel, an Arlan 650™ must have the same System Identifier (SID) to communicate with another Arlan 650™. There are over 8 million possible SID
settings which provide a good degree of privacy and security for the radio-frequency ALD.

The range of the Arlan 650™ in a given indoor environment depends the following factors.

- data rate (lower bit rate channels have an advantage over higher bit rate channels, since there is approximately a 6 to 7 dB decrease in receiver threshold as the data rate is increased from 200 Kbps to 1 Mbps [30].)

- the building material of the indoor environment and the number of obstacles (people, furniture, walls, partitions) in the direct path between the transmitter and the receiver. Floor to floor penetration also depends on the material used between the floors.

- type and placement of antenna. (Telesystem has an optional high gain Yagi antenna for the Arlan 650™ wireless network card.)

Typical range is 300 feet indoor and 1000 feet outdoor line-of-sight operation [30].

Software drivers are supplied by the manufacturer for operating the Arlan 650™ under Novell Netware™ 286 and 386. A Packet Driver is also available for operation with third party TCP/IP protocols.
Fig. 3.2 Radio frequency spectrum of Arlan 650™ Channel 7 to 9.
3.4 Codec Interface

The block diagram of the Codec Interface is shown in Fig. 3.3. The Codec Interface is designed so that it can be used in the radio-frequency ALD transmitter and the radio-frequency ALD receiver. In the transmitter, the Codec Interface passes digitized voice data from the $\mu$-law codec to the PC. In the receiver, the Codec Interface passes received digitized voice data from PC to the codec.

As the conventional PC serial port cannot handle serial data at 64 Kbps, the serial digitized voice data is passed to and from the PC via a digital signal processor board. The digital signal processor board also acts as a buffer between the PC and the Codec.
Interface. A TMS320C30 Processor Board from Spectrum Signal Processing [29] was used in the design. This is an IBM PC™ compatible plug-in board with two synchronous serial ports capable of input/output serial data at speeds of up to 8.3 Mbps. The PC reads the data received via the serial port of the TMS320C30 Processor Board, and writes data to this board for transmission via its serial port.

Serial port 0 and serial port 1 in the TMS320C30 Processor Board are totally independent. Each serial port can be configured to transmit and receive 8, 16, 24, or 32 bits of data per word. The clock for each serial port can be either internal or external. The signal description of the TMS320C30 Processor Board serial port is shown in Table 3.3.

<table>
<thead>
<tr>
<th>Signal</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLKX0/1</td>
<td>Serial port 0/1 transmit clock. This input clock signal determines the data output rate at DX0/1.</td>
</tr>
<tr>
<td>DX0/1</td>
<td>Serial port 0/1 data output.</td>
</tr>
<tr>
<td>FSX0/1</td>
<td>Frame synchronisation pulse for transmit. The input pulse initiates the data transmit process in DX0/1.</td>
</tr>
<tr>
<td>CLKR0/1</td>
<td>Serial port 0/1 receive clock. This input clock signal determines the data input rate at DR0/1.</td>
</tr>
<tr>
<td>DR0/1</td>
<td>Serial port 0/1 data input.</td>
</tr>
<tr>
<td>FSR0/1</td>
<td>Frame synchronisation pulse for receive. The input pulse initiates the receive data process in DR0/1.</td>
</tr>
</tbody>
</table>

In the radio-frequency ALD design, the serial port 0 of the TMS320C30 Processor Board is configured for fixed data-rate, burst mode, external timing, 8-bit word operation. Transfers of data are separated by periods of inactivity on the serial port. Each transfer
involves a single word, and is initiated by either a Frame Synchronization Transmit (FSX) pulse or a Frame Synchronization Receive (FSR) pulse as shown in Fig. 3.4. In the receive operation, FSR must be low during the last bit, or another transfer will be initiated.

![Serial Port Timing of TMS320C30 Processor Board](image)

Fig. 3.4 TMS320C30 Serial Port Timing [32]

Implementation of the Codec Interface is based on the Motorola MC14402 Codec-Filter PCM-Mono-Circuit [18]. Signal descriptions of the MC14402 that are important for interfacing with the TMS320C30 Processor Board are shown in Table 3.4.
Chapter 3 Hardware Design & Implementation

Table 3.4 Signal description of the MC14402 [18].

<table>
<thead>
<tr>
<th>Signal</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>TDC</td>
<td>TDC is a clock signal input to the MC14402 which determines the transmit data bit rate at TDD.</td>
</tr>
<tr>
<td>TDE</td>
<td>TDE is a 8 KHz input signal. The leading edge of TDE initiates the shifting out of an 8-bit word at TDD at a rate determined by the TDC clock signal.</td>
</tr>
<tr>
<td>TDD</td>
<td>TDD is the digital data output.</td>
</tr>
<tr>
<td>RDC</td>
<td>RDC is the receive data clock. It operates in sync with RCE and RDD to produce all receive data timing.</td>
</tr>
<tr>
<td>RCE</td>
<td>RCE is a 8 KHz input signal. The leading edge of RCE initiates the receiving of an 8-bit word at RDD. The received word is shifted in RDD at a rate determined by the RDC clock signal.</td>
</tr>
<tr>
<td>RDD</td>
<td>RDD is the digital data input.</td>
</tr>
</tbody>
</table>

In the radio-frequency ALD design, the MC14402 is configured to digitize and recover voice using the \( \mu \)-law format. Transfers of digitized voice data into and out of MC14402 is shown in Fig. 3.5. The leading edge of TDE initiates data transmission at TDD, the rate of data output being determined by the clock signal at TDC. The leading edge of RCE initiates data reception in RDD, the rate of data input being determined by the clock signal at RDC.
Chapter 3 Hardware Design & Implementation

Fig. 3.5 Transmit and receive timing diagram for MC14402 [18]

In the radio-frequency ALD transmitter, input voice is sampled by the MC14402 at a rate of 8 KHz and each sample is represented by an 8-bit word in the $\mu$-law format. Every 125 $\mu$s, an 8-bit word is transferred from the MC14402 to serial port 0 in the TMS320C30 Processor Board at a rate of 128 Kbps. Transfer of the 8-bit words are separated by periods of inactivity of the serial port. The signal timing necessary for the data transfer is shown in Fig. 3.6. The data transfer rate is controlled by a 128 KHz clock signal CK2 input to TDC of the MC14402 and CLKX0 of serial port 0 in the TMS320C30 Processor Board.

Data transfer is initiated by the leading edges of a 8 KHz clock CK3 inputs to FSR0 of the serial port 0 in the TMS320C30 Processor Board, which prepares the serial port to accept the 8-bit word at the next CK2 clock pulse. At the next CK2 clock pulse, the leading edge of CK1, a one clock cycle delay of CK3 inputs to TDE initiates the 8-bit word transfer from the MC14402 to serial port 0 in the TMS320C30 Processor Board.
In the radio-frequency ALD receiver, received data are passed from the serial port to the MC14402 and converted back to an analog voice signal. Every 125 μs, an 8-bit word is transferred from serial port 0 of the TMS320C30 Processor Board to the MC14402. Data transfer occurs in a similar way as in the radio-frequency ALD transmitter. The data transfer rate is also controlled by the 128 KHz clock signal CK2 input to TDE of the MC14402 and CLKR0 of the serial port 0 in the TMS320C30 Processor Board.

Data transfer is initiated by the leading edge of an 8 KHz clock CK3 inputs to FSX0 of the serial port 0 in the TMS320C30 Processor Board, which prepares the serial port to transmit the 8-bit word at the next CK2 clock pulse. At the next CK2 clock pulse, the leading edge of CK1 input to RCE initiates the 8-bit word transfer from the serial port 0 of the TMS320C30 Processor Board to the MC14402.

Implementation of the timing signal, CK1, CK2, and CK3 is shown in Fig. 3.7.
A crystal CRY1 is used to generate a 4.096 MHz clock signal for the flip-flop U2:B. The flip-flop acts as a divide-by-2 device. The 2.048MHz clock output by U2:B is input to U3, a 14-Stage Ripple-Carry Binary Counter/Divider. The Q4 output of the counter divides the 2.048 MHz input clock by $2^4$, generating the 128 KHz signal CK2. The Q8 output of the counter divides the 2.048 MHz input clock by $2^8$ generating the 8 KHz signal CK3. The flip-flop U2:A is used to delay CK3 by one CK2 clock cycle to generate CK1.

Connection of the MC14402 to the three clock signals and the serial port 0 of the TMS320C30 Processor Board is shown in Fig. 3.8. The analog and digital inputs and outputs of the Codec Interface in Fig. 3.3 is also shown in Fig. 3.8. Analog input voice is amplified by the operational amplifier U4:A before being passed to the MC14402 and converted to digital data. Received digitized voice data is converted back to analog voice by the MC14402.
Chapter 3 Hardware Design & Implementation

Fig. 3.8 MC14402 to TMS320C30 Processor Board interface circuit.
Chapter 4  Software Design & Implementation

4.1  Overall design

The software design was divided in two parts as shown in Fig. 4.1; one part controls the TMS320C30 Processor Board, and the other controls the PC. Data exchanges between the PC and the TMS320C30 Processor Board are achieved using the dual-access memory on the TMS320C30 Processor Board. In the radio-frequency ALD transmitter, the software in the TMS320C30 Processor Board reads the voice data from the Codec Interface, puts it in the dual-access memory and groups the voice data into a fixed length voice packet. The software in the PC reads the voice packet from the dual-access memory, and encapsulates the voice packet with suitable protocol headers for the Arlan 650™ wireless network card to transmit.

![Diagram of software design](image)

Fig. 4.1  Overall software design of the radio-frequency ALD system.
In the radio-frequency ALD receiver, the software in the PC reads the data received from the Arlan 650\textsuperscript{TM}, extracts the voice packet and puts it in the dual-access memory. The software in the TMS320C30 Processor Board reads the voice packet from the dual memory and writes the voice data to the Codec Interface.

The Arlan 650\textsuperscript{TM} wireless network card supports the Novell SPX/IPX\textsuperscript{TM} protocol and the Internet TCP/IP protocol. The TCP/IP and the SPX/IPX\textsuperscript{TM} protocols are rules that co-ordinate exchange of information between connected computers. While SPX/IPX\textsuperscript{TM} is developed by Novell Inc., TCP/IP is an open system and its specifications are freely available. More importantly, TCP/IP is designed to facilitate communication between machines with diverse hardware architectures, to use almost any packet switched network hardware, and to accommodate multiple computer operating systems. Thus, anyone can write the software needed to communicate across different computer networks. For this reason, the TCP/IP protocol is used to transmit the voice packets in the radio-frequency ALD system.

4.2 TCP/IP protocol

TCP/IP is a set of protocols that allow computers from different vendors to share resources across connected networks e.g. transferring files via the file transfer protocol (FTP), remote login via the network terminal protocol (TELNET) or sending electronic mail. They are sets of rules that co-ordinate the exchange of messages between computers and make the exchange more efficient. The most accurate name for the set of protocols is the "Internet protocol suite". Transmission Control Protocol (TCP) and Internet Protocol (IP) are two protocols in the suite needed for many applications e.g. TELNET and FTP.
Because TCP and IP are the best known of the protocols, it is common to use the term TCP/IP to refer to the whole family of protocols.

The general concept of data transmission in the TCP/IP network is described in the following section to illustrate the software implementation of the RF ALD.

Internet is a packet-switching network, where information in the network is transmitted in small segments, known as packets. The TCP/IP protocols define the format of these packets including the origin of the packet, the length of packet, and the type of packet, as well as the way computers on the network are to receive and re-transmit packets.

TCP/IP provides three sets of services as shown in Fig. 4.3. At the lowest level, a connectionless delivery service provides a foundation on which everything rests. At the next level, a reliable transport service provides a higher level platform on which applications depend.

![Diagram](image)

**Fig. 4.2** The three level of service provided by TCP/IP.

Connectionless delivery is an abstraction of the service that most packet-switching networks offer. This service routes packet from one node to another based on address...
information carried in the packet. (a node on the Internet is any device that understands TCP/IP protocols) Because the connectionless service routes each packet separately, it does not guarantee reliable, in-order delivery. Connectionless packet delivery is the basis for all Internet services and makes the TCP/IP protocols adaptable to a wide range of network hardware.

Most applications need more than just connectionless packet delivery because they require the communication software to recover automatically from transmission errors, lost packets or failures of intermediate switches along the path of the sender and receiver. The reliable transport service handles such problems. It allows an application on one node to establish a 'connection' with an application on another node, and then to send a large volume of data across the connection as if it were a permanent, direct hardware connection. In practice, the service divides the stream of data into small packets and make use of the connectionless packet delivery service to send them one at a time, and waiting for the receiving host (small or large computers) to acknowledge reception of packets.

The TCP/IP services can be modelled with four functional layers that build on a fifth layer of hardware. The model shown in Fig. 4.4 presents a framework to describe protocol characteristics and functions. Fig. 4.4 also shows some of the TCP/IP protocols in relation to the reference model.
Each layer in the model has various functions, which are independent of the other layers. Each layer, however, expects to receive certain services from the layer beneath it, and each layer provides certain services to the layer above it. Each layer on the transmit host communicates with that same layer (peer layer) on the destination host. The model is generally used as a framework to describe the functions and characteristics of different protocols in the Internet.

At the highest level, users invoke application programs that access services available across the Internet. An application interacts with the transport level protocol(s) to send or receive data. Each application program chooses the style of transport needed, which can be either a sequence of individual messages or a continuous stream of bytes. The application program passes the data in the required form to the transport level for delivery.

The prime duty of the transport layer is to provide communication from one appli-
cation program to another. The transport layer may regulate flow of information. It may also provide reliable transport service, ensuring that data arrives without error and in sequence. The transport layer may accept data from several user programs and send them to the next lower level. To do so, it adds additional information to each packet, including codes that identify which application program sent it and which application program should receive it. The receiving node uses the destination code to identify the application program to which it should be delivered.

The Internet layer handles communication from one node to another. It accepts a packet from the transport layer along with an IP address to which the packet should be sent. The basic transfer unit in the Internet is an Internet datagram, sometimes also known as an IP datagram. The Internet layer accepts a packet from the transport layer and encapsulates the packet in an IP datagram, fills in the datagram header, uses the routing algorithm to determine whether to deliver the datagram directly or send the datagram to a gateway (gateways are dedicated computers that are attached to two or more networks and forward data from one network to another), and passes the datagram to the appropriate network interface for transmission. The Internet layer also handles incoming datagrams, checks their validity, and uses the routing algorithm to determine whether the datagram should be processed locally or forwarded. For datagrams addressed to the local host, software in the Internet layer deletes the datagram header and chooses the appropriate transport protocol that will handle the packet. The Internet layer does not limit datagrams to a small size nor does it guarantee that large datagrams will be delivered without fragmentation. Fragmentation is a process in the transmitting node of dividing a large datagram into several pieces. The receiving node will re-assemble the
datagram based on the information in the fragments.

At the network interface layer, the nodes on a network communicate with other nodes on the network using physical addresses specific to that network. Each node has a unique physical address for the hardware device that connects it to the network. The IP address for a node is a logical address, which is independent of the physical address. Because IP addresses are not dependent on any particular network interface hardware, they can be used to send datagrams from one network to another network. The network interface layer encapsulates an IP datagram in a network interface frame, maps the IP address of a node or a gateway on the same network into a physical address, and uses the network hardware interface to delivery the network interface frame.

An IP address is mapped onto a physical address using the Address Resolution Protocol (ARP). The sending node broadcasts an ARP packet containing a IP address. The node with the IP address sends its physical address back to the requesting node.

To speed packet transmissions and reduced the number of ARP request, each time the node broadcasts an ARP request and receives a response, it creates an entry in the address resolution cache in the sending node. The entry maps the IP address to the physical address. When the node needs to send another IP datagram, it looks up the IP address in its cache. If it finds that IP address, the node uses the corresponding physical address for its network interface frame. The node broadcasts an ARP only if the IP address is not in its cache.

Fig. 4.5 shows how messages from a host are transferred across different networks. At the sending host, user data are presented by a user application at the upper (application) layer. Each layer encapsulates its protocol control information (header) to the user data
and passes its header and user data to the next lower layer which repeats the process.

Fig. 4.4 Communication process in the TCP/IP protocols.

In the Internet layer, the sending host has to determine whether the destination IP address in the datagram is within the same physical network or on other network. If the destination IP address is not on the same network, the host sends the datagram to a gateway. The sending host has a table of IP addresses for one or more hosts that serve as gateways to other networks. It looks for the IP address of a gateway that leads to the destination network, maps the IP address of the gateway to a physical address using ARP, and sends a network interface frame containing the IP datagram to the gateway. When the gateway receives the IP datagram, it uses the destination IP address in the datagram
to send the message to its destination host using an appropriate network interface frame.

If the destination IP address is within a single network, the host can map the destination IP address into a physical address using ARP, and sends the IP datagram in a network interface frame to its destination directly.

The fully encapsulated data are transported across the physical networks to the receiving host. Here the process is reversed. The data go from the lower layers to the upper layers, and the header created by the transmitting peer layer is used by the receiving peer layer to invoke a service function for the transmitting site and the upper layers of the receiving site. As the data go up through the layers, the headers are stripped away after they have been used.

4.3 Voice packet transmission in the radio-frequency ALD system

In the radio-frequency ALD design, the radio-frequency ALD transmitter and receivers are grouped in a simple physical network by the use of Arlan 650™ wireless network cards over which TCP/IP operates. The Arlan 650™ wireless network card is the network interface hardware that connects the radio-frequency ALD transmitter and receiver to the same network. Access to the network is governed by the Institute of Electrical and Electronics Engineers 802.3 Media Access Control (IEEE 802.3 MAC) procedure [3] in the network interface layer which uses a Ethernet frame for communication between different nodes.

The communication process between the radio-frequency ALD transmitter and the radio-frequency ALD receiver can be modelled as shown in Fig. 4.6.
The application layer reads the voice packet from the dual-memory in the TMS320C30 Processor Board and passes it to the transport layer on the radio-frequency ALD transmitter. To reduce overhead in the packet, the transport layer adds no header to the voice packet. The transport layer ensures that the voice packet from the application layer is passed to the internet layer on a first-come-first-out basis. The time interval between transmission depends on the length of the voice packet (e.g. if the voice packet size is 64 bytes, the transmitter needs to transmit once every 8 ms to make up the codec rate of 64 Kbps; if the voice packet size is 32 bytes, the transmitter needs to transmit once every 4 ms; etc.). The Internet layer creates an IP datagram with a data portion containing...
the voice packet and passes the IP datagram down to the network interface layer. The network interface layer encapsulates the IP datagram in a Ethernet frame and passes that to the Arlan 650™ network card for transmission.

At the radio-frequency ALD receiver, the network interface layer computes the checksum of the Ethernet frame. If the checksum contained in the Ethernet frame does not match the checksum computed by the network interface layer, it discards the frame. If the checksums match, the network interface layer passes the IP datagram to the internet layer. The internet layer computes the checksum in the datagram header. If the checksum contained in the header does not match the checksum computed by the internet layer, it discards the IP datagram. If the checksums match, the internet layer passes the voice packet to the transport layer. The transport layer ensures a voice packet is passed to the application layer in a timely order. If the voice packet is lost during transmission, a simple packet replacement strategy is used to recover the lost voice data. The application layer writes the received voice packet in the dual-memory of the TMS320C30 Processor Board.

The structure of the IP datagram is shown in Fig. 4.7. An IP datagram is divided into a header and a data area. The IP datagram does not specify the format of the data area, it can be used to transport any data of length up to 65,515 bytes.
Chapter 4 Software Design & Implementation

Fig. 4.6 Structure of an IP datagram.

The fields in the IP header have the following meaning:

Version Number (VERS) is a 4-bit field which specifies the version number of the IP. It is used to verify that the sender, receiver, and any gateway in between them agree on the format of the datagram. The version used in the radio-frequency ALD implementation is 4.

Length (HLEN) specifies the length of the IP protocol header in 32-bit words. The minimum IP protocol header contains five words (20 bytes). The length of the protocol header may be increased by the addition of optional fields, but the exact length must be known for the purpose of interpretation. In the radio-frequency ALD implementation, the minimum length (20 bytes) is used.

Type of service (SERVICE TYPE) is a 8-bit field that specifies how the datagram should be handled. It is broken down to five subfields as shown in Fig. 4.8.
Fig. 4.7 The five subfields that comprise the Type of Service field.

Precedence indicates the relative importance of the datagram from 0 (routine) to 7 (network control), allowing the sender to indicate the importance of each IP datagram. Bits D, T, and R specify the type of transport the IP datagram desires. When set (1), the D bit is a request for low delay, the T bit is a request for high throughput, and the R bit is a request for high reliability. In practice, the value 0 is always used since most host and gateway software ignore the Type of Service field. A value of 0 is used in the field in the radio-frequency ALD implementation, and the receiver ignores the content in this field.

Total Length contains the length of the datagram including the header in bytes. This entry is used to establish the data length. This field allows the length of a datagram to be up to $2^{16} - 1$ or 65,535 bytes. In the radio-frequency ALD implementation, the value in this field depends on the size of the voice packet.

Identification is a 16-bit value assigned by the sender to aid in assembling the fragments of a datagram. Since fragmentation of IP datagram is not expected in the radio-frequency ALD application, a value of 0 is used in this field.

The field following Identification is a 3-bit value, which controls the handling of datagrams in the case of fragmentation. The first bit is unused, the second and third bits are DF (Don’t fragment) and MF (More Fragments). If DF bit is set (1), the IP datagram
is not fragmented under any circumstances, even if it can no longer be forwarded and must be discarded. The MF bit shows whether or not the IP datagram is followed by more sub-packets (0 indicates no more sub-packet, 1 indicates more sub-packets). Since fragmentation of IP datagram is not expected in the radio-frequency ALD application, a value of 0 is used in DF and MF.

Fig. 4.9 shows the content of various fields in a IP datagram during the course of fragmentation process so that the resulting IP datagram would adapt to a network with a maximum packet size of 128 bytes.

If the MF bit is set, the fragment offset specifies the offset in the original datagram of the data being carried in the fragment. It is measured in units of 8 bytes, starting at offset zero for the first fragment. The receiving host can use this information to re-assemble the original message correctly. Since fragmentation of IP datagram is not expected in the radio-frequency ALD application, a value of 0 is used in this field.
Time to Live specifies how long the datagram may remain in the network before it is discarded. The time to live is usually equal to the maximum number of gateways that a datagram may pass through. Each gateway along the path from source to destination is required to decrement the Time to Live field by 1 when it processes the datagram header. If this field is zero, the datagram must be discarded by the current gateway. This prevents a datagram from circulating endlessly in the network. In the radio-frequency ALD design, a value of 100 is used in this field.

Protocol contains the ID of the transport protocol to which the datagram has to be handed over. An arbitrary value of 35, not used by other transport protocols, is used in this field for the radio-frequency ALD application.

Header Checksum contains the checksum for the protocol header fields. It prevents node or host from working with false data. For efficiency, the user data in the IP datagram is not checked. The Internet Checksum is formed by treating the entire header as a sequence of 16-bit integers (starting from field Vers), adding them together using one's complement arithmetic, and then taking the one's complement of the result. For the purpose of computing the checksum, the Header Checksum field is initially assumed to contain zero.

Source and Destination Address are 32-bit Internet addresses which provide an unambiguous description of the access to a host in a network. Each IP address is divided into two parts: a network portion which identifies the network, and a host portion which identifies the node. This division can fall at one of three locations within the 32-bit address, corresponding to the three Internet address classes: Class A, Class B and Class C as shown in Fig. 4.10. Regardless of the address class, all nodes on any single network
share the same network portion, and each node has a unique host portion.

Fig. 4.9 IP address classes.

Given an Internet address, its class can be determined from the three high-order bits. Class A addresses devote 7 bits to the network portion and 24 bits to the host portion. Class B addresses devote 14 bits to the network portion and 16 bits to the host portion. Class C addresses devote 21 bits to the network portion and 8 bits to the host portion.

Internet addresses are usually written as four decimal integers separated by decimal points, where each integer gives the value of one byte of the Internet address. Thus the 32-bit internet address

10001001 01010010 00111001 00100011

is written as 137.82.57.35.

The Internet addressing rules reserve the following type of Internet addresses for special purposes:

- Network addresses are internet addresses in which the host portion is set to all zeros. These are addresses of networks rather than nodes on a network. By convention, no node is ever assigned a host portion consisting of all zeros.
• Broadcast addresses are addresses in which the host portion is set to all ones. A broadcast is destined for every node on the network. By convention, no node is ever assigned a host portion consisting of all ones.

• Addresses in which the network portion consists of all zeros means this network. Using network address 0 is important in those cases where a host wants to communicate over a network but does not know the network IP address. It allows the host to communicate temporarily, but once the host learns its correct network and IP address, it must not use network address 0.

• The class A network address 127 is reserved for "loopback" and is designed for testing inter-process communication on the local host. When any host uses the loopback address to send data, the protocol software in the host returns the datagram without sending it across any network.

In the radio-frequency ALD implementation, each node is assigned an arbitrary Internet address. A broadcast address is put in the datagram when a voice packet is transmitted from the radio-frequency ALD transmitter.

The IP protocol header is extended to include options for rarely used fields such as Time stamp, Security, etc, so that the IP protocol header is kept as small as possible [2]. The number of 8-bit words in an IP header is always a multiple of 4, padding character must be inserted in the option field if necessary.

In the radio-frequency ALD transmitter, each voice packet is put in an IP datagram and the IP datagram is encapsulated in an Ethernet frame and transmitted to the radio-frequency ALD receivers through the Arlan 650™ wireless network card.

The structure of an Ethernet frame is shown in Fig. 4.11. Ethernet frames are of
variable length, with no frame smaller than 64 bytes or larger than 1518 bytes. The minimum frame size is specified by the IEEE 802.3 standard and is required for correct protocol operation.

The preamble to an Ethernet frame is a stream of 7 bytes of alternating 0s and 1s which serves to synchronize the receiving station.

The Start Frame Delimiter (SFD) field immediately follows the preamble pattern. It contains a fixed 8–bit sequence 10101011 that indicates the start of a frame.

The source and destination addresses field in the Ethernet frame are 48-bit long. The source address field specifies the Ethernet address of the sending network interface hardware. Each Ethernet address is fixed in every Ethernet network interface hardware. Since the IEEE distributes numbers to the manufacturers of Ethernet network interface hardware, it is certain that every station world-wide has a unique address. Ethernet frame can be sent to a specific Ethernet network interface hardware or to all Ethernet network interface hardware actively connected to the physical network. This is done by putting a specific Ethernet address or the Broadcast address in the destination address field. All ones in the 48–bit long Ethernet address is defined as the Broadcast address.
In the radio-frequency ALD implementation, the radio-frequency ALD transmitter uses its own Ethernet address in the source address field and the broadcast address is put in the destination address field when a voice packet is transmitted.

The Protocol type field contains a 2-byte integer that identifies the type of data carried in the frame. Thus, various protocol layers above the network interface layer can use this information to determine which protocol should be used to process the data in the frame. In the radio-frequency ALD implementation, the higher level protocol is IP and 8 is used in that field.

The Frame data contains an arbitrary sequence of data of length from 38 up to 1492 bytes. If less data are to be sent, the data field is extended by appending extra bytes (pads) to satisfy the minimum frame size requirement. In the radio-frequency ALD implementation, an IP datagram containing the voice packet is put in this field.

After the data section comes a Frame Check Sequence field. It contains a 32-bit cyclic redundancy check (CRC) value which helps the interface to detect transmission errors. The value is computed as a function of the content of the Destination Address, Source Address, Protocol type, Frame data and pads (if exist). The encoding is defined by the following generating polynomial.

\[ G(x) = x^{32} + x^{26} + x^{23} + x^{22} + x^{16} + x^{12} + x^{11} + x^{10} + x^{8} + x^{7} + x^{5} + x^{4} + x^{2} + x + 1 \]

The CRC value corresponding to a given frame is computed by the following procedure:

1. The first 32 bits of the frame are complemented.
2. The n bits of the frame are then considered to be the coefficients of a polynomial \( M(x) \) of degree n-1. (The first bit of the Destination Address corresponds to the \( x^{n-1} \)
term and the last bit of the data field corresponds to the $x^0$ term.

3. $M(x)$ is multiplied by $x^{32}$ and divided by $G(x)$, producing a remainder $R(x)$ of degree $< 31$.

4. The coefficients of $R(x)$ are considered to be a 32-bit sequence.

5. The bit sequence is complemented and the result is the CRC.

The sender computes the CRC as a function of the Ethernet frame, and the receiver recomputes the CRC to verify that the frame has been received intact.

4.4 Voice packet size in the radio-frequency ALD

The voice packet size plays a significant role in the radio-frequency ALD design. It has a significant effect on the end-to-end transmission delay of the system, and affects the quality of speech perceived by the listener when a packet loss occurs. Its choice is also influenced by a number of parameters in the network such as codec data rate, header overhead and network data rate. Their effects ought to be considered in order to determine an optimal packet length for the radio-frequency ALD.

The transmission delay depends on the codec data rate, the packet size, the network data transmission rate, and the header overhead. As the network data transmission rate, header overhead and codec data rate are fixed in the radio-frequency ALD design, the transmission delay depends mainly on the voice packet size. This delay occurs because before transmission can occur, the transmitter has to buffer the voice data until a voice packet size is full. When a speaker communicates with a listener through the radio-frequency ALD, and expects a response from the listener, the speaker will expect the response to come within a time-width known as the expectation time window. If the
transmission delay is larger than this expected time-width, the speaker will notice the delay. This could adversely affect the efficiency of information exchange between the speaker and the listener. Studies in the telephone industry have found that a transmission delay in the range of 100 - 200 ms is acceptable [6]. Given the data rate of the \( \mu \)-law codec is 64 Kbps, that would make 800 bytes (corresponding to 100 ms) the maximum voice packet size.

The effect of packet losses in terms of their size is categorized by Jayant and Christensen [22] as shown in Table 4.1.

<table>
<thead>
<tr>
<th>Voice packet size (ms)</th>
<th>≤ 4</th>
<th>16-32</th>
<th>≥ 64</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nature of distortion</td>
<td>crackles</td>
<td>glitches</td>
<td>phoneme losses</td>
</tr>
</tbody>
</table>

Jayant and Christensen [22] also suggest that the shortest possible phoneme (i.e. the smallest distinguishable unit of speech) is roughly 20 ms, and when the voice packet length is greater than 20 ms, the probability of totally losing a phoneme will increase rapidly. Based on their results, it is desirable to make the voice packet as short as possible in order to minimize the effects of packet loss and transmission delay.

However, as the voice packet size decreases, the number of voice packets needed to be transmitted per second has to increase in order to sustain the 64 Kbps codec data rate. Since there is a fixed amount of protocol overhead associated with each voice packet, as the voice packet transmission rate increases, the minimum network transmission rate required will also increase. The overhead for each voice packet is shown in Fig. 4.12, where \( L \) is the voice packet size in bytes.
Fig. 4.11 Overhead accompanying each voice packet.

The packet overhead is 46 bytes in the radio-frequency ALD design.

For $L \geq 18$, the minimum data transmission rate required, $R$, in bits per second is given by:

$$R = \text{Frame size (byte)} \times 8 \times \text{Number of voice packet transmission per second}$$

$$R = (46 + L) \times 8 \times \left( \frac{64,000}{8L} \right)$$

$$R = (46 + L) \times 64,000 \div L \text{ (bps)}$$

For $L < 18$, the Frame size is restricted to the minimum Ethernet frame size (64 bytes), $R$ is given by:

$$R = 64 \times \left( \frac{64,000}{L} \right)$$

$$R = 4,096,000 \div L \text{ (bps)}$$

The plot of minimum data transmission rate versus the voice packet size is shown in Fig. 4.13. The data rate of the $\mu$-law codec and the data transmission rate of Channel 0–6 in the Arlan 650$^\text{TM}$ wireless network card is also shown in the diagram.
As the packet size decreases, the data rate required to deliver the voice packet increases. Channel 0–6 in the Arlan 650™ wireless network card has the lowest data transmission rate (215 Kbps) among the 13 available channels. Since it is desirable for the radio-frequency ALD to be able to transmit voice packets in all 13 available channels, the voice packet size selected should be greater or equal to 20 bytes (2.5 ms) which requires a rate of 211.2 Kbps to transmit. Although the radio-frequency ALD should be capable of supporting voice transmission using 2.5 ms voice packet, it is highly inefficient as the overhead (46 bytes) accompanying the voice packet is more than double the length

Fig. 4.12 Data transmission rate versus packet size.
of the voice packet.

In order to minimize both the transmission delay and the perceptual effect of lost packet on the voice quality at the receiver, packets should be as short as possible. On the other hand, in order to maintain high channel utilization, it is desirable to keep the size of the packet as large as possible.

As a compromise, the radio-frequency ALD uses a 8 ms (64 bytes) voice packet. It is shorter than the shortest possible phoneme (20 ms) observed by Jayant and Christensen [22], but it is longer than the overhead of the packet (46 bytes). The required transmission rate is 110 Kbps, which is well within the capability of the Arlan 650™ wireless network card.

4.5 Lost packet replacement strategies

A common feature of packet-switching communication systems is that they cannot guarantee accurate and prompt delivery of every packet. In large networks, this could be due to network congestion or transmission impairments. In the radio-frequency ALD, this is mainly caused by transmission impairments which lead to occasional, random packet losses.

Packet losses cause distortion by introducing gaps in the speech sequence. The perceptual effect of such losses depends on the voice encoding scheme, packet size, the packet loss rate and the speech segment affected. Distortion may vary from negligible during silent period to unacceptable during high level voiced speech.

Speech transmission in a packet switching network can tolerate some loss of packets without an adverse effect on the quality of the received speech perceived by the lis-
Chapter 4 Software Design & Implementation

tener. Different techniques have developed to replace lost PCM voice packets in packet switching networks [23], [7], [22]. They all aim to increase the tolerance of the packet voice system to missing packets. The choice of a particular packet replacement technique will depend on the packet loss rate and signal processing power of the system. Two relatively simple methods as shown in Fig. 4.14 were considered in the design of the radio-frequency ALD system.

![Utterance of 'ga' with no packet loss (389 ms)](image1)

![Loss of 8ms packet, silence substitution applied](image2)

![Loss of 8 ms packet, packet repetition applied](image3)

Fig. 4.13 Packet replacement techniques for the radio-frequency ALD.

The simplest way of dealing with the gaps caused by packet losses is known as silence substitution or zero-stuffing. It requires no signal processing at the transmitter
or receiver. Every lost packet is treated as a silent interval in the transmitted speech. There is no published information on the effect of silence substitution on 8 ms voice packets. For 16 ms voice packets, silence substitution is tolerable (MOS >3.5) only for small packet loss rates (2 percent maximum) [23].

The next simplest technique is known as packet repetition. It requires the receiver to store the contents of the most recently received packet. When one or more subsequent packets are missing, the receiver sends this stored information to the PCM decoder. This approach is more attractive because it is likely that the missing packet will resemble the immediately preceding packet. Unfortunately, there is no published information on the effect of packet repetition on 8 ms voice packets. For 16 ms packets, this technique can extend the maximum tolerable packet loss rate from 2 percent to 5 percent [23].

Informal listening with the radio-frequency ALD system indicates that the effect of packet loss is not noticeable using either silence substitution or packet repetition replacement. It is reported that the effect of a 1 percent packet loss rate with 16 ms packet using silence substitution is noticeable only to critical listeners [22]. For the radio-frequency ALD system, the measured packet loss rate under normal circumstances is only $3.05 \times 10^{-4}$ (see Table 5.1 on p. 84), and the radio-frequency ALD system also uses a shorter voice packet (8 ms instead of 16 ms). The effect of packet loss should be minimal for most listeners.

There are other more sophisticated techniques which require more complicated signal processing that can extend the maximum tolerable packet loss rate to approximately 10 percent [23], [22]. Application of these packet replacement techniques is not necessary given the low packet loss rate and the small packet size used.
Chapter 5 Evaluation of the radio-frequency ALD

5.1 Packet error rates of the radio-frequency ALD

The packet error rate of the radio-frequency ALD achieved in a given environment will depend on the transmission conditions, location of the antenna, the type of building and the number of obstacles (people, furniture, walls, partitions, etc.) in the direct path between the transmitter and the receiver. Since these are variables it is difficult to predict how well the radio-frequency ALD will operate in any specific situation.

To find the packet error rates of the radio-frequency ALD, we measured the packet loss rate in different environments. The tests were carried out in the Hector Macleod Building — a large multiuse building on the UBC campus.

In the packet error measurement, a radio-frequency ALD transmitter and receiver arrangement was simulated by using Channel 12 of the Arlan 650™ wireless network card.† This channel has the highest data rate (946 Kbps) and is more susceptible to packet loss than other lower bit rate channels [30]. Instead of 64-byte voice data (corresponding to 8 ms of voice data), 64-byte sequences of data known to the receiver are transmitted every 8 ms. The receiver compares the received sequences with the expected sequences, and records the packets lost during transmission.

† Note that current Canadian DOC regulations prohibit the use of Channel 13 and 14 [30].
The packet error measurements were carried out with the transmitter and the receiver placed in line-of-sight at four different locations in the Hector Macleod Building as shown in Fig. 5.1 and Fig. 5.2.

- At either end of the short corridor on 4th Floor (outside Room 459 and Room 441).
- At either end of the long corridor on 4th Floor (outside Room 439 and Room 402).
- In the Communication Laboratory. (Room 458; approx. 30’ x 40’).
- In a lecture room (Room 228; approx. 37’ x 60’).

Measurements in different locations took place during the period of June 28 to July 17, 1993. Fig 5.3 shows the cumulative number of erroneous packets plotted against the number of packets transmitted for seven measurements.

The packets are transmitted at 8ms intervals, and the number of packets transmitted is related to the time elapsed since the start of the measurement by:

\[
\text{Time elapsed} = (\text{number of packet transmitted}) \times 8 \text{ ms}
\]

The seven measurements in Fig. 5.1 lasted from 19 hrs. 52 min. (Measurement # 6) to just over 48 hrs (Measurement # 5). The packet loss rate varied from location to location but it was more or less constant throughout each measurement.
Fig. 5.1 Location of radio-frequency ALD transmitter and receiver in Measurement #1, 2, 3, 6, 7, 8, 9 and 10.
Chapter 5 Evaluation of the radio-frequency ALD

Fig. 5.2 Location of radio-frequency ALD transmitter and receiver in Measurement #4 and 5.
Chapter 5 Evaluation of the radio-frequency ALD

Packet error measurement in the Hector Macleod Building, UBC.

1. In the short corridor on 4th Floor (outside Rm 459 and Rm 441)
2. In the long corridor on 4th Floor (outside Rm 439 and Rm 402)
3. In a lecture room (Rm 228)
4. In the Communication Lab. (Rm 458)

Fig. 5.3 Packet error measurement of the radio-frequency ALD.

The mean packet error rate (number of lost packets / total packets transmitted), mean number of correct packets transmitted between losses (1 / packet error rate) and mean time between packet losses (mean number of correct packets transmitted between losses $\times 8$ ms) are shown in Table 5.1.
Chapter 5 Evaluation of the radio-frequency ALD

Table 5.1 Mean packet error rate in measurement # 1-7.

<table>
<thead>
<tr>
<th>Measurement #</th>
<th>Mean packet error rate</th>
<th>Mean correct packets transmitted between loss</th>
<th>Mean time between packet loss (seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.64 x 10^{-4}</td>
<td>2750.49</td>
<td>22.00</td>
</tr>
<tr>
<td>2</td>
<td>3.50 x 10^{-4}</td>
<td>2858.67</td>
<td>22.87</td>
</tr>
<tr>
<td>3</td>
<td>3.18 x 10^{-4}</td>
<td>3146.89</td>
<td>25.18</td>
</tr>
<tr>
<td>4</td>
<td>3.10 x 10^{-4}</td>
<td>3228.91</td>
<td>25.83</td>
</tr>
<tr>
<td>5</td>
<td>2.89 x 10^{-4}</td>
<td>3457.78</td>
<td>27.66</td>
</tr>
<tr>
<td>6</td>
<td>2.59 x 10^{-4}</td>
<td>3861.59</td>
<td>30.89</td>
</tr>
<tr>
<td>7</td>
<td>2.43 x 10^{-4}</td>
<td>4116.87</td>
<td>32.93</td>
</tr>
<tr>
<td>mean</td>
<td>3.05 x 10^{-4}</td>
<td>3345.89</td>
<td>26.77</td>
</tr>
</tbody>
</table>

Packet error rates fluctuated during some of the measurements. Measurements # 8, # 9 and # 10 in Fig. 5.4 show large increases in packet error rates for brief periods.

The sudden rise in packet error rates could be the result of external interference or other transmission impairments. Measurement # 10 shown in Fig. 5.4 and 5.5, was obtained in the Communication Laboratory. During the course of measurement, another Arlan 650™ wireless network card was deliberately set to transmit data over the same data channel (Channel 12) as the radio-frequency ALD. (i.e. two spread spectrum transmitters with the same spread code, modulation method and transmitting frequency operating in the same area.) The packet error rate increases with such interference.
Chapter 5 Evaluation of the radio-frequency ALD

Fig. 5.4 Packet error measurements under interference

There were several Arlan 650\textsuperscript{TM} wireless network cards installed in the Communication Laboratory connecting PCs and workstations. They used the same data channel (Channel 12) as that used in the packet error measurement. If they were operating during the packet error measurement, the interference generated by them would account for the sudden increase in packet error rates in Measurements # 8 and # 9.
Chapter 5 Evaluation of the radio-frequency ALD

Packet error measurement in the Communication Lab.

The kind of interference experienced in Measurement # 10 is unlikely to happen in the radio-frequency ALD application. If there would be several radio-frequency ALDs operating in the same area, each would use a different spreading code to prevent interference from each other.

Fig. 5.5 Packet error measurement under interference.
5.2 Miller and Nicely audiometric test

Packet losses in the radio-frequency ALD cause distortions in the transmitted speech. The distortion would depend on the packet size and the speech segments affected. A single 8 ms speech segment lost in a word or sentence would have little or no effect on the score of a word or sentence intelligibility test because the lost information can most likely be reconstructed from the context of the word or the sentence.

The Miller and Nicely test [11] was used to investigate the effect of packet losses on intelligibility of speech transmitted by the radio-frequency ALD. This is a syllable test with little contextual information presented. The Miller and Nicely test is designed to give an articulatory analysis of the types of errors made by the listener. There are 16 consonants (\{p, t, k, f, θ, s, j, b, d, g, v, ʃ, z, ʒ, m, n\}) used in the Miller and Nicely test. In the Miller and Nicely test, each of the 16 consonants are spoken before the vowel \(a\) to form a syllable (pronounced as “pa”, “ta”, “ka”, “fa”, “tha” as in “thank”, “sa”, “sha”, “ba”, “da”, “ga”, “va”, “tha” as in “that”, “za”, “Zha”, “ma”, and “na” respectively.). These 16 syllables make up almost three quarters of the consonants in normal speech and about 40 percent of all phonemes [11].

Each test consists of a number syllable presentations arranged in a random order. After each syllable presentation the listener is asked to indicate which of the 16 syllables was presented.

The 16 syllables can be classified according to the articulatory process used to generate the sounds. There are five articulatory features that serve to characterize and distinguish the different phonemes: voicing, nasality, affrication, duration and place of articulation. These features of speech production are reflected in certain specific acoustical
Chapter 5 Evaluation of the radio-frequency ALD

characteristics which are used by the listener to differentiate the syllables. The following set of features are used as a basis for classification.

1. Voicing. In articulatory terms, the vocal cords do not vibrate when
consonants \(|p|, |k|, |f|, |\theta|, |s|, and |f|\) are produced, but they do vibrate for
\(|b|, |d|, |g|, |v|, |\delta|, |z|, |\zeta|, |m|, and |n|\). Acoustically, this means that the voiceless
consonants are aperiodic or noisy in character, whereas a periodic or line-spectrum
component is superimposed on the noise for voiced consonants.

2. Nasality. To articulate \(|m|\) and \(|n|\) the lips are closed and the pressure is released
through the nose by lowering the soft palate at the back of the mouth. The nasal
resonance introduced in this way provides an acoustical clue.

3. Affrication. If the articulator (i.e. tongue and lips) close completely, the consonant
may be a stop or nasal, but if they are brought close together and air is forced
between them, the result is a kind of turbulence or friction noise that distinguishes
\(|f|, |\theta|, |s|, |f|, |v|, |\delta|, |z|, and |\zeta|\) from \(|p|, |t|, |k|, |b|, |d|, |g|, |m|, and |n|\).

4. Duration. This is the name that designates the difference between \(|s|, |f|, |z|, and |\zeta|\)
and the other 12 consonants. These four consonants are long, intense, high-frequency
noises, but the most effective features in setting them apart is their extra duration.

5. Place of Articulation. This feature has to do with where in the mouth the major
construction of the vocal passage occurs. The 16 consonants can be separated into
three groups with \(|p|, |b|, |f|, |v|, and |m|\) as front (0), \(|t|, |d|, |\theta|, |s|, |\delta|, |z|, and |n|\)
as middle (1), and \(|k|, |g|, |f|, and |\zeta|\) as back (2) consonants.
The classification of the 16 syllables is summarized in Table 5.2

Table 5.2 Classification of consonants used to analyze confusion

<table>
<thead>
<tr>
<th>Consonant</th>
<th>Voicing</th>
<th>Nasality</th>
<th>Affrication</th>
<th>Duration</th>
<th>Place</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>t</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>k</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>f</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>θ</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>s</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>j</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>b</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>d</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>g</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>v</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ʒ</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>z</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>ʃ</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>m</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>n</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

The results of the Miller and Nicely test are presented in a confusion matrix shown in Fig. 5.6. The syllables presented to the listener are indicated by the consonants listed vertically in the first column, and the responses of the listener are indicated horizontally across the top of the matrix. The number in each cell is the frequency with which each stimulus-response pair that was observed. The number of correct responses can be obtained by totalling the number of entries along the main diagonal. The overall articulation score is obtained by dividing the number of correct responses by the number of syllables presented to the listener.
In order to analyze the articulation score, the scores of different syllables in the confusion matrix can be combined according to their articulatory features. Combining syllables in the confusion matrix creates a smaller confusion matrix that shows the confusions between groups, and the sum along the diagonal of the smaller confusion matrix gives a new articulation score for the articulatory feature. The new score will be greater than the original score, since all the responses that were originally correct remain so and in addition all the confusions within each group are now considered to be correct in the new score.
Fig. 5.7 A confusion matrix grouped by voiced and voiceless consonants

Using the classifications in Table 5.2, the confusion matrix in Fig. 5.6 can be grouped into four portions dividing the voiced and voiceless consonant as shown in Fig. 5.7. The scores of each portion can then be summed to form the confusion matrix for the voicing feature as shown in Fig. 5.8. The probability that the voicing feature will be perceived correctly can be calculated by summing the diagonal cell in the confusing matrix (i.e. the articulation score for voicing ).


<table>
<thead>
<tr>
<th></th>
<th>voiceless</th>
<th>voiced</th>
</tr>
</thead>
<tbody>
<tr>
<td>voiceless</td>
<td>140</td>
<td>0</td>
</tr>
<tr>
<td>voiced</td>
<td>5</td>
<td>175</td>
</tr>
</tbody>
</table>

articulation score for voicing = 315 / 320

Fig. 5.8 confusion matrix for voiced and voiceless consonant

Using a similar method, a new set of articulation scores for each of the articulatory feature in Table 5.2 can be obtained by combining the syllables in the confusion matrix. This set of scores will indicate how well different articulatory features will be perceived correctly.

5.2.1 Miller and Nicely Test procedures

The purpose of the experiment was to determine whether the intelligibility of transmitted speech will decrease when the speech material processed by the $\mu$-law encoding algorithm is subjected to various lost packet conditions.

Ten recordings of the 16 Miller and Nicely syllables by two Canadian English speakers processed by the $\mu$-law encoding algorithm were stored in a Next\textsuperscript{TM} computer. From the ten recordings, the shortest and the longest sounds of the 16 syllables from each speaker were selected as the control set (a total of 64 syllables). The measured length of these syllables in the control set varies from 225 ms to 612 ms with a mean length of 448 ms.

Simulated packet losses of various packet sizes were applied to the 64 syllables in the control set to make new sets of syllables. The speech segment chosen to correspond to the simulated packet loss was the segment in the syllable that contains the maximum
amount of voice information (i.e. the region where the end of the consonant meets the start of the vowel as shown in Fig. 5.9).

The following eight test conditions were used.

1. Control Set (original syllables).
2. 8 ms packet loss replaced by silence.
3. 8 ms packet loss replaced by the last received 8–ms-packet.
4. 16 ms packet loss replaced by silence.
5. 16 ms packet loss replaced by the last received 8–ms-packet repeating two times to fill the 16 ms packet loss.
6. 32 ms packet loss replaced by silence.
7. 32 ms packet loss replaced by the last received 8–ms-packet repeating four times to fill the 32 ms packet loss.
8. 64 ms packet loss replaced by silence.
Chapter 5 Evaluation of the radio-frequency ALD

The arrangement shown in Fig. 5.10 was used to conduct the tests. The Next\textsuperscript{TM} computer outside a sound-proof room was used to present the syllables to the subjects. Sound output from the Next\textsuperscript{TM} computer was administered to the subjects through headphones at 50 dBHL. Sound level control was accomplished using the audiometer. The subject responded through the Macintosh\textsuperscript{TM} Computer in the sound-proof room. After each syllable presentation, the Next\textsuperscript{TM} computer waited for the subject to respond before presenting the next syllable.

![Fig. 5.10 Set up for the Miller and Nicely test](image)

For each test condition, the 64 syllables in a set were repeated five times (i.e. each test had a total of 320 syllables) and presented to the subjects in a random order. After each syllable presentation, the listener was asked to indicate (to guess, if necessary) which of the given 16 syllables was heard.

The experiment was divided into three stages. In the first stage, a normal hearing pilot subject was used to check out the experimental procedures and to obtain primarily results in the experiment. Results and feedbacks in the first stage were used as guidelines in the
second stage and third stage where normal and hearing impaired subjects were tested.

5.2.2 Results of the Miller and Nicely Test

Results of the pilot subject

A normal hearing subject was tested with the Control set; 8 ms, 16 ms, 32 ms and 64 ms simulated packet lost replaced with silence substitution. The overall score of the tests are shown in Table 5.3.

Table 5.3 Overall Scores of the pilot subject

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>8 ms</th>
<th>16 ms</th>
<th>32 ms</th>
<th>64 ms</th>
</tr>
</thead>
<tbody>
<tr>
<td>overall score</td>
<td>81.9</td>
<td>82.8</td>
<td>74.7</td>
<td>70.3</td>
<td>51.5</td>
</tr>
</tbody>
</table>

The results show that the score for silence substitution in 8ms packet lost has a slightly higher score (0.9%) than the Control set, which indicates a single 8 ms packet loss has small or no effect to the subject in terms of the intelligibility. The results also indicate that as the length of speech segment lost increases, the overall score decreases.

The pilot subject indicated that great concentration and effort is needed for the 64 ms silence substitution test, and there is a lot of uncertainty in getting the correct syllables. To avoid introducing any hardship on the subjects, it was decided that 64 ms silence substitution test will be dropped from the second and third stages of the experiment.
Chapter 5 Evaluation of the radio-frequency ALD

Results using normal hearing subjects

Six young normal hearing subjects with training in phonetic pronunciation (students of the Department of Audiology and Speech Sciences) were tested. The reason for using normal hearing subjects is that they have more sensitive hearing than the hearing impaired listeners in detecting the minor changes in the syllables resulting from the packet replacement strategies applied to small segments of the speech materials.

Seven tests, the control set; silence substitution of 8 ms, 16 ms and 32 ms packet; and packet repetition of 8 ms, 16 ms and 32 ms were applied to each subject in random order. The overall score of the tests were shown in Table 5.4.

Table 5.4 Miller and Nicely test results for the normal hearing subjects.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Control</th>
<th>8 ms</th>
<th>16 ms</th>
<th>32 ms</th>
<th>8 ms</th>
<th>16 ms</th>
<th>32 ms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>86.9</td>
<td>90.0</td>
<td>84.7</td>
<td>80.6</td>
<td>89.1</td>
<td>85.3</td>
<td>76.6</td>
</tr>
<tr>
<td>2</td>
<td>88.8</td>
<td>89.7</td>
<td>86.9</td>
<td>78.4</td>
<td>88.4</td>
<td>86.6</td>
<td>81.3</td>
</tr>
<tr>
<td>3</td>
<td>85.0</td>
<td>80.6</td>
<td>82.8</td>
<td>75.9</td>
<td>81.6</td>
<td>83.4</td>
<td>75.6</td>
</tr>
<tr>
<td>4</td>
<td>82.2</td>
<td>82.5</td>
<td>81.9</td>
<td>75.3</td>
<td>86.6</td>
<td>83.4</td>
<td>74.3</td>
</tr>
<tr>
<td>5</td>
<td>91.3</td>
<td>90.0</td>
<td>89.4</td>
<td>74.7</td>
<td>88.1</td>
<td>87.2</td>
<td>83.1</td>
</tr>
<tr>
<td>6</td>
<td>87.8</td>
<td>81.6</td>
<td>82.2</td>
<td>74.1</td>
<td>84.4</td>
<td>80.6</td>
<td>72.5</td>
</tr>
<tr>
<td>Mean</td>
<td>87</td>
<td>85.7</td>
<td>84.7</td>
<td>76.5</td>
<td>86.4</td>
<td>84.4</td>
<td>77.2</td>
</tr>
</tbody>
</table>

The mean results show that the scores for silence substitution and packet repetition in 8ms packet loss has a slightly lower score (within 1.3%) than the Control set, which indicates a single 8 ms packet loss has small effect in degrading the intelligibility of the syllables. The results also indicate that in terms of intelligibility scores, the difference
between the two packet replacement techniques are within 0.7% of each other.

Results using hearing impaired subjects

Since the results in the normal hearing subjects show that the two packet replacement techniques has similar effects. The tests on hearing impaired subjects concentrated on the silence substitution technique, Four tests, the control set; silence substitution on 8 ms, 16 ms and 32 ms packet were applied to each subject in random order.

Three senior hearing impaired subjects with mild hearing loss at high frequency were tested. Their audiogram is shown in Fig 5.11. The overall score of the tests are shown in Table 5.5.

Table 5.5 Miller and Nicely test results for the hearing impaired subjects.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Control</th>
<th>8 ms</th>
<th>16 ms</th>
<th>32 ms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>73.8</td>
<td>76.9</td>
<td>69.1</td>
<td>59.1</td>
</tr>
<tr>
<td>2</td>
<td>69.7</td>
<td>71.3</td>
<td>68.1</td>
<td>60.3</td>
</tr>
<tr>
<td>3</td>
<td>82.8</td>
<td>76.6</td>
<td>74.7</td>
<td>68.4</td>
</tr>
<tr>
<td>Mean</td>
<td>75.4</td>
<td>74.9</td>
<td>70.6</td>
<td>62.6</td>
</tr>
</tbody>
</table>

The mean results show that the scores for silence substitution with 8ms packet loss are slightly lower score (within 0.5%) than the Control set, which indicates a single 8 ms packet loss has a small degradation effect on intelligibility. The results also indicate that as for the normal hearing subjects, as the length of speech segment lost increases, the overall score decreases.
Fig. 5.11 Audiogram of the hearing impaired subjects
5.2.4 Effect of packet lost on speech intelligibility

The mean results of the experiment for the pilot subject, the normal, and the hearing impaired subjects are summarized in Fig. 5.12. (0 ms speech segment losses corresponds to the scores of the control set)

![Miller and Nicely Test (Overall score)](image)

Fig. 5.12 Results of the Miller and Nicely Tests

For the normal hearing subjects, 8 ms packet loss have small effect on the overall score (within 1.3% of control set) with silence substitution or packet repetition. The hearing impaired subjects have lower mean scores than the normal hearing subjects. The
overall scores of each hearing impaired subject are shown in Fig. 5.13. The overall score decreases as the length of packet segment increases, but 8 ms packet loss has greater effect in subject 3 than the other two subject.

![Miller and Nicely Test for hearing impaired subjects (Overall score)](image)

**Fig. 5.13** Overall scores for the hearing impaired subjects.

Fig. 5.14, Fig. 5.15 and Fig. 5.16 show the different mean articulatory scores obtained under different packet loss conditions using different packet replacement strategies for the two group of listeners (0 ms speech segment losses corresponds to the scores of the control set). The results show that some of the articulatory features are better preserved when packets are lost. In general, as the length of speech segment increases, scores for different articulatory features decrease.
Chapter 5 Evaluation of the radio-frequency ALD

Fig. 5.14 Results of different articulatory scores in silence substitution for normal hearing subjects.

For the normal hearing subjects, the articulatory scores obtained using silence substitution or packet repetition techniques are similar, except that the voicing articulatory feature is more well preserved using the packet repetition technique.
The articulatory scores for different hearing impaired subjects are shown in Fig. 5.16. The scores of different articulatory features vary from individual to individual. In general, the scores for different articulatory features decrease as the length of packet lost increases.

The individual scores of the 16 syllables for the normal hearing subjects and hearing impaired subjects under different packet loss conditions are shown in Fig. 5.17. The results indicate that some of the syllables like “pa”, “ta”, “ka”, “fa”, “Sha”, “ga”, “ma”, “na” is better preserved than others.

Fig. 5.15 Results of different articulatory scores in packet repetition for normal hearing subjects.
Chapter 5 Evaluation of the radio-frequency ALD

Fig. 5.16 Results of different articulatory scores in silence substitution for hearing impaired subjects.
Fig. 5.17 Scores of the 16 syllables under different packet lost conditions.
Chapter 5 Evaluation of the radio-frequency ALD

Syllable tests are the most difficult for intelligibility because no contextual information is presented in the stimuli. When words or sentences are presented to the listener as in a normal conversation, there is more contextual information and the lost information present in a small segment of speech is likely to be recovered from the contextual information in the rest of the utterance. In addition, the packet losses occur in a more or less random fashion, and not always occurs in the place that would affect the word or syllable most. The degradation effect of small speech segment losses on speech intelligibility should become smaller. The effect of single (8 ms) or double packet (16 ms) speech segment loss in the radio-frequency ALD should be minimal for normal or hearing impaired listeners.
Chapter 6 Conclusions

6.1 Summary

This thesis documents the design, implementation and testing of a secure radio-frequency Assistive Listening Device for hard of hearing listeners. The design combines the use of spread spectrum and digital voice encoding technology to provide secure and high quality voice reception for hard of hearing users. The devices offer hearing impaired listeners a better communication channel in difficult listening conditions than hearing aids alone can provide.

In spite of advances in digital voice encoding technology, it has not enjoyed widespread use in devices for the hard of hearing community. The results from our intelligibility tests indicate that in terms of intelligibility the CCITT G.722 and the $\mu$-law encoding algorithms are comparable to undigitized voice for hard of hearing listeners.

In order to investigate the feasibility of transmitting digitized voice using the direct sequence spread spectrum method, a radio-frequency ALD transmitter and receiver set was built based on the Arlan 650\textsuperscript{TM} wireless network card.

The radio-frequency ALD was tested in different areas of a large multi-use building. The results of these tests indicate that the device performs well under normal circumstances with a packet error rate around $3 \times 10^{-4}$.

Despite the low packet error rate, each packet lost will result in the loss of a 8 ms long speech segment. Two simple lost packet replacement methods, silence substitution and packet repetition, were investigated to see if speech intelligibility would degrade substantially in the event of single and multi-packet losses. The results indicate that
Chapter 6 Conclusions

Intelligibility loss is minimal under single or double packet losses. Intelligibility begins to drop when the lost speech segment is longer than 16 ms. We also found that packet repetition is slightly better than silence substitution in preserving speech intelligibility.

6.2 Suggestions for further work

The development of the radio-frequency ALD was accomplished on personal computers. The physical size of the device severely limits its mobility. Further work is needed to produce a device which could be used by the hard of hearing community.

Advances in digital voice processing have provided more efficient encoding algorithms, in terms of bit rate required than the $\mu$-law algorithm used in this work. Application of such voice encoding algorithms in the radio-frequency ALD would reduce the data transmission rate and hence the cost of the final system. Further study is needed to determine whether these algorithms are acceptable for hard of hearing listeners.
### Appendix A
Example of a Revised SPIN Test Form

Form #4 of the Revised SPIN Test (12/83)

<table>
<thead>
<tr>
<th>Name</th>
<th>(#)</th>
<th>Marker</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>S/B</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>#C.</th>
<th>#C.</th>
<th>ACCEPT?</th>
<th>Y/N</th>
<th>Percent Hrg.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

108
26. She made the bed with clean SHEETS.          H
27. I've been considering the CROWN.            L
28. The team was trained by their COACH.        H
29. I've got a cold and a sore THROAT.          H
30. We've spoken about the TRUCK.               L
31. She wore a feather in her CAP.              H
32. The bread was made from whole WHEAT.        H
33. Mary could not discussed the TACK.          L
34. Spread some butter on your BREAD.           H
35. The cabin was made of LOGS.                 H
36. Harry might consider the BEEF.              H
37. We're glad Bill heard about the ASH.         L
38. The loin gave an angry ROAR.                H
39. The sandal has a broken STRAP.              H
40. Nancy should consider the FIST.             L
41. He's employed by a large FIRM.              H
42. They did not discuss the SCREEN.            L
43. Her entry should win first PRIZE.           H
44. The old man think about the MAST.           L
45. Paul wants to speak about the BUGS.         L
46. The airplan dropped a BOMB.                 H
47. You're glad she called about the BOWL.      L
48. A zebra has black and white STRIPES.        H
49. Miss Black could have discussed the ROPE.   L
50. I hope Paul asked about the MATE.            L
Bibliography


1990.


