

Manuel G. de Oliveira Gericofa

Ajudas Técnicas à
Comunicação para Pessoas
com Paralisia Cerebral

DEEC
FEUP
1995



DEPARTAMENTO DE
ENGENHARIA ELECTROTÉCNICA E DE COMPUTADORES

Ajudas Técnicas à
Comunicação para Pessoas
com Paralisia Cerebral

FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

Rua dos Bragas, 4099 Porto Codex – PORTUGAL

engenharia

FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Departamento de Engenharia Electrotécnica e de Computadores

Ajudas Técnicas à Comunicação para Pessoas com Paralisia Cerebral

Manuel Gradim de Oliveira Gericota

Licenciado em Engenharia Electrotécnica e de Computadores pela
Faculdade de Engenharia da Universidade do Porto

Dissertação submetida para satisfação parcial dos requisitos do
programa do curso de

Mestrado em Engenharia Electrotécnica e de Computadores

0437
GERm/AGU
PORTO, Junho de 1995

UNIVERSIDADE DO PORTO
Faculdade de Engenharia
BIBLIOTECA M
N.º <u>27071</u>
CDU <u> </u>
Data <u>24 / 09 / 1998</u>

DEEC
FEUP

681.2(043)
GERm/AGU

Dissertação realizada sob a supervisão do
Professor Doutor José Manuel Martins Ferreira,
Professor Auxiliar do
Departamento de Engenharia Electrotécnica e de Computadores da
Faculdade de Engenharia da Universidade do Porto

*O rosto da criança com paralisia cerebral
poderá ser inexpressivo e verter saliva,
mas não porque seja deficiente mental,
senão, simplesmente, porque os músculos
não cumprem a sua função normal.*

M. F. Evans

Este trabalho é dedicado a todas elas

Resumo

Comunicar é fundamental para todos os seres humanos. Por motivos vários muitos de nós apresentam problemas ao nível da comunicação. De entre outros, a paralisia cerebral é um dos motivos que frequentemente conduz a graves distúrbios comunicativos.

O objectivo da comunicação aumentativa e alternativa é permitir a comunicação a pessoas incapazes de comunicar através dos meios naturais como a fala, o gesto ou a escrita. A introdução da electrónica veio melhorar significativamente a eficácia desses sistemas. Todavia o seu elevado custo tem constituído uma barreira intransponível para um largo número de potenciais utilizadores.

Iniciado com uma identificação dos requisitos dos utilizadores, realizada em cooperação com terapeutas do Centro de Reabilitação de Paralisia Cerebral do Porto, e recorrendo à utilização de tecnologias desenvolvidas para aplicação em sistemas de gravação/reprodução com larga expansão no mercado, este trabalho conduziu à criação de um conjunto de ajudas técnicas à comunicação, tornando os seus benefícios acessíveis a um número maior de pessoas com problemas de comunicação.

Abstract

Communication is fundamental for all human beings. For several reasons many of us display expressive communication disorders. Among other examples, cerebral palsy is often one reason that leads to severe communication impairments.

The aim of augmentative and alternative communication (AAC) is to enable communication to all those who are incapable to communicate by natural means such as speech, gesture, or in written form. The introduction of electronics into AAC systems significantly improved their performance/effectiveness, but cost remains a barrier for a large number of potential users.

Starting from an identification of user needs which was done in co-operation with therapists from the Cerebral Palsy Rehabilitation Centre of Porto, and using widely available off-the-shelf technology developed for general purpose voice record/playback applications, the work described in this document led to the development of a variety of voice-output communication aids, potentially extending their benefits to a much larger number of persons with communication impairments.

Agradecimentos

Este trabalho implicou uma colaboração multidisciplinar entre várias pessoas e instituições a quem aqui quero deixar uma palavra de agradecimento.

Em particular ao meu orientador científico, Professor Doutor José Manuel Martins Ferreira pelo incentivo e pela confiança que ao longo destes anos de colaboração sempre demonstrou pelo meu trabalho.

Aos meus colegas do grupo PTSE (Projecto Testável de Sistemas Electrónicos) no INESC - Porto (Instituto de Engenharia de Sistemas e Computadores) pela sua permanente disponibilidade para me ajudar.

Ao INESC - Porto pela disponibilização dos meios necessários para realização da tese.

Aos terapeutas do CRPCP (Centro de Reabilitação de Paralisia Cerebral do Porto), em especial à terapeuta Lourdes Lourenço e ao terapeuta Joaquim Faias, pela sua colaboração.

Uma palavra especial de carinho aos utentes do CRPCP a quem este trabalho é dirigido.

Aos meus pais pela sua inabalável confiança em mim, pelo seu constante estímulo, pelo seu carinho, pela paciência e por todas as possibilidades que me deram para vencer.

Aos meus sonhos e a todos aqueles que de uma ou outra forma contribuíram para os alimentar e os tornar realidade.

Índice

1. Introdução.....	1
1.1 Organização da tese	2
2. Comunicação Aumentativa e Alternativa	4
2.1 Conceitos básicos e áreas de aplicação	5
2.2 Objectivos principais dos sistemas de CAA.....	11
2.3 Principais sistemas de CAA.....	14
3. Ajudas Técnicas à Comunicação.....	22
3.1 Técnicas de selecção	23
3.2 Ajudas técnicas com saída de voz	25
3.3 Técnicas de codificação de voz: o ADPCM	31
3.3.1 Características da fala.....	32
3.3.2 Digitalização da fala	33
3.3.2.1 Amostragem.....	33
3.3.2.2 Quantificação das amostras.....	39
3.3.2.3 Codificação de sinal	44
3.3.2.4 Medidas da qualidade da fala	52
3.4 O armazenamento directo em analógico: A tecnologia DAST.....	54

4. O Projecto “Fala-Barato”: Da Identificação de Requisitos à Especificação Técnica	58
4.1 A paralisia cerebral	59
4.2 Identificação de requisitos	62
4.3 Especificação técnica	64
5. Implementação da Especificação Técnica	71
5.1 “Fala-Barato I”: máxima simplicidade, satisfação mínima de requisitos	72
5.2 “Fala-Barato II”: satisfação completa de requisitos	78
5.2.1 Arquitectura da solução desenvolvida	78
5.2.2 As diferentes possibilidades de acesso	90
5.2.2.1 <i>Interface</i> por teclado	91
5.2.2.2 <i>Interface</i> por matriz luminosa	91
5.2.2.3 <i>Interface</i> numérica	92
5.2.3 Considerações gerais	93
6. Avaliação	95
6.1 Critérios de avaliação	96
6.2 Procedimentos	99
6.3 Resultados obtidos	101
7. Conclusão	103
7.1 Principais dificuldades encontradas	104
7.2 Perspectivas de trabalho futuro	104
Bibliografia	106

Lista de Figuras

Figura 2.1 : Quadro de comunicação.....	6
Figura 2.2 : Exemplos de símbolos <i>Bliss</i>	16
Figura 2.3 : Exemplo dum quadro com alfabeto.....	17
Figura 2.4 : Exemplos de símbolos SPC.....	18
Figura 2.5 : Exemplo de criação de símbolos <i>Picsyms</i>	18
Figura 2.6 : Exemplos de símbolos PIC.....	19
Figura 2.7 : Apontador luminoso para cabeça com cinto ajustável.	20
Figura 3.1.: Exemplo de um botão de pressão accionável com um movimento lateral da cabeça.	24
Figura 3.2 : Exemplo de utilização do <i>Liberator</i>	29
Figura 3.3 : Sistema multi-modal de acesso a meios informáticos.....	29
Figura 3.4 : Segmentos de sinais acústicos de fala vocalizada e não-vocalizada.	33
Figura 3.5 : Envolvente espectral de um segmento de fala vocalizada.....	34
Figura 3.6 : Sinal analógico.	35
Figura 3.7 : Espectro de frequência dum sinal analógico.	35
Figura 3.8 : Espectro de frequência dum sinal analógico amostrado a $f_s > 2f_m$	35
Figura 3.9 : Espectro de frequência dum sinal analógico amostrado a $f_s < 2f_m$	36
Figura 3.10 : Espectro de frequência típico dum sinal de voz.	37
Figura 3.11 : Aplicação de um filtro <i>anti-aliasing</i>	37
Figura 3.12 : Função de amostragem.....	38
Figura 3.13 : Sinal analógico de entrada.	38
Figura 3.14 : Versão amostrada do sinal de entrada.	38
Figura 3.15 : Funções de <i>anti-aliasing</i> , amostragem e retenção.	39
Figura 3.16 : Definição de 4 níveis de quantificação.....	39
Figura 3.17 : Forma de onda obtida após a quantificação em 4 níveis.	40
Figura 3.18 : Função de transferência de um conversor analógico-digital com quantificação linear ou uniforme.	40

Figura 3.19 : Sinal de voz típico e esquema de quantificação uniforme.	41
Figura 3.20 : Sinal de voz típico e esquema de quantificação não-uniforme.	41
Figura 3.21 : Função de transferência da μ -law.	42
Figura 3.22 : Função de transferência da A-law.	43
Figura 3.23 : Conversor analógico-digital por aproximações sucessivas.	43
Figura 3.24 : Técnica de codificação PCM.	44
Figura 3.25 : Aplicação da técnica de codificação <i>Delta Modulation</i> a um sinal analógico amostrado.	45
Figura 3.26 : Modulador e desmodulador <i>Delta</i>	46
Figura 3.27 : Saída digitalizada de um circuito de amostragem-retensão e sua codificação usando a <i>Delta Modulation</i>	46
Figura 3.28 : <i>Delta Modulation</i> com sobrecarga na inclinação.	47
Figura 3.29 : <i>Delta Modulation</i> com granularidade.	48
Figura 3.30 : Modulador e desmodulador DPCM.	48
Figura 3.31 : Exemplo de quantificação adaptativa dependente da amplitude de entrada do sinal.	50
Figura 3.32 : ADPCM: Modulador e desmodulador para a estimação adiantada (a) e atrasada (b).	50
Figura 3.33 : Esquemas do modulador (a) e desmodulador (b) definidos pela norma CCITT G.721.	51
Figura 3.34 : Fases do processo de armazenamento directo em analógico.	55
Figura 3.35 : Esquema construtivo duma célula de memória analógica.	56
Figura 4.1 : Modelo genérico de interacção homem-máquina.	65
Figura 4.2 : Possibilidades a oferecer pelo segundo sistema.	70
Figura 5.1 : Diagrama de blocos do ISD1416.	72
Figura 5.2 : Esquema eléctrico do "Fala-Barato I".	75
Figura 5.3 : O "Fala-Barato I".	77
Figura 5.4 : Configuração base dum sistema baseado no DS2271 em modo <i>software</i>	80
Figura 5.5 : Organização da memória de voz após a formatação.	81
Figura 5.6 : Organização do <i>byte</i> de controlo.	82
Figura 5.7 : Diagrama de blocos geral do "Fala-Barato II".	87
Figura 5.8 : Protótipo do módulo de base do "Fala-Barato II".	88
Figura 5.9 : Esquema eléctrico do módulo de base do "Fala-Barato II".	89
Figura 5.10 : Configurações possíveis do teclado.	92

Lista de Tabelas

Tabela 3.1 : Escala MOS	53
Tabela 5.1 : Pinos de selecção da velocidade de transferência da porta série.	80
Tabela 5.2 : Capacidade de gravação disponível (segundos).	83
Tabela 5.3 : Lista de comandos do DS2271.....	85

Abreviaturas

ADPCM - *Adaptive Differential Pulse Code Modulation*

ASCII - *American Standard Code for Information Interchange*

ASHA - *American Speech-Language-Hearing Association*

CAA - *Comunicação Aumentativa e Alternativa*

CCITT - *International Telegraph and Telephone Consultative Committee*

CRPCP - *Centro de Reabilitação de Paralisia Cerebral do Porto*

DAM - *Diagnostic Acceptability Measure*

DAST - *Direct Analog Storage Technology*

DM - *Delta Modulation*

DPCM - *Differential Pulse Code Modulation*

DRT - *Diagnostic Rhyme Test*

EEPROM - *Electrically Erasable Programmable Read Only Memory*

LDM - *Linear Delta Modulation*

LED - *Light Emitter Diode*

MOS - *Mean Opinion Score*

NVSRAM - *Non-Volatile Static Random Access Memory*

PCM - *Pulse Code Modulation*

PIC - *Pictogram Ideogram Communication*

SPC - Símbolos Pictográficos de Comunicação

TIDE - *Technology Initiative for Disabled and Elderly People*

TTL - *Transistor-Transistor Logic*

Glossário

Afasia - defeito ou perda da palavra, consequente a lesão dos centros cerebrais.

Amiotrofia - atrofia muscular.

Apraxia - perda da faculdade de reproduzir movimentos coordenados.

Célula piramidal - cada uma das células que se encontram no córtex cerebral colocadas a várias alturas, constituindo camadas, entre as quais se encontram outras de forma diversa.

Cerebelo - parte posterior e inferior do encéfalo, situada atrás e para baixo do cérebro.

Cognitivo, processo - o processo cognitivo abrange uma vasta gama de funções incluindo a memória, a atenção, a aprendizagem, a resolução de problemas e a linguagem.

Comunicação aumentativa e alternativa - ampliação ou substituição das formas naturais de comunicação como a fala, o gesto ou a escrita.

Córtex - camada superficial do cérebro.

Disartria - dificuldade na pronúncia e articulação das palavras.

Disfasia global - dificuldade ou perturbação global da comunicação.

Ditongo - reunião de duas vogais que se pronunciam numa só emissão de voz, pertencendo, portanto, à mesma sílaba.

Doença - falta ou perturbação de saúde.

Doença de Parkinson - síndrome composta por um grupo de distúrbios motores que incluem tremor, rigidez, redução no movimento e diminuição nos reflexos.

Encéfalo - parte do sistema nervoso contida na cavidade do crânio e que compreende o cérebro, o cerebelo e o bulbo raquidiano.

Esclerose lateral amiotrófica - doença degenerativa progressiva de etiologia desconhecida, envolvendo os neurónios do cérebro e da espinhal medula responsáveis pela parte motora.

Esclerose múltipla - doença degenerativa do sistema nervoso central.

Espasmo - contracção convulsiva dos músculos.

Etiologia - parte da medicina que estuda as causas das doenças.

Fala - a faculdade de emitir palavras.

Feixe piramidal - feixe que entra na composição da via motriz.

Fricativa - som consonântico que se produz com estreitamento, mas sem contacto das partes do tubo vocal, como o v e o f.

Glotologia - estudo científico das línguas.

Inteligível - que se compreende bem.

Laringectomia - ablação total ou parcial da laringe que se pratica, por exemplo, no caso de cancro da laringe.

Linguagem - conjunto de sinais, nascidos da vida social, de que os homens se servem para a expressão e comunicação das suas ideias ou dos seus sentimentos.

Meningite - inflamação da membrana envolvente do sistema nervoso central.

Prosódia - pronúncia regular das palavras.

Semântica - estudo dos elementos da linguagem considerados nas suas significações.

Síndrome - conjunto de sintomas ou sinais que ocorrem em determinado estado mórbido (de doença).

Sintaxe - parte da glotologia que estuda a forma de ligar as palavras, a fim de, por essa forma, se exprimir com a maior precisão o pensamento lógico; relaciona, portanto, essas palavras consoante o nexó lógico das ideias que elas traduzem, estudando não só as frases como o sentido dos diferentes elementos que elas possam contar.

Subcortical - zona imediatamente inferior ao córtex.

Tonicidade - manifestação particular da elasticidade dos tecidos vivos que se manifesta por resposta normal aos estímulos, em especial do tecido muscular.

Tono ou tónus - estado normal de tensão dos elementos anatómicos.

Traqueotomia - incisão cirúrgica com que se estabelece a comunicação entre a traqueia e o exterior.

Voz - produção de sons na laringe por meio do ar emitido pelos pulmões.

1. Introdução

Uma das mais importantes necessidades do ser humano é o seu relacionamento com os outros. Para expressão das suas necessidades, da sua vontade, para troca de pontos de vista, para um aumento do conhecimento mútuo, para fazer amigos, para a sua realização profissional, a comunicação é factor essencial.

As estimativas apontam que cerca de 10% da humanidade apresenta um qualquer tipo de deficiência. Desse grupo uma porção significativa é afectada por deficiências ao nível comunicativo [BM92].

A deficiência, entendida como a parte dum distúrbio relacionada com a redução da capacidade de um indivíduo prover às necessidades da sua vida diária, resulta em desvantagens sociais que se devem procurar atenuar ao máximo.

Ser portador duma deficiência comunicativa grave é uma incapacidade que pode ser reduzida lançando mão de modernas tecnologias, tornando-as acessíveis a essas pessoas, proporcionando-lhes as ferramentas e o treino correcto para a sua eficaz utilização.

A forma como uma pessoa com problemas de comunicação enfrenta a sua dificuldade, varia de indivíduo para indivíduo, conforme a sua personalidade e igualmente conforme a origem dessa dificuldade. Se aparentemente para um indivíduo com um problema congénito é mais fácil o confronto com essa situação, nos casos em que o problema é adquirido tal parece provocar um elevado grau de frustração. Para o homem, ser intrinsecamente social, a comunicação é factor de primordial importância para o seu próprio bem-estar. Daí a quase obrigação de nos esforçarmos, empregando todos os meios ao nosso alcance, para proporcionarmos àqueles de entre nós com problemas de comunicação, meios para se poderem expressar e fazer compreender pelos que os rodeiam. A comunicação aumentativa e alternativa tem como objectivo o oferecimento de métodos de ampliação das capacidades remanentes de comunicação, ou de substituição, no caso da

ausência de qualquer forma de expressão comunicativa perceptível.

Foi sem dúvida a motivação resultante do facto de se ter consciência de que se poderia contribuir para incrementar significativamente a ajuda que em Portugal é proporcionada a essas pessoas, e em especial ao grupo de pessoas com paralisia cerebral, que levou ao lançamento do projecto “Fala-Barato”, o qual esteve na base deste trabalho. O principal objectivo deste projecto é a concepção de sistemas de comunicação aumentativa e alternativa com saída de voz, de baixo custo, tornando-os acessíveis aos nossos deficientes, fomentando uma sua mais plena integração na sociedade.

O desenvolvimento de sistemas deste tipo a custos reduzidos tornou-se possível graças à utilização de tecnologias de gravação/reprodução de voz, as quais ultimamente têm tido uma larga expansão no mercado, nomeadamente em brinquedos, atendedores de chamada ou memorandos.

1.1 Organização da tese

Para além deste capítulo introdutório, esta tese possui mais seis capítulos.

O capítulo dois pretende situar o leitor na área da comunicação aumentativa e alternativa, apresentar alguns conceitos, demonstrar a sua necessidade, os seus objectivos, dar a conhecer a sua evolução bem como alguns dos sistemas existentes.

No terceiro capítulo analisaremos o impacto da introdução da electrónica nas ajudas à comunicação, principalmente o aparecimento de sistemas com saída de voz. Resumidamente são apresentados os conceitos básicos da digitalização de sinais analógicos e as técnicas de codificação frequentemente empregues nesses sistemas. Por último, descreve-se brevemente uma nova tecnologia que permite o armazenamento directo de sinais analógicos.

O capítulo quatro inicia-se com uma caracterização sumária da paralisia cerebral. A partir dos conhecimentos adquiridos são identificados os requisitos que os sistemas de ajuda à comunicação a conceber devem possuir e apresentada a sua especificação técnica.

O capítulo cinco apresenta as soluções desenvolvidas: dois sistemas de características técnicas e finalidades diferentes. O primeiro, designado por “Fala-Barato I”, possui um tipo único de *interface* e permite o armazenamento e reprodução de uma só mensagem. O

segundo, o "Fala-Barato II", é um sistema modular que permite, dependendo da escolha da *interface* com o utilizador de entre as três apresentadas, o acesso a um máximo de 254 mensagens. As várias *interfaces* implementam diferentes possibilidades de acesso e de controlo da operação do sistema permitindo uma melhor adaptação destes a cada caso concreto.

O sexto capítulo é dedicado à avaliação dos sistemas. São discutidos e apresentados os critérios que presidiram à avaliação, bem como os procedimentos adoptados, finalizando com a apresentação dos resultados da avaliação.

Por último no capítulo sete apresentam-se as conclusões principais do trabalho realizado e discutem-se as perspectivas de futuro que os resultados obtidos permitem antever.

2. Comunicação Aumentativa e Alternativa

As ajudas técnicas à comunicação inserem-se numa área mais vasta designada por Comunicação Aumentativa e Alternativa (seguidamente designada apenas pela abreviatura CAA). É objectivo desta área proporcionar a pessoas incapazes de prover às suas próprias necessidades diárias de comunicação através de meios naturais como a fala, o gesto ou a escrita, sistemas alternativos ou aumentativos de expressão.

Embora o enfoque principal deste trabalho seja a concepção de ajudas técnicas com saída de voz para apoiarem a comunicação de pessoas com paralisia cerebral, veremos que os distúrbios de comunicação resultam de uma vasta variedade de condições, doenças e síndromas que podem afectar pessoas de todas as idades.

A CAA é um campo multidisciplinar por excelência, no qual os utilizadores de sistemas de CAA e suas famílias, conjuntamente com educadores, terapeutas ocupacionais e da fala, fisioterapeutas, psicólogos, engenheiros e programadores informáticos contribuem com práticas e conhecimentos tendo como objectivo a concepção do sistema que melhor se adapte e preencha as necessidades do utilizador.

Na primeira parte deste capítulo será proporcionado ao leitor o necessário enquadramento dentro desta área, identificando-se a sua necessidade. Numa segunda parte são expostos os objectivos que se visam atingir com o seu emprego e consequentemente as regras que devem presidir à sua concepção. Por último são revistos alguns dos sistemas de ajuda existentes.

2.1 Conceitos básicos e áreas de aplicação

Segundo a *American Speech-Language-Hearing Association* (ASHA) a “comunicação aumentativa e alternativa é a área da prática clínica que tenta compensar (quer temporária, quer permanentemente) dificuldades ou incapacidades demonstradas por indivíduos com distúrbios graves da expressão comunicativa (isto é, dificuldades severas da fala, da linguagem e/ou da escrita)” [BM92].

Dois aspectos a salientar desta definição, e que requerem um maior aprofundamento, são as diferenças entre comunicação aumentativa e comunicação alternativa, e entre fala e linguagem. Por comunicação aumentativa entende-se o uso de um sistema tendente a facilitar a comunicação pessoal [Edw94] que complemente capacidades vocais ou verbais existentes. O termo comunicação alternativa refere-se aos métodos usados por uma pessoa sem qualquer capacidade vocal.

É de realçar que as intervenções na área da CAA devem, sempre que possível, complementar e mesmo incentivar o uso das capacidades residuais de comunicação do indivíduo, incluindo qualquer vestígio de fala ou vocalização, gestos ou sinais, e não procurar substituí-las integralmente, o que poderá contribuir para o atrofiamiento dessas formas residuais.

Embora fala e linguagem sejam termos usados correntemente com significação indistinta, a sua distinção na área da CAA é importante. Por fala entende-se a capacidade de emitir sons (por ex.: p, b, m, ch, nh) e de os aglutinar por forma a gerar palavras. A linguagem implica o conhecimento do significado das palavras e o seu uso correcto, tendente a construir frases com sentido [Tat93].

Ainda segundo a ASHA, um sistema de CAA “é um grupo integrado de componentes, incluindo símbolos, ajudas, estratégias e técnicas usadas por indivíduos para melhorar a comunicação” [BM92].

De realçar que esta definição abarca o uso de múltiplos componentes ou modos de comunicação não dissociáveis. O termo símbolos refere-se, aqui, a representações visuais, auditivas e/ou tácteis de conceitos convencionais, como gestos, fotografias, sinais gestuais,

pictogramas, ideogramas, palavras escritas ou faladas, objectos, *Braille*. Incluída está por conseguinte a comunicação gestual identificada com expressões faciais, de olhar e posturas corporais, e que normalmente enriquecem o processo comunicativo. É de salientar esta ideia, uma vez que as ajudas desenvolvidas durante a prossecução deste trabalho, não são de alguma forma exclusivas, mas complementares de um todo que inclui igualmente os referidos modos naturais remanentes nos indivíduos com problemas de comunicação, permitindo inclusive, como será referido na parte da avaliação dos resultados, melhorar a interacção visual entre os intervenientes na comunicação.

O termo ajuda abrange um grupo vasto de objectos ou dispositivos físicos usados para transmitir ou receber mensagens como livros, carteiras ou quadros de comunicação, gráficos, dispositivos mecânicos ou electrónicos, ou computadores. De referir que estes últimos, os dispositivos mecânicos ou electrónicos e os computadores, são normalmente denominados ajudas técnicas, devido à componente tecnológica envolvida na sua concepção. A figura 2.1 ilustra um quadro de comunicação com selecção indirecta (cada célula do quadro inferior é identificada por intermédio de um par de coordenadas cuja escolha é efectuada com recurso ao quadro ilustrado em primeiro plano).

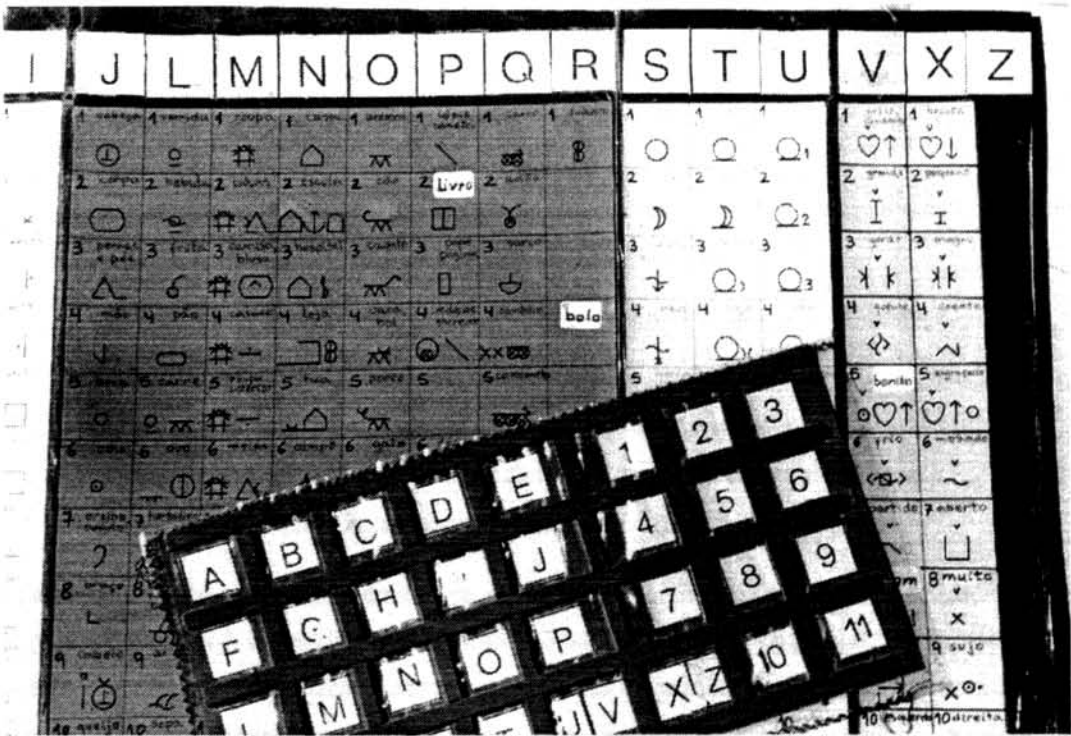


Figura 2.1 : Quadro de comunicação.

Perante um sistema de CAA é necessária a definição de uma estratégia, isto é, uma maneira específica de usar as ajudas e/ou técnicas de forma eficiente, que conduza à melhoria da comunicação. Uma estratégia, quer ensinada, quer apreendida, resultado de um processo de auto-descoberta, é um plano que visa facilitar o alcance de um elevado grau de desempenho do conjunto indivíduo/sistema de CAA.

O termo técnica refere-se ao método de selecção/transmissão das mensagens. As técnicas de selecção incluem a selecção directa, na qual o indivíduo indica directamente um símbolo, que por sua vez transmite ao parceiro de comunicação uma determinada mensagem, ou a selecção por varrimento, na qual o parceiro de comunicação ou um método electrónico vai possibilitando consecutivamente várias escolhas, permitindo ao indivíduo assinalar a sua opção quando o símbolo desejado aparece. O assinalar da escolha, quer na selecção directa, quer na selecção por varrimento, pode ser efectuada de várias formas, como o uso de sinalética, gesticulação, ou o accionar de qualquer dispositivo por meio de um movimento corporal não intempestivo.

Estes quatros componentes - símbolo, ajuda, estratégia e técnica - são elementos críticos, e como veremos ao longo deste trabalho, omnipresentes, que enquadram todas as intervenções na área da CAA.

O papel que os sistemas de CAA devem desempenhar, pode ser sintetizado em três pontos principais:

- proporcionar um meio temporário de comunicação até que se estabeleça a fala ou a mesma se torne apropriada, isto é, funcional e inteligível;
- proporcionar um meio de comunicação permanente, quando a aquisição da fala seja totalmente impossível ou improvável;
- proporcionar um meio para facilitar (aumentar) o desenvolvimento da fala propriamente dita e/ou, em alguns casos, das capacidades cognitivas e comunicativas para aquisição da linguagem.

À parte estes três pontos principais, os sistemas de CAA podem cumprir muitas outras funções orientadas à normalização do processo educativo, à ocupação laboral, ao ócio e ao controlo do espaço físico envolvente dos indivíduos com incapacidade, principalmente se esta se estender à área motora, como é o caso das pessoas com paralisia cerebral. A capacidade de intervir sobre o meio envolvente, mediante actos comunicativos, é

importante para o desenvolvimento da capacidade comunicativa intencional e facilitadora do desenvolvimento global do indivíduo.

Muitas das perturbações da comunicação resultam de problemas congénitos. Pessoas com atrasos de desenvolvimento exibem por vezes distúrbios ao nível da fala. Em casos de atraso severo ou profundo, a capacidade de comunicar através da fala pode ser grandemente afectada, implicando o uso de técnicas alternativas de comunicação. Estão neste caso os autistas. O autismo caracteriza-se por perturbações de linguagem, incapacidade de comunicar e estabelecer relações normais com os outros, isolamento, desinteresse pelo meio envolvente, incapacidade de resposta adequada a estímulos externos e fenómenos ritualistas. As pessoas com autismo podem beneficiar grandemente dos sistemas de CAA para desenvolvimento da comunicação e tentativa de quebra do seu auto-isolamento [BM92].

Outro grupo que pode colher grandes benefícios dos sistemas de CAA são as pessoas com paralisia cerebral. A paralisia cerebral é um distúrbio incapacitante, permanente mas não progressivo, resultante de danos cerebrais ocorridos antes, durante, ou imediatamente após o nascimento. Os danos cerebrais resultam numa perda do controlo voluntário sobre os músculos conjuntamente com uma tonicidade muscular inadequada. Tipicamente, esta população apresenta pouco ou nenhum uso das mãos ou outras extremidades para tarefas motoras precisas. Alguns membros deste grupo são incapazes de falar. O resultado é que os seus possíveis canais de comunicação estão extremamente limitados. Na sua maioria, no entanto, as pessoas com paralisia cerebral apresentam um bom potencial cognitivo cujo desenvolvimento depende, em grande parte, da possibilidade de serem quebradas as barreiras impostas pela impossibilidade física de comunicarem com a envolvente. Os sistemas de CAA desempenham neste caso papel fundamental na quebra desse isolamento e no desenvolvimento do indivíduo [Bar84].

Incapacidades sensoriais (isto é, a cegueira e a surdez) podem igualmente afectar o desenvolvimento das aptidões vocais ou verbais. Outros distúrbios são específicos da linguagem, como a apraxia evolutiva da fala, caracterizada por uma perda da faculdade de reproduzir os movimentos coordenados necessários à produção dos sons constituintes da fala.

As situações descritas têm origem em problemas congénitos. No entanto, não sendo portador de um desses problemas, não está livre de o adquirir, quer de forma temporária,

quer mesmo permanente.

“Imagine acordar num hospital, rodeado de doutores, enfermeiras e familiares com aspecto preocupado. Imagine não ser capaz de entender o que está a ser dito, nem ser capaz de falar. Imagine a vida sem palavras...sem linguagem”[Bla91]. O quadro clínico descrito é resultado de um acidente vascular cerebral que afectou o hemisfério esquerdo do cérebro, onde reside a faculdade da linguagem, conduzindo a uma afasia grave. Esta perturbação pode igualmente manifestar-se de diversas outras formas. A mais comum é caracterizada por um falar hesitante, estilo telegrama. O afásico ouve os seus erros e tenta corrigi-los. O praguejar é comum devido ao facto da linguagem emocional estar associada ao hemisfério direito do cérebro. Outra forma é caracterizada por um falar fluente mas incompreensível para o ouvinte. O afásico inventa novas palavras sem se aperceber disso. A afasia pode igualmente ser acompanhada de dificuldades de leitura, escrita e aritmética. A recuperação, quando possível, requer um treino prolongado, durante o qual o uso de sistemas de CAA permite de certa forma contornar o problema da falta de comunicação [Lev94]. De salientar que as capacidades intelectuais e mentais não são, na maioria dos casos, afectadas.

Os danos cerebrais podem igualmente resultar em alguns casos numa disfasia global. A disfasia global é caracterizada pela perda da fala, falta da capacidade de escrever, pouco ou nenhum reconhecimento das letras ou de palavras completas, gestos limitados e incerteza nas respostas sim/não. Todavia, algumas funções cognitivas podem ser poupadas resultando em pacientes com alguma capacidade para reconhecer, entender e apontar algumas palavras ou figuras [GME93].

As dificuldades de comunicação podem ainda surgir como resultado de algumas outras doenças ou devido a acidentes ou intervenções cirúrgicas. A doença de *Parkinson*, um síndrome composto por um conjunto de sintomas motores que incluem tremores quando em descanso, rigidez, redução de movimentos e alterações nos reflexos posturais, é um desses casos. Esta doença afecta sobretudo pessoas de idade avançada. O aparecimento de disartria (dificuldade na articulação das palavras) em indivíduos com a doença de *Parkinson* é bastante comum. No entanto, os distúrbios de comunicação entre estes doentes não são uniformes. Alguns tornam-se difíceis de entender devido à extrema rapidez com que falam, enquanto outros tendem a falar demasiado baixo. Com o avançar da doença muitos pacientes apresentam combinações várias destes distúrbios.

Outro caso passível de gerar perturbações comunicativas é a esclerose lateral amiotrófica, uma doença degenerativa progressiva de etiologia desconhecida envolvendo os neurónios motores do cérebro e da medula espinal. A esclerose lateral amiotrófica não conduz a dificuldades de linguagem, mas a disartria, resultado da fraqueza e dos movimentos convulsivos dos músculos inerente a essa doença.

A esclerose múltipla, uma doença degenerativa do sistema nervoso central, produz igualmente disartria na quase totalidade dos seus pacientes.

Como resultado de vários tipos de acidentes é frequente encontrar-se sinistrados com lesões ao nível da coluna cervical (lesões na área em redor do pescoço) que apresentam dificuldades na escrita e, ocasionalmente, igualmente distúrbios da fala. As lesões da medula espinal ocorrem em resultado de uma contusão ou esmagamento provocado por uma fractura, ou deslocamento, causada pela ruptura dos ligamentos entre vértebras individuais, ou ambos. Acidentes automóveis e quedas são os grandes responsáveis por este tipo de lesões [BM92].

Lesões traumáticas a nível cerebral resultando em danos cerebrais temporários ou permanentes são igualmente bastante comuns. Os distúrbios de comunicação associados às lesões traumáticas cerebrais variam consoante o grau e localização da lesão, podendo ser consequência de problemas cognitivos, dano de áreas específicas do cérebro ligadas à linguagem, resultando em sintomas idênticos à afasia, ou danos nas áreas cerebrais de controlo motor, resultando em disartria.

Intervenções cirúrgicas, como a laringectomia, podem conduzir igualmente a incapacidades temporárias ou permanentes a nível da fala.

Um caso normalmente esquecido é o de pacientes sujeitos a traqueotomia ou necessitando de ventilação automática durante a sua permanência nas unidades de cuidados intensivos [KS94]. O pessoal hospitalar não parece muito alertado para as necessidades de comunicação do paciente, preocupando-se mais com outras necessidades consideradas básicas como a nutrição e a higiene. Talvez esse o motivo da quase inexistência de sistemas de CAA nos hospitais [BSB93, BBS94].

A área de aplicação dos sistemas de CAA é, como visto, bastante extensa. Apesar disso não parece ter havido um fio condutor na sua evolução. A primeira forma de comunicação alternativa a surgir foi a linguagem gestual para surdos, cuja origem remonta à França

do séc. XVIII. No entanto, o sistema francês não se universalizou, tendo surgido nos últimos dois séculos múltiplos sistemas gestuais em diferentes países, com o mesmo objectivo. Alguns desses sistemas viriam também a ser usados com indivíduos com perturbações intelectuais, estando reportados algumas intervenções em que o seu uso foi bem sucedido [BM92].

Um dos sistemas de ajuda à comunicação que mais sucesso obteve foi o *Braille*. Desenvolvido no século passado por um francês, *Louis Braille*, consiste num sistema de símbolos tácteis que permite o acesso à leitura e à escrita por parte de pessoas cegas. Os caracteres *Braille* são formados por combinações de seis pontos em relevo, organizados em “células” de duas colunas verticais com três pontos cada. Os caracteres representam letras, partes de palavras, ou palavras inteiras. Cada carácter é formado de acordo com um modelo normalizado.

Os sistemas simbólicos representacionais usados actualmente, e de que veremos alguns exemplos mais adiante, surgiram, na sua maioria, já na 2ª metade deste século. O aparecimento dos vários sistemas não resulta de qualquer esforço conjunto e sistemático de colmatagem de várias necessidades, mas antes de experiências empíricas de vários grupos dispersos, lidando muitas das vezes com problemas comuns, mas sem estabelecerem muitos contactos entre si.

Um dos sistemas simbólicos actualmente em uso, os símbolos *Bliss*, criados por *Charles Bliss* em 1949, só viria a ser usado para colmatar deficiências de comunicação já no início dos anos 70. A ideia inicial era o de criar um sistema simbólico internacional, comparável ao *Esperanto*, que minimizasse as dificuldades de entendimento entre diferentes culturas e nações [MB94].

2.2 Objectivos principais dos sistemas de CAA

Os sistemas de CAA devem preencher não só certas necessidades básicas de comunicação do indivíduo, mas igualmente permitir a sua inclusão na sociedade, como membro de pleno direito, e a livre expressão das suas ideias. Um objectivo fundamental a perseguir no desenvolvimento de sistemas de CAA deverá ser o de tornar a comunicação mais natural e extensiva às várias actividades da vida. Intervenções bem sucedidas na área da CAA devem proporcionar uma melhoria das relações sociais e da segurança, um maior

auto-controlo e auto-domínio, a participação na educação, na vida familiar e da comunidade, aumentar as possibilidades de emprego e consequentemente a independência [Dem94].

Os sistemas de CAA devem não só proporcionar ao indivíduo um meio para expressar as suas necessidades e pedidos, mas igualmente permitir a transferência interactiva de informação e contribuir para uma maior aproximação, em termos sociais, do indivíduo, permitindo o desenvolvimento de relações inter-pessoais. Se no primeiro caso, o da expressão de uma necessidade, o conteúdo da mensagem é o mais importante e o vocabulário mais ou menos previsível, sendo a sua elaboração em termos sintácticos apenas relativamente importante, nos casos seguintes encontramos situações diversas. A transferência de informação é de mais difícil satisfação uma vez que implica o uso de um vocabulário mais vasto que permita ao indivíduo expressar-se sobre um leque alargado de temas. A sintaxe assume nestes casos uma maior importância visto poder afectar a boa compreensão da mensagem que se pretende veicular. No último caso manifesta-se igualmente o problema do vocabulário, uma vez que o conteúdo das mensagens não é previsível. No entanto, nestes casos, o ênfase deve colocar-se mais na interacção que deve ser mantida entre os participantes na conversação, por forma a não se perder o seu interesse, de que propriamente no conteúdo ou no maior ou menor grau de correcção sintáctica da mensagem.

A maior parte da pesquisa e dos desenvolvimentos técnicos na área da CAA têm-se focado na expressão das necessidades do indivíduo desprezando-se a indispensável parte social, devido principalmente ao facto dos sistemas com elevado conteúdo tecnológico estarem limitados na sua capacidade de armazenamento de vocabulário e apresentarem algumas dificuldades de acesso a um grande número de mensagens, o que origina longos períodos de tempo morto entre mensagens, propiciando o desinteresse dos participantes nos casos em que é exigida uma elevada interacção da conversação.

Ao dispensar a ajuda de terceiros para comunicarem entre si, é concedido aos utilizadores um maior grau de privacidade no seu relacionamento pessoal. Por tal motivo, aos sistemas de CAA deve igualmente ser adicionado vocabulário que possibilite a expressão de sentimentos e interesses do utilizador e que permita manter uma conversação afectiva mais profunda. A constrição provocada por limitações na expressão afectiva não só frustra como limita o potencial do utilizador desenvolver relações afectivas [SK94].

Na concepção dos sistemas de CAA há que ter em conta alguns aspectos para que a interacção entre o utilizador do sistema e a sociedade seja realmente efectiva, contribuindo para o sucesso da intervenção [Tre94].

A aprendizagem da utilização de um sistema de CAA deve ser fácil. A lista de regras para a operação de um dispositivo de comunicação deve ser mínima e simples. As regras não devem conter termos complexos, relações complexas e dependências contextuais. É essencial que as funções mais importantes sejam acedidas facilmente. A necessidade de procura de informação num manual ou o acesso a menus de ajuda deve ser evitada. Qualquer destas actividades quebra a interacção comunicativa conduzindo a sentimentos de frustração em ambos os parceiros de comunicação. O dispositivo deverá permitir que o indivíduo o possa utilizar imediatamente sem necessidade de um longo período de aprendizagem e treino. Para um indivíduo letrado, um sistema baseado na linguagem escrita deverá poder ser efectivamente usado sem treino especial. A possibilidade de usar imediatamente o dispositivo contribuirá para um aumento da sua auto-confiança numa área de actividade, a comunicação, que não domina.

Deve ser mantida uma constante consistência na operação do sistema de CAA. Os princípios apreendidos no início da utilização devem ser generalizados a estádios mais avançados da aprendizagem. Alterações nas funções de controlo da selecção requerem a mudança de procedimentos já intuídos, conduzindo a situações de desperdício de tempo causadoras de frustração no indivíduo. A manutenção da consistência no acesso do indivíduo a outros meios de controlo da envolvente ambiental é igualmente benéfica pelos mesmos motivos.

Desejável é, igualmente, a redução ao mínimo da necessidade de tradução da mensagem veiculada pelo utilizador. Por exemplo, o sistema de linguagem gestual usado pelos surdos-mudos é limitativo da sua integração social visto apenas os próprios, ou o reduzido número de pessoas que com eles convivem, serem capazes de o entender.

A comunicação interactiva envolve contacto visual. Quer falando ou gesticulando, os participantes numa conversação tendem a olhar-se nos olhos. O olhar é um complemento da comunicação revelando sentimentos e atitudes. A possibilidade de usar o dispositivo comunicativo sem necessidade de procurar os comandos desejados permitirá que o utilizador mantenha o contacto visual com o seu parceiro de comunicação.

A necessidade do uso da capacidade motora para accionamento dos comandos deve

igualmente ser mínima. A exigência de um esforço suplementar de concentração que permita o controlo dos comandos, provoca uma dispersão da atenção da conversação, mesmo nos casos em que não se tenha de seguir visualmente a acção [Lig93].

Um axioma que deve estar sempre presente quando se concebe um sistema de CAA é que este deve ir de encontro às necessidades do indivíduo. Muitas das vezes é esperado que os utilizadores modifiquem o seu comportamento em presença da tecnologia, em vez de esta ser concebida em conformidade com as suas necessidades, capacidades e comportamentos. Se um dispositivo de CAA impõe ao utilizador exigências que este julga excessivas, evitará ou procurará reduzir o seu contacto com ele. Os efeitos deste comportamento na efectivação do processo comunicativo serão, claro, altamente negativos.

A melhoria da fluência da comunicação tem sido ultimamente um dos focos principais da investigação e do desenvolvimento nesta área. A fluência é habitualmente tão parca que exclui efectivamente o indivíduo da interacção social em termos verbais, não sendo claramente conforme com o comportamento humano.

Tendo em vista o alcançar dos objectivos enunciados para as várias áreas de aplicação referidas, têm sido ao longo dos anos desenvolvidos um elevado número de sistemas de CAA, alguns dos quais, os mais relevantes, são apresentados na secção seguinte.

2.3 Principais sistemas de CAA

Os sistemas de CAA podem ser divididos em dois grandes grupos: os sistemas sem ajuda; e os sistemas com ajuda.

Os sistemas sem ajuda são aqueles que, como deriva da definição de ajuda apresentada anteriormente, não requerem o uso de qualquer objecto ou dispositivo físico. Os movimentos da cara, da cabeça, das mãos, dos braços e doutras partes do corpo são os únicos mecanismos físicos para a transmissão das mensagens. Incluem-se neste grupo os gestos de uso comum e os sistemas de sinais manuais usados pelos surdos. São exemplos de gestos de uso comum a afirmação e a negação com movimentos da cabeça, bem como todos os outros usados de forma mais ou menos inconsciente que enfatizam a comunicação oral [CSR94].

Os denominados sistemas com ajuda requerem algum tipo de assistência externa,

instrumentos ou ajudas técnicas para que possa ter lugar a comunicação. Por exemplo, pode necessitar-se de papel e lápis, livros ou quadros de comunicação, máquina de escrever, computador pessoal e outras ajudas técnicas de complexidade diversa, especialmente desenhadas segundo as necessidades especiais do utilizador.

Estes sistemas cobrem um amplo repertório quanto ao tipo de elementos de representação, que podem ir dos mui iconográficos aos totalmente abstractos, e quanto à complexidade das suas regras formais e combinatórias. Os sistemas mais simples são compostos por um conjunto de objectos, miniaturas, fotografias ou desenhos fotográficos, que o indivíduo assinala ou indica de alguma forma, com fins comunicativos. Os mais complexos utilizam, como elementos de representação, as palavras ou frases impressas do alfabeto gráfico tradicional ou codificado através de sistemas como o *Braille* e o *Morse*. Entre estes extremos existem diversos sistemas baseados em símbolos pictográficos, ideográficos (que mantêm uma relação conceptual com aquilo que representam) e arbitrários, que representam conceitos ou palavras.

Muita da magia destes sistemas recai precisamente no vasto conjunto de símbolos e sinais, que não os empregues na fala, e que podem ser usados para enviar mensagens. Especialmente para os indivíduos que não são capazes de escrever e/ou ler, a possibilidade alternativa de representar mensagens e conceitos é central para a comunicação [BM92, Pap94]. É o reconhecimento dessa importância que está por detrás de todo o esforço dedicado ao estudo e desenvolvimento de sistemas de símbolos que sejam fáceis de usar e entender.

Várias definições são usadas para descrever e classificar as várias formas de representação simbólica. Basicamente, e segundo [LB92], um símbolo é uma representação de um referente, isto é, algo que alude a qualquer outra coisa.

É importante enfatizar que um símbolo está profundamente ligado às raízes culturais de uma sociedade. Um mesmo símbolo pode aludir a dois referentes diferentes em duas sociedades com bases culturais distintas.

Nalguns casos, a relação entre o símbolo e o seu referente, quando este é conhecido, é facilmente perceptível. O símbolo é altamente translúcido. O termo iconicidade refere-se ao grau de relação visual entre um símbolo e o seu referente. Num dos extremos desta iconicidade encontram-se os símbolos transparentes, os quais são similares aos seus referentes, e como tal facilmente adivinháveis na ausência deste. Geralmente um símbolo

transparente descreve a forma ou a função do referente. Os símbolos transparentes são frequentemente classificados como altamente translúcidos. No outro extremo estão os símbolos opacos, cuja relação visual com o seu referente não é óbvia e pode ser bastante arbitrária. Por exemplo, uma fotografia de um copo é bastante transparente, mas a palavra escrita “copo” é opaca.

Os símbolos podem ser divididos em dois grupos segundo a sua tangibilidade. Símbolos tangíveis são aqueles que se podem discriminar com base nas sensações tácteis que transmitem, como por exemplo, forma, textura e consistência. Segundo esta divisão, símbolos bi-dimensionais (símbolos gráficos) são não-tangíveis. Os símbolos tangíveis são tipicamente usados com indivíduos com problemas sensoriais (visuais e/ou auditivos) ou com deficiências intelectuais graves, mas podem igualmente ser apropriados para outras populações. Objectos reais ou miniaturas de objectos idênticos, similares ou associados aos seus referentes, são exemplos de símbolos tangíveis. Os símbolos bi-dimensionais incluem fotografias ou desenhos, e podem ser usados para representar vários conceitos. Existe actualmente uma multiplicidade de diferentes conjuntos de símbolos.

Originalmente desenvolvidos com o objectivo de servirem como uma verdadeira linguagem internacional, os símbolos *Bliss* consistem em desenhos a traço negro sobre fundo branco. Este sistema tem como base um conjunto de cerca de 2250 símbolos representando aproximadamente 4000 palavras. Os símbolos são formados a partir do arranjo em várias formas de cerca de 50 símbolos básicos. Os símbolos *Bliss* incluem símbolos pictográficos, no sentido em que representam palavras ou ideias visualmente associáveis ao referente (transparentes), ideográficos, representando conceitos abstractos através de formas associadas com o referente, e abstractos, cuja relação com o referente não é perceptível mas apenas convencional. A figura 2.2 ilustra alguns desses símbolos.

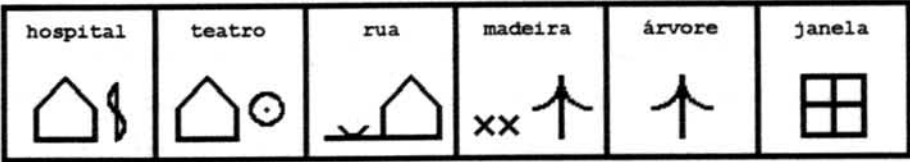


Figura 2.2 : Exemplos de símbolos *Bliss*.

Devido à simplicidade da sua estrutura, o sistema tem ganho bastante popularidade em muitos países, tendo vindo a ser usado com sucesso junto de pessoas com paralisia cerebral [BF93, SKR+94]. Muitos sistemas, embora efectivos noutros casos, apresentavam limitações para este grupo particular de pessoas. Por exemplo, o uso de

quadros com o alfabeto, conforme se ilustra na figura 2.3, pressupõem capacidade para ler e escrever, e tornam a comunicação extremamente lenta. O parceiro de comunicação tenta aperceber-se, uma a uma, sucessivamente, das letras para onde o utilizador “aponta” com o olhar, indo assim construindo as palavras.

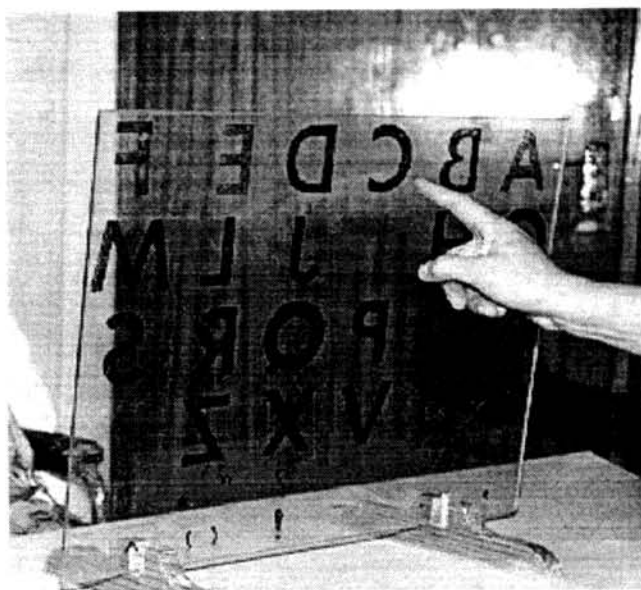


Figura 2.3 : Exemplo dum quadro com alfabeto.

Os símbolos *Bliss* permitem, devido aos seus conceitos relacionados, quer pictográficos, quer ideográficos, ultrapassar as restrições impostas à informação transmissível pela utilização dos quadros de imagens. Os conceitos abstractos podem ser simbolizados duma maneira sistemática e lógica, ultrapassando as limitações de imagens concretas referenciadas. Cada símbolo, quando exposto num quadro, pode ser acompanhado pela palavra apropriada, permitindo a sua fácil compreensão por estranhos ao sistema. Os princípios e estratégias de combinação dos símbolos permitem a expressão de pensamentos não presentes no quadro de comunicação. Os símbolos são conceptualmente baseados e construídos usando regras consistentes e sistemáticas. Aliás, como visto anteriormente, a manutenção da consistência ao longo da aprendizagem é um dos aspectos importantes a ter em conta na concepção de sistemas de CAA.

Para além dos símbolos *Bliss*, encontramos outros sistemas como os SPC (Símbolos Pictográficos de Comunicação) ilustrados na figura 2.4, símbolos *Rebus*, *Picsyms* e símbolos PIC (*Pictogram Ideogram Communication*).

Os símbolos SPC consistem em desenhos a traço negro que representam substantivos,

verbos e adjetivos. O fundo sobre o qual o símbolo é desenhado pode ser branco ou de outra cor. É normal o emprego de cores comuns aos vários géneros (substantivos, verbos e adjetivos). Os *Picsyms* consistem em cerca de 1800 símbolos igualmente a traço negro e sobre fundo branco, indexados alfabeticamente. Existem ainda instruções para a criação de símbolos adicionais, a partir de uma grelha de seis linhas por oito colunas. A figura 2.5 ilustra a criação de dois símbolos *Picsyms*.

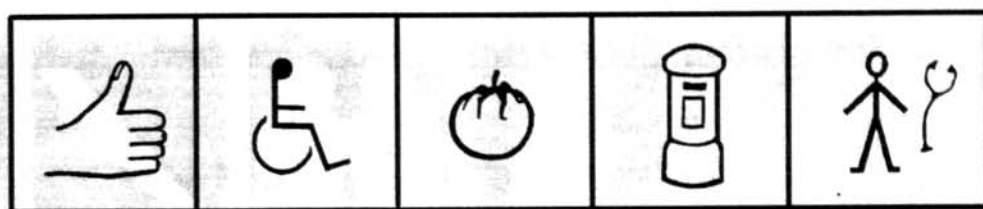
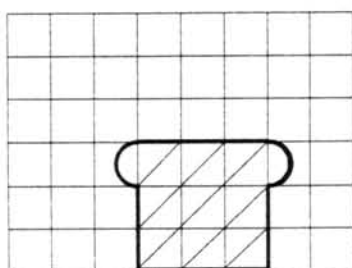
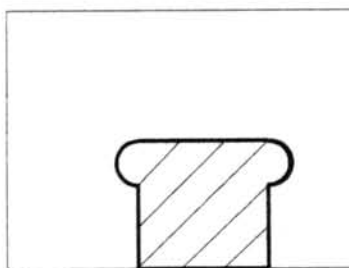


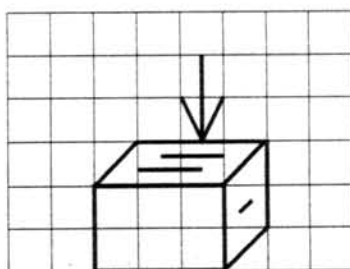
Figura 2.4 : Exemplos de símbolos SPC.



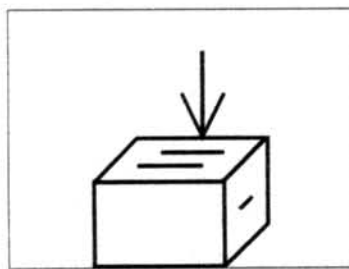
TORRADA



TORRADA



TORRAR



TORRAR

Figura 2.5 : Exemplo de criação de símbolos *Picsyms*.

Os sistemas diferem entre si quanto à transparência e translucidez dos símbolos, tendo em conta a população que visam atingir. Estudos indicam que os símbolos SPC e *Picsyms* são mais transparentes que os símbolos *Bliss* para indivíduos adultos com deficiências

intelectuais. Em geral os símbolos SPC são mais facilmente apreendidos que os símbolos *Bliss* e *Picsyms*. Os símbolos SPC têm sido usados com sucesso em casos de pessoas com deficiências intelectuais, paralisia cerebral e autismo, entre outras [BM92].

Muitas vezes confundidos com os *Picsyms*, os símbolos PIC consistem num conjunto de 400 símbolos em preto e branco concebidos para reduzir dificuldades de discriminação entre figura e fundo. A figura 2.6 ilustra alguns desses símbolos.



Figura 2.6 : Exemplos de símbolos PIC.

Todos estes sistemas de símbolos organizam-se tradicionalmente em carteiras, livros ou quadros de comunicação. As carteiras são facilmente transportáveis mas não permitem que um grande número de símbolos seja mostrado simultaneamente. Os livros possibilitam o armazenamento de um grande número de símbolos, vantajoso para utilizadores com um vocabulário extenso. Permitem igualmente o arranjo dos símbolos segundo a sua frequência de utilização, categoria semântica, ou qualquer outro formato lógico. No entanto, tanto as carteiras como os livros exigem o mínimo de capacidade motora que permita a sua manipulação. Os quadros de comunicação são habitualmente superfícies planas aonde é afixada uma matriz de símbolos bi-dimensionais. O tamanho dos quadros é determinado quer pelas necessidades de comunicação do utilizador, quer pelas suas limitações físicas.

O uso da ortografia tradicional é desejável se ambos, utilizador e parceiro de comunicação, são capazes de ler. Quadros consistindo em palavras frequentemente usadas, letras do alfabeto e números, permitem ao utilizador criar um número ilimitado de mensagens. A associação entre os símbolos e a ortografia tradicional é bastante útil quando o parceiro de comunicação não conhece o conjunto de símbolos usados.

Um indivíduo que comunique através do uso de símbolos gráficos deve indicar de uma qualquer forma, ao parceiro de comunicação, qual o símbolo que quer seleccionar. Indivíduos com capacidades motoras relativamente boas, podem simplesmente apontar para os vários símbolos. No entanto, para utilizadores com incapacidades motoras

acentuadas, este método de selecção não é o mais adequado, sendo necessário o recurso a técnicas alternativas, as quais implicam a participação activa do parceiro de comunicação na selecção dos símbolos.

A selecção directa, na qual o utilizador aponta o símbolo desejado, pode ser realizada pelo toque do símbolo com um dedo ou um apontador. Utilizadores com pouco ou nenhum movimento voluntário dos membros, mas um bom controlo dos movimentos da cabeça, podem usar apontadores adaptados a capacetes ou cintos ajustáveis. A figura 2.7 ilustra um apontador luminoso montado num cinto ajustável. A técnica de selecção descrita para o quadro alfabético ilustrado na figura 2.3, é igualmente uma técnica de selecção directa.

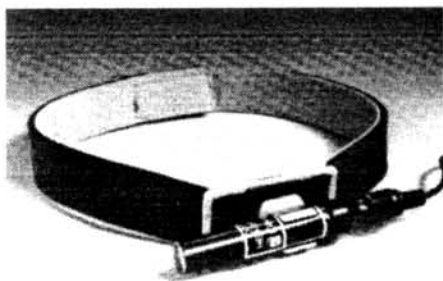


Figura 2.7 : Apontador luminoso para cabeça com cinto ajustável.

A selecção directa de símbolos, com ou sem o uso de apontadores ou outros dispositivos, é quase sempre o método mais rápido de selecção. Se o utilizador possui um controlo adequado dos movimentos voluntários, este será provavelmente o método de selecção a escolher. Caso contrário, será necessário recorrer a outros métodos de selecção, quer em exclusividade, quer como possível complemento da selecção directa.

Um dos métodos alternativos é o da selecção por varrimento. Na técnica de varrimento, são oferecidas sucessivamente ao utilizador, normalmente por parte do parceiro de comunicação, um conjunto de possíveis alternativas de escolha. O utilizador indicará, por meio de um gesto, sinal, ou de qualquer outra forma, ao parceiro de comunicação, a escolha do símbolo desejado quando este lhe for mostrado.

O uso exclusivo de varrimento manual resulta numa forte dependência em relação aos parceiros de comunicação para iniciar e manter a interacção, uma vez que o utilizador não pode comunicar a menos que lhe sejam fornecidas as possíveis alternativas de escolha.

Para especificação de uma determinada mensagem é por vezes necessário que o utilizador

selecione sucessivamente mais do que um símbolo. Esta técnica é referida como codificação. Algumas estratégias de codificação de mensagens incluem o uso de números, letras ou ícones. A codificação tem como objectivo o aumento da fluência da comunicação, principalmente no caso das mensagens mais frequentemente usadas. Um caso de codificação numérica, em que o utilizador selecciona primeiramente o algarismo das dezenas e posteriormente o das unidades, correspondendo o número a uma determinada mensagem, é descrito em [FTR+94]. Outra técnica de codificação que tem tido bastante sucesso é o da codificação iconográfica, normalmente referida como compactação semântica ou *Minspeak*¹. Símbolos pictográficos, denominados ícones, representando objectos e acções habitualmente presentes no dia-a-dia do utilizador, são seleccionados em diferentes combinações, resultando em alterações do seu significado. Por exemplo, uma laranja pode ser combinada com um ponto de interrogação para significar “Quando é que comemos?”, ou combinada com um copo para significar “sumo de laranja”. Isto permite a representação de uma vasta quantidade de vocabulário num único quadro, eliminando a necessidade de vários quadros e resultando numa comunicação mais fluente [Tat93].

Todas as ajudas descritas implicam uma grande participação do parceiro de comunicação na conversação, nomeadamente a sua constante atenção, por forma a ver qual o símbolo seleccionado e a tentar perceber qual o sentido da mensagem. Se o parceiro está distraído, dificilmente o utilizador conseguirá despertar-lhe a atenção.

Se um sistema de CAA deve proporcionar um meio de comunicação para pessoas com distúrbios da fala, o ideal seria que o sistema falasse. Como veremos no próximo capítulo, a introdução da electrónica e da informática nos sistemas de CAA veio tornar esse sonho realidade.

¹ *Minspeak* é uma marca registada da *Semantic Compaction Systems*.

3. Ajudas Técnicas à Comunicação

A introdução da componente tecnológica nos sistemas de CAA contribuiu de forma decisiva para o aumento da sua eficácia, tornando mais independentes os utilizadores e alargando o espectro de potenciais parceiros de comunicação. Se a utilização de símbolos obrigava ao conhecimento do seu significado e a uma grande atenção por parte do parceiro de comunicação, a introdução da saída de voz nos sistemas de CAA veio pôr fim a estes problemas. Ao usar a linguagem falada, a comunicação aproxima-se dos seus moldes naturais. Mais importante, as ajudas técnicas reduzem a responsabilidade comunicativa do utilizador ao permitirem a geração de mensagens bastante mais sofisticadas que aquelas permitidas pelas suas próprias capacidades. Para tornar tal possível, muito contribuíram as recentes evoluções nas técnicas de codificação e armazenamento dos sinais de voz. Os desenvolvimentos verificados nesta área permitiram melhorar a qualidade da voz reproduzida pelos sistemas de CAA, aumentando ao mesmo tempo o vocabulário disponibilizado ao utilizador. Importante na implementação dessas novas técnicas o aumento crescente das densidades de integração permitindo a construção de circuitos mais complexos, ocupando espaços mais reduzidos e consumindo menos energia, ponto fundamental para o desenvolvimento de sistemas portáteis.

A introdução da electrónica permitiu igualmente um alargar dos métodos de selecção disponibilizados, conduzindo à criação de *interfaces* que possibilitam uma maior eficácia na interacção entre utilizador e sistema.

Na primeira parte deste capítulo será analisada a influência da electrónica nas técnicas de selecção e serão vistos alguns dos sistemas disponibilizados comercialmente. Numa segunda parte abordam-se as novas técnicas de codificação dos sinais de voz e do seu armazenamento, e as suas possíveis influências sobre os sistemas de CAA.

3.1 Técnicas de selecção

O varrimento manual pode ser bastante útil como técnica de iniciação ao uso de um dispositivo com varrimento electrónico. O uso exclusivo de varrimento manual implica uma extrema dependência do utilizador em relação ao seu parceiro de comunicação para iniciar e manter a interacção, uma vez que o utilizador não pode comunicar a menos que o parceiro lhe ofereça certas escolhas. Para além do mais, o varrimento manual é uma maneira relativamente lenta de seleccionar símbolos. Num dispositivo com varrimento electrónico, a superfície do quadro de comunicação contém uma matriz de pequenas lâmpadas (uma para cada símbolo). As lâmpadas iluminam-se automaticamente, uma a uma, sucessivamente, ao longo do quadro, com um intervalo de tempo definido. O ponto iluminado é usualmente designado por cursor. Quando o cursor ilumina o símbolo desejado, o utilizador pressiona um botão, parando o seu deslocamento.

A técnica mais simples é a do varrimento linear, em que o cursor se move, símbolo a símbolo, da esquerda para a direita, sucessivamente ao longo das várias linhas. Esta técnica tem a vantagem de ser bastante previsível e de requerer apenas um único movimento para activar e desactivar o cursor. É todavia bastante lenta, o que pode ser uma vantagem para alguns utilizadores, mas é uma desvantagem para outros. Se o utilizador tiver dificuldades motoras, um varrimento linear, combinado com um deslocamento lento do cursor, permite-lhe planear atempadamente o movimento requerido para actuação do botão e sua consequente paragem. No entanto, se o utilizador falha a escolha, terá de esperar que o cursor dê uma volta completa ao quadro, o que pode ser bastante moroso. Um utilizador capaz de seguir um cursor com movimento rápido, e que possua uma rápida resposta motora, poderá usar esta técnica eficientemente.

Em utilizadores com uma resposta motora razoável, mas com dificuldades cognitivas profundas que os impossibilitem de antever a evolução da posição do cursor e a escolha que pretendem fazer, recorre-se a dispositivos em que o movimento do cursor não é automático, necessitando os utilizadores de pressionar o botão cada vez que pretendam que este avance para o símbolo seguinte.

Outro tipo de varrimento bastante usado é o varrimento linha-coluna. Neste caso, são iluminadas sucessivamente, de cima para baixo, as linhas do quadro. O accionamento do

botão implica que o cursor passe a deslocar-se, coluna a coluna, ao longo da linha seleccionada. Quando o símbolo desejado é iluminado, o utilizador pressiona de novo o botão para assinalar a sua escolha.

Tipos diferentes de varrimento podem ser usados consoante as capacidades do utilizador. Todavia, há que ter em conta que se o varrimento linha-coluna aumenta a eficiência dos sistemas com selecção por varrimento, implica igualmente um maior número de movimentos por parte do utilizador, uma vez que para uma escolha é necessário mais que um accionamento do botão, e uma maior capacidade cognitiva que lhe permita encadear uma sequência directa de respostas. Por exemplo, no caso do varrimento linha-coluna, o utilizador tem de ser capaz de visualmente localizar a linha na qual se encontra o símbolo desejado, assinalando que é essa que deve passar a ser varrida linearmente. Posteriormente o utilizador deverá localizá-lo dentro dessa linha e accionar o botão por forma a parar o cursor nesse símbolo.

A vantagem das técnicas electrónicas de varrimento é a de permitir, a utilizadores com incapacidades motoras significativas, uma selecção mais eficiente dos símbolos. O seu uso implica, no entanto, que o utilizador seja capaz de realizar pelo menos um movimento motor com um mínimo de confiança e repetibilidade. Uma vez identificado um comportamento motor apropriado, é necessário seleccionar o melhor tipo de botão para a sua captação. Os mais comuns são os botões de pressão. Estes podem ser actuados por qualquer parte do corpo e ter sensibilidade variável. A figura 3.1 ilustra o uso de um botão de pressão accionável com um movimento lateral da cabeça.

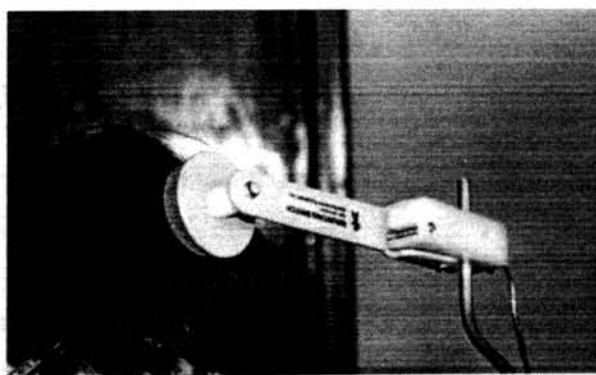


Figura 3.1.: Exemplo de um botão de pressão accionável com um movimento lateral da cabeça.

Uma alternativa, no caso de não ser possível a identificação dum movimento motor que permita a actuação dum botão de pressão, é o uso de botões activáveis por pressão de ar.

O utilizador segura entre os lábios a ponta de um tubo flexível e activa o botão soprando ar para o interior do tubo.

Embora o uso de varrimento electrónico permita uma maior liberdade do utilizador na escolha do símbolo desejado, a compreensão da mensagem continua a estar dependente da atenção e interesse do parceiro de comunicação. A introdução de sistemas com saída de voz veio trazer à CAA um enorme potencial permitindo que a comunicação se processe em moldes mais naturais.

3.2 Ajudas técnicas com saída de voz

As ajudas técnicas à comunicação com saída de voz baseiam-se quer em fala sintetizada, quer em fala digitalizada. A fala sintetizada é produzida por um dispositivo que converte informação matemática pré-programada acerca dos parâmetros do tracto vocal, numa reprodução audível da voz humana [Edw91, Cor91]. A fala sintetizada é normalmente usada nos chamados conversores texto-voz, nos quais o utilizador constrói expressões a partir de uma entrada de texto (por exemplo, um teclado com acesso directo ou por varrimento) [Sag90]. A entrada é analisada e convertida por forma a simular informação fonética por um conjunto de algoritmos matemáticos pré-programados. A fala sintetizada difere da fala natural quanto à quantidade limitada de informação redundante que possui, como por exemplo a entoação, e que não raras vezes permite ao ouvinte discernir o seu significado mesmo que a inteligibilidade seja baixa [CHA+95]. Consequentemente, muitas aplicações com fala sintetizada são significativamente menos inteligíveis que a fala natural [LOT93].

Uma das tentativas para ultrapassar os problemas apresentados pela fala sintetizada, mantendo a possibilidade de reproduzir um vocabulário ilimitado, é o da concatenação fonética. Neste caso são guardados no sistema digitalizações dos fonemas constituintes da linguagem, os quais, quando sequenciados apropriadamente, podem gerar uma qualquer palavra. Um problema que surge é o da ligação entre os fonemas por forma a que as transições entre eles soem o mais natural possível. Outro ponto a ter em conta é que a quantidade e tipo de fonemas difere de idioma para idioma, e que, dentro de cada um, existem palavras que fogem às regras de pronúncia desse idioma.

No caso da fala digitalizada, a fala humana é amostrada e quantificada em conversores analógico-digitais, sendo a informação guardada em memória. Para a sua reprodução é realizada a operação inversa. Uma vez que a saída produzida resulta do processamento da fala humana, esta é aproximadamente tão inteligível como a fala natural. Para além do mais, a prosódia natural e outras características da voz do falante (como por ex. idade, sexo, entoação) são mantidas. Um senão desta técnica é a grande quantidade de memória requerida, limitando por tal o vocabulário ou frases que podem ser guardadas. As técnicas de codificação actuais, e que veremos na próxima secção, tendem a atenuar este problema.

Temos pois que quanto maior as unidades (fonemas, palavras ou frases), maior a naturalidade alcançada na reprodução, mas menor o grau de flexibilidade. Por exemplo, se todas as frases forem guardadas, então a pronúncia das palavras que as compõem estará correcta, mantendo-se igualmente a prosódia da frase. No entanto, estaremos confinados a expressões muito específicas. Por exemplo, a frase “Quero uma chávena de café”, pode ser guardada e reproduzida exactamente, mas o utilizador sentir-se-à frustrado se realmente preferir chá. Se em vez de usarmos a frase como unidade, usarmos unidades mais pequenas, palavras ou fonemas, aumentamos a flexibilidade mas diminuímos a naturalidade aquando da sua reprodução. Assumindo que cada uma das palavras da frase referida anteriormente foi guardada individualmente, mais a palavra “chá”, o utilizador pode agora escolher entre chá e café. Todavia, haverá uma perda geral de qualidade, pois a frase que resulta da junção das palavras não tem em conta a prosódia natural da fala.

Nas ajudas à comunicação com saída de voz a realização duma selecção provoca a reprodução audível da mensagem a ela associada. A *interface* com o utilizador é do mesmo tipo usado nos quadros de comunicação, mas agora o parceiro de comunicação não necessita de saber interpretar os símbolos escolhidos, visto o sistema reproduzir, em linguagem comum, a mensagem desejada [Hun94, Hek94].

Enquanto nas ajudas não electrónicas era necessário obter primeiramente a atenção do parceiro de comunicação, antes de se seleccionarem os símbolos, agora a comunicação poder-se-à iniciar mesmo que este esteja distraído. A reprodução audível da mensagem seleccionada despertará a sua atenção.

Existem disponíveis comercialmente alguns sistemas de CAA com saída de voz, quer baseados em fala sintetizada, em concatenação de fonemas ou em fala digitalizada.

Exemplos desses sistemas são o *Liberator*¹, o *Touch Talker*² e o *Light Talker*³. A base destes três produtos é o *DECtalk*⁴, um dos melhores sistemas de geração de voz disponíveis para língua inglesa, baseado na técnica da concatenação fonética [Edw91]. Uma das suas características mais importantes é a de oferecer uma gama de vozes masculinas e femininas e de diferentes idades, seleccionáveis pelo utilizador segundo a sua preferência. O *DECtalk* é basicamente um conversor texto-voz que aceita ficheiros de texto em código ASCII (*American Standard Code for Information Interchange*). Uma das razões que tornam a qualidade da fala produzida tão elevada, é o seu dicionário interno que contém cerca de 6 000 palavras de excepção (palavras que não seguem as regras normais de pronúncia). Considerando que o vocabulário médio activo de um indivíduo, com um grau académico ao nível da licenciatura, está estimado em 15 000 palavras, existe uma elevada probabilidade de a palavra a ser pronunciada estar no dicionário. Se tal não acontecer, a palavra é "passada" pelos algoritmos de conversão de texto para voz. A elevada qualidade da fala é, neste caso, parcialmente devida ao extensivo conjunto de regras aplicadas nesta fase [Edw91].

Os três sistemas, *Liberator*, *Touch Talker* e *Light Talker*, são baseados em matrizes de 128 células, as quais não são mais que teclas usadas para selecção directa. Para além deste tipo de selecção, o *Liberator* e o *Light Talker* oferecem ainda outras possibilidades como o varrimento linear, para o qual cada tecla dispõe de uma pequena lâmpada encarnada, podendo o controlo ser efectuado por um botão de pressão ou um apontador de infravermelhos. Este será colocado na cabeça do utilizador, ou em qualquer outra parte do corpo, sobre o qual tenha suficiente controlo. Quando o apontador é dirigido para uma das células, esta ilumina-se. Desta forma o utilizador pode seleccionar qualquer das células do dispositivo. A acção de selecção de uma dada célula pode ser executada de duas maneiras diferentes: quer deixando o apontador dirigido para a célula pretendida durante um tempo superior a um determinado intervalo pré-programado; ou pelo accionamento de um botão enquanto a célula está iluminada. A matriz pode ser configurada em qualquer dos dispositivos para oferecer 8 ou 32 células, por junção de células adjacentes, obtendo-se

¹ *Liberator* é uma marca registada da *Prentke Romich Company*.

² *Touch Talker* é uma marca registada da *Prentke Romich Company*.

³ *Light Talker* é uma marca registada da *Prentke Romich Company*.

⁴ *DECtalk* é uma marca registada da *Digital Equipment Corporation*.

desta forma células de maior tamanho permitindo responder às necessidades do utilizador com incapacidades visuais, cognitivas e/ou motoras graves.

Qualquer dos sistemas emprega a técnica da compactação semântica ou *Minspeak*, tendente a minimizar o número de selecções requeridas para construir uma expressão [Tat93]. A adição de uma nova expressão é uma operação bastante simples. O programador selecciona a célula ou a sequência de células onde deseja guardar a nova expressão, seguida da expressão teclada letra a letra. O alfabeto é uma das funções alternativas de algumas das células presentes nos três dispositivos. De igual maneira, o utilizador poderá formar uma palavra não programada previamente em qualquer célula. Esta possibilidade torna virtualmente ilimitado o vocabulário disponível. De salientar que, sendo a operação de programação de uma nova expressão idêntica à usada para a selecção, o utilizador não depende de terceiros para programar o seu próprio sistema, o que o torna totalmente independente quanto ao conteúdo das mensagens permitindo-lhe manter um elevado grau de privacidade.

Os sistemas possuem ainda um mostrador de cristais líquidos onde a mensagem a ser reproduzida aparece escrita. A figura 3.2 ilustra uma utilização do *Liberator* em que um apontador de infravermelhos é usado para seleccionar a célula pretendida.

Um problema que ainda se mantém é o de o número de mensagens acedível de forma rápida e directa, ser limitado pelo tamanho do quadro que o utilizador consegue visualmente explorar. A introdução da informática, com o uso de quadros dinâmicos, permitiu ultrapassar esta limitação. Vários quadros, contendo cada um um número determinado de símbolos, vão sendo visualizados sucessivamente pelo utilizador. Quando o quadro que contém o símbolo correspondente à mensagem desejada aparece, o utilizador indica o facto ao sistema, que dá início ao varrimento desse quadro. Quando seleccionada, a mensagem é reproduzida [GME93, Wol94].

A informática apresenta no campo dos sistemas de CAA uma fonte inesgotável de possibilidades que só agora começam a ser exploradas. A composição guiada de mensagens em que, a partir de duas ou mais letras digitalizadas pelo utilizador, o sistema selecciona e apresenta um conjunto de possíveis palavras ou mensagens, vem incrementar grandemente a fluência da comunicação [ANA93, GKP+93, HP93, MMP+94]. No entanto, exige que o utilizador seja capaz de interagir com um computador

pessoal. Necessário se torna pois, a criação de *interfaces* multi-modais em que a captura de informação se processe por vários canais de entrada que incluam outros aspectos da linguagem corporal [RHF93, CSR94]. A figura 3.3 esquematiza algumas possibilidades alternativas de acesso a meios informáticos.



Figura 3.2 : Exemplo de utilização do *Liberator*.

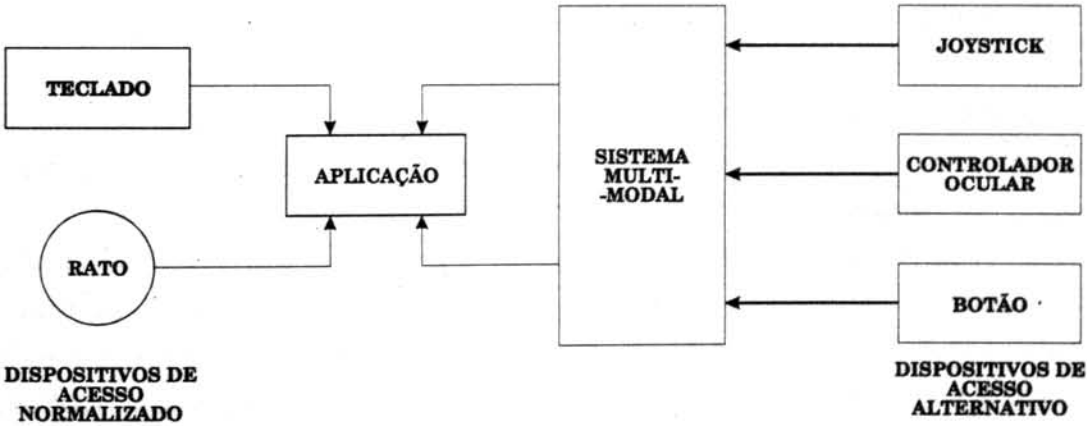


Figura 3.3 : Sistema multi-modal de acesso a meios informáticos.

Um dos sistemas de CAA baseado em meios informáticos, neste caso um vulgar

computador pessoal, é o *Equalizer*¹. Este sistema apresenta um menu principal onde se encontra um conjunto de palavras vulgarmente usadas pelo utilizador, e o alfabeto, letra a letra. Desta forma o utilizador poderá seleccionar uma palavra pré-programada, ou gerar no momento a palavra desejada. A técnica de selecção usada é a do varrimento, podendo o seu tipo ser programado em função das possibilidades do utilizador. A selecção é efectuada por meio de um único botão. O sistema “aprende” a linguagem normalmente usada pelo utilizador, o que lhe permite antecipar as prováveis próximas palavras que completarão a frase [Edw91].

Para além de poder ser usado como sintetizador de voz, o *Equalizer* pode ser usado como sistema de entrada para um processador de texto. Isto significa que o utilizador pode não só escrever discursos que serão mais tarde sintetizados pelo sistema, mas também preparar texto para ser imprimido.

O *Equalizer* é o sistema usado por *Stephen Hawking*, provavelmente o mais famoso utilizador dum dispositivo de CAA. *Stephen Hawking* é considerado por muitos como o maior físico desde *Albert Einstein*, mas uma doença degenerativa do sistema nervoso central limitou-lhe progressivamente os movimentos e obrigou-o a uma laringectomia. Graças ao *Equalizer* pode continuar o seu trabalho académico como professor e escritor. Para além de muitos prémios como reconhecimento pela sua brilhante carreira, teve o raro privilégio de, por mais de um ano, ter um livro seu na lista dos mais vendidos em todo o mundo. Esse livro, “Uma breve história do tempo”, foi escrito no seu *Equalizer*[Edw91].

A utilização de meios informáticos é no entanto limitada devido à pouca portabilidade dos sistemas. Para além do mais, os sistemas descritos são extremamente caros e não se encontram disponíveis com sintetizadores que permitam a sua expressão em língua portuguesa. Uma área a explorar é a da criação de sistemas portáteis com saída de voz de qualidade apreciável, de baixo custo e com possibilidade de expressão em português. Como vimos anteriormente, para um bom nível de qualidade da voz, a fala digitalizada é preferível à fala sintetizada. O problema a ultrapassar é o da quantidade de memória necessária para o armazenamento digital das mensagens. O que os recentes avanços nas

¹ *Equalizer* é uma marca registada da *Words+, Inc.*

técnicas de codificação de sinal e de armazenamento podem fazer para resolver esse problema, é o que exploraremos na próxima secção.

3.3 Técnicas de codificação de voz: o ADPCM

Os desenvolvimentos verificados na área das telecomunicações têm conduzido ultimamente a uma crescente oferta de novos serviços, implicando uma cada vez maior necessidade de transferência de elevados volumes de informação. O rápido esgotamento da capacidade dos canais de comunicação em alguns mercados é já uma possibilidade latente, justificando o crescente interesse nas tecnologias de codificação, especialmente compressão de voz e de imagem [Spa94]. Aliás, a motivação que conduziu há cerca de 50 anos atrás ao início da pesquisa na área da codificação da voz, era o desenvolvimento de um sistema que possibilitasse a transmissão de voz através da estreita largura de banda dos cabos telegráficos. Pioneiro neste trabalho, *Homer Dudley* dos *Bell Telephone Laboratories*, demonstrou a redundância existente nos sinais de voz e criou o primeiro método de codificação de voz. A ideia básica por trás do codificador de voz, ou *vocoder*, era a de analisar a voz em termos do seu tom (frequência fundamental) e do seu espectro, e sintetizá-la pela excitação de um banco de filtros passa-banda analógicos. O “vocoder” centrou atenções durante a II Guerra Mundial devido ao seu potencial na transmissão eficiente da voz encriptada. Implementações melhoradas do “vocoder” de *Dudley* estão referenciadas ao longo dos anos 50 e 60 [Spa94].

Embora as primeiras implementações do *vocoder* tivessem sido baseadas em representações analógicas do sinal de voz, a representação digital ganhou rapidamente interesse. A pesquisa na área da codificação de voz digitalizada tem-se intensificado. A percepção de que voz de alta qualidade pode ser alcançada com taxas de compressão que há alguns anos atrás seriam consideradas irrealistas, tem atraído a atenção dos investigadores [Dim93]. Resultados recentes indicam que se pode conseguir uma qualidade boa a excelente com taxas de 1 *bit*/amostra para voz e 2 *bit*/amostra para música. Reduções na taxa para 0,5 e 1 *bit*/amostra, respectivamente, podem ser esperadas na próxima década [Nol93].

O nosso interesse nas técnicas de codificação reside na possibilidade de, codificando-se o sinal de voz, aumentar a eficiência do armazenamento da voz digitalizada. Uma taxa de

1 *bit*/amostra permitiria, numa conversão analógico-digital de 8 *bits*, octuplicar o tempo disponível para as mensagens num sistema de CAA com saída de voz, baseado em voz digitalizada.

Deve aproveitar-se para salientar desde já a clara opção tomada pelos sistemas com voz digitalizada em detrimento dos sistemas baseados em voz sintetizada. Tal deve-se à pobre qualidade da voz produzida pelos sistemas sintetizados: na melhor das hipóteses, o resultado soaria robótico; na pior, seria ininteligível [Gal94].

Ao longo desta secção analisaremos alguns aspectos da voz humana, estudaremos a sua amostragem e quantificação, e alguns métodos actuais de codificação. Finalizaremos com a referência a algumas propostas de avaliação da qualidade da voz codificada, tendente a possibilitar uma análise comparativa das diferentes técnicas.

3.3.1 Características da fala

Antes de apresentarmos os métodos existentes para codificação da voz é importante analisarmos algumas das suas características de modo a melhor compreendermos o porquê de certas opções nas fases de amostragem, quantificação e codificação.

Os sinais de voz são não-estacionários, e o melhor que se pode considerar é o serem estacionários ao longo de pequenos segmentos, tipicamente 5 a 20 ms. As propriedades estatísticas e espectrais dos sinais são por tal definidas sobre estes pequenos segmentos.

A fala pode geralmente ser classificada como vocalizada, não-vocalizada, ou mista. O processo de produção da fala é iniciado pela libertação de ar dos pulmões para o tracto vocal. O ar pode simplesmente passar através da laringe, ou ser modulado por uma acção rápida e repetitiva das cordas vocais. Do primeiro caso, resulta um som não-vocalizado (tipicamente consoantes, como /s/, /t/ ou /f/), e do segundo, um som vocalizado (vogais ou consoantes vocalizadas, como /m/, /z/ ou /g/). A diferença entre estes dois tipos de sons é evidente quando analisamos o sinal acústico resultante, ilustrado na figura 3.4 [Kel94].

A fala vocalizada é quasi-periódica no domínio temporal e estruturada harmonicamente na frequência (uma frequência fundamental e vários harmónicos), enquanto a fala não-vocalizada é aleatória temporalmente e distribuída na frequência. Em adição, a energia

dos segmentos vocalizados é geralmente mais elevada que a energia dos segmentos não-vocalizados.

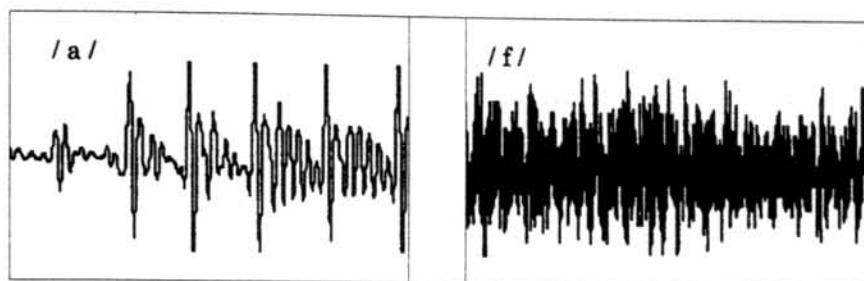


Figura 3.4 : Segmentos de sinais acústicos de fala vocalizada e não-vocalizada.

O tracto vocal consiste na laringe, na faringe e na cavidade vocal. A forma da envolvente espectral (envolvente do espectro de potência) está relacionada com as características de transferência do tracto vocal: Essa envolvente é caracterizada por um conjunto de picos chamados formantes. Os formantes indicam as frequências de ressonância do tracto vocal. Em média existem 3 a 5 formantes abaixo dos 5 kHz. A amplitude e localização dos primeiros 3 formantes, os quais ocorrem usualmente abaixo dos 3 kHz, são bastante importantes para a percepção e inteligibilidade da voz [Spa94]. A figura 3.5 ilustra a forma típica da envolvente espectral dum segmento de fala vocalizada. Este ponto é de importância fundamental na correcta escolha das frequências de amostragem a usar na digitalização da voz. Os formantes localizados acima dos 3 kHz são usualmente de baixa amplitude e resultante de sons não-vocalizados.

3.3.2 Digitalização da fala

3.3.2.1 Amostragem

Os sinais de voz são, obviamente, sinais analógicos, isto é, sinais contínuos no tempo. Para os podermos guardar numa memória é necessário convertê-los em sequências numéricas. O primeiro passo para a representação digital da forma de onda da fala não envolve uma medida constante da sua amplitude, mas uma amostragem a intervalos fixos. Por outras palavras, a amplitude é medida, e após um determinado período de tempo, a operação é repetida. A amostragem consiste assim na tomada de valores instantâneos do sinal.

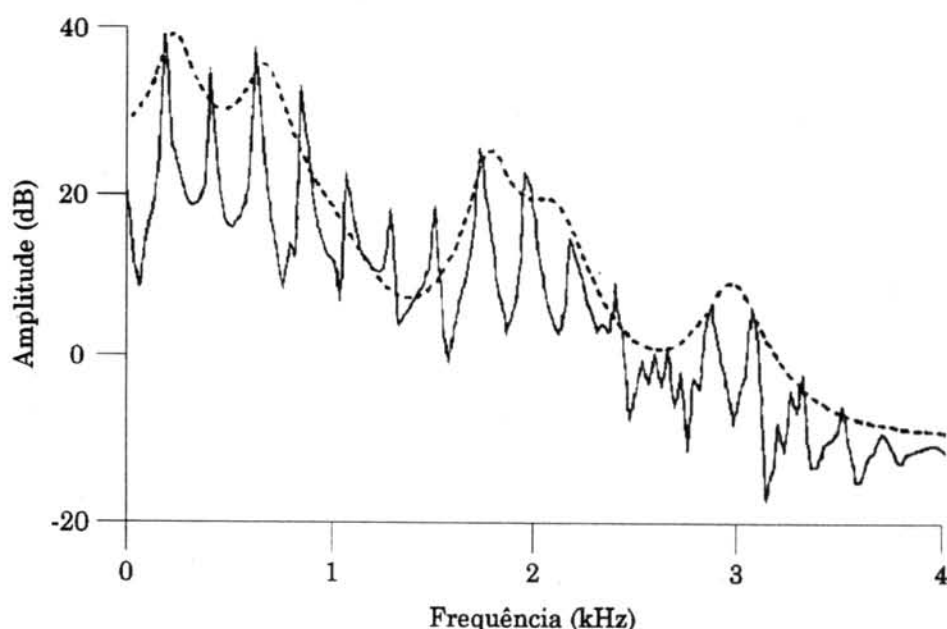


Figura 3.5 : Envolvente espectral de um segmento de fala vocalizada.

Um elemento crítico na correcta captura da informação existente no sinal de voz é a frequência com que essa amostragem é realizada. O Teorema da Amostragem de *Nyquist* define a frequência mínima requerida para representar correctamente um sinal analógico. Quando amostramos um sinal contínuo no tempo de modo a produzir uma sequência de amostras desse sinal, queremos assegurar que a informação do sinal original é retida pelas amostras. Não haverá perda de informação se conseguirmos recuperar exactamente o sinal original a partir das suas amostras [Kuc88]. O teorema define que para tal ser possível, a frequência de amostragem, f_s , deverá ser maior ou igual a duas vezes a maior das componentes frequenciais, f_m , do sinal analógico amostrado.

O efeito da amostragem dum sinal é similar ao da multiplicação dum sinal por uma série de ondas sinusoidais com frequências múltiplas da frequência de amostragem f_s . Isto significa que a amostragem é similar à modulação analógica – o espectro do sinal de entrada é repetido para cada f_s . Consideremos o sinal analógico ilustrado na figura 3.6. O espectro desse sinal está representado na figura 3.7.

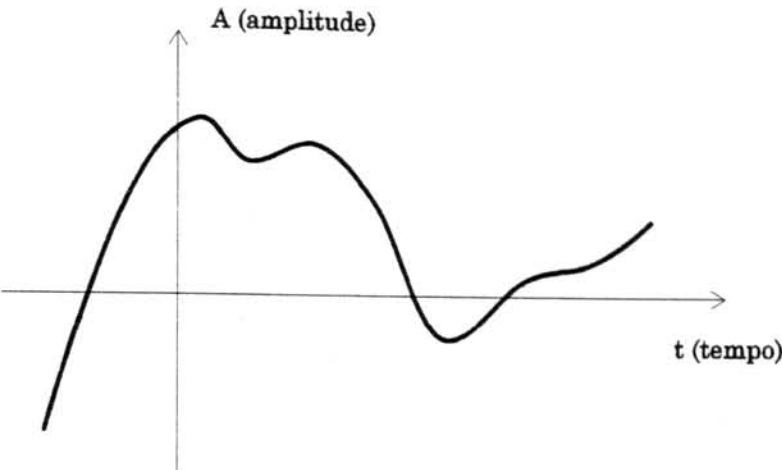


Figura 3.6 : Sinal analógico.

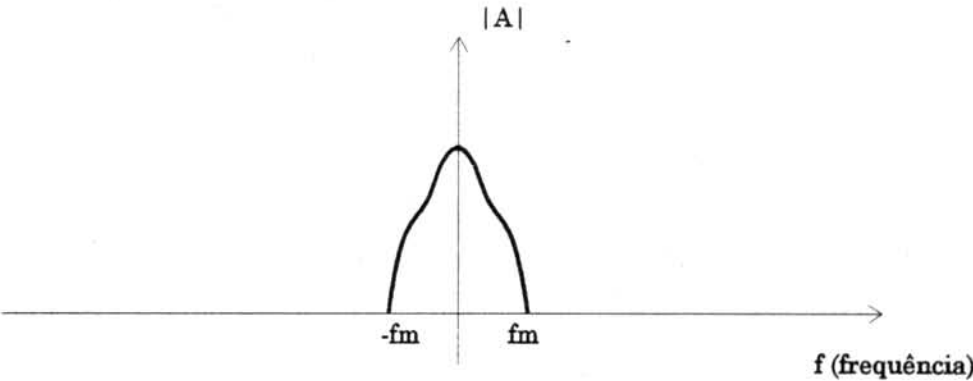


Figura 3.7 : Espectro de frequência dum sinal analógico.

Se amostrarmos esse sinal a uma frequência f_s , tal que f_s seja maior que duas vezes a frequência máxima do sinal analógico, f_m , o espectro do sinal amostrado terá o aspecto ilustrado na figura 3.8.

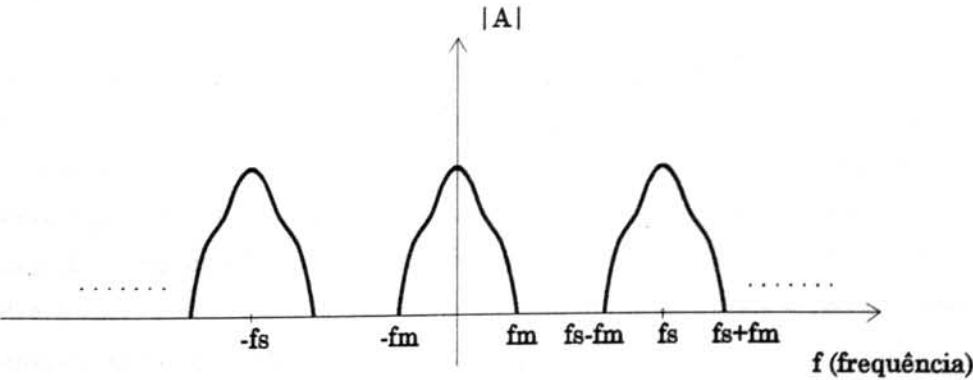


Figura 3.8 : Espectro de frequência dum sinal analógico amostrado a $f_s > 2f_m$.

Consideremos agora o caso de amostrarmos o sinal analógico a uma frequência f_s tal que $f_s < 2f_m$. O espectro do sinal amostrado terá a forma representada na figura 3.9.

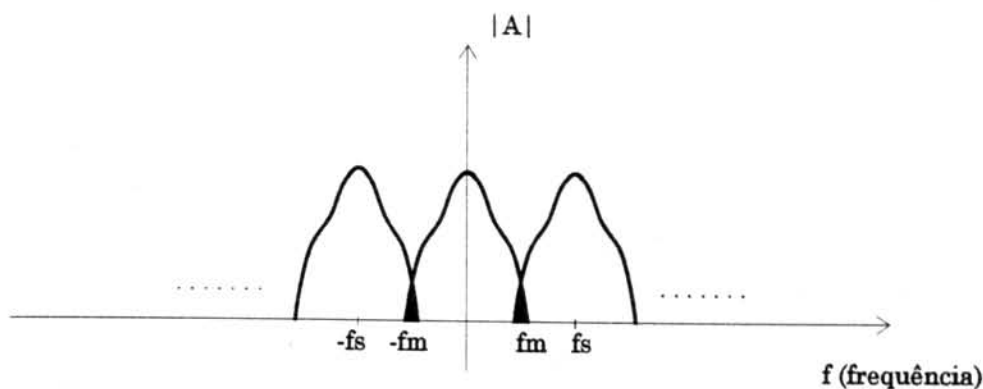


Figura 3.9 : Espectro de frequência dum sinal analógico amostrado a $f_s < 2f_m$.

As bandas de frequência sobrepõem-se o que causa interferências no sinal de saída. Este efeito é conhecido por *aliasing*. Na prática o *aliasing* impede a reconstrução correcta do sinal de entrada a partir das suas amostras.

Para evitarmos o aparecimento de *aliasing* o sinal de entrada deve ser passado por um filtro passa-baixo, denominado filtro *anti-aliasing*, com frequência de corte igual ou inferior a $f_s/2$.

Se usarmos uma frequência de amostragem muito superior à máxima frequência do sinal amostrado ($f_s \gg 2f_m$), mais correcta será a representação do sinal de entrada. Não nos devemos no entanto esquecer que quanto maior o número de amostras, maior será o espaço necessário para a sua posterior retenção em memória.

Um sinal de voz contém componentes frequenciais que se estendem acima dos 5 kHz. O espectro de frequência típico dum sinal de voz está ilustrado na figura 3.10.

Todavia, e como observado anteriormente, a informação mais relevante encontra-se abaixo dos 3 kHz. De modo a que essa informação não se perca, assume-se para frequência de corte do filtro passa-baixo (*anti-aliasing*) os 3,4 kHz para uma frequência de amostragem de 8 kHz. Esta frequência de corte garante uma boa qualidade da voz digitalizada e satisfaz o Teorema da Amostragem de Nyquist para a frequência de 8 kHz. O filtro passa-baixo evita a sobreposição do espectro e conseqüente aparecimento de *aliasing*, conforme ilustrado na figura 3.11.

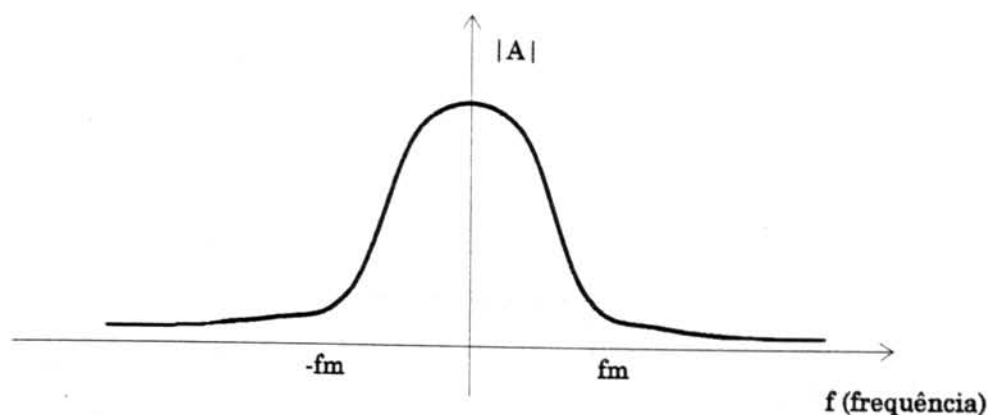


Figura 3.10 : Espectro de frequência típico dum sinal de voz.

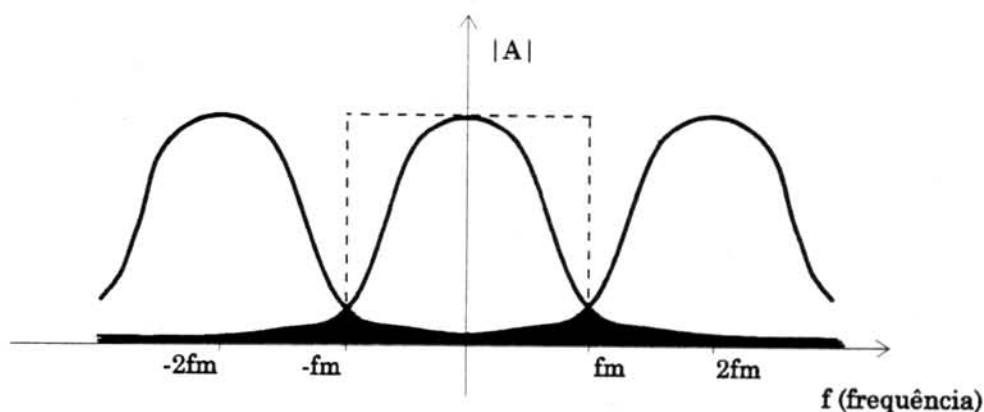


Figura 3.11 : Aplicação de um filtro *anti-aliasing*.

A característica do filtro *anti-aliasing* ideal encontra-se representada a tracejado na figura 3.11. O filtro apresenta atenuação nula em toda a banda passante e atenuação infinita na banda de corte. O filtro ideal apresentaria igualmente uma resposta em fase linear. Uma resposta em fase pobre conduz a distorção harmónica.

Num sistema ideal a forma de onda do sinal amostrador consiste num trem de impulsos igualmente espaçados por um período t_s , conforme se ilustra na figura 3.12.

Se então multiplicarmos estes impulsos pelo nosso sinal analógico, $f(t)$, ilustrado na figura 3.13, obtemos um trem de impulsos cuja amplitude é igual à amplitude de $f(t)$ num determinado instante de tempo (aquele em que se realiza a amostragem).

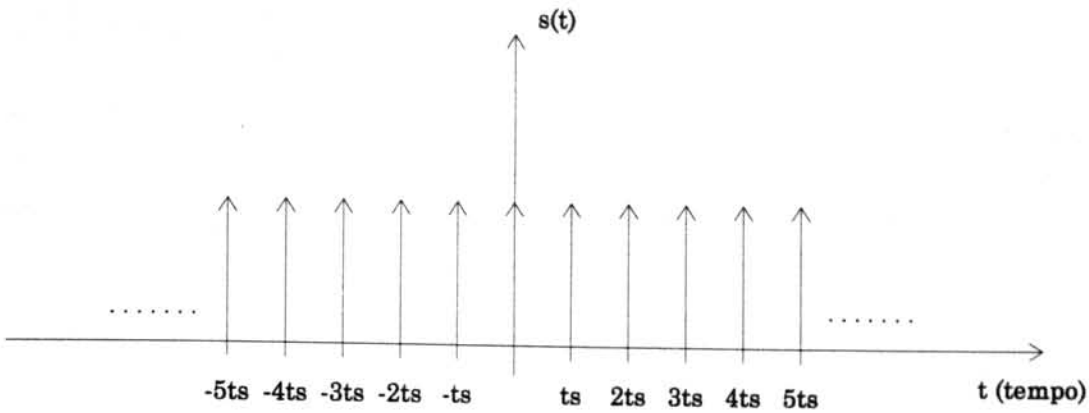


Figura 3.12 : Função de amostragem.

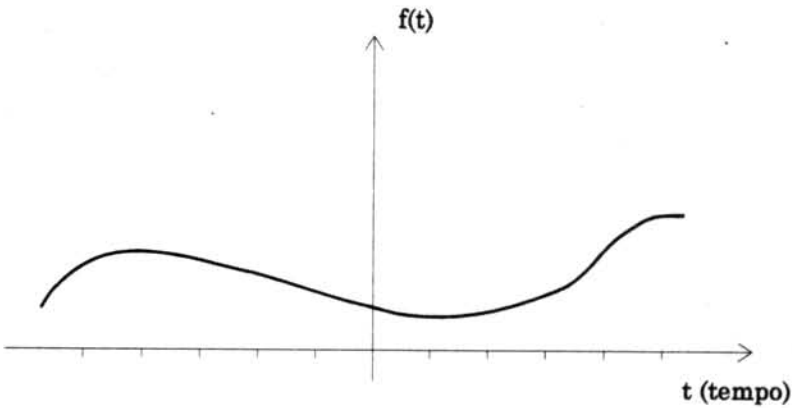


Figura 3.13 : Sinal analógico de entrada.

Obtemos assim a onda amostrada ilustrada na figura 3.14.

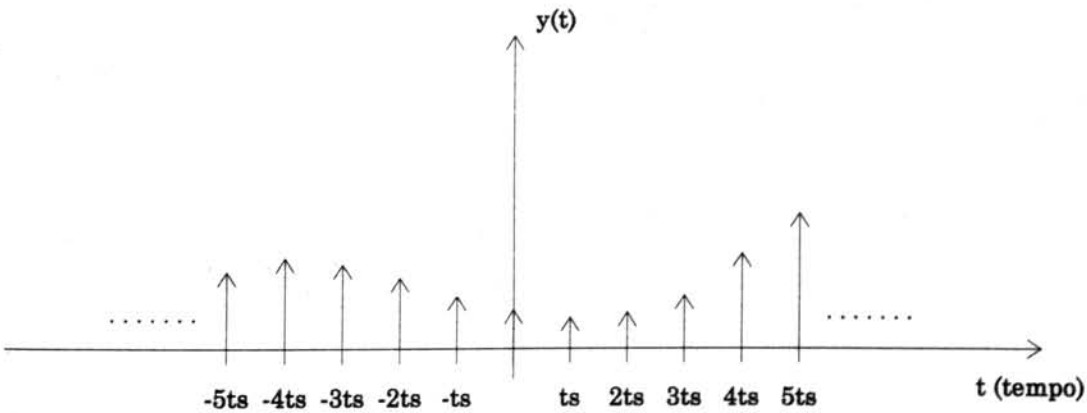


Figura 3.14 : Versão amostrada do sinal de entrada.

3.3.2.2 Quantificação das amostras

Na realidade, a função de amostragem representada na figura 3.14 é substituída por um circuito de amostragem-retenção, o qual mantém o nível amostrado até que a próxima amostra seja tomada, resultando numa forma de onda em escada conforme representado na figura 3.15.

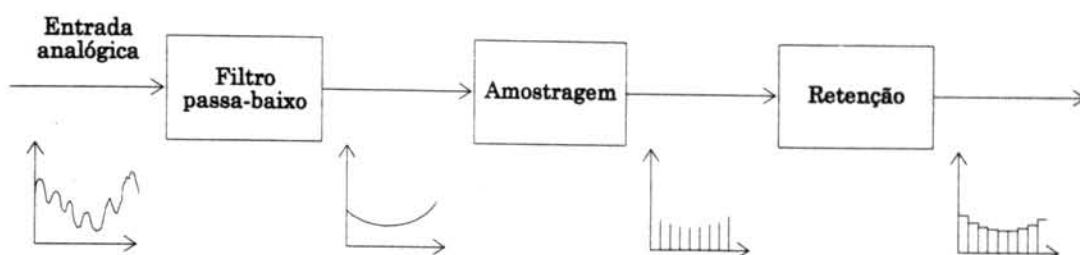


Figura 3.15 : Funções de *anti-aliasing*, amostragem e retenção.

O passo seguinte é o da especificação dum método para representar os valores amostrados da nossa onda analógica. Esta função é denominada quantificação e é realizada pelos conversores analógico-digitais.

A quantificação pode ser vista como a classificação do sinal dentro de certas bandas. Por exemplo, usando dois *bits* para definir quatro níveis igualmente espaçados (00, 01, 10, 11), conforme se ilustra na figura 3.16, e classificando o sinal dentro desses quatro grupos. A forma de onda obtida após essa quantificação está ilustrada na figura 3.17. Obviamente, quantos mais níveis forem usados, maior será a precisão com que o sinal analógico é descrito. Em geral na digitalização da voz são usados 256 níveis de quantificação (8 *bits*).

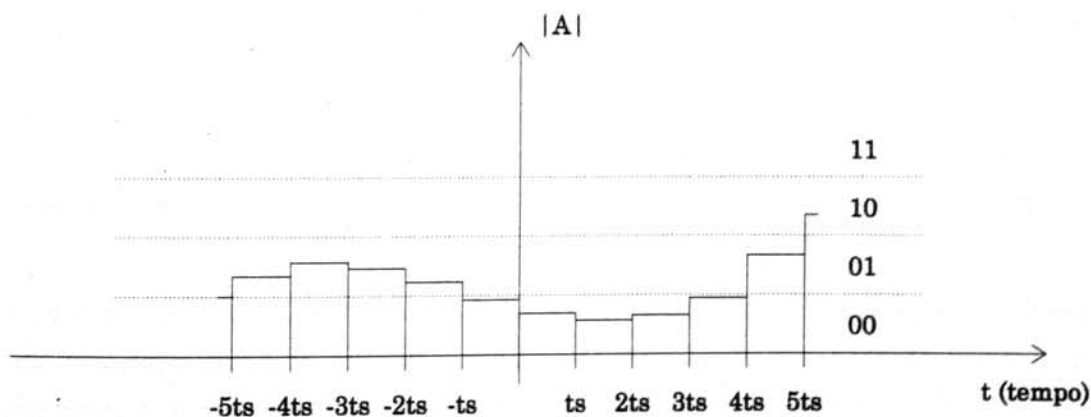


Figura 3.16 : Definição de 4 níveis de quantificação.

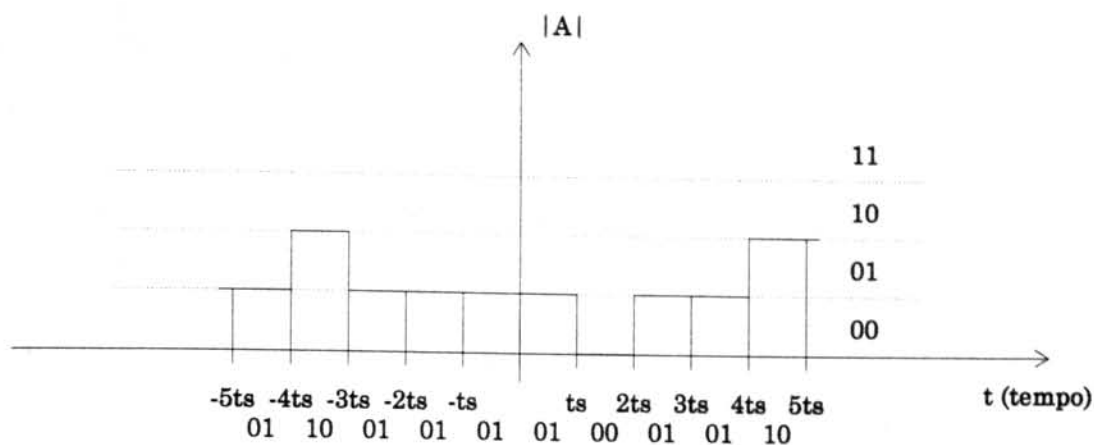


Figura 3.17 : Forma de onda obtida após a quantificação em 4 níveis.

Comparando o sinal quantificado com o sinal original, determina-se qual o erro introduzido. Esse erro produz um efeito designado por ruído de quantificação. Em aplicações de voz esse erro traduz-se pelo aparecimento de ruído na saída. Na maior parte dos casos esse ruído é negligenciável quando comparado com outras fontes de ruído no sistema.

No exemplo apresentado na figura 3.16, os níveis de quantificação definidos encontram-se igualmente espaçados. A figura 3.18 ilustra a função de transferência dum conversor analógico-digital com quantificação linear ou uniforme.

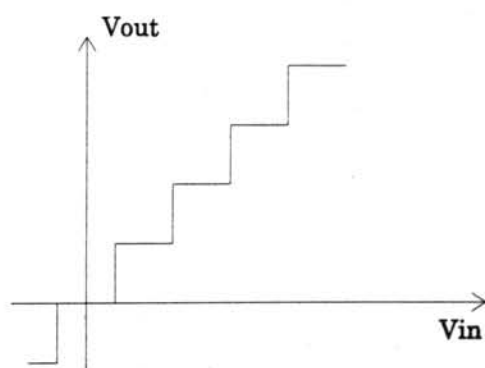


Figura 3.18 : Função de transferência de um conversor analógico-digital com quantificação linear ou uniforme.

Nos sinais de voz existem tipicamente sons altos, por exemplo as vogais, e sons muito mais suaves e baixos, as consoantes. Uma quantificação uniforme, conforme ilustrado na figura 3.19, asseguraria uma representação adequada dos sons mais altos, mas aos mais baixos seria dado o mesmo valor binário [ME93].

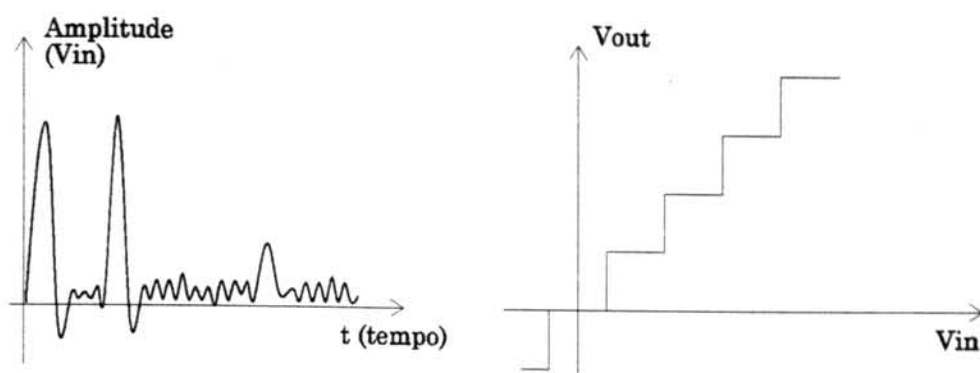


Figura 3.19 : Sinal de voz típico e esquema de quantificação uniforme.

Isto significa que não seria possível uma distinção adequada entre os sinais baixos, o que afectaria grandemente a qualidade da voz digitalizada. Para contornar este problema adoptam-se normalmente esquemas de quantificação não-uniformes, cujos níveis têm espaçamentos variáveis. No caso dos nossos sinais de voz devemos assegurar um maior número de níveis de quantificação nas baixas amplitudes, conforme ilustrado na figura 3.20.

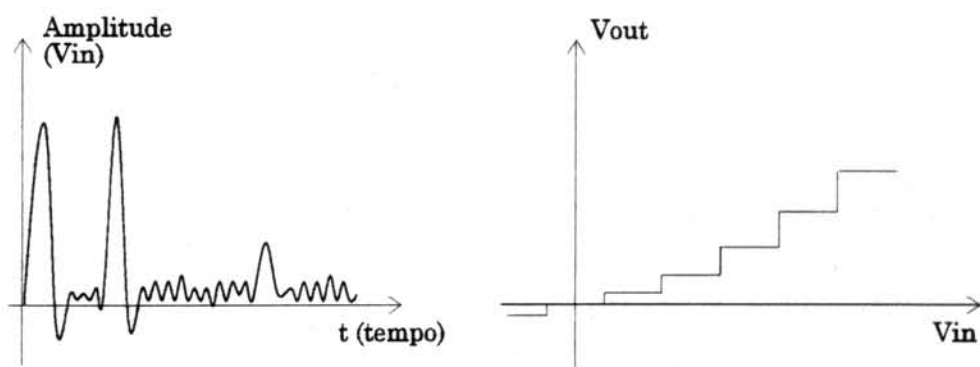


Figura 3.20 : Sinal de voz típico e esquema de quantificação não-uniforme.

Na prática o quantificador tem um passo uniforme, sendo o sinal de entrada sujeito a compressão. No caso dos sinais de voz, os sons de elevada amplitude são comprimidos, enquanto os de baixa amplitude se mantêm quasi inalteráveis. O efeito final é idêntico a uma quantificação não-uniforme. Na reconstrução do sinal aplica-se a mesma técnica mas de forma inversa, isto é, o sinal é então expandido. Os dois métodos mais divulgados de expansão e compressão são a μ -law, usada nos Estados Unidos, e a A-law, usada na Europa [ME93]. A função de transferência da μ -law encontra-se representada na figura 3.21. A função tem um andamento simétrico pelo que os sinais negativos e positivos são tratados de igual forma.

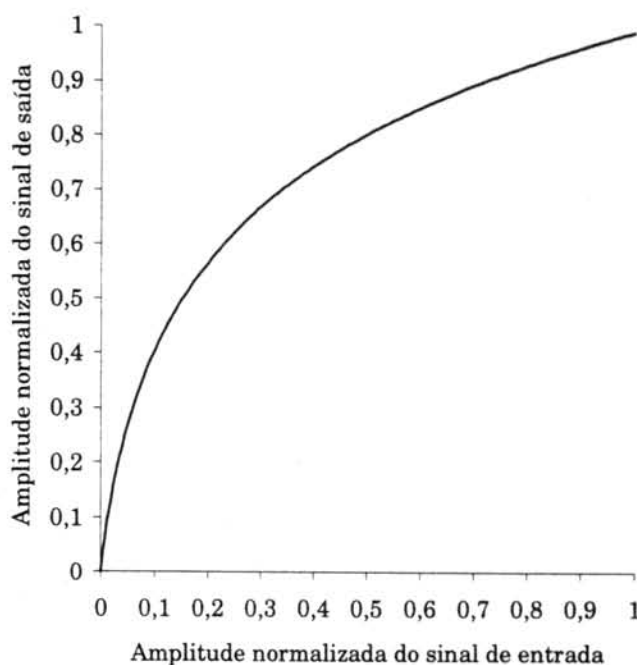


Figura 3.21 : Função de transferência da μ -law.

Qualquer lei de compressão deve proporcionar uma relação linear ou quasi-linear para pequenos sinais. Esses sinais são não só os mais comuns, mas também os mais importantes para a inteligibilidade da voz [Owe82]. A naturalidade da voz pode ser associada com as amplitudes maiores e a curva de compressão escolhida deve ter isso em consideração. Leis de compressão logarítmica como a μ -law, adaptam-se perfeitamente a esse critério. Ao ter um andamento linear para pequenos sinais a A -law satisfá-lo exactamente. A figura 3.22 ilustra a compressão produzida por esta lei.

De entre os vários tipos de conversores analógico-digitais disponíveis para realizar a quantificação do sinal analógico, impõe-se escolher o que melhor se adapta ao caso da digitalização da fala. Dada a relativamente baixa frequência que, como vimos anteriormente, é usada para a amostragem dos sinais de voz, 8 kHz, o nosso conversor não necessita de ser muito rápido. É lógico pois que a escolha recaia sobre os conversores analógico-digitais por aproximações sucessivas. Estes conversores produzem uma saída de n bits em n ciclos de relógio, por comparação da onda de entrada com a saída dum conversor digital-analógico, conforme se ilustra na figura 3.23.

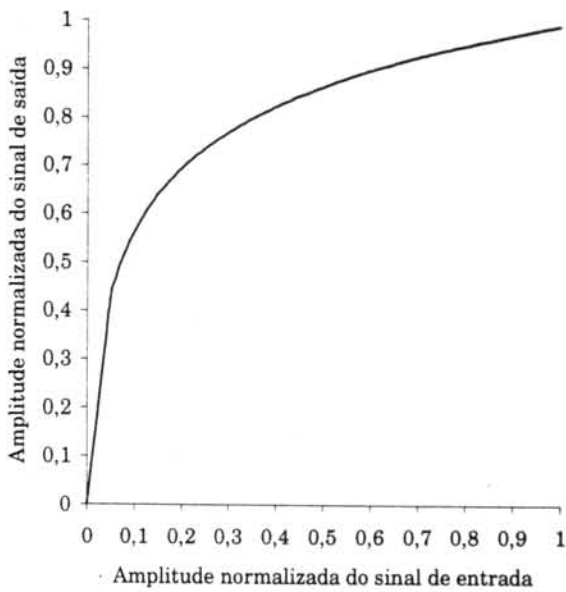


Figura 3.22 : Função de transferência da A-law.

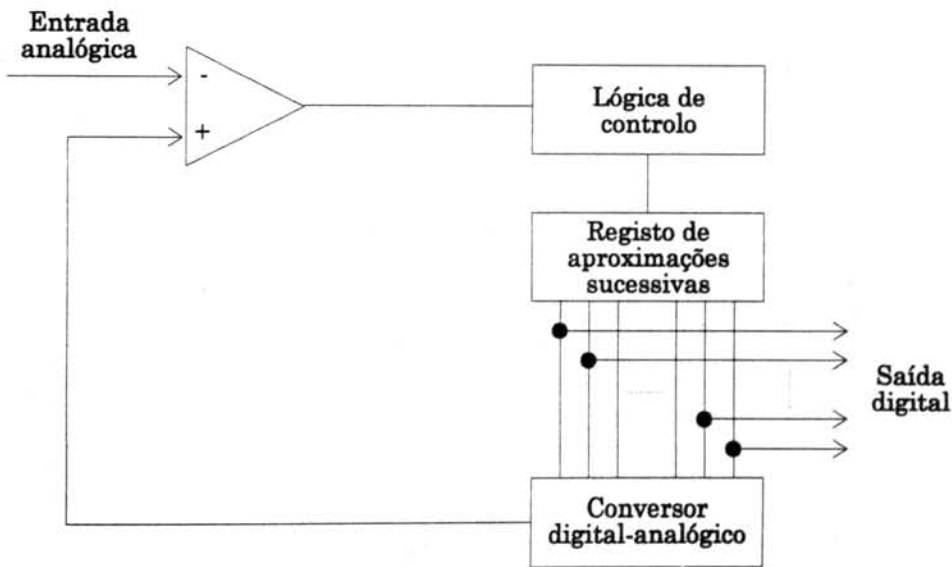


Figura 3.23 : Conversor analógico-digital por aproximações sucessivas.

A técnica usada por este conversor é a de ir dividindo sucessivamente a meio a gama de tensões, e determinar a qual dos intervalos pertence o sinal de entrada. Estes conversores são relativamente baratos e em geral de boa precisão. Os tempos de conversão variam entre $<1\ \mu\text{s}$ e $50\ \mu\text{s}$, com resoluções entre 8 e 12 bits[ME93].

A diferença entre níveis adjacentes de quantificação (largura do degrau) nos esquemas

uniformes é constante. Mesmo nos casos em que se empregam técnicas de compressão e expansão (μ -law e A-law), estas são aplicadas sobre os sinais anteriormente à sua quantificação, e não propriamente sobre a função de transferência do quantificador (o conversor analógico-digital).

3.3.2.3 Codificação de sinal

A amostragem e quantificação dum sinal analógico tem como resultado final um trem de impulsos cujo valor, alto ou baixo, representa os valores binários (0's ou 1's) resultantes da quantificação. Na figura 3.24 é apresentado o exemplo dum sinal analógico quantificado em 8 níveis (3 bits) e o respectivo trem de impulsos, $y(t)$, gerado a partir dessa quantificação. Este método de codificação do sinal é denominado *Pulse Code Modulation* (PCM).

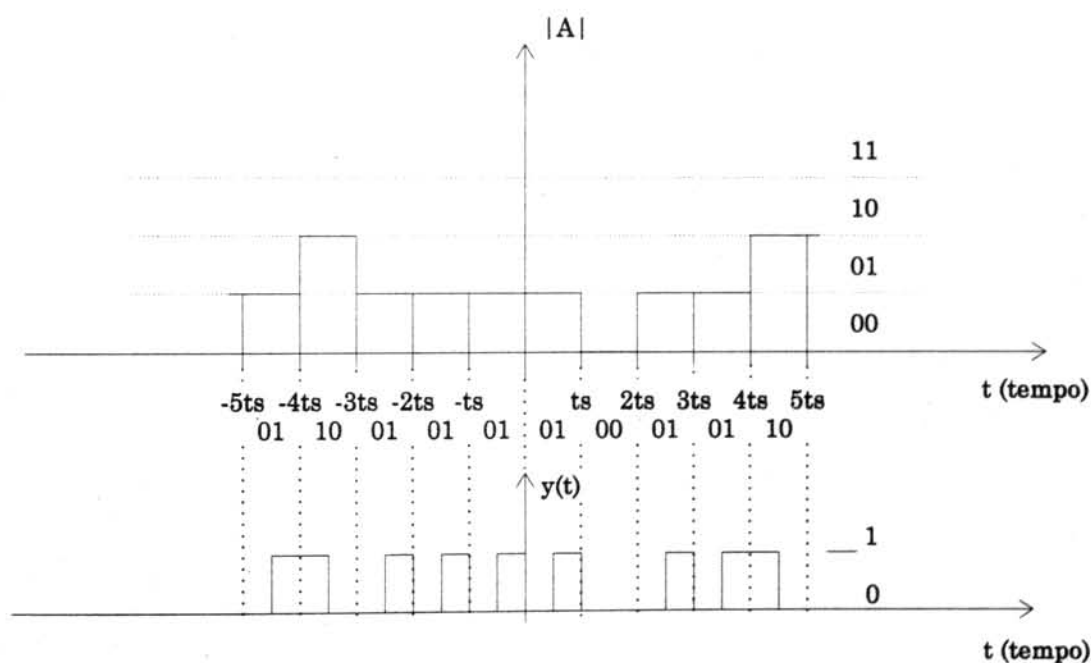


Figura 3.24 : Técnica de codificação PCM.

A codificação de sinais de voz em PCM, com frequência de amostragem 8 kHz e 256 níveis de quantificação (8 bits por amostra), empregando um dos dois algoritmos de compressão/expansão, μ -law ou A-law, encontra-se normalizado pela União Internacional de Telecomunicações. Os esquemas de digitalização de voz seguem esta norma. Todavia, embora sendo o método mais simples de codificação digital, é também o mais dispendioso em termos de informação produzida, uma vez que não possui nenhum mecanismo que

explore a redundância do sinal. De facto, os sinais sonoros da fala contêm um elevado grau de redundância, pelo que são relativamente lentos, embora bastante ricos. O limite superior da velocidade da fala é imposto não pela necessária inteligibilidade, mas por toda a informação extra associada, como por exemplo, a prosódia e a entoação [Edw91].

Um tipo diferente de codificação, denominado *Delta Modulation* (DM), permite remover alguma da redundância presente na codificação PCM, explorando a lentidão da variação do sinal. Se o sinal de entrada é de maior amplitude que o valor digital acumulado num ponto de amostragem, então o valor digital é incrementado de 1. De forma idêntica, se o sinal de entrada é de menor amplitude, o valor digital é decrementado de 1. A figura 3.25 ilustra esta técnica de codificação. A palavra de saída, $y(t)$, para cada período de amostragem do sinal de entrada, $x(t)$, é só um zero ou um um, consoante o valor acumulado seja decrementado ou incrementado.

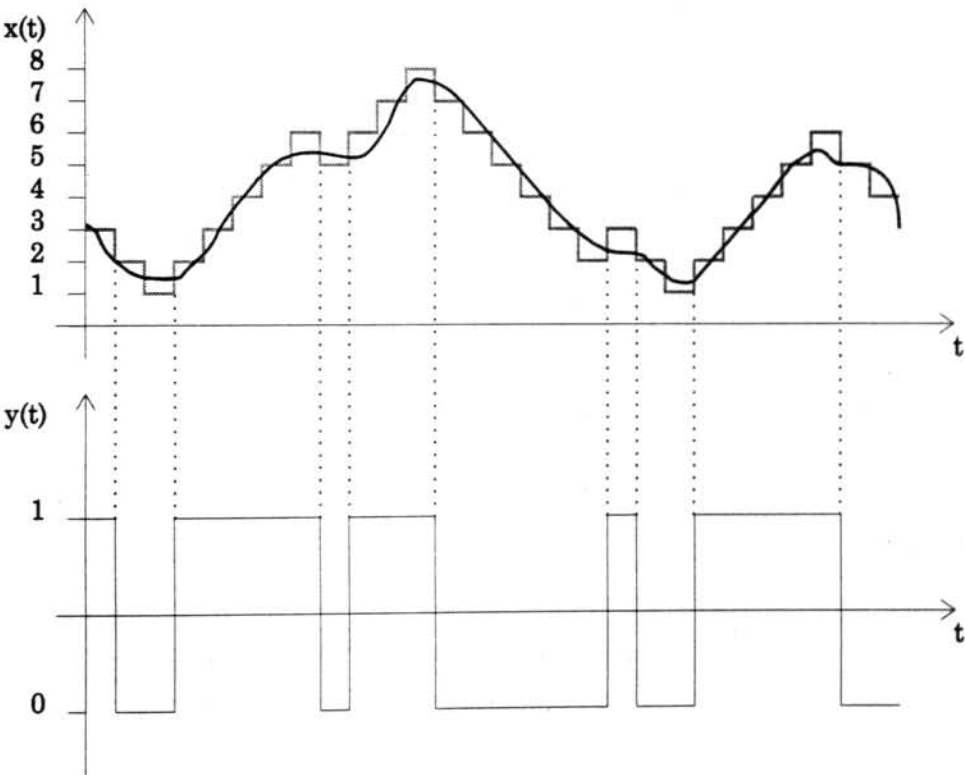


Figura 3.25 : Aplicação da técnica de codificação *Delta Modulation* a um sinal analógico amostrado.

O resultado final é a diminuição em $1/n$ do número de *bits* requeridos em relação à codificação PCM. Por exemplo, uma quantificação de 256 níveis implica a geração duma palavra de 8 *bits* por cada amostra em codificação PCM, enquanto que em DM teremos

apenas 1 *bit* por amostra, diminuindo em 1/8 a informação que será necessário armazenar para guardar o sinal digitalizado. A figura 3.26 ilustra o diagrama de blocos do modulador e do desmodulador *Delta*.

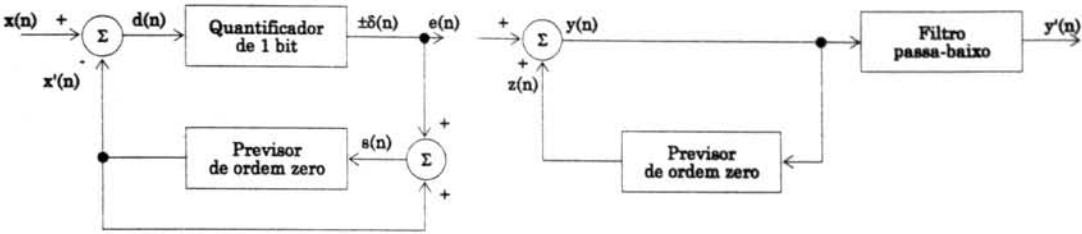


Figura 3.26 : Modulador e desmodulador *Delta*.

Antes de ser aplicado ao modulador o sinal analógico é passado por um circuito de amostragem e retenção que produz uma forma de onda discreta $x(n)$, conforme ilustrado na figura 3.27. Assumindo a não existência de um sinal anterior a $t=0$, $x(n)$ é quantificado em '1' ou '0'. O elemento de realimentação produz uma previsão do próximo valor de entrada esperado. Num previsor de ordem zero a previsão da amostra actual é realizada com base na amostra anterior [Cap85].

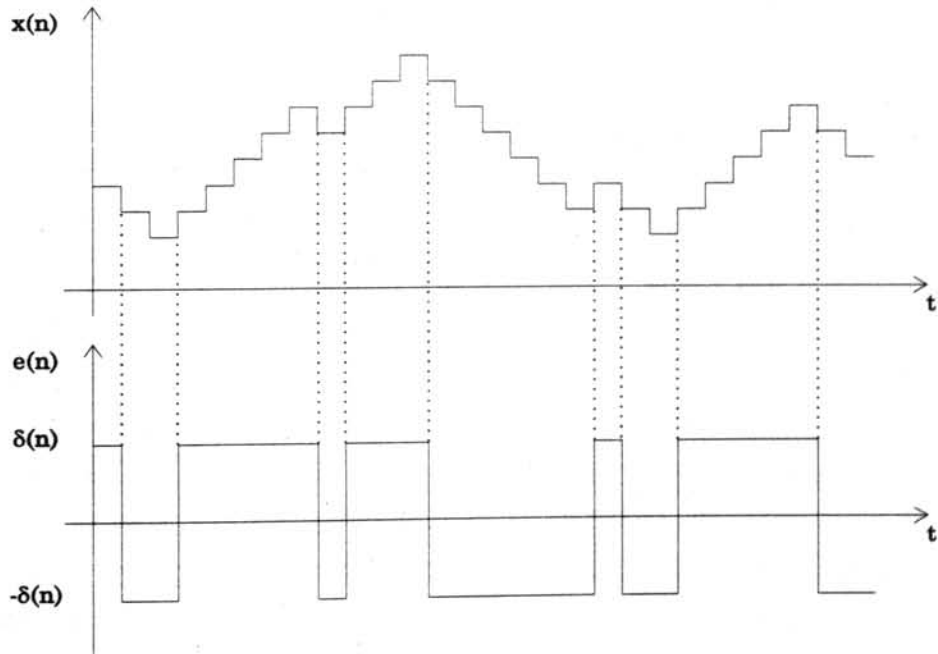


Figura 3.27 : Saída digitalizada de um circuito de amostragem-retenção e sua codificação usando a *Delta Modulation*.

Se assumirmos na realimentação que $x'(n)=s(n)$, então o sinal de saída $e(n)$ será o representado na figura 3.27, assumindo $\delta(n)=\pm 1$. O mesmo previsor é usado no

desmodulador para reconstruir o sinal.

Codificadores DM com largura de degrau constante (referidos vulgarmente como LDM - *Linear* DM) apresentam no entanto duas fontes de erro que podem afectar gravemente a codificação — sobrecarga na inclinação e granularidade. Se a largura do degrau, $\delta(n)$, é demasiado pequena, os incrementos no sinal à saída do desmodulador não permitem o acompanhamento da rápida variação que se verifica no sinal analógico original. Neste caso teremos uma sobrecarga na inclinação, como ilustrado na figura 3.28.

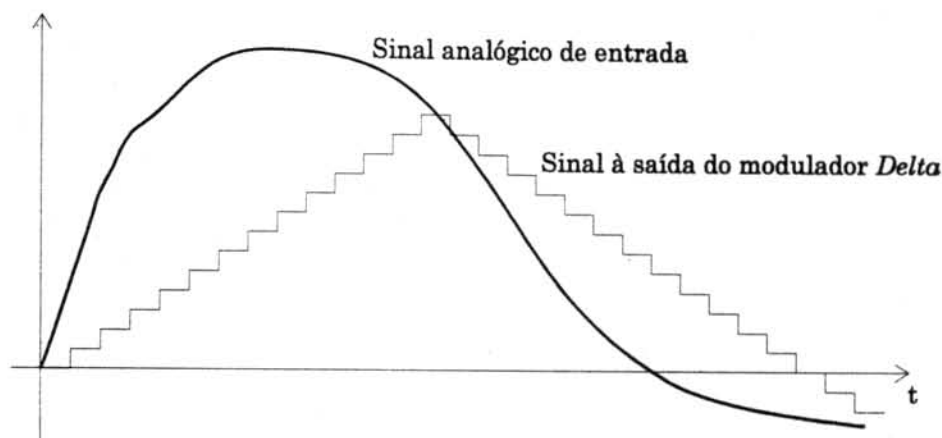


Figura 3.28 : *Delta Modulation* com sobrecarga na inclinação.

Se, em oposição, $\delta(n)$ for demasiado grande, conduz a que uma pequena variação do sinal analógico de entrada produza variações demasiado elevadas no sinal à saída do desmodulador, originando granularidade, conforme se ilustra na figura 3.29.

Para evitar estes dois problemas é necessário amostrar o sinal a uma maior frequência que no caso da codificação PCM, pelo que muita da vantagem deste tipo de codificação, a diminuição do número de *bits* necessários para codificar um sinal analógico, se perde. Consequentemente é necessário um elevado cuidado quando se escolhe um valor para δ . Se se conhecer as características da forma de onda de entrada, nomeadamente a sua rapidez de variação, isso não será problema visto se poder escolher um valor de δ que evite quer o aparecimento de sobrecarga na inclinação, quer de granularidade. Caso contrário dever-se-à optar por outra solução. Um exemplo onde a aplicação de codificação DM é inapropriada é a voz humana, devido à grande variação de amplitude existente entre vogais e consoantes. As primeiras produzem tipicamente sons de amplitude muito mais elevada que as segundas. Nestes casos devem ser usados previsores adaptativos que fazem depender a largura do degrau dos valores dum conjunto anterior de amostras.

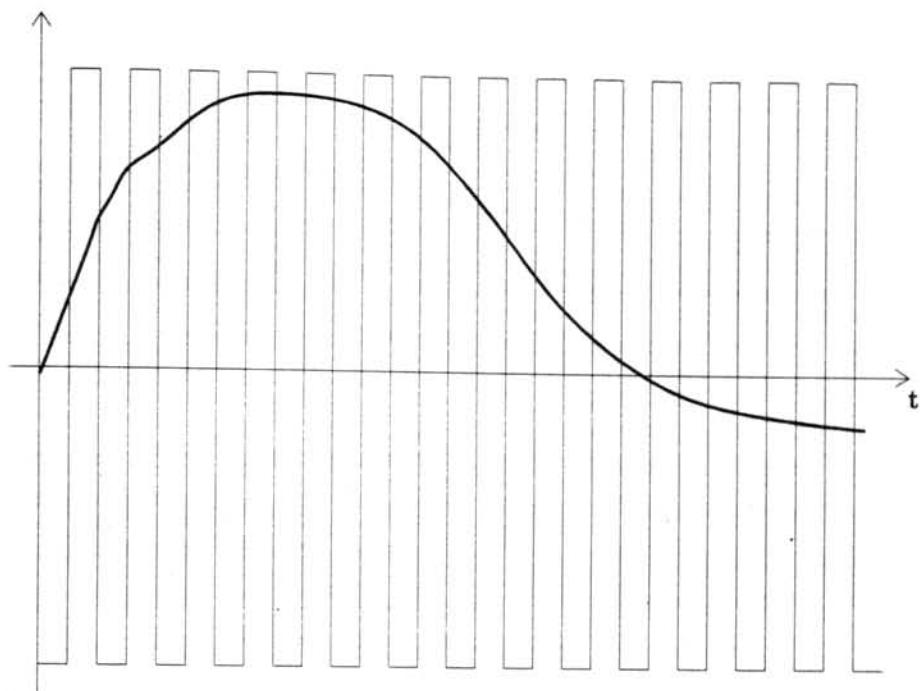


Figura 3.29 : Delta Modulation com granularidade.

O PCM diferencial (DPCM — *Differential PCM*) é uma extensão da codificação DM, que utiliza igualmente a redundância presente nos sinais analógicos. Na codificação diferencial, a diferença entre o sinal discreto de entrada e o valor previsto é quantificada em um de vários níveis. A complexidade do sistema está directamente relacionada com a complexidade do previsor. Um diagrama ilustrativo está representado na figura 3.30.

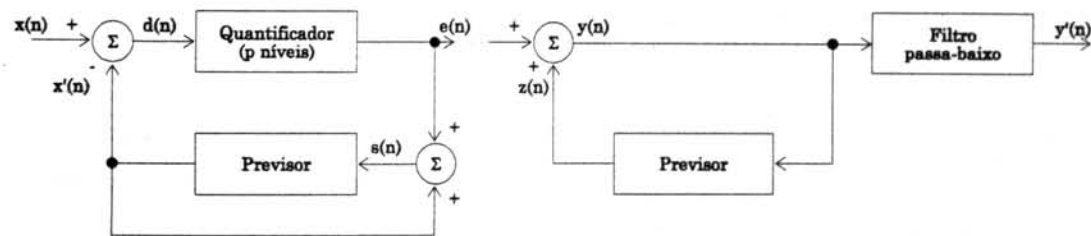


Figura 3.30 : Modulador e desmodulador DPCM.

Repare-se que, enquanto na figura 3.26 tínhamos um quantificador com apenas dois níveis de quantificação (1 bit), aqui temos um quantificador com p níveis de quantificação.

O previsor DPCM pode ser geralmente expresso pela equação:

$$x'(n) = \sum_{R=1}^N [h_R s(n-R)]$$

onde h_R é o conjunto de coeficientes do previsor e N o número de amostras anteriores que se toma em consideração para cálculo do valor do previsor. De notar que se $h_1=1$ e $N=1$, isto é, se se tomar em conta apenas o valor da última amostra, sem alteração, temos a codificação DM.

A vantagem do uso do DPCM sobre o PCM é o ganho na relação sinal-ruído, melhorando a qualidade do sinal de saída. Como no DPCM apenas é quantificada a diferença entre sinais e não o seu valor absoluto, incorre-se num muito menor erro de quantificação, em $e(n)$, apesar de usarmos o mesmo número de *bits* que nos sistemas PCM. O DPCM é muito mais preciso para o mesmo número de *bits* por amostra.

A qualidade dos algoritmos de codificação está relacionada com o conhecimento que se tiver do sinal. Se o sinal estiver claramente definido, podem usar-se previsores fixos (invariantes no tempo). Todavia, na maior parte dos casos, os sinais afastam-se significativamente do previsto por curtos períodos de tempo. Nestes casos é de toda a vantagem o uso de esquemas adaptativos de codificação. O termo DPCM adaptativo (ADPCM — *Adaptive DPCM*) é usado geralmente para titular dois esquemas diferentes — adaptação do quantificador e adaptação do previsor. A figura 3.31 exemplifica o uso de um quantificador adaptativo dependente da gama de amplitudes do sinal de entrada.

A estimação da largura do degrau pode ser realizada adiantada ou atrasada. Na estimação adiantada, as amostras são armazenadas inicialmente e usadas para estimar o nível da entrada antes da quantificação, reduzindo os efeitos do ruído de quantificação. A estimação atrasada evita a necessidade de armazenamento de um determinado número de amostras à entrada, sendo por isso a mais usada [ME93]. A figura 3.32 esquematiza a constituição do modulador e do desmodulador para estes dois casos.

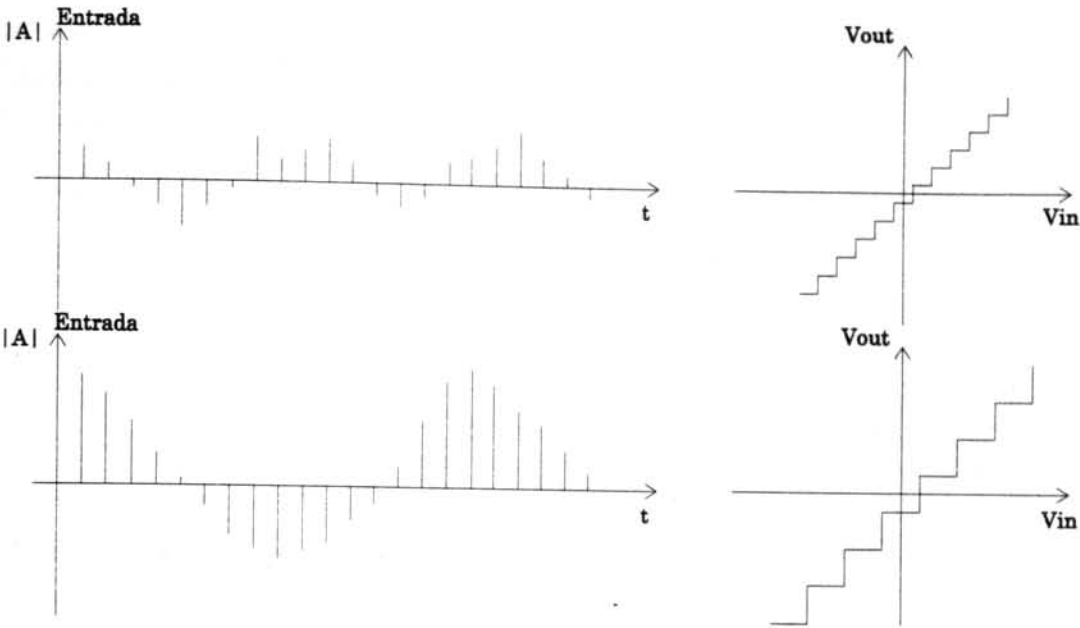


Figura 3.31 : Exemplo de quantificação adaptativa dependente da amplitude de entrada do sinal.

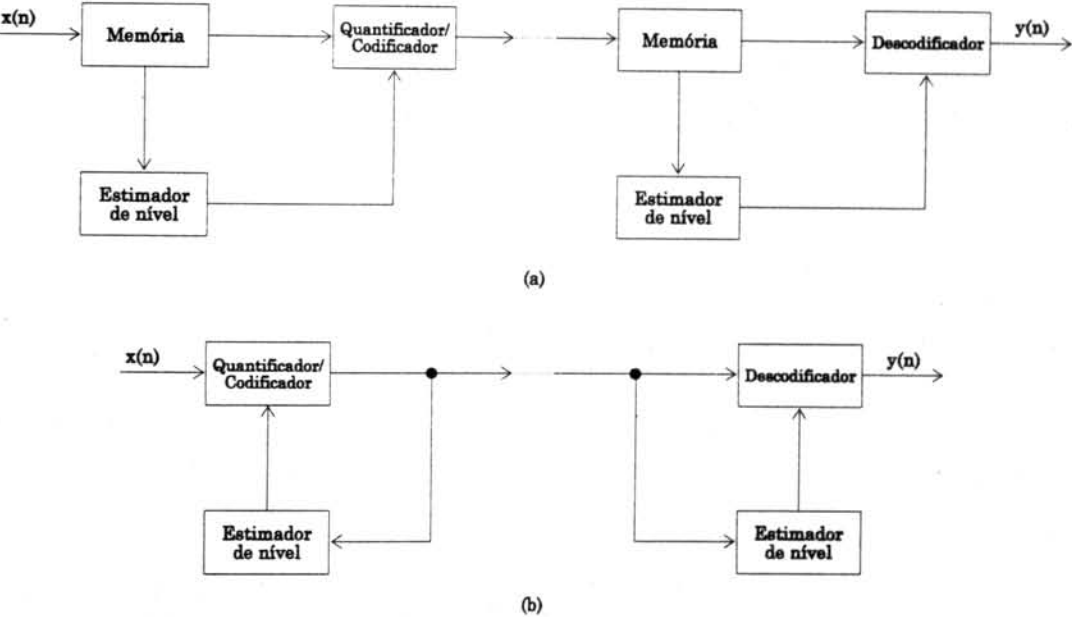


Figura 3.32 : ADPCM: Modulador e desmodulador para a estimação adiantada (a) e atrasada (b).

Os quantificadores adaptativos oferecem uma apreciável melhoria na relação sinal-ruído em comparação com os esquemas não-adaptativos, inclusive com algoritmos de codificação de voz de 3 bits por amostra. Normalmente, os sistemas de codificação de voz combinam a quantificação adaptativa com previsores adaptativos, o que se traduz numa acentuada diminuição do erro de quantificação. A norma CCITT G.721 [CCI84] emprega esquemas

de adaptação do quantificador e do predictor baseados em estimação atrasada. A figura 3.33 ilustra o diagrama de blocos deste esquema para o modulador e para o desmodulador.

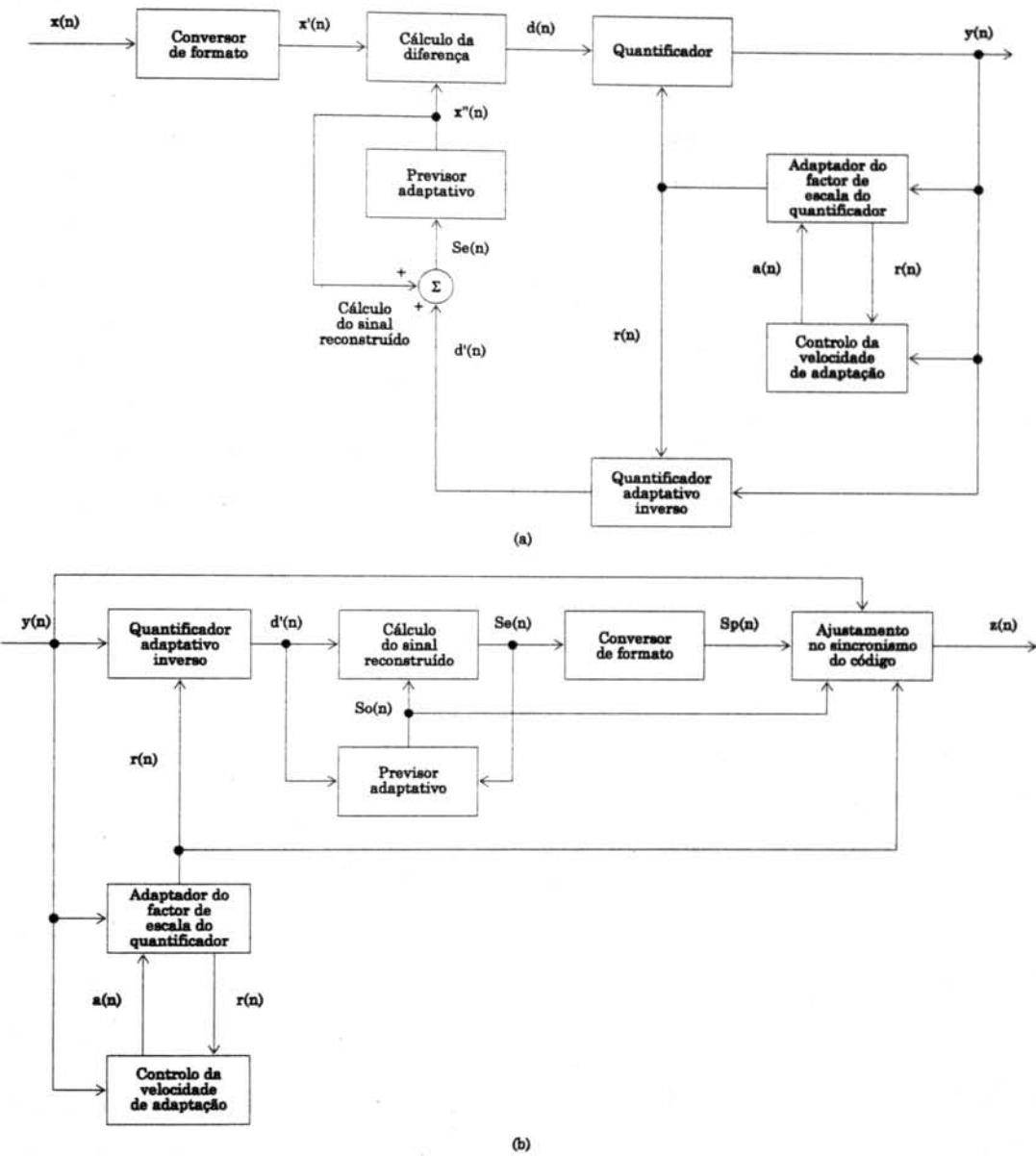


Figura 3.33 : Esquemas do modulador (a) e desmodulador (b) definidos pela norma CCITT G.721.

O sinal à entrada do modulador, $x(n)$, segue a norma CCITT de codificação de voz em PCM a 64 kbit/s. A norma usa um algoritmo de expansão/compressão, o μ -law para os Estados Unidos e a A-law para a Europa. O modulador converte a entrada PCM em ADPCM a 32 kbit/s. O conversor de formato do modulador transforma o sinal PCM

resultante da quantificação do sinal comprimido e presente na entrada do desmodulador, em PCM linear, eliminando essa compressão. O conversor de formato do desmodulador realiza a função inversa. O quantificador adaptativo usa 4 *bits* para quantificar a diferença, sendo que o primeiro é de sinal e os três restantes correspondentes a 8 níveis diferentes de quantificação [CCI84].

A transformação de PCM a 64 *kbit/s* em ADPCM a 32 *kbit/s* reduziu para metade a área necessária ao armazenamento da voz digitalizada. Posteriormente, a norma CCITT G.726 [CCI90] que permite a passagem de PCM a 64 *kbit/s* para ADPCM a 24 e 16 *kbit/s*, veio possibilitar uma maior diminuição do espaço de armazenamento necessário. Esta contribuição da codificação de sinal é de primordial importância para a melhoria dos sistemas de CAA com saída de voz, permitindo o alargamento do vocabulário disponível, mantendo a elevada qualidade da voz reproduzida por estes sistemas.

3.3.2.4 Medidas da qualidade da fala

A medida da qualidade da fala é um problema de difícil solução e que se arrasta já há alguns anos [Jay90]. No entanto, torna-se necessária a criação de uma escala que nos permita comparar os vários métodos de codificação em termos da qualidade final da voz, para que se possa aferir do seu interesse prático.

A medida da relação sinal-ruído não levanta grande polémica por ser uma medida objectiva. Todavia, o seu emprego na medida da qualidade da voz não se demonstra apropriado, uma vez que não tem em conta as propriedades perceptivas do ouvido humano. Por isso, o recurso que se tem verificado às avaliações subjectivas baseadas em gravações da fala, visto a concepção dos algoritmos de codificação se fundamentar, na sua maior parte, em critérios perceptuais [Spa94].

Procedimentos de teste subjectivos como o *Diagnostic Rhyme Test* (DRT), o *Diagnostic Acceptability Measure* (DAM) e o *Mean Opinion Score* (MOS), são baseados em médias de avaliação de um conjunto de ouvintes. O DRT é uma medida da inteligibilidade onde a tarefa do ouvinte é reconhecer uma de duas palavras possíveis num conjunto de pares de palavras rimadas (por ex.: cascar - mascar). O DAM é baseado nos resultados de vários métodos de teste que pretendem avaliar a qualidade dum sistema de comunicação com base na aceitabilidade da fala por parte dum ouvinte devidamente treinado segundo um determinado padrão.

O MOS é o método mais usado nas medidas de qualidade da fala codificada. Envolve usualmente 12 a 24 ouvintes (32 a 64 em testes formais organizados pela União Internacional de Telecomunicações), os quais são instruídos para relacionarem entre si várias gravações e classificá-las segundo uma escala de qualidade de 5 níveis, de acordo com a tabela 3.1.

Tabela 3.1 : Escala MOS

Escala MOS	Qualidade da fala
1	Má
2	Pobre
3	Razoável
4	Boa
5	Excelente

Uma qualidade excelente implica que a fala codificada seja indistinguível do original e sem ruído perceptível. Por outro lado, uma qualidade má (inaceitável) implica a presença de ruído incomodativo na fala codificada.

No teste MOS os ouvintes estão “calibrados” no sentido em que estão familiarizados com as condições de audição e com a gama de qualidade da fala que vão encontrar. Os valores finais são obtidos pela média dos valores numéricos atribuídos a algumas centenas de gravações da fala. De notar que as médias MOS podem diferir significativamente de teste para teste, e logo não podem ser consideradas como medidas absolutas para comparação de diferentes códigos.

Apesar das muitas áreas de aplicação em que as técnicas de codificação de voz descritas encontram actualmente lugar, merece ainda referência neste capítulo uma técnica recente de armazenamento directo em analógico dos sinais de voz, que oferece excelentes perspectivas de baixo custo para sistemas de menor complexidade.

3.4 O armazenamento directo em analógico:

A tecnologia DAST¹

Os sinais analógicos, como a voz, têm de ser convertidos para formato digital antes de poderem ser armazenados em memórias de estado sólido. Um sistema completo de gravação requer um conversor analógico-digital, uma memória, um circuito de endereçamento adequado (usualmente um decodificador incorporado no próprio circuito integrado da memória e um contador externo) e um gerador de sinais de relógio. A reprodução do sinal guardado requer ainda o uso de um conversor digital-analógico.

Desde os anos 60 que se tem tentado alterar esta situação. O aparecimento das memórias magnéticas é disso exemplo. Ao permitirem o armazenamento directo de sinais analógicos, evitavam todo o processo de digitalização. No entanto, a sua baixa gama dinâmica, levou ao seu abandono [Goo91]. Recentemente, uma nova tecnologia denominada DAST (*Direct Analog Storage Technology*) veio alterar esta situação. Um novo método baseado na tecnologia das EEPROMs (*Electrically Erasable Programmable Read Only Memory*) permite que valores analógicos sejam directamente “escritos” numa única célula de memória sem necessidade de conversão para digital. Isto resulta num aumento de densidade sobre os métodos digitais equivalentes e o armazenamento não-volátil de dados analógicos [ISD94].

Embora a memorização seja analógica na sua natureza, técnicas de amostragem continuam a ser necessárias, visto não ser possível o armazenamento do sinal de forma contínua no tempo, mas apenas do seu valor em determinados instantes de tempo. Como tal, continua a ser indispensável o uso de filtros *anti-aliasing* para atenuação das componentes de frequência presentes no sinal de entrada que sejam superiores a metade da frequência de amostragem empregue. Este procedimento é necessário por forma a satisfazer o Teorema da Amostragem de Nyquist, já referido na secção anterior, e aplicável a todos os sistemas de amostragem.

Depois de amostrado, o valor do sinal é sujeito a um processo de deslocamento do nível de

¹ DAST é uma marca registada da ISD (Information Storage Devices, Inc.).

tensão por forma a obter-se o valor elevado requerido pela operação de escrita na célula de memória analógica. O relógio empregue na amostragem serve igualmente para incrementar o decodificador da matriz de células por forma a que as amostras do sinal de entrada sejam armazenadas sequencialmente.

Durante a reprodução os sinais analógicos de tensão são lidos sequencialmente à mesma cadência da frequência de amostragem, reconstruindo-se a forma de onda amostrada. Um filtro passa-baixo permitirá a remoção da componente devida à frequência de amostragem, restaurando o sinal inicial. O diagrama de blocos apresentado na figura 3.34 ilustra as várias fases deste processo.

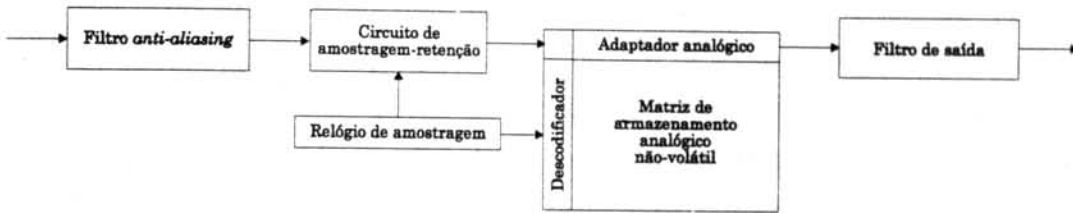


Figura 3.34 : Fases do processo de armazenamento directo em analógico.

Um circuito de amostragem e retenção fornece, à entrada dum comparador, a tensão analógica a ser armazenada. A saída da célula de memória liga à outra entrada. Os electrões são “bombeados” em pequenas parcelas para a célula de memória. O valor de tensão resultante é, em cada instante, aplicado ao comparador. Quando o nível de tensão na célula iguala a tensão presente no circuito de amostragem e retenção, a injeção de electrões na célula cessa. A quantidade mínima de carga injectada na célula de cada vez determina a resolução do sistema.

Presentemente, as memórias analógicas disponíveis no mercado permitem guardar um de 230 níveis de tensão possíveis. Este valor aproxima-se da resolução digital conseguida pelos conversores analógico-digitais de 8 *bits*. Todavia, na memorização digital necessitaríamos de oito células para guardar tal valor, enquanto que usando memórias analógicas armazenamos directamente o sinal em apenas uma célula. Qualquer técnica de codificação digital que vise atenuar este fosso resulta numa pior qualidade de reprodução posterior do sinal.

O funcionamento destas memórias baseia-se nas propriedades de armazenamento analógico das células não-voláteis de porta flutuante das EEPROMs. A figura 3.35 esquematiza a constituição duma destas células.

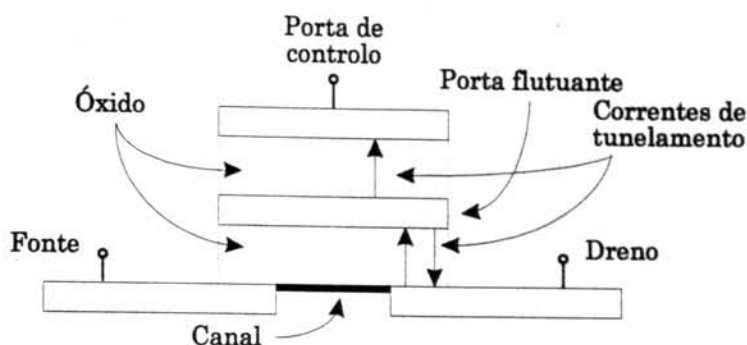


Figura 3.35 : Esquema construtivo duma célula de memória analógica.

Até agora, a informação digital era guardada nestas células, *bit a bit*. Se a porta flutuante era programada com uma forte carga positiva, o canal do transistor tornava-se condutor. Este estado corresponderia, digamos, a um valor lógico 1. Se a célula era apagada, a porta flutuante recebia uma forte carga negativa. Desta forma, a célula tornava-se não condutora, correspondendo ao valor binário complementar, o zero lógico. Programando ou apagando as células numa matriz, criava-se uma memória digital [Goo91].

As memórias analógicas tiram partido dos valores intermédios de condutividade existentes entre os extremos condutor e não-condutor para criar uma célula analógica não-volátil. Uma célula duma EEPROM é um transistor MOS (*Metal-Oxide Semiconductor*) que armazena carga numa placa condutora electricamente isolada — ou porta flutuante. A porta flutuante, localizada acima do canal do transistor, cria, quando carregada, um campo eléctrico que modifica a condutividade desse canal.

A carga na porta, composta de electrões, é aumentada ou decrementada por um mecanismo de condução de corrente denominado tunelamento. A corrente de tunelamento é conduzida através da camada de óxido, o qual é um excelente isolante sob níveis baixos de tensão de polarização, mas que, quando sujeito a tensões de polarização elevadas, torna-se suficientemente condutor permitindo a passagem dos electrões, carregando ou descarregando a porta flutuante. A corrente de tunelamento varia exponencialmente com a tensão de polarização aplicada. A polaridade dessa tensão determina a carga ou descarga da porta. O potencial da porta flutuante e os potenciais através do óxido são manipulados por acoplamento capacitivo electrostático, por meio da porta de controlo e pelos outros terminais do transistor. Quando os valores de polarização são reduzidos, os óxidos voltam a comportar-se como excelentes isolantes. Sob esta condição, a quantidade de carga na porta flutuante fica retida, mantendo-se neste estado por muitos anos, mesmo

se sujeita a temperaturas elevadas [Goo91]. A tensão na porta flutuante pode agora ser “lida” sem perturbar a carga, e como tal, sem alterar a tensão armazenada. O valor da condutividade da célula corresponde ao valor do nível analógico armazenado.

A simplicidade introduzida por este tipo de memória nos sistemas de gravação e reprodução de voz aumenta a sua portabilidade, devido ao menor número de componentes necessários, o que consequentemente conduz a uma redução dos níveis de consumo, permitindo uma maior autonomia. A sua relativamente recente introdução no mercado é factor preponderante na sua ainda pouca divulgação e utilização, apesar das suas aparentes potencialidades.

Ao longo deste capítulo foi abordada a influência da electrónica nos sistemas de CAA e revistas algumas técnicas de codificação de sinal, bem como a possibilidade do uso dum novo tipo de memória que permite guardar directamente valores analógicos. Necessário se torna agora um aprofundar de conhecimentos sobre a paralisia cerebral, de modo que, identificando correctamente as limitações, dificuldades e necessidades dos utilizadores, se torne possível a especificação dum conjunto de ajudas a desenvolver que, baseadas nestas evoluções tecnológicas, permitam melhorar a qualidade de vida destas pessoas.

4. O Projecto “Fala-Barato”: Da Identificação de Requisitos à Especificação Técnica

É desde há muito consensual a ideia de que as novas tecnologias têm tido um impacto tremendo na organização da vida diária de todos nós ao longo das últimas décadas [Sch93]. Muitos dispositivos e sistemas contribuíram para tornar as nossas vidas mais confortáveis quer em casa, no trabalho, quando viajamos ou durante actividades de recreio. A mesma evolução fez-se sentir de igual forma na vida de pessoas com necessidades especiais. Não no sentido de proporcionar um maior luxo, mas na maioria dos casos, para a realização de actividades diárias elementares que anteriormente lhes estavam vedadas, como a comunicação.

No entanto, a medicina de reabilitação apresentou igualmente progressos significativos na melhor compreensão do corpo e da mente humanas. As causas e consequências de desordens nas funções sensoriais, cognitivas e motoras têm sido discutidas em contextos multidisciplinares, contribuindo para a integração de diferentes perspectivas sobre um mesmo tema, conduzindo a intervenções mais específicas e eficientes [Spae94]. A chave para uma intervenção com sucesso recai num correcto entendimento das funcionalidades e possibilidades da pessoa deficiente. Essencial se torna pois um conhecimento mais aprofundado sobre a paralisia cerebral para que se possa identificar os requisitos sobre os quais assentará a especificação técnica dos sistemas de CAA a desenvolver. Basicamente são estes três pontos, a paralisia cerebral, a identificação de requisitos e a especificação técnica, que serão focados ao longo deste capítulo.

4.1 A paralisia cerebral

Foi o inglês *Little* quem em 1862 descreveu, pela primeira vez, a forma espasmódica da paralisia cerebral, à qual atribuiu o nome de “mal de *Little*”. O paralítico cerebral distingue-se dos demais pela sua marcha descoordenada e precária, pelos exagerados movimentos dos braços e mãos, pela sua dificuldade em comer e falar. No entanto, e em geral, o paralítico cerebral é uma pessoa com possibilidades intelectuais e mentais, com ilusões, com vontade de viver, de ser amada, compreendida e atendida como as outras.

O termo paralisia cerebral, ao contrário daquilo que pode sugerir, não significa a existência de uma paralisia mas tão só dum transtorno motor devido a lesões ao nível do encéfalo, que conduzem a alterações do movimento e da postura. Pela sua generalidade este termo abarca vários transtornos específicos unidos entre si por uma característica comum, a perda de controlo motor.

A paralisia cerebral está englobada dentro das enfermidades do sistema nervoso central. Os três síndromas motores, piramidal, extrapiramidal e cerebeloso, representam transtornos ao nível do sistema nervoso central, e são determinantes dos vários tipos de paralisia cerebral. A área motora piramidal, ou neurónio motriz superior, é responsável pela mobilidade voluntária. As lesões nos neurónios motrizes superiores afectam os movimentos voluntários, e como tal a fala, originando anomalias na função muscular: espasmodicidade; debilidade; limitações no alcance; e lentidão dos movimentos. A área motora extrapiramidal está determinada pelos centros motores subcorticais controlados pelo córtex. A lesão dos neurónios da área extrapiramidal afecta os movimentos involuntários e impulsivos, assim como o controlo automático da marcha e da corrida.

A paralisia cerebral pode ser definida como uma sequela duma afecção encefálica que se caracteriza primordialmente por um transtorno persistente, mas não invariável, do tónus, da postura e do movimento, que aparece na primeira infância, e não só é consequência directa desta lesão não evolutiva do encéfalo, mas também da influência que tal lesão exerce no amadurecimento neurológico [Bar84].

Três características principais permitem delimitar a paralisia cerebral:

- os transtornos são devidos a uma lesão cerebral que interfere com o

desenvolvimento normal da criança;

- a paralisia cerebral distingue-se primordialmente pelos danos causados nas funções motoras, os quais afectam o tónus (contração muscular em repouso), a postura (equilíbrio do indivíduo) e o movimento (acção motora voluntária);
- a lesão não é evolutiva, mas as suas consequências são variáveis ao longo do desenvolvimento da criança; as células nervosas (neurónios) lesionadas não crescem, não se espalham, nem afectam as restantes células nervosas vivas.

Os distúrbios motores afectam na maioria dos casos os órgãos motores da boca e dificultam o desenvolvimento da alimentação e da fala.

Um grande número de causas pode ser responsável pela paralisia cerebral. Ponto assente é o seu carácter não hereditário. As lesões são adquiridas, fruto de um acidente, e não herdadas. As lesões cerebrais podem ocorrer durante a gravidez, no parto ou no período pós-natal, em particular nos primeiros seis meses de vida. As causas que podem conduzir à paralisia cerebral nos primeiros quatro meses de gestação têm a ver com infecções embrionárias devidas a rubéola, sífilis, herpes ou hepatite, e dão lugar a malformações cerebrais, oculares, auditivas ou cardíacas. Desde o quarto mês até final da gravidez, o feto tem já uma certa autonomia e são observáveis algumas manifestações do seu sistema nervoso, como os movimentos de braços e pernas. A contração pela mãe de alguma enfermidade intra-uterina que não resulte na morte do feto usualmente deixa sequelas que se podem manifestar sob a forma de hipertonia, transtornos oculares e deficiências mentais.

O parto é o período mais crítico uma vez que é no seu decurso que a probabilidade de ocorrência de lesões cerebrais que conduzam à paralisia cerebral é maior [Bar84].

Algumas das razões são:

- o choque técnico; a criança quando nasce vê-se bruscamente obrigada a adaptar-se a condições vitais muito distintas (respirar, adaptar o seu sistema cardíaco, oxigenar o sangue, passar de um meio aquático de 37°C de temperatura a outro aéreo de menor temperatura);
- anoxia (falta de oxigénio no sangue) em virtude dum parto demasiado prolongado, voltas do cordão umbilical, dificuldades para passar a cabeça devido à estreiteza da pélvis, ou aplicação de fórceps;

- prematuridade;
- hipermaturidade.

No período pós-natal a paralisia cerebral surge devido a enfermidades metabólicas, meningite, desidratações, traumatismos devidos a acidentes graves ou acidentes anestésicos.

A sensibilidade individual de cada encéfalo aos agentes nocivos é muito diferente, pelo que as causas enumeradas, não sendo determinantes, podem levar ao aparecimento de paralisia cerebral.

As alterações motoras na parálítico cerebral podem assumir diferentes formas dependendo da localização da lesão cerebral. Nalguns casos os movimentos tornam-se rígidos e bruscos. A debilidade dos movimentos voluntários afecta também os lábios e a língua. O rosto apresenta-se inexpressivo e a saliva escorre pelos cantos da boca. Noutras manifesta-se sob a forma de movimentos involuntários e incessantes que se agravam com a fadiga e as emoções, e se atenuam em repouso desaparecendo durante o sono. Esses movimentos involuntários afectam os músculos articuladores e faríngeos, e a língua, originando igualmente contracções irregulares do diafragma e doutros músculos ligados à respiração, atingindo a produção da fala. Casos há em que o parálítico cerebral não consegue aperceber-se da força nem controlar a direcção dos seus movimentos. Os movimentos livres dos braços provocam instabilidade do tronco causando uma grande insegurança. Estas diferentes formas raras vezes aparecem isoladas sendo o mais comum uma mistura de várias delas.

Outros distúrbios podem igualmente aparecer associados à paralisia cerebral. Os mais comuns são os distúrbios sensoriais, de percepção, intelectuais e de linguagem. Os distúrbios visuais são bastante frequentes e variados no parálítico cerebral [Bar84]. Uma grande maioria apresenta estrabismo, capacidade visual enfraquecida e campo visual reduzido. A dificuldade no controlo dos movimentos manifesta-se igualmente a nível da visão realçando-se os desvios involuntários do olhar, dificuldade de coordenação visual, de fixação do olhar e dos movimentos necessários ao acompanhamento de pessoas ou objectos em deslocamento. As dificuldades de percepção alteram a apreensão e a interpretação correcta das informações referentes ao próprio corpo e ao meio envolvente, manifestando-se na criança com paralisia cerebral através da incapacidade demonstrada de brincar e jogar de forma construtiva, apesar da sua destreza e inteligência. Os distúrbios psíquicos

são normalmente consequência do meio no qual a criança com paralisia cerebral se move. A sua situação motora e o atraso geral na evolução fazem com que persista a fase de dependência ou de simbiose com a mãe, o que conduz a um atraso no início do funcionamento psíquico diferenciado. Um ambiente familiar tenso, o não tratar da criança com normalidade e a sobreprotecção, levam a alterações do estado psíquico da criança.

As alterações motoras são o factor responsável pelos distúrbios de linguagem. A linguagem é uma função que permite ao homem comunicar com os seus semelhantes. A fala, como expressão oral da linguagem, é um acto individual de vontade e de inteligência, no qual convém distinguir as combinações através das quais o sujeito utiliza o código da língua com vista a expressar o seu pensamento pessoal, do mecanismo psico-físico que permite exteriorizar essas combinações. O paralítico cerebral encontra-se privado, total ou parcialmente, deste segundo ponto, já que os seus transtornos motores o impedem de utilizar o tal mecanismo psico-físico e exteriorizar aquelas combinações que, em última análise, constituem a fala.

O objectivo deste trabalho é exactamente o de dotar o paralítico cerebral de meios economicamente acessíveis que aumentem, da forma mais eficaz possível, a sua capacidade de comunicar, proporcionando-lhe uma forma de expressão dos seus pensamentos e da sua vontade. No entanto, no alcançar deste objectivo, há que ter em conta todos os outros problemas expostos anteriormente, como as dificuldades motoras e de acuidade visual, para que o objectivo seja plenamente alcançado. A secção seguinte apresenta uma súmula desses problemas, das possibilidades e limitações do paralítico cerebral, concluindo com um conjunto de pontos que devem ser tidos em conta na especificação técnica.

4.2 Identificação de requisitos

Da breve descrição efectuada sobre a paralisia cerebral na secção anterior, existem dois factores principais a ter em conta na concepção dos sistemas de CAA para os utilizadores-alvo: os problemas de mobilidade; e os distúrbios visuais apresentados pelo paralítico cerebral. A importância destes factores advém do facto de, sendo os sistemas de CAA externos ao indivíduo, este necessitar de interagir com eles através dos seus meios sensoriais, com especial relevo para a visão, meio sensorial por excelência para apreensão da envolvente. Mas a interacção não se pode restringir a essa apreensão. É necessário

proporcionar ao indivíduo formas de actuar sobre essa envolvente permitindo-lhe expressar a sua vontade. Dadas as limitações motoras apresentadas pelo parálítico cerebral, essa forma de actuação reveste-se de especial importância. Dificuldades de interacção podem conduzir a situações de frustração e, em última instância, de rejeição da própria ajuda.

Nos casos mais problemáticos, a capacidade motora restringe-se à efectivação dum único movimento motor. Como visto na secção 3.1, a forma de se conseguir aproveitar esse movimento é através dum simples botão de pressão. Como tal os sistemas devem ser de operação suficientemente simples, de modo a possibilitar o seu controlo através desse único botão. De salientar que, sem prejuízo duma evolução na aprendizagem, tal situação se deve manter, garantindo a consistência na utilização da ajuda.

Devido aos distúrbios visuais apresentados pelos indivíduos com paralisia cerebral, alguns cuidados devem ser tidos em consideração na criação das *interfaces* visuais com os sistemas de CAA. O uso de símbolos, letras, palavras ou números, deve ter em conta a deficiente acuidade visual e o reduzido ângulo de visão apresentado por uma grande parte destes indivíduos. Como tal, a área ocupada pela *interface* visual não deverá ser demasiado grande. Normalmente a dimensão dos símbolos empregues nos sistemas de CAA utilizados pelos parálíticos cerebrais ronda os 25×25 mm, por forma a serem correctamente percebidos por estes, sendo que o número máximo numa *interface* não excede os 32, disposto numa matriz de 4 por 8. Tendo em conta estes pontos, torna-se necessário o recurso a esquemas de paginação ou codificação que permitam o acesso a um número maior de mensagens que aquelas directamente acessíveis através dessas 32 células. Esses esquemas devem ser facilmente perceptíveis por forma a não provocar um exagerado esforço de concentração e memorização, provocando uma dispersão da atenção e conduzindo a situações de fadiga. Uma outra possibilidade é o recurso a técnicas de codificação do tipo descrito em [FTR+94] e já referida na secção 2.3. Esta técnica evita os problemas resultantes da necessidade de acompanhamento do movimento de um cursor nos dispositivos que empregam técnicas de selecção por varrimento, embora conduza a uma maior carga cognitiva sobre o utilizador.

Todas as ajudas concebidas devem ter como resultado final a locução da mensagem seleccionada. Como visto na secção 3.2, quanto maior forem as unidades (fonemas, palavras ou frases), maior a naturalidade alcançada na reprodução, mas menor o grau de flexibilidade. No sentido dum aumento dessa flexibilidade é desejável que as expressões

sejam construídas a partir de palavras pré-guardadas, embora isso conduza a uma perda geral da qualidade. Correspondendo a esse desejo, as ajudas a conceber devem possibilitar a concatenação de várias palavras de molde a ser possível a sua sequenciação e posterior reprodução como uma única frase.

Outro ponto a ter em conta, não directamente relacionado com as limitações colocadas pela paralisia cerebral mas comum a todos os sistemas de CAA tendo em vista garantir a maior facilidade e versatilidade possíveis na sua utilização, é a sua portabilidade. A dimensão e o peso do sistema devem ser tais que permitam o seu fácil transporte, nomeadamente o seu acoplamento a uma cadeira de rodas. A não esquecer igualmente os problemas relacionados com o consumo energético dos sistemas. Este deve ser o mais baixo possível permitindo o uso de pilhas ou baterias, sem penalizar excessivamente a sua autonomia.

Por último, mas não menos importante, é o custo dos sistemas comercialmente disponíveis [FGL95]. Como facilmente se percebe, sistemas com custos da ordem das várias centenas de milhares de escudos, com alguns mesmo a ultrapassar o milhão, não estão ao alcance do comum cidadão deficiente português. A escassez de apoios para a aquisição de material de ajuda que, apesar do seu preço elevado, não é de modo algum um luxo mas uma necessidade, não permite uma maior integração do cidadão deficiente na sociedade e maiores possibilidades de alcançar a sua própria auto-sustentação. Daí que o problema do custo tenha sido desde início um dos vectores principais de orientação deste trabalho.

Tendo em conta estes pontos, procedeu-se à especificação técnica dos sistemas a qual constitui o tema da próxima secção.

4.3 Especificação técnica

A concepção e desenvolvimento de qualquer sistema deve ter como base um modelo que represente as interacções entre o homem e a máquina. A figura 4.1 apresenta um modelo genérico dessa interacção [Dem94].

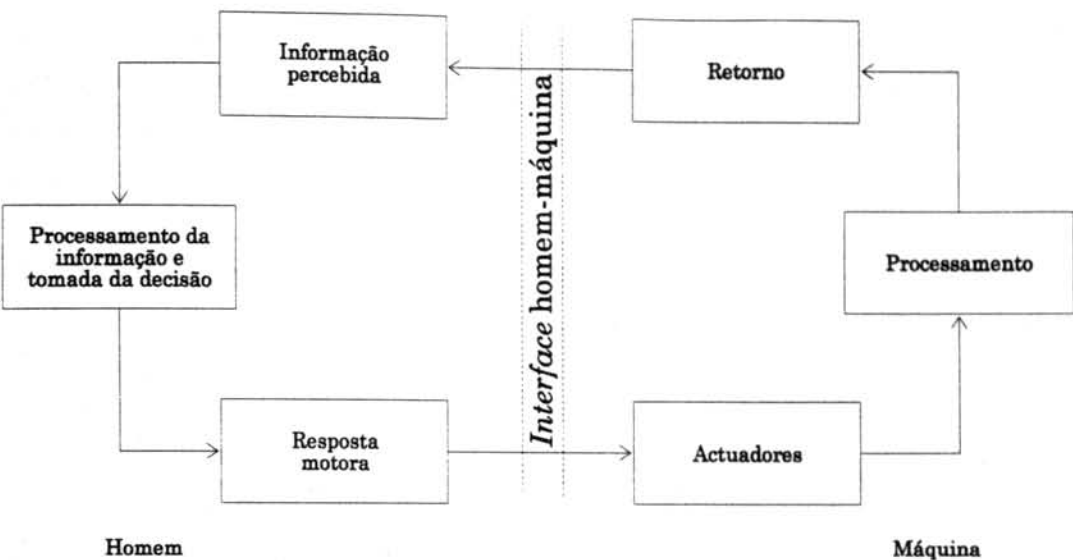


Figura 4.1 : Modelo genérico de interação homem-máquina.

O modelo apresenta uma clara simetria entre a parte esquerda, correspondente ao homem, e a direita, respeitante à máquina. As acções do homem (por ex. a resposta motora) interagem com as entradas da máquina (isto é, os actuadores), e as acções da máquina (por ex. um sinal sonoro) interagem com os sentidos do homem. Entre os blocos de entrada e de saída, directamente relacionados com a interacção homem-máquina, encontra-se a componente de processamento, a qual, no caso do homem, são os seus processos cognitivos, e que no caso da máquina poderá ser um microprocessador. Este modelo, embora simples, pode ser aplicado com eficácia a uma vasta gama de máquinas. Em termos dos sistemas de CAA, a componente actuadores corresponde tipicamente ao acesso físico directo a teclados ou botões de pressão. A componente retorno corresponde à saída de voz resultante da locução da mensagem escolhida, ou à emissão dum tom que confirme o sucesso da execução dum comando do qual não resulte a reprodução imediata duma mensagem. A componente processamento mapeia as acções do utilizador nos comandos e nas mensagens seleccionadas através da técnica de selecção. Acessoriamente poderá também ser responsável pela oferta de escolhas de comandos e mensagens ao utilizador, no caso do uso de técnicas de selecção por varrimento.

Os processos sensoriais a considerar na *interface* com a máquina são, no caso dos sistemas de CAA, essencialmente a audição, a visão, e o tacto, embora este último só seja de considerar no caso em que o utilizador possui a necessária coordenação motora para conseguir aceder directamente a um teclado. No caso particular da visão há a considerar os vários requisitos identificados na secção anterior, como sejam a acuidade visual, o

campo visual e a dificuldade de realização de movimentos de acompanhamento. Na audição devem ser considerados alguns pontos como amplitude, frequência e ruído ambiente. O factor inteligibilidade não se coloca no caso presente visto se ter optado, pelas razões já apontadas na secção 3.3, pela digitalização em detrimento da síntese da voz. Quanto à interacção física temos a considerar uma vasta gama de possíveis interactuadores como a cabeça, os braços, as pernas, as mãos ou os dedos, e vários parâmetros de actuação como sejam a força, a velocidade e a precisão. Para as pessoas com paralisia cerebral que utilizam sistemas de CAA, as limitações físicas são a mais importante das barreiras que impedem a efectivação da comunicação. Por último, mas não menos importante, há a referir os aspectos cognitivos afectados pela utilização dos sistemas de CAA. Ponto acente é o de que a carga cognitiva derivada da utilização dum sistema deste tipo deve ser mínima por forma a não desviar a atenção do utilizador do seu objectivo principal, a conversação. Essa carga pode manifestar-se de diversas formas, como sejam um esforço acrescido para a coordenação motora, uma atenção especial em acontecimentos que ocorram na *interface* visual, um esforço demasiado grande de memorização devido ao uso de técnicas de codificação inadequadas ou emprego de regras de utilização demasiado complexas, originando um *stress* adicional e conduzindo rapidamente a situações de fadiga. O retorno assume neste caso papel importante ao proporcionar uma confirmação do sucesso ou falhanço da acção.

O sistema deve apresentar um modelo conceptual simples que permita ao utilizador a fácil previsão do resultado da interacção. A relação entre um comando e os seus efeitos deve ser clara e produzir os resultados esperados.

Tendo em conta os factores e os requisitos expostos, optou-se pela concepção de dois sistemas, ambos com saída de voz, mas com diferentes graus de complexidade.

O primeiro, bastante simples, destinado a cobrir os requisitos de iniciação ao uso de sistemas de CAA bem como os de utilizadores com baixo potencial cognitivo, deverá consistir num sistema capaz de armazenar apenas uma mensagem de curta duração, da ordem da dezena de segundos, com possibilidade de gravação e reprodução incorporadas. O retorno da acção será a reprodução audível da mensagem. Aliás, esta será a única função disponibilizada directamente ao utilizador, uma vez que a gravação terá de ser efectuada por uma pessoa não deficiente. A interacção com o sistema deverá ser realizada por acesso directo através de uma tecla e com possibilidade de ligação dum botão de pressão externo, como por exemplo do tipo apresentado na figura 3.1. A sua operação,

tendo em conta a simplicidade do dispositivo, não implicará longos períodos de aprendizagem. Este sistema não colocará qualquer tipo de problema relacionado com a *interface* visual. Um símbolo de dimensões adequadas, tipicamente 25×25 mm, poderá facilmente ser apostado na caixa que protegerá o sistema. As dimensões deste deverão ser reduzidas, optando-se o mais possível por componentes de elevada integração, o que conduzirá, como desejável, a um baixo consumo, permitindo o uso de pilhas para a sua alimentação e simultaneamente a obtenção de uma grande autonomia, garantindo-se desta forma a sua portabilidade. Obviamente, e tendo em conta as suas características, o seu custo deverá ser baixo, da ordem dos poucos milhares de escudos.

O segundo sistema, mais avançado, deverá manter uma facilidade de operação comparável, mas possuir características que permitam ao utilizador uma maior proximidade aos meios naturais de comunicação. Neste sentido, o sistema deverá permitir a gravação, armazenamento e reprodução de um elevado número de mensagens, da ordem das poucas centenas. As mensagens deverão poder ser gravadas e apagadas independentemente ou na sua totalidade, sendo a sua duração individual arbitrária, embora obviamente limitada à capacidade total do sistema, próxima da dezena de minutos.

O elevado número de mensagens levanta problemas de acesso ao tentar manter-se a portabilidade do sistema, advindo daí a necessidade de serem usadas técnicas de codificação do tipo *Minspeak* ou numérica, referidas na secção 2.3, ou de paginação.

Tendo em conta o leque de utilizadores e as diferentes características por estes apresentadas, é aconselhável o desenvolvimento de vários tipos de *interfaces*, mas tentando manter uma base comum por forma a reduzir custos de desenvolvimento e, no futuro, de posterior produção. Esse módulo de base deverá conter toda a electrónica necessária à aquisição, tratamento, armazenamento e reprodução dos sinais de voz, bem como ao controlo dessas operações, implementando igualmente as funções de ligação com as entradas das diferentes *interfaces* com o utilizador.

Por tal deverá ser criada uma *interface* com teclado, para acesso directo, destinada a utilizadores com coordenação motora razoável. Acautelando possíveis dificuldades visuais o número de células será limitado a 32 (matriz de 4×8), correspondendo cada célula a uma tecla que, permitindo a afixação de símbolos com a dimensão de 25×25 mm, não conduz a uma área exagerada e ao comprometimento da portabilidade do sistema. Obviamente tal condição implica uma limitação do número de mensagens directamente acessíveis. No

sentido de ultrapassar essa limitação deverão ser implementados dois esquemas diferentes de codificação. O primeiro será baseado na criação de várias páginas, sendo o acesso às mensagens dentro destas realizada através da mesma matriz de células. A comutação entre as várias páginas deverá ser efectuada através do recurso a uma célula auxiliar, acessível da mesma forma que as restantes células. O segundo caso corresponderá a seleccionar duas células em sequência para aceder a uma mensagem. Este esquema permite a implementação de técnicas tipo compactação semântica. De modo a haver um retorno para o utilizador da acção executada, e uma vez que a selecção da primeira célula não implica a reprodução imediata da mensagem, o sistema deverá gerar um sinal audível, de curta duração, que confirme o seu sucesso. Em qualquer dos dois casos deverá ser possível a reconfiguração do tamanho das células, e consequentemente do seu número, pela associação de várias células adjacentes sobre uma mesma mensagem. Esta possibilidade conduz à satisfação de dois requisitos: o da criação de células de maior tamanho que permitam que o sistema possa ser utilizado por pessoas com uma coordenação motora mais reduzida; e a de numa fase inicial de aprendizagem o número de mensagens acessíveis ser menor, aumentando com o evoluir da aprendizagem sem que com isso haja perda de consistência ou alteração das regras de operação do dispositivo. A junção das células em dois, quatro, oito ou dezasseis grupos permite a gravação de outras tantas mensagens, conduzindo a um aumento do número de mensagens disponíveis à medida que o utilizador se adapta ao sistema.

Para qualquer das situações expostas anteriormente deverá estar sempre disponível uma função que permita a concatenação de várias mensagens e a sua reprodução de uma só vez. No caso de se usar paginação deverá ser possível a concatenação de mensagens pertencentes a diferentes páginas. Uma célula auxiliar permitirá indicar ao sistema que se pretende iniciar a concatenação duma sequência de mensagens. Uma vez que da escolha das mensagens a associar não resulta a sua imediata reprodução, o sistema deverá gerar um retorno audível que indique o sucesso da escolha. A reprodução da sequência deverá ser realizada após indicação por parte do utilizador de que terminou a escolha das mensagens a concatenar.

Alguns deficientes são capazes de controlar com eficácia e precisão os movimentos da cabeça, sendo nesses casos vulgar o uso de capacetes com ponteiro para acesso a teclados. Por tal, o teclado a usar deverá ser do tipo membrana, não exigindo pressões elevadas para a sua actuação e eliminando possíveis espaçamentos entre teclas, nos quais o ponteiro poderia ficar preso.

Uma parte dos paráliticos cerebrais não possui no entanto um suficiente controlo motor que lhe permita o uso de técnicas de acesso directo. Nos casos em que o utilizador apenas é capaz de realizar um único movimento com um mínimo de precisão e repetibilidade é necessário o emprego de técnicas de varrimento. Como tal é essencial a implementação duma *interface* em que a matriz de teclas seja substituída por uma matriz de lâmpadas. Cada célula, para além do símbolo, passará a conter num dos cantos uma pequena lâmpada, possivelmente do tipo LED (*Light Emitter Diode*). A selecção deverá ser realizada através dum único botão.

As técnicas de varrimento, já descritas na secção 3.1, implicam uma maior concentração durante a operação do dispositivo por parte do utilizador. Atendendo a esse ponto não é justificável a implementação da possibilidade de emprego de técnicas de codificação baseadas na compactação semântica, pois tal iria conduzir a exigências cognitivas sobre o utilizador demasiado grandes, provocando um desvio daquele que deve ser o ponto central de interesse, ou seja, a conversação. Por tal deverá ser apenas disponibilizada a possibilidade do uso de paginação e de concatenação de mensagens. O acesso a qualquer uma destas funções deve ser possibilitado da mesma forma que o acesso às mensagens, isto é, a cada uma corresponderá uma célula que deve poder ser seleccionada num ciclo completo de varrimento, tal e qual como uma vulgar célula de mensagem. O sistema deverá possibilitar a implementação de vários tipos de varrimento.

Por último, deverá igualmente ser implementada uma *interface* para codificação numérica do tipo referido na secção 2.3 e descrita em [FTR+94]. O uso de dois algarismos possibilitará o acesso a um total de cerca de 100 mensagens. No entanto, e no intuito de manter a coerência na operação do dispositivo, um dos códigos deve ser reservado para activação duma função que permita realizar a concatenação de mensagens, tal e qual como nos casos anteriores. Apesar do seu mais baixo consumo, não deverão ser usados mostradores de cristais líquidos para apresentação dos algarismos devido ao baixo contraste que apresentam, o que dificulta a sua percepção. O total controlo do sistema deverá ser igualmente possível através de um único botão, seleccionando-se sequencialmente o algarismo das dezenas e o das unidades.

O esquema apresentado na figura 4.2 ilustra as diferentes possibilidades que deverá ter este segundo sistema.

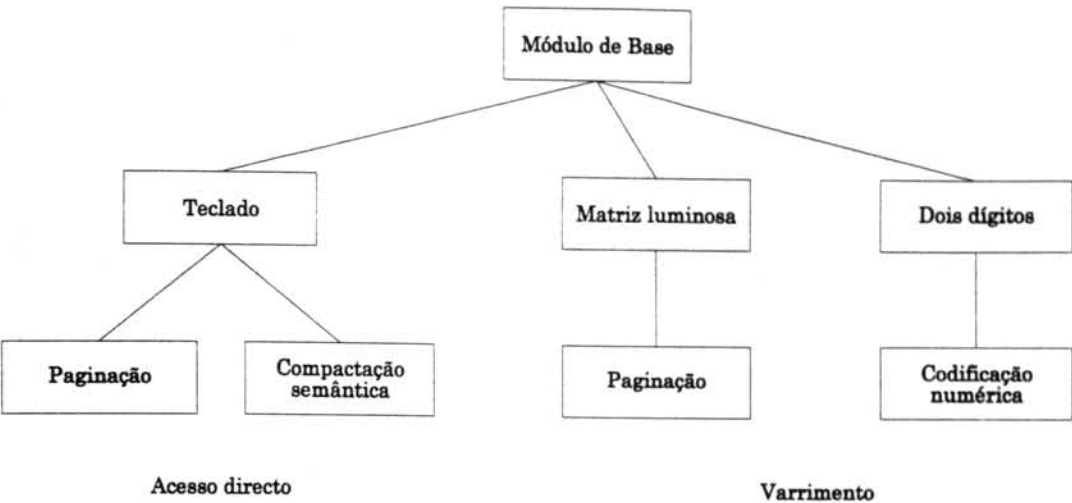


Figura 4.2 : Possibilidades a oferecer pelo segundo sistema.

A função de apagamento da totalidade das mensagens não deverá ser acessível ao utilizador através da mesma *interface*. A sua inadvertida actuação, resultado de qualquer movimento intempestivo, conduziria à perda da totalidade das mensagens gravadas. A sua manipulação apenas deve poder ser efectuada por uma pessoa não deficiente.

O consumo do sistema deverá ser o mais baixo possível de molde a poder ser alimentado a baterias, garantindo-se a sua portabilidade e uma elevada autonomia de funcionamento. Por tal o número de componentes deverá ser o mais reduzido possível, devendo-se optar por circuitos de elevada integração e memórias para as mensagens que retenham a informação mesmo na total ausência de alimentação. Evita-se deste modo a necessidade de se manter o circuito permanentemente alimentado, o que levaria a uma diminuição drástica do tempo útil de utilização e a um constante cuidado com a carga das baterias para que a informação armazenada não se perdesse.

A escolha e organização do vocabulário é obviamente da competência dos terapeutas da fala e dos próprios utilizadores e seus familiares, não sendo por tal realizada qualquer referência nesta especificação a esse ponto. Apenas se procura oferecer um leque de possibilidades técnicas que permitam a implementação de esquemas com vários níveis de complexidade, adequados a cada caso.

Com base nestas especificações partiu-se para a implementação prática dos sistemas, cuja descrição é assunto do próximo capítulo.

5. Implementação da Especificação Técnica

Depois de no capítulo anterior se terem identificado os requisitos e realizado a especificação técnica dos sistemas que se pretendiam ver desenvolvidos, resta expor como se procedeu à sua implementação e quais as soluções adoptadas.

Numa tentativa de se cobrirem o maior número de requisitos foram, no âmbito do projecto “Fala-Barato”, desenvolvidos dois sistemas de CAA. O primeiro, bastante simples, possibilita a gravação/reprodução de apenas uma mensagem de curta duração. O segundo permite a gravação/reprodução dum número muito maior de mensagens e o uso de diferentes *interfaces* com o utilizador. Consoante os casos é pois possível escolher aquela que melhor se adapte às suas características.

O capítulo encontra-se dividido em duas grandes partes, correspondendo à descrição de cada um dos sistemas. Para o segundo sistema, uma subdivisão permite aprofundar o funcionamento do módulo de base e de cada uma das diferentes *interfaces*.

5.1 “Fala-Barato I”: máxima simplicidade, satisfação mínima de requisitos

O “Fala-Barato I” é o mais simples dos dois sistemas desenvolvidos. Baseado num integrado com memória analógica incorporada, é capaz de gravar e reproduzir uma única mensagem com uma duração máxima de 16 segundos. O sistema destina-se sobretudo a ser usado como ferramenta de iniciação ao uso de dispositivos de CAA com saída de voz, bem como por utilizadores com capacidades cognitivas que não lhes permitam o uso de sistemas mais complexos. O seu baixo custo permite o desenvolvimento de diferentes abordagens estratégicas, nomeadamente com recurso ao uso simultâneo de vários destes sistemas, permitindo a manutenção da consistência na utilização em estágios de aprendizagem mais avançados. A interacção com o sistema é bastante simples. Para a reprodução da mensagem previamente gravada basta accionar um botão de pressão incorporado. O sistema possibilita igualmente o acoplamento dum botão externo. O microfone incorporado permite a gravação da mensagem de uma forma simples e rápida. A base deste sistema é um integrado da ISD (*Information Storage Devices*), baseado na tecnologia DAST descrita na secção 3.4. O diagrama de blocos interno deste integrado está representado na figura 5.1.

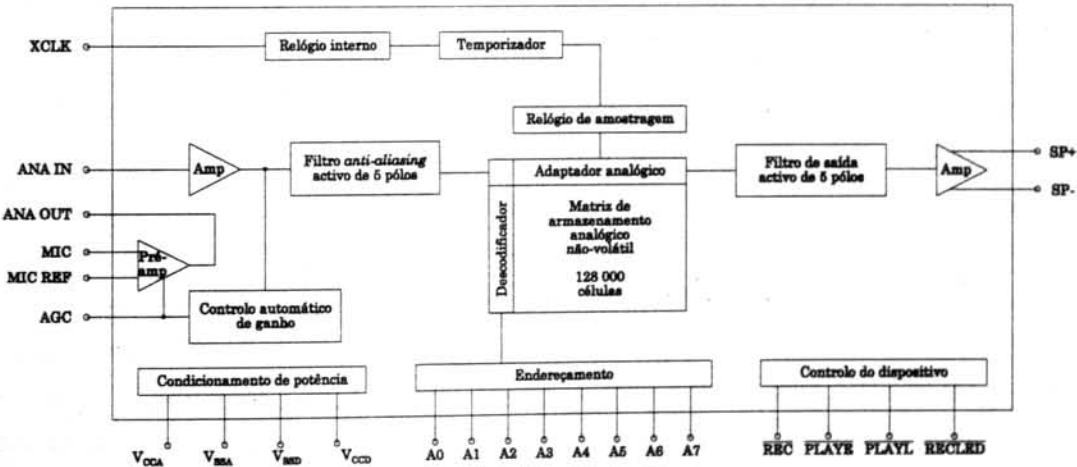


Figura 5.1 : Diagrama de blocos do ISD1416.

Este integrado incorpora todas as funções de condicionamento de sinal necessárias quer na gravação quer na reprodução, bem como a matriz de células para o armazenamento

não volátil dos níveis analógicos.

Durante a gravação, os blocos de pré-amplificação, de amplificação e de controlo automático de ganho adaptam a amplitude do sinal de entrada à gama dinâmica do circuito de armazenamento. O pré-amplificador deve ser ligado ao microfone por meio dum condensador por forma a remover a componente contínua do sinal. A amplificação é realizada em duas fases: inicialmente pelo pré-amplificador de entrada; e posteriormente pelo amplificador. A ligação entre a saída do pré-amplificador e a entrada do amplificador é implementada externamente através dum condensador (pinos ANA OUT e ANA IN). Esta arquitectura permite uma grande flexibilidade na concepção dos sistemas, principalmente para aplicações fora da área da voz, permitindo igualmente a introdução dum pólo adicional para baixas frequências de corte. O circuito de controlo automático de ganho monitoriza permanentemente o nível do sinal à saída do amplificador, gerando uma tensão que continua e automaticamente ajusta o ganho do pré-amplificador, de forma a manter num nível óptimo o sinal à entrada do filtro. Esta implementação permite conservar o sinal num nível elevado, evitando distorções ou cortes na gravação devidas a grandes variações na sua amplitude. O tempo de resposta do circuito de controlo automático de ganho a um aumento ou redução do sinal podem ser ajustados através dum condensador e duma resistência externas, ligadas ao pino AGC. A sua acessibilidade permite ao projectista configurar o sistema adaptando-o às suas necessidades específicas. Os cerca de 20 dB de variação do ganho do pré-amplificador permitem compensar variações no nível do volume da voz bem como adaptar-se à variedade de características dos microfones.

O passo seguinte no condicionamento do sinal é realizado pelo filtro. Conforme já visto na secção 3.4, embora o armazenamento seja, por natureza, analógico, continua a ser necessário o emprego de técnicas de amostragem. Como tal, o uso dum filtro *anti-aliasing* continua a ser necessário para remover, ou pelo menos reduzir a um nível insignificante, os componentes de frequência acima de metade da frequência de amostragem, satisfazendo o Teorema da Amostragem de *Nyquist*. Uma qualidade de voz superior à conseguida pela rede telefónica é alcançada com uma frequência de amostragem de 8 kHz. A frequência superior de corte escolhida para o filtro, 3,4 kHz, deixa uma largura de banda suficientemente extensa para permitir uma boa qualidade na reprodução da voz. O filtro é do tipo *Chebyshev*, de 5 pólos, com uma atenuação máxima na banda de passagem inferior a 2 dB, construído com base em técnicas convencionais (condensadores

não-comutados) sem recurso a componentes externos. O declive alcançado ronda os 40 dB por oitava aos 3,4 kHz. Após a filtragem o sinal está pronto a ser amostrado para posterior “escrita” na matriz de armazenamento analógica.

O sinal é amostrado a uma frequência de 8 kHz, sendo as amostras ulteriormente sujeitas a um processo de deslocamento do nível de tensão por forma a serem alcançados os elevados valores de tensão requeridos pelo procedimento de escrita. O relógio usado na amostragem é igualmente utilizado para incrementar o descodificador da matriz, sendo as amostras mapeadas sequencialmente dentro desta.

Durante a reprodução as tensões analógicas gravadas são “lidas” igualmente de forma sequencial, à mesma cadência a que foi realizada a amostragem, reconstruindo-se desta maneira a forma de onda amostrada. Um filtro passa-baixo colocado na saída remove a componente frequencial introduzida pelo processo de amostragem, restaurando o sinal original. Na realidade, o filtro *anti-aliasing* de entrada e o filtro passa-baixo de saída são um só, sendo atribuído através dum *multiplexer* analógico a uma ou outra função, conforme necessário para a gravação ou a reprodução.

Um amplificador de potência colocado na saída permite a ligação directa dum altifalante, proporcionando uma potência eficaz de 12,5 mW sobre uma impedância de 16 Ω , suficiente para obtermos um sinal audível em boas condições numa sala de dimensões médias.

Uma matriz de 128 000 células de memória baseadas na tecnologia DAST, descrita na secção 3.4, proporciona um tempo total de gravação de 16 segundos, e uma retenção dos dados garantida por dez anos, sem alimentação [ISD94].

Ao fim dum ciclo de gravação ou reprodução, o dispositivo entra automaticamente num modo de espera, reduzindo o consumo para valores típicos aproximados da ordem dos 0,5 μ A. Esta característica é de extrema importância pois, reduzindo o consumo, aumenta significativamente a portabilidade dos sistemas baseados neste dispositivo. A figura 5.2 ilustra o circuito completo implementado para o protótipo desta ajuda técnica. De notar que quer a entrada quer a saída do sinal estão acessíveis externamente, sendo possível a substituição do altifalante e do microfone internos, por partes idênticas colocadas externamente, ou a ligação a amplificadores de potência ou fontes alternativas de sinal, respectivamente.

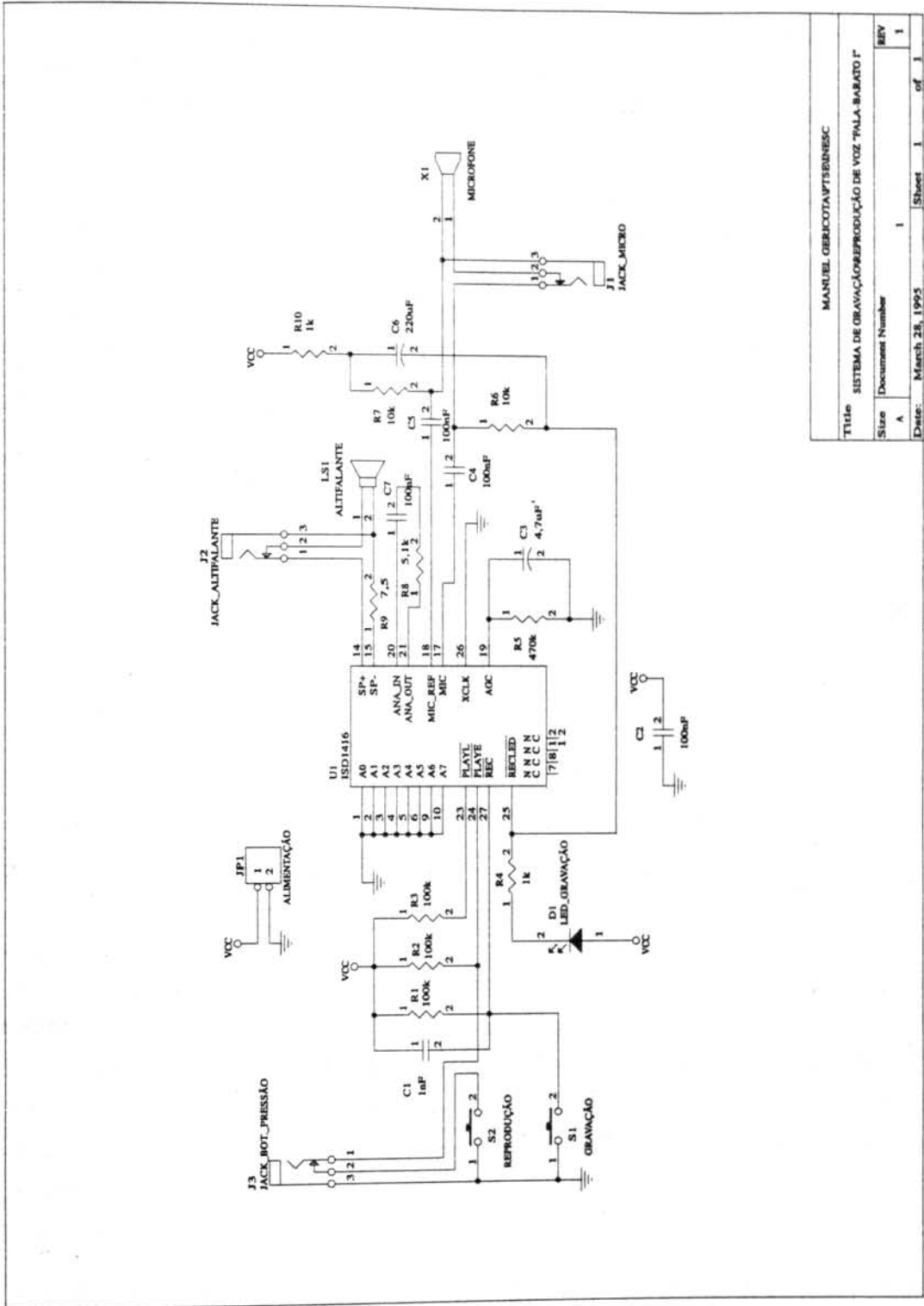


Figura 5.2 : Esquema eléctrico do “Fala-Barato I”.

Durante a gravação duma mensagem o botão de gravação deverá manter-se pressionado. O ciclo de gravação termina quando o botão for libertado ou quando o espaço de memória estiver totalmente ocupado. No primeiro caso o dispositivo coloca automaticamente no fim da gravação uma marca de fim de mensagem. Um LED vermelho proporciona um retorno visual enquanto decorre o ciclo de gravação. Este LED está ligado através duma resistência ao pino RECLEL. Este pino fica em nível baixo durante todo o tempo que durar a gravação. Quando o botão de reprodução é accionado, o sistema inicia a reprodução da mensagem. Este ciclo termina quando for encontrada uma marca de fim de mensagem, caso em que o sinal RECLEL passa momentaneamente ao nível baixo provocando o piscar do LED, ou quando for alcançado o fim do espaço de memória. De notar que basta accionar o botão, não sendo necessário mantê-lo pressionado até ao fim da reprodução da mensagem. A gravação duma nova mensagem provoca automaticamente o desaparecimento da anterior, independentemente do tempo de duração desta.

O pré-amplificador incorporado foi concebido para amplificar sinais na gama de 1 a 20 mV. Este pré-amplificador de transcondutância de ganho controlado possui uma impedância de entrada de aproximadamente 10 k Ω e um ganho máximo de 24 dB. O microfone de electrete usado fornece um nível adequado para atacar este pré-amplificador. Uma vez que a impedância de entrada é inalterável, a frequência inferior de 3 dB é determinada pela fonte sonora e pelo condensador de acoplamento. O uso dum condensador de acoplamento de 0,1 μ F conduz, pela equação

$$f_{3dB} = \frac{1}{2\pi RC}$$

a uma frequência inferior de corte de 159 Hz. Da mesma forma, a capacidade colocada entre os pinos ANA OUT e ANA IN determina a frequência inferior de 3 dB nesta secção do circuito. O amplificador possui uma impedância de entrada de aproximadamente 3 k Ω . O baixo valor da capacidade coloca a frequência inferior de 3 dB em cerca de 530 Hz. Reduzindo os baixos, maior será a energia transferida nas altas frequências, durante a reprodução, para o altifalante, fazendo com que o som pareça mais alto e claro, aumentando a inteligibilidade da mensagem.

A matriz de armazenamento analógica está organizada internamente em 160 colunas. Cada coluna proporciona um tempo de armazenamento de 100 ms. As colunas são endereçáveis individual e directamente, através dos pinos A0 a A7, podendo a gravação e a reprodução ter lugar a partir da posição inicial de qualquer uma das colunas,

continuando automaticamente através das seguintes. O contador que gera os endereços é interno ao dispositivo. Os pinos de endereçamento actuam somente como entradas, permitindo apenas definir a posição de início da gravação ou reprodução, mas não disponibilizam para o exterior o endereço "corrente". Na implementação realizada os pinos A0 a A7 estão todos a "zero". Temos desta forma um total aproveitamento da memória disponível, limitando no entanto a uma o número de mensagens. A divisão do espaço de memória em duas ou mais partes, com endereçamento directo de cada uma delas, permitiria o armazenamento de mais mensagens. No entanto, não se considerou tal necessário neste sistema, tendo em conta a população alvo.

A portabilidade deste sistema é garantida quer pelo baixo consumo (praticamente nulo quando fora de um ciclo de gravação ou reprodução), quer pela gama de tensões de alimentação que aceita, entre 4,5 V e 6,5 V, sem recurso a qualquer regulador, possibilitando o uso de pilhas ou baterias de pequeno tamanho e baixo custo. A figura 5.3 ilustra o sistema completo, montado numa placa de circuito impresso.

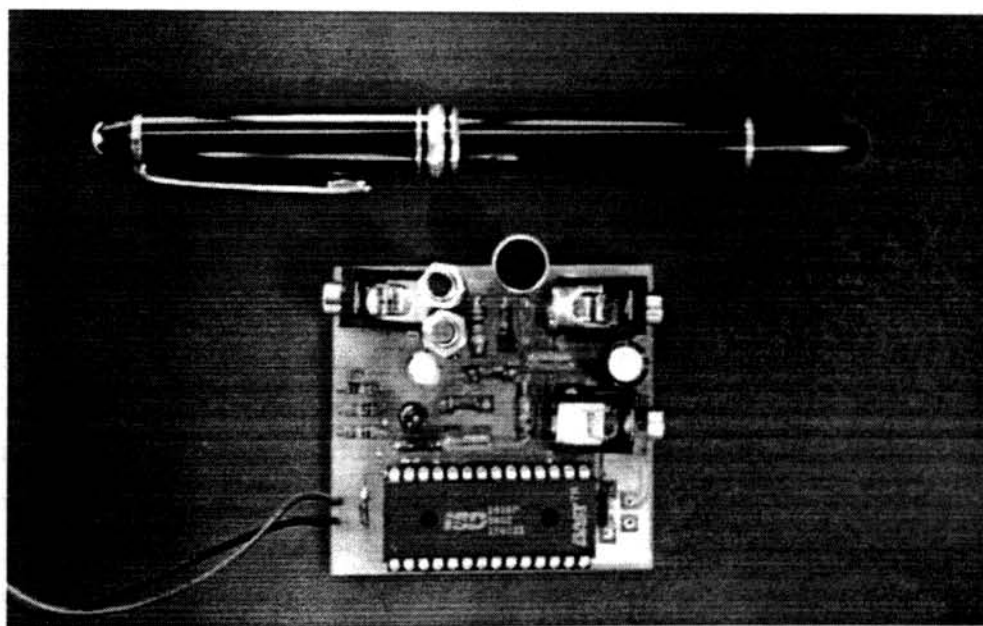


Figura 5.3 : O "Fala-Barato I".

A qualidade sonora do sistema constituiu uma agradável surpresa, dadas as suas dimensões e simplicidade.

Embora as possibilidades do sistema sejam limitadas, foi possível já no decorrer do trabalho identificar outras áreas em que a sua aplicação é possível e mesmo desejável,

quer como sistema de CAA, servindo outros destinatários que não as pessoas com paralisia cerebral, quer em áreas não relacionadas com este tema, nomeadamente em brinquedos, sistemas de alarme ou outros em que o emprego dum dispositivo com saída de voz torna menos impessoal o contacto com a tecnologia.

5.2 “Fala-Barato II”: satisfação completa de requisitos

Os sistemas de CAA portáteis actualmente disponíveis no mercado, alguns dos quais se encontram descritos na secção 3.2, apresentam preços extremamente elevados, na ordem das várias centenas de milhares de escudos. A especificidade dos sistemas contribui certamente para essa situação. Uma das formas de tentar contornar esse ponto é empregar tecnologias não desenvolvidas especificamente para esse fim, mas que, servindo outras áreas no vasto mercado da electrónica de consumo, possam ser obtidas a preços mais baixos. Outra seria a de criar um módulo de base que permitisse o desenvolvimento de várias variantes sobre o mesmo sistema, garantindo uma elevada versatilidade e simultaneamente minimizando custos de desenvolvimento e de posterior produção.

Mais do que um sistema individual, o “Fala-Barato II” é um conjunto de sistemas que, apresentando as mesmas características base em todos eles, se distinguem pelo seu modo de operação e *interface* de acesso, permitindo diferentes escolhas conforme as especificidades do utilizador. A sua versatilidade permite igualmente a fácil criação de outras variantes sobre o mesmo módulo de base, respondendo a necessidades individuais dum determinado utilizador. Cumpre-se desta forma um axioma que deve estar sempre presente quando se concebe um sistema de CAA: o de este ir de encontro às necessidades do indivíduo e não o de se esperar que este modifique o seu comportamento em face da tecnologia.

5.2.1 Arquitectura da solução desenvolvida

O coração do módulo de base é o DS2271 *Speech Stik* da DALLAS SEMICONDUCTOR, um módulo electrónico que é um subsistema completo de gravação/reprodução de som [SS91]. Este módulo digitaliza e armazena automaticamente os sinais de voz numa

memória externa. O uso de memórias do tipo NVSRAM (*Non-Volatile Static Random Access Memory*) permite manter esses sinais em memória mesmo que lhe seja retirada a alimentação. O sistema suporta um número máximo de 254 mensagens. A distinção entre elas é realizada pela atribuição dum número entre 001 e 254 (inclusive), colocado num cabeçalho criado pelo DS2271 e memorizado conjuntamente com cada uma das mensagens digitalizadas. A atribuição desse número é da responsabilidade do utilizador e não do subsistema. A atribuição a uma mensagem de um número já anteriormente outorgado provoca a substituição da anterior mensagem pela nova. O comprimento de cada mensagem é arbitrário, estando apenas limitado ao espaço de memória livre.

O DS2271 comunica com a memória das mensagens por intermédio de um barramento de dados de 8 linhas e de um barramento de endereços de 19 linhas, permitindo o endereçamento de até 512 *kbytes*.

O DS2271 possui dois modos de operação: *hardware* e *software*. Quando é aplicada a alimentação o DS2271 monitoriza um dos seus pinos (pino MODE) para determinar qual o modo de operação que deve seguir. O modo de *software* é assumido por omissão.

O modo de *hardware* não requer nenhum controlador externo para a sua operação. O controlo é realizado simplesmente deixando os pinos de controlo no ar ou ligando-os directamente à massa do circuito. Embora se tenha acesso a todas as funções básicas do DS2271 este modo de operação é pouco versátil, sendo útil especialmente em aplicações onde se pretenda apenas gravar e reproduzir um número reduzido de mensagens. Daí que para implementação do "Fala-Barato II" se tenha optado pelo modo de *software*. Neste modo é necessário o uso dum controlador externo. No nosso caso a escolha recaiu sobre um microcontrolador de 8 *bits* da família MCS¹-51 da *Intel*, como será analisado mais adiante. Os comandos são transmitidos ao DS2271 por meio dum porto série assíncrono que existe nos pinos RXD e TXD. Um modo de eco permite saber se os comandos enviados foram correctamente recebidos e executados pelo subsistema.

A figura 5.4 esquematiza em traços gerais a configuração base dum sistema neste modo.

¹ MCS é uma marca registada da *Intel Corporation*.

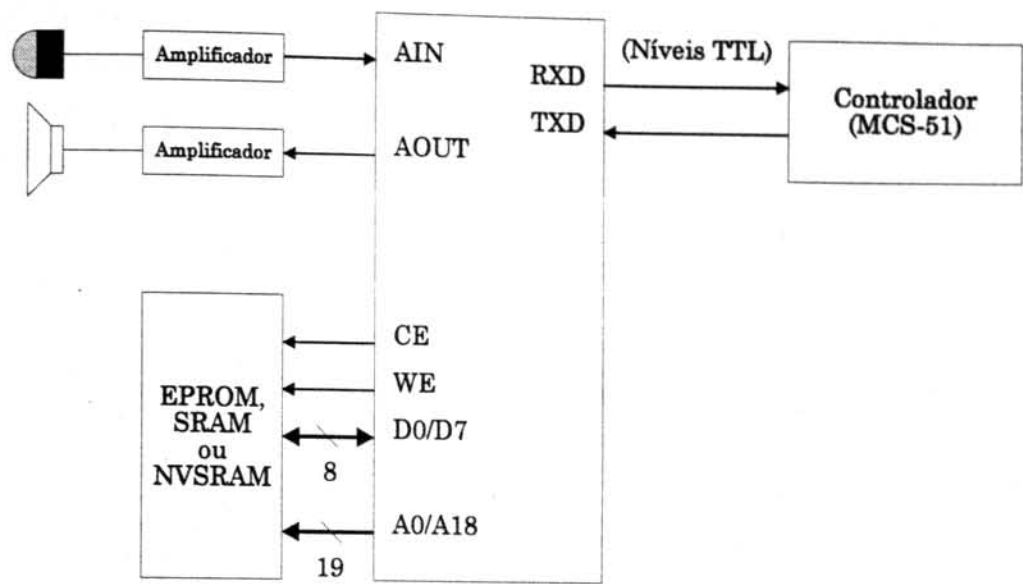


Figura 5.4 : Configuração base dum sistema baseado no DS2271 em modo *software*.

A comunicação série entre o DS2271 e o controlador segue o protocolo normalizado de 8 bits com um STOP bit e sem bit de paridade. Quando é aplicada a alimentação o subsistema monitoriza os pinos RATE0 e RATE1 por forma a determinar qual a velocidade de transferência usada no porto série. A tabela 5.1 mostra a configuração desses pinos em função da velocidade pretendida. Por omissão é assumida a velocidade de 1200 bit/s.

Tabela 5.1 : Pinos de selecção da velocidade de transferência da porta série.

Velocidade de transferência da porta série (bit/s)	RATE1	RATE0
57600	0	0
9600	0	1
2400	1	0
1200	1	1

No “Fala-Barato II” a velocidade de transferência implementada é de 9600 bit/s. Embora operando dentro do protocolo RS-232C, os níveis de tensão aplicados ao porto série são níveis TTL, sob pena de destruição do subsistema.

Antes da memória adstrita ao DS2271 poder ser usada para gravar sinais de voz

digitalizada, tem de ser formatada. A formatação da memória disponível consiste na sua divisão em blocos de 128 *bytes*. Uma mensagem gravada no DS2271 ocupará um número inteiro desses blocos. A organização da memória de voz e de cada um dos seus blocos após a formatação, está ilustrada na figura 5.5.

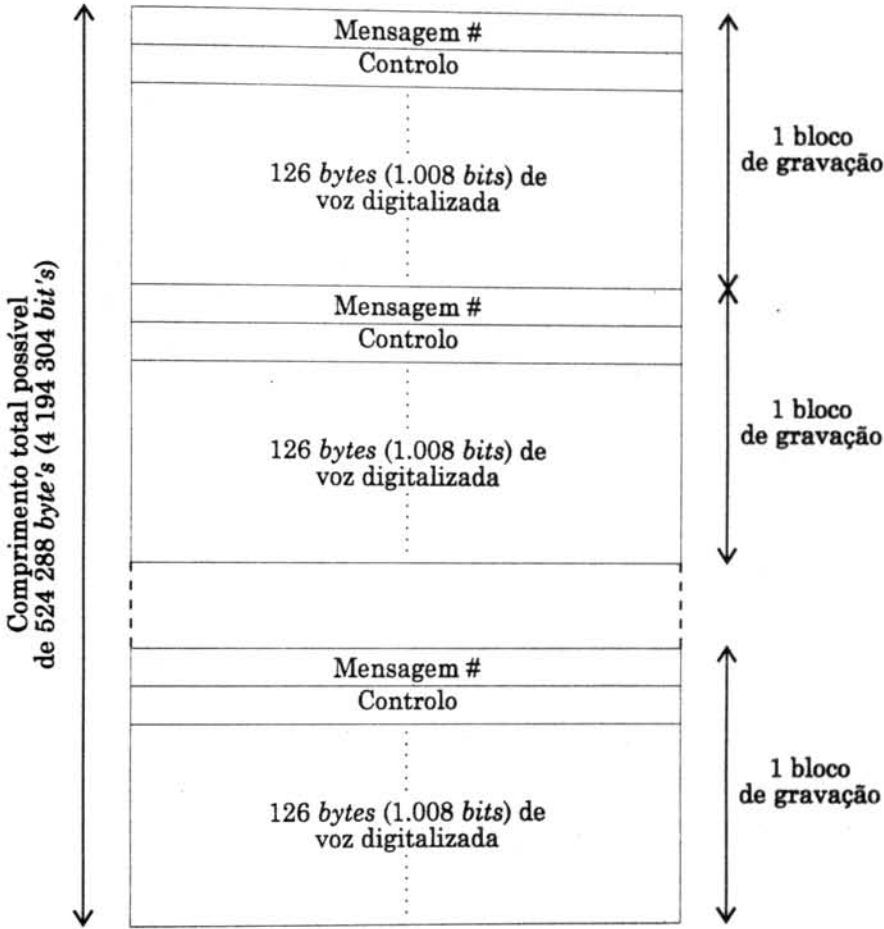


Figura 5.5 : Organização da memória de voz após a formatação.

O primeiro *byte* de cada bloco contém o número da mensagem. O segundo contém um *byte* de controlo, cuja constituição está ilustrada na figura 5.6, o qual inclui três tipos de informação:

- os *bits* R0 e R1 indicam a taxa de compressão à qual a voz é armazenada: (*k*bps - *k*bits por segundo)
R1='0', R0='0', 8 *k*bps;
R1='0', R0='1', 12 *k*bps;
R1='1', R0='0', 16 *k*bps;

$R1='1', R0='1', 24 \text{ kbps}$.

- o *bit* RST, caso esteja a '1', indica que esse bloco é o primeiro daquela mensagem;
- os *bits* L0 e L1 indicam qual o nível máximo de gravação atingido pelo sinal no bloco anterior:

$L1='0', L0='0', AIN > -15 \text{ dBm0}^1$;

$L1='0', L0='1', -15 \text{ dBm0} > AIN > -25 \text{ dBm0}$;

$L1='1', L0='0', -25 \text{ dBm0} > AIN > -40 \text{ dBm0}$;

$L1='1', L0='1', -40 \text{ dBm0} > AIN$.

Os restantes 126 *bytes* em cada bloco são preenchidos com os dados provenientes da digitalização da voz.

bit mais significativo				bit menos significativo			
L1	L0	x	x	RST	R1	R0	x

Figura 5.6 : Organização do *byte* de controlo.

O sinal de voz aplicado ao pino AIN é comprimido após a passagem pelo filtro *anti-aliasing*. O esquema de compressão usado pelo DS2271 é a μ -law, já referido na secção 3.3.2.2, a que não será alheio o facto de a sua proveniência ser os Estados Unidos da América. Ulteriormente o sinal é amostrado a uma frequência de 8 kHz e quantificado em 256 níveis por intermédio dum conversor analógico-digital de 8 *bits*. A frequência de amostragem é obtida por divisão da própria frequência de operação, de 11,0592 MHz, gerada internamente e disponibilizada para o exterior através do pino CLK. O resultado é um trem de impulsos série, a 64 *kbps*, em formato PCM normalizado.

Como exposto na secção 3.3.2.3, embora o PCM seja o método mais simples de codificação digital é também o mais dispendioso em termos de informação produzida, uma vez que não possui nenhum mecanismo que explore a redundância do sinal. Por tal facto, é de todo o interesse o uso de técnicas de codificação que, tirando partido da redundância, diminuam a quantidade de informação que é necessário armazenar, optimizando o espaço de memória disponível. O DS2271 tem disponíveis quatro taxas de compressão diferentes:

¹ 0 dBm0 corresponde a um valor eficaz de 1,231 V.

8 *kbps*; 12 *kbps*; 16 *kbps*; e 24 *kbps*. A taxa de 24 *kbps* segue a norma G.726 [CCI90], convertendo o sinal PCM a 64 *kbps* num sinal ADPCM a 24 *kbps*. Esta taxa de compressão é a que possibilita uma melhor qualidade sonora na reprodução do sinal, mas é também a mais dispendiosa em termos de memória ocupada. A tabela 5.2 mostra a relação entre o tempo de gravação disponível, em segundos, consoante a taxa de compressão usada, em função da capacidade da memória utilizada.

Tabela 5.2 : Capacidade de gravação disponível (segundos).

Capacidade de memória (bits)	Taxa de compressão seleccionada			
	8 <i>kbps</i>	12 <i>kbps</i>	16 <i>kbps</i>	24 <i>kbps</i>
64 k	8	5	4	3
128 k	16	11	8	5
256 k	32	21	16	11
512 k	64	43	32	21
1024 k	128	85	64	43
2048 k	256	171	128	85
4096 k	512	341	256	171

À medida que a taxa de compressão aumenta, aumenta igualmente o tempo de gravação disponível para os sinais de voz, embora à custa duma maior degradação da qualidade sonora. Todavia, mesmo usando a taxa de compressão mais elevada (8 *kbps*) a qualidade sonora é bastante boa, sendo perfeitamente inteligível. Aliás, e de forma a permitir uma maior duração das mensagens, foi esta a taxa escolhida para implementação no “Fala-Barato II”. Os algoritmos de codificação para as taxas de 8 *kbps*, 12 *kbps* e 16 *kbps* são proprietários da marca.

Na reprodução o processo é percorrido de forma inversa. O sinal armazenado na memória é descodificado e passado a um conversor digital-analógico, sendo posteriormente descomprimido e filtrado para remoção da componente frequencial introduzida pelo processo de amostragem. O sinal analógico resultante é disponibilizado no pino AOUT.

A sequência de passos necessária para a gravação e reprodução das mensagens é a seguinte:

- envio do comando FORMAT para formatação da NVSRAM adstrita ao DS2271;
- selecção da taxa de compressão a usar no armazenamento das mensagens;
- escolha do número da mensagem a gravar, a ser enviado com o respectivo comando;
- envio do comando para parar a gravação;
- envio do comando para reprodução acompanhado do número da mensagem pretendida.

As mensagens são gravadas de forma sequencial nos blocos de gravação definidos pela formatação. Sempre que se pretende gravar uma mensagem com o mesmo número de outra anteriormente gravada, deve enviar-se primeiramente um comando de ERASE dessa mensagem. Isso provocará um rearranjo automático das mensagens na memória de forma a que não fiquem "buracos" livres entre blocos de gravação ocupados, aproveitando ao máximo a capacidade disponível. Se este procedimento não for executado e se a nova mensagem tiver uma duração inferior à anteriormente gravada com o mesmo número, os blocos de gravação não "reescritos" continuam a ser considerados ocupados, embora contenham agora apenas informação sem interesse. No caso de ter uma duração superior, a nova mensagem irá ocupar blocos de gravação pertencentes a mensagens colocadas subsequentemente na memória, implicando a sua perda. Para evitar esta situação na implementação do "Fala-Barato II", cada comando de gravação é automaticamente precedido dum comando de ERASE.

Dois outros comandos não directamente relacionados com o processo de gravação/reprodução, mas utilizados em algumas das variantes do "Fala-Barato II", são o da geração de tonalidades e o da listagem do número de mensagens e do respectivo espaço ocupado, em número de blocos de gravação. O primeiro permite a geração dum tom com frequência à escolha entre três possíveis, e de duração variável. O segundo implica o envio por parte do DS2271 duma lista das mensagens presentemente guardadas na memória externa e do seu respectivo comprimento em número de blocos de gravação.

Todos os comandos enviados pela porta série bem como a informação recebida do subsistema, como eco ou como resposta a esses comandos, são codificados em formato ASCII. A tabela 5.3 lista os comandos usados pelo DS2271.

Tabela 5.3 : Lista de comandos do DS2271.

Nome do comando:	Gravação duma mensagem.
Código do comando:	R
Função:	O envio deste comando despoleta o início da gravação do sinal presente na entrada AIN. Esta situação manter-se-à até ser recebido um comando STOP ou até preencher completamente a memória.
Exemplo:	Gravação duma mensagem sob o número 15.
Enviado:	[R] [0] [1] [5] [CR]
Recebido:	[R] [0] [1] [5] [CR] [LF] [>]
Nome do comando:	Reprodução duma mensagem.
Código do comando:	P
Função:	O envio deste comando provoca a reprodução da mensagem pré-gravada. Esta situação manter-se-à até ao fim da mensagem.
Exemplo:	Reprodução da mensagem número 5.
Enviado:	[P] [0] [0] [5] [CR]
Recebido:	[P] [0] [0] [5] [CR] [LF] [>]
Nome do comando:	Apagamento duma mensagem.
Código do comando:	E
Função:	O envio deste comando provoca o apagar da mensagem da memória e a sua reorganização para que não sejam deixados "buracos" e a máxima capacidade de gravação esteja disponível.
Exemplo:	Apagar a mensagem número 126.
Enviado:	[E] [1] [2] [6] [CR]
Recebido:	[E] [1] [2] [6] [CR] [LF] [>]
Nome do comando:	Formatação da memória.
Código do comando:	Z
Função:	Formata a memória adstrita ao DS2271, dividindo-a em blocos de 128 bytes.
Bytes opcionais:	Dois caracteres indicam o número de blocos de 32 kbytes disponíveis na memória externa. O número máximo de blocos é de 16. A memória será totalmente apagada se a opção for 0.
Exemplo:	Formatação duma memória de 64kx8 bit.
Enviado:	[Z] [0] [2] [CR]
Recebido:	[Z] [0] [2] [CR] [LF] [>]
Nome do comando:	Paragem de gravação.

Código do comando:	S
Função:	Este comando força o fim da gravação.
Exemplo:	Fim da gravação duma mensagem.
Enviado:	[S] [CR]
Recebido:	[S] [CR] [LF] [>]
Nome do comando:	Geração de tonalidades.
Código do comando:	T
Função:	Este comando levará o DS2271 a gerar um tom na saída AOUT. O tom será gerado continuamente até ser enviado um comando de cessação da geração. Os tons de 350 Hz e 620 Hz são transmitidos a um nível de -12 dBm0 e o de 1004 Hz a um nível de 0 dBm0 (0dBm0 = 1,231 V eficazes)
Bytes opcionais:	Um carácter indica o tom a ser gerado. D = 350 Hz B = 620 Hz T = 1004 Hz 0 = cessa a geração do tom
Exemplo:	Geração dum tom com a frequência de 350 Hz.
Enviado:	[T] [D] [CR]
Recebido:	[T] [D] [CR] [LF] [>]
Nome do comando:	Seleção do algoritmo de compressão.
Código do comando:	A
Função:	Este comando permite seleccionar a taxa de compressão a usar no armazenamento dos dados em memória.
Bytes opcionais:	Dois caracteres indicam qual a taxa de compressão a ser usada. 08 = 8 kbps 16 = 16 kbps 12 = 12 kbps 24 = 24 kbps
Exemplo:	Escolha de uma taxa de compressão de 12 kbps.
Enviado:	[A] [1] [2] [CR]
Recebido:	[A] [1] [2] [CR] [LF] [>]
Nome do comando:	Listagem das mensagens gravadas.
Código do comando:	D
Função:	Este comando causa o envio por parte do DS2271 de uma lista de todas as mensagens actualmente guardadas e seu respectivo comprimento em número de blocos de gravação.
Exemplo:	Supondo gravadas as mensagens números 5 e 14 com 49 e 122 blocos de gravação, respectivamente.
Enviado:	[D] [CR]

Recebido: [D] [CR] [LF] [0] [0] [5] [] [0] [0] [4] [9] [CR] [LF] [0] [1]
[4] [] [0] [1] [2] [2] [CR] [LF] [CR] [LF] [>]

Nota: todos os caracteres entre [] são caracteres ASCII.

A possibilidade de se poder controlar o subsistema exclusivamente através de comandos enviados pelo porto série torna extremamente simples a sua operação. Para controlo do subsistema e geração desses comandos, bem como para permitir a interligação entre o módulo de base e as várias *interfaces* com o utilizador, optou-se pelo emprego dum microcontrolador da família MCS-51, o 80C51. A escolha deste microcontrolador ficou a dever-se a vários factores, dos quais são de salientar a sua grande divulgação (o microcontrolador interno do DS2271 é desta mesma família), o seu baixo custo, o possuir um porto série, 126 *bytes* de RAM interna e 4 *kbytes* de memória ROM incorporada, o que evita o emprego de memórias externas para armazenar o programa de controlo e liberta 16 linhas de entrada/saída, das 32 disponíveis, normalmente usadas como barramento de endereços e de dados para acesso a essas memórias. Obtém-se desta forma um sistema bastante compacto e com um consumo bastante baixo.

A figura 5.7 apresenta um diagrama de blocos geral do "Fala-Barato II".

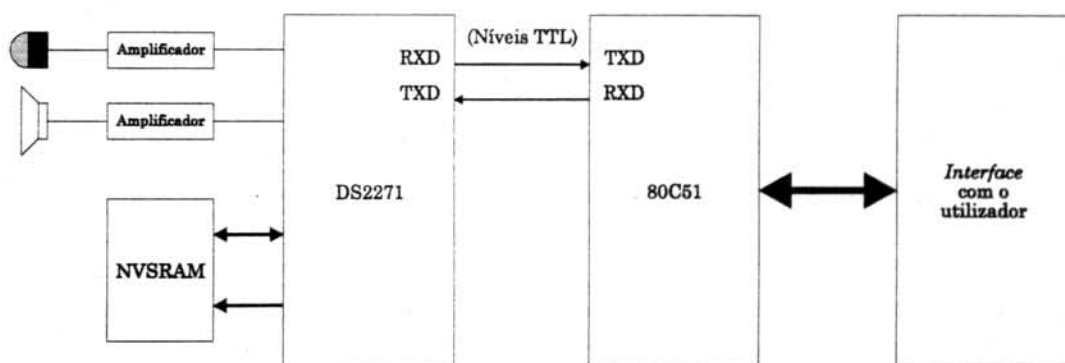


Figura 5.7 : Diagrama de blocos geral do "Fala-Barato II".

A figura 5.8 ilustra o protótipo concebido para o módulo de base do "Fala-Barato II". De salientar que este protótipo está preparado para receber qualquer uma das *interfaces* concebidas, sendo apenas necessário alterar o programa de controlo contido na memória interna do microcontrolador.

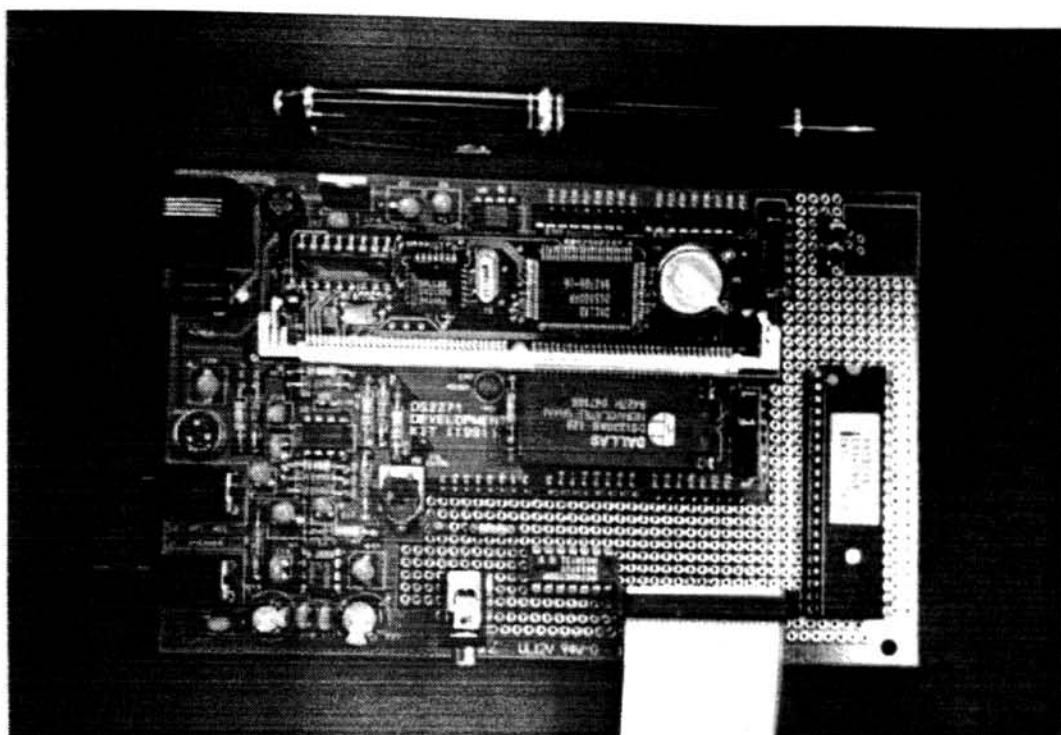


Figura 5.8 : Protótipo do módulo de base do "Fala-Barato II".

O 80C51 opera à frequência de 11,0592 Mhz obtida a partir do pino CLK do DS2271. É ele que interpreta as ordens recebidas do utilizador, através das diferentes *interfaces* de acesso disponibilizadas, e que em função do programa de controlo, envia através do porto série para o DS2271 os comandos necessários à execução dessas ordens. A figura 5.9 ilustra o circuito completo implementado para o protótipo do módulo de base desta ajuda técnica.

Devido à pouca memória disponível para armazenamento do programa de controlo, foi necessário optimizá-lo ao máximo, evitando-se qualquer tipo de redundância. Embora à custa de um maior tempo de desenvolvimento, o programa foi directamente escrito em linguagem de baixo nível e testado com a ajuda dum emulador. As diferentes *interfaces* com o utilizador implicam a alteração de algumas sub-rotinas e a introdução de outras, bem como a reformulação do programa principal.

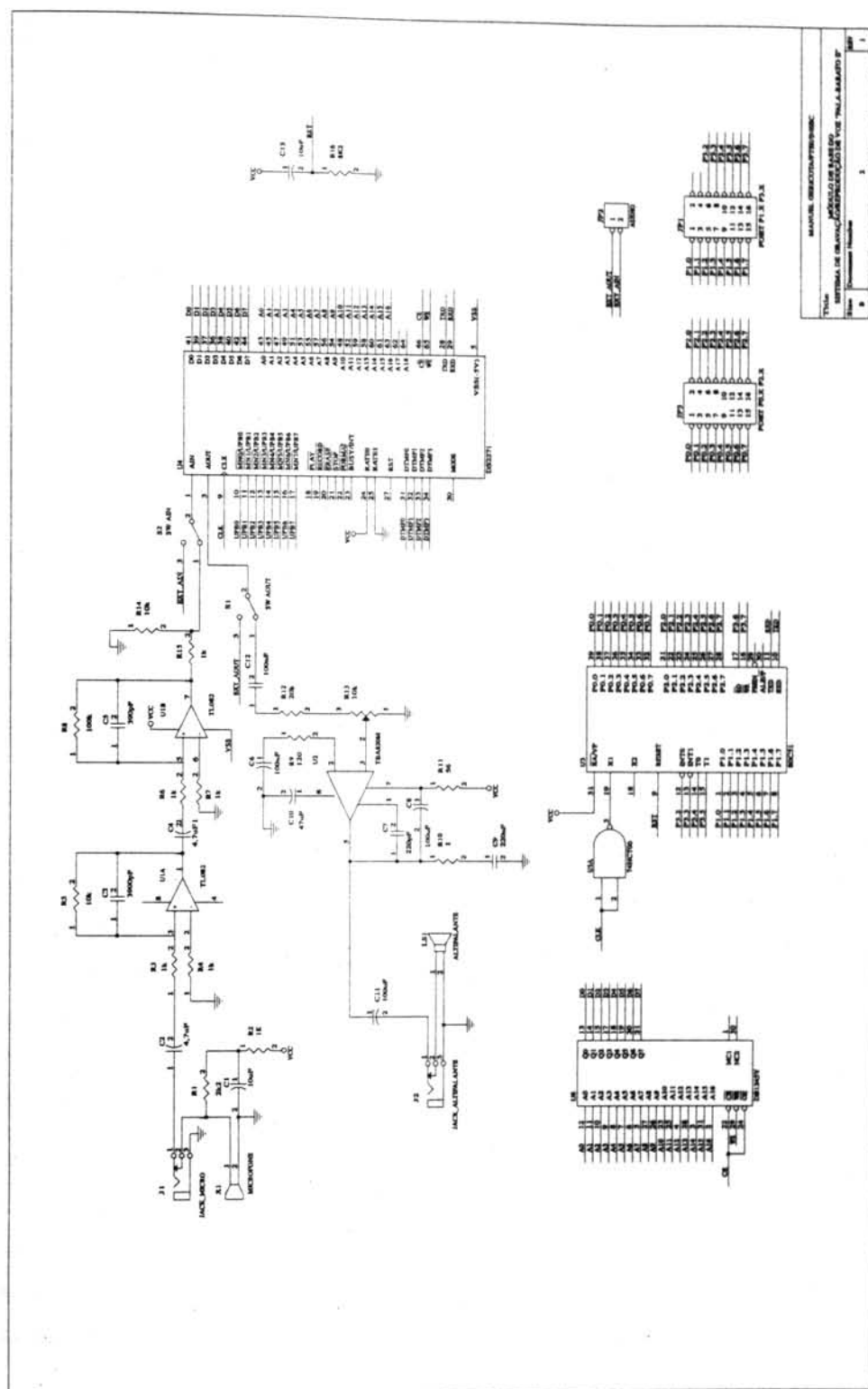


Figura 5.9 : Esquema eléctrico do módulo de base do "Fala-Barato II".

5.2.2 As diferentes possibilidades de acesso

Basicamente foram concebidas três *interfaces* com o utilizador, conforme especificado na secção 4.3, e cuja descrição é apresentada nas próximas secções.

5.2.2.1 *Interface* por teclado

A pensar nos utilizadores que apresentam um domínio razoável dos movimentos dos membros superiores ou da cabeça (utilização de um ponteiro de cabeça para a selecção), foi concebida uma *interface* por teclado que, permitindo a selecção directa, implementa a forma mais rápida de se atingir o objectivo final: a reprodução duma mensagem. Trata-se basicamente duma matriz de 32 teclas que permitem o acesso directo a 30 mensagens e a duas funções de comando disponibilizadas para o utilizador, no caso de se usar paginação, ou 31 teclas de mensagem e apenas uma de comando, no caso da implementação dum esquema do tipo compactação semântica, conforme definido na secção 4.3. A tecla de comando comum aos dois casos possibilita a concatenação de várias mensagens e a sua reprodução em sequência.

No caso da paginação, a outra tecla de comando permitirá a alteração da página de mensagens disponível, concedendo acesso a mais 30 mensagens. O sistema permite a implementação dum máximo de oito páginas acessíveis sequencialmente e de forma rotativa. Visto que da actuação da tecla de mudança de página não resulta a reprodução de nenhuma mensagem, o sistema gera um tom de breve duração que indica ao utilizador que a mudança de página foi efectuada com sucesso. Este processo é seguido sempre que da acção executada não resulte a reprodução imediata duma mensagem, por forma a que o utilizador obtenha um retorno que lhe confirme o sucesso da operação.

No segundo caso a selecção duma mensagem implica a escolha sucessiva de duas teclas. Permite-se desta forma a implementação de técnicas do tipo compactação semântica. Dado o número elevado de mensagens que podem ser guardadas desta forma a partir das teclas disponíveis, não é usada paginação, disponibilizando-se por tal mais uma tecla para acesso a mensagens.

Por razões de espaço de memória disponibilizada internamente pelo microcontrolador, o número máximo de mensagens que podem ser reproduzidas em sequência está limitada a dezasseis. Quando o utilizador pretender reproduzir numa única frase várias mensagens,

deve pressionar a tecla de concatenação, seguida das teclas das mensagens desejadas. No final da escolha deve voltar a pressionar a mesma tecla para dar início à reprodução. Durante a selecção das mensagens estas não são reproduzidas mas é gerado um tom de retorno que indica o sucesso da selecção. No caso da paginação podem ser escolhidas mensagens duma ou várias páginas para integrarem a frase.

Para além das funções disponibilizadas para o utilizador existem ainda mais três botões que permitem a configuração do sistema. O botão de gravação possibilita a gravação das mensagens. Para executar a gravação duma mensagem deve pressionar-se primeiramente este botão e posteriormente a tecla à qual se pretende que a mensagem fique associada. Esta deverá manter-se premida até final da gravação. No caso em que a selecção duma mensagem implica a escolha sucessiva de duas teclas, pressiona-se momentaneamente a primeira, o que originará a produção dum tom de retorno, devendo manter-se posteriormente a segunda premida até ao fim da gravação. Um LED vermelho mantém-se acesso enquanto decorrer a gravação. Caso este se apague antes da tecla ser libertada, é sinal de que a memória externa está totalmente ocupada.

Um segundo botão possibilita a reconfiguração do tamanho das células, permitindo agrupar as teclas em grupos de duas, quatro, oito ou dezasseis sobre uma mesma mensagem, conforme se ilustra na figura 5.10.

Por último existe ainda um botão que possibilita o apagar de todas as mensagens previamente gravadas.

5.2.2.2 *Interface por matriz luminosa*

No caso da interface por matriz luminosa, as teclas são substituídas por pequenos LEDs, passando a ser usadas técnicas de varrimento, sendo a selecção realizada através dum botão de pressão externo. Este sistema destina-se sobretudo a utilizadores com dificuldades motoras acentuadas, incapazes de utilizarem sistemas com acesso directo.

O varrimento implementado no protótipo é do tipo linha-coluna, descrito na secção 3.1. A velocidade de varrimento é ajustável por intermédio dum selector rotativo com saída binária cujo valor é lido pelo microcontrolador ao fim de cada ciclo de varrimento. O ajuste permite que a velocidade de varrimento seja aumentada à medida que o utilizador se adapta ao funcionamento do sistema.

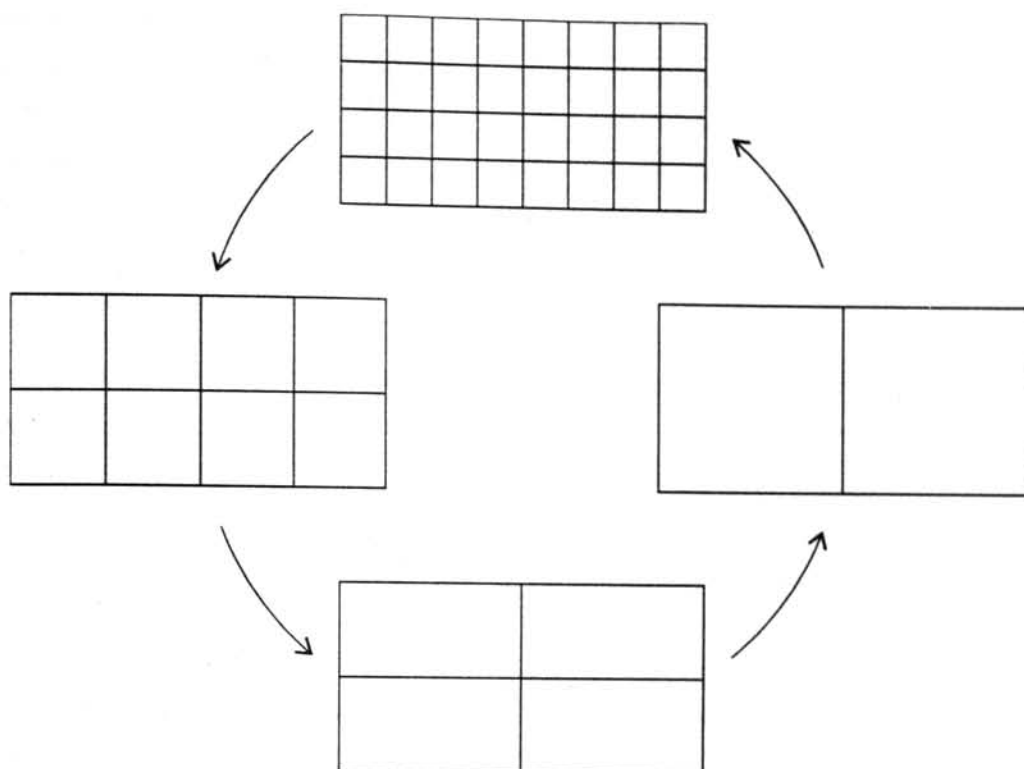


Figura 5.10 : Configurações possíveis do teclado.

Acessíveis ao utilizador através da mesma *interface* estão duas funções, a paginação e a concatenação, cujo princípio de funcionamento é idêntico ao descrito na secção 5.2.2.1 para a *interface* por teclado. Dois botões permitem a gravação e o apagamento de todas as mensagens. Para gravar uma mensagem prime-se momentaneamente o botão de gravação. O sistema inicia imediatamente o varrimento da matriz. Deve então escolher-se, por intermédio do botão de pressão externo, qual a célula sob a qual se pretende gravar a mensagem. No final da gravação deve premir-se de novo o botão de pressão. Um LED vermelho mantém-se aceso durante todo o tempo em que decorrer a gravação.

5.2.2.3 Interface numérica

A *interface* numérica permite a implementação de codificação numérica no acesso às mensagens. A cada mensagem corresponde um número entre 01 e 99, num total máximo de 99 mensagens. O código 00 corresponderá à função de concatenação, cujo princípio de funcionamento é idêntico ao dos dois casos vistos anteriormente. A selecção é efectuada através dum botão de pressão externo, tal como no caso da matriz luminosa.

Os códigos são apresentados ao utilizador por intermédio de dois dígitos de 7 segmentos. O sistema promove inicialmente o varrimento do dígito correspondente às dezenas, de forma crescente e rotativamente, começando no 0. Quando o algarismo pretendido aparece o utilizador deve pressionar o botão de pressão externo, fixando o seu valor, e dando início ao varrimento, da mesma forma, do dígito das unidades. Pressionando de novo o botão fixa-se o dígito das unidades e a mensagem correspondente ao código seleccionado é reproduzida.

O sistema determina automaticamente após cada gravação qual o mais alto dos códigos que contém uma mensagem gravada, sendo o varrimento efectuado apenas até esse código. Se por exemplo o sistema tiver 52 mensagens gravadas, sequencialmente do código 01 ao código 52, o sistema apenas varrerá no dígito das dezenas os algarismos entre 0 e 5, e, caso seja o 5 o seleccionado, apenas serão varridos nas unidades os números entre 0 e 2. Ao evitar-se que sejam varridos códigos que não contêm qualquer mensagem, diminui-se o tempo de cada ciclo de varrimento, aumentando-se a eficiência da selecção e consequentemente a fluência da comunicação. As mensagens devem ser gravadas em códigos sucessivos para que este esquema seja eficaz, caso contrário, se forem deixados "buracos", serão varridos códigos aos quais não corresponde nenhuma mensagem. A velocidade de varrimento é ajustável da mesma forma que no caso da matriz luminosa.

Quando se pretender concatenar mensagens escolhe-se inicialmente o código 00 seguido das mensagens desejadas, finalizando-se de novo com o código 00, o que dará início à reprodução da sequência.

Tal como no caso anterior, dois botões permitem a gravação e o apagamento das mensagens. Depois de premido o botão de gravação, selecciona-se a partir da mesma *interface* de utilização, à semelhança do que acontecia para a matriz luminosa, o código sob o qual se pretende que a mensagem fique gravada. No final da gravação deve premir-se de novo o botão de pressão. Tal como nos casos anteriores um LED vermelho estará acesso enquanto decorrer a gravação.

5.2.3 Considerações gerais

Após a descrição do funcionamento de cada uma das *interfaces* resta fazer algumas considerações gerais sobre o funcionamento e desempenho do sistema. Em todos os casos vistos anteriormente a gravação da mensagem pode ser realizada através do microfone

incorporado ou recorrendo a um externo ou a partir duma qualquer outra fonte de sinal. Igualmente é disponibilizado para o exterior uma saída de som caso se pretenda ligar o sistema a um amplificador externo. O altifalante incorporado proporciona um sinal audível em boas condições na maioria das situações de utilização.

O consumo do sistema difere naturalmente consoante a *interface* usada. A que apresenta menor consumo é a *interface* por teclado. As duas outras vêem o seu consumo aumentado devido quer à matriz luminosa quer aos dígitos empregues nas respectivas *interfaces* com o utilizador. Mesmo nestes casos o consumo não é no entanto exagerado, não comprometendo de forma alguma a opção pelo emprego de baterias na sua alimentação e consequentemente a portabilidade do sistema. O consumo pode ser reduzido se recorrermos a um processo de corte automático da alimentação decorrido algum tempo após a última acção sobre o sistema, voltando a estabelecer-se assim que nova acção seja detectada [FTR+94]. De salientar que o emprego de NVSRAMs para armazenamento das mensagens permite que se corte a alimentação ao sistema sem que estas se percam, reduzindo o consumo de energia e aumentando o tempo entre cada carga.

Em termos de dimensões o sistema com *interface* para codificação numérica é aquele que apresenta as menores dimensões. Os outros dois têm as suas dimensões subordinadas ao tamanho do quadro da *interface*. Com uma matriz de 4×8 e com as dimensões de 25×25 mm para cada célula, a dimensão do quadro situa-se nos 100×200 mm, ao que há que acrescentar um pequeno espaçamento entre células e um natural rebordo.

Por último, realce para o baixo custo do sistema, por diversas razões já apontadas ao longo deste capítulo, e que, não diferindo grandemente consoante as *interfaces*, se situa na meia centena de milhares de escudos, o que comparado com sistemas comerciais com características idênticas, lhe confere excelentes perspectivas de viabilidade.

6. Avaliação

Um trabalho deste tipo não estaria completo se, após a fase de implementação, não se procedesse à avaliação, no campo, dos protótipos desenvolvidos. Certamente que não é difícil perceber do bom ou mau acolhimento dos dispositivos junto da população a que se dirigem, mas, são melhores ou piores que os sistemas anteriormente usados? Se o utilizador não tinha anteriormente tido acesso a um dispositivo com saída de voz, não será difícil imaginar o seu entusiasmo face a um sistema deste tipo. No entanto é necessário verificar se o desempenho do conjunto homem-máquina é bom ou se poderá ser melhorado, se as metas inicialmente definidas foram atingidas, ou se deverão ser introduzidas alterações. Para tal é indispensável a existência duma fase de avaliação que possibilite aferir da validade do trabalho desenvolvido.

Um dos maiores problemas que se coloca é a dificuldade de estabelecimento de critérios objectivos de avaliação. A quantificação dum determinado comportamento não é tarefa simples pela ausência de referentes objectivos. No entanto, mesmo que com falhas, é sempre importante observar de que modo os novos sistemas influem na actividade e desenvolvimento dos seus utilizadores, até porque permite a detecção de falhas de concepção e implementação que devem ser corrigidas, antes de se passar a outras fases como a produção em larga escala.

É sobre esta fase de avaliação que versa este capítulo. Nas secções subsequentes serão expostos os critérios que presidiram à avaliação, os procedimentos seguidos e os resultados obtidos.

Alguns dos dados apresentados neste capítulo são provenientes dum projecto anterior referido em [FTR+94]. No entanto, e atendendo a que as diferenças se situam essencialmente no que respeita aos dois modos de *interface* adicionais proporcionados pelo “Fala-Barato II”, os elementos já reunidos são extensíveis ao projecto aqui descrito.

6.1 Critérios de avaliação

A primeira e mais evidente razão para despendar tempo e recursos na avaliação é o teste dos sistemas em condições reais de utilização documentando os resultados para que outros lhes possam ter acesso. Do ponto de vista da engenharia, a nossa atenção centrar-se-á no descobrir de falhas na concepção e implementação bem como nas possibilidades da sua correcção.

Dada a evidência doutros estudos, é de acreditar que os sistemas desenvolvidos no âmbito do projecto "Fala-Barato" se mostrarão eficazes e facilitarão o desenvolvimento comunicativo dos seus utilizadores. Mas claro, a nossa crença não é suficiente e não é com certeza muito informativa para outros potenciais interessados.

Está demonstrado que muitos sistemas e técnicas falham quando em condições reais de utilização [CNO+93]. As dificuldades surgem muitas vezes não de falhas na concepção técnica, mas porque a identificação de requisitos falhou na correcta avaliação das condições sociais, as quais suportam o uso efectivo da tecnologia. O sucesso ou insucesso individual é determinado por um conjunto complexo de factores que muitas das vezes não são evidentes numa fase de identificação de requisitos, só se revelando quando o utilizador é confrontado com a situação real do uso da tecnologia, já que a experiência leva frequentemente à descoberta de novos requisitos. Neste sentido haverá certamente aspectos a corrigir e melhorar.

De notar igualmente que uma rejeição da tecnologia pode não ser resultado da tecnologia em si, mas da forma como esta é apresentada ao utilizador, pelo que cabe aqui um papel importante, no caso dos sistemas de CAA, aos terapeutas da fala e aos seus colaboradores.

Há muitas maneiras de se definir um benefício ou ganho funcional. Duma forma mais concreta, pode significar que um indivíduo é agora capaz de fazer algo que queria fazer antes mas que não lhe era possível, ou que é capaz de fazer melhor que anteriormente. Não surpreendentemente, é muito difícil definir e medir um ganho funcional resultado do uso de uma nova técnica de comunicação [Tre94]. Se um utilizador dum sistema de CAA com uma *interface* baseada em codificação numérica é capaz de decorar 20 códigos em uma hora de treino e lembrar-se deles uma semana depois, será que se pode derivar daí alguma conclusão acerca do benefício funcional dessa técnica? Um dos problemas neste

tipo de avaliação é a ausência dum padrão de referência. Decorar 20 códigos não é em si um bom ou mau resultado. Quando comparado com outras técnicas de codificação, pode-se dizer se esta técnica é mais ou menos bem sucedida que a outra, pelo menos quanto à facilidade com que o utilizador apreende o código. Mas qual é a taxa de apreensão adequada para um determinado código? É a apreensão de 20 códigos boa, ou 50, ou 100? Inclusive a observação do desempenho numa situação controlada não nos diz se uma determinada técnica é a melhor para um particular indivíduo. Necessário será uma comparação com outras opções. Outra técnica ou combinação de técnicas pode conduzir ou não a resultados muito superiores. Muitas das vezes, em presença de uma nova técnica existe uma inegável tendência para subvalorizar técnicas concorrentes. As desvantagens das técnicas antigas são enfatizadas e as virtudes do novo método foco das atenções.

A selecção dum dispositivo de comunicação requer muito mais informação que aquela que resulta duma avaliação clínica do indivíduo e da literatura disponível sobre o dispositivo. Experiências controladas e a observação dos indivíduos em actividades diárias representativas, como a indicação duma direcção, a interacção com um parceiro para resolução dum problema, o pedir um favor a alguém ou o encomendar duma refeição são importantes para se perceber da real eficácia da ajuda. O sucesso seria definido mais em termos da qualidade e bom cumprimento das tarefas que da velocidade ou exactidão na produção da mensagem. Um dos aspectos resultantes duma comunicação eficiente é a criação duma imagem positiva sobre nós próprios, com o objectivo de arranjar um emprego, fazer amigos ou inspirar respeito. A necessidade de serem levados a sério e respeitados como membros plenos da sociedade é especialmente importante para as pessoas deficientes, as quais se encontram frequentemente confrontadas com um conjunto de preconceitos negativos enraizados ao longo de séculos de obscurantismo.

No sentido de se determinar se um dispositivo exerce um efeito benéfico sobre o utilizador, é importante seleccionar antecipadamente as variáveis a serem medidas. Quais as actividades que o dispositivo visa melhorar? Como eram essas actividades realizadas anteriormente? Quais as medidas de desempenho válidas para essas actividades, e quais os resultados obtidos quando essas medidas foram aplicadas às técnicas anteriormente usadas? Após um determinado período de treino e uso foram os resultados obtidos com o dispositivo superiores aos resultados obtidos com as anteriores técnicas, por exemplo em termos de fluência e percepção das mensagens que o utilizador pretendia veicular?

Uma medida objectiva do ganho funcional conseguido com uma determinada ajuda técnica seria a melhor forma de avaliar e comparar vários dispositivos. A frequência de utilização é sem dúvida um bom índice para aferição desse ganho, mas apenas de utilidade relativa em comparação com outras opções. A frequência de utilização é dependente da envolvente e da utilidade que o dispositivo apresenta em cada situação para essa pessoa.

A selecção das actividades é igualmente crítica. Essas actividades devem ser representativas da vontade e necessidades do utilizador. Os resultados podem ser enganosos se a amostra das actividades não for representativa dos interesses do indivíduo, e não significarem um benefício funcional real.

Uma visão global dos benefícios deveria ter em conta os ganhos alcançados pelo indivíduo em termos de factores como a qualidade de vida, a independência ou o desenvolvimento pessoal. Estes são, em última análise, os factores mais importantes.

Após um determinado período de treino é necessário analisar quais as variáveis sob observação em que se registaram melhorias e quais aquelas que permaneceram inalteradas.

A perspectiva do utilizador e do engenheiro que concebeu a ajuda têm sido tradicionalmente diferentes. O engenheiro centra a sua atenção sobretudo na tecnologia empregue e na interacção homem-máquina. O utilizador preocupa-se principalmente com a influência da ajuda na sua vida diária e em que medida esta vai ao encontro das suas necessidades. Inevitavelmente, e apesar de todos os esforços para tentar compreender a perspectiva do utilizador, muita informação escapará mesmo ao observador mais atento.

Apesar de todas as dificuldades e das incorrecções que se possam cometer, a realização duma avaliação é sempre importante. No caso do projecto "Fala-Barato" a avaliação foi realizada tendo como guia um conjunto de parâmetros pré-definidos, e sobre os quais se procurou aferir da evolução verificada em cada utilizador, comparativamente à sua situação anterior.

A avaliação decorre no Centro de Reabilitação de Paralisia Cerebral do Porto (CRPCP). Um grupo de utentes entre os 7 e os 15 anos foi seleccionado tendo em conta as suas características cognitivas e motoras. Integram o grupo sobretudo pessoas com dificuldades acentuadas na articulação das palavras, e em alguns casos de total incapacidade, devido a problemas motores que igualmente lhes limitam os movimentos

que são capazes de executar. Numa parte deles esses movimentos limitam-se a um movimento lateral da cabeça, único capaz de executarem com um mínimo de confiança e repetibilidade.

De acordo com os terapeutas do CRPCP, decidiu-se seleccionar o seguinte conjunto de parâmetros a observar durante o período de avaliação:

- a rapidez da aprendizagem;
- o número de ciclos de conversação por minuto, isto é, ciclos de interacção com o parceiro de comunicação;
- a atenção colocada na conversação;
- o desejo de comunicar;
- a razão à qual o vocabulário de cada indivíduo aumenta;
- as melhoras na estrutura comunicativa, ou seja, o grau de elaboração das frases que veicula;
- a fluência da comunicação;
- o desenvolvimento verificado na linguagem;
- o nível de interacção familiar e social;
- o grau de independência alcançado pelo indivíduo;
- a estimulação intelectual verificada;
- o desenvolvimento cognitivo;
- o aumento da auto-estima.

Como facilmente se pode verificar, a maior parte dos parâmetros são de avaliação subjectiva, e embora outros sejam passíveis de quantificação deve-se no entanto ter em conta a sua relativa validade por ausência de padrões de referência.

6.2 Procedimentos

Uma avaliação prévia individual realizada pelos terapeutas ocupacionais e da fala e por fisioterapeutas permite definir para cada caso uma estratégia de intervenção,

nomeadamente a selecção, de entre o leque de *interfaces* possíveis, daquela que melhor se adapta às características do utilizador, bem como se é necessário ou não o uso do "Fala-Barato I" numa fase introdutória, para iniciação da aprendizagem. Após esta fase, iniciam-se sessões diárias individuais de aprendizagem e adaptação do indivíduo ao sistema, avaliando-se desde logo a facilidade e rapidez com que essa aprendizagem é realizada. Outro ponto importante é o da determinação de qual o vocabulário que o indivíduo desejava ver acessível na ajuda, em função certamente das que ele considera serem as suas necessidades. Interessante é ainda observar em que medida é que a colocação à sua disposição dum maior vocabulário afecta o desenvolvimento da sua estrutura comunicativa.

Após a fase de aprendizagem individual e de comunicação entre terapeuta e utilizador, passa-se a uma fase de grupo. Vários utilizadores são reunidos numa sala tendo como objectivo a estimulação da conversação entre eles. Um primeiro ponto que se constata ter lugar é uma monopolização da conversação por parte daqueles que, devido a uma melhor capacidade motora, conseguem uma maior fluência na comunicação. Talvez devido ao entusiasmo resultante da novidade e de um desejo de comunicar contido ao longo de tantos anos, a conversação parece algo atabalhoada.

Os participantes do grupo com maiores dificuldades de coordenação motora demonstram inicialmente uma certa inibição causada pelo receio de falharem a selecção duma mensagem em presença dos restantes intervenientes. Na realidade, e quando o grupo é deixado sozinho, sendo o seu comportamento registado por intermédio duma câmara de vídeo, essa inibição praticamente desaparece, verificando-se um aumento do número de ciclos de comunicação.

Esta avaliação não estaria no entanto completa se não fosse avaliado o grau de contribuição dos sistemas para a integração do deficiente na família e na sociedade. Indispensável e preciosa a pronta disponibilidade demonstrada pelos familiares em colaborar nesta fase do projecto, comentando as alterações que verificam a nível do comportamento, quer em casa, quer quando os utilizadores se vêem confrontados com situações sociais externas, e expressando a sua própria opinião.

Por último, e sem dúvida da maior importância, a opinião do próprio utilizador acerca da nova ajuda.

6.3 Resultados obtidos

Do ponto de vista técnico, notou-se a necessidade de correcção de alguns aspectos, nomeadamente no sentido de aumentar a comodidade de utilização, a possibilidade de os dígitos na *interface* baseada em codificação numérica poderem ser colocados inclinados em relação à horizontal, permitindo uma melhor visualização, e alguns reparos quanto à posição das saídas para os botões de pressão externos.

Do ponto de vista dos utilizadores, os sistemas constituem um verdadeiro sucesso. Dos resultados obtidos é de salientar alguns aspectos positivos:

- um maior desejo de comunicar;
- uma maior estabilidade de comportamento;
- um maior interesse nas tarefas e participação no grupo.

A eficiência do sistema fica demonstrada pelo elevado número de ciclos de conversação e pela elevada participação dos utilizadores em conversação com pessoas com e sem dificuldades de comunicação.

Ao dispensar a interacção de terceiros, o uso destes sistemas com saída de voz permite uma maior independência, resultando numa interacção mais activa entre os seus utilizadores, notando-se um aumento acentuado do tempo dedicado à conversação, o que possibilita uma maior troca de ideias e conduz a desenvolvimentos mais rápidos das estruturas comunicativas e da própria linguagem.

Observa-se igualmente o desenvolvimento de relações inter-pessoais mais fortes e profundas, estreitando laços de amizade, sem dúvida resultado da possibilidade de comunicarem entre si com uma maior privacidade.

Um dos aspectos evidenciados pelos terapeutas é que o uso dum sistema com saída de voz possibilita que durante a conversação, após a selecção e enquanto decorre a reprodução da mensagem escolhida, os intervenientes estabeleçam entre si contacto visual, permitindo uma maior interacção e reacções expressivas mais ricas do que quando usam quadros de comunicação.

De salientar a opinião dos familiares, que testemunharam o entusiasmo sentido pelos utilizadores por poderem mais facilmente intervir em família, expressando as suas



necessidades e vontades.

Por último, a opinião dos próprios utilizadores. Mais do que traduzidas em palavras, as suas reacções constituíram o melhor reconhecimento do sucesso do trabalho realizado. No decorrer duma das sessões de avaliação, quando deixados sózinhos, um dos utilizadores envolvidos usou o protótipo para dizer "Que seca!". A maior facilidade de comunicação que lhe permitiu esta expressão da sua opinião, constitui sem dúvida um sucesso do ponto de vista da CAA.

7. Conclusão

A constante evolução que se verifica na área das telecomunicações, com o incremento da necessidade de transmissão de cada vez maiores volumes de informação, tem contribuído de forma decisiva para o surgir de técnicas de codificação cada vez mais eficientes. Do aproveitamento dessas técnicas podem os sistemas de CAA baseados em voz digitalizada retirar grandes benefícios. Não são no entanto de descurar os eventuais desenvolvimentos que certamente se verificarão nas técnicas de síntese de voz, pela maior versatilidade que estas podem aportar aos sistemas de CAA. Vital é portanto, para o desenvolvimento das ajudas técnicas à comunicação, que haja um constante acompanhamento das evoluções tecnológicas que se verifiquem nessas áreas. Importante igualmente o crescente aumento das densidades de integração conduzindo a circuitos mais complexos com áreas de ocupação mais pequenas e menores consumos energéticos, permitindo a ampliação do leque das funcionalidades disponibilizadas ao utilizador e contribuindo para uma simplificação da sua operação, sem penalizar a autonomia e consequentemente a portabilidade dos sistemas.

Outro aspecto a ter em conta no futuro é a necessidade de uma maior divulgação deste tipo de ajudas, por forma a que pessoas com problemas de comunicação temporários ou permanentes possam beneficiar destes sistemas. Seguindo as leis de mercado, um alargamento do número de utilizadores levará a uma maior diluição de custos de concepção, desenvolvimento e produção, baixando os preços e tornando os sistemas mais acessíveis.

7.1 Principais dificuldades encontradas

A concepção e desenvolvimento de sistemas de CAA exige uma colaboração multidisciplinar entre engenheiros e outros grupos profissionais como terapeutas ocupacionais e da fala, e a própria classe médica e de enfermagem. Para se conseguirem alcançar bons resultados é necessário seguir uma metodologia que comece por identificar os requisitos dos utilizadores, para que sobre esses dados se possa trabalhar na concepção dos sistemas. Este foi um dos problemas que se levantou no decorrer deste projecto. A escassez de oportunidades para se desenvolver uma plataforma de entendimento comum relativamente às oportunidades proporcionadas pelas novas tecnologias, bem como a necessidade do estabelecimento de uma terminologia que possibilitasse uma melhor troca de informação entre profissionais de áreas tão distintas como a engenharia e a reabilitação, entre outras, foram factores que requereram uma atenção especial e um envolvimento conjunto desde o início deste trabalho.

Um trabalho deste tipo exige, como demonstrado no capítulo 6, a realização de uma experimentação alargada dos sistemas desenvolvidos, para se poder aferir do seu desempenho junto dos utilizadores-alvo. O pouco tempo disponível prejudicou sem dúvida esta fase e as conclusões apresentadas, bem como a realimentação que daí se poderia recolher para melhoria dos protótipos.

Um outro problema que se mostra de difícil resolução é a passagem dos protótipos de laboratório para a produção em massa. Não basta desenvolver os sistemas, é preciso produzi-los, tornando os frutos deste trabalho acessíveis aos que dele podem colher benefícios.

7.2 Perspectivas de trabalho futuro

Na sequência do trabalho desenvolvido estabeleceram-se já contactos, quer a nível nacional quer internacional, que abrem boas perspectivas para a concretização de novos projectos de investigação e desenvolvimento nesta área, nomeadamente a colaboração com terapeutas da fala ligados ao acompanhamento de pessoas com afasia em ambiente hospitalar. Esperemos que esta colaboração possa abrir perspectivas que conduzam ao uso

generalizado de sistemas de CAA em ambientes hospitalares, o que contribuiria certamente para uma diminuição da angústia e frustração experimentadas pelos doentes privados da fala na sequência de actos médico-cirúrgicos, e atenuaria o choque causado pela sua perda devido a ataques cerebrais ou acidentes.

De referir aqui a existência, no plano de trabalho para as novas propostas no âmbito do programa europeu TIDE (*Technology Initiative for Disabled and Elderly People*), de duas tarefas na área da CAA, cujo objectivo é o desenvolvimento, adaptação e integração de tecnologias que facilitem a comunicação inter-pessoal de todos os tipos, desde a conversação frente-a-frente à comunicação à distância. Um dos pontos que se pretende ver contemplados nas propostas é exactamente o da concepção de sistemas flexíveis e com potencial de crescimento, partindo de uma abordagem modular. Outro dos pontos salientados é o da criação de sistemas portáteis [Eur95].

Importante é sem dúvida a continuação deste trabalho, explorando novas áreas de intervenção, aperfeiçoando métodos e técnicas, tentando responder a novos desafios, contribuindo para a melhoria das condições de vida de todos aqueles que por variados motivos ou situações se vêm confrontados com problemas de comunicação.

A CAA tornou possível que muitas pessoas com deficiências comunicativas aspirassem a uma melhor qualidade de vida, que de outra forma estaria seriamente comprometida.

Bibliografia

- [ANA93] Norman Alm, Alan F. Newell, John L. Arnott. An Integrated Research Strategy for Improving Communication Systems for Severely Physically Impaired Non-speaking People. In *Rehabilitation Technology - Strategies for the European Union, Proceedings of the 1st TIDE Congress*, pp. 249-253, IOS Press, April 1993.
- [Bak94] Bruce R. Baker. Translating an Augmentative Communication Language System. In *Tapping Technology*, pp. 14-15, June 1994.
- [Bar84] M. del Carmen Busto Barcos. *Reeducacion del Habla y del Lenguaje en el Parálítico Cerebral*. Ciencias de la Educacion Preescolar y Especial, S.A., 1984.
- [BBS94] Zitta Bülman, Åse Brandt, Birgit Seelen. Why do Hospitals Fail to Use Technological Communication Aids for Severely Motor-Impaired, Nonspeaking Patients?. In *Proceedings of the 6th ISAAC Conference*, pp. 327-328, October 1994.
- [BF93] Serenella Besio, Lucia Ferlino. Blissymbolics Software Worldwide: from Prototypes towards Future Optimized Products. In *Proceedings of the 2nd ECART Conference*, 1993.
- [Bla91] Sarah W. Blackstone. Persons with Severe Aphasia: What does AAC have to Offer?. In *Augmentative Communication News*, pp. 1-3, Vol. 4, N° 1, January 1991.
- [BM92] D. R. Beukelman, P. Mirenda. *Augmentative and Alternative Communication*. Paul H. Brookes Publishing Co., Inc., 1992.

- [BSB93] Åse Brandt, Birgit Seelen, Zitta Büllman. Why Don't Hospitals Use Electronic Communication Aids for Patients who are Severely Motor Impaired and who Can't Speak?. In *Proceedings of the 2nd ECART Conference*, 1993.
- [Cap85] V. Cappellini, eds. *Data Compression and Error Control Techniques with Applications*. Academic Press Inc. Ltd., 1985
- [CCI84] CCITT Recommendation G.721. 32 kbit/s Adaptive Differential Pulse Code Modulation (ADPCM). 1984.
- [CCI90] CCITT Recommendation G.726. 40-, 32-, 24-, and 16 kbit/s Adaptive Differential Pulse Code Modulation (ADPCM). 1990.
- [CHA+95] P. A. Cudd, S. Hunnicutt, J. Arthur, B. Granström, S. Aguilera, B. Waernulf, P. Dalsgaard, G. Wilson. Voices, Attitudes and Emotions in Speech Synthesis. In *Proceedings of the 2nd TIDE Congress*, pp. 344-347, IOS Press, April 1995.
- [CNO+93] I. Craig, P. Nisbet, P. Odor, M. Watson. Evaluation Methodologies for Rehabilitation Technology. In *Rehabilitation Technology - Strategies for the European Union, Proceedings of the 1st TIDE Congress*, pp. 238-243, IOS Press, April 1993.
- [Cor91] Toshiba Corporation. Speech LSIs Products Guide, 1991.
- [CSR94] A.Y. Cairns, W. D. Smart, I. W. Ricketts. Alternative Access to Assistive Devices for the Physically Impaired. In *Proceedings of the 6th ISAAC Conference*, pp. 449-451, October 1994.
- [Dem94] Patrick Demasco. Human Factors Considerations in the Design of Language Interfaces in AAC. In *Assistive Technology*, Vol. 6.1, pp. 10-25. RESNA Press, 1994.
- [Dim93] Spiros Dimolitsas. Standardizing Speech-Coding Technology for Network Applications. In *IEEE Communications Magazine*, pp. 26-33, November 1993.
- [Edw91] Alistair D. N. Edwards. *Speech Synthesis: Technology for Disabled People*. Paul Chapman Publishing Ltd., 1991.

- [Edw94] Alistair D. N. Edwards. Speech Synthesis. *Information Technology and Disabilities*, Vol. 1, N° 2, April 1994 (jornal editado sob a forma electrónica disponível através do endereço: <gopher://sjumusic.stjohns.edu:70/11/disabled/easi/easijrnl>).
- [Eur95] European Commission. *Background to the Workplan for the IVth Framework Programme Telematics for the Integration of Disabled and Elderly people (TIDE)*, March 1995.
- [FD93] Prue Fuller, Mick Donegan. The Role of Specialist AAC Centres: Doers or Enablers?. In *Proceedings of the 2nd ECART Conference*, 1993.
- [FGL95] J. Martins Ferreira, M. Gradim Gericota, M. L. Lourenço. The Cost×Performance Tradeoff for Voice Output AAC Aids: Is it Real? In *Proceedings of the 18th Annual Resna Conference*, June 1995.
- [FTR+94] J. Martins Ferreira, A. C. Teixeira, J. L. Ramalho, M. L. Lourenço. A Portable Device to Improve the Communication Ability of People with Cerebral Palsy. In *Proceedings of the 6th ISAAC Conference*, pp. 446-448, October 1994.
- [Gal94] John Gallant. Speech-synthesis and -recognition chips personalize consumer products. In *EDN Europe*, pp. 51-57, October 20, 1994.
- [GKP+93] Franz Guenther, Karin Krüger-Thielmann, Robert Pasero, Paul Sabatier. Communication Aids for Handicapped Persons. In *Proceedings of the 2nd ECART Conference*, 1993.
- [GME93] Eileen Grist, Jane Mortley, Jayne Easton. Communication Opportunities for Global Dysphasics. In *Human Communication*, August 1993.
- [GÖ93] Karin Guttman, Bräcke Östergård. AAC-Talking to Family and Strangers. In *Proceedings of the 2nd ECART Conference*, 1993.
- [Goo91] Frank Goodenough. IC Holds 16 Seconds of Audio without Power. In *Electronic Design*, January 31, 1991.

- [Gra94] José Paulo Graça. As Novas Tecnologias no Apoio à Comunicação para Pessoas com Necessidades Especiais. II Conferência "Comunicações para Populações com Necessidades Especiais", Lisboa, Dezembro 1994.
- [Gray94] Caroline Gray. Switch Control for the Young Person. In *Proceedings of the 6th ISAAC Conference*, pp. 444-445, October 1994.
- [Hek94] Ir A. D. Hekstra. Advantages and Disadvantages of Highly Flexible AAC-Software, Focused on the BlissProcessor. In *Proceedings of the 6th ISAAC Conference*, pp. 102-104, October 1994.
- [HP93] Marianne Hickey, Colin Page. Polyvox: Flexible Message Selection in a Communication Prosthesis for Non-speakers. In *Proceedings of the 2nd ECART Conference*, 1993.
- [HTN94] Mikael Heimann, Tomas Tjus, Keith E. Nelson. Delta/Messages: A Rewarding, Fun, and Motivating Multimedia Computer Program for Teaching Reading Skills to Children with Language Delays. In *Proceedings of the 6th ISAAC Conference*, pp. 404-406, October 1994.
- [Hun94] Sheri Hunnicutt. Development and Evaluation of Communication Aids Supported by Speech Technology: The AAC Group, Stockholm. In *Tapping Technology*, pp. 20-21, June 1994.
- [ISD94] Application Notes and Design Manual for ISD's Single-Chip Voice Record/Playback Devices. Information Storage Devices, Inc., July 1994.
- [Jay90] N.S. Jayant. High-Quality Coding of Telephone Speech and Wideband Audio. In *IEEE Communications Magazine*, pp. 10-20, January 1990.
- [JC93] Nikil Jayant, Richard V. Cox. Speech Processing. In *IEEE Communications Magazine*, pp. 24-25, November 1993.
- [Kel94] Eric Keller. *Fundamentals of Speech Synthesis and Speech Recognition*. John Wiley & Sons Ltd., 1994.

- [KS94] Yvonne Karlsson, Kersti Samuelsson. Augmentative Communication Systems in Intensive Care. In *Proceedings of the 6th ISAAC Conference*, pp. 517-518, October 1994.
- [Kuc88] Roman Kuc. *Introduction to Digital Signal Processing*. McGraw-Hill International Editions, 1988.
- [LB92] Lyle L. Lloyd, Doreen M. Blischak. AAC Terminology Policy and Issues Update. In *AAC Augmentative and Alternative Communication*, Vol. 8, pp. 104-109, June 1992.
- [Lev94] Willem J. M. Levelt. What can a Theory of Normal Speaking Contribute to AAC?. In *Proceedings of the 6th ISAAC Conference*, pp. 18-20, October 1994.
- [Lig93] Janice Light. A Vision of Simplicity for Augmentative Communication Systems. In *AAC Augmentative and Alternative Communication*, Vol. 7, N° 2(20), pp. 9-11, August 1993.
- [LOT93] S. E. Levinson, J. P. Olive, J. S. Tschirgi. Speech Synthesis in Telecommunications. In *IEEE Communications Magazine*, pp. 46-53, November 1993.
- [LS93] Mats Lundälv, Dag Svanæs. Comspec-Towards a Modular Software Architecture and Protocol for AAC Devices. In *Rehabilitation Technology - Strategies for the European Union, Proceedings of the 1st TIDE Congress*, pp. 55-59, IOS Press, April 1993.
- [MB94] Jón Hjaltalín Magnússon, Stefán Briem. Application of IsBliss and BlissGrammar Symbolic Processing Programs. International Conference "Beyond Normalization Towards One Society for All", Reikjavik, June 1994.
- [ME93] Craig Marven, Gillian Ewers. *A simple approach to Digital Signal Processing*. Texas Instruments, 1993.
- [Mim91] Forrest M. Mims III. A Tape Recorder Chip. In *Computer Craft Magazine*, September 1991.

- [MJr76] J. D. Markel, A. H. Gray, Jr.. *Linear Prediction of Speech*. Springer-Verlag, 1976.
- [MMP+94] K. F. McCoy, W. M. McKnitt, D. M. Peischl, C. A. Pennington, P. B. Vanderheyden, P. W. Demasco. AAC-User Therapist Interactions: Preliminary Linguistic Observations and Implications for Compansion. In *Proceedings of the 17th Annual Resna Conference*, pp. 129-131, June 1994.
- [Mur94] Joan Murphy. Advantages and Disadvantages of AAC Systems: Perceptions of Users and Partners. In *Proceedings of the 6th ISAAC Conference*, pp. 504-505, October 1994.
- [NL93] Gillian Nelms, Andrew Lysley. A Communication System for David. In *Proceedings of the 2nd ECART Conference*, 1993.
- [Nol93] Peter Noll. Wideband Speech and Audio Coding. *IEEE Communications Magazine*, pp. 34-44, November 1993.
- [Owe82] Frank F. E. Owen. *PCM and Digital Transmission Systems*. McGraw-Hill Book Company, 1982.
- [Pap94] Marina-Elvira Papangelou. Trying to Communate Desperation leds to Augmentative Communication. In *Proceedings of the 6th ISAAC Conference*, pp. 87-88, October 1994.
- [PS94] Patricia Politano, Janice E. Scharre. Functional Vision Skills for AAC Systems. In *Proceedings of the 6th ISAAC Conference*, pp. 61-63, October 1994.
- [RHF93] David M. Roy, William Harwin, Robert Fawcus. The Enhancement of Computer Access for People with Cerebral Palsy through the Computer Recognition of Imprecise Gestures. In *Proceedings of the 2nd ECART Conference*, 1993.

- [RR93] Parimala Raghavendra, Elisabet Rosengren. Effectiveness of Multi-Talk II as a Voice Output Communication Aid. In *Rehabilitation Technology - Strategies for the European Union, Proceedings of the 1st TIDE Congress*, pp. 50-54, IOS Press, April 1993.
- [RRo93] Parimala Raghavendra, Elisabet Rosengren. Evaluation of Multi-Talk II: Initial Communicative Interaction Experiences of an AAC User. In *Proceedings of the 2nd ECART Conference*, 1993.
- [Sag90] Yoshinori Sagisaka. Speech Synthesis from Text. *IEEE Communications Magazine*, pp. 35-41, November 1993.
- [Sch93] Marcia J. Scherer. *Living in the State of Stuck: How Technology Impacts the Lives of People with Disabilities*. Brookline Books, 1993.
- [SK94] Lynn A. Sweeney, Clifford A. Kushler. Development, Programming and Concealment of Private Messages in AAC Systems. In *Proceedings of the 17th Annual Resna Conference*, pp. 109-111, June 1994.
- [SKR+94] Ralf W. Schlosser, Rajinder Koul, Parimala Raghavendra, Lyle L. Lloyd. State-of-the-Art in Blissymbolics Research: Implications for Practice. In *Proceedings of the 6th ISAAC Conference*, pp. 121-123, October 1994.
- [SOH93] Osamu Sueda, Masami Oguri, Miwako Hanabusa. Stability and Instability of Dysarthric Sound of Cerebral Palsy Children. In *Proceedings of the 2nd ECART Conference*, 1993.
- [Spa94] Andreas S. Spanias. Speech Coding: A Tutorial Review. In *Proceedings of the IEEE*, Vol. 82, N° 10, pp. 1541-1582, October 1994.
- [Spae94] A. Spaepen. Can We Avoid Putting People with an Impairment at a Disadvantage with New Technology in Daily Life: A Preventative Control for Future Research. In *Proceedings of the 6th ISAAC Conference*, pp. 26-27, October 1994.
- [SS91] DS2271 Speech Stik Data Sheet. *Dallas Semiconductor Teleservicing Design Handbook*, pp. 212-232, July 1991.

- [TAE+94] John Todman, Norman Alm, Leona Elder, Portia File. Talk: A Computer-Based Conversation Aid Using Pre-stored Text. In *Proceedings of the 6th ISAAC Conference*, pp. 490-492, October 1994.
- [Tat93] G. Van Tatenhove. *What is Minspeak™?*. Prentke Romich Company, 1993.
- [Tre94] Cheryl Goodenough-Trepagnier. Design Goals for Augmentative Communication. In *Assistive Technology*, Vol. 6.1, pp. 3-9. RESNA Press, 1994.
- [VMP91] DS2130Q Voice Messaging Processor Data Sheet. *Dallas Semiconductor Teleservicing Design Handbook*, pp. 190-211, July 1991.
- [VPP+94] P. B. Vanderheyden, C. A. Pennington, D. M. Peischl, W. M. McKnitt, K. F. McCoy, P. W. Demasco, H. van Balkom, H. Kamphuis. Developing AAC Systems that Model Intelligent Partner Interactions: Methodological Considerations. In *Proceedings of the 17th Annual Resna Conference*, pp. 126-128, June 1994.
- [Wol94] Walter S. Woltosz. *Dynamic Displays: the Changing Face of Augmentative Communication*. Words+, Inc., 1994.

