

CORDIAL-SIN

Corpus Dialectal para o Estudo da Sintaxe
(PRAXIS/P/PLP/113046/1998; POSI/1999/PLP/33275 e POCTI/LIN/46980/2002)

Normas de Transcrição

**Catarina Magro (org.)
Ernestina Carrilho
André Eliseu
Maria Lobo
Ana Maria Martins
Sandra Pereira**

Julho • 2007

A. Versão Conservadora	3
1. Cabeçalho.....	3
2. Procedimentos de identificação dos interlocutores (na transcrição).....	4
3. Marcação de sobreposições de produção	4
4. Transcrição de numerais	5
5. Capitalização.....	5
6. Pontuação	6
7. Representação ortográfica de palavras compostas.....	7
8. Transcrição de variantes fonéticas e morfofonológicas.....	7
8.1. Variantes fonéticas e morfofonológicas: código "PH" (<i>phonetic</i>).....	8
8.2. Formas contraídas: código "CT" (<i>contracted</i>).....	8
8.3. Formas com a parte inicial truncada: código "IP" (<i>initial partial</i>)	9
9. Transcrição de palavras estrangeiras: código "FR" (<i>foreigner</i>).....	10
10. Marcação, na transcrição, de disfluências na produção do discurso.....	10
10.1. Formas ou sequências abandonadas: código "AB" (<i>abandoned</i>)	10
10.2. Formas incompletas: código "RC" (<i>reconstructed</i>).....	11
10.3. Formas e sequências repetidas: código "RP" (<i>repeated</i>)	11
10.4. Pausas não preenchidas com material sonoro – {pp}	12
10.5. Pausas preenchidas com material sonoro – {fp}.....	13
11. Problemas de transcrição decorrentes de percepção deficiente	13
11.1. Dúvidas de audição – (xxx)	13
11.2. Divergências de audição – (xxx) /yyy\	14
11.3. Sequências imperceptíveis — (...)	14
12. Informação extra-linguística	15
13. Anonimização dos intervenientes nos eventos registados nos textos transcritos	16
Apêndice I — Lista de Assuntos	17
Apêndice II — Transcrição Ortográfica de Onomatopeias Não Dicionarizadas.....	18
B. Versão Normalizada.....	19
1. Formato word/pdf	19
2. Formato ASCII.....	21
Bibliografia	25

A. Versão Conservadora

1. Cabeçalho

Cada excerto da transcrição é introduzido por um cabeçalho que contém o código de identificação do ficheiro, informação relativa à proveniência geográfica do texto, à caracterização dos informantes, à identidade dos inquiridores, à identificação do inquérito no arquivo sonoro do Grupo de Estudos de Dialectologia do CLUL, à localização do texto transcrito no arquivo sonoro do CORDIAL-SIN, à identidade dos transcritores e à data da transcrição. Os campos relativos à caracterização dos informantes nem sempre podem ser preenchidos na sua totalidade dada a existência de lacunas de informação na fonte. Por outro lado, a identidade dos informantes é protegida através da atribuição aleatória de um nome próprio fictício a cada informante (cf. Secção A. 13). O campo relativo ao *assunto* de cada unidade transcrita é preenchido de acordo com uma lista de assuntos (cf. Apêndice I) estabelecida com base no índice do questionário linguístico do projecto Atlas Linguístico-Etnográfico de Portugal e da Galiza, desenvolvido pelo Grupo de Estudos de Dialectologia do CLUL (Gottschalk et alii 1974). Excepcionalmente, o campo *assunto* pode ser preenchido com a expressão "não aplicável", nos casos pontuais de excertos de curta dimensão ou carácter generalista. O formato do cabeçalho é o que se apresenta a seguir.

Código de identificação do ficheiro:	
Localidade: Distrito:	Concelho: Data:
Informante1: Idade:	Sexo: Escolaridade:
Informante2: Idade:	Sexo: Escolaridade:
Informante3: Idade:	Sexo: Escolaridade:
Fonte: Inquiridor1: Cassete nº: lado: min:	Inquiridor2:
Assunto:	
Tipo de transcrição: Autor da primeira transcrição: Autor da revisão final: CD nº: faixa:	Data da primeira transcrição: Data da revisão final:

2. Procedimentos de identificação dos interlocutores (na transcrição)

As alternâncias de intervenção dos interlocutores no diálogo correspondem, na transcrição, a mudanças de parágrafo. Cada tomada de palavra inicia, pois, um parágrafo e é precedida por "INQ" ou "INF", conforme se trate de um enunciado produzido, respectivamente, pelo inquiridor ou pelo informante. Nos casos em que existe mais do que um inquiridor e/ou informante, os elementos "INF" ou "INQ" são seguidos de um algarismo que indica a ordem de intervenção no diálogo por parte de cada locutor. Não existe qualquer sinal de pontuação entre estes indicadores e o início dos enunciados. Os enunciados produzidos pelos inquiridores são transcritos em itálico.

Exemplo:

INQ1 Sim, mas depois também tem... tem um buraco no outro lado, quando é para, para remar...

INF1 Sim, que é por causa do remo {PH|nũ=não} andar {fp} a dançar {fp}. Assim fica sempre com (uma e outra), (...).

INQ1 Pois, fica, entra nos dois, pois...

INF2 O senhor já viu, o senhor já viu o que isto mata?

INQ1 Como é que chama a este?

INF2 Faneca.

INF1 É fodãozinho.

INQ2 A faneca...

(VPA02-C)

3. Marcação de sobreposições de produção

As situações de fala sobreposta são assinaladas sublinhando os fragmentos produzidos em simultâneo. A indicação de sobreposição é marcada de forma tão precisa quanto possível; no entanto, se o excerto de fala sobreposta tiver início a meio de uma palavra e acabar a meio de outra, por exemplo, o sublinhado é marcado a partir do princípio da primeira palavra até ao fim da última.

Exemplo:

INQ Quem é que fazia as rocas, eram as mulheres ou eram os homens, os rapazes?

INF1 As mulheres. {pp} as mulheres.

INF2 E os homens. [AB|E os ho-]

(CTL02-C)

Os sublinhados, juntamente com a marcação de pausas (cf. Secção A. 9.4 e 9.5) permitem reconstituir a cronologia das produções de diferentes informantes nos casos em que a transcrição ortográfica, necessariamente sequencial, retrata menos fielmente essa cronologia.

4. Transcrição de numerais

A utilização de algarismos é reservada à transcrição de datas. Todos os outros numerais são transcritos por extenso.

Exemplo:

INF Eu fui {CT|pa=para a} pesca em **45**, quando acabou [AB|a guerra de] a guerra. Quando acabou a guerra {pp}, fui eu {CT|pa=para a} pesca do bacalhau, tinha eu **vinte e cinco** anos.

(VPA04-C)

5. Capitalização

A transcrição ortográfica é sempre feita em minúsculas, reservando-se as maiúsculas para os seguintes casos:

- (i) a representação de siglas e acrónimos;
- (ii) a primeira letra de:
 - antropónimos e topónimos,
 - nomes dos meses e estações do ano,
 - nomes de instituições e de sectores ou órgãos das mesmas,
 - títulos de livros, publicações periódicas e produções artísticas de qualquer género,
 - formas de tratamento institucionais e de cortesia;
- (iii) o início de frase – em início de tomada de palavra ou depois de sinal de pontuação forte, i. e. ponto final, ponto de exclamação, ponto de interrogação, reticências.

Exemplo:

INF Salmão! Ai, ajudei a caçar muitos! [AB|Nesse... No dia treze de Maio deste]
No dia treze de Maio deste mês, da Nossa Senhora de Fátima, caçámos nós um
salmão e três 'sáveis'. Um salmão de dez quilos. Palavra de honra, então!
(VPA08-C)

Note-se que em início de tomada de palavra a maiúscula não ocorre nos casos em que falas dos inquiridores interrompam frases completas dos informantes.

Exemplo:

INF1 Eu já {PH|li=lhe} disse a você {pp} – {PH|nũ=não} ({PH|li=lhe}) disse já?
– {pp}
INQ2 O quê?
INF1 que no bacalhau (que) caçámos duas {pp}
INQ2 Ah! Sim, sim.
INQ1 Já disse.
INF1 como daqui a acolá.
(VPA32-C)

6. Pontuação

Os sinais de pontuação são utilizados de acordo com as normas de pontuação na ortografia corrente, sem se procurar estabelecer (sistematicamente) equivalências entre estes e as pausas do discurso oral. A pontuação marca fronteiras de frases e fronteiras de constituintes, sendo por isso os sinais de pontuação indicadores importantes para a análise e anotação sintácticas. São utilizados os seguintes sinais:

- . (ponto)
- , (vírgula) ¹
- ; (ponto e vírgula)
- ? (ponto de interrogação)
- ! (ponto de exclamação)
- ... (reticências)

¹ Existem casos em que a vírgula é colocada entre o sujeito e o predicado. Trata-se, no entanto, de casos em que o sujeito se encontra topicalizado, sendo a sua natureza de tópico marcado prosodicamente identificável.

- : (dois pontos)
- " " (aspas)
- (travessões)²

7. Representação ortográfica de palavras compostas

As palavras compostas são transcritas com hífen em todos os casos. A inexistência de uniformidade entre os dicionários e vocabulários consultados na utilização de hífen nos compostos (cf. por exemplo, MORAIS [1789] 1948 e GONÇALVES 1966) e a ocorrência no *corpus* de compostos não dicionarizados (por exemplo, *caranguejo-cão*, *par-de-gatas*, *fecho-pedreiro*, etc.) aconselham a adoção sistemática do uso de hífen na transcrição destas palavras.

8. Transcrição de variantes fonéticas e morfofonológicas

{PH/CT/IP|transcrição fonética=forma ortográfica}

A transcrição é feita de acordo com a ortografia oficial portuguesa – excepto nos casos explicitados em 7, 8 (final) e 9 da secção A e no Apêndice II – sem preocupação de reproduzir realizações regionais ou individuais (decorrentes da existência de diferentes sistemas fonológicos dialectais, da ocorrência de fenómenos regulares de monotongação, ditongação, alteração de timbre de vogais, etc., bem como de fenómenos esporádicos mas tipologicamente predizíveis).

Exemplos:

- [ˈtʃujβɐ] é transcrito como "chuva";
- [simẽˈterɐ] é transcrito como "sementeira";
- [ˈkwɐɾpu] é transcrito como "corpo";
- [ʒɐˈzuʃ] é transcrito como "Jesus".

² O travessão emprega-se exclusivamente para delimitar sequências parentéticas, intercaladas (travessão duplo) ou adicionados no final da frase (travessão simples). Não se utiliza o travessão para indicar, no diálogo, o início das falas dos interlocutores.

No entanto, existem casos particulares em que pode ser relevante, para futura análise morfológica e/ou sintáctica, representar fielmente as realizações dos informantes. Para dar conta destes casos, utiliza-se um sistema duplo que combina transcrição fonética e transcrição ortográfica e que obedece à seguinte estrutura: {código identificador do tipo de fenómeno|forma fonética=forma gráfica}.

Deve esclarecer-se que, no que diz respeito aos enunciados dos inquiridores, estes são sempre transcritos apenas segundo a ortografia oficial ainda que apresentem realizações particulares. As convenções estabelecidas nesta secção, bem como nas secções 9, 10 e 11 abaixo, destinam-se, portanto, exclusivamente à transcrição dos enunciados dos informantes.

8.1. Variantes fonéticas e morfofonológicas: código "PH" (*phonetic*)

Exemplos:

- realização de "não" como [nũ] ou [nɐ]: {PH|nũ=não}, {PH|nɐ=não};
- realização visível ou não e realizações não *standard* de determinados morfemas flexionais: {PH|taj=tais}, {PH|ʃegĩ=chegam};
- realização do pronome "lhe" como [li]: deu-{PH|li=lhe};
- realização de segmentos de ligação entre palavras: {PH|'puɲẽw̃nɐ'agwɐ=punham a água}, {PH|'tẽẽnuʒ=têm os};
- realizações não *standard* da ligação entre verbo e pronome exclítico: {PH|ʃ'timulɐ=estimo-a}, {PH|'fizjɐ=fi-la}.

8.2. Formas contraídas: código "CT" (*contracted*)

Exemplos:

- {CT|pra=para a}; {CT|prɔ=para o};
- {CT|ɔ=a o};

- {CT|kwɐ=com a};
- {CT|nɛ=não é};
- {CT|vɔmi'se=vocemecê}.

8.3. Formas com a parte inicial truncada: código "IP" (*initial partial*)

Exemplos:

- {IP|ta=está}, {IP|'tivi=estive};
- {IP|'pɛrɐ=espera}.

Estas transcrições fonéticas são eliminadas na Versão Normalizada da edição dos textos (cf. Secção B), conservando-se apenas a transcrição ortográfica oficial correspondente. Existem, contudo, casos pontuais de realizações particulares dos informantes, sem grafia oficial prevista, cujo registo se quer preservar em todas as versões dos textos.

Estas formas são transcritas apenas ortograficamente, entre aspas simples.

Exemplos:

- (i) variantes morfológicas resultantes de processos de regularização analógica³

Exemplos:

- 'ouvisto'
- 'prefirava'
- 'fizestes'
- 'diferentas'
- 'sáveles' (por sáveis)

- (ii) itens lexicais não dicionarizados

Exemplo:

- 'lubis' (por lobisomem)

³ Não são transcritas entre aspas simples as variantes morfológicas resultantes de processos de derivação ou flexão produtivos. Exemplos: lavagatinho; grandezinho; grãozinhas; compradeira, estandem (gerúndio flexionado).

- (iii) formas de participio presente (não reanalisadas como Nomes ou Adjectivos)

Exemplo:

INF [AB|Era] Olhe {fp}, [AB|'acabante' se] 'acabante' se [AB|pa-] paria, já estava a galinha a cozer. Já a mulher estava [AB|para] para parir, [AB|já] já estava a galinha a cozer no pote.

(CTL32-C)

9. Transcrição de palavras estrangeiras: código "FR" (*foreigner*)

{FR|transcrição fonética=forma ortográfica}

As palavras estrangeiras são transcritas segundo o mesmo sistema duplo que combina transcrição fonética e transcrição ortográfica. A forma gráfica é apresentada segundo a ortografia oficial da língua em questão, entre aspas simples. O código identificador para as formas estrangeiras é "FR" (*foreigner*).

Exemplos:

INF Mas naquele tempo que nós fomos {pp} o gelo dava por aqui. O {FR|si'no='snow'}. Como caiu aqui há dias. Vocês, {PH|nũ=não} caiu lá em Lisboa?

(VPA43-C)

INF (E dizia-se-lhe) /Era dizer-se-lhe\ : "Eu vou {CT|prɔ=para o} médico para onde a Dom Albano". "Passe. Passe {FR|uʃ'te='usted'}, passe {FR|uʃ'te='usted'}" – os de lá.

(CTL15-C)

Nos casos em que apenas seja possível identificar a categoria morfológica da palavra estrangeira, utiliza-se a estratégia descrita em 11.3.

10. Marcação, na transcrição, de disfluências na produção do discurso

10.1. Formas ou sequências abandonadas: código "AB" (*abandoned*)

[AB|forma ortográfica dos segmentos produzidos-]

As formas ou sequências não concluídas, em contexto de reformulação do discurso, são transcritas entre parênteses rectos e identificadas através do código "AB" (*abandoned*). Nos casos em que uma sequência abandonada termina a meio de uma

palavra, transcrevem-se ortograficamente apenas os segmentos efectivamente produzidos, seguidos de hífen.

Exemplos:

INF É a única coisa que eu vejo [AB|que {pp} co-] que faz mal aos coelhos, é isto.

(AAL01-C)

INF [AB|Se {PH|nũ=não} aca-] Se o nosso governo {PH|nũ=não} acaba {CT|kw'isu=com isso}, aqui a pesca [AB|do {pp}] artesanal daqui da nossa praia está perdida.

(VPA01-C)

10.2. Formas incompletas: código "RC" (*reconstructed*)

[RC|forma ortográfica dos segmentos produzidos=-forma ortográfica completa]

Existem casos em que a reconstituição de formas incompletas pode ser vantajosa para a análise sintáctica. Nestes casos, são apresentadas duas transcrições da mesma unidade lexical: a transcrição ortográfica dos segmentos efectivamente produzidos seguidos de hífen e a transcrição ortográfica da forma completa com o material não produzido mas facilmente reconstituível. Estas duas transcrições são apresentadas entre chavetas, separadas por um "sinal de igual" e identificadas com o código "RC" (*reconstructed*).

Exemplo:

INF E depois é que veio, mais tarde – isso já era eu {RC|casa=-casado}, depois já era casado eu.

(VPA18-C)

10.3. Formas e sequências repetidas: código "RP" (*repeated*)

[RP|forma ou sequência repetida]

As repetições consecutivas da mesma forma ou sequência são transcritas ortograficamente entre parênteses rectos e identificadas com o código "RP" (*repeated*).

Exemplo:

INF Vai-se começar [AB|por a]{fp} **por a vinha** {pp} [RP|por a vinha].

(PST01-C)

As repetições intensivas ou aditivas não são, no entanto, assinaladas pelo código "RP".

Exemplos:

INF1 Então depois era **amassar, amassar, amassar** ali: "Ó mãe, isto já {IP|ta=está} bom"? "Não". "Ó mãe, então quando é que {IP|ta=está} bom"? "Quando [AB|as {fp}] as pernas da trapeça suarem"! Boa! As pernas da trapeça {PH|nẽ=não} suavam, aquilo é que era **amassar, amassar!**

(LVR35)

INF1 Era muito gordo! Era era muito gordo, muito gordo, **muito, muito.**

(VPA33)

INF Porque, antigamente, uma pessoa caçava **milhares e milhares e milhares** de centos de sardinhas {pp}, e a sardinha, ia-se para comer {pp}, branquinha na espinha, branquinha!

(VPA26)

INF A gente tinha-se que (ir), outra vez, a **remar, remar, remar**, para fora, para elas.

(CLC28)

INF Às vezes, começo eu assim a olhar – formou-se, além, umas nuvenzinhas – começo eu **a olhar, a olhar, a olhar, a olhar, a olhar** e **vai, vai**: [AB|se] umas, se é para se formar, vão-se [AB|f-] formando {PH|mẽju'rijẽ}=maiorzinhas} e outras **vai, vai, vai, vai, vai**, desfaz-se em vento e {fp} pff, ficou sem nada.

(PAL19)

INF E depois **ia andando, ia andando** [AB|até] {pp} até [AB|fazi-] formar parreira [AB|e] {pp} e dar uvas.

(PST01)

10.4. Pausas não preenchidas com material sonoro – {pp}

As rupturas ou hesitações no desenvolvimento discursivo correspondentes a uma pausa, maior ou menor, não preenchida por material sonoro são indicadas através do código "{pp}" (*plain pause*).

Exemplos:

INF1 E para esfregar as colheres {pp}

11.2. Divergências de audição – (xxx) /yyy\

As audições divergentes ou alternativas são transcritas sequencialmente. A primeira alternativa (aquela que se considera mais provável) é transcrita entre parênteses curvos "()" e a(s) seguinte(s) entre barras oblíquas "/ \". As transcrições das diferentes percepções auditivas – por diferentes transcritores ou por um mesmo transcritor – são separadas por um espaço.

Exemplos:

INF Hoje trazem (**alguma**) /logo uma\ bagatela de peixe, (**ganham-me**)
/ganham\ um dinheirão! (**Ganham-me**) /Ganham\ um dinheirão!
(VPA06-C)

INF2 (**Põe-a em**) /Põe a\ de vinha de alho, com muitos preparos, ou que é, e tal.
(VPA30-C)

11.3. Sequências imperceptíveis — (...)

As sequências não transcritas por serem auditivamente imperceptíveis são assinaladas ortograficamente por reticências entre parênteses curvos "(...)".

Exemplo:

INF (...) E eu tinha aquela. Quando ia lá, ala, pronto.
(VPA40-C)

As palavras de percepção deficiente cuja categoria morfológica é possível identificar são indicadas por reticências seguidas de barra oblíqua e etiqueta morfológica, entre parênteses curvos: "(.../N)", "(.../VB)", "(.../INTJ)", etc. Tanto o formato "barra oblíqua – etiqueta" como as etiquetas morfológicas utilizadas nestes casos estão de acordo com o sistema adoptado nas versões dos textos etiquetados morfológicamente (cf. Manual de Anotação Morfológica).

Exemplos:

INF Até por acaso esta, eu [AB|que] o que hei-de fazer [AB|é{fp}] é tirar a
(.../N), que é porque{fp} a copa {pp} voltou ao contrário.
(AAL44-C)

INF2 Depois o caçapo tem que se (.../VB). Rhã-rhã!
(CLC26-C)

INF4 Dê-me aí [AB|o {fp}cu-] o (cutelo), (.../INTJ)!

(CLC25-C)

12. Informação extra-linguística

Inclui-se informação extralinguística nos textos ortograficamente transcritos nos casos em que essa informação é indispensável para que possam interpretar-se adequadamente certas sequências discursivas. A informação extralinguística é dada entre parênteses rectos "[]".

Exemplos:

INQ1 Peixe-lua?

INF É.

INQ [Mostra outra fotografia]

INF Isso [AB|é o {fp}] é o que a gente chama – que mata-se muito cá – peixe-porco.

(CLC34-C)

INF2 As mãos [AB|inch-] inchadas, (que) não se poder {RC|traba-=trabalhar}.

[AB|Nã-] A gente pôr a molhar! A gente pôr as mãos assim! **[som de palmadas nas mãos]**

INF1 Mas no frio. Tinha de ser assim mesmo {PH|'komɐ=como} tábuas. **[de novo, som de palmadas]** {pp} [AB|D-, d-] Do remo, ficava-se como {fp} calos.

(CLC15-C)

INQ1 Porque, para vocês, a tamagueira é, é tipo cedro.

INF1 É cedros.

INQ2 Ah!

INQ1 Está bem. Bom.

[Corte na gravação original]

INQ1 Claro, para a gente saber. Portanto, a zona do vinho do Porto Santo, era esta aqui, da Camacha. Ou havia outras zonas aqui da ilha onde também havia, onde também se fazia?

(PST08-C)

13. Anonimização dos intervenientes nos eventos registados nos textos transcritos

Por forma a proteger a identidade dos informantes e de todos os indivíduos nomeados nos textos transcritos, os antropónimos relevantes são sempre substituídos por nomes fictícios. Para efectivar este processo de anonimização, que consiste na atribuição aleatória de falsos nomes, usam-se, respeitando a ordem alfabética, os nomes próprios portugueses que constam simultaneamente da lista estabelecida em Bergström e Reis 1997 e do Dicionário Onomástico de José Pedro Machado (Machado 1984). Excluem-se, no entanto, os nomes próprios que apenas existem como apelidos e aqueles que são ambíguos quanto ao género.

Apêndice I — Lista de Assuntos

O céu e os corpos celestes
A atmosfera e as condições climáticas

Os rios e os mares
O terreno, configuração e constituição

Ervas, arbustos e flores
As árvores

Os animais domésticos
Os animais bravios
As aves
Os peixes e outros animais marinhos
Os moluscos e crustáceos
Os insectos e outros invertebrados
Os batráquios e répteis

*

O corpo humano – generalidades
A cabeça e o cabelo
Os órgãos sensoriais e a sua actividade
O tronco e os órgãos internos
Os membros superiores e inferiores
A pele e as infecções da pele
A saúde e as doenças

O vestuário
A alimentação

*

A agricultura – generalidades
Preparação do terreno e rega
Os cereais
A ceifa e a debulha
A desfolhada
As alfaias agrícolas
A horta e os produtos hortícolas
As árvores de fruto e os frutos
Produtos não cultivados utilizados na alimentação

Aproveitamento dos produtos vegetais – generalidades
A vinha e o vinho
A oliveira e o azeite
A farinha: moinho e panificação
O linho e o tear
A lã e o tear
O sobreiro e a cortiça
A azinheira
O pinheiro e a resina
O lenhador e o fabrico de carvão

A criação de gado – generalidades
O gado vacum
O gado ovino
O gado caprino
O leite e o queijo
O porco e a matança
O gado equino
As aves de capoeira
As abelhas e o mel

Os barcos e a pesca

Ofícios, profissões e outras actividades – generalidades
O ferreiro e o ferrador
O carpinteiro e o pedreiro
O oleiro
O cesteiro
O moliço e as salinas
A caça
Ocupações domésticas

A casa de habitação: aspecto exterior e construção
O quarto de dormir
O lar e a cozinha

*

A vida humana: nascimento, vida e morte
A família: relações de parentesco
As relações sociais

A língua e a comunicação
Fórmulas de tratamento
Fórmulas de saudação, despedida, agradecimento e súplica

A sociedade: organização e situações marginais
A vida espiritual
A religião e as superstições
As festas religiosas e profanas
Os jogos e as diversões

*

Numerais e medidas
Unidades monetárias

*

A passagem do tempo – generalidades
As fases do dia
Os dias da semana
Os meses do ano

B. Versão Normalizada

1. Formato word/pdf

A versão normalizada da transcrição obtém-se executando automaticamente as seguintes operações:

- (i) Eliminação dos sublinhados (cf. Secção A. 3);
- (ii) Eliminação dos sinais identificadores de pausa – {pp}, {fp} (cf. Secção A. 10.4 e 10.5);
- (iii) Supressão das sequências delimitadas e marcadas pelos códigos AB (*abandoned*) e RP (*repeated*) (cf. Secção A. 10.1 e 10.3), sendo os lugares de elisão assinalados por "(...)".
- (iv) Eliminação das transcrições em alfabeto fonético que representam variantes dialectais e palavras estrangeiras, bem como da transcrição ortográfica dos segmentos efectivamente produzidos de formas reconstruídas (cf. Secção A. 8, 9 e 10.2). Concomitantemente são suprimidos os códigos PH (*phonetic*), CT (*contracted*), IP (*initial partial*), FR (*foreign*), RC (*reconstructed*) e respectiva formatação;

Exemplo 1:

INQ Como é que chamavam à ponta? Tinha algum nome, a ponta?

INF [AB|Era o, era o {fp}] {PH|ʃɐma'βamuʒɫi=Chamávamos-lhe} fuso. O fuso de {CT|fi'alu=fiar o} linho {pp} era o fuso de ferro; e {PH|ɔ=o} de {CT|fi'alɐ=fiar a} lâ era o fuso de pau.

INQ Hum, sim senhor.

INF Só que o de {CT|fi'alu=fiar o} linho {pp} era: tinha um bocadinho de pau cá no fundo {pp} e outro bocadinho de ferro. E {PH|ɔ=o} da lâ era todo de pau.

(CTL01-C)

INQ Como é que chamavam à ponta? Tinha algum nome, a ponta?

INF (...) Chamávamos-lhe fuso. O fuso de fiar o linho era o fuso de ferro; e o de fiar a lâ era o fuso de pau.

INQ Hum, sim senhor.

INF Só que o de fiar o linho era: tinha um bocadinho de pau cá no fundo e outro bocadinho de ferro. E o da lâ era todo de pau.

(CTL01-N-word/pdf)

Exemplo 2:

INQ Quem é que fazia as rocas, eram as mulheres ou eram os homens, os rapazes?

INF1 As mulheres, as mulheres.

INF2 E os homens. [AB|E os ho-]

INQ E os rapazes não costumavam oferecer à rapariga umas rocas todas enramadas?

INF1 [AB|As.os ra-] Os rapazes também faziam.

INF2 E fusos? Diga-lhe quem {PH|nu|=os} fazia.

INF1 {fp}. Mas os fusos faziam-nos os homens. Isso faziam-nos os homens.

[AB|E os rapazes depois]

INF2 Os fusos de pau.

INF1 Sim. E [AB|os] o homem que era (.../ADJ) e o rapaz é que fazia [AB|o{fp}] o fuso [AB|para lhe] para lhe dar.

INQ Pronto.

INF1 Mas as rocas, [AB|eu] nós fizemo-las sempre nós. Roquinhas que aquilo davam... Era uma maravilha!

(CTL02-C)

INQ Quem é que fazia as rocas, eram as mulheres ou eram os homens, os rapazes?

INF1 As mulheres, as mulheres.

INF2 E os homens. (...)

INQ E os rapazes não costumavam oferecer à rapariga umas rocas todas enramadas?

INF1 (...) Os rapazes também faziam.

INF2 E fusos? Diga-lhe quem os fazia.

INF1 Mas os fusos faziam-nos os homens. Isso faziam-nos os homens. (...)

INF2 Os fusos de pau.

INF1 Sim. E (...) o homem que era (.../ADJ) e o rapaz é que fazia (...) o fuso (...) para lhe dar.

INQ Pronto.

INF1 Mas as rocas, (...) nós fizemo-las sempre nós. Roquinhas que aquilo davam... Era uma maravilha!

(CTL02-N-word/pdf)

2. Formato ASCII

Os ficheiros em formato ASCII da Versão Normalizada dos textos destinam-se a ser submetidos a ferramentas de tratamento de texto, nomeadamente, a programas de concordância, etiquetadores morfológicos ou anotadores sintácticos.

Os ficheiros neste formato obtêm-se a partir dos seus correspondentes em formato Word, executando automaticamente as seguintes operações:

- (i) Simplificação do cabeçalho que antecede cada excerto de transcrição (cf. Secção A. 1), que passa a conter apenas o código de identificação do ficheiro;
- (ii) Delimitação das sequências a ignorar pelas ferramentas de tratamento de texto, nomeadamente:
 - delimitação de "(...)" – correspondente a sequências elididas na Versão Normalizada (cf. Secção B. 1. iii) ou a sequências não transcritas por serem imperceptíveis (cf. Secção A. 11.3) – pelos códigos `<break>` `</break>`;
 - delimitação de "(.../ETIQ)" – correspondente a palavras imperceptíveis cuja categoria morfológica é possível identificar (cf. Secção A. 11.3) – pelos códigos `<break>` `</break>`;
 - delimitação de informação extralinguística transcrita entre parênteses rectos "[]" (cf. Secção A. 13) pelos códigos `<break>` `</break>`;
 - delimitação dos cabeçalhos simplificados pelos códigos `<header>` `</header>`;
 - delimitação das falas dos inquiridores, e respectiva identificação por "INQ", bem como do rótulo "INF" que identifica o início das falas dos informantes (cf. Secção A. 2) pelos códigos `<inq>` `</inq>` e `<inf>` `</inf>`, respectivamente.
- (iii) Eliminação dos parênteses curvos delimitadores da transcrição da primeira alternativa nos casos em que existem divergências de audição (cf. Secção A. 11.2);
- (iv) Eliminação das barras oblíquas delimitadoras da transcrição da segunda alternativa (ou seguintes) nos casos em que existem divergências de audição (cf. Secção A. 11.2) e sua substituição pelos códigos `<alt>` `</alt>`;

- (v) Eliminação dos parênteses curvos delimitadores da transcrição de segmentos discursivos cuja percepção é duvidosa (cf. Secção A. 11.1) e introdução dos códigos <alt> </alt> a seguir a esta. Neste caso, os códigos delimitam um espaço em branco – esta estratégia indica a existência de uma dúvida de transcrição naquele ponto do excerto sem que, no entanto, se sugira uma alternativa à transcrição apresentada;

Exemplo 1:

INQ Como é que chamavam à ponta? Tinha algum nome, a ponta?

INF (...) Chamávamos-lhe fuso. O fuso de fiar o linho era o fuso de ferro; e o de fiar a lã era o fuso de pau.

INQ Hum, sim senhor.

INF Só que o de fiar o linho era: tinha um bocadinho de pau cá no fundo e outro bocadinho de ferro. E o da lã era todo de pau.

(CTL01-N-word/pdf)

```
<header> CTL01 </header>
```

```
<inq> INQ Como é que chamavam à ponta? Tinha algum  
nome, a ponta? </inq>
```

```
<inf> INF </inf> <break> (...) </break> Chamávamos-lhe  
fuso. O fuso de fiar o linho era o fuso de ferro; e o  
de fiar a lã era o fuso de pau.
```

```
<inq> INQ Hum, sim senhor. </inq>
```

```
<inf> INF </inf> Só que o de fiar o linho era: tinha um  
bocadinho de pau cá no fundo e outro bocadinho de  
ferro. E o da lã era todo de pau.
```

(CTL01-N-ASCII)

Exemplo 2:

INQ Quem é que fazia as rocas, eram as mulheres ou eram os homens, os rapazes?

INF1 As mulheres, as mulheres.

INF2 E os homens. (...)

INQ E os rapazes não costumavam oferecer à rapariga umas rocas todas enramadas?

INF1 (...) Os rapazes também faziam.

INF2 E fusos? Diga-lhe quem os fazia.

INF1 Mas os fusos faziam-nos os homens. Isso faziam-nos os homens. (...)

INF2 Os fusos de pau.

INF1 Sim. E (...) o homem que era (.../ADJ) e o rapaz é que fazia (...) o fuso (...) para lhe dar.

INQ Pronto.

INF1 Mas as rocas, (...) nós fizemo-las sempre nós. Roquinhas que aquilo davam... Era uma maravilha!

(CTL02-N-word/pdf)

```
<header> CTL02 </header>
```

```
<inq> INQ Quem é que fazia as rocas, eram as mulheres
ou eram os homens, os rapazes? </inq>
<inf> INF1 </inf> As mulheres, as mulheres.
<inf> INF2 </inf> E os homens. <break> (...) </break>
<inq> INQ E os rapazes não costumavam oferecer à
rapariga umas rocas todas enramadas? </inq>
<inf> INF1 </inf> <break> (...) </break> Os rapazes
também faziam.
<inf> INF2 </inf> E fusos? Diga-lhe quem os fazia.
<inf> INF1 </inf> Mas os fusos faziam-nos os homens.
Isso faziam-nos os homens. <break> (...) </break>
<inf> INF2 </inf> Os fusos de pau.
<inf> INF1 </inf> Sim. E <break> (...) </break> o homem
que era <break> (.../ADJ) </break> e o rapaz é que
fazia <break> (...) </break> o fuso <break> (...)
</break> para lhe dar.
<inq> INQ Pronto. </inq>
<inf> INF1 </inf> Mas as rocas, <break> (...) </break>
nós fizemo-las sempre nós. Roquinhas que aquilo
davam... Era uma maravilha!
```

(CTL02-N-ASCII)

Exemplo 3:

INQ1 Porque, para vocês, a tamagueira é, é tipo cedro.

INF1 É cedros.

INQ2 Ah!

INQ1 Está bem. Bom.

[Corte na gravação original]

INQ1 Claro, para a gente saber. Portanto, a zona do vinho do Porto Santo, era esta aqui, da Camacha. Ou havia outras zonas aqui da ilha onde também havia, onde também se fazia?

(PST08-N-word/pdf)

```
<inq> INQ1 Porque, para vocês, a tamagueira é, é tipo
cedro. </inq>
<inf> INF1 </inf> É cedros.
<inq> INQ2 Ah! </inq>
<inq> INQ1 Está bem. Bom. </inq>
```

<break> [Corte na gravação original] </break>
<inq> INQ1 Claro, para a gente saber. Portanto, a zona do vinho do Porto Santo, era esta aqui, da Camacha. Ou havia outras zonas aqui da ilha onde também havia, onde também se fazia? </inq>

(PST08-N-ASCII)

Exemplo 4:

INF (O) meu pai era lavrador. Se chovesse, que (o) meu pai tivesse trigo e cevada para o ano inteiro, a gente tinha a nossa fartura de pão. Mas se não havia para o ano inteiro, era deste milho! Ia-se buscar este milho, havia moinhos de vento, aqui – que agora só há um, mas o dono até morreu – e moía-se e a minha mãe fazia milho, como agora se faz deste, destas arepas, da farinha da Venezuela.

(PST10-N-word/pdf)

<inf> INF </inf> O <alt> </alt> meu pai era lavrador. Se chovesse, que o <alt> </alt> meu pai tivesse trigo e cevada para o ano inteiro, a gente tinha a nossa fartura de pão. Mas se não havia para o ano inteiro, era deste milho! Ia-se buscar este milho, havia moinhos de vento, aqui – que agora só há um, mas o dono até morreu – e moía-se e a minha mãe fazia milho, como agora se faz deste, destas arepas, da farinha da Venezuela.

(PST10-N-ASCII)

Exemplo 5:

INQ Tinha uma escada por fora. De madeira?

INF Há de madeira e também havia em pedra, metida no patamal, que a gente (chama) /se chama\ /lhe chama\, também havia, mas a maioria era tudo de escada de pau, de madeira.

(PST26-N-word/pdf)

<inq> INQ Tinha uma escada por fora. De madeira? </inq>
<inf> INF </inf> Há de madeira e também havia em pedra, metida no patamal, que a gente chama <alt> se chama </alt> <alt> lhe chama </alt>, também havia, mas a maioria era tudo de escada de pau, de madeira.

(PST26-N-ASCII)

Bibliografia

BERGSTRÖM, M. & N. Reis 1997. *Prontuário Ortográfico e Guia da Língua Portuguesa*. Actualização linguística de M. H. Costa Campos; Coordenação da edição de M. T. Brocardo et alii. Lisboa: Caminho. (33ª edição)

BLANCHE-BENVENISTE, Claire e C. Jeanjean 1987. *Le Français Parlé: Transcription et Édition*. Paris: Didier.

GONÇALVES, R. 1966. *Vocabulário da Língua Portuguesa*. Coimbra: Coimbra Editora.

GOTTSCHALK, F. et alii 1974. *Questionário Linguístico do Atlas Linguístico-Etnográfico de Portugal e da Galiza*. Lisboa: Instituto de Linguística.

MACHADO, J. P. 1984. *Dicionário Onomástico e Etimológico da Língua Portuguesa*. Lisboa: Editorial Confluência.

MORAIS, A. [1789] 1948. *Grande Dicionário da Língua Portuguesa*. Lisboa: Editorial Confluência (10ª ed.).

NASCIMENTO, M. F. Bacelar do, M. L. Garcia Marques e M. L. Segura da Cruz 1987. *Português Fundamental: Métodos e Documentos*. Tomo 1. Lisboa: Instituto Nacional de Investigação Científica & Centro de Linguística da Universidade de Lisboa.

STROUD, Christopher e Perpétua Gonçalves 1997. *Panorama do Português Oral de Maputo*. Vol. I: *Objectivos e Métodos*. Maputo: República de Moçambique

VIANA, Céu, I. Trancoso e I. Mascarenhas 1999. CORAL - Corpus de Diálogo Etiquetado, Doc: Coral/99/03: Transcrição Ortográfica (Tarefa T3). URL: http://www.speech.inesc.pt/~imt/priv/coral/doc/cor99_03.html