A description of an annotation scheme to analyse anaphora in dialogues

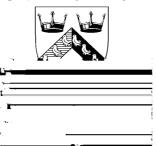
M arco Antonio Estevesda R ocha

427

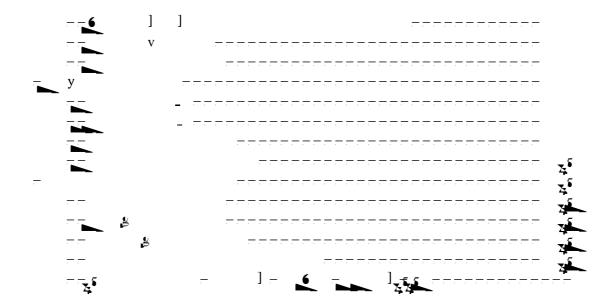
February, 1998

ISSN 1350-3162

UNIVERSITY OF

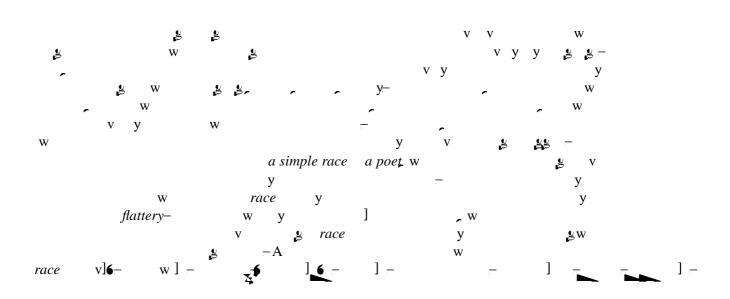


Cognitive Science Research Papers



v Contents



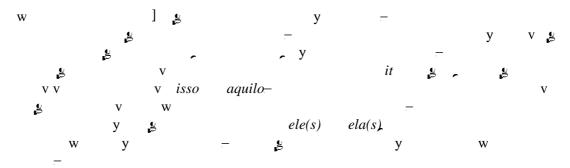


tr: It comforts, perhaps; but it also makes us diffferent from others

PQ: cava abismos entre os homens...

gl: digs abysses between the-MASCp men

tr: it digs chasms between men



(2) S: Eu vou dar a despedida

gl: I go give-INF the-FEMs goodbye

tr: I'm going to say goodbye

S: como deu o bacurau

gl: as gave-3rds the-MASC bacurau (a bird)

tr: as the bacurau did

S: uma perna no caminho

gl: one-FEM leg on-the-CONTR road

tr: one leg on the road

S: a outra no galho de pau

gl: the-FEM other-FEM on-the-CONTR branch of wood

tr: the other on the tree branch

v dar w

У

v

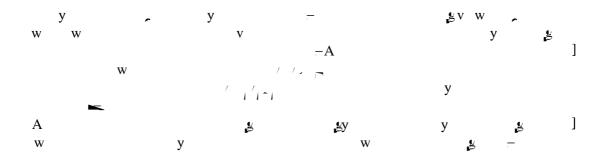
I did tr: C1: compraste? bought-2nds gl: tr: did you buy it? C2: comprei bought-1sts gl: I did tr: me diz quanto foi C1: me tell how much was-3rds gl: tell me how much it was tr: C2: foi quinhento-réis was-3rds five hundred réis (old Brazilian currency) gl: it was five hundred réis tr:] gyy y g y g w quinhento-réis, g g y W gg y g y g] v \mathbf{w} g] y W y] y y y У W] g gW g у] g g gg g y g y w W v g g] y -Ay y W W W y]]] y W] W

went-1sts

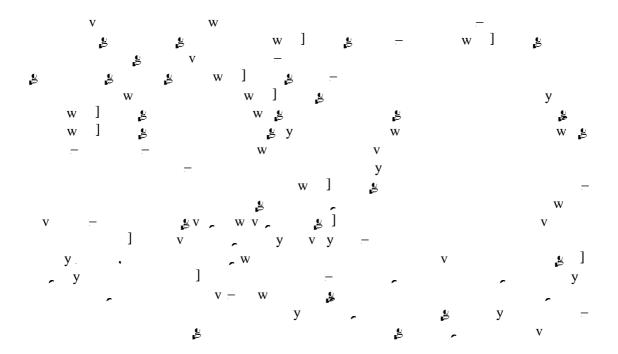
v

g y

gl:



2.1. Anaphora processing and discourse



] g g -

- (6) I was supposed to get up at about seven o'clock
 - C What do you mean you were supposed to

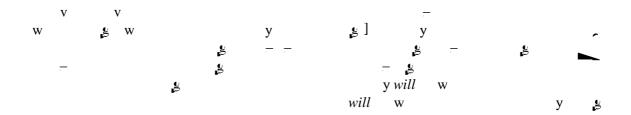
W 💆 V V

3.1. The notion of topic

v g - g g w g - g - A **6** Chapter 3. Methodology

– v





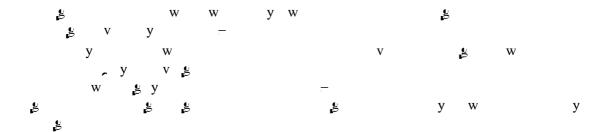
w Ford , w –
y y y bibliography bibliography Ford

| y | į | g w |] | w Ford | | | |
|-----|------------|------------|---|---------------------|------|-------------|------|
| , V | V | W | | g university | | Lor | ıdon |
| | university | | | v y | | g g | _ |
| | • | Ford g | | y | Ford | y | |
| g | , | , W | | g | g | W | - |

Chapter 3. Methodology y v y faculty board board of faculty w
faculty of arts faculty

 \mathbf{W}

8 -



Chapter 3. Methodology

thesis— y w 🛂

у , у

w yl doctor phone 6 1 6 x y - x

v v g

0016 B | didn't s- | didn't see

0017 A you didn't see it

0018 B well

0019 B no I didn't

0020 BIII all I know

 $0021~\mbox{B}$ was my mother was having her lunch

0022 B when I arrived

0023 A and

0024 A how did she seem then

0025 A at two o'clock

0026 B well

0027 B she seemed all right

0028 B I think she was a little tired

0029 A and how long did it take

0030 A for her to complete her lunch

0031 B oh I would think

0032 B probably

0033 B fifteen minutes

0034 A was it any a meal of any substance

0035 B she had erm chicken

0036 B she didn't eat very much of it

0037 A did you sit with her

0038 A whilst

0039 A she completed the meal

0040 B I was in the room

0041 B while she was having it

0042 B yes

0043 A and then uh did she have it on a tray

0044 B yes

0045 A somebody took the tray out presumably

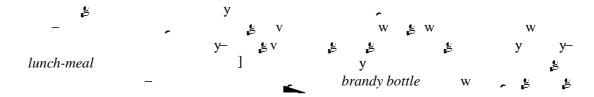
0046 B er my wife took it out

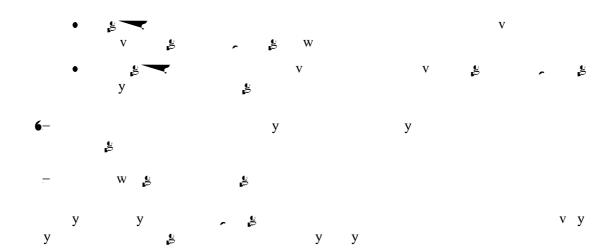
0047 A and uh that's then about two fifteen

0048 B uh yes

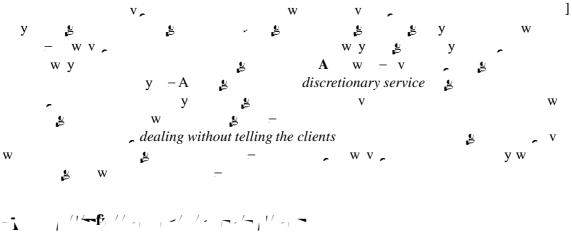
0049 B i- yes

0050 B it would be

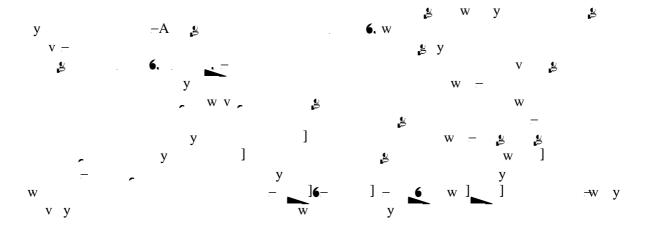




called a discretionary service um whereby people sign a little chit and that we deal for them without telling them

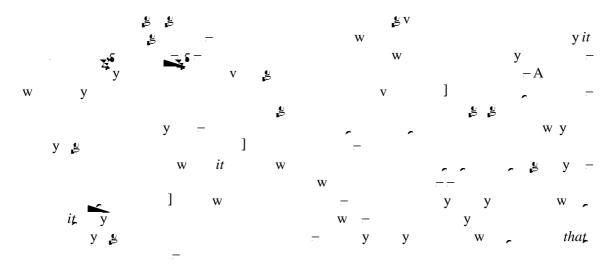


- (9) **B:** erm in the sort of general outline that I sent you of the of the project how did it strike you
 - A: oh I think it's good

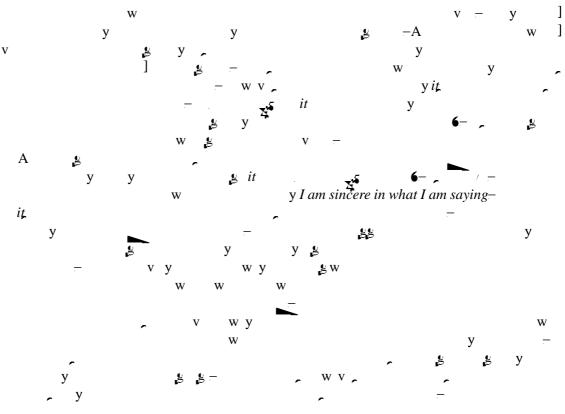


signed the will

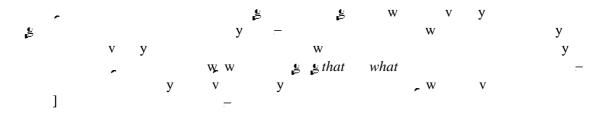
B: well it didn't really register that Mr Coleman was her solicitor to me



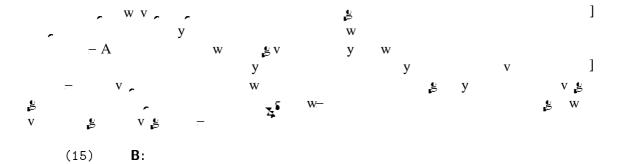
- (12) B: I hope you'll accept my word on this
 - A: yes
 - B: because I I mean it
 - A: all right I will



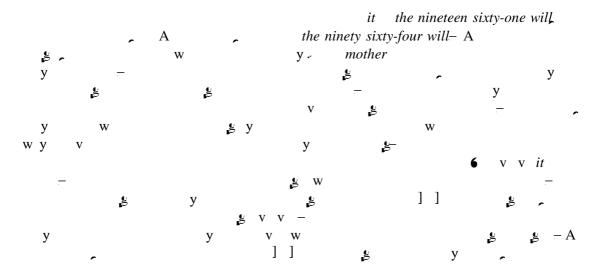
(13) **B**: it was very shortly after that interview that I sent my circular letter around to various scholars and I sent you a copy



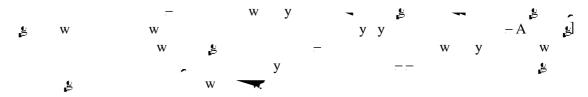
(14) **B:** it's the academic structure of the university that that uh we're concerned about



A: when you read it did you read the whole thing through



g g



(17) * s19 'B's statement through solicitors'

A: your solicitors furnished a statement made by you to the defendant's, is that within your knowledge

B: yes yes

A: you know that

B: yes

A: mm

ss1/s19 'contents of statement'

A: did you know that in that statement furnished by your solicitor to the defendant's, it's stated: both Elsie and I had suggested this to mother before I phoned her doctor who was out but arranged

B: what it was all about yes

A: I gather you've been at it for nine years

B: erm by golly that's true yes yes it's not a long time of course in the uh in this sort of work (FNP; im_5; thel; SK;) you know



(20) A: was there any time between your arrival at two o'clock and your departure after she had signed the will when she had any alcoholic drink

B: no

A: are you sure (AdjAn; ex_162; p_st; VMm;)

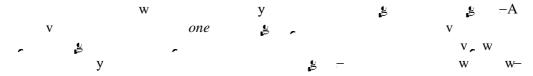
B: I'm absolutely positive (AdjAn; ex_162; p_st; VMm;)

A: very good

(21) A: I just took it out of the shelf that particular volume because it was the smallest book

B: mm mm

A: you know I just go into uh a stationer and buy whatever happens to be there you see and that happened to be the smallest (AdjAn; ex_29; dt; AM;)



(22) A: o peso menor que a senhora pode atingir

gl: the-MASC weight smaller that the lady can reach

tr: the lowest weight you should reach

A: são quarenta e seis e oitocentos

gl: are forty and six and eight humdred

tr: is forty-six eight hundred

A: quase quarenta e sete quilos

gl: almost forty and seven kilos

tr: almost forty-seven kilos

B: o menor (AdjAn; ex

Chapter 4. Description of the annotation scheme

A it he she they g w y
y it w -A
ele ela eles elas y v



(27) A: the the the sort of Harold Macmillan the the um Harold Nicholson type who write their (Pos; ex_22; sst; FtC;) diary because they are aware of having their (Pos; ex_22; sst; FtC;) pulse on the on the goings on of the time



gl: calcium is not going to be so well absorbed

A: se a senhora fizer as refeições assim (AdvM; ex_

15/11/C

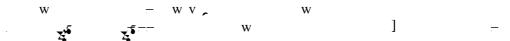
A: the whole will

B: yes

A: in (PP; ex_19; p_st; VMm;) one go or how many turns

A v g ,

pediu – v



- (48) A: I wasn't asked is the answer if it was a legal document and never mentioned having witnessed a will on an envelope had I done so (DPA; ex_115; p_sst; CK;) it would have been a lie
- (49) **B**: this wouldn't work properly the first time and mother signed again

tr: all right

A: nunca na mesma refeição

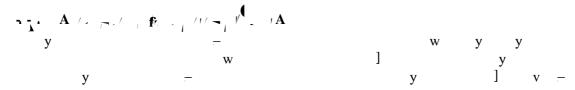
gl: never in-the-CONTR same-FEM meal

tr: never in the same meal

A: ou deixa para fazer (DPA; im_49; p_sst; VMm;) na próxima refeição

gl: or leave to do in-the-CONTR next-FEM meal

tr: or else do that in the next meal



(53) **B:** in the hope that they would do two things firstly - to give (NF-ClAn; ex_68; p_sst; SetMb;) me uh ay- small Ford Ford Foundation travelling grant to visit a number of key centres and universities to explore the land so to speak - and uh when that has been done to submit (NFClAn; ex_68; p_sst; SetMb;) to them a full documented report with the backing of virtually every major library and every major philologist in the world to get them to give me a substantial sum of money to enable me to finish it



- (56) A: diminuiu novecentos gramas
 - gl: diminished-PAST3rds nine hundred grams
 - tr: you lost nine hundred grams
 - A: mas é uma boa coisa, né, em um mês
 - gl: but is a-FEMs good-FEM thing, not is, in a-MASC month
 - tr: but it is a good thing, isn't it, in a month
 - **B**: é (LV; ex_6; p_dt; FtCCh;)
 - gl: is
 - tr: yes, it is
- (57) A: e aí você fez uma uma pequena cirurgia (2syl)
 - gl: and then you made a a small surgery
 - tr: did you have a small surgery then?
 - **B**: é
 - gl: is
 - tr: yes, I did



- (58) A: I just took it out of the shelf, that particular volume because it was the smallest book
 - B: mm mm
 - A: you know I just go into uh a stationer and buy whatever happens to be there you see and that happened to be the smallest and was (CopFNP; ex_29; dt; PI;) the most convenient to carry



- (59) A: sim mas aqui fruta uma fruta né?
 - gl: yes but here fruit a-FEM fruit not is?
 - tr: yes but here fruit is any fruit isn't it?
 - A: então quer dizer a senhora
 - gl: then want say-INF the-FEM lady

tr: then I mean you

A: pode ser a banana (CopFNP; ex_52; st; FtC;)

gl: may be the-FEM banana

tr: it may be a banana

A: pode ser a laranja (CopFNP; ex_52; st; FtCCh;)

gl: may be the-FEM orange

tr: it may be an orange

tr: and do you have any idea of how your pressure has been in the last days

B: não 'tava a a quinze por oito (CopPP; ex_16; st; FtC;)



- (64) A: but there's no indication there of who the writer is
 - **B:** no and we may well mother is mother uh coughs if she signs herself (REF; ex_64; thel; FtCCh;) mother

(65) A: Mr Potter did you arrive about two o'clock on the Sunday the date (FNP; im_2; thel; SK;) the will (FNP; im_1; dt; SK;) was signed

B: yes (AdvR; ex_198; p_st; VMm;)

A: and did you go and see your mother (FNP; im_3; dthel; SK;) straight away

B: yes I did (OPT; ex_4; p_st; VMm;)

A: what was she (SP; ex_3; dthel; FtC;) then doing

B: she (SP; ex_3; dthel; FtCCh;) was having her (Pos; ex_3; dthel; FtCCh;) lunch

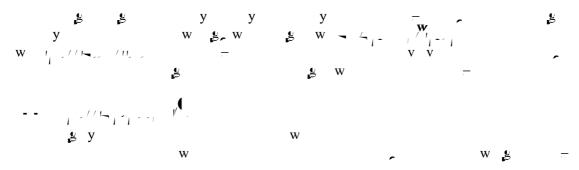
the Sunday the date the will was signed— y
the date the will was signed

W W

y

-- -/, --, / - /<u>·</u>

A: but you have applied er for monies (FNP; im_12; st; LS;) | keep hearing wherever | go



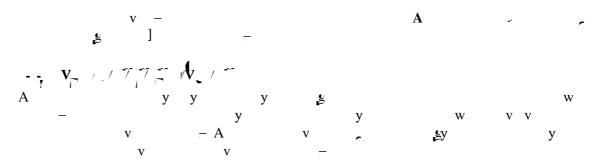
(77) **B**: and erm I don't knoldon't knonwee

w g y – A gv —

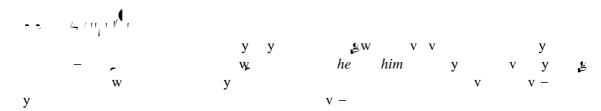
(79) A: how's the thesis going

B: uh I'm typing it (OP; ex_1; dt; FtC;) up now

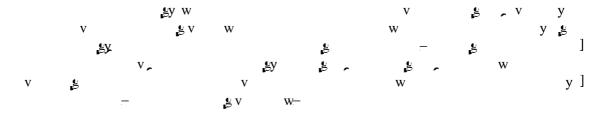
-- -- 1_z, -_z, -_z,



- (82) A: I gather you've been at it for nine years
 - **B:** erm by golly that's true yes yes it's not a long time of course in the uh in this sort of work you know
 - A: well no but it's quite a long time by any standards
 - **B**: yes suppose so (SoAn; ex_6; p_st; VMm;)

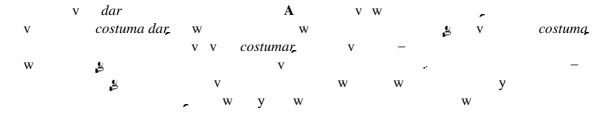


(83) A: well of course a stockbroker doesn't do that he merely takes on Mr Y as a client and he (SP; ex_220; dt; PI;) does his best for him (OP; ex_149; dthel; PI;)



- (84) A: você tem gases, costuma ter, assim?
 - gl: you have gases, wont-PRES3s to have, so?
 - tr: do you usually have trapped air?
 - A: porque, costuma dar, né, uma uma um incômodo
 - gl: because, wont to give, not-is-CONTR, a a a discomfort
 - tr: because it usually causes a discomfort
 - B: diz que dá
 - gl: says that gives
 - tr: it is said it does
 - A: é, dor mesmo, dá dor mesmo
 - gl: is, pain same, gives pain same
 - tr: yes, real pain it causes real pain
 - B: é diz que dá (VerbAn; ex_110; sst; PI;) muita dor até

gl: is , says that gives , much pain even tr: yes, it is said it does, a lot of pain



any trouble is she (SP; ex $_221$; dthel; DK;) wrong in that (De; ex $_108$; p $_st$; DK;)

B: yes I would take it that she didn't remember it (OP (cataph); ex_109; p_dthel; DK;)



- (88) A: but I've always been told that diarists are crazy as well
 - **B:** um well there may be of course something in this (De; ex_19; p_st; DK;) but

- (89) **B**: we replace all the proper names including place names
 - A: mm yes mhm
 - B: by fictitious ones (One_an; ex_33; thel; SetMb;)

- (90) **B:** if you want to have philosophy and uh mathematics as your your two possible subjects (FNP; ex_26; thel; SetCr;) as an undergraduate then you can do those
 - A: oh no

(92) A: and I said if this is what you (SP; ex

```
- Natural language understanding-
                                               y
                         - Directions in corpus linguistics.
      -Agreement and anaphora-A
                   g – Discourse analysis–
      8-
                   y – Nova Gramática do Português Contemporâneo-
           , Directions in corpus linguistics, &
   – A y y v y Advances in Spoken Discourse Analysis–
             - v g
                  , New directions in English language corpora: methodology,
results, software development,
            y , w
         - Discourse structure and anaphora-
   8-
          , A v
                             -Computational Linguistics
    -Computational Linguistics
 -A-- y - - Cohesion in English-
            , Readings in Natural Language Processing-
```

y–Patterns of lexis in text– v y

y —A university grammar of English— g

y v v y — A comprehensive grammar of the English language— g

y —Anaphora and semantic interpretation—

y —Anaphora and semantic interpretation—

v g w]

operational factors in practical, robust anaphora resolution for unrestricted texts g

A — y —

— A corpus-based study of anaphora in dialogues in English and Portuguese—

y y y

A v A A gv

A v V

A v V

A v V

V V

V V

A v A V

g Av

 $]_{\mathbf{W}}$

g

V

§ V

v

w §

w g

w g