

A description of an annotation scheme to analyse anaphora in dialogues

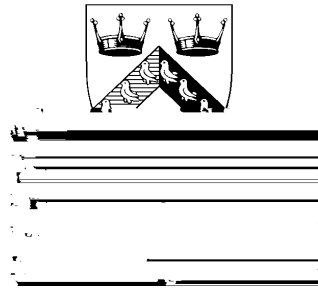
Marco Antonio Esteves da Rocha

427

February, 1998

ISSN 1350-3162

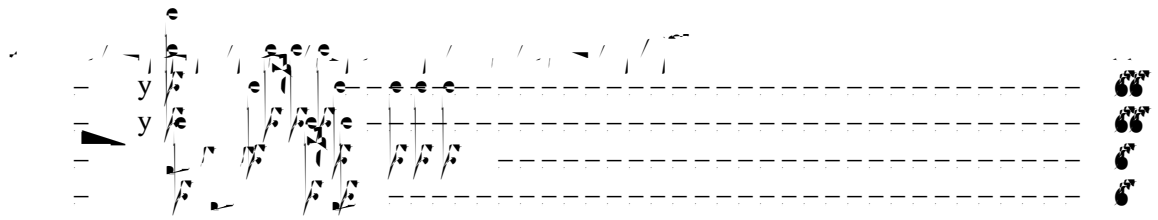
UNIVERSITY OF



Cognitive Science
Research Papers

A musical score consisting of five staves. The notation includes various note values (quarter, eighth, and sixteenth notes), rests, and dynamic markings such as *mf* and *ff*. The score is written in a standard musical notation style with a treble clef on the first staff. The notation is dense and includes many slurs and ties. The page is numbered 'v' in the top right corner.

v. Contents



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

PQ: cava abismos entre os homens...

gl: digs abysses between the-MASCp men

tr: it digs chasms between men

Musical score for the first passage. The score consists of two staves of music. The lyrics are written below the notes. The lyrics are: *isso aquilo ele(s) ela(s)*. The music is in a minor key and features a complex rhythmic pattern with many sixteenth and thirty-second notes.

(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

Musical score for the second passage. The score shows the beginning of a vocal line with the word *dar*. The music is in a minor key and features a complex rhythmic pattern with many sixteenth and thirty-second notes.

gl: went-1sts

tr: I did

C1: compraste ?

gl: bought-2nds

tr: did you buy it ?

C2: comprei

gl: bought-1sts

tr: I did

C1: me diz quanto foi

gl: me tell how much was-3rds

tr: tell me how much it was

C2: foi quinhento-réis

gl: was-3rds five hundred réis (old Brazilian currency)

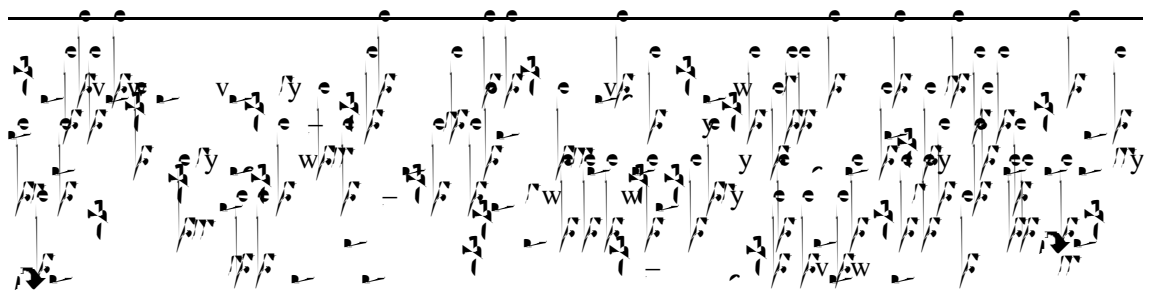
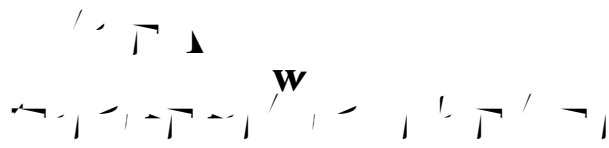
tr: it was five hundred réis

quinhento-réis

The image shows a musical score for the phrase "quinhento-réis". It features a vocal line with lyrics and a piano accompaniment. The lyrics are "quinhento-réis" and the translation is "it was five hundred réis". The score is written in a standard musical notation with a treble clef and a key signature of one flat. The lyrics are written below the vocal line, and the translation is written below the piano accompaniment.

Chapter 1. Introduction

The image displays a musical score for the introduction of a chapter. It consists of three staves. The top staff is a vocal line with lyrics: "y w w e A v w y e". The middle staff is a piano accompaniment line with lyrics: "H y". The bottom staff is another piano accompaniment line with lyrics: "A W y w". The music is written in a complex, rhythmic style with many notes and rests. The lyrics are interspersed with the musical notation.



The image displays a complex musical score consisting of approximately 12 staves. The notation is dense, featuring a variety of note values, rests, and dynamic markings. Interspersed throughout the score are several lines of lyrics, including the words "W", "y", "V", and "H". The overall appearance is that of a highly detailed and intricate musical composition, possibly a choral or orchestral work, with a focus on rhythmic and melodic complexity.

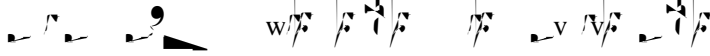
Chapter 2. *Brief overview of the literature*



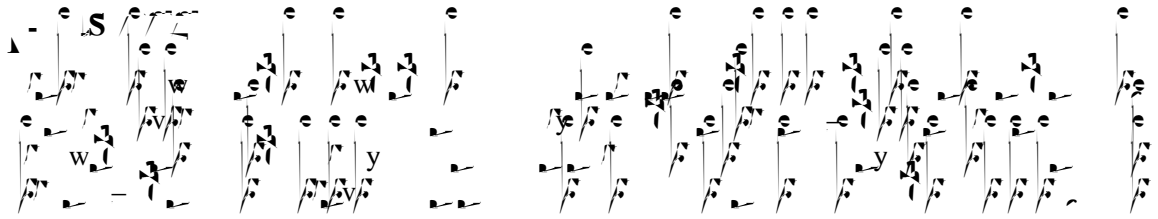
2.3. Corpus-based approaches

(6) I I was supposed to get up at about seven o'clock

C What do you mean you were supposed to



Chapter 2. Brief overview of the literature



3.1. The notion of topic



Chapter 3. Methodology

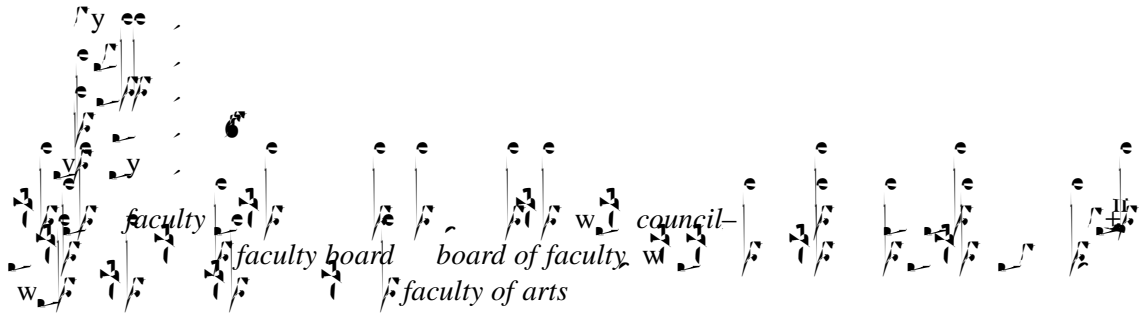
W v e e v e e W y y will W y

The image shows a musical score for a vocal line. The lyrics are: "W v e e v e e W y y will W y". The music is written on a single staff with a treble clef. The notes are mostly eighth and sixteenth notes, with some rests. The lyrics are placed below the notes, with some words like "y will" appearing on a lower line. The score is a snippet of a larger piece, as indicated by the ellipsis in the lyrics.

Chapter 3. Methodology

The image shows a musical score for a piece titled "Ford University London". The score is written on a single staff with a treble clef and a key signature of one flat (B-flat). The melody is composed of eighth and sixteenth notes, with some rests. The lyrics "Ford university London" are written below the staff, with "Ford" appearing twice. The word "y" is written below the staff in two places, once under a note and once under a rest. The score is enclosed in a rectangular box with a double-line border. There are small black triangles at the top left and bottom right corners of the box, possibly indicating the start and end of the score. The overall appearance is that of a musical score for a song or a short piece of music.

Chapter 3. Methodology



A musical score with lyrics. The lyrics are: "faculty board of faculty with council faculty of arts". The score consists of a single line of music with a treble clef and a key signature of one flat. The melody is written in a simple, rhythmic style. The lyrics are placed below the notes. There are some additional markings above the notes, including a "y" and some "e" characters. The score ends with a double bar line.

3.1. The notion of topic

The image shows a musical score for voice and piano. The lyrics are "we're very well". The score is written in a complex, multi-measure format with many notes and rests. The lyrics are written below the notes. The score is divided into two systems. The first system contains the first two lines of music, and the second system contains the next two lines. The lyrics are: "we're very well". The score is written in a complex, multi-measure format with many notes and rests. The lyrics are written below the notes. The score is divided into two systems. The first system contains the first two lines of music, and the second system contains the next two lines. The lyrics are: "we're very well".

3.1. The notion of topic

The image shows musical notation for the words "doctor" and "phone". The word "doctor" is written below the first staff, and "phone" is written below the second staff. The notation consists of a series of notes on a staff, with pitch accents (marked with 'e') above certain notes. The notes are connected by lines, and there are various musical symbols like stems, beams, and slurs. The notation is somewhat abstract and appears to be a phonetic or prosodic representation of the words.

- 0016 B I didn't s- I didn't see
0017 A you didn't see it
0018 B well
0019 B no I didn't
0020 B I I I all I know
0021 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



3.1. The notion of topic

The image shows a musical score for a piece titled "lunch". The score is written on a single staff with a treble clef. The melody is composed of eighth and sixteenth notes, with some rests. The lyrics "lunch sw" are written below the first few notes, and "w w w" are written below the last few notes. The score is marked with a "3.1. The notion of topic" at the top right. The music is in a simple, rhythmic style, possibly a folk or pop song. The notes are mostly on the lower half of the staff, suggesting a lower register. The overall feel is light and melodic.

Chapter 3. Methodology

The image displays a complex musical score, likely for a multi-instrument ensemble or a vocal piece. It consists of several staves of music. The top two staves are a grand staff, with the upper staff containing a treble clef and the lower staff containing a bass clef. Below the grand staff are several single-line staves, some of which begin with a treble clef. The notation is dense, featuring numerous notes, rests, and various symbols such as 'y' and 'w' interspersed throughout the score. The overall appearance is that of a highly technical and detailed musical composition.

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

discretionary service
dealing without telling the clients

(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



3.3. Features of the annotation

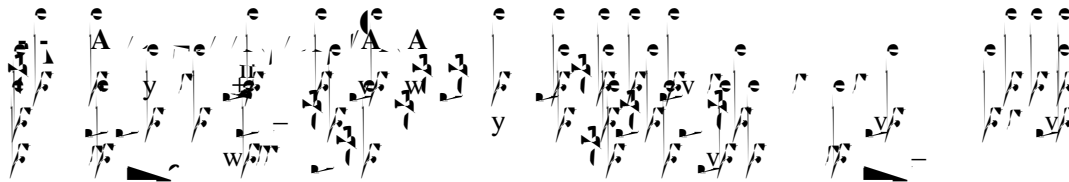


v A w

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; theI; 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;)

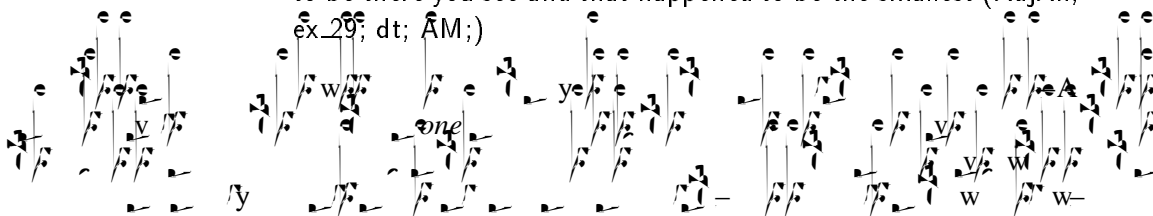
B: I'm absolutely positive (AdjAn; ex_162; p_st; VMm;)

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 hundred

tr: is forty-six eight hundred

A: quase quarenta e sete quilos

gl: almost forty and seven kilos

tr: almost forty-seven kilos

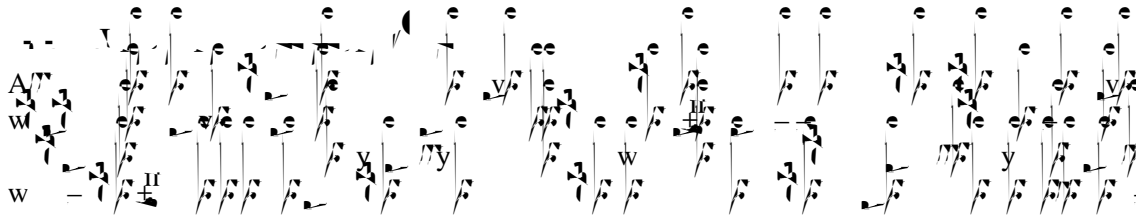
B: o menor (AdjAn; ex

Chapter 4. Description of the annotation scheme

The image shows a musical score for the sentence "it is she they will be with you". The score is written on a single staff with a treble clef and a key signature of one flat (B-flat). The melody is written in a simple, rhythmic style. The lyrics are written below the notes. The annotations consist of small letters and symbols placed above and below the notes. The annotations are: "S" above the first note, "A" above the second note, "e" above the third note, "e" above the fourth note, "e" above the fifth note, "e" above the sixth note, "e" above the seventh note, "e" above the eighth note, "e" above the ninth note, "e" above the tenth note, "e" above the eleventh note, "e" above the twelfth note, "e" above the thirteenth note, "e" above the fourteenth note, "e" above the fifteenth note, "e" above the sixteenth note, "e" above the seventeenth note, "e" above the eighteenth note, "e" above the nineteenth note, "e" above the twentieth note, "e" above the twenty-first note, "e" above the twenty-second note, "e" above the twenty-third note, "e" above the twenty-fourth note, "e" above the twenty-fifth note, "e" above the twenty-sixth note, "e" above the twenty-seventh note, "e" above the twenty-eighth note, "e" above the twenty-ninth note, "e" above the thirtieth note, "e" above the thirty-first note, "e" above the thirty-second note, "e" above the thirty-third note, "e" above the thirty-fourth note, "e" above the thirty-fifth note, "e" above the thirty-sixth note, "e" above the thirty-seventh note, "e" above the thirty-eighth note, "e" above the thirty-ninth note, "e" above the fortieth note, "e" above the forty-first note, "e" above the forty-second note, "e" above the forty-third note, "e" above the forty-fourth note, "e" above the forty-fifth note, "e" above the forty-sixth note, "e" above the forty-seventh note, "e" above the forty-eighth note, "e" above the forty-ninth note, "e" above the fiftieth note, "e" above the fifty-first note, "e" above the fifty-second note, "e" above the fifty-third note, "e" above the fifty-fourth note, "e" above the fifty-fifth note, "e" above the fifty-sixth note, "e" above the fifty-seventh note, "e" above the fifty-eighth note, "e" above the fifty-ninth note, "e" above the sixtieth note, "e" above the sixty-first note, "e" above the sixty-second note, "e" above the sixty-third note, "e" above the sixty-fourth note, "e" above the sixty-fifth note, "e" above the sixty-sixth note, "e" above the sixty-seventh note, "e" above the sixty-eighth note, "e" above the sixty-ninth note, "e" above the seventieth note, "e" above the seventy-first note, "e" above the seventy-second note, "e" above the seventy-third note, "e" above the seventy-fourth note, "e" above the seventy-fifth note, "e" above the seventy-sixth note, "e" above the seventy-seventh note, "e" above the seventy-eighth note, "e" above the seventy-ninth note, "e" above the eightieth note, "e" above the eighty-first note, "e" above the eighty-second note, "e" above the eighty-third note, "e" above the eighty-fourth note, "e" above the eighty-fifth note, "e" above the eighty-sixth note, "e" above the eighty-seventh note, "e" above the eighty-eighth note, "e" above the eighty-ninth note, "e" above the ninetieth note, "e" above the ninety-first note, "e" above the ninety-second note, "e" above the ninety-third note, "e" above the ninety-fourth note, "e" above the ninety-fifth note, "e" above the ninety-sixth note, "e" above the ninety-seventh note, "e" above the ninety-eighth note, "e" above the ninety-ninth note, "e" above the hundredth note. The annotations are: "S" above the first note, "A" above the second note, "e" above the third note, "e" above the fourth note, "e" above the fifth note, "e" above the sixth note, "e" above the seventh note, "e" above the eighth note, "e" above the ninth note, "e" above the tenth note, "e" above the eleventh note, "e" above the twelfth note, "e" above the thirteenth note, "e" above the fourteenth note, "e" above the fifteenth note, "e" above the sixteenth note, "e" above the seventeenth note, "e" above the eighteenth note, "e" above the nineteenth note, "e" above the twentieth note, "e" above the twenty-first note, "e" above the twenty-second note, "e" above the twenty-third note, "e" above the twenty-fourth note, "e" above the twenty-fifth note, "e" above the twenty-sixth note, "e" above the twenty-seventh note, "e" above the twenty-eighth note, "e" above the twenty-ninth note, "e" above the thirtieth note, "e" above the thirty-first note, "e" above the thirty-second note, "e" above the thirty-third note, "e" above the thirty-fourth note, "e" above the thirty-fifth note, "e" above the thirty-sixth note, "e" above the thirty-seventh note, "e" above the thirty-eighth note, "e" above the thirty-ninth note, "e" above the fortieth note, "e" above the forty-first note, "e" above the forty-second note, "e" above the forty-third note, "e" above the forty-fourth note, "e" above the forty-fifth note, "e" above the forty-sixth note, "e" above the forty-seventh note, "e" above the forty-eighth note, "e" above the forty-ninth note, "e" above the fiftieth note, "e" above the fifty-first note, "e" above the fifty-second note, "e" above the fifty-third note, "e" above the fifty-fourth note, "e" above the fifty-fifth note, "e" above the fifty-sixth note, "e" above the fifty-seventh note, "e" above the fifty-eighth note, "e" above the fifty-ninth note, "e" above the sixtieth note, "e" above the sixty-first note, "e" above the sixty-second note, "e" above the sixty-third note, "e" above the sixty-fourth note, "e" above the sixty-fifth note, "e" above the sixty-sixth note, "e" above the sixty-seventh note, "e" above the sixty-eighth note, "e" above the sixty-ninth note, "e" above the seventieth note, "e" above the seventy-first note, "e" above the seventy-second note, "e" above the seventy-third note, "e" above the seventy-fourth note, "e" above the seventy-fifth note, "e" above the seventy-sixth note, "e" above the seventy-seventh note, "e" above the seventy-eighth note, "e" above the seventy-ninth note, "e" above the eightieth note, "e" above the eighty-first note, "e" above the eighty-second note, "e" above the eighty-third note, "e" above the eighty-fourth note, "e" above the eighty-fifth note, "e" above the eighty-sixth note, "e" above the eighty-seventh note, "e" above the eighty-eighth note, "e" above the eighty-ninth note, "e" above the ninetieth note, "e" above the ninety-first note, "e" above the ninety-second note, "e" above the ninety-third note, "e" above the ninety-fourth note, "e" above the ninety-fifth note, "e" above the ninety-sixth note, "e" above the ninety-seventh note, "e" above the ninety-eighth note, "e" above the ninety-ninth note, "e" above the hundredth note.



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



Chapter 4. Description of the annotation scheme

gl: calcium is not going to be so well absorbed

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

4.1. *The type of anaphor*

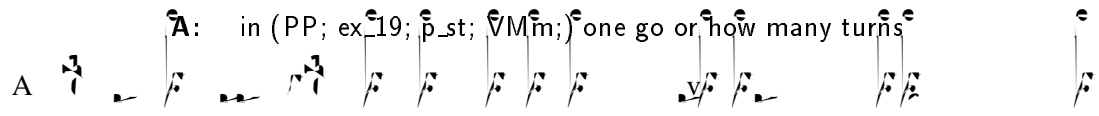
... / ...

Chapter 4. Description of the annotation scheme

A: the whole will

B: yes

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



The image shows a musical score for the phrase "one go or how many turns". The notation is written on a single staff with a treble clef. Above the notes, there are several annotations: "in (PP; ex_19; p_st; VMm;)" is written above the first part of the phrase, and "one go or how many turns" is written above the rest. The notes are mostly quarter and eighth notes, with some rests. The annotations appear to be part of a larger system, possibly related to the "Description of the annotation scheme" mentioned in the chapter title.

pediu

W

This block contains musical notation for the word "pediu". It consists of a series of notes on a staff, with the word "pediu" written below the first few notes. The notes are connected by stems and have various rhythmic values. A "W" is written below the final note.

4.1. The type of anaphor

This block contains musical notation for the section "4.1. The type of anaphor". It features a series of notes on a staff, with the word "pediu" written below the first few notes. The notes are connected by stems and have various rhythmic values. A "W" is written below the final note.

Chapter 4. Description of the annotation scheme



(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

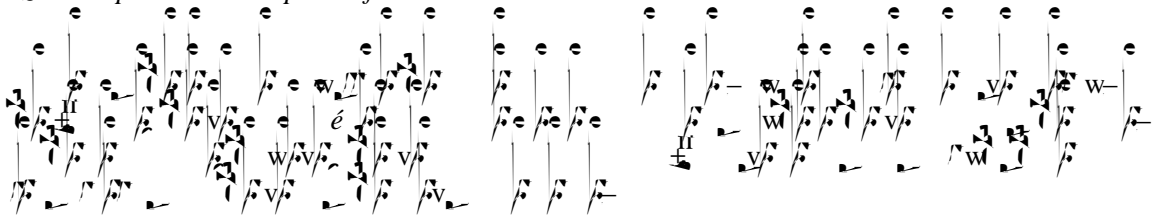


(53)

B: in the hope that they would do two things firstly - to give (NFCIAN; 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 (NFCIAN; 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



Chapter 4. Description of the annotation scheme



- (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; FtCCCh;)
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

4.1. The type of anaphor

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;)

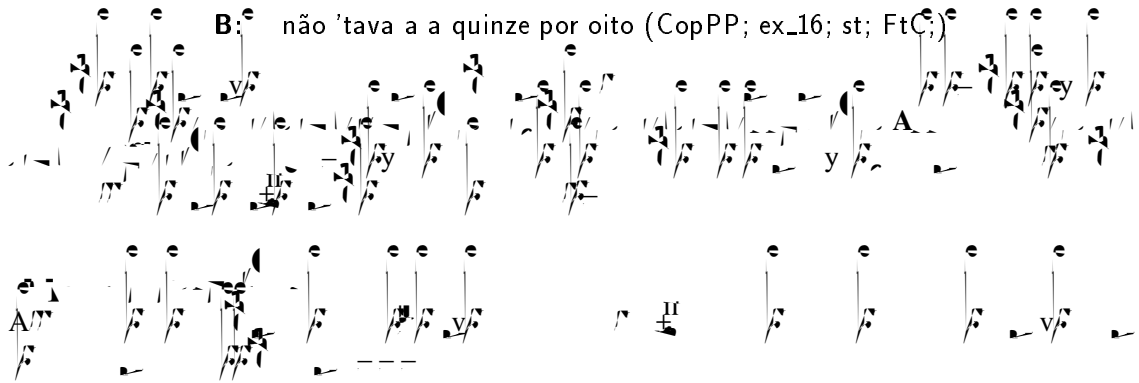
gl: may be the-FEM orange

tr: it may be an orange

→ *... A...*

Chapter 4. Description of the annotation scheme

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; theI; FtCCh;) mother

A:  each other v - 

(65) **A:** Mr Potter did you arrive about two o'clock on the Sunday the date (FNP; im₂; the_l; SK;) the will (FNP; im₁; dt; SK;) was signed

B: yes (AdvR; ex₁₉₈; p_{st}; VMm;)

A: and did you go and see your mother (FNP; im₃; dthel; SK;) straight away

B: yes I did (OPT; ex₄; p_{st}; VMm;)

A: what was she (SP; ex₃; dthel; FtC;) then doing

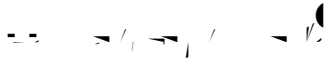
B: she (SP; ex₃; dthel; FtCCh;) was having her (Pos; ex₃; dthel; FtCCh;) lunch

The image shows musical notation for the text. The first line of text is "the Sunday the date the will was signed" with a long horizontal line under "signed" indicating a note extension. The second line of text is "the date the will was signed". The notation consists of a single melodic line with various note values (quarter, eighth, and sixteenth notes) and rests, corresponding to the syllables of the text.

Chapter 4. Description of the annotation scheme

The image displays a musical score with five systems of staves. The notation includes notes, rests, and various annotations. The first three systems are relatively sparse, with notes and rests on a single staff. The fourth system is significantly more complex, featuring multiple staves with dense notation and numerous annotations. The fifth system also shows multiple staves with dense notation and annotations. The annotations include letters such as 'S', 'y', 'w', 'v', 'A', and 'y' placed above or below the notes. Some annotations are enclosed in boxes or have lines pointing to specific notes. The overall appearance is that of a technical document illustrating a specific annotation scheme for musical notation.

4.4. Processing strategy



Chapter 4. Description of the annotation scheme

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;)

Musical notation for example (88) showing two staves, A and B, with lyrics 'y' and 'y-'.

(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

Musical notation for example (89) showing two staves, A and B, with lyrics 'y' and 'y'.

(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;)

Musical notation for example (90) showing two staves, A and B, with lyrics 'y' and 'y'.

(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

Musical notation for example (90) showing two staves, A and B, with lyrics 'that's it so put it mildly'.

♩ Chapter 4. Description of the annotation scheme

A musical score snippet showing two staves. The top staff contains a melodic line with various note values and rests. Above several notes, the letter 'e' is written, indicating a specific annotation. The bottom staff contains a bass line with notes and rests. Above some notes in the bass line, the letter 'v' is written, indicating another annotation. The lyrics 'y', 'w', 'ye', 've', 'w', 'y' are written below the notes in the bottom staff, corresponding to the notes they are positioned under.

Natural language understanding

Directions in corpus linguistics

Agreement and anaphora

Discourse analysis

Nova Gramática do Português Contemporâneo

Directions in corpus linguistics

Advances in Spoken Discourse Analysis

New directions in English language corpora: methodology, results, software development

Discourse structure and anaphora

Computational Linguistics

Computational Linguistics

Cohesion in English

Readings in Natural Language Processing

Patterns of lexis in text

-A university grammar of English-

-A comprehensive grammar of the English language-

-Anaphora and semantic interpretation-

Operational factors in practical, robust anaphora resolution for unrestricted texts

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

A *verb* A

verb

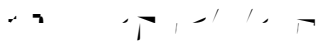
A *verb* y w v A y *verb*

A *verb*

A v A *verb*

A w v w y w y w v w v

A



Musical notation for the first system, featuring a treble clef, a key signature of one flat, and a 4/4 time signature. The notation includes a series of eighth notes and quarter notes, with dynamic markings such as 'A' and 'v'.

Musical notation for the second system, featuring a treble clef, a key signature of one flat, and a 4/4 time signature. The notation includes a series of eighth notes and quarter notes, with dynamic markings such as 'A' and 'v'.

Musical notation for the third system, featuring a treble clef, a key signature of one flat, and a 4/4 time signature. The notation includes a series of eighth notes and quarter notes, with dynamic markings such as 'A' and 'v'.



A

v

A

The first system of the musical score consists of several staves. The notation is dense and complex, with many notes and rests. A large brace on the right side of the system groups several of the staves together. The notation includes various note values, stems, and rests, suggesting a complex rhythmic and melodic structure.

The second system of the musical score continues the complex notation. It features several staves with notes and rests. There are some text labels interspersed within the notation, including 'W', 'y', and 'w'. The notation is dense and complex, with many notes and rests. The system concludes with a large brace on the right side, similar to the first system.