Resolving discourse anaphors -- what can the brain tell us?

Jos J.A. van Berkum

Max Planck Institute for Psycholinguistics
University of Amsterdam, Psychology Department
FC Donders Centre for Cognitive Neuroimaging

Some of you asked me whether there’s a written paper corresponding to my DAARC2007 talk. Unfortunately, the talk was brand new! However, two things get reasonably close:

Van Berkum, J.J.A., Koornneef, A.W., Otten, M., & Nieuwland, M.S. (in press). Establishing reference in language comprehension: An electrophysiological perspective. *Brain Research*. This is a review of (most of) the *Nref effect* studies, and includes a full description of the implicit causality ERP experiment.

Van Berkum, J.J.A. (2004). *Sentence comprehension in a wider discourse: Can we use ERPs to keep track of things?* In M. Carreiras & C. Clifton Jr. (Eds.), *The on-line study of sentence comprehension: Eyetracking, ERPs and beyond* (pp. 229-270). New York: Psychology Press. This is a methods-oriented chapter discussing the pro’s and con’s of ERP research at the text level, including the types of inferences one can draw from ERPs. It is aimed at the beginner in the ERP&discourse field.

You can download these overview papers (as well as all detailed experimental papers) from [www.josvanberkum.nl](http://www.josvanberkum.nl)
Overview

(1) Tracking referential processing by means of ERPs
   lots of ERP data

(2) What factors control this stuff?
   more ERP data

(3) Beyond the text
   more ERP data

(4) So what?
   (and where to go next?)
Only **EEG** (and ERPs, average bits of EEG timelocked to some event X)
The EEG reflects neuronal communication

The scalp-recorded EEG reflects...
the joint activity of large groups of neurons
... which are all simultaneously activated
... which have a particular shape
... and which are oriented similarly
The EEG to a particularly critical event

Gazzaniga, Ivry, and Mangun (1998)
The event-related brain potential

Gazzaniga, Ivry, and Mangun (1998)
Computing the ERP to events of type A (e.g. synt gender violation)

The ERP waveform associated with an event of type A

Average the EEG segments of 50 type A events

EEG segment of interest

I saw a nice\textsubscript{neu} cloud\textsubscript{com} in the sky
Computing the ERP to events of type B (e.g. comparable control)

1. Average the EEG segments of 50 type B events.
2. The ERP waveform associated with an event of type B.
3. EEG segment of interest.

I saw a nice cloud in the sky.
ERPs to a syntactic gender violation (vs. control)

Van Berkum et al (unpublished)
ERPs to a syntactic gender violation (vs. control)

Van Berkum et al (unpublished)
ERPs to a syntactic gender violation (vs. control)
Snowwhite kissed a dwarf before she left.

Incremental syntactic analysis

P600 effect
The P600 reflects a problem with the syntactic analysis

Snowwhite kissed a dwarf before she left.
Snowwhite kissed a *dwarf* before she left.

Incremental semantic analysis

**cloud**
The N400 reflects increasing difficulty in semantic integration.

Snow White kissed a dwarf before she left.
Can we use ERPs to keep track of referential analysis?

Once upon a time, a bunch of dwarfs stumbled across…

When do readers and listeners work out to what or whom particular expressions refer?

Snowwhite kissed a dwarf before she left.
When does the brain use discourse context?

Once upon a time, a bunch of dwarfs stumbled across...

Snowwhite kissed a dwarf before she left.
Two kinds of answers

Common sense:

“Right away, of course!”

Many cognitive scientists:

“Sentence-internal processing is fast...

*but discourse-level processing is slow*”
Study 1&2: Discourse-referential ambiguity, nouns

2-ref: David had asked the two girls to clean up their room before lunchtime. But one of the girls had stayed in bed all morning, and the other had been on the phone all the time. David told the...

1-ref: David had asked the boy and the girl to clean up their room before lunchtime. But the boy had stayed in bed all morning, and the girl had been on the phone all the time. David told the...


Study 1&2: Discourse-referential ambiguity, nouns

2-ref: David had asked the two girls to clean up their room before lunchtime. But one of the girls had stayed in bed all morning, and the other had been on the phone all the time. David told the...

1-ref: David had asked the boy and the girl to clean up their room before lunchtime. But the boy had stayed in bed all morning, and the girl had been on the phone all the time. David told the...


Study 1&2: Discourse-referential ambiguity, nouns

Implications:

- Readers and listeners immediately look for referents in the earlier discourse
  - right at the referring noun
  - within only 300-400 ms after onset of this noun

- The presence of two equally eligible referents for a singular definite noun has differential processing consequences, reflected in a frontally dominant sustained negative shift ("Nref effect")
  - Increased working memory load due to 2 competing referential options?
  - Additional inferencing or search in discourse memory?
  - Anticipation of additional information that will disambiguate?
  - Confound: Prediction of a restrictive postnominal relative clause?
Study 3&4: Sentence-referential ambiguity, pronouns

2-ref: David shot at John as...
1-ref: David shot at Linda as...

Nieuwland & van Berkum, (2006) *Brain Research*
Study 3&4: Sentence-referential ambiguity, pronouns

2-ref: David shot at John as...

Paper-and-pencil pretest ("referential cloze test")

David shot at John as he...
Max called Robert when he...

Complete the sentence and then encircle whom (of the two men) you had in mind for the pronoun

Nieuwland & van Berkum, (2006) Brain Research
Study 3&4: Sentence-referential ambiguity, pronouns

Nieuwland & van Berkum, (2006) *Brain Research*

van Berkum, Zwitserlood, Bastiaansen, Brown, & Hagoort & (2004, CNS)
The Nref effect -- an ERP index of referential ambiguity

2-ref: David had asked the two girls to clean up their room before lunchtime. But one of the girls had stayed in bed all morning, and the other had been on the phone all the time. David told the...

1-ref: David had asked the boy and the girl to clean up their room before lunchtime. But the boy had stayed in bed all morning, and the girl had been on the phone all the time. David told the...

C

Nref effect

2-ref: David shot at John as...
1-ref: David shot at Linda as...

D

Nref effect
Problems with reference and problems with word meaning recruit (at least partially) non-overlapping neuronal ensembles.

Some problems with reference...

Nieuwland, Petersson & Van Berkum (in press) Neuroimage
Study 3&4: Sentence-referential ambiguity, pronouns

2-ref: David shot at John as...
1-ref: David shot at Linda as...

Nieuwland & van Berkum, (2006) Brain Research

van Berkum, Zwitserlood, Bastiaansen, Brown, & Hagoort & (2004, CNS)
Study 3&4: Sentence-referential failure, pronouns

1-ref: David shot at Linda as...

0-ref: Anna shot at Linda as...

written pronoun  A

Nieuwland & van Berkum, (2006) Brain Research
Study 3&4: Sentence-referential failure, pronouns

Nieuwland & van Berkum, (2006) Brain Research

van Berkum, Zwitserlood, Bastiaansen, Brown, & Hagoort & (2004, CNS)
The P600 reflects a problem with the syntactic analysis.

Snowwhite kissed a dwarf before she left.

![Graph showing the P600 effect](image)
Study 3&4: Sentence-referential failure, pronouns

The comprehension system only considers locally available (or strongly implied) referents, so with two women, the pronoun “he” is taken to have the wrong morpho-syntactic marking → P600 effect

Osterhout & Mobley, 1995

Nieuwland & van Berkum, (2006) Brain Research

van Berkum, Zwitserlood, Bastiaansen, Brown, & Hagoort & (2004, CNS)
**Study 3&4: Sentence-referential ambiguity & failure**

2-ref: David shot at John as...

1-ref: David shot at Linda as...

written pronoun

C

Nref effect

400 - 1100 ms

spoken pronoun

D

Nref effect

400 - 1100 ms

written pronoun

A

P600 effect

400 - 1100 ms

spoken pronoun

B

P600 effect

400 - 1100 ms

0-ref: Anna shot at Linda as...
• The Nref effect
  • Emerges with referential ambiguity
  • Nouns and pronouns
  • Spoken and written
  • Referential ambiguity vs semantic integration difficulty (N400)
  • Not necessarily reference- or language-specific!
  • Not the only ERP trace of referential processing!

• The P600 effect
  • Emerges with referential failure on gender/number-marked pronouns
Overview

(1) Tracking referential processing by means of ERPs
lots of ERP data

(2) What factors control this stuff?
more ERP data

(3) Beyond the text
more ERP data

(4) So what?
(and where to go next?)
**Study 4: More referential ambiguity, bigger $N_{ref}$ effect?**

*David shot at John as he...*

Referential cloze test:

<table>
<thead>
<tr>
<th>Sentence</th>
<th>% NP1</th>
<th>% NP2</th>
<th>Contextual bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anton forgave Michael the problem because his ... ... car was a wreck.</td>
<td>50</td>
<td>50</td>
<td>0</td>
</tr>
<tr>
<td>The chemist hit the historian while he ... ... was laughing hard.</td>
<td>53</td>
<td>47</td>
<td>6</td>
</tr>
<tr>
<td>Bruce Willis told Al Pacino that he ... ... was a bit promiscuous.</td>
<td>37</td>
<td>63</td>
<td>26</td>
</tr>
<tr>
<td>Linda invited Anna when her ... ... irritating brother wasn’t around.</td>
<td>70</td>
<td>30</td>
<td>40</td>
</tr>
<tr>
<td>The businessman called the dealer just as he ... ... left the trendy club.</td>
<td>20</td>
<td>80</td>
<td>60</td>
</tr>
</tbody>
</table>

Nieuwland & Van Berkum (2006) *Brain Research*
Study 4: More referential ambiguity, bigger Nref effect?

Moderate ambiguity

*Linda invited Anna when her ...*

70% 30%

Strong ambiguity

*The chemist hit the historian while he ...*

53% 47%

so you need ‘real’ ambiguity
(approx. equibiased options)

Nieuwland & Van Berkum (2006) *Brain Research*
Study 5: Discourse-model referential ambiguity

2-ref: David had asked the two girls to clean up their room before lunchtime. But one of the girls had been sunbathing in the front yard all morning, and the other one had been on the phone all the time.

1-ref: David had asked the boy and the girl to clean up their room before lunchtime. But the boy had been sunbathing in the front yard all morning, and the girl one had been on the phone all the time.

Study 5: Discourse-model referential ambiguity

2-ref: David had asked the two girls to clean up their room before lunchtime. But one of the girls had been sunbathing in the front yard all morning, and the other one had been on the phone all the time.

1.5-ref: David had asked the two girls to clean up their room before lunchtime. But one of the girls had been sunbathing in the front yard all morning, and the other had actually just driven off in his car for some serious downtown shopping. As he gazed at the empty driveway...

1-ref: David had asked the boy and the girl to clean up their room before lunchtime. But the boy had been sunbathing in the front yard all morning, and the girl one had been on the phone all the time.

The Nref effect only emerges when there is true discourse-model referential ambiguity.

At the family get-together, Jim had been talking to his nephew who was very much into politics and his uncle who was really into history. But Jim himself was only interested in sports, cars, girls etc. The uncle who was into history left early, but the nephew who was into politics kept rambling on. Jim didn’t understand one bit and got rather bored. He told the nephew who was into politics that politicians should not systematically neglect delightful and important subjects like sports and girls.
Study 5: Discourse-model referential ambiguity

2-referent

At the family get-together, Jim had been talking to one nephew who was very much into politics and another one who was really into history. But Jim himself was only interested in sports, cars, girls etc. The nephew who was into history kept telling boring stories, and the other one also kept rambling on. Jim didn’t understand one bit and got rather bored. He told the nephew who was into politics that politicians should not systematically neglect delightful and important subjects like sports and girls.
Study 5: Discourse-model referential ambiguity

1.5-referent

At the family get-together, Jim had been talking to one nephew who was very much into politics and another one who was really into history. But Jim himself was only interested in sports, cars, girls etc. The nephew who was into history left early, but the nephew who was into politics kept rambling on. Jim didn’t understand one bit and got rather bored. He told the nephew who was into politics that politicians should not systematically neglect delightful and important subjects like sports and girls.
Study 6: Implicit causality

David praised Hans because he...

had done a great job.

David praised Hans because he...

was proud of him.

David apologized to Hans because he...

had been behaving badly.

David apologized to Hans because he...

had not received an invitation.
Study 6: Implicit causality

Interpersonal verbs differ in the expectations they raise as to which person is the ‘cause’ of the event being described.

This may affect the initial resolution of “he” in the “because”-clause.

David praised Hans because he...

NP2-biased verb (NP2-verb)

David apologized to Hans because he...

NP1-biased verb (NP1-verb)

Interpersonal verbs differ in the expectations they raise as to which person is the ‘cause’ of the event being described.
Or does it? When do readers use this information?

David praised Hans because he...

had done a great job.

Hypothesis 1: During reading of the sentence, as soon as the info is relevant

Hypothesis 2: After reading the sentence, when main and subordinate clauses are being combined (Stewart et al 2000, JML)
Study 6: Implicit causality

Linda *praised* Hans *because* he

Hans *praised* Linda *because* he

**P600 effect**

If the verb-induced referential bias is such that readers expect the because-clause to inform us about NP2 (Linda), then the pronoun “he” should initially be perceived to have the wrong morphosyntactic marking.

Congruent with verb bias (P2)

Incongruent with verb bias (P2)
So implicit causality is brought to bear rapidly enough to affect initial resolution of the pronoun in the “because”-clause.

The resulting referential bias is so strong that the system initially takes a formally correct pronoun as a syntactic error!

In other research (Van Berkum et al., 1999, JML) referential bias can even briefly overrule a syntactic gender agreement constraint.
So what controls this stuff?

• **Nref effects**
  • …get bigger when there is more **referential ambiguity**
  • …reflect **deep ambiguity** at the level of the situation model

• **P600 effects**
  • …that were observed before by **referentially failing** pronouns ("he" in a sentence with 2 women) can also be induced via verb-based implicit causality ("David praised Linda because he")

• Hence, we have at least two ERP effects that can be used to track referential processing during text comprehension
• They testify to the speed and power of referential processing
Overview

(1) Tracking referential processing by means of ERPs
   lots of ERP data

(2) What factors control this stuff?
   more ERP data

(3) Beyond the text
   more ERP data

(4) So what?
   (and where to go next?)
When does the brain ‘index’ the interlocutors?

the text
i.e., what has been told so far

the reader
i.e., who’s taking it in

the speaker
i.e., who is telling it

I think I am pregnant because I feel sick

I think abortion should be forbidden
When does the brain ‘index’ the speaker?

the speaker
i.e., who is telling it

“I want you out!”
When does the brain ‘index’ the speaker?

the speaker
i.e., who is telling it

“I want you out!”
Study 7: When do listeners take the speaker into account?

“I think I am pregnant”

When and how do listeners relate what’s being told to who is telling it?

When do we take into account that somebody is a compulsory liar?

a politician before elections?
a soccer referee?

“I think I am pregnant”
The standard model: 2-step interpretation

Wider theoretical context

Lattner & Friederici, 2003: speaker info is used late

Compute what it really means in context

after ~600 msec

~200-500 msec

I think I am ... pregnant
Wider theoretical context

The alternative: *immediately contextualized* (‘1-step’) interpretation

Compute what it *really* means in context

- the speaker
- the current situation
- the prior discourse
- lexical semantics (‘linguistic word meaning’)

*I think I am ... pregnant*
Study 7: When do listeners take the speaker into account?

(1) speaker-consistent (80 sentences)
   e.g. female voice: “I think I am pregnant because I feel sick”

(2) speaker-inconsistent (80 sentences)
   e.g. male voice: “I think I am pregnant because I feel sick”

Van Berkum, Van den Brink, Kos, Tesink & Hagoort (accepted pending revisions), Journal of Cognitive Neuroscience
Study 7: When do listeners take the speaker into account?

(1) speaker-consistent (80 sentences)
  e.g. female voice: “I think I am pregnant because I feel sick”

(2) speaker-inconsistent (80 sentences)
  e.g. male voice: “I think I am pregnant because I feel sick”
  female voice: “I think I am pregnant because I feel sick”
  male voice: “I think I am pregnant because I feel sick”
  mature voice: “I always like to have a cigar after dinner”
  young voice: “I always like to have a cigar after dinner”
  lower-class accent: “I have a big tattoo on my back”
  upper-class accent: “I have a big tattoo on my back”
Study 7: When do listeners take the speaker into account?

A  Speaker inconsistency effect

male / female: “If only I looked like Britney Spears in her latest video”

upper-/lower-class: “I have a large tattoo on my back”

young child / adult: “Every evening I drink some wine before I go to sleep”

200-300 ms!
The standard model: 2-step interpretation

Compute what it really means in context

Wider theoretical context

Lattner & Friederici, 2003: speaker info is used late

after ~600 msec

~200-500 msec

I think I am ... pregnant
Study 7: *When do listeners take the speaker into account?*

Listeners need only some 200-300 ms to relate what’s being told to who’s telling it...and presumably this partly hinges on resolving the 1st person pronoun

"I want you out!"
When does the brain ‘index’ itself, i.e., the reader?
I think abortion should be forbidden categorically

If my child were homosexual, I’d find this easy to accept

I think the emancipation of women is a positive development

Watching TV to relax is wrong in my eyes

agree

don’t agree
While answering attitude questions, when do people take their attitudes into account? During or after incremental question interpretation?

Suppose the respondent is strongly opposed to abortion and is then reading:

“I think abortion is bad in general.”

vs

“I think abortion is good in general.”

Suppose the respondent is strongly in favor of abortion and is then reading:

“I think abortion is bad in general.”

vs

“I think abortion is good in general.”

differential ERP effect?
I think abortion should be forbidden categorically

agree  don’t agree
Attitude inconsistency effect (pooled) -- prelim data

e.g., I think abortion should be

- forbidden

+ allowed

positivity between 400-700 ms

attitude-consistent word

attitude-inconsistent word
e.g., I think abortion should be forbidden/allowed ...
e.g., I think abortion should be forbidden/allowed ...
e.g., I think abortion should be

allowed

forbidden

Attitude inconsistency effect (pooled)

positivity

between

400-700 ms

attitude-consistent word

attitude-inconsistent word
1. The **differential ERP effect** reveals that attitudes are related to questionnaire statements immediately, on a word-by-word basis, and within 400-700 ms after the relevant word.

   Presumably, this doesn’t critically hinge on “I think abortion should be”

   We predict comparable effects on “Abortion should be”

   But something self-referential is bound to be critical (“**Others think**…”)

2. ERPs can be applied to complex situations where language is used for a purpose
Overview

(1) Tracking referential processing by means of ERPs
  lots of ERP data

(2) What factors control this stuff?
  more ERP data

(3) Beyond the text
  more ERP data

(4) So what?
  (and where to go next?)
When does the brain resolve discourse anaphors?

Once upon a time, a bunch of dwarfs stumbled across…

Immediately!

Snowwhite kissed a dwarf before she left.
When does the brain ‘index’ the interlocutors?

- **The text**: i.e., what has been told so far
- **The speaker**: i.e., who is telling it
- **The reader**: i.e., who’s taking it in

*Immediately!*

- *I think I am pregnant because I feel sick*
- *I think abortion should be forbidden*
As for methods...

- Tracking referential processing via ERPs is feasible/profitable

- There is no **simple** mapping from linguistic levels to ERPs
  - Semantics - **N400 effect**, ...
  - Syntax - **P600 effect**, LAN effect,...
  - Reference - **Nref effect**, ... **P600 effect**??
In sentence comprehension, listeners incrementally build up interconnected structures at multiple levels

The little star
The little star
The little star is beside
The little star is beside a big star
The little star is beside a big star.

phonological unification

syntactic unification
The little star is beside a big star

phonological unification

syntactic unification

semantic unification

situation model building

Spatial structure
The little star is beside a big star.
The little star is beside a big star

phonological unification

syntactic unification

semantic unification

situation model building
The little star is beside a big star
The little star is beside a big star.
The little star is beside a big star.
No simple levels-to-ERPs mapping

The little star is beside a big star

phonological unification

syntactic unification

semantic unification

situation model building

P600
LAN
N400
Nref effect
Study 6: Implicit causality

Linda praised David because he ...
David praised Linda because he ...

bias-con: Linda praised David because...
bias-inc: David praised Linda because...

A pronoun with a morphosyntactic marking that is inconsistent with a referential bias due to verb semantics

P600 effect

Van Berkum, Koornneef, Otten & Nieuwland (in press) Brain Research
The little star is beside a big star

Implicit causality

Syntactic unification

Semantic unification

Referential attractor

Semantic attractor

P600

Phonological unification

Situation model building

Morphology

Syllabic structure

Segmental structure

Structure
If there is tension between levels, where to put the blame?

A strong attractor at level X might cause the system to reanalyze level Y or Z.
As for methods...

- Tracking referential processing via ERPs is feasible/profitable
- There is no **simple** mapping from linguistic levels to ERPs
  - Semantics - N400 effect, ...
  - Syntax - P600 effect, LAN effect, ...
  - Reference - Nref effect, ...
- These experiments aren’t simple (see Van Berkum 2004)
- There is more in the EEG than ERPs!
Spoken pronoun interpretation

1-ref referentially successful (1 suitable referent)
David shot at Linda as he jumped over the fence

2-ref referentially ambiguous (2 suitable referents)
David shot at John as he jumped over the fence

0-ref referentially failing (0 suitable referents)
Anna shot at Linda as he jumped over the fence

1. When do listeners find out that “he” is problematic in 2-ref and 0-ref?
2. Can we replicate various ERP effects observed before?
3. Oscillatory changes associated with referential success, ambiguity & failure?
The ERP data

referentially induced negative shift

referentially induced P600 effect

van Berkum, Zwitserlood, Bastiaansen, Brown, & Hagoort & (2004, CNS)
Referentially successful binding reflected in gamma?
Spoken pronoun interpretation

1-ref referentially successful (1 suitable referent)
David shot at Linda as he jumped over the fence

2-ref referentially ambiguous (2 suitable referents)
David shot at John as he jumped over the fence

0-ref referentially failing (0 suitable referents)
Anna shot at Linda as he jumped over the fence

- Only when the pronoun “he” is referentially **successful** can the surrounding information be **bound to a known discourse entity**...

- **Increased mid-gamma power (40-55 Hz)** may well reflect this!
Finally...

- I haven’t told you how the brain resolves discourse anaphors...

- …because we don’t know yet!

- Neuroimaging provides additional constraints, but it’s not going to solve the issue by itself
  - We need theoretical constraints
  - Behavioral data are privileged (the brain evolved for behavior, not for its capacity to generate ERP or fMRI-BOLD signals…)

- So there’s a lot of work for all of us!
My collaborators

Amsterdam
- Marte Otten
- Mante Nieuwland
- Peter Hagoort
- Colin Brown

Utrecht
- Bregje Holleman
- Dannie van den Brink
- Cathelijne Tesink

Nijmegen
- Arnout Koornneef
- Jaap Murre
- Pienie Zwitserlood
- Miriam Kos

Münster
- Arnout Koornneef