UNIVERSITY OF EXETER

Centre for Research in Writing

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

The Growth in Grammar Corpus: On Working with Children (but not Animals)

Mark Brenchley
Phil Durrant
Debra Myhill

?

Growth in Grammar Project

CONTEXT

- Increasing curricular emphasis on grammatical development
- Not much contemporary evidence to guide this emphasis
- Especially re: quality
- No interest in grammatical "accuracy" (at least for now)
- Grammar approached as a resource for "meaning-making"

Growth in Grammar Project

RESEARCH QUESTIONS

1) How does children's written language change as they get older?
2) How does it change depending on the quality of the writing?
3) How does it change according to the kinds of writing asked of students?
4) Are there differences between how children's texts group "grammatically" and how they are grouped conceptually?

Growth in Grammar Project

THE CORPUS

- Build & (MD-)analyse corpus of "school writing"
- Naturalistic: already produced as part of normal classwork
- England-wide: north/south, rural/urban, PP/¬PP
- 3 curricular areas: English, Science, Humanities (History)
- 4 ages: 6-7, 10-11, 13-14, 15-16
- 3 attainment levels: "weak", "average", "good"
- Not "EAL" (i.e. not not-native speakers!)

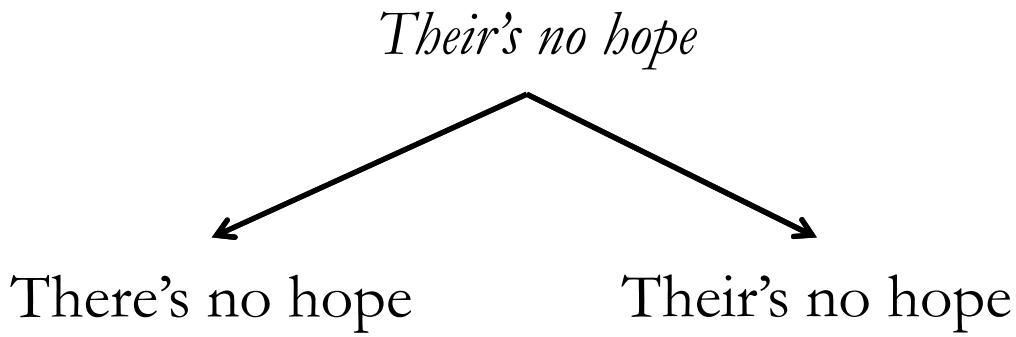
Where We Are Now

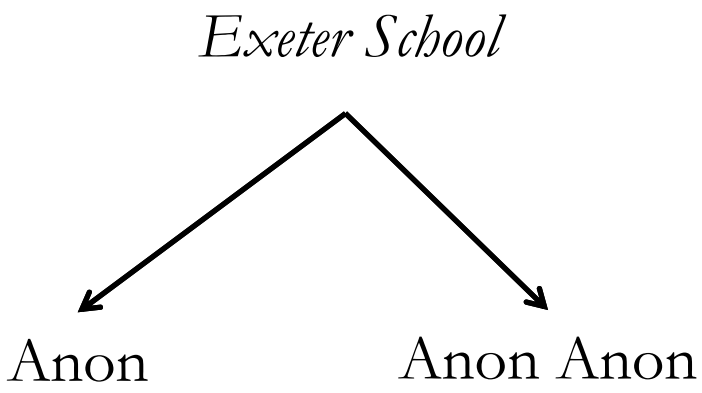
- 1,200-1,500 texts (out of 6,000)
- Mix of primary and secondary
- Good mix of primary genres
- OK mix of secondary genres
- Begun transcription and annotation process

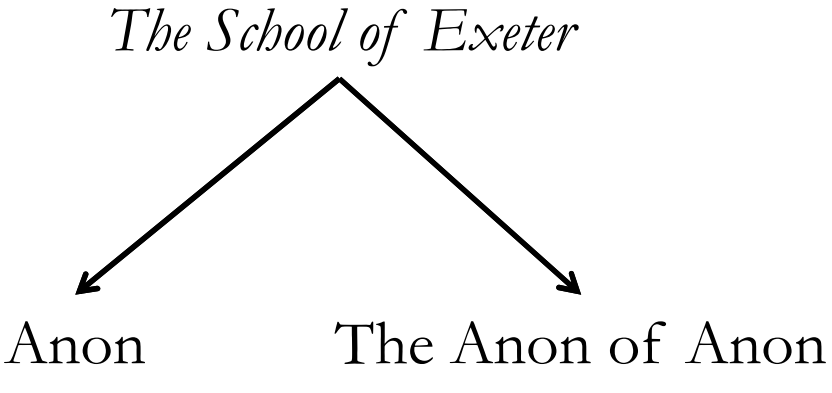
Challenges

- "Authentic" vs. "Relevant" vs. "Ethical"

Challenges


Their's no hope
There's no hope
Their's no hope


Exeter School
Anon
Anon Anon


The School of Exeter
Anon
The Anon of Anon

Challenges

- “Want” vs. “Can”

Challenges

- ? Subordinate Clause Types
- ? Modifiers per NP
- ? Adverbial Placement
- ? Appositional Structures
- ? Finiteness
- ? Subject-Verb Inversions
- ? Depth of Embedding
- ? Ellipsis
- ? Cohesive Ties
- ? Relative Clause Gap Position

Challenges

- "Raw" material (i.e. ¬adult, ¬published, ¬typed)

Challenges

It wasn't just the support of the Vietnamese people which was important, America and France in particular needed support from their citizens. However as people soon realised the Vietnam war was a lost cause and this support gradually slipped.

content

Source 7 shows a girl running naked after having Napalm dropped on her. At the time of the war this was a very famous image to reach the American public and when they saw this they were horrified by what was going on in Vietnam and did not think it right to support such brutality.

context

Source 19 shows an Anti-war demonstration in Washington. Demonstrations like these were not uncommon as death rates rose as well as taxes, people weren't willing to support a war which couldn't be won. This source could be seen as subjective because there only appears to be white women and men of a similar age protesting. Therefore we don't know the general feel of the American Population.

content

typicality

corroborative

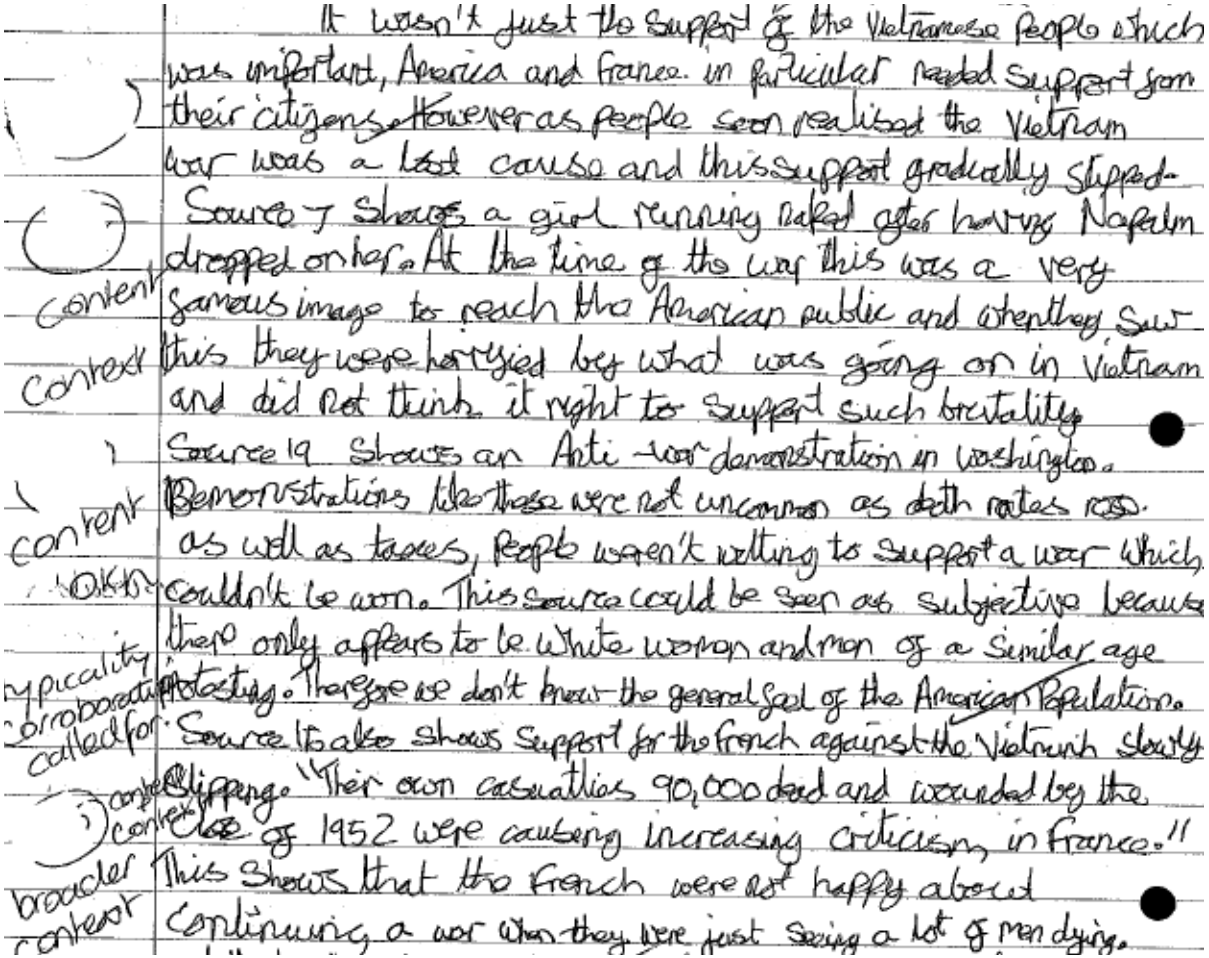
called for

Source 16 also shows support for the French against the Vietminh slowly slipping. "Their own casualties 90,000 dead and wounded by the close of 1952 were causing increasing criticism in France." This shows that the French were not happy about continuing a war when they were just seeing a lot of men dying.

context

context

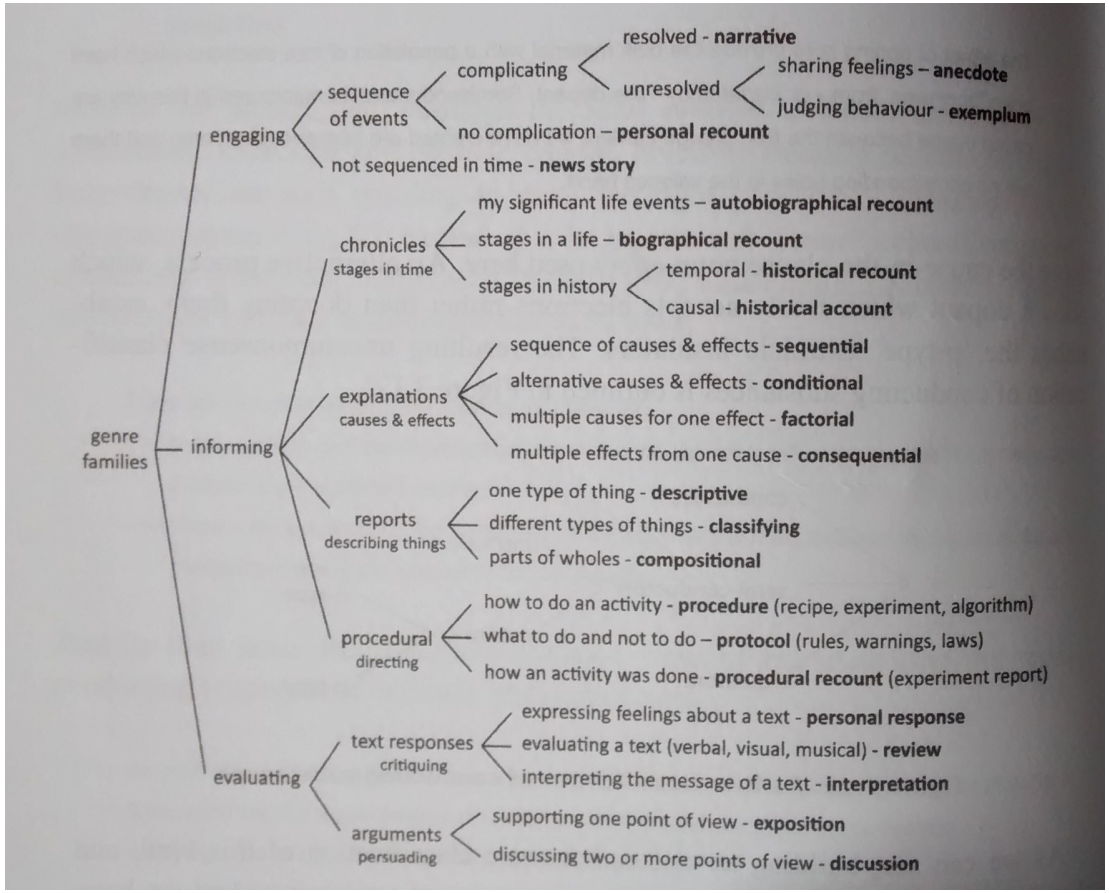
broader context



Challenges

- Material that varies in kind

Challenges



Rose & Martin (2012)
Learning to Write, Reading to Learn

Challenges

- Material that varies in quality

Challenges

It wasn't just the support of the Vietnamese people which was important, America and France in particular needed support from their citizens. However as people soon realised the Vietnam war was a lost cause and this support gradually slipped.

content

Source 7 shows a girl running naked after having Napalm dropped on her. At the time of the war this was a very famous image to reach the American public and when they saw this they were horrified by what was going on in Vietnam and did not think it right to support such brutality.

context

Source 19 shows an Anti-war demonstration in Washington. Demonstrations like these were not uncommon as death rates rose as well as taxes, people weren't willing to support a war which couldn't be won. This source could be seen as subjective because there only appears to be white women and men of a similar age protesting. Therefore we don't know the general feel of the American Population.

content

typicality

corroborative

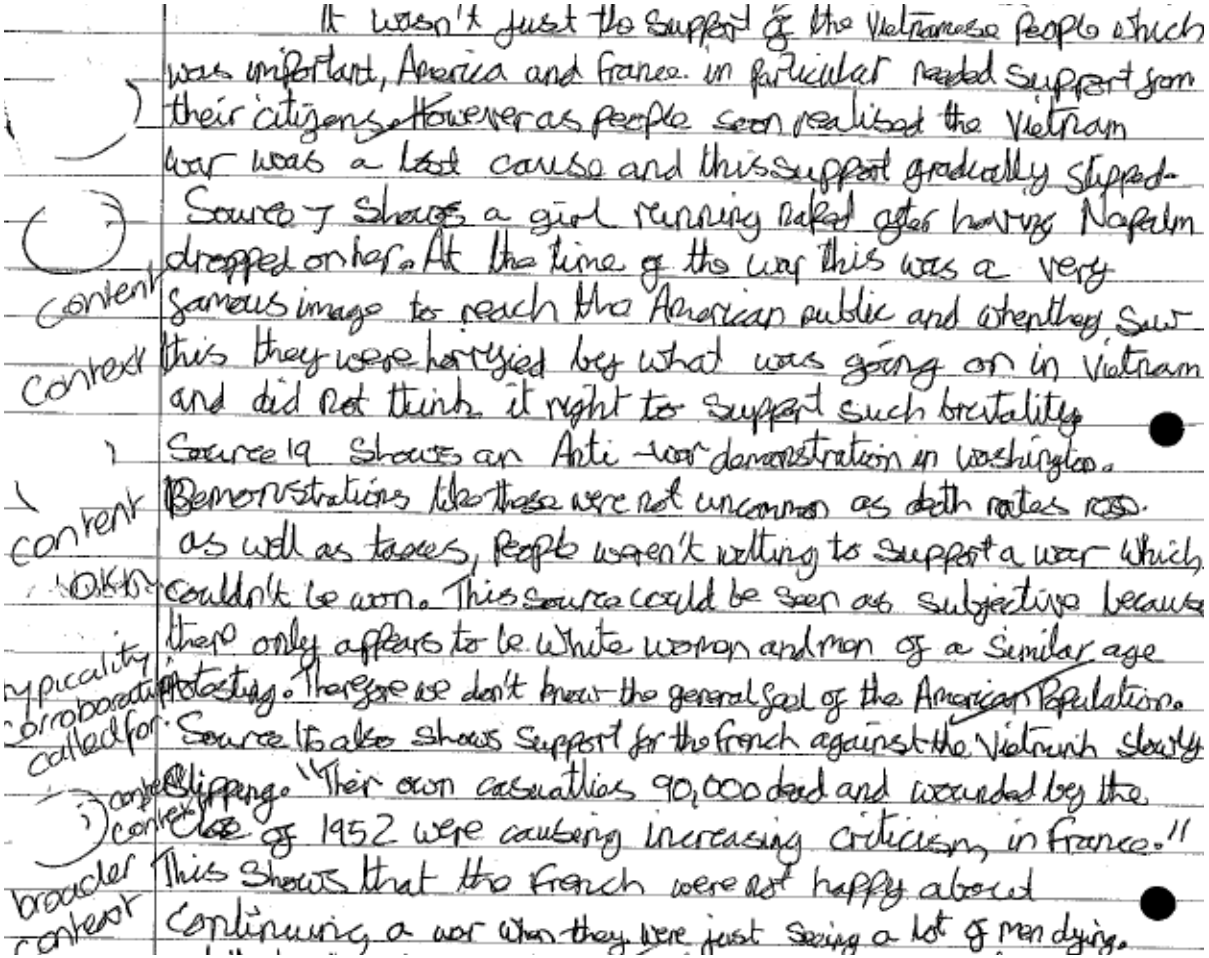
called for

Source 16 also shows support for the French against the Vietminh slowly slipping. "Their own casualties 90,000 dead and wounded by the close of 1952 were causing increasing criticism in France." This shows that the French were not happy about continuing a war when they were just seeing a lot of men dying.

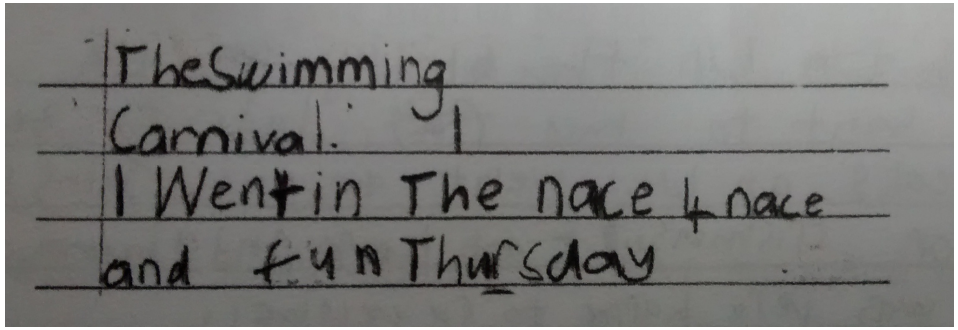
context

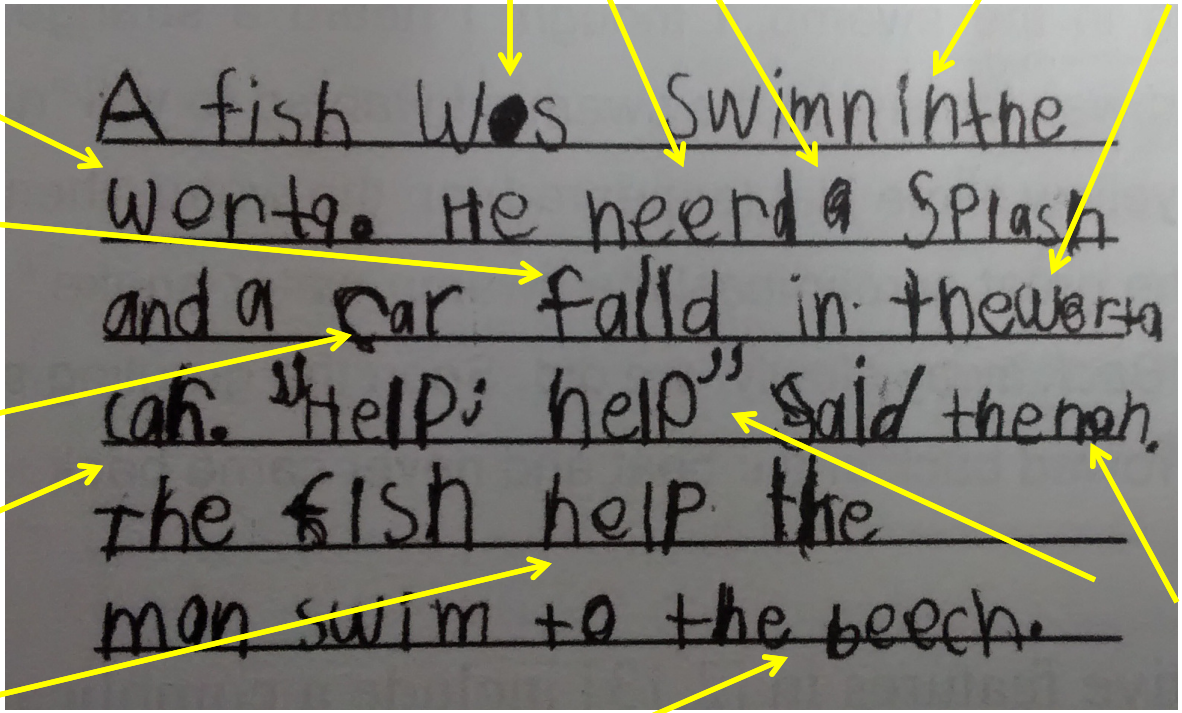
context

broader context



Challenges


TheSwimming
Carnival. I
I Wentin The nace 4 nace
and fun Thursday


A fish Wos Swimnin the
worta. He heerd a Splash
and a car falld in the worta
car. "Help; help" said the man.
The fish help the
man swim to the beech.

Some "Simple" Solutions ?

- Anonymise the whole name:
  - *Exeter School* → Institution_name
  - *The School of Exeter* → Institution_name

- Mark up spelling "errors":
  - *worta* → <sp worta>water</sp>

Some Interim Solutions

- Mark up grammatically "incoherent" stretches:
  - *They went into to town*

    → <garble>They went into to town</garble>

- Mark up verb "errors":
  - *He does loves her*

    → He does <gram>loves</gram> her

Some Less Simple Problems ?

PUNCTUATION

I lost. She won.	→	ROOT; ROOT
I lost, she won.	→	ccomp(won, lost)
I lost she won.	→	ccomp(lost, won)
I lost. But she won.	→	ROOT; ROOT
I lost, but she won.	→	conj(lost, won)
I lost but she won.	→	ccomp(lost, won)
I lied. Then I died.	→	ROOT; ROOT
I lied, then I died.	→	parataxis(lied, died)
I lied then I died.	→	ccomp(lied, died)

Some Less Simple Problems ?

PUNCTUATION

- Piloted full stop insertion after "independent" clauses
- *I lost she won* → I lost<sent>.</sent> She won.

- Definite differences between two versions
- Mostly not critical
- But some are…

Some Less Simple Problems ?

- ***This isn't coming*** *from taxpayers' money either, it is entirely fundraised.*

  → ccomp(fund-raised, coming)

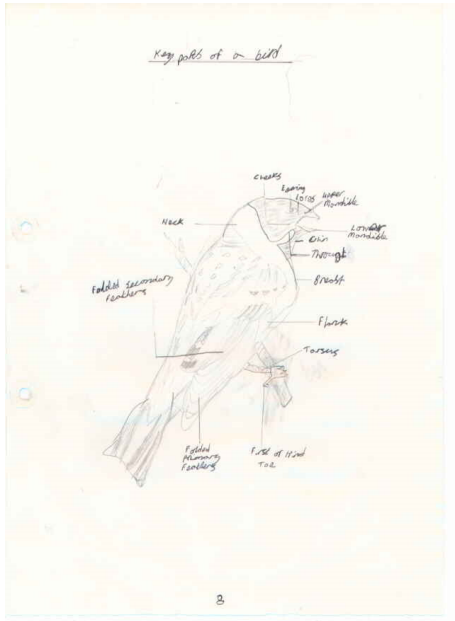
Some Less Simple Problems ?

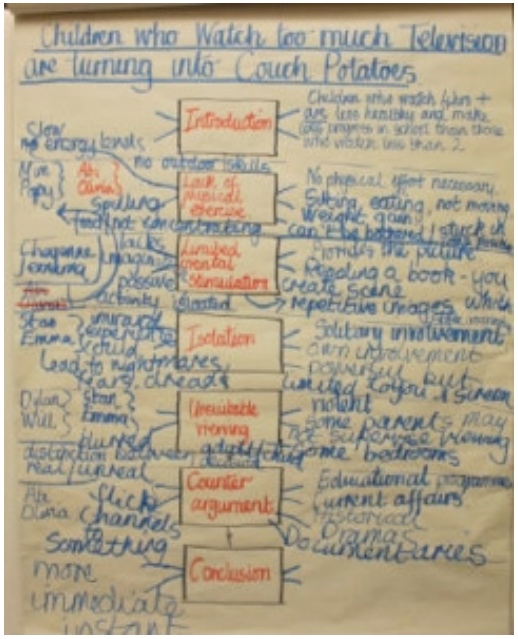
PUNCTUATION

- "Interpretive" Issues
  - I hope your readers remember that travel broadens the mind<sent ,>.</sent> and that trips like these have been proven to work in the past.
  - *I think he's great and she's great.*
  - *I think that's terrible and we should do better.*
- Make our peace with it?

Some Less Simple Problems ?

GRAPHICAL TEXT





Moves	Bottles A	B	C	What we do:
	10	7	3	
1st	10	0	0	Fill the Bottle A.
2nd	7	0	3	Fill the bottle C.
3rd	7	3	0	Empty the index of C ~~to~~ B.
4th	4	3	3	Refill C.
5th	4	6	0	Empty C in B.
6th	1	6	3	Refill C from A.
7th	1	7	2	Refill B from C.
8th	8	0	2	Empty the index of B ~~to~~ in A.
9th	8	2	0	Empty the index of C in B.
10th	5	2	3	Refill C from A.
11th	5	5	0	Empty C in B.

Some Less Simple Problems ?

GRAPHICAL TEXT

- Laborious to transcribe
- Error-prone

```
<table rows="4" cols="4”>
<head>Features of Birds</head>
        <row  role="label">
                <cell/>                                 <cell>Main colour</cell>
                <cell>Has wings</cell>
                <cell>Can fly</cell>
        </row>
        <row>
                <cell role="label">Blackbird</cell>
                <cell>black</cell>
                <cell>yes</cell>
                <cell>yes</cell>
        </row>
        <row>
                <cell role="label">Parrot</cell>
                <cell>loads!</cell>
                <cell>yes</cell>
                <cell>yes</cell>
        </row>
        <row>
                <cell role="label">Penguin</cell>
                <cell>black and white</cell>
<cell>no</cell>
        </row>
        <row>
                <cell role="label">Seagull</cell>
                <cell>white</cell>
                <cell>yes</cell>
                <cell>yes</cell>
        </row>
</table>
```

	Main Colour	Has wings	Can fly
Blackbird	black	yes	yes
Parrot	loads!	yes	yes
Penguin	black and white	yes	no
Seagull	white	yes	yes

<cell>yes</cell>

Some Less Simple Problems ?

GRAPHICAL TEXT

- Laborious to transcribe
- Error-prone
- Grammatically Awkward

Some Less Simple Problems ?

- *folded secondary feathers*
  root(ROOT, folded-VBN)
  dobj(folded, feathers)
- *twitching ears*
  root(ROOT, twitching-VBG)
  dobj(twitching, ears)
- *lower beak*
  root(ROOT, lower-JJR)
  dep(lower, beak)
- *lower beak*
  nsubj(beak, lower)
  root(ROOT, beak)

Some Less Simple Problems ?

"CHARACTERISTIC" FEATURES

- Need to track the grammatical bases of writing development
- Requires identifying two characteristic features:
    1) Different types of discourse – the "genres" of school writing
    2) General later language development, especially re: "quality"
- Many such structures not all that straightforward

Some Less Simple Problems ?

FORMULAIC MARKERS

- *Dear X,*
- *Yours sincerely/faithfully, Y.*

Some Less Simple Problems ?

- *Yours sincerely/faithfully*

root(ROOT, yours-PRP$)
advmod(yours, sincerely/faithfully)

- *yours sincerely/faithfully*

Some Less Simple Problems ?

- *Dear Editor,*
  compound(Editor-NNP, Dear-NNP)
  nsubj(MAIN CLAUSE, Editor)
- *Dear Sir,*
  compound(Sir-NNP, Dear-NNP)
  nsubj(COMP. CLAUSE, Sir)
- *Dear Sir or madam,*
  compound(Sir-NNP, Dear-NNP)
  nmod:tmod(MAIN CLAUSE, Sir)
- *dear editor/ sir/ sir or madam,*
  amod(sir-NN, dear-RB)
  ccomp(MAIN CLAUSE, editor/sir)

Some Less Simple Problems ?

ISOLATED NPs

- *folded secondary feathers*
- *twitching ears*
- *lower beak*

- *Clouds of dust as blinding as fog and the sound of animal roars dancing around the arena.*
- *The sound of two strong, sturdy, swords clashing together.*
- *The sound of the gladiators, declaring war on each other*

Some Less Simple Problems ?

ISOLATED NPs

- Potentially characteristic of:
  - narrative fiction
  - scientific descriptions
  - poetry
  - ?
  - "sophisticated" writers

Some Less Simple Problems ?

- *folded secondary feathers*

  root(ROOT, folded-VBN)
  dobj(folded, feathers)

- *twitching ears*

  root(ROOT, twitching-VBG)
  dobj(twitching, ears)

- *lower beak*

  root(ROOT, lower-JJR)
  dep(lower, beak)

Some Less Simple Problems ?

- *Clouds of dust as blinding as fog and the sound of animal roars dancing around the arena.*

  nsubj(roars-VBZ, clouds)
  root(ROOT, roars-VBZ)
  xcomp(roars-VBZ, dancing)

- *The sound of two strong, sturdy, swords clashing together.*

  amod(sturdy, strong)
  amod(sound, sturdy)
  dep(sound, clashing)

- *The sound of the gladiators, declaring war on each other.*

  nsubj(declaring, sound)
  root(ROOT, declaring)
  root(ROOT, sound)
  acl(gladiators, declaring)

Some Less Simple Problems ?

"DISPLACED" AdjPs

- *The beast, monstrous, ravenous, roamed the house.*
- *He's a great student, dedicated, hardworking and ambitious.*
- *He is a terrible student – nasty, lazy, stupid.*
- *Monstrous, ravenous, the beast roamed the house.*
- *The beast roamed the house, monstrous, ravenous.*
- *Green, bronze, and golden, it flowed through the weeds.*
- *John chuckled, highly amused.*

Some Less Simple Problems ?

"DISPLACED" AdjPs

- Potentially characteristic of:
  - Fiction
  - ?
  - "sophisticated" writers

Some Less Simple Problems ?

- *The beast, monstrous, ravenous, roamed the house.*

  appos(beast, monstrous)
  appos(monstrous, ravenous)

- *Monstrous, ravenous, the beast roamed the house.*

  nsubj(roamed, monstrous—JJ)
  appos(monstrous, ravenous)
  appos(ravenous, beast)

- *The beast roamed the house, monstrous, ravenous.*

  nsubj(ravenous, house)
  appos(house, monstrous)
  xcomp(roamed, ravenous)

Some Less Simple Problems ?

- *Green, bronze, and golden, it flowed through the weeds*

  dep(flowed, Green-NNP)
  conj(Green-NNP, bronze-NN)
  conj(Green-NNP, golden)

- *John chuckled, highly amused.*

  xcomp(chuckled, amused-VBN)

- *He's a great student, dedicated, hard-working and ambitious.*

  acl(student, dedicated-VBN)
  xcomp(dedicated, hardworking-VBG)
  conj(hardworking-VBG, ambitious)

- *He is a terrible student, nasty, lazy, stupid.*

  amod(stupid, nasty-JJ)
  amod(stupid, lazy-JJ)
  amod(student, stupid-JJ)

Some Less Simple Problems ?

- Maybe not all that much of a surprise – what you'd expect when working with a highly variable, even "deviant" corpus
- And maybe we can't "count" these more problematic features
- And maybe that's not a major problem -
- Perhaps too sparse for substantive, reliable counts anyway
- BUT

Some Less Simple Problems ?

- This isn't something we yet know, which raises two issues:
- First, even if they aren't pervasive across the corpus generally, they might be pervasive for certain *kinds* of texts
    - Science reports
    - High level science reports
- In which case, we will lose our capacity to pick up on some core developmental differences
- Perhaps even be *the* core differences

Some Less Simple Problems ?

- Second, the annotations marking these more obviously problematic features also implicated in other features that we might like to measure
- But suppose we can't reliably separate out these annotations when they the mark the problematic features from when they mark these other features
- Then these other features also become compromised
- In which case, we'll lose even more of our capacity to pick out those grammatical features that really count in the development of school writing

?

UNIVERSITY OF EXETER

Centre for Research in Writing

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

http://socialsciences.exeter.ac.uk/education/research/centres/centreforresearchinwriting/projects/growthingrammar/