

# First Steps in Marking Up a Corpus of Cryptic Crosswords

Bethan Tovey

1. A Brief Introduction to Cryptic Crosswords
2. The Cryptic Crossword Corpus - Why?
3. The Cryptic Crossword Corpus - What?
4. The Cryptic Crossword Corpus - How?
5. Progress and Next Steps

# 1. A Brief Introduction to Cryptic Crosswords

9

# Capital city (6)

--	--	--	--	--	--

# Capital city (6)

L O N D O N

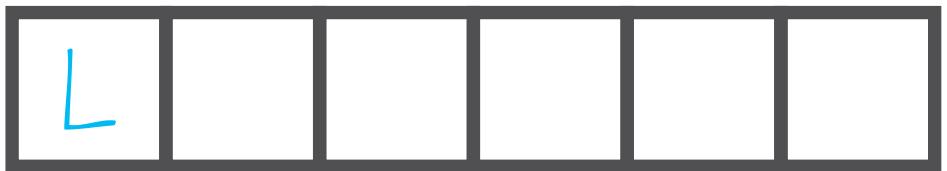
**ONE DOES NOT SIMPLY  
DO THE QUICK CROSSWORD**

[imgflip.com](http://imgflip.com)

Left Ontario with DeLillo  
for UK city (6)



9  
10  
**Left** Ontario with DeLillo  
for UK city (6)



9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50

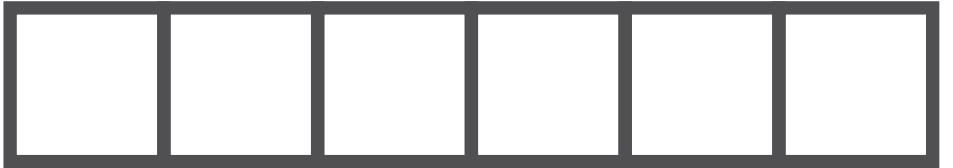
Left **Ontario** with DeLillo  
for UK city (6)



9  
10  
**Left Ontario with DeLillo**  
for UK city (6)

L	O	N	D	O	N
---	---	---	---	---	---

City bird around nine doves,  
initially (6)



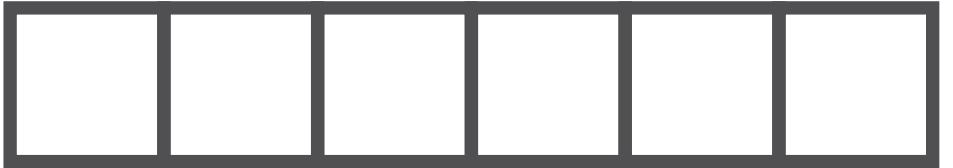
City **bird** around nine doves,  
initially (6)

L O      O N

City **bird** around **nine doves**,  
initially (6)

L O N D O N

Melon donut in possession  
of Julie (6)



Melon donut in possession  
of Julie (6)

L	O	N	D	O	N
---	---	---	---	---	---

Melon donut in possession of  
Julie (6)

City bird around nine doves,  
initially (6)

## Definition - Source - Metalanguage

City bird around nine doves, initially  
City **bird** around **nine doves**, initially  
City bird **around** nine doves, **initially**

## 2. The Cryptic Crossword Corpus - Why?

The simple answer: Linguistics

LINGUISTICS

ME

REAL  
LIFE



The complicated answer: Linguistics

“the cryptic crossword can provide excellent source material for illuminating many aspects of how a language can be operated as a creative system by its users.”

(Cleary 1996, p. 14)

Cryptic clues “bend or break a vast range of linguistic rules. They do so, however, in a controlled fashion and with the awareness of the reader. [...] Misunderstanding, in the sense of arriving at the wrong solution, is very rarely a problem.”

(Coffey 1998, p. 18)

The language of cryptic clues “is shaped by two diametric requirements: it must, on the one hand, encrypt the answers, yet, on the other hand, point to them unambiguously.”

(Pham 2016, p. 121)

“La communication cruciverbiste n'est pas *apoétique*, mais *anti-poétique* [...] Ainsi, sa lecture se propose, à partir du sens, à obtenir du non-sens, à partir de la manifestation libre que revèle l'inventaire inarticulé de ses définitions, à retrouver un code contraignant.”

(Greimas 1967, p. 800)

“Cruciverbalist communication is not *apoetic*, but *anti-poetic*. [...] Thus, the reading of it intends to obtain nonsense from meaning; to recover a restrictive code from the free expression that reveals the inarticulate inventory of its definitions.”

(Greimas 1967, p. 800)

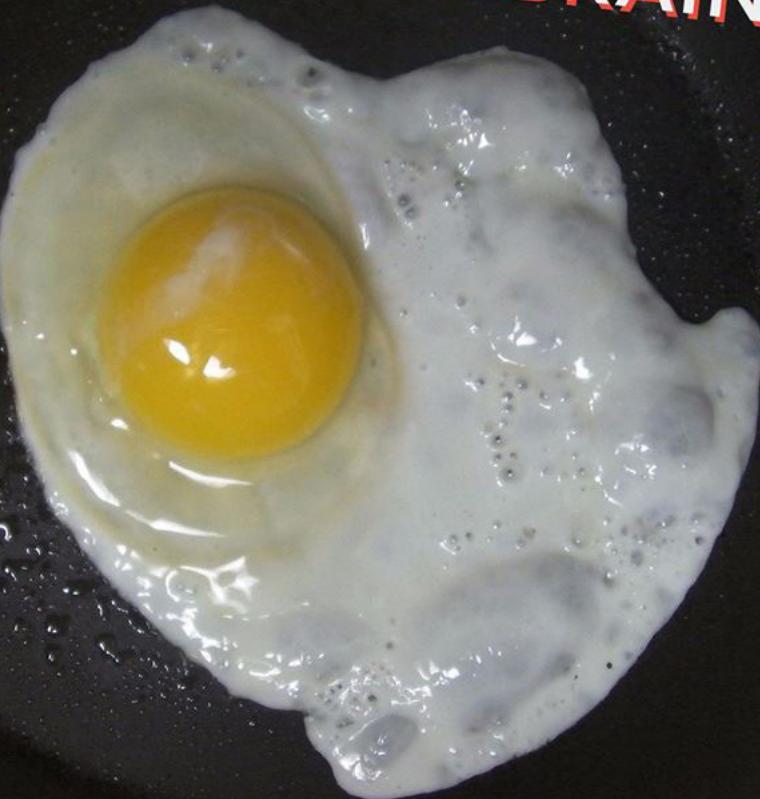
9

6

18

**STOP FRYING YOUR BRAIN**

THIS IS YOUR  
BRAIN ON  
SEMIOTICS



- What are the linguistic differences between cryptic clues and other comparable genres?
- Is clue difficulty related to identifiable linguistic features?
- Do individual crossword compilers have a linguistic “signature”?
- Which metalinguistic terms are used most commonly in which newspapers?
- Do anagrams require the use of less-common words than other clue types?

### 3. The Cryptic Crossword Corpus - What?

Feeble type adopts the French headdress

clue

subsidiary indication

definition

Feeble type adopts the French headdress

clue

subsidiary indication

definition

**Feeble type** adopts **the** French headdress

source

source

clue

subsidiary indication definition

Feeble type adopts the French headress

source

meta

source

meta

clue

subsidiary indication

definition

**Feeble type** **adopts** the **French** headdress

source  
(synonym)

meta  
(locator)

source  
(translation)

meta  
(translation)

clue

subsidiary indication

definition

Feeble type **adopts** the **French** headdress

Publication: The Independent  
Setter: Aelred  
Source: The Independent Book  
of Cryptic Crosswords 1(2005)

WIMPLE

answer

Feeble type adopts the French headdress

adjective

noun

verb

determiner

adjective

noun

WIMPLE

noun

## 4. The Cryptic Crossword Corpus - How?

```
<publication pub="independent" type="anthology" ISBN="0550101756" pubdate="2005">
    <puzzle id="1" setter="Aelred">
        <item id="4">
            <clue>
                <subsidiary>
                    <source class="def" id="1">
                        <word pos="JJ">Feeble</word>
                        <word pos="NN">type</word>
                    </source>
                    <meta class="locator" type="concatenation">
                        <word pos="VVZ">adopts</word>
                    </meta>
                    <source class="translation" id="2">
                        <word pos="DT">the</word>
                    </source>
                    <meta class="operator" type="translation">
                        <word pos="JJ">French</word>
                    </meta>
                </subsidiary>
                <def type="hypernym">
                    <word pos="NN">headdress</word>
                </def>
            </clue>
            <solution words="1" letters="6" text="wimple" pos="NN">
                <unit id="1">WIMP</unit>
                <unit id="2">LE</unit>
            </solution>
        </item> [...]
    </puzzle> [...]
</publication>
```

<publication  
pub="independent"  
type="anthology"  
ISBN="0550101756"  
pubdate="2005">

<puzzle id="1"  
setter="Aelred">

definition subsidiary indication

```
<clue>
  <subsidiary>
    <source class="synonym" id="1">
      <word pos="JJ">Feeble</word>
      <word pos="NN">type</word>
    </source>
    <meta class="locator" type="concatenation">
      <word pos="VVZ">adopts</word>
    </meta>
    <source class="translation" id="2">
      <word pos="DT">the</word>
    </source>
    <meta class="operator" type="translation">
      <word pos="JJ">French</word>
    </meta>
  </subsidiary>
  <def type="hypernym">
    <word pos="NN">headdress</word>
  </def>
</clue>
```

```
<solution words="1"  
letters="6" text="wimple"  
pos="NN">  
    <unit id="1">WIMP</unit>  
    <unit id="2">LE</unit>  
</solution>
```

```
<clue>
  <def type="descriptive">Irish leader</def>
  <subsidiary>
    <source class="anagram" id="1">is a cheat</source>
    <punct>,</punct>
    <meta class="operator" type="anagram">worried</meta>
    <meta class="locator" type="container">about</meta>
    <source class="abbreviation" id="2">nothing</source>
  </subsidiary>
</clue>
<solution words="1" letters="9" text="Taoiseach" pos="NP">
  <unit id="1">TA<unit id="2">O</unit>ISEARCH</unit>
</solution>
```

Source: The Independent

9

```
<clue>
  <quick>On fire</quick>
  <subsidiary>
    <source class="literal" id="1">A</source>
    <source class="synonym" id="2">fair</source>
  </subsidiary>
  <def type="synonym">land</def>
</clue>
<solution words="1" letters="6" text="alight"
pos="VV">
  <unit id="1">A</unit><unit id="2">LIGHT</unit>
</solution>
```

6

Source: The S\*n

## 5. Progress and Next Steps

9

```
<clue>
  <def type="synonym">Guess</def>
  <def type="descriptive">there's no
    proof for it</def>
</clue>
<solution words="1" letters="10"
  text="conjecture" pos="ambiguous">
  <unit>CONJECTURE</unit>
</solution>
```

6

Source: The D\*ily M\*il

9

```
<clue>
    <def type="synonym">Alert</def>
    <def type="descriptive">goalkeeper may
dive thus</def>
</clue>
<solution words="3" letters="9" text="on
the ball" pos="PHRASE">
    <unit>ON THE BALL</unit>
</solution>
```

6

Source: The Times

9  
<clue>

  <def type="oblique">Academic stream</def>

  <subsidiary>

    <source class="literal" id="1">is</source>

    <meta type="locator">beside</meta>

    <source class="ref\_intra\_id1" id="2">itself</source>

  </subsidiary>

</clue>

<solution words="1" letters="4" text="Isis" pos="NP">

  <unit id="1">IS</unit>

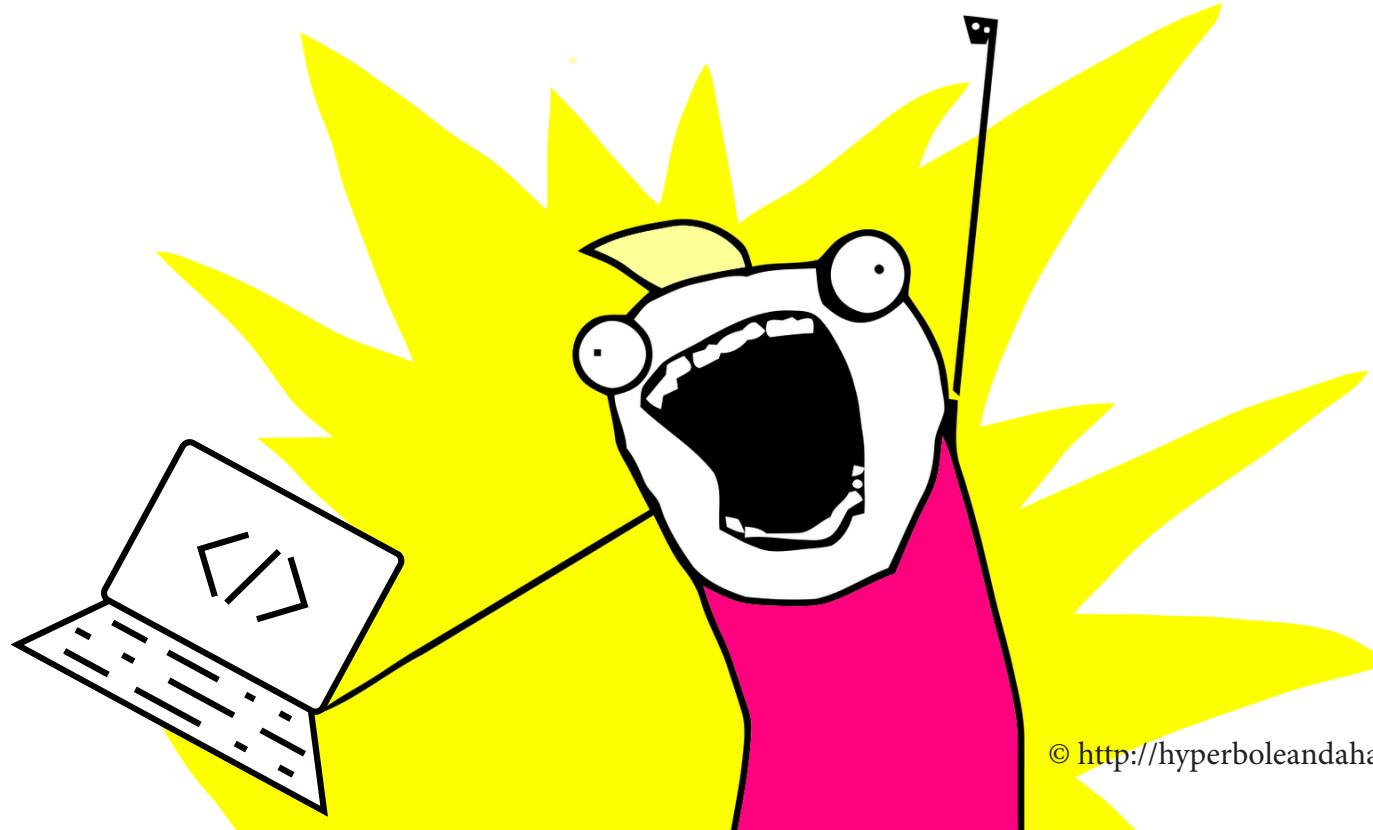
  <unit id="2">IS</unit>

</solution>

Source: The Times

9

Mark up ALL THE THINGS!

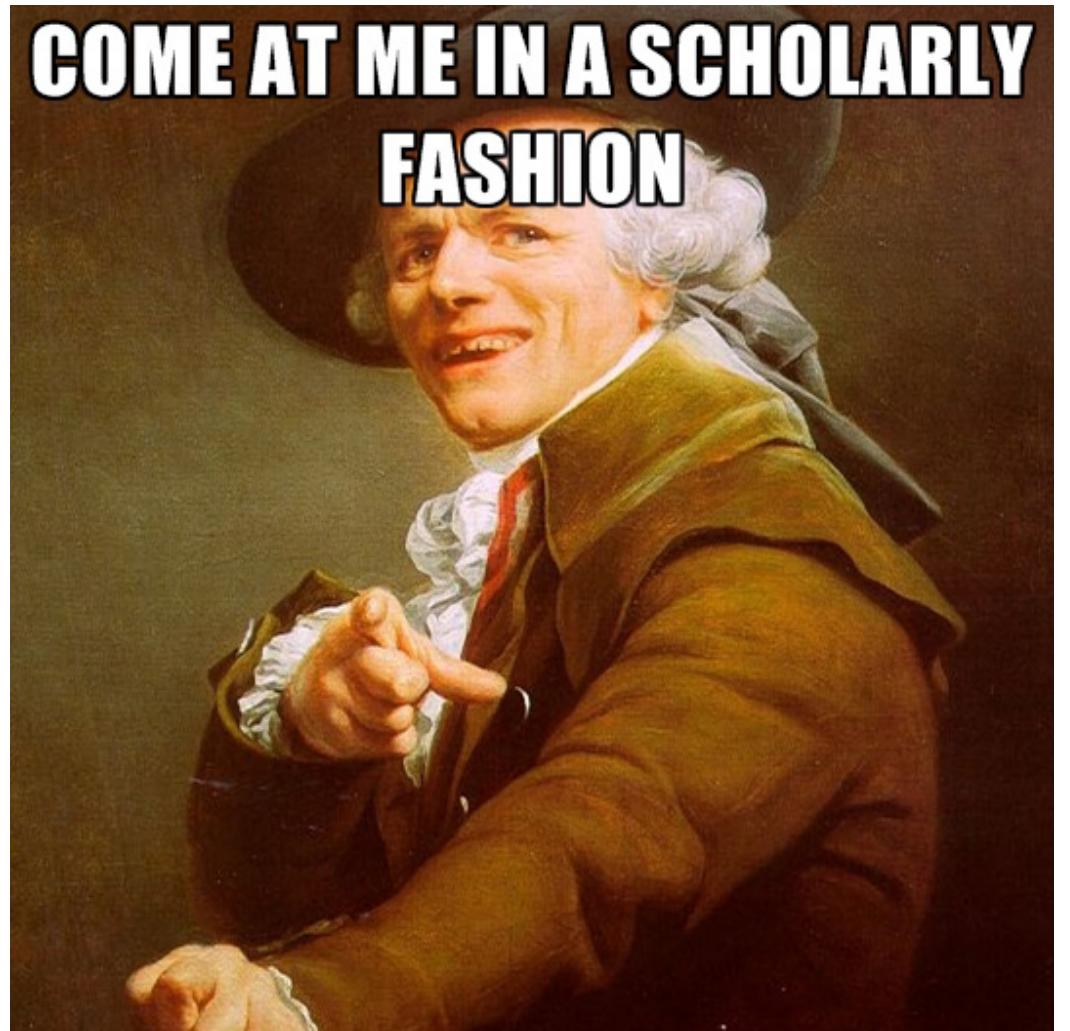


© <http://hyperboleandahalf.blogspot.com/>

6

18

49 of 50



# Questions?

Or you can find me at:



bethan@linguacelta.com



@linguacelta



[www.linkedin.com/in/linguacelta](https://www.linkedin.com/in/linguacelta)