RESEARCH   **Open Access**

'A man who revels in his own ignorance, racism and misogyny': Identifiable referents trump indefinite grammar

Katy Jones

Correspondence: Jonesks1@cardiff.ac.uk
School of English, Communication and Philosophy, Cardiff University, John Percival Building, Colum Drive, Cardiff CF10 3EU, UK

Abstract

Typically, a noun phrase beginning with the indefinite article introduces a referent assumed to be unknown to the addressee. But in newspaper opinion journalism, this is not always the case. In 'instead of hailing its first female president, it [the US] seems poised to hand the awesome power of its highest office to *a man who revels in his own ignorance, racism and misogyny*' (*The Guardian*, 9/11/16), 'a man who…' *can* be understood as a new referent or type. But once seen in context, where the identity of the man is known, it becomes clear that it is signalling something different. This paper examines how this sort of reference works by challenging existing accounts of 'late' indefinites and the meaning relation of co-extension. It is shown that lexical cohesive ties between the expression and preceding text and context create a shared space which allows these expressions to function 'definitely'.

Keywords: (In)definite reference, Identifiability, Lexical cohesion, Co-extension, Similarity chains

Introduction

While much attention has been devoted to conventional referring and (in)definiteness, i.e. the use of an indefinite expression for something the addressee is not expected to be able to identify and a definite expression to refer to something identifiable (e.g. Hawkins 1978; Ariel 1988,1990; Givón 1983,1993a,1993b; Gundel, Hedberg and Zacharski 1993; Cornish 1999, 2010; Schiffrin 2006; Abbott 2010), atypical uses of referring expressions remain largely unexplored. In fact, the conventional rules of reference are often exploited for effect, as the indefinite expression in [1], from *The Guardian* opinion article *David Beckham: how this crock of a footballer can still woo the French*, demonstrates:

[1] It seems that France, like Spain and the United States before it, is poised to be charmed by *a man who, with his un-British attention to grooming, muscle tone and non-novelty underwear, may become an honorary Frenchman before his six months in Paris are up* (01/02/13)[1]

In principle, the use of the indefinite article signifies that the writer does not expect the reader to be able to identify the man with these very specific traits. However, in reality, a reader with an interest in British football (and we can assume that someone

reading an article about David Beckham does) would be able to identify the man referred to in [1] as David Beckham, especially if we consider that the expression is part of an on-going narrative about Beckham, where reference has been made to him 18 times previously. This process of identification perhaps works because, in the spirit of cooperation, the speaker intends to convey marked meaning by means of an implicature (Grice 1975), which needs to be inferred by the addressee.

The use of the indefinite article in such expressions could be to indicate that the text producer is talking about a type of entity with particular qualities. This use can be clearly seen in the italicised expression in [2], from an opinion piece entitled *Why is Theresa May still in No 10? Because the Tories need a human shield.*

[2] Why is she still here? The Tories, after all, have a well-publicised talent for regicide: *a leader who has failed is generally found standing in front of No 10 or Tory HQ within 24 hours, making a statement to the press.* John Major, William Hague, David Cameron: all departed in this way. (*The Guardian* 20/06/17)

Example [2], has the structure A(n) + NOUN+RESTRICTIVE RELATIVE CLAUSE (A(n) + N + RRC forthwith). This expression is undoubtedly referring to a type of leader, made clear by the use of the adverb *generally* to introduce what typically happens in this scenario. The subsequent sentence then provides instances of the type (i.e. a list of leaders with the qualities presented in the relative clause). The A(n) + N + RRC expressions in [3], [4] and [5] however, appears to be signalling something different:

Today the United States stands not as a source of inspiration to the rest of the world but as a source of fear. Instead of hailing its first female president, it seems poised to hand the awesome power of its highest office to [3] *a man who revels in his own ignorance, racism and misogyny....* [4] *A man with no control of his impulses* will be unrestrained, the might of a superpower at the service of his ego and his id. (*The Guardian* 09/11/2016)

But you have to watch out for those Trumpites who pop up to call you "fake news" and who frighten radio station editors. The media's continuing respect for "fair play" when discussing [5] *a president who is self-evidently a dangerous and racist xenophobe (as opposed, for example, to the Arab variety)* should one day be examined. (*The Independent* 20/01/18)

The indefinite expressions in [3], [4] and [5] are singling out a particular man (i.e. Donald Trump), one who is readily identifiable in the given context and who has been mentioned directly several times in the preceding texts (e.g. in [3]: *the new president elect, Trump, he*; in [4]: *Donald Trump, the booby who thinks he's running the United States*). The chain of reference to the individual is carried on in the subsequent co-text (e.g. [3] *his impulses, his ego, his id*; [4] *Trump, this infantile person*), reinforcing the argument that the indefinite expression is signalling Trump. Also, in each of these examples, it would be felicitous (semantically and syntactically, at least) to replace the

indefinite expression with a proper name. Clearly, this usage is not simply an example of typical indefinite reference.

In a study based on a corpus of 40 journalistic opinion articles, Jones (2014) identified several key features of this particular use of the A(n) + N + RRC expression, which occurs in non-initial position in a reference or identity chain (c.f. Martin 1992; Halliday 1994). Jones (2014) carried out a reader interpretation experiment, which showed that readers largely do not interpret the expression (when seen in context) as referring to a type or a newly introduced referent as convention would suggest, but rather to the previously mentioned, fully-identified entity. The results of this experiment also showed that the amount and level of detail of lexical information in the relative clause plays a role in how the expressions are interpreted. That is, the more lexical detail in the relative clause, the more likely it is that the reader interprets the expression as referring to the definite referent. For example, readers were almost equally divided about whether the expression in [6] referred to the specific MP, Nadine Dorries, or a type of MP with those qualities, but almost all readers interpreted the expression in [7] as referring to the referent, Louis Kahn.

[6] an MP who can spread such inaccuracies

[7] a man who died in a public lavatory in a low-grade public building, whose corpse lay unrecognised in a New York City morgue for three days, and who flitted from one family home to another

Further, in an analysis of the expressions' Accessibility[2] (Ariel, 1988, 1990) using Toole's (1996) framework, Jones (2014) found that 93% of the expressions in her corpus achieved a high or mid degree of Accessibility, making these expressions comparable to demonstratives, pronouns and zero anaphora in terms of how accessible the text producer assumes the referent is in the mind of the reader. What this means is the A(n) + N + RRC expressions in Jones' (2014) study are not produced as if the referents are new in the discourse, but rather given and highly accessible.

So the question is how do these formally indefinite expressions become functionally identifiable within their specific context of British English journalistic opinion writing? The aim of the remainder of this paper is to answer this question.

This paper examines the use of indefinite referring expressions, such as those in [1, 3, 4, 5, 6, 7] in their context, and considers how the writer creates an adequately 'definite' shared space with the reader to allow for an indefinite expression to be understood as referring to the established entity. In the section on "Referring and (in)definiteness" I discuss the notions of reference and (in)definiteness. In the "Co-extensional analyses" section, I challenge and expand upon existing accounts of 'late' or 'second-mention' indefinites, as well as Hasan's (1985) co-extensional framework. I then exploit the notions of co-extension and similarity chains (Hasan 1985) in the "Discussion" section to analyse the cohesive ties in four journalistic opinion articles and show that sufficient cotextual and contextual scaffolding is put in place by the writer for the indefinite expression to be understood as referring to the identifiable entity. Finally, in the "Conclusion", thought will be given to the

reasons for this switch from definite to indefinite reference, when we would expect definiteness to be maintained through the use of a definite expression.

Referring and (in)definiteness

Reference can be seen as a four-way relation between a speaker using an expression to identify an entity so that the addressee is able to recognise the entity in question. Under this conception, we can assume that reference is largely a pragmatic phenomenon as it concerns the speaker's use of linguistic expressions (Abbott 2010:2). Linguistic, cognitive, psycholinguistic and discourse factors contribute to the speaker's choice of the form of the linguistic item selected to carry out this task, as well as to the addressee's interpretation of the expression.

Conventionally, only structurally 'definite' expressions are used to refer to entities that are considered to be identifiable to the addressee (in English anyway) (Abbott 2010; Givón 1993a; Fawcett 1980; Hawkins 1978; Halliday and Hasan 1976). In English, these expressions can be realised by a NP with a definite or demonstrative determiner (*the/this apple*), pronoun (*it*) or proper noun (*Bardsey Island Apple*).

Indefinite expressions are assumed to be used for referents which the addressee is not expected to be able to identity (Halliday and Hasan 1976; Givón 1993b; Martin 1992; Hawkins 1978) and are typically realised by NPs with indefinite determiners and pronouns (e.g. *an apple*; *some apples*; *any apples; any*). It is generally accepted that the conventional pattern of article use is indefinite first mention to introduce the referent into the discourse, with subsequent mentions being signalled by a definite marker. For example:

> [8] *A German policewoman* was shot in the head when a man grabbed a police gun
> at a suburban station in Munich. *The woman*, 26, was critically wounded …
> (*BBC* 13/06/17)

It is also acknowledged that there are exceptions to this pattern (e.g. Givón 1993a:235; Du Bois 1980:207; Martin 1992:94; Jones 2014), as we shall see below.

Du Bois (1980:219) suggests that definiteness involves "a tracing of the constant idea (referent) through links with the shifting words (references) used to refer to the idea", which speakers have control over. What is interesting here though, is that even if the identity of an object is known to both interlocutors, they are under no obligation to use the available definite expression. For example, a mother says to her small child *someone left their bike out in the rain last night*: Both speaker and addressee (the culpable 'someone') are well-aware of the fully identifiable culprit, but for pragmatic purposes, the speaker chooses not to explicitly identify them and uses an indefinite expression. This obfuscation is done, however, while remaining in the spirit of cooperation; that is, the speaker is well aware that the addressee is able to identify the referent. In Gricean terms, the speaker is flouting the maxim of quantity, i.e. make your contribution as informative as required; in this case, the speaker is not being as informative as required (which would involve explicit identification of the wrongdoer), but the addressee, assuming speaker cooperation, is able to work out the implied meaning.

So, even though the conditions for the use of a definite expression are in place, the speaker can choose whether to oblige. Du Bois (1980:219) concludes that "speakers have facultative control of definiteness". This conclusion is echoed by Fries (2001:83):

> Of course, speakers always have the **choice** of how to present new information so while information which is presented as structurally New is usually new in fact to the listener, individual speakers may choose to present as New, information which is obvious to the listener. [emphasis in the original]

The notion of choice is crucial here; ultimately is it under the speaker's control *how* a referent is presented. Speakers make a choice about how they wish to depict the referent, and have to ensure that it fits with the assumed discourse representation held by the addressee. These claims suggest there are other parameters beyond identifiability which govern the choice of an (in)definite expression. Indeed, as Fries (2001:87) points out:

> [W]e regularly choose the wordings of our nominal groups (including Heads and Modifiers) to establish features of the referents which are relevant to the discourse.

The 'wordings' of referring expressions are not always necessarily intended to assist in the identification of the referent, but rather to fulfil a discourse function. Several scholars have examined this idea. For example, Vonk, Hustinx and Simons (1992:303) showed that referring expressions which are more specific than necessary for identification of the antecedent indicate "an episode boundary". That is, when an expression is used that is more specific than needed for identification, it also has a discourse structuring function. They use the following extract [9] as an example (p303):

1. Sally Jones got up early this morning.
2. She wanted to clean the house.
3. Her parents were coming to visit her.
4. She was looking forward to seeing them.
5. **She** weighs 80 kg.
6. She had to lose weight on her doctor's advice.
7. So she planned to cook a nice but sober meal.

The use of the pronoun *she* in line 5 may not cause any identification problems, but certainly disrupts the coherence of the passage and sounds a little odd. In fact, as Vonk et al. point out, the more specific *Sally* would make "the sentence sound better" (p304), because it marks the beginning of a new theme concerning the same discourse referent, i.e. a shift from talking about Sally's parents' visit to her weight. The idea of episode boundaries contradicts coherence-based views of reference (e.g. Kehler 2002), which hold that the coherence of a text involving referential terms depends on the salience of the referents. In these accounts, a text is considered coherent if the resolution of reference does not require too much cognitive effort. In line 5 of [9], however, the referent is very clear (i.e. salient) but the use of 'she' disrupts the coherence of the text. We can see here that the identifiability of a referent is not the whole story.

Another discourse function of a referring expression which exploits conventional uses is the use of *the* to indicate discourse prominence (Epstein 2002). This usage occurs when a writer employs a definite expression "to introduce an important entity at the start of a narrative, for the purpose of calling the reader's attention to that entity" (p349), when the entity itself is not yet identifiable. We can see this technique in extract [10] from the first page of the novel *All the Pretty Horses* by Cormack Macarthy.

> The candleflame and the image of the candleflame caught in the pierglass twisted and righted when **he** entered the hall and again when **he** shut the door. **He** took off **his hat** and came slowly forward. The floorboards creaked under **his boots**. In **his black suit he** stood...

It is not until 14 paragraphs later that we learn the identity of the protagonist, John Grady Cole. What is interesting about this extract is that if definite determiners and pronouns mark identifiability, then in theory, the reader should be able to identify the referent. As the narrative unfolds, the discourse representation of the referent develops in that we learn about his actions upon arriving home, his clothing, and the time in which the scene was set. The reader, therefore, has been involved in a gradual process of constructing a more and more detailed mental representation of the referent, but identification is still not possible in terms of referent resolution, and the referential content of the discourse representation of the individual is not ratified until we encounter a lexical expression.

'Late' or 'second-mention' indefinites

We saw above that indefinite expressions are typically used for non-identifiable referents, and definite expressions for something that the addressee is expected to be able to identify, but also that there are exceptions to this tendency. We have also seen examples of a particular exception to this convention; the use of A(n) + N + RRC to refer an entity previously established in the discourse. The purpose of this section is to review extant accounts of these 'late' indefinite expressions and show that while they provide a useful framework from which to guide an analysis of the targeted A(n) + N + RRC expressions in this paper, they cannot fully account for their use and so are ultimately rejected.

There have been a few attempts to account for the use of indefinite expressions which refer to a previously mentioned entity (Du Bois 1980; Ushie 1986; Epstein 1994; Schouten and Vonk 1995; Jones 2014). In the case of these 'late' or 'second-mention' indefinites, the speaker chooses a structurally indefinite expression for a noninitial mention, the referent of which is potentially uniquely identifiable to the addressee. The speaker does this in defiance of the typical conventions governing the use of the indefinite article, thereby creating an implicature which conveys something different from what has strictly speaking been expressed.

Schouten and Vonk (1995:4) view the use of indefinite expressions for known referents as *marked* (i.e. the expressions function in ways that are inconsistent with the conventional use of indefinite expression, e.g. to introduce a new referent into the discourse). The forms of referring expressions are part of a more general pragmatic principle based on speaker-addressee knowledge, and in the case of marked referring

expressions, the speaker violates the principle in order to convey additional information, which needs to be inferred by the addressee (i.e. flouting an implicature (Grice 1975), which conveys that the speaker expects the hearer to identify their true intentions (i.e. resolve the implicature)). The speaker relies on the addressee to accept that the expression is intended as marked and thus infers an interpretation that is meaningful in the specific discourse context. Further, for successful interpretation, marked indefinites necessitate the establishment of a relationship between an existing referent and the referent of the indefinite expression. It is the combination of the indefinite form and the current mental discourse representation which influences the addressee's interpretation of the marked form.

Schouten and Vonk (1995:6) also argue that the interpretation of marked referential expressions is derived from the conventional function of the marked form when observed in an unmarked way. That is, the choice of an indefinite or definite form is based on whether the entity is or can be inferred to be a unique member of a set (i.e. a set in the current referential domain). The set may be present in (or inferrable from) the surrounding discourse, situational or world knowledge (p12). So the choice of form of a referring expression is based on the value for unique identifiability and the accessibility of the referent in the existing discourse representation. Once a non-uniquely identifiable referent has been introduced into the discourse, it then becomes uniquely identifiable and accessible (p13). If either of these factors is transgressed (or flouted, in Gricean terms), an expression is being used in a marked way. At first glance, this general explanation works for expressions in this paper. But when we delve deeper into the types of expressions that these scholars discuss, we will see that there remains something exceptional about the A(n) + N + RRC expressions used in their specific context, which is unaccounted for in these accounts.

Ushie (1986) characterises late indefinites as *corepresentational*. That is, indefinite expressions which identify known referents have underlying representations which are identical to an already established referent. An indefinite expression which refers to an already established referent relates to "a certain degree of 'newness'" (p440) that results from a particular way in which the text producer organises and presents the content of a text. This presentation can be triggered by the speaker's (re)interpretation of an event or 'a shift/discontinuity in the point of view' (ibid). In terms of the re-interpretation of an event, this could be to emphasise a new feature of an established referent because of its relevance to the particular situation, representing the entity in a new light. Ushie cites the following text as an example of re-interpretation (ibid:430):

[11] In 1974, the American magazine *Rolling Stone* invited *Jan Morris* to write a series of travel articles. The fruits of that collaboration between *a romantic traditionalist Welsh author* and *a lively and innovative American paper* appears in Destinations (Oxford Books, S 1980, cited in Oda (1982:171))

In [11], the new features highlighted show the referents in contrast to one another – *traditionalist* as opposed to *lively and innovative*; an effect that would not have been achieved had pronouns been used (which would be the expected form given the salience of the referents).

While this account may capture the use of some A(n) + N + RRC expressions, we will see in the co-extensional analyses in the "Co-extensional analyses" section that, 'newness' cannot account for the use of these particular expressions because the information in the RRCs is either 'discourse-old' or 'hearer-old' (Prince 1992), i.e. the information is either given in the discourse or in the knowledge of the addressee.

There is also a type of expression when "the point of view is shifted" (Du Bois 1980:259) from the speaker to someone else involved in the text. In such "reintroductory noun phrases" (Schouten and Vonk 1995:14), the indefinite expression reflects a new, subjective interpretation of the referent or a shift in perspective from which the referent is based. In [12], an indefinite expression is used to refer to an entity which has been previously introduced and is uniquely identifiable to the addressee:

[12] *A 35-year-old citizen of Nijmegen*, caught in the act during a break-in in a pizzeria on St. Annastraat, appears to have much more on *his plate*. *The man* has confessed to being an accomplice of the armed robbery [...]. *The man* was being sought by the police for several days [...]. Yesterday night the police suddenly received a call from a pizzeria owner, who had caught *a burglar*. It turned out to be *the 35-year-old citizen of Nijmegen who was on the wanted list of the police*. (from Schouten and Vonk 1995:15)

At the moment of the occurrence of the indefinite *a burglar*, the reader might expect a definite expression to be used, based on the current context, and indeed a definite expression would work well here (*... a pizzeria owner, who had caught the burglar/the man*), even though the coherence may be slightly disrupted. However, the referent has been introduced "into the subordinated perspective of a character in the text" (Schouten and Vonk 1995:15). From the point of view of the pizzeria owner, the identity of the burglar is unknown and the use of an indefinite expression conveys this fact. Because of the clash between the expected expression (a definite expression) and the actual expression used (an indefinite expression), the expression is *marked*. The entity denoted in the indefinite expression (a burglar) is not uniquely identifiable to the pizzeria owner, thus the intended referent is reintroduced, despite the fact that the reader would be able to identify the referent, thus creating a narrative effect.

In reintroductory indefinites the unique identifiability of the referent is not measured against the knowledge structures of the addressee, but rather the knowledge structures of another participant in the text. However, the expressions in this paper do not fit into this explanation; for example, even though Trump (in [3] and [5]) and Beckham (in [1])) have been reintroduced into the text from another's perspective, it would still be difficult to argue that the particular participants in the text do not know the identity of the referent singled out in the expression. To illustrate this, let's look again at the expressions with their immediate co-text (renumbered for ease of referral):

[13] Today the United States stands not as a source of inspiration to the rest of the world but as a source of fear. Instead of hailing its first female president, it seems

poised to hand the awesome power of its highest office to *a man who revels in his own ignorance, racism and misogyny....*

[14] But you have to watch out for those Trumpites who pop up to call you "fake news" and who frighten radio station editors. The media's continuing respect for "fair play" when discussing *a president who is self-evidently a dangerous and racist xenophobe* (as opposed, for example, to the Arab variety) should one day be examined.

[15] It seems that *France*, like Spain and the United States before it, is poised to be charmed by *a man who, with his un-British attention to grooming, muscle tone and non-novelty underwear, may become an honorary Frenchman before his six months in Paris are up.*

In [13], the US knows the identity of the man they elected. They did, after all, do the electing. In [14], the media knows who they are discussing as they are the ones doing the discussing. In [15], the French know who may become an honorary Frenchman because they are, in fact, poised to be charmed by him. In addition, it is likely that the reader also relies on the discourse representation of the referent they have already built up to guide their interpretations as well as the general principle of relevance to infer that the indefinite expression refers to the identifiable entity. Once all these factors are combined, it is hardly surprising that the indefinite expression allows for identification of the referent.

On the other hand, marked predicative indefinites express properties of an established referent which do not uniquely identify that particular referent. The properties are, however, attributed to it (Schouten and Vonk 1995). This can be seen in opposition to definite expressions, which are used when the properties are uniquely tied to a referent. For example,

[16] Erik has been totally out of it since he of all people, on Monday, had to find that horrible couple: his old friend Robert an insane murderer and *Magda, a woman who in my opinion never meant very much to him,* in the most absurd state a human being can be in, dead. (in Schouten and Vonk 1995:24)

This type of expression appears to be "a curious mixture of the referential function and the predicative function for which indefinites can be used" (Schouten and Vonk 1995:23), where the expressions refer to already identifiable referents (the referential function) but also attribute information which is not uniquely tied to the referent (the predicative function). However, if we consider again the expressions in this paper, the properties in the relative clauses can be uniquely tied to the referents in the contexts in which they appear. The semantic content of the relative clause contains information which is either discourse-old or hearer-old (Prince 1992) and so newness cannot be the sole explanation for the use of the indefinite article. We will see how this connection is achieved in the co-extensional analyses in the "Co-extensional analyses" section.

Marked classifying indefinites, on the other hand, introduce the intended referent as "a member of a subclass of a basic category, with certain properties" (Schouten and

Vonk 1995:26), where new information is conveyed in the comment part of the topic-comment structure of the sentence (also Epstein 1994:223). At first glance, this kind of marked indefinite shares some features with the A(n) + N + RRC expression, as is evident in [17]:

> [17] In addition to the spiritual suffering of loneliness, of having to leave behind him 'the world which had made him what he was', *Hammarskjöld* had to endure [..] the plain physical suffering of constant nervous strain and overwork. If, as the reader goes through the entries between 1953 and 1957, he finds himself impatient [..] with their relentless earnestness and not infrequent repetitiousness, let him remember that most of them must have been written by *a man at the extreme limits of mental and physical exhaustion.*

[A]t the extreme limits of mental and physical exhaustion is both hearer-new and discourse-new (Prince 1992). However, marked classifying indefinites deviate from the A(n) + N + RRC expressions because, as we will see in the "Co-extensional analyses" section, the information in the relative clauses in the A(n) + N + RRC expressions is not new (i.e. the content can be traced back to the preceding text or on-going discourse about the entity in question), and so the use of the indefinite article cannot be explained solely through the newness function of the indefinite article.

This section, "'Late' or 'second-mention' indefinites", has provided a brief discussion of existing accounts of late indefinites and has shown that, while they provide a useful background from which to guide an analysis of the A(n) + N + RRC expressions, they cannot fully account for their use. The A(n) + N + RRC expressions do not contain new information and the content in the RRCs can be tied uniquely to the referent. For a more detailed review of the literature discussing late indefinites, see Jones (2014).

Lexical cohesion

It was noted above that the results of Jones' (2014) reader interpretation experiments suggested that the content of the relative clauses in the A(n) + N + RRC expressions appears to lead the reader to either a referring or non-referring interpretation. Therefore, a closer examination of what ties the relative clauses to the established referent would be useful to determine how the reader makes this interpretation. To do this, I exploit and extend Halliday and Hasan's well-established approach to lexical cohesion (1976, 1985), which identifies semantic fields and the logical relations between words within these fields. This particular approach is employed because it provides a transparent way of linking together various elements of the text and discourse. After a critical overview of Halliday and Hasan's approach to lexical cohesion, a detailed analysis of the semantic ties of four A(n) + N + RRC expressions is carried out.

Hasan (1985) merges Halliday and Hasan's (1976) two principle aspects of lexical cohesion, reiteration and collocation, into one broad meaning relation: co-extension. Co-extension, however, is best understood if contrasted with Halliday and Hasan's two other meaning relations: co-reference and co-classification.

Co-reference for Halliday and Hasan (1976:31) is the relationship of situational identity that exists between members of a cohesive tie. In a cohesive tie, the two (or more)

terms are anchored together through a meaning relation (Hasan 1985:73). Halliday and Hasan (1976:31) maintain that co-referential items are not interpreted semantically in their own right but rather are "directives" which signal that the information is to be recovered from elsewhere. They provide the following example:

[18] Three blind mice, three blind mice.

See how they run! See how they run!

They refers to *three blind mice,* and as a "directive", signposts that the meaning of *they* needs to be retrieved from *three blind mice.*

Halliday and Matthiessen (2014:624) broaden this definition and comment that the identity of the entity being referred to is recoverable from "the instantial system of meanings that is built up by the speaker and listener as the text unfolds". Thus, the characteristics of the entity being referred to are not static and the representation of the entity in the mind of the addressee evolves as the discourse develops.

Co-classification, on the other hand, relates to the things, processes or circumstances which belong to an identical class of items, but in which each tie refers to a distinct member of the class. The following example [19] (uttered by me) illustrates co-classification:

[19] There's *Lego*® everywhere! *It*'s in the kitchen, in the hallway and *it*'s even in the bathroom!

The relationship between each member of the tie in [19] is not that of referential identity as each instance of *Lego* refers to a distinct member of the class of items 'Lego'. The pile of Lego in the kitchen is different to that in the hallway and both are different to that in the bathroom. The meaning relations of co-reference and co-classification are different, but what ties them together is that in both cases, the meaning of the second item (and any subsequent items) in the chain is implicit and needs to be retrieved from elsewhere. So in the case of [19] the referent of the pronoun *it* in *it's in the kitchen*, the nominal ellipsis in *in the hallway* and the second pronoun *it* can only be retrieved from the initial mention of *Lego*.

Co-extension, the third meaning relation, and the one we are most concerned with here, is a semantic relation between members of a cohesive tie, where members refer to something within the same general field of meaning (Hasan 1985:74). For example, *lunch, restaurant* and *meal* are all cohesively tied to the semantic field of *food,* and if presented in a text, would form a 'similarity chain' (as opposed to an identity chain). Co-extension differs from co-classification and co-reference in that the meaning of one member is not elicited only by reference to its relation to another, but rather the relationship arises from some contiguity of meaning. Co-extension is usually realised by lexical items or content words, which form a semantic grouping (a similarity chain) within the context of a specific text items can refer to related actions, events and objects and their attributes and the semantic grouping they form is genre-specific (Hasan 1985:85). The field of discourse[3] determines, to a large extent, what items may form the semantic groupings which make up the similarity chains.

It is important to note here that grammatical cohesion and lexical cohesion are in a relationship which is interdependent and reciprocal, and in a typical text they work "hand-in-hand, the one supporting the other" (Hasan, 1985:83). Hasan (ibid) provides the following text to illustrate this point:

[20]

1. once upon a time there was a little *girl*
2. and she went out for a *walk*
3. and she saw a lovely little *teddybear*
4. and so she took it *home*
5. and when she got it home she washed it

Hasan then analyses this simple example in terms of the threads of continuity or chains highlighted: *girl*, *walk*, *teddybear* and *home* in Fig. 1.

The four separate chains (marked with different kinds of lines) link together items which are related to each other in some way, and demonstrate the simultaneity of cohesive chains. *Girl* and *she*, and *teddybear* and *it* are part of two separate identity chains, and the relation between the members of each chain is that of co-reference.

The chain initiated by *walk* is an example of a "similarity chain" (Hasan 1985:84), a chain composed of items that refer to "non-identical members of the same class of things, events etc., or to members of non-identical but related classes of things events etc." (ibid). Items in a similarity chain are related by co-classification or co-extension. So with the example of *went-walk-got,* the items lie within the same general area of meaning; that is, "walking is a kind of going, and going is an important part of getting anywhere" (p85) and the relation is that of co-extension. Whereas in identity chains the relationship between items is identity of reference, the relationship between items in a similarity chain is that of similarity of reference (Hasan 1985:85).

As noted above, similarity chains are formed by items in a text which are in the same 'general field of meaning'. However, to delimit the vagueness of 'general field of meaning', Hasan (1985:80) proposes that cohesive ties between items are only established when they stand in the sense-relations of synonymy, antonymy, meronymy and hyponymy. Whilst repetition could be argued to be absolute synonymy, Hasan (p81) claims that the cohesive tie is not "strictly-speaking" established on the basis of a sense-relation, but nevertheless contributes to the cohesion of a text as a "similar experiential meaning is encoded in each repeated occurrence of the lexical unit". For the purposes of this paper, repetition of a lexical unit is considered to be one of the ways of realising co-extension. According to Hasan (1985:80), the sense-relation restriction prevents the formation of chains such as the following: *flower, petal, stem, stalk,*
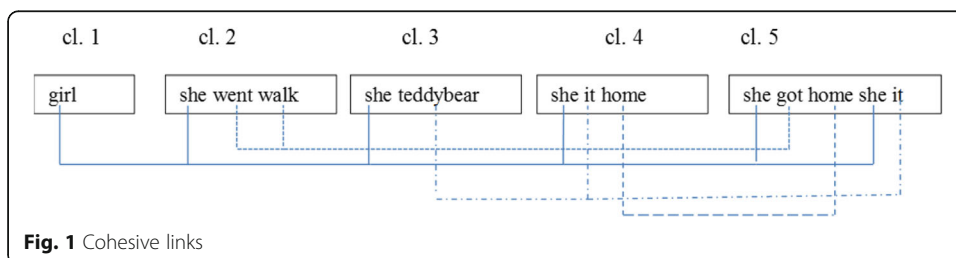

Fig. 1 Cohesive links

twig, branch, trunk, tree, wood, log, faggot, tinder, fire and *flame.* In this chain, all members of the chain are semantically related to the immediately preceding item, but not necessarily to those further away. That is, we can see the meaning relation between *flower* and *petal* (i.e. that of hyponymy) but the connection between *flower* and *flame* is unclear. Thus, by restricting the 'general field of meaning' to items which have sense-relations, this problem is avoided.

However, delimiting the notion of 'general field of meaning' by relying on sense-relations between 'content words' or 'lexical items' creates a problem for texts in which the co-extensional ties are provided by units larger than content words.[4] As we will see in co-extensional analysis 3 in the "Co-extensional analyses" section (a text entitled *Israel's Royal Welcome*), the textual cohesion comes from co-extensional ties between chunks of text which refer to events or circumstances (in the case of co-extensional analysis 3, events or circumstances which demonstrate the overall 'leitmotif' of the text, i.e. Israel's discrimination against Palestinians and non-Jews). In texts such as this, the sense-relation restriction is problematic as cohesion is formed by items which cannot be categorised into sense-relations. In fact the cohesion here goes beyond lexical cohesion; these chunks (referred to forthwith as 'propositions') are made up of highly complex noun-phrases, clauses and sentences and so the cohesion can be argued to be more 'thematic' (see Section "Co-extensional analyses of the use of A(n) + N + RRC expressions in four journalistic opinion articles", Analysis 3).

So perhaps rather than delimiting the 'general field of meaning' by sense-relations, it would be more logical for items to form a similarity chain if items in a text are in the same semantic field; this may well correspond to the initial item in the chain or the main theme of a text. So in the *flower, petal* etc. example above, the semantic field would be *flowers* and the unrelated items of *fire, tinder* etc. would simply not be accepted as part of that grouping. The formation of similarity chains in this way would preclude the need for the sense-relation restriction and therefore allow for larger units of meaning to be included.

This section has discussed Halliday and Hasan's (1985) approach to lexical cohesion. It was suggested that the co-extensional framework should be broadened to include utterances larger than 'content words' or 'lexical items', as cohesion can be created by units (i.e. propositions) which illustrate events and circumstances pertaining to the overall leitmotif of a text. This change, would however, mean doing away with the sense-relation delimitation and instead require that similarity chains are formed by items which are in the same semantic field. In the following section, we will see these changes in practice.

Co-extensional analyses of the use of a(n) + N + RRC expressions in four journalistic opinion articles

Out of the meaning relations discussed above, co-extension is the most relevant to the analysis of the A(n) + N + RRC expressions. Therefore, to explore the potential and parameters of Halliday and Hasan's approach to lexical cohesion, and to show that the use of these expressions cannot be fully captured by existing accounts of late indefinites, four examples of A(n) + N + RRC expressions and their texts are analysed. Each expression and its co-text offers something unique to the analysis or highlights a

distinct aspect of these expressions. These special features are discussed in the analyses, but first, we need to establish how the similarity chains in the texts were identified for this analysis.

In any text there are manifold similarity chains running through it, but what we are concerned with in this study is the similarity chains which show how the writer provides a shared *definite* context in the text so that the indefinite expression can be understood as referring to an established referent. I identified the key similarity chains in the texts by dividing the relative clause of each A(n) + N + RRC expression into component 'chunks' of information, which hold some kind of semantic content in their own right. In some cases, this corresponds to an entire clause. The result of this was that sometimes the relative clauses were not divided at all as there was only one semantic unit and therefore just one similarity chain (see Fig. 4), and others are very complex relative clauses and contain several semantic units, and thus several similarity chains (see Fig. 3). Members identified as being part of the similarity chains are based on information related to the RRC given in the text preceding the occurrence of the A(n) + N + RRC expression.

Key for the analyses

To better understand the following analyses, descriptions of the terms used are provided below.

> **Identity chain**: the referent being tracked and the number of times the referent has been mentioned in the text preceding the occurrence of the A(n) + N + RRC expression
>
> **Similarity chain**: the members of the chain which create lexical cohesion through mentions of items in the same general field of meaning as the items identified in the RRCs. For this analysis, only items which have occurred in the text preceding the occurrence of the A(n) + N + RRC expression are traced because how the writer creates a sufficiently 'shared' context for the indefinite expression to get interpreted as definite is of interest. In this way, these items are 'old' in the discourse model of the addressee (Prince 1992:303) and textually traceable.
>
> **Shared cultural knowledge:** the information in the relative clauses can also be "hearer-old" (Prince 1992: 301) or "copresent" (Clark and Marshall 1981: 38), which means that the information is considered old with respect to the text producer's assumptions about what the hearer already knows (Prince 1992: 301). To put this into context, it is necessary to consider the kind of texts being analysed here, as well as the assumed readership. The texts are opinion pieces from the British broadsheet newspaper, *The Guardian*. They are about a specific, well-known, entity in a particular field of discourse. We can assume that readers of such articles have an interest in the particular entity being discussed and are familiar with the field of discourse, and therefore have some previous, culturally-shared knowledge about the thing (or at the very least, the writer is behaving as if they do). Thus, the connection is outside the text, in the broader field of discourse, but nevertheless helps add to the overall cohesion which ties the attributes in the RRCs to the specific participant in the text (c.f. co-reference and co-classification, which are not considered cohesive if they occur outside the text (Halliday and Hasan 1976)).

Fig. 2 Reader's Digest
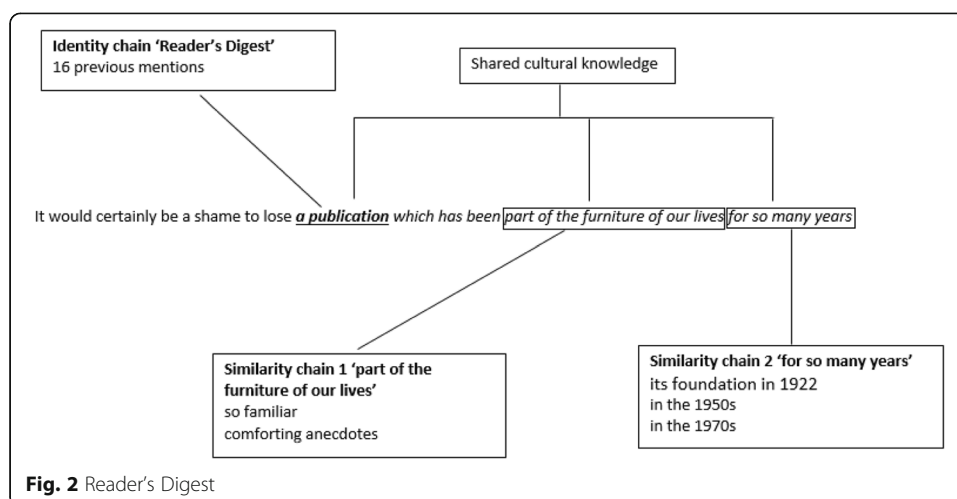
Co-extensional analysis 1: In praise of…Reader's digest (*The Guardian*, 21/08/09) (Fig. 2)

The A(n) + N + RRC expression in this text corresponds most closely to Hasan's notion of lexical items or content words forming co-extensional ties, and illustrates a relatively simple co-extensional meaning relationship between the A(n) + N + RRC expression and the text. To recap, the members of the similarity chains (SCs) appear in the text before the occurrence of the A(n) + N + RRC expression and are considered to be in the same general field of meaning as the similarity chain. SC1 'part of the furniture' has two members, one of which, *so familiar*, relates to the notion that if something is part of the furniture, it is therefore familiar. The remaining member, *comforting anecdotes*, semantically and cohesively links the idea that something that is part of the furniture, and is therefore familiar, is also usually comforting. SC2 'for so many years' has three members (*its foundation in 1922; in the 1950s; in the 1970s*) which all relate to past time. But not only does the expression have co-textual cohesive ties, it also contains shared cultural knowledge. The Reader's Digest is a well-known general interest 'family' magazine, and even if readers of the article have not read it personally, they would likely be aware of its existence as an established 'institution' in the UK and that it has been 'part of the furniture for so many years'.

Co-extensional analysis 2: David Beckham: How this crock of a footballer can still woo the French (*The Guardian*, 01/02/13) (Fig. 3)

This text demonstrates how complex A(n) + N + RRC expressions can be. This expression contains three similarity chains, each of which has several members, and as well as illustrating co-textual cohesive ties, it shows a link to shared cultural knowledge that the writer assumes the reader brings with them to the text. SC1, 'appearance', has two textual members and is also shared cultural knowledge; Beckman's looks and physique get considerable media attention in the UK. The second, 'charming the French' has five members and ties together the core argument of the text; that although Beckham is old ('age' being another similarity chain, which is dealt with and dismissed as an argument in paragraph 3), his personal attributes may still help him win over the French. SC3, 'move to France', has five members

Fig. 3 David Beckham
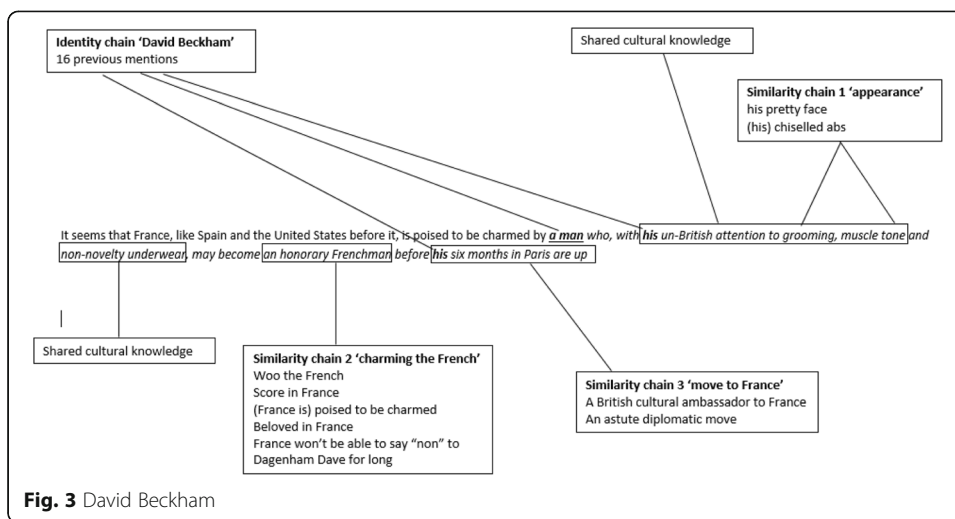
and makes repeated reference to the situation described in the text; that Beckham has been signed to play for Paris St-Germain. Two of the members of SC3, *a British cultural ambassador to France* and *an astute diplomatic move* also relate cohesively to SC2, as they suggest that one of the reasons for his move to France is his charm. We are also provided with some textually new information, *non-novelty underwear*, which can certainly be considered as culturally-shared knowledge (in the UK at least). At the time this article was written, it was difficult to step outside one's house without seeing images of David Beckham modelling his own underwear range. Thus, the connection is outside the text, to the reader's assumed cultural knowledge of Beckham, but it nevertheless creates cohesion.

Something that is particularly striking in this example is that the A(n) + N + RRC expression contains two direct definite references. That is, the possessive pronoun *his* occurs twice in the expression, which can only be interpreted as referring to the previously mentioned participant in the text, David Beckham.

Co-extensional analysis 3: Israel's Royal Welcome (*The Guardian* 25/03/08) (Fig. 4)

Hasan's (1985:80–81) claim that co-extension is usually realised by lexical items or content words is problematic for this example. The textual cohesion in this article is created by units larger than content words or lexical items. It is clear that these larger chunks of text (i.e. propositions) about discrimination against Palestinians and non-Jews are crucial to the continuity of the argument running through the article, and therefore to the links of the preceding text to the A(n) + N+ RRC expression. It would be impossible to create a similarity chain without them because it is reference to events and circumstances in the preceding text which creates the textual cohesion. More specifically, the members of the similarity chain in this text are all illustrations of actions taken by the JNF which discriminate against Palestinians and non-Jews; the relative clause sums up the overall 'leitmotif' of the text by providing an overview of what all these occurrences exemplify. Without the inclusion of propositions as items which are able to create cohesive ties, it would not have been possible to demonstrate how the preceding text is cohesively linked to the RRC in this text.
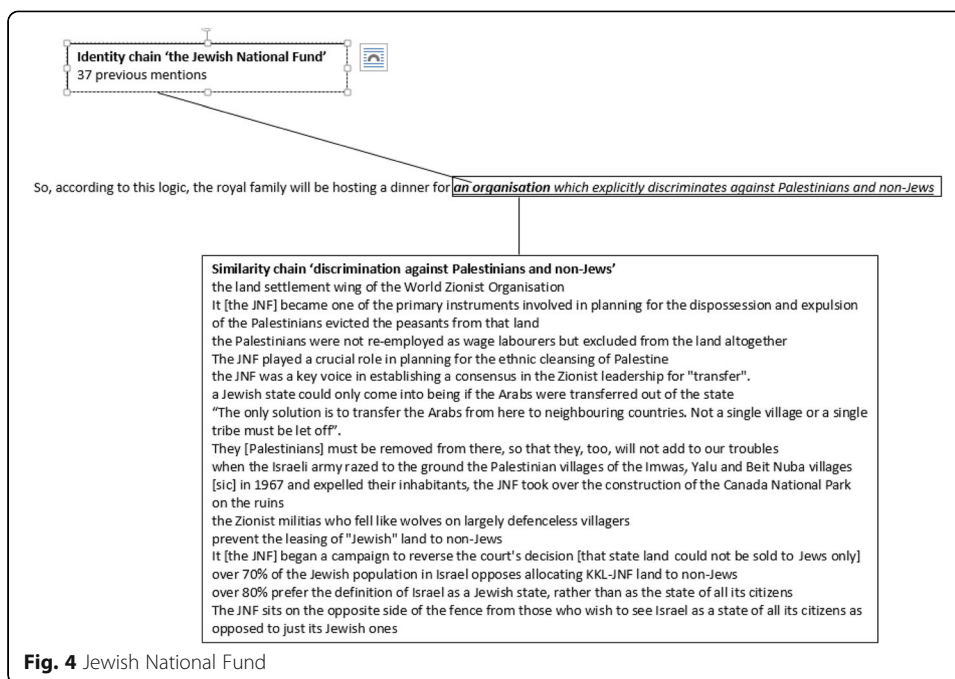
Fig. 4 Jewish National Fund

Co-extensional analysis 4: The man who came in from the cold (*The Guardian*, 29/01/04) (Fig. 5)

This text illustrates that an A(n) + N + RRC expression without textual co-extensional ties. The expression in this text occurs in the second paragraph, after five previous mentions to the participant, Greg Dyke. There are no lexical links to the preceding co-text, so the question is how can the connection to the entity in the indefinite expression to the definite referent be made. There seem to be several strategies. First, the identity chain of references to the participant is continued in the RRC, with references to **his** *defence* and **he** *took*. Furthermore, this RRC contains other features of definiteness; a specific episodic event and a definite past time.

Moreover, the content of the RRC can be considered shared cultural knowledge, but the time of writing and context must be taken into account. This article was written on 29th January 2004, the day that Greg Dyke resigned from the BBC after an extremely well-publicised enquiry (The Hutton Report) into errors of judgement made by the BBC when checking news stories. The media furore surrounding this affair was intense,
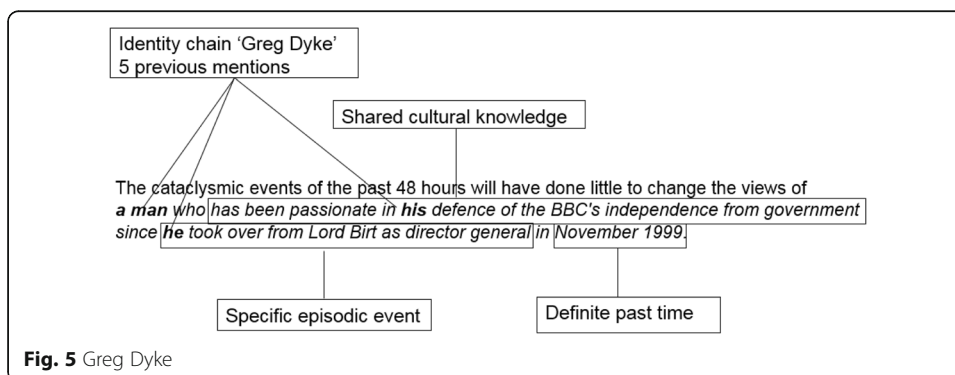

Fig. 5 Greg Dyke

and much of what was discussed was about the government's interference in the BBC and Dyke's resistance to this.

So even though there are no preceding textual ties between the text and the RRC in this instance, the connection to the referent is made by continuity of reference and cohesive ties to shared cultural knowledge of the time. There are ties, not to the text but to the on-going discourse surrounding the events described in the text. So again, there is no newness in the expression which can account for the use of the indefinite article.

Concluding comments about the analyses

The analyses in this section have illustrated various features of the A(n) + N + RRC expressions. We have seen how these features operate within their texts to create cohesion so that the indefinite expressions can be interpreted as referring to the definite referent. More specifically, the analyses have shown that the content of the RRCs is connected to the preceding discourse via textual co-extensional ties, and/or to the shared cultural knowledge that the reader is assumed to bring with them to the text. The following section considers how and why this exploitation of traditional referring conventions might occur.

Discussion: 'Dual' reference

This analysis has revealed the limitations of existing accounts of late indefinites and challenged and extended Hasan's co-extensional framework. Co-extensional analysis can provide a means of illustrating how the referents of these complex indefinite expressions are cohesively tied to the preceding text and on-going discourse. The writer does not simply rely on textual cohesion, but also taps into the cultural knowledge assumed as shared by the readership, i.e. the writer makes out that the information is 'hearer-old', given the likely readership of this of kind of article. Another strategy to guide interpretation is the continuation of the identity chain of reference. In each of the expressions, there is something to link the expression to the specific participant, whether it is textual ties, contextual/discoursal ties or continuity of reference. Thus, there is an assumption of cooperation and adherence to the principle of distant responsibility towards the reader (Clark and Wilkes-Gibbs 1986:34), through the provision of a platform for the interpretation of the expression as referring to the established participant.

These expressions can be highly complex and have manifold ties to the preceding text and surrounding discourse. These ties allow the expressions to function in a *definite* manner, but perhaps it is not only the ties that perform this role. The reader may be happy to be persuaded by the presence of specific detail in the relative clause that the expression is referring to the previously mentioned entity. The amount of detail in the relative clause may well indicate how much the writer intends for the reader to interpret the expression as referring to the definite entity. Why writers might do this is considered next.

One explanation for why the chain of definite reference is broken and a switch to in-definite reference is made is that there is some kind of *duality* in this type of reference; reference to the fully identified referent and to a *virtual* referent with the same

qualities. Rather than the expression being understood as a generalisation to a type or sub-class of entity or, on the other hand, definite reference to the identified individual, the two factors are considered simultaneously. Thus, the writer intends to refer to *both* the individual *and* a type, and the type is the generalisable form associated with the entity, creating what Epstein (1994: 226) calls a "generalising effect". For Epstein (1994: 227), these expressions have two simultaneous functions: they are reclassified in light of the *new* information in the relative clause, and the ascribed property is presented as both characteristic of the whole class and implicitly portrayed as characteristic of a specific uniquely identifiable entity. The addressee recognises that these expressions are not simply referring to an arbitrary member of the particular category and is aware of the identity of the referent 'thanks to the (virtual) link between it and its previous mention' (ibid). However, Epstein, like Schouten and Vonk (1995) and Ushie (1986), explains the use of the indefinite article through some element of newness in the expression, but as we have seen from the co-extensional analyses, the expressions do not contain new information, whether it is discourse-new or hearer-new (c.f. Prince 1992), so the newness hypothesis does not hold. However, the duality explanation should not be abandoned because of this.

An explanation for why the switch from definite to indefinite reference occurs in these particular texts comes from journalism academic Howard Barrell (personal communication, 2012)[5] who suggests that:

> [A]n important feature of our Western rational tradition is the assumption that we best explain life by identifying and referring to regularities. […] Journalists who recognise this […] often wish they were involved in something less random in its collection of data. They find themselves dealing […] with individual stories — stories that are saleable to audiences precisely because they are individual, a departure from the norm, or sensational in some way. Yet journalists simultaneously yearn to be able to extrude some kind of 'pattern', 'regularity' or 'rule' from one or a collection of these stories. Indeed, 'analysis' (and perhaps 'comment' as well) would seem to demand that they do so.

One way journalists are able to achieve this is by using an indefinite expression to refer to something already established. Using a definite expression initially allows the journalist to establish the parameters and relationships which characterise their story and make it unique. When the journalist wishes to argue that the story they have told about a specific individual "may be governed by some rule or regularity of politics or of existence" (Barrell 2012) or that their story may enable the reader to identify some previously unrecognised rule or regularity, the journalist "may very well start referring to '*a* prime minister who does this/fails to do that'". So the intention of the journalist here is to abstract from the specificities and limitations of their story a more general truth. The use of an indefinite expression makes of the story an instance or example of some truth or regularity beyond itself, suggesting that the journalist intends to refer to the specific individual and at the same time, establish an abstract principle: 'dual' reference.

The effect of this *categorical principle* is more powerful when the expression is less explicit: the broader the regularity, the more abstract and generalisable it becomes, and

therefore the stronger the impact. But the more explicit the details in the RRC (and therefore the more ties), the stronger the association is with the specific individual, but as a consequence of this explicitness, the regularity or general truth is less powerful.

This explanation is certainly reasonable for the expression in the context of opinion journalism, but further exploration into other genres or registers may reveal data that leads to other explanations. Future research should investigate whether the A(n) + N + RRC expression occurs in spoken data and other written genres, and if so, what function it has.

Conclusion

I have analysed the use and function of the expression A(n) + N + RRC in journalistic opinion writing to challenge existing accounts of late indefinites and Hasan's co-extensional framework (1985). Explanations of late indefinites tend to rely on the newness factor, but the co-extensional analyses showed that the information in the expressions in not actually new, given the context and readership of the texts. It was also suggested that the co-extensional framework be broadened to accommodate linguistic units larger than lexical items, because cohesive ties can be formed by clauses and sentences depicting events and actions. However, this extension is problematic for Hasan's framework because she limits what can be included in a similarity chain to members which stand in sense-relations, but sentences and clauses cannot be categorised in this way. Further, broadening what can count as a lexical cohesive tie in similarity chains and thus doing away with the sense-relation restriction runs the risk of an infinite number of (sometimes unrelated) items being included in any one chain. But as we saw, limiting the members of a grouping to those semantically related to the overall theme or initial member of the chain, rather than to a preceding item, would circumvent this problem.

English does not have an explicit linguistic way of exhibiting simultaneously the relationship between old and identifiable and new and non-identifiable, so the relation has to operate outside of language, perhaps in cognition only. It is difficult to engineer a way in which language is responsive to this relation (in English anyway), as it only responds to the construal of the 'virtual' referent as non-identifiable (c.f. Du Bois: Definiteness, reference and analogues, unpublished). However, the writers of the texts analysed have found a way of ensuring that the association to the old, identifiable referent remains explicit, through the use of a RRC which contains already established information which is tied cohesively to the cotext, context or on-going discourse about the specific entity.

There are three components which carry meaning in the A(n) + N + RRC expressions which guide interpretation: the indefinite article, the detail and level of specificity of the information in the relative clause, and the surrounding cotext, context and ongoing discourse about the specific entity. To interpret the expressions in a meaningful way, the three components cannot be separated. Only when they are considered as three interrelated parts which make up the whole, can we begin to understand the use and function of the A(n) + N + RRC expression.

The encoding of a referring expression is important, but it is not the only factor which contributes to its (in)definite status and interpretation. When the expression encodes both definite and indefinite information, the entire discourse event may ultimately determine how the expression is interpreted.

Endnotes

[1]Texts reproduced under Open Licence terms and conditions. Copyright owned by The Guardian News & Media Ltd. and Independent Digital News and Media Ltd.

[2]The fundamental principle of Accessibility Theory is that by using a certain expression, the speaker instructs the addressee to search in their memory for a piece of given information. The choice of referring expression thus indicates how accessible the speaker considers this piece of information to be for the addressee at a particular point in the discourse.

[3]Where discourse is defined as 'the total event, in which the text is functioning, together with the purposive activity of the speaker or writer' (Halliday and Hasan 1974:22).

[4]In fact, support for including larger lexical elements comes from Halliday (1994:311), who includes 'wordings having more than one lexical item in them, such as *maintaining an express locomotive at full steam*' (italics in the original) in his description of what constitutes reiteration and collocation, both of which are 'relations between lexical elements', i.e. features of lexical cohesion.

[5]Interview with Dr. Howard Barrell at the School of Journalism, Media and Cultural Studies, Cardiff University, November 2012.

Availability of data and materials
The data used in the study are from texts reproduced under Open Licence terms and conditions. Copyright owned by Guardian News & Media Ltd. and Independent Digital News and Media Ltd.
All data generated or analysed during this study are included in this published article [and its Additional files].

Authors' contributions
KJ is responsible for all aspects of the work. The author read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References
Abbott, B. 2010. *Reference*. Oxford and New York: Oxford University Press.
Ariel, M. 1988. Referring and accessibility. Journal of Linguistics 24(1): 65–87. https://doi.org/10.1017/S0022226700011567
Ariel, M. 1990. *Accessing noun phrase antecedents*. London/New York: Routledge.
Clark, H.H., and C.R. Marshall. 1981. Definite reference and mutual knowledge. In *Elements of discourse understanding*, ed. A.K. Joshi, B. Webber, and I. Sag, 10–63. Cambridge: Cambridge University Press.
Clark, H.H. and Wilkes-Gibbs, D. 1986. Referring as a collaborative process. Cognition 22:1-39. https://doi.org/10.1016/0010-0277(86)90010-7.
Cornish, F. 1999. *Anaphora, discourse and understanding*. Oxford: Oxford University Press.
Cornish, F. 2010. Anaphora: Text-based or discourse- dependent? Functionalist vs. formalist accounts. Functions of language 17: 207–241 https://doi.org/10.1075/fol.17.2.03cor.
Du Bois, J. 1980. The trace of identity in discourse. In The pear stories: cognitive, cultural, and linguistic aspects of narrative production, ed Chafe, W. 203–274, Norwood: Ablex.
Epstein, R. 1994. *Discourse and definiteness: Synchronic and diachronic perspectives*. San Diego: University of California.
Epstein, R. 2002. The definite article, accessibility, and the construction of discourse referents. Cognitive linguistics, 12(4): 333–378 https://doi.org/10.1515/cogl.2002.007
Fawcett, R.P. 1980. *Cognitive linguistics and social interaction: Towards an integrated model of systemic functional grammar and the other components of a communicating mind*. Bamberg: Julius Groos Verlag/University of Exeter.
Fries, P. 2001. Issues in modeling the textual metafunction: A constructive approach. In *In Patterns of text: In honour of Michael Hoey*, ed. M. Scott and G. Thompson, 83–107. Amsterdam: John Benjamins.

Givón, T. 1983. In *Topic continuity in discourse: an introduction. In Topic continuity in discourse: a quantitative cross-language study*, ed. T. Givón, 4–41. Amsterdam/Philadelphia: John Benjamins.

Givón, T. 1993a. *English grammar 1. A functional-based introduction*. Amsterdam/Philadelphia: John Benjamins.

Givón, T. 1993b. *English grammar 2. A functional-based introduction*. Amsterdam/Philadelphia: John Benjamins.

Grice, P. 1975. Logic and conversation, in Syntax and semantics, 3: *Speech Acts*, eds Cole, P. and Morgan, J. 41–58. New York: Academic Press.

Gundel, J.K., N. Hedberg, and R. Zacharski. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69 (2): 274–307.

Halliday, M.A.K. 1994. *An introduction to functional grammar*. 2nd ed. London: Edward Arnold.

Halliday, M.A.K., and R. Hasan. 1976. *Cohesion in English*. London: Longman.

Halliday, M.A.K., and R. Hasan. 1985. *Language, context and text*. Oxford: Oxford University Press.

Halliday, M.A.K., and C.M.I.M. Matthiessen. 2014. *Halliday's introduction to functional grammar*. 4th ed. London: Routledge.

Hasan, R. 1985. The texture of a text. In Language, context and text. In, ed. M.A.K. Halliday and R. Hasan. Oxford: Oxford University Press.

Hawkins, J.A. 1978. *Definiteness and indefiniteness*. London: Humanities Press/Croom Helm.

Jones, K. 2014. *Towards an understanding of the use of indefinite expressions for definite reference. PhD thesis*. Cardiff University Available at: http://orca.cf.ac.uk/71386/.

Kehler, A. 2002. Coherence, reference and the theory of grammar. Stanford, CA: CSLI Publications.

Martin, J.R. 1992. *English text: System and structure*. Amsterdam: John Benjamins.

Prince, E. 1992. Subjects, definiteness and information status. In *Discourse description: diverse linguistic analyses of a fund-raising text*, ed. W.C. Mann and S.A. Thompson, 295–325. Amsterdam/Philadelphia: John Benjamins.

Schiffrin, D. 2006. *In other words. Variation in reference and narrative*. Cambridge: Cambridge University Press.

Schouten, C.H. and Vonk, W. 1995. On the use of marked indefinite noun phrases. The perspectival nature of indefinite noun anaphors. Technical Report, ESPRIT Project 6665 DANDELION. Tilburg/Nijmegen etc.: DANDELION Consortium and CEC.

Toole, J. 1996. The effect of genre on referential choice. In *Referent and referent accessibility*, ed. J. Gundel and T. Fretheim, 263–289. Amsterdam/Philadelphia: John Benjamins.

Ushie, Y. 1986. 'Corepresentation' – A textual function of the indefinite expression. Text 6(4):427-446 https://doi.org/10.1515/text.1.1986.6.4.427.

Vonk, W., Hustinx, L.G.M.M. and Simons, W.H.G. 1992. The use of referential expressions in structuring discourse, language and cognitive processes, 7(3-4): 301-333 https://doi.org/10.1080/01690969208409389.