

ENGLISH LANGUAGE
& LINGUISTICS



CAMBRIDGE
UNIVERSITY PRESS

**English article usage as a window on the meanings of same,
identical and similar**

Journal:	<i>English Language and Linguistics</i>
Manuscript ID	ELL-ART-Mar-15-007.R2
Manuscript Type:	Original Article
Keywords:	English articles, definite descriptions, same, identical, similar

SCHOLARONE™
Manuscripts

English article usage as a window on the meanings of *same*, *identical* and *similar*¹**Luna Filipović¹ & John A. Hawkins^{2,3}**¹University of East Anglia, ²University of California Davis, ³Cambridge University

For Peer Review

¹ We are grateful to two anonymous referees and to Ekkehard König for detailed comments on earlier versions of this paper which improved the first draft considerably. Our gratitude also goes to Bernd Kortmann for helpful and efficient editorial work. We are solely responsible for any remaining shortcomings in the current version.

ABSTRACT

We propose an explanation for a traditional puzzle in English linguistics involving the use of articles with the nominal modifiers *same*, *identical* and *similar*. *Same* can only take the definite article *the*, whereas *identical* and *similar* take either *the* or *a*. We argue that there is a fundamental difference in the manner in which a comparison is made with these modifiers. *Identical* and *similar* involve direct comparisons between at least two entities and an assertion of either full property matching (*identical*), or partial property matching (*similar*). The comparison with *same* proceeds differently: what is compared is not linguistic entities directly, but definite descriptions of these entities that can be derived through logical entailments. *John and Mary live in the same house* entails *the house that John lives in is the (same) house that Mary lives in*. There must be a pragmatic equivalence between these entailed definite descriptions, ranging from full referential equivalence to a possibly quite minimal overlap in semantic and real-world properties shared by distinct referents. These differences in meaning and article co-occurrence reveal the sensitivity of syntax to semantic and pragmatic properties, without which all and only the grammatical sentences of a language cannot be predicted.

No two persons ever read the same book.
Edmund Wilson

1 INTRODUCTION

In 1991 Hawkins proposed an integrated semantic-pragmatic-syntactic theory of various ungrammaticalities involving definite and indefinite articles in English in combination with other items in the noun phrase. He demonstrated that certain grammaticality distinctions are “extremely fine-tuned to the semantics and pragmatics” (Hawkins 1991: 434). For instance, a noun modifier that normally carries a uniqueness entailment (and thus would require a co-occurring definite article) may on occasion not do so, and the indefinite article can then occur grammatically, as in *a best buy* and *a first course in German* versus **a wisest king* (ibid.). Further contrasts that he was able to account for were *an only child* versus **an only student*, *a colour like red* versus **a colour red*, and *I recalled a sweet little child that Mary used to be like* versus **I recalled a sweet little child that Mary used to be*. This is our starting point in this paper – we use information about the grammaticality of definite and indefinite articles with the modifiers *same*, *identical* and *similar* to shed light on some semantic and pragmatic properties of these items. We focus on these three modifiers and their respective interactions with the articles in English because they reveal some intriguing differences in meaning and usage that require a more detailed analysis than was given in Hawkins’ earlier work. We will also argue that despite the addition of numerous rich and informative studies during the last 25 years concerned with English articles and with these adjective modifiers, two essential questions have not been satisfactorily answered, involving the grammaticality of these noun phrases on the one hand, and their meanings on the other.

First, why is the indefinite article ungrammatical with *same* (**a same house/the same house*), whereas both *a* and *the* are grammatical with *identical* (*an identical house/the identical house*), as they are with *similar* (*a similar house/the similar house*)? We are not aware of any convincing explanation for this in the literature, going beyond mere observation and stipulation, and yet it is a rather fundamental fact about the syntax of this area of English grammar which does need to be accounted for. It is also a rather surprising fact, since to assert of two or more objects that they are identical seems to involve a claim that “the objects belong to one and the same common type, exactly as is the case with *same*” (Hawkins 1978: 251). Hawkins (1978:247-253, 1991) appealed to the uniqueness of the ‘type’ of entity to which *same* + noun refers, with or without the uniqueness and referential identity of tokens referred to, in order to explain the required co-occurrence with the uniqueness-entailing definite article. For *identical* he proposed that its meaning involved no such notion of abstract type for *house*, but instead full property for property matching between distinct referential tokens. The selection of articles with *identical* would be based in the usual way, he argued, on whether one of the referential tokens of *house* was or was not unique within a pragmatically restricted domain of interpretation (a “P-set”, see Hawkins 1991). We will not continue this line of explanation here for the ungrammaticality of **a same* but will propose a different account that makes no appeal to ‘types’.

Second, there are some key differences in meaning and usage between *same* and *identical* whose theoretical significance has not been fully appreciated in the recent literature. For example, in her insightful and empirically detailed study of *same* and *identical* Breban (2010) gives numerous examples, including actual corpus data and usage statistics, of their use and meaning and concludes that “The postdeterminers *same*

and *identical* clarify that the hearer can identify the instance by means of a phoric relation of identity or co-referentiality”. She adds that the “postdeterminers *same* and *identical* in fact signal identity of reference by invoking the idea of non-identity: ‘it is the same instance and not another one’. The determiner unit conjures up a second possible referent, only to deny it and to confirm that the referent the hearer has in mind, is the right one.” Her notion of “identity” allows for “generalized” instances of e.g. a certain kind of house, following Langacker (1991, 2005; cf. also Breban 2011), in addition to strict referential identity and co-reference between tokens of the house in question, and she documents a rich array of uses for noun phrases containing *same* and *identical*. What is missing in her account, however, is an appreciation for the precise **difference** between them. This difference can be seen when one tries to replace *same* with *identical* in the illustrative corpus examples she gives in Breban (2010).

On p.212 (op cit.), for example, she cites the following attested use (her (7.41)):

- (1) A dog which plunged 400ft down a mountainside had its fall broken by two climbers who had plunged through **the same ice hole**.

If we replace *the same ice hole* in (1) with *the identical ice hole* the reference changes to a quite different ice hole from the one that the two climbers had plunged through and the sentence will receive a very different pragmatic interpretation:

- (2) A dog which plunged 400ft down a mountainside had its fall broken by two climbers who had plunged through **the identical ice hole**.

A dog which plunges through one ice hole cannot normally in our world have its fall broken by climbers who are in another, albeit “identical”, ice hole. If the climbers were formerly in another, identical ice hole and are now somehow in the “same” one as the plunging dog, then the sentence may be interpretable and avoid pragmatic anomaly. But this contrast makes clear that *same* and *identical* pattern differently in English definite NPs with respect to their referents, their co-reference and the manner in which “a second possible referent” is evoked (Breban 2010:212). Contrary to what Breban claims in the quotes above, this second referent is not denied in (2), but is actually asserted. Hence whatever similarities and overlaps there are in English definite NPs containing *same* and *identical*, there are profound differences as well which are not being accounted for.

In earlier joint work with Davidse and Van Linden (Davidse et al. 2008), Breban also equated *same* with *identical* and pointed out that these post-determiners “merely emphasise the coreferentiality and inclusive reference conveyed by the primary determiners”. This was in line with other studies (Barker 2007; subsequently Brasoveanu 2011) that have tried to equate the meanings of the definite article and *same* and to blend them into one. But definite descriptions with *same* are not used merely for the purpose of emphasis or inclusive/unique reference. They involve instead a form of reference that can be best understood by going back to the basic logic of Russell’s (1905) theory of definite descriptions and to the kind of pragmatic extension of his theory that was developed in Hawkins (1978, 1991). Hawkins’ (op cit) proposal was that the existence and uniqueness of definite referents needs to be interpreted relative to different pragmatic sets (his “P-sets”) within a pragmatically structured universe of

discourse that can make sense of everyday uses of definite descriptions going far beyond *the present king of France*.

The anaphoric nature of many definite NPs is well-known and well-documented (see e.g. Birner 2013: ch.4), as are their deictic properties (see Schwarz 2009 for a particularly interesting recent proposal for describing these in terms of “strong” definites). Many other appropriate uses of the definite article in English, for example various “situational” and “associative” uses, have been summarised in C. Lyons (1999), in Hawkins (1978, 1991), and in more traditional works such as Christophersen (1939) and Jespersen (1949). What is new in the present study is an explanation for the semantically different uses of *same*, as well as the reason we shall propose for its obligatory use with the definite article. Another original contribution is the contrastive discussion of the various meaning possibilities for *same*, *identical* and *similar*. Our point of departure is that grammaticality distinctions in article usage provide an independent piece of evidence, in the form of syntactic wellformedness judgments, for subtle semantic features of these adjectives of comparison. Our claim will be that there is a difference in the semantics of the comparison and in what exactly is being compared, and that this is what underlies the article co-occurrence differences.

Briefly, *identical* and *similar* in combination with singular nouns refer to a single entity (a pragmatically unique one if preceded by *the* and generally a non-unique one if preceded by *a*) whose existence is entailed, and this single entity is compared with a second distinct entity or entities, whose existence is also entailed. The nature of this comparison when *identical* is chosen is claimed to involve full matching of all properties between what we can call the ‘direct’ referent of e.g. *(I saw) the/an identical house*, and the second or ‘indirect’ referent or referents, with which the direct referent is

being compared. In the case of (*I saw*) *the/a similar house* the existence of a direct referent is entailed, a second indirect referent or set of referents is also entailed, and the comparison is now claimed to involve only a partial match between their respective properties. Crucially, for both *identical* and *similar* the comparison is between at least two distinct entities, the direct referent and the indirect referent, the comparison involves either full or partial property matching between them, and the direct referent can either be definite (*the identical/similar house*) or indefinite (*an identical/similar house*).

Same behaves differently. The comparison does **not** now apply at the level of distinct entities and between direct and indirect referents, as it does for *identical* and *similar*, since there may be only one entity that figures in the comparison. If I say, for example, that *John and Mary live in the same house*, this is logically equivalent to asserting that *the house that John lives in is the same house that Mary lives in*. What is being compared, in effect, is one definite description *the house that John lives in* with another, *the house that Mary lives in*, and it is being claimed in what is probably the most common interpretation that the two are referentially identical. This is the first difference between *the same house* and *the identical house*. In *John and Mary live in the identical house* there is necessarily a comparison between distinct houses. The most likely interpretation compares the direct referent that John and Mary are claimed to live in with some referentially distinct house, the indirect referent. In *John and Mary live in an identical house* John and Mary may live in a single house, the direct referent that is being compared with some other, the indirect referent, or they may live in distinct houses that are being compared with one another, with full property for property matching being asserted between all the houses in question. For *the same house* the

comparison need not involve such distinct referents, direct or indirect, however, and does not apply at the level of the items referred to, checking for their matched properties. Instead it involves a more subtle comparison over partially distinct definite descriptions, *the house that John lives in* and *the house that Mary lives in*, and these definite descriptions may refer, we have seen, to just one single entity.

But there is also another interpretation for *John and Mary live in the same house* in which John and Mary do live in distinct houses and in which the meaning is, for example, that they live in the same model in a housing estate, but in different tokens of this model. Here the meaning of *the same house* encroaches on that of *identical*, and this state of the world could be captured by saying that *John and Mary live in an identical house*. This less usual interpretation for *John and Mary live in the same house* is also present in the logical paraphrase comparing definite descriptions, *the house that John lives in is the same house that Mary lives in* (i.e. the same model of house in both cases). The interpretation here involves the process of referential “generalization” for *house* that has been insightfully discussed by Langacker (1991, 2005) and Breban (2010, 2011). The key point for now is that *same* permits an interesting variability in its interpretive possibilities, it does not require distinct direct and indirect referents like *identical* and *similar* do, and the comparison that is evoked does not apply at the level of distinct referential tokens but rather it applies to partially distinct definite descriptions that can be seen in logically equivalent paraphrases of sentences containing *the same + Noun*. We will argue here that *the same + Noun* involves a comparison over alternative partially overlapping definite descriptions, derived through logical entailments, and that it asserts a pragmatic (not a logical) equivalence between them, along the lines of *the house that John lives in is the (same) house that Mary lives in*.

In what follows we explain this idea further and we exemplify the overlaps as well as crucial differences in article usage between *same*, *identical* and *similar* and in the semantic interpretations that characterise these expressions. We will also account for the fact that *same* is not semantically vacuous in combination with the definite article. *Same* is a relational term. It has a more restrictive meaning than the definite article, but this latter is required in co-occurrence with *same* because *same* involves a form of double definiteness, a comparison over two definite descriptions, each of which requires *the*, with at least some partial equivalence between them either at the level of their referents, or (for non-identical referents) at the level of the semantic properties within the respective definite descriptions. This partial equivalence gives *the same* features in common with both *similar* and *identical*, as we shall see, but crucially *the same* differs from these latter over what exactly is being compared, two definite descriptions with pragmatic equivalence in the one case, versus two distinct referential tokens with full or partial property matching in the other. Since what is being compared is two definite descriptions, and since an equivalence is being asserted between them, the single NP in the entailing sentence can be no less definite than the definite descriptions being compared, and hence it has to be *the same* not **a same*. The definite article on its own captures uniqueness of singular entities in general and in a whole variety of contexts (previous discourse, situation of utterance, ‘association sets’ or frames, etc., see Hawkins 1991). In order to explain our theory more fully, we need to go back to basics: Russell’s (1905) theory of definite descriptions.

2 RUSSELL’S (1905) THEORY OF DEFINITE DESCRIPTIONS

Consider the definite description in sentence (3):

(3) The house was sold.

According to Russell's (1905) analysis, its logical translation would be (4) (ignoring the semantics of the past tense):

(4) $\exists x(H(x) \ \& \ \sim\exists y(H(y) \ \& \ x \neq y) \ \& \ S(x))$
i.e. there is an x which is a house and there is no y such that y is a house and non-identical to x and x was sold

Example (3) accordingly makes three claims:

- (5) (a) Existence: *there is a house*
(b) Uniqueness: *there is only one house*
(c) Predication: *this entity was sold*

If sentence (3) is true, then each of the italicised sentences in (5a), (5b) and (5c) will be true, and hence (3) entails the conjunction of the italicised sentences (5a)-(5c).

The crucial distinction between (3) and the corresponding indefinite description

(6) lies in the uniqueness claim:

(6) A house was sold.

(7) $\exists x (H(x) \ \text{and} \ S(x))$
i.e. there is an x which is a house and x was sold

The truth of sentence (6) requires that there should be at least one house that was sold. It is logically compatible with there being more than one such, or with one only, and so (6) is logically neutral to uniqueness and does not actually contradict it. Existence and predication entailments are shared between (3) and (6) and hence (3) entails (6) but is not entailed by it.

For any account of natural language Russell's theory of definite descriptions raises the question: what exactly does it mean for an entity to be unique? There are millions of houses out there so how do speaker and hearer co-operate and coordinate

their references so as to understand a given unique one on a given occasion, in the manner of Grice (1975)? The answer given in Hawkins (1978, 1991) appeals to the pragmatic structuring of the universe of discourse and to the existence of pragmatic sets (P-sets) within which uniqueness is achieved in everyday discourse. Hawkins (1991) also explains why indefinite descriptions are neutral to uniqueness on some occasions, but contrast with *the* on other occasions and “implicate” non-uniqueness (*a senator* can refer to one of the 100 senators of the US senate, but *a president* cannot refer to the unique president of the USA).

Consider now the definite article + modifier combinations of the present paper, starting with (8), (9) and (10):

- (8) The same house was sold.
- (9) The identical house was sold.
- (10) The similar house was sold.

Compare (8) and (9) first. The interesting, and at first apparently contradictory, point to note about *the identical house* in (9) is that it makes an existence and a uniqueness claim about the house in question in accordance with Russell’s semantics in (4) and (5), i.e. that there is some house *x* and there is no house *y* non-identical to *x*. But at the same time we have seen that the semantics of *identical* does assert the existence of some other house *y* non-identical to *x* (recall example (2) with *the identical ice hole*)! The apparent contradiction is resolved by appealing to a pragmatically more fine-tuned universe of discourse within which the semantics of *the identical house* is interpreted, for example the “previous discourse set” shared by a given speaker-hearer pair and containing entities that they have talked about (cf. Hawkins 1991). Appropriate usage of *the identical house* can be achieved if this set contains just one house with the property of

being *identical* to a second one whose existence is entailed, i.e. the uniqueness of the one is achieved by being the only one within the relevant set that has the property of being *identical* to some other house.

For *the same house* we have seen, in *John and Mary live in the same house*, that there does not have to be actual referential distinctness and that John and Mary could both be living in a single existing and unique house. This is the interpretation for *same* that appears to merely emphasise Russell's semantics for *the* whereby there is a house x and there is no other house y non-identical to x . The more explicit and logically entailed paraphrase for this is *The house that John lives in is the (same) house that Mary lives in* (the paraphrase goes through both with and without *same*) which compares and equates two definite descriptions, *the house that John lives in* and *the house that Mary lives in*, each of which requires *the*.

Let $\{H_j\}$ stand for the set of properties associated with *house that John lives in*, and let $\{H_m\}$ stand for the set associated with *house that Mary lives in*, and let us represent *the house that John lives in* informally as 'the x ($\{H_j\} x$)' and *the house that Mary lives in* as 'the y ($\{H_m\} y$)'. The former achieves its uniqueness in the usual way by entailing that there is no other y distinct from x of which $\{H_j\}$ holds, i.e. there is only one house that John lives in. The latter achieves its uniqueness also by entailing that there is no other x apart from y of which $\{H_m\}$ holds, i.e. there is only one house that Mary lives in. Now, the crucial additional claim made by *same* is, in the more usual and straightforward interpretation for *John and Mary live in the same house*, that $x = y$, i.e. the unique x of which $\{H_j\}$ holds is in fact identical to the unique y of which $\{H_m\}$ holds. There is an equivalence between these two definite descriptions that figure in the

more explicit and logical paraphrase for this sentence, i.e. between ‘the x ($\{H_j\} x$)’ and ‘the y ($\{H_m\} y$)’, resulting in the possible use of *same*.

But the equivalence between ‘the x ($\{H_j\} x$)’ and ‘the y ($\{H_m\} y$)’ does not have to involve the actual identity of x and y for *the same* to be used appropriately, as we have seen. *John and Mary live in the same house* could mean that John and Mary live in distinct referential tokens of a single house model, and this interpretive possibility is also present in the logically equivalent paraphrase *The house that John lives in is the (same) house that Mary lives in* as well. This meaning involving distinct referents will be pragmatically preferred in a sentence such as *The house that John lives in was sold and so was the same house that Mary lives in* since there is a much more direct description in the event that only one referential token is intended, namely *The house that John and Mary live in was sold*, and its absence in ... *and so was the same house that Mary lives in* will lead to the inference that the referential tokens are different. The meaning of *same* is closer to that of *identical* when the comparison does involve distinct referential tokens, as we have seen. More generally, the form of equivalence between definite descriptions in logical paraphrases that are entailed by appropriate uses of *the same* is an equivalence in pragmatic referential possibilities, not a full logical or semantic equivalence, as we shall illustrate in more detail in the next section. It is also an equivalence that requires only a partial overlap between the definite descriptions being compared, ‘the x ($\{H_j\} x$)’ and ‘the y ($\{H_m\} y$)’, either at the level of the properties that figure in the definite description, $\{H_j\}$ and $\{H_m\}$ etc, or at the level of the referents x and y .

The semantic and pragmatic analysis of *the similar house* in (10) proceeds as for *the identical house* in (9). There is an existing and unique house x within some

pragmatically defined set and the definite description achieves its uniqueness by being the only *x* in that set with partial property matching to some other referent *y* whose existence is entailed and also possibly within the same pragmatic set, resulting in appropriate sequences such as *The house that John owns was sold and so was the similar house that Mary owns.*

Notice finally in this section that sentences corresponding to (8)-(10) with the indefinite article, namely (11)-(13), involve ungrammaticality in the case of **a same*, as we have mentioned, and a different semantics and pragmatic interpretation for *an identical* and *a similar* compared with their definite counterparts:

(11) *A same house was sold.

(12) An identical house was sold.

(13) A similar house was sold.

The indefinite article is incompatible with *same*, we claim, because the semantics of *same* involves a comparison and a pragmatic equivalence between two definite descriptions, each of which already requires *the*, e.g. *The house that John lives in is the (same) house that Mary lives in.* The product of this comparison and equivalence and its reduction to *John and Mary live in the same house* cannot be any less definite and less uniquely referring in the single NP, *the same house*, than it is in the two compared definite descriptions *the house that John lives in* and *the house that Mary lives in* that are being claimed to be pragmatically equivalent (see the further discussion of this in the next section). For *identical house* and *similar house* the description can be either definite as in (9)-(10), or indefinite as in (12)-(13), depending on whether there is a unique direct referent with full or partial property matching to the indirect referent(s). If no uniqueness in the manner of Russell's analysis is claimed for the direct referent, then

the indefinite article will be used for *identical* and *similar* as in (12)-(13), and it will be asserted that there is at least one house (the direct referent) that is fully or partially matching in properties to the indirect referent(s).

3 SAME AND THE PRAGMATIC EQUIVALENCE OF DEFINITE DESCRIPTIONS

Same is inherently relational and comparative. So are *identical* and *similar*. They link and compare one entity with another. But they do so in different ways. Our proposal is that the comparison with *same* is between two logically entailed definite descriptions, with an assertion of “pragmatic equivalence” between them. This pragmatic equivalence can result from the fact that the two definite descriptions have identical reference tokens (a single house token referred to by both, for example) or they may have different reference tokens but a sufficient sharing of semantic properties combined with real-world pragmatic knowledge (of models of house, for example, with their distinctive and criterial features) to justify being called *the same*. The comparison made with *identical* and *similar* is, we claim, more straightforward and less abstract and applies to linguistic and real-world entities directly, not to logically entailed definite descriptions of these entities. The difference between *identical* and *similar* is then one of full versus partial property matching between these entities. The greater variability in the interpretation of *the same* + Noun, which we shall now illustrate, is a consequence of this more abstract and linguistically specified comparison between descriptions of entities, as opposed to the direct comparison between entities themselves which is characteristic of *identical* and *similar*.

One consequence of this form of comparison with *the same* is that it allows for considerable ambiguity with respect to the unique entity that it refers to. This entity can

be realised as one individual token, in the more usual interpretation of *John and Mary live in the same house*. Consider also the more usual interpretation of *Mary was wearing the same dress as yesterday*. Most plausibly yesterday's dress token worn by Mary was also the one worn subsequently. Alternatively there may be different house or dress tokens in these sentences, in their less usual interpretations. In the example *John has the same nose as his father* (example provided by Ekkehard König) the normal interpretation will be that we are talking about two distinct noses.

This ambiguity has sometimes been described in terms of token versus type meanings (and was so described in Hawkins 1978). The notion of a type is not straightforward, however. It is often loosely defined in the literature, and its very existence has been fiercely debated (see e.g. Wetzel 2009, Kearns 2010, Bromberger 1992). Kearns (op. cit.) observes that once we list all the facts about each individual token (for example, in the case of US grizzly bears, how each one eats and behaves and what it looks like, etc.) and the generalizations about each of them, evoking the notion of a unifying type (i.e. the US grizzly bear) seems superfluous.

In the present context it is not helpful to talk about a "type" meaning for *the same* in *John has the same nose as his father*, or in *John and Mary live in the same house* meaning the same model of house. This is in part because *the same* does not oblige us to consider all relevant tokens and compare their properties, which is what we have to do for *identical*, checking that all properties are shared. Moreover, if we try to impose a "type" analysis on distinct tokens we soon encounter theoretical and descriptive problems involving the very definition of the type, what features define it, what properties the entities that belong to it must possess, how it is delimited and how type membership is determined. These are the kinds of problems for which Langacker

(1991, 2005) and Breban (2010, 2011) have proposed an alternative analysis in terms of referential “generalization”, an insight that we believe can be incorporated here for examples such as *John has the same nose as his father*.

What *the same* invites us to compare is two logically entailed definite descriptions, for example *the house that John lives in*, ‘the x ($\{H_j\} x$)’, and *the house that Mary lives in*, ‘the y ($\{H_m\} y$)’. The claim made by *same* is that there is at least some equivalence between these two that is pragmatically sufficient for the comparison and equivalence to be made. In the extreme case $x=y$ and all the properties of $\{H_j\}$ will be identical to those of $\{H_m\}$. This is full logical equivalence between definite descriptions and we see it realised when one and the same house token is involved. But when there are different house tokens ($x \neq y$) there must then be some equivalence between entities at the level of their properties, i.e. between $\{H_m\}$ and $\{H_j\}$, and just how much equivalence there needs to be at the property level, in order to describe relevant items as *the same*, seems to be pragmatically highly variable and context-dependent. That is why we claim that the kind of equivalence between definite descriptions that is required, in general, for the appropriate use of *the same*, may be full-bodied logical equivalence at the one end between referential tokens and their properties, but only a much looser and pragmatically sanctioned equivalence at the other, between some of the properties of $\{H_j\}$ and $\{H_m\}$.²

² Notice that the analysis proposed here for *the same house* extends readily to anaphoric uses in which a definite description refers back to a first-mention indefinite. One of our reviewers raises the following example: *Yesterday I saw a man with a blue jacket on the bus, and today I saw the same man on the train*. In this discourse *the man* (*on the train that I saw today*) is referentially identical and pragmatically equivalent to a number of alternative definite descriptions for this man, based on information given in the

Consider some further examples that illustrate this variability in the interpretive possibilities for *the same* and that justify our position that the kind of equivalence that is required between logically entailed definite descriptions is pragmatic in nature, i.e. an equivalence deemed sufficient in context to justify the claim of “sameness” based on some sharing at least of referential and/or semantic and/or real world properties, not a stricter form of logical equivalence between definite descriptions.

Consider the following:

(14) John and Mary saw the same white rhino.

This example can be interpreted in many different ways depending on the situation being described. The speaker may be referring to just one token (a singular entity ‘white rhino_i’ that appeared either once or twice and that both John and Mary saw, together or on separate occasions) or he could be referring to more than one token of a common white rhino species. If it happens to be clear that there is only one such animal in the relevant context (in a zoo for example), then this single token will be linked to the two

preceding context: *the man I saw yesterday, the man I saw yesterday on the bus*, etc. Hence *the man on the train that I saw today is the same man that I saw yesterday on the bus*. There is a pragmatic equivalence between two entailed definite descriptions in this example, just as there is with *the same house* in the main text. More generally, recall the important point made by Kempson (1988) that the logical form for sentences containing discourse-sensitive items like anaphoric pronouns and definite descriptions must of necessity include contextually given information about individuals and their properties, in order for truth conditions to be assigned. If we don’t know who the pronoun *he* refers to in context, we cannot assign a truth value to a sentence that contains it. Similarly for *the same man*, alternative logically entailed definite descriptions are made possible by contextual information, and pragmatic equivalence between them is then required, we claim, for appropriate uses of *the same* in the usual way.

individuals, John and Mary, who may have seen it on one or more than one occasion. If this is not clear, on a safari for example, then the interpretation will be quite vague with respect to the number of white rhino tokens seen by these two individuals. In fact, neither the speaker nor John nor Mary may have a clue whether there was one or more than white rhino token that they saw.³ The crucial point is that the referential status of *the same white rhino* may be irrelevant or unknowable in a given real-world situation, and this example highlights the flexibility and variability that has to exist pragmatically with respect to the possible referents of definite descriptions entailed by *the same* (*the white rhino that John saw was the same white rhino that Mary saw*). When our world knowledge is more constrained and there is just one white rhino (in the zoo), this indeterminacy is much less.

With respect to the descriptive content of the definite descriptions entailed by *the same*, in the event that referential tokens are shared (*the house that John lives in is the same house that Mary lives in*) then the real-world properties of {Hj} and {Hm} will, of course, be identical too. In the event that referential tokens are not shared (*John has the same nose as his father*) then the normal expectation may also be that the properties of John's nose {Nj} are identical to those of his father {Nf}, and this may even be an implicature in the sense of Grice (1975), Sperber & Wilson (1995) and

³ This indeterminacy is readily compatible with an intensional semantic account, rather than the extensional approach developed here. The present paper builds on Russell (1905) (see Section 2) and on the pragmatically structured universe of discourse proposed by Hawkins (1978) within which Russell's theory of definite descriptions is interpreted. For a clarifying summary and discussion of the relationship between extensional and intensional approaches to definite descriptions, see Fitting (2015).

Levinson (2000), requiring cancellation in different contexts in the event that the implicature is not intended, as seen in the following examples involving jackets:

- (15) (a) John and Bill were wearing the same jacket but for the buttons.
 (b) John and Bill were wearing the same jacket but for the colour.
 (c) ?John and Bill were wearing the same jacket but for the style.
 (d) *This jacket and this dress are the same.

Cancellation of the implicature ('all properties of the coreferential tokens are shared with *same*') is possible in (15a) and even in (15b). Two jacket tokens worn by two individuals (or by one individual at different times or by one or more than individual at different times and places) can qualify as *the same*, whether or not the buttons are property for property identical. Even the colour can vary, as in (15b). Example (15c) is less felicitous, however, since the style of a jacket is a more inherent part of it, and the absence of this match renders it less plausible to view one of these entities as pragmatically equivalent to the other. (15d) is completely unacceptable because two entities that are as different as a jacket and a dress cannot be judged to be *the same*.

An extreme case illustrating the role of real-world knowledge in sanctioning an acceptable equivalence between two definite descriptions at the property level is the following. Every massive California redwood tree begins its life cycle as a tiny seed. Comparing a token of each, the fully grown tree and the seed, it would be possible to say: *These are the same (tree)*; thereby establishing the link across times between the fully grown redwood and the seed that is, despite appearances, *the same tree*. The entities in question are known to change radically over time, but their very different appearances at different stages are not sufficient to block the pragmatic equivalence of two definite descriptions (*the redwood tree that is fully grown is the same tree as the*

seed) describing exemplars at extreme ends of the life cycle with minimal property overlap but a form of identity provided by real-world knowledge. Notice, by the way, that we could not say that *The California redwood and the tiny seed are identical* (nor even that they are *similar*, see section 5), since hardly any of their properties match. But we can say that they are *the same (tree)*, because of the time and life cycle link that is known to connect the definite descriptions describing these tokens.

Returning to jackets and the like, notice that whereas different colours may not be sufficient to block pragmatic equivalence between definite descriptions for these items, they may do so for others. Compare (16a) and (16b):

(16) (a) ?Their eyes were the same except for the colour.

(b) The two women were the same except for the colour of their eyes.

In (16a) colour is an important and salient property of any pair of eyes, and one pair would not normally be deemed equivalent to another if that feature contrasts. Two women as a whole, however, possess many more properties than their respective eye colours, and so the comparison of one woman token with another can overlook this particular feature and establish sufficient equivalence between them at the level of their properties, despite their different eye colours. In the more limited context of eyes alone colour is too important a difference to render an equivalence plausible.

Notice finally in this section the interesting compound form *selfsame* in Modern English, which through the addition of the reflexive and intensifier form *self* to *same* results in preferred readings for *the same* that stress referential identity and co-reference rather than distinct referential tokens: *John and Mary live in the selfsame house* has this preferred interpretation. So does *John and Mary live in one and the same house*, where the addition of *one* also encourages the single referent interpretation. Interestingly, the -

self of Modern English *himself*, *herself*, etc, was first used in the form of an adjective modifier of nouns and written *self* or *sylf* in Old English with the meaning of Modern English ‘same’, as in *se sylfa God* ‘the same God’ (from König & Siemund 1999: 57, see also Farr 1905: 18). The cognate form *selb-* in Modern German still translates *same* in Modern English (see König & Siemund *ibid* and Gast 2006: 3). For *selfsame* in English, see more generally König & Siemund (2000) and Keenan (2002) on the origins of reflexive pronouns in English and on their related intensifier uses (*The Queen_i herself_i was there*), and for other lexical overlaps across languages between items expressing identity between referential tokens see König & Siemund (2005).

4 IDENTICAL

Semantically *same* and *identical* seem at first to be close (recall Section 1). Their historical origins are different. *Same* is inherited in English from Common Germanic (via Norse, see Faarlund & Emonds 2014), whereas *identical* came from French (cf. Davidse et al. 2008: 479).

Identical occurs with both *a* and *the*, as we have seen, though the use of the latter is pragmatically restricted to cases where there have been previous mentions, as in (17a) and also (18b) below.

- (17) (a) Jackie was wearing an identical dress to the one that I was wearing. The identical dress caused a lot of trouble at the party!
- (b) Jackie was wearing a similar dress to the one I was wearing. The similar dress caused a lot of trouble at the party!

We mentioned in Section 1 that a central difference between *the same* and *identical* lies in the number of entities referred to, namely reference to at least two for *identical* vs.

one or more for *the same*. For expressions with *identical* the identity is actually a predication (of full property-for-property matching) applied directly to these two distinct entities. For *the same* the comparison is between two logically entailed definite descriptions that describe the single or plural referential tokens in question. The co-occurrence restrictions of articles provide independent syntactic evidence for saying that *identical* does not in fact overlap semantically with *same*: *identical* pairs better with *similar* since they can both co-occur with *a*.

Identical and *similar* share a fundamental semantic feature: they predicate their description, namely full versus partial property matching respectively, of at least two separate entities. When they occur within referential expressions they maintain this plurality of reference. The meaning of *identical* can be described as universal quantification over all the properties of these distinct entities (i.e. all their properties are shared). The meaning of *similar* involves existential quantification (some of their properties are shared). Plurality of entities is inherent in the meaning of phrases containing *identical* and *similar*, whereas *same* occurs in NPs that can have a single referent. This contrast can be seen again graphically in the following minimal pair:

(18) (a) I have just seen the same twin (that I saw yesterday).

(b) I have just seen the identical twin (to the one I saw yesterday).

(18a) refers to ‘the twin_i’ that I have just seen and links this entity to the coreferential ‘the twin_i’ that I saw yesterday. (18b), by contrast, links ‘the twin_i’ that I have just seen to the other, referentially distinct twin ‘the twin_j’, that I saw yesterday. The use of *identical*, which is appropriate and commonplace when referring to twins, necessarily involves a referential plurality and a reference to two quite distinct individuals; *the same twin* involves coreference here to a unique one. Notice, interestingly, that *the identical*

twin in (18b) is appropriate as a first-mention definite NP referring back to the other twin (*Yesterday I saw one twin: Today I saw the identical twin*). This further confirms the plurality of reference inherent in *identical*: (18b) refers directly to ‘the twin_i’ that I have just seen while at the same time asserting the existence of another ‘twin_j’ and it acquires its anaphoric definiteness by reference to this other, previously mentioned ‘twin_j’. Since *identical*, like *similar*, involves reference to a plurality, these modifiers can combine with *a* to refer non-uniquely to just one of a plural set (*an identical twin* like *an identical coat*, also *a similar coat* [to the one Mary was wearing]). In a parallel way *a prince* can refer to one of the princes of England, *a senator* to one of the US senators, *a window* to one of the house’s windows, and so on (Hawkins 1978, 1991).

Consider now some further examples that highlight this difference between *identical* and *the same*:

- (19) (a) Jane was wearing an identical jacket to the one Mary wore yesterday; in fact, it was the same [one]!
- (b) Jane was wearing an identical jacket, but it was not the same [one that Mary wore yesterday].
- (c) Jane was wearing the same jacket, not an identical one.
- (d) Jane was wearing the same jacket that Mary wore yesterday, or rather, an identical one.

In (19a) the speaker first believes that there are two jacket tokens and then realises that there was just a single one and so corrects the reference to *the same*, after first using *an identical*. Example (19b) clearly signals that the jacket in question was a separate token from the token in an earlier reference (to the jacket that Mary wore the day before). Example (19c) emphasises the singularity of the jacket token, i.e. there was only one

jacket, not two that were identical. In (19d) we have the same situation as in (19b), but in reverse. The speaker first thinks there was only one token, and then realises that there were actually two and decides to correct himself.

When the pragmatic interpretation of *the same* involves distinct referential tokens, however, the contrast with *identical* is much less and both (20a) and (b) can be used:

(20) (a) Mary and Jane wore an identical jacket.

(b) Mary and Jane wore the same jacket.

In (20a) two distinct jacket tokens are being referred to, and it is asserted that they match one another in all their properties, hence they are *identical*. (20b) with *the same* has the more usual interpretation that there are again two jacket tokens and not just one, as there might be in a circus act with two people in one jacket, and it achieves this similar referential effect through the comparison and pragmatic equivalence between logically entailed definite descriptions, *the jacket that Mary wore is the same jacket that Jane wore*. It is commonplace in everyday language use for one and the same situation to be describable in different ways (e.g. *the candle is on the candle holder* vs. *the candle holder is under the candle*) and this is what is going on in (20). A single situation or event can be conceptualised and lexicalised in different ways. The manner in which similar reference is achieved in (20a) and (b) is also different, as we have seen.

Notice that even though *identical* involves the matching of all properties across different tokens in the normal case, a certain latitude in its descriptive meaning is also permitted. The following are appropriate uses of *identical* even when the sharing of all properties is explicitly denied:

(21) (a) Jackie was wearing an identical dress to mine, except for the colour.

(b) [pointing to a pair of gloves] I bought him an identical pair of gloves, just a bigger size.

Sentences (21a) and (b) are understood as qualifications of, or exceptions to, the full property-for-property matching claim, but not as contradictions. Universally quantified sentences with *all* also permit such exceptions without apparent contradiction, as in *All the boys were having a good time, except for Charlie*. The exception qualifies and does not contradict the universal claim made about all the boys, just as (21a) and (b) qualify and do not contradict the claim that all properties are shared between the two dresses and two pairs of gloves. What we wish to highlight here is simply the difference between *identical* and *same* with respect to the number of entities referred to, a necessary plurality for *identical* versus a possibly single one with *same*, as well as the manner in which the comparison is made, through direct reference to entities on the one hand versus a comparison of definite descriptions on the other. The plurality of reference tokens required for *identical* makes this item closer to *similar* than to *same*, and this is reflected in their parallel article co-occurrence possibilities contrasting with *the/*a same*. On the other hand the pragmatic equivalence between definite descriptions logically entailed by the use of *the same* may involve a sharing of all referential and semantic properties between definite descriptions, or it may involve only a partial sharing of these properties (recall *the same white rhino* and *the same redwood tree* examples above), which means that *same* can overlap semantically with both *identical* and *similar* on different occasions of use.

5 SIMILAR

Similar is usually preceded by the indefinite article, except when it is previously mentioned and the context allows the use of the definite article (recall (17b)), as was the case for *identical* (in (17a)). Grammatically, *similar* patterns like *identical*, therefore. Semantically, it differs from *identical* in that the concept of similarity involves the sharing of only some criterial properties and not of all. For example:

- (22) (a) Jackie was wearing a similar dress to the one I was wearing.
 (b) Bill bought a similar car to ours.
 (c) He saw a similar tree to the one we saw. The similar tree was by the orange house down the road.

In examples (22a-c) we have different tokens that share a number of properties (some but not all), by virtue of which they can be referred to as similar to one another. This near-identity together with the potential variability with regard to which properties are shared and which are not allows for an infinity of similarity descriptions and hence for the indefinite article (unless prior reference is available, as in 22c). *The same*, by contrast, does not have a focus on property-matching between distinct individuals but rather on the pragmatic equivalence between definite descriptions that are logically entailed by the sentence containing *the same*.

The reason why *the* is uncomfortable with *similar* is because the sharing of only some criterial properties makes the existence of other, similar entities inevitable, in general, which conflicts with the uniqueness of definiteness. When the nature and number of other similar entities can be brought under pragmatic control, as in (22c), the definite article becomes possible with *similar*. The interpretation of the definite description in (22c) is relativised to a set of just two trees and because there are just two entities in the pragmatic set under consideration, each becomes unique compared to the

other (see also Hawkins 1978: 250). This conflict between *the* and *similar* is also present between *the* and *identical*. In contrast to *identical* and *similar*, the uniqueness of definiteness maps perfectly onto the meaning of *same*, since *same* involves a form of double definiteness and pragmatic equivalence between two definite descriptions that are logically entailed by a sentence containing *the same*. Since each compared definite description is uniquely referring and requires *the*, the product of the comparison, *the same house* etc. can be no less definite and uniquely referring.

Notice finally in this section that the adjective *different* is in many ways the negative counterpart to *similar* in English, predicating of two or more entities that 'at least some properties are NOT shared between them', as in *John's jacket is different from Bill's*.

6 CONCLUSION

We have given an analysis for some of the different meanings and uses of the semantically related modifiers *same*, *identical* and *similar* in English and accounted for differences in their combinability with articles. *Same* and *identical* differ with respect to differences in their combinability with articles. *Same* and *identical* differ with respect to the number of referents they compare and the manner of the comparison. With *identical* the comparison applies directly to referential entities: there must be two or more such entities, of which the sharing of all properties is predicated in the normal case. This feature of *identical* is shared with *similar*, though only some properties are now asserted to be shared. With *same* the comparison is more abstract and applies to linguistic descriptions of the relevant entities. Specifically *same* involves a comparison and a claimed pragmatic equivalence between two definite descriptions that are logically

entailed by sentences containing *the same* + Noun. Since the entailed descriptions are definite, and since there is an assertion of pragmatic equivalence between them, the single Noun Phrase with *same* in the sentence that entails them must also be definite and *same* must be preceded by *the*, and only by *the*. Hence the ungrammaticality of **a same*.

Identical and *similar*, by contrast, can both occur with both articles in English since their (direct) referents can be potentially non-unique or unique. Despite the fact that *identical* is semantically closer to *same* in terms of the sharing of all properties (this being implicated for *same*, unless cancelled recall (15), and entailed for *identical*, unless the entailment is explicitly denied, recall (21)), unique identifiability of the referent is not a part of the meaning of *identical* or *similar*, as it is with *same*. There is no semantic and grammatical requirement for a co-occurring definite article with *identical* and *similar*, therefore, as there is for *same*, and the choice of *a* versus *the* will reflect the availability or otherwise of appropriate reference tokens in the relevant pragmatic set (cf. Hawkins 1991) containing the entities to which *the identical house* or *an identical house* refer.

We can conclude that the English articles, and syntactic grammaticality judgements involving their co-occurrence, provide a unique insight into both the semantics of *same*, *identical* and *similar*, and into the extreme sensitivity of the syntax to their semantic properties. These article+modifier ungrammaticalities pattern like the many others discussed in Hawkins (1978, 1991) that are all fine-tuned to semantic differences. It would be impossible to write syntactic rules predicting all and only the grammatical sentences of English in this area without the grammar having access to these semantic and pragmatic distinctions in some form, as argued in Hawkins (ibid).

Since this conclusion is inescapable for this area of English, once subtle details of the syntax and semantics of the noun phrase are properly exposed and analysed, it is prima facie plausible to assume that the same relationship between the syntax and semantics holds for all other areas of English grammar, and indeed for the grammars of all languages.

Finally, the quote from Edmund Wilson at the beginning of this paper, *No two persons ever read the same book*, gives us interesting confirmation for the essential idea proposed in this paper. On this occasion the negative universal claim in *no two persons* quantifies over as many definite descriptions as there are people reading the relevant book: *the book in question that the first person reads is not the same book that the second person reads and each is not the same book that the third person reads*, and so on for as many readers as there are of this book. Edmund Wilson's point is that even though people may literally be reading what can be called *the same book*, they are not really doing so since they each bring their own background, knowledge base, historical context, opinions, feelings and attitudes to the content, and consequently they can each have a very different understanding of, and different reactions to, what they read. Putting this in the terms of this paper, there is a pragmatic equivalence between *the book in question that the first person reads* and *the book that the second person reads*, etc, sufficient to justify calling them *the same book*, but the real-world differences between readers and the general context of their reading are so significant as to make it a different reading experience for each one.

Author I's Address:

Department School of Politics, Philosophy, Language and Communication Studies

Institution University of East Anglia

Full postal address including zip or postal code:

Norwich Research Park, Norwich, NR4 7TJ, UK

Email: l.filipovic@uea.ac.uk

Author 2's Address:

Department Linguistics

Institution University of California Davis

Full postal address including zip or postal code:

Kerr Hall, 1 Shields Avenue, Davis, 95616, California, USA

Email: jhawkins@ucdavis.edu

REFERENCES

- Barker, Chris. 2007. Parasitic scope. *Linguistics and Philosophy* 30 (4), 407–44.
- Birner, Betty J. 2013. *Introduction to Pragmatics*. Chichester: Wiley-Blackwell.
- Brasoveanu, Adrian. 2011. Sentence-internal different as quantifier-internal anaphora. *Linguistics and Philosophy* 34 (2), 93-168.

- Breban, Tine. 2010. *English Adjectives of Comparison: Lexical and Grammaticalized Uses* (Topics in English Linguistics Series 63). Berlin and New York: Mouton de Gruyter.
- Breban, Tine. 2011. Secondary determiners as markers of generalized instantiation in English noun phrases. *Cognitive Linguistics* 22 (3), 511-33.
- Bromberger, Sylvain. 1992. Types and tokens in linguistics. In Sylvain Bromberger (ed.), *Essays on What We Know We Don't Know*, 170–208. Chicago: University of Chicago Press.
- Christophersen, Paul. 1939. *The Articles: A Study of their Theory and Use in English*. Copenhagen: Munksgaard.
- Davidse, Kristin, Tine Breban & Ann Van linden. 2008. Deictification: The development of secondary deictic meanings by adjectives in the English NP. *English Language and Linguistics* 12 (3), 475-503.
- Faarlund, Jan Terje & Joseph E. Emonds. 2014. *English: The Language of the Vikings*. Olomouc: Palacký University.
- Farr, James M. 1905. *Intensives and Reflexives in Anglo-Saxon and Early Middle-English*. Baltimore: J.H. Furst Company.
- Fitting, Melvin. 2015. Intensional logic. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2015 Edition), URL = <http://plato.stanford.edu/archives/sum2015/entries/logic-intensional/>.

- Gast, Volker. 2006. *The grammar of Identity: Intensifiers and Reflexives in Germanic Languages*. London and New York: Routledge.
- Grice, Paul. 1975. Logic and conversation. In Peter Cole & Jerry Morgan (ed.), *Speech Acts*, 41-58. New York: Academic Press.
- Hawkins, John A. 1978. *Definiteness and Indefiniteness*. London: Croom Helm, and Atlantic Highlands, New Jersey: Humanities Press.
- Hawkins, John A. 1991. On (in)definite articles: Implicatures and (un)grammaticality prediction. *Journal of Linguistics* 27: 405-442.
- Jespersen, Otto. 1949. *A Modern English Grammar On Historical Principles*, vol. 7. Copenhagen: Munksgaard.
- Kearns, Stephen. 2010. Book review of *Types and Tokens* by Linda Wetzel. *Notre Dame Philosophical Reviews*. Retrieved from ndpr.nd.edu/news/24272-types-and-tokens-on-abstract-objects/
- Keenan, Edward L. 2002. Explaining the creation of reflexive pronouns in English. In Donka Minkova & Robert Stockwell (eds.), *Studies in the History of English: A Millennial Perspective*, 325-55. Berlin and New York: Mouton de Gruyter.
- Kempson, Ruth. 1988. Grammar and conversational principles. In Frederick Newmeyer (ed.), *Linguistic Theory: Extensions and Implications (Linguistics: The Cambridge Survey, vol 2.)*, 139-63. Cambridge: Cambridge University Press.
- König, Ekkehard & Peter Siemund. 1999. Intensifiers and reflexives: A typological perspective. In Zygmunt Frajzngier & Traci S. Curl (eds.), *Reflexives: Forms and Functions*, 41-74. Amsterdam/Philadelphia: John Benjamins.

- König, Ekkehard & Peter Siemund. 2000. The development of complex reflexives and intensifiers in English. *Diachronica* XVII (1), 39-48.
- König, Ekkehard & Peter Siemund. 2005. Intensifiers and reflexives. In Bernard Comrie, David Gil & Martin Haspelmath (eds.), *The World Atlas of Language Structures*, 194-97. Oxford: Oxford University Press.
- Langacker, Ronald W. 1991. *Foundations of Cognitive Grammar, vol.2 : Descriptive Application*. Stanford: Stanford University Press.
- Langacker, Ronald W. 2005. Dynamicity, fictivity, and scanning: The imaginative basis of logic and linguistic meaning. In Diane Pecher & Rolf A. Zwaan (eds.), *Grounding Cognition: The Role of Perception and Action in Memory, Language and Thinking*, 164-97. Cambridge: Cambridge University Press.
- Levinson, Stephen C. 2000. *Presumptive Meanings*. Cambridge, MA: MIT Press.
- Lyons, Christopher. 1999. *Definiteness*. Cambridge: Cambridge University Press.
- Russell, Bertrand. 1905. On denoting. *Mind* 14, 479-493.
- Schwarz, Florian. 2009. Two Types of Definites in Natural Language, PhD dissertation, University of Massachusetts, Amherst.
- Sperber, Dan & Deirdre Wilson. 1995. *Relevance: Communication and Cognition*, 2nd edition. Oxford: Basil Blackwell.
- Wetzel, Linda E. 2009. *Types and Tokens: On Abstract Objects*. Cambridge, MA: MIT Press.