The German Definite Article and the ‘Sameness’ of Indices

Emily A. Hanink
University of Chicago

This paper is posted at ScholarlyCommons. http://repository.upenn.edu/pwpl/vol23/iss1/9
For more information, please contact repository@pobox.upenn.edu.
The German Definite Article and the ‘Sameness’ of Indices

Abstract
The German definite article may contract with a preceding preposition under certain circumstances; the contracted form is referred to in the literature as weak, while the non-contracted form is referred to as strong. Schwarz (2009) gives an analysis of this contrast according to which the weak form is required when the referent of an NP is unique, while the strong form is required when it is also anaphoric, i.e., when it refers back to an antecedent. However, as Schwarz himself points out, anaphoric uses in which the anaphoric NP is modified by the adjective same surprisingly surface with the weak form, and not the strong. The use of the weak form with the clearly anaphoric uses of same pose a challenge to the generalization that anaphoric uses of the definite article always require the strong form. I provide an account of the strong/weak distinction in the German definite article that explains the puzzling use of the weak form in anaphora involving same by proposing the following. P-D contraction in the general case is achieved through P-D Lowering (Embick and Noyer 2001). In the strong form however, D selects for the index-hosting head idx, to which it may lower and bleed the environment for P-D contraction. However, D may optionally not lower to idx, in which case P-D contraction freely occurs while idx spells out as same. Same in this account is therefore treated as an allomorph of an otherwise non-exponed anaphora-encoding head that is usually occupied by D. This account draws support from cross-linguistic evidence from English and Hebrew that same may undergo alternations with pronominal expressions.
The German Definite Article and the ‘Sameness’ of Indices

Emily A. Hanink

1 Introduction

The German definite article may contract with a preceding preposition under certain conditions, as demonstrated by the contrast in (1). The contracted form is referred to in the literature as weak (1a), while the non-contracted form is referred to as strong (1b).

(1)  a. Hans ging zum Haus.
    Hans went to the house
    ‘Hans went to the house.’
  b. Hans ging zu dem Haus.
    Hans went to the house
    ‘Hans went to the house.’

Schwarz (2009: 7)

Schwarz (2009) gives an analysis of this contrast according to which the weak form is required when the referent of an NP is unique, while the strong form is required when it is also anaphoric, i.e., when it refers back to an antecedent. However, as Schwarz (2009) himself points out, anaphoric uses involving the modifier same like (2) surprisingly surface with the weak form, and not the strong: 1

(2)  a. Es hängt an einem Haus. #Am/an dem Haus findet ihr eine Jahreszahl...
    It hangs on a house. On the house you’ll find a date.
    ‘It’s hanging on a house. On the house you’ll find a date...’
  b. Es hängt an einem Haus. Am/#an dem selben Haus findet ihr eine Jahreszahl...
    It hangs on a house. On the same house find you a date.
    ‘It’s hanging on a house. On the same house you’ll find a date...’

What is puzzling about (2) is that the modifier selb- (same) triggers contraction of the definite article even though it clearly signals an anaphoric relation, posing a challenge for the generalization that anaphoric uses require the strong form. In order to account for the puzzle in (2), I argue for the following proposal. The strong form of the article in German selects for an index-hosting head that intervenes between D and N (cf. Schwarz 2009). This head either spells out together with the definite article, blocking contraction in the general case, or is itself realized by the modifier same, forcing contraction even in anaphoric uses. The weak form has no extra structure, just as in Schwarz (2009). The benefit of this structural proposal is that it both accounts for the surprising behavior of the modifier selb- and allows for a single denotation for the definite article across all of its uses. The proposal is moreover supported by cross-linguistic evidence from English and Hebrew that same may alternate with pronominal expressions.

The structure of this paper is as follows. Section 2 lays out a truncated version of the distribution of the strong and weak form. Section 3 reviews Schwarz’s (2009) seminal account of the German definite article. Section 4 then presents the analysis, section 5 provides a morphological account as well as cross-linguistic evidence, and section 6 concludes.

2 Distribution

This section describes various uses of both forms of the definite article in German as put forth by Schwarz (2009). There are in fact other uses of the strong form discussed in that work that I do not

---

1I express my many thanks to Karlos Arregi, Rajesh Bhatt, Amy Rose Deal, Itamar Francez, Anastasia Giannakidou, Julian Grove, Chris Kennedy, Ruth Kramer, Jason Merchant, Line Mikkelsen, Kjell Johann Sebo, Florian Schwarz, and Ming Xiang for discussion of the data and analysis presented here, as well as to the anonymous reviewers of PLC 40 and the audience there for their constructive feedback.

1http://www.geocaching.com/geocache/GC47BF5_mmm-schmeiss-weg/?guid=b467c1c0-2552-4fc5-8481-d58c9908da3f
discuss here, due to limitations of space. The main purpose of this section is therefore simply to illustrate the general contrast between anaphoric and non-anaphoric uses of the definite article.

2.1 Strong Form

The canonical use of the strong form is in anaphoric environments, as in the case of cross-sentential anaphora like (3), in whose second clause the phrase von dem Haus (from the house) refers back to the indefinite antecedent einem großen Haus (a big house).

(3) Fritz wohnt seit Jahren in einem großen Haus. Er schwärmt immer noch von dem Haus.

‘Fritz has lived in a big house for years. He still raves about the house.’

Another use of the strong form is found in some cases of bridging (Hawkins’ (1978) associative anaphora), where a definite description is introduced in virtue of its relation to some antecedent, as in (4), where Autor (author) is introduced in reference to the previously-mentioned Theaterstück (play) and surfaces with the strong form (Schwarz 2009: 53).

(4) Das Theaterstück missfiel dem Kritiker so sehr, dass er in seiner Besprechung kein gutes Haar an dem Autor ließ.

‘The play displeased the critic so much that he tore the author to pieces in his review.’

Schwarz (2009) analyzes these instances of bridging too as anaphoric; the bridged noun is a relational noun, whose internal argument may be supplied by the antecedent (e.g., the author of the play, above). More on this construction as well as Schwarz’s treatment of such examples follows in section 3.

Finally, the strong form is used alongside the external head of restrictive relative clauses. Below, the matrix determiner in in dem Haus (in the house) must surface as the strong form.

(5) Fritz wohnt jetzt in dem Haus, von dem er immer noch schwärmt.

‘Fritz is now living in the house he’s still raving about.’

Restricted relatives are not accounted for in Schwarz (2009), but see Simonenko (2014, 2015) for a non-anaphoric explanation of the use of the strong form in the matrix clause, as well as Hanink and Grove (To appear) and Grove and Hanink (To appear) for accounts assimilating relative clauses to instances of anaphora.

2.2 Weak Form

The weak form is generally used when the referent of a noun phrase is unique but not anaphoric (Schwarz 2009). Because it is also used in idiomatic expressions and in other situations where uniqueness does not necessarily hold, the weak form should be treated essentially as the elsewhere form (though Schwarz 2009 does not cast its distribution in these terms).
The second is the larger situation use, in which the referent picked out by the noun phrase Bürgermeister (mayor) is contextually salient in a larger sphere, which in (7) would consist of a town with exactly one mayor (Schwarz 2009: 40).

(7) Der Empfang wurde von dem Bürgermeister eröffnet.
the reception was by+the/bys the mayor opened 
‘The reception was opened by the mayor.’

In addition to situational uses, the weak form is likewise used in the majority of bridging contexts. Specifically, it is used when the bridged noun is not relational, as in (8), where the Fenster (window) is contextually related to its antecedent Zimmer (room), though the latter cannot be construed as the internal argument of the former (Schwarz 2009: 55).

(8) Das Zimmer war angenehm eingerichtet und das Fenster gab es sogar Vorhänge.
The room was pleasantly decorated and on+the/on the window gave it even curtains.
‘The room was pleasantly decorated and there were even curtains on the window.’

3 Schwarz 2009

In Schwarz’s (2009) semantic proposal, the strong and weak forms of the article require different presuppositions to be satisfied. The weak form carries a presupposition of uniqueness, while the strong form carries a presupposition of anaphoricity.3 The two forms also exhibit structural differences. The weak form on the one hand has a simple syntax and makes use of a Strawsonian denotation for D, as shown in (9).

(9) a. [[the\textsubscript{weak}]]: $\lambda P(e,t) : \exists x(P(x)) \& \exists y[P(x) \& x=y]$  

b. 

\[DP \quad \underline{D} \quad NP\]

The strong form on the other hand has a more complex structure, and requires a different denotation for D that encodes an extra argument. Along the lines of Neale (2004); Elbourne (2005, 2008), Schwarz (2009) posits an index in the strong form as a specifier and therefore must encode anaphoricity into the denotation of the definite article itself, as in (10).

(10) a. [[the\textsubscript{strong}]]: $\lambda P(e,t) : \lambda y : \exists x(P(x) \& x=y)$  

b. 

\[DP \quad \underline{1} \quad \underline{D'} \quad D \quad \underline{NP}\]

The specifier 1 denotes an individual, which saturates the y argument of D and enters an equivalence relation with x. 1 is assigned a meaning by the Traces and Pronouns rule of Heim and Kratzer (1998).

(11) $[[1]]^g$: g(1)

The example in (12) demonstrates how composition works in the case of the strong form.

Fritz lives since years in a big house. He raves always still from the Haus.
‘Fritz has lived in a big house for years. He still raves about the house.’

\[3\]In Schwarz (2009), uniqueness is derived via situations, building on Kratzer (1989). For ease of presentation I do not include situation variables here.
In order to extend the account to strong-form bridging, Schwarz shows that only relational nouns are allowed to surface with the strong form, which is explained if they are anaphoric to their internal argument.

(13) a. \[
[\text{[author of the play]}]: \lambda x \lambda y [\text{author}(x)(y)]
\]

b. Das Theaterstück missfiel dem Kritiker so sehr, dass er in seiner Besprechung kein gutes Haar an dem Autor ließ.

‘The play displeased the critic so much that he tore the author to pieces in his review.’

To account for strong-form bridging, he encodes a relation directly into the meaning of the definite article, requiring a third denotation for D.

(14) \[
[\text{[the relational-noun]}]: \lambda R_{(e, (e))}. \lambda z_t x_e [R(y)(x) \& y = z]
\]

3.1 The Same Problem

Returning now to the focus of this paper, examples like (2) (repeated below in (15)) pose problems for Schwarz’s account. Such examples show that the weak form is used felicitously in anaphoric uses of the definite article, as long as the modifier same is present.

(15) Es hängt an einem Haus. Am selben Haus findet ihr eine Jahreszahl...

‘It hangs on a house. On the same house you’ll find a date...’

The co-occurrence of the anaphoric modifier with the weak form poses a problem for the anaphoricty/uniqueness dichotomy. However, because same quite clearly signals an anaphoric relation, Schwarz suggests that the function of same might in fact be related to that of the index. He proposes a tentative property denotation for the modifier instead of an individual meaning, which essentially does the work of the specifier selected for in strong uses of the article:4

(16) \[
[\text{selbig}_1] = \lambda x [x = g(1)]
\]

However, this modifier is structurally lower than the specifier would be otherwise, and, as he points out, there is no clear connection between the two.

4 Schwarz (2009) discussed the outdated adjective selbig-, though I discuss the more modern selb- in this paper. The two variants display the same contraction effects. The more colloquial gleich- does not seem to act like selb-, but its behavior warrants further investigation.

4 Proposal

My proposal in a nutshell is that we can account straightforwardly for the contraction facts involving selb- if we treat same as the overt realization of an index-hosting head that intervenes between D
and N in anaphoric uses of the definite article. More specifically, I propose that D in the strong form selects for a property-denoting head idx, housed in its own projection, which hosts an index. This index introduces discourse familiarity (Heim 1982; Roberts 2002) and encodes anaphora. The immediate benefits of this revised syntax are that it explains the contraction puzzle of same as well as the connection between same and the index, and it requires just one denotation for the definite article across both forms. Importantly, this analysis is also an improvement on Schwarz (2009) in that it unites the semantic interpretation with a complete morphosyntactic analysis of contraction.

4.1 Encoding Anaphoricity

In the strong form alone, a phrase idxP intervenes between DP and NP (cf. Schwarz 2009). Because idx is property-denoting, it undergoes Predicate Modification with the NP, leaving the Strawsonian denotation for the definite article available for the anaphoric use. The composition of a DP involving the strong form is as in (17c), given the denotations for idx and D in (17a) and (17b), respectively.

(17) *Strong form*
   a. \([idx] = \lambda x_e[x = g(i)]\)
   b. \([D]: \lambda P_{e:t} : \exists ! x P(x).t x_e[P(x)]\)
   c. DP
           \(D\)
           \(idxP\)
           ideX\(\delta\)
       NP

The anaphora-enabling index hosted by idx is what Heim (1998) termed the *inner* index, which is essentially a bindable index, as opposed to what she called the *outer* index, which may act as a binder. A full derivation is exemplified in (18) (cf. Fox’s (2002) trace conversion).

(18) a. Fritz wohnt seit Jahren in einem großen Haus. Er schwärmt immer noch von dem Haus. ‘Fritz lives since years in a big house. He still raves about the house.’  
   b. DP
           \(tx_x[house(x) & (x) = g(i)]\)
           \(\lambda x_e[house(x)]\)
           \(\lambda x_e[x = g(i)]\)
           \(idx\)
       NP

Just as in Schwarz (2009), the weak form in contrast has no extra structure and selects only for N, as in (19).

(19) *Weak form*
   a. \([D]: \lambda P_{e:t} : \exists ! x P(x).t x_e[P(x)]\)
   b. DP
           \(D\)
           \(NP\)

An example of complete composition involving the weak form is given in (20).

(20) a. Der Empfang wurde vom Bürgermeister eröffnet.  
    the reception was by+the mayor opened  
    ‘The reception was opened by the mayor.’
Returning to cases of bridging, the relational noun selected for by \( \text{idx} \) is now of the wrong type \( (\langle e, \langle e, t \rangle \rangle) \) for composition to proceed. This however can be remedied if \( \text{idx} \) undergoes the IOTA type-shift of Partee (1987) (see also Simonenko (2014, 2015) for the proposal that an index-hosting head may type-shift). Once \( \text{idx} \) denotes an individual rather than a property, it can saturate the internal argument of the relational noun, as in (21): \text{Autor (author)} as a whole then becomes anaphoric because by virtue of its internal argument.

(21) a. Das Theaterstück mißfiel dem Kritiker so sehr, dass er in seiner Besprechung kein gutes Haar an dem Autor ließ.
The play displeased the critic so much, that he in his review no good hair on the author left.

‘The play displeased the critic so much that he tore the author to pieces in his review.’

b. \[
\begin{array}{l}
\text{DP} \\
\text{tx}[\text{author-of}(g(i))] \\
\lambda P_{(e,t)}: \exists ! x P(x). \text{tx}[P(x)] \\
\lambda x_{g(i)} \text{[mayor(x)]}
\end{array}
\]

In this account, there is no additional denotation required to account even for bridging contexts involving the strong form of the article. It is notable that among languages differentiating between two forms of the definite article (see Schwarz (2012) for an overview), the choice between the strong and weak form is not always consistent (e.g., Arkoh and Matthewson 2013; Kang 2014; Simpson and Biswas 2015).

5 The Morphosyntax of Contraction and the ‘Sameness’ of \( \text{idx} \)

With the structure and interpretation of the two forms in mind, I now turn to an account of the morphological process of contraction. I propose that the cross-linguistically common phenomenon of preposition-determiner (P-D) Contraction in the general case is best captured by the post-syntactic movement operation \text{Lowering} (Embick and Noyer 2001) in the framework of \text{Distributed Morphology} (Halle and Marantz 1993) (cf. Svenonius 2012). When adjacent, P always lowers to D in German, resulting in the weak form:

(22) P-to-D Lowering
In the strong form on the other hand, D lowers to \( idx \) before Spell-Out and, as a result, bleeds the environment for P-D contraction:

(23) a. \( D\text{-}to\text{-}idx \) Lowering

\[
\begin{array}{c}
PP \\
P \\
\downarrow \\
DP \\
\downarrow \\
idxP \\
\downarrow \\
D+idx \\
\downarrow \\
NP
\end{array}
\]

b. \( P\text{-}to\text{-}D \) Lowering (vacuous)

(24) Es hängt an einem Haus. **Am\/^{\mathbb{A}}n\ dem** selben Haus findet ihr eine Jahreszahl...

It hangs on a house. On\/+the/on the same house find you a date.

‘It’s hanging on a house. On the same house you’ll find a date...’

I propose that, while \( idx \) has no over realization when D lowers into this position, it spells out as \( same \) when on its own. This flexibility lends itself to the idea that D-to-\( idx \) lowering is in fact optional – a process that I call \( same \) allomorphy, as informally formulated in (25).

(25) **Same-allomorphy:**

When \( idx \) is not occupied by D, insert \( same \).

The two options in cases of anaphora are illustrated in (26). The corresponding vocabulary entries moreover are given in (27).\(^5\)

(26) a. \( D\text{-}to\text{-}idx \) Lowering

\[
\begin{array}{c}
PP \\
P \\
\downarrow \\
DP \\
\downarrow \\
idxP \\
\downarrow \\
D+idx \\
\downarrow \\
NP
\end{array}
\]

---

\(^5\)These entries are highly simplified and are meant simply to expose the idea that \( same \) is an allomorph of \( idx \). Details about pronominal structure and inflection are not fleshed out.
b. No D-to-idx Lowering

\[
\begin{array}{c}
\text{PP} \\
\text{DP} \\
P+D \\
\text{idxP} \\
\text{idx} \\
\text{NP}
\end{array}
\]

(27)  
a. \([an] \leftrightarrow an\) \quad P
b. \([D+\text{dat}+\text{neut}] \leftrightarrow \text{dem}\) \quad D
c. \([an+D+\text{dat}+\text{neut}] \leftrightarrow am\) \quad P+D
d. \([D+\text{dat}+\text{neut}+\text{idx}] \leftrightarrow \text{dem}\) \quad D+\text{idx}
e. \([\text{idx}] \leftrightarrow \text{selb-}\) \quad \text{idx}

_Same_ is essentially the elsewhere form whose contextual allomorph (i.e., the allomorph surfacing in the context of a lowered definite article) is null. It is important to note however that _same_ inflects like an adjective in German, though I abstract away from adjectival endings here. The connection between the category _idx_ and the ability to inflect of course necessitates further investigation.

Aside from the contraction facts, there is moreover further evidence that D does in fact lower to _idx_. Another morphological strategy in German is to pronounce the definite article together with _same_ as a single prosodic word, which I take as overt evidence that D has lowered to _idx_.

(28)  
Es gab in _denselben_ Haus auch eine Welt hinter den Kulissen...

_It was in the same house also a world behind the stage_

_In the house there was also a world behind the stage_...

A clear prediction of the proposal that _selb- _is an allomorph of _idx _is that it be felicitous in exactly those environments where the strong form is otherwise required: this prediction is borne out. First, as we have already seen, _selb- _is licensed in cross-sentential anaphora:

(29)  
Es hängt an einem Haus. _Am selben_ Haus findet ihr eine Jahreszahl...

_It hangs on a house. On the same house find you a date._

_It's hanging on a house. On the same house you'll find a date._

_Same_ is also licensed in strong-form bridging:6

(30)  
Wenn ich ein Buch mochte und _vom selben_ Autor noch eins lese, bin ich sehr oft enttäuscht.

_When I a book liked and from the same author yet one read, am I very often disappointed._

_When I've liked a book and read another by the same author I'm often very disappointed._

Additionally, _selb- _is also licensed as modifier of the external head in a restrictive relative:7

(31)  
Die andere Person folgt der Richtung meines Fingers und sieht genau _dasselbe_ Haus,

_The other person follows the direction of my finger and sees exactly the same house_

_that I indicate_

_The other person follows the direction of my finger and see exactly the same house that I'm indicating._

---


7[Das innere Verbum in Gadamers Hermeneutik: pg. 234.](#)
The felicity of *selb-* in these constructions contrasts clearly with its infelicity in non-anaphoric uses of the weak form, as demonstrated by the following examples of an immediate situation use (32) and a weak-form bridging use (33).

(32) #Das Buch, das du suchst, steht im selben Glasschrank.
   ‘The book that you are looking for is in the same glass-cabinet.’

(33) #Der Kühlschrank war so groß, dass der Kürbis problemlos im selben Gemüsefach untergebracht werden konnte.
   ‘The fridge was so big that the pumpkin could easily be stowed in the same crisper.’

5.1 Hebrew

There is cross-linguistic support for the inherent relation between *same* and pronominal-like expressions. In Hebrew, the third-person accusative pronoun is *oto* (34), though it can also take on the meaning of the modifier *same* (35).

(34) Ra’iti oto.
    saw.1.sg him
    ‘I saw him.’

(35) Karati sefer. Itamar kara et oto (ha)-sefer.
    read.1.sg book. Itamar read.3.sg.m ACC same (the)-book
    ‘I read a book. Itamar read the same book.’

The ability of a pronoun to act as *same* lends preliminary support to the proposal that the two are allomorphs of the same underlying head, though *oto* in Hebrew never looks like an adjective as it does in German (see Matushansky (2010) for a different claim about the Hebrew facts).

English too displays a pronominal-like use of *same*, suggestive of its use as an anaphora-hosting element. For instance, Hardt et al. (2012) and Hardt and Mikkelsen (2015) point out that *the same* is used in elliptical contexts, as in (36).

(36) John told Mary he loved her, but unfortunately she couldn’t say the same/it/that to him.

The use of *same* in such contexts implies a broader cross-linguistic generalization that *same* may take the place of or at least act like an index.

6 Conclusion

The German definite article displays curious behavior when co-occurring with the anaphoric modifier *selb-* (*same*). While the strong, non-contracted form is generally used in anaphoric environments, the weak form is required when *same* is present. This behavior is accounted for if the structure of anaphoric definites necessarily houses an extra phrase, *idxP*, that intervenes between DP and NP. If *selb-* is not present, D generally lowers to *idx* and contraction with P is blocked; otherwise contraction occurs freely. If contraction does occur before D has lowered to *idx*, then *selb-* acts as the overt realization of the index, explaining why the contracted form surfaces even in some anaphoric environments.

The benefits of this analysis are both theoretical and empirical. First, the account relies on just one denotation for D (cf. Schwarz 2009). Second, the account extends to explain the puzzling behavior of the modifier *same* in addition to capturing the fundamental difference between the weak and strong forms. Third, this proposal is supported (preliminarily) by cross-linguistic data from Hebrew and English suggesting that pronominal expressions and the modifier *same* may alternate with one another. The exact contribution of *same* remains to be explored in future research, as does the nature of cross-linguistic variation outside of the languages presented in this paper.

---

*Note that these become good when interpreted anaphorically.*
References


Schwarz, Florian. 2009. Two Types of Definites in Natural Language. Doctoral dissertation, University of Massachusetts, Amherst, MA.


Simpson, Andrew, and Priyanka Biswas. 2015. Bare nominals, classifiers, and the representation of definiteness in Bangla. Handout from the 5th Formal Approaches to South Asian Languages (FASAL) workshop, Yale University, New Haven, CT.


Department of Linguistics
University of Chicago
1115 E. 58th Street
Rosenwald Hall, Room 224
Chicago, IL 60637
eahanink@uchicago.edu