



Philosophische Fakultät

Heike Wiese

So as a focus marker in German

Suggested citation referring to the original publication:
Linguistics 49–5 (2011), 991–1039
DOI: <http://dx.doi.org/10.1515/LING.2011.028>

Postprint archived at the Institutional Repository of the Potsdam University in:
Postprints der Universität Potsdam
Philosophische Reihe ; 102
ISSN 1866-8380
<http://nbn-resolving.de/urn:nbn:de:kobv:517-opus4-93592>

So as a focus marker in German*

HEIKE WIESE

Abstract

This paper discusses a hitherto undescribed usage of the particle so as a dedicated focus marker in contemporary German. I discuss grammatical and pragmatic characteristics of this focus marker, supporting my account with natural linguistic data and with controlled experimental evidence showing that so has a significant influence on speakers' understanding of what the focus expression in a sentence is. Against this background, I sketch a possible pragmaticalization path from referential usages of so via hedging to a semantically bleached focus marker, which, unlike particles such as auch 'also'/'too' or nur 'only', does not contribute any additional meaning.

1. Introduction

In contemporary German, particularly in spoken language, one finds examples like the following that point to a usage of the particle *so* 'so, such' that deviates from the usages commonly described for *so*.¹

- (1) [*Es war ein ganz anderes Turnier als die WM 2006.*]
Wir haben nie so den Schwung kreieren können, der uns
we have never so the impetus_{ACC} create can_{INF} that us
durch das Turnier getragen hat.
through the tournament_{ACC} carried has
[*Wir mussten uns hier durchkämpfen, durchbeißen.*]
'It was a tournament that was totally different from the 2006 World Cup.
We were never able to create the impetus that carried us through the
tournament. We had to fight our way through here, struggle our way
through it.'
(Interview with soccer player Christoph Metzelder in *Die ZEIT*
27/2008²)

- (2) [A: *suche nette sie mit msn bitte melden*
 B: *AHA WAS SUCHST DU IM CHAT?*]
 A: *so coole leute zum kennenlernen*
 so cool people_{ACC} to.the get.to.know_{INF}
 ‘A: seeking nice ‘her’ with msn please get in touch
 B: I SEE WHAT ARE YOU LOOKING FOR IN THE CHAT?
 A: cool people to get to know
 (Chat Community, Dec 5, 2008)
- (3) J. B. Kerner: *So auch auf der STRAÙe kommen die Leute,*
 so also on the street come the people
oder sind das hauptsächlich dämliche
 or are that mainly stupid
Journalisten?
 journalists
 C. Roche: *Ja, also es sind hauptsächlich männliche*
 yes well it are mainly male
Journalisten so, die mit mir über
 journalists so who with me about
sowas reden wollen,
 such.things speak want
[aber es gibt auch Leute auf der STRAÙe, die dann TIPPS
haben wollen, wie die besser im Bett miteinander sind.]
 ‘J. B. Kerner: Is it also in the street that people approach you, or is it
 mainly stupid journalists?
 C. Roche: Well, yes, it is mainly male journalists who want to talk
 to me about such things, but there are also people in the
 street, who then want tips on how to be better in bed
 together.’
 (Conversation in a TV talk show between host Johannes B. Kerner and
 author Charlotte Roche about people asking her about her new book on
 sexual experiences (Oct 21, 2008) (underlining marks passages of
 simultaneous speech))
- (4) *Das sieht so INDISCH aus.*
 that looks so Indian VPRT
 ‘That looks Indian.’
 (Conversation between young people in Berlin-Kreuzberg about a TV
 movie that they are watching)³
- (5) Conversation between young people in Berlin-Kreuzberg about
 European soccer cup (from the same corpus as Example (4), see Note 3).
 [A: *Frankreich gegen Rumänien: null-null.*]
 B: *Aber Rumänien hat rischtisch geil verteidigt so.*
 but Romania has really awesomely defended so

‘A: France against Romania: nil-nil.

B: But Romania {had a really awesome defense / defended really well}.’

- (6) Doctor explaining application of two kinds of cream to a patient.⁴

Die ist für die NACHT, und diese so für TAGSÜBER so.

this is for the night and this.one so for day.over so

‘This one is for the night, and this one, for daytime.’

In these examples, *so* remains unaccented, but stands adjacent to the expression that receives the main sentence stress (cf. the examples where prosodic information is available, i.e., (3) to (6)). Semantically, it is peculiar in not contributing to the meaning of the sentences in which it appears here, and accordingly it does not show up in the final translations, but only in the glosses. There is no good way of translating *so* into English here — except maybe by colloquial *like*: in a sentence such as ‘She’s, like, really nice’, *like* does not contribute additional content to the semantic representation, either, and similarly to *so* in the examples above, it stands in an unusual syntactic position. *like* in such a usage has been analyzed as a focus marker (Underhill 1988; Meehan 1991), and this is what I am going to argue for *so* as well: in what follows, I am going to show that *so* can take on the function of a focus marker in German.

This function of *so* has not yet received linguistic attention, and is not described in major grammars of German such as those of the Institute for German Language (IdS Mannheim) (Zifonun et al. 1997), Duden (Duden 4; 2009), or the Akademie-Grammatik (Heidolph et al. 1981), and it does not appear in scholarly encyclopedias such as those going back, respectively, to Jacob and Wilhelm Grimm (DWB), and to Hermann Paul (Paul 1992), whose objective it is to provide comprehensive, historically oriented listings for lexical entries.

Yet, as one reviewer of this paper pointed out, as the 1961 data in (7) shows, *so* as a focus marker can already be found in spoken language from 50 years ago.⁵

- (7) Recalling a family visit over the Christmas holidays as a child:

[[. . .]und ich hab mit meinem Vati Skat gelernt! Also zwar Anfänger-Skat, nicht?]

Nicht mit so Re und all so'n Käse,

not with so Re and all such.a_{CL} cheese

aber es war ganz prima.

but it was entirely brilliant

Und da haben wir auch all so Tischfeuerwerk,

and there have we also all so table.crackers

haben wir so Stimmungskanonen losgelassen,

have we so party.gags let.go_{INF}

(Lachen) 'swar ganz prima. Und (Pause) so
 laughter it was entirely brilliant and pause so
 Knallfrösche haben wir im Zimmer laufen lassen.
 jumping.crackers have we in.the room run let
 'And I learned Skat [card game] with Daddy! Well, beginners' Skat,
 though, right? Not with "Re" [special bid in Skat] and all that stuff, but
 it was really brilliant. And then we let off table crackers, let off party
 gags (laughs), it was really brilliant. And we set off firecrackers in the
 room.'

In fact, focus-marking usages of *so* might go back quite a bit longer. In the *Grimm Wörterbuch*, the following passage from Lessing's *Nathan der Weise* from 1779 is interpreted as a use of *so* for emphasis ("zur Hervorhebung"), but this might actually be an example of focus marking *so* from 18th century German (DWB: *so*, II.A.2).

- (8) Lessing, *Nathan der Weise*; Act I, Scene 3:
 [Derwisch: [. . .] Lässt sich
 Aus einem Derwisch denn nichts, gar nichts machen?
 Nathan: *Ei wohl, genug! — Ich dachte mir nur immer,*
 Der Derwisch — so der rechte Derwisch — woll'
 the dervish — *so* the proper dervish — want'
 Aus sich nichts machen lassen.
 out.of himself nothing make_{INF} let_{INF}
 'Dervish: [. . .] Can nothing then be made out of a dervish?
 Nathan: Yes, surely; but I have been wont to think a dervish, that's to
 say a thorough dervish, will allow nothing to be made of him.'⁶

In this paper, I will leave open the question of how old, or novel, the use of *so* as a focus marker is,⁷ and will instead concentrate on its characteristics in contemporary German, and its relation to other usages of *so*.

What is special about its usage as a focus marker, is that *so* is a functional, rather than referential, element: it is semantically bleached and reduced to its pragmatic / information-structural function as a marker of the focus constituent in a sentence. This adds another dimension to focus marking in German, which is normally achieved by means of intonation and word order, rather than by functional elements. Note that what is commonly described as "focus particles" in German are particles like *auch* 'also'/'too', *nur* 'only', or *sogar* 'even' (e.g., Jacobs 1983; König 1991), which are not functional elements in this sense. Rather, they are focus-sensitive particles that interact with focus, but contribute their own (restrictive, additive, or scalar) semantics to the sentence, and trigger characteristic presuppositions.⁸

In contrast to this, in *so*, I am going to argue, we see a genuine, dedicated focus marker, similar to those better known from African languages,⁹ and also from creoles.¹⁰ The appearance of *so* as a genuine marker of focus in German, taken together with the evidence for *like* in English, then, indicates that the usage of such functional focus markers is a valid option in Germanic languages as well.

In what follows, I am going to explore this in some more detail. I first summarize conventional usages for *so* in German and show that in constructions like the ones illustrated above, we see an additional, hitherto undescribed function for *so* (Section 2). I then present evidence from a study that tested interpretations of sentences that contained *so* as a putative focus marker, and showed that *so* has a significant influence on the identification of the focus expression in a sentence (Section 3). Against this background, I discuss a possible pragmaticalization path for the emergence of *so* as a focus marker, and suggest a representation capturing its role in the organization of information in a sentence (Section 4). The final section summarizes our results and their contributions to different domains of linguistic analysis (Section 5).

2. *So* in German

So is a multifunctional item in German, which is reflected in the range of different classifications given for it in major descriptive grammars, including deictic adverb, modal adverb, degree particle, intensity particle, augmentive particle, complementizer, subordinating and coordinating conjunction, and correlative to conjunctions.¹¹ In what follows, I will concentrate on the commonalities between the different usages, and in particular on shared aspects of meaning, describing the semantic contribution of different conventional usages of *so* and the semantic relations between them. This will set the stage for a characterization and distinction of its semantically bleached, pragmatic usage as a focus marker, and will allow us to discuss its relation to other usages within a pragmaticalization scenario accounting for its development.

So goes back to an adverb with the meaning ‘in this way’, with cognates such as English *so* or Swedish *så*. This meaning is still at the core of modern German *so*, with a semantic contribution that answers to ‘how’ and is indexical in the sense that *so* is always used as pointer, be it to other linguistic elements (e.g., in anaphoric usages), or to extralinguistic elements (e.g., in manner deixis) that can be present in the communicative situation (e.g., in direct pointing) or just part of the common ground of hearer and speaker. The semantic content answering to ‘how’ covers indexicality to sorts of objects (‘this kind’), activities (‘this manner’), and properties (‘this quality / degree’) alike. I will account for this by calling *so*, in a generalized way, a *modal* indexical. This indexical

modal meaning subsumes a number of different specialized meanings for *so* described in the literature.¹²

In order to show how they come together, I will start from a core meaning of *so* in comparative contexts, which directly reflects this indexical modal status, and describe how other usages are semantically related to it. When doing so, I will show how more peripheral meanings can be derived from the core meaning — without, though, implying any diachronic ordering here.

2.1. Conventional usages of *so*: *so* as an indexical modal element

A central usage of *so* is that as a comparison particle. In this usage, *so* refers to the comparison of properties (9a), activities (9b), or objects (9c), where the object of comparison can be either explicit (introduced by *wie* ‘how’/‘as’) or implicit.

- (9) a. *Annas Hund ist so groß wie ein Kalb.* | ... *ist so groß*
 Anna’s dog is *so* big as a calf | ... is *so* big
 [pointing].
 [pointing].
 ‘Anna’s dog is as big as a calf.’ / ‘Anna’s dog is this big [pointing].’
- b. *Er bellt so, wie mein Großvater lacht.*
 he/it barks *so* as my grandfather laughs
 ‘It barks like my grandfather laughs.’
- c. *Wir machen das so.*
 we make that *so*
 ‘We do it {so / this way}.’
- d. *Paul will auch so einen Hund (wie Anna) haben.*
 Paul wants also *so* a dog as Anna have_{INF}
 ‘Paul wants to have such a dog (a dog like Anna’s) as well.’

In the last usage, (9d), *so* is always followed by the indefinite article *ein-*. For mass nouns, which do not combine with the indefinite article, this means that they have to undergo coercion from mass to count, and accordingly refer to sorts of a substance, rather than to the substance itself (e.g., *so ein Bier* ‘such a kind/sort of beer’). In constructions with plural nouns, which likewise do not combine with the indefinite article, *so* is substituted by *solch-* in standard German, the pronominal counterpart of *so*.¹³ In spoken German, the indefinite article cliticizes with *so*, leading to a form *son-*, which has been described as a type-token article that identifies an indefinite token of a definite type (cf. Hole and Klumpp 2000). In some dialects, this article has plural forms as well, and can hence be used instead of *solch-* (e.g., *sonne Hunde* ‘such dogs’).

A special case of the comparison usage is the combination of *so* with the subordinate conjunction *dass* ‘that’. In this case, the comparison targets events, leading to a resultative meaning ‘such that’:¹⁴

- e. *Der Hund hat so gebellt, dass er heiser geworden ist.*
 the dog has *so* barked that he/it hoarse become has
 ‘The dog barked in a way that it has become hoarse.’
- f. *Er benimmt sich jetzt besser, so dass er bleiben kann.*
 he/it behaves itself now better *so* that he/it stay
kann.
 can
 ‘It behaves better now, so that it can stay.’

so can also function as an intensity marker with evaluative expressions. This usage can be derived from that as a comparison marker if we assume an implicit object of comparison whose quality (or degree) is emphasized. In these constructions, *so* is always stressed.

- (10) a. *Annas Hund ist so dumm.*
 Anna’s dog is *so* stupid.
 ‘Anna’s dog is {so/really} stupid.’
- b. *Er ist so ein Trottel.*
 he/it is *so* a dumbo
 ‘It is {such a / a real} dumbo.’

The indexical modal meaning of *so* is further accessed in somewhat more formal constructions such as (11), where *so* sets up an indexical relation between attributes (11a) or between a general phenomenon and an instance of it (11b).

- (11) a. *So schön er ist, so dumm ist er leider auch.*
so pretty he/it is *so* stupid is he/it unfortunately too
 ‘Unfortunately, it is as dumb as it is pretty.’
- b. *Es gibt hier weitaus klügere Hunde, so der des Milchmanns.*
 it gives here far smarter dogs *so* that of.the
 milkman’s
 ‘There are much smarter dogs around, for instance the milkman’s.’

Another function of *so* is that of a quotative marker in informal speech, similar to English *like*.¹⁵

- (12) *Sie so ‘Sei ruhig!’.*
 she *so* be quiet
 ‘She was like ‘Be quiet!’’

As (12) illustrates, *so*, in contrast to *like*, fulfills the quotative function by itself, without an additional verb such as ‘to be’. Similar to *like*, *so* in this usage is a quote-oriented quotative index in the sense of Gldemann (2008), that is, an indexical element that is oriented towards the quoted material (rather than towards the speaker). As Golato (2000) pointed out, again similar to *like*, this quoted material is not restricted to reported speech, but often involves sound effects, body movements, and gestures as well.

This function of *so* can be linked to that as a comparison marker by way of its indexical character: semantically, the use of *so* as a quotative marker can build on the ‘activity’ aspect of its meaning in comparisons (‘in this way / manner’). In the quotative usage, this yields a modal indexical orientation towards the quote, relating it to a speaker S in the sense of “S acted in this way/manner”. In contrast to constructions with *verba dicendi* such as “S says”, this indexical modal contribution of *so* is not specified for linguistic utterance as the quoted material and thus covers nonverbal actions as well.¹⁶

Related to this informal quotative usage of *so*, but mostly restricted to more formal written styles, often in newspapers, is the use of *so* illustrated in (13), where *so* is backwards rather than forwards indexical, that is, it follows, rather than precedes the quoted passage (from an article on the internet company Fon in the newspaper *Die Zeit* 08/2009).¹⁷

- (13) [„Wir glauben, dass jeder berall Zugang zum Internet haben sollte.“]
so US-Geschftsfhrer Ejovi Nuwere.
so US-CEO Ejovi Nuwere
 ‘We believe that everyone should have internet access everywhere’,
 said US CEO Ejovi Nuwere.’

A further step away from the semantics of a comparison marker is the use of *so* to indicate approximation, as illustrated in (14).

- (14) *Der Hund hat so 20 Euro gekostet.*
 the dog has so 20 Euros cost
 ‘The dog cost around 20 Euros.’

In the intended reading here, *so* does not compare the amount of 20 Euros to something (say, in the sense of “such 20 Euros as those”), but marks it as approximate: we are not talking about an exact amount, but rather a range that extends below and above 20. We can relate this usage to the core meaning of *so* by assuming that the indexicality that *so* has as a comparison marker is lost, with the effect that *so* denotes a generalized, vague modal extension. On the discourse level, this supports a use of *so* as a hedging device.¹⁸

2.2. *An additional usage of so: so as a semantically bleached focus marker*

If we put together the different usages of *so*, we hence get different semantic contributions that access, in a more or less direct way, the central modal indexical meaning of *so* answering to ‘How?’, with its usage as a comparison marker at the core, clustering around it the closely related meanings for resultative, intensifying, and relational usages, a further derived usage as a quotative marker, and, finally, as an approximation marker.

Where do our examples from the beginning come in, then? In comparison to the usages described here, *so*, as mentioned above, is semantically bleached in those examples, it does not contribute any meaning and does not affect the truth value of a sentence: if we removed *so* from the examples in (1) through (7), the meaning of the sentences would not change. In contrast to that, removing *so* from the examples in the previous section (= Examples (9)–(14)) would change the sentence meaning¹⁹ and in a lot of instances (in resultative, relational, quotative, and some comparison contexts) lead to ungrammatical constructions.

This, then, adds a further usage to the ones previously described for *so* in German: in this usage, *so* does not provide any additional content to the sentences in which it appears, but has shifted from lexical-semantic meaning to a purely pragmatic / discourse contribution, namely that of a focus marker.

As such, it is systematically unaccented, as opposed to its other usages (which are ranging from optional stress, e.g., as a comparison particle, to consistent stress, as in intensifier usage),²⁰ while its co-constituent carries the main sentence stress as the focus expression of the sentence, or the phrase containing the focus expression. As Manfred Krifka (p.c.) mentioned, the contrast between unaccented *so* and its co-constituent makes the sentence stress on the latter more salient, which might further support its identification as a focus expression.

Further evidence for its focus marker status comes from the distribution of *so* from a syntactic point of view. *so* shows a remarkable syntactic promiscuity, with all major lexical categories (or their immediate functional projections) attested for its co-constituents in our examples from the beginning, namely NP/DP (Examples (1), (2), and (3)), AP (Example (4)), VP (Example (5)), and PP (Example (6)). For easier access, I briefly repeat the English translations here, which are close enough to the German constructions to illustrate the point, with ‘so’ inserted at the relevant positions (for the German originals and interlinear glosses, see Section 1 above).

- (1’) ‘It was a tournament that was totally different from the 2006 world cup. We were never able to create so [the impetus]_F that carried us through the tournament. We had to fight our way through here, struggle our way through it.’

- (2') 'A: looking for nice 'her' with msn please get in touch
 B: I SEE WHAT ARE YOU LOOKING FOR IN THE CHAT?
 A: so [cool people]_F that I can get to know'
- (3') 'J. B. Kerner: Is it so [also in the street]_F that people approach you, or
 is it mainly stupid journalists?
 C. Roche: Well, yes, it is mainly [male journalists]_F so who want
 to talk to me about such things, but there are also
 people in the street, who then want tips on how to be
 better in bed together.'
- (4') 'That looks so [Indian]_F.'
- (5') 'A: France against Romania: nil-nil.
 B: But Romania [defended really well]_F so.'
- (6') 'This one is for the night, and this one, so [for daytime]_F so.'

Sure enough, in some of the examples, *so* appears in positions that would also work for it in a nonfocus marker usage, at least with different stress patterns, e.g., if *so* received main stress in (4), it could work as an intensifier. But in quite a few of the data, *so* is in positions that would be peculiar or even ungrammatical if it was not used as a focus marker, e.g., in (3), (5), and (6). In its status as a focus marker, though, these positions are not surprising: in order to mark focus, *so* should be expected to vary its position in accordance with the location of focus and thus combine with phrases of different syntactic categories. Not surprisingly, then, *so* shares this characteristic with traditional focus particles, which König (1993: Ch. 4.1) describes as syntactically variable, cross-categorical operators.

so as a focus marker can both precede and follow the focused constituent, adding another dimension of syntactic variability. This, again, is something that can also be observed for traditional focus particles: as König (1993: 982) points out in a crosslinguistic overview, "particles adjacent to their focus may often either precede or follow this constituent", with additive particles in particular usually having the option to follow the focus constituent. The following examples illustrate this for German *auch* 'also'/'too'.

- (15) a. *Sie isst auch in der U-BAHN Erdnüsse.*
 she eats also in the subway peanuts
 'She eats peanuts also [on the subway]_F.'
- b. *Sie isst in der U-BAHN AUCH Erdnüsse.*
 she eats in the subway also peanuts
 'She eats peanuts [on the subway]_F, too.'
- c. *Sie isst in der U-Bahn auch ERDNÜSSE.*
 she eats in the subway also peanuts
 'She eats also [peanuts]_F on the subway.'

Note, though, that, as the contrast between (15b) and (15c) shows, in such sentences, where there are two focus candidates for *auch*, postpositioning *auch* as in (15b) only works if the focus particle itself receives stress as well, and not only its co-constituent.

As our data shows, *so* is much freer in this option and — in accordance with its desemanticized status — remains unstressed in all distributions. What is more, not only can *so* precede as well as follow its co-constituent, but it can even, by way of duplication, form a bracket around it (cf. Example (6) above). So far, the functional distinctions between these options are not altogether clear. Some of the variation between pre- and post-*so* might be induced by the stress pattern of the sentence and support a more regular rhythm, while *so*-bracketing might serve to unambiguously mark the borders of the focus constituent. This will need further investigation, though. For the time being, I will concentrate on the status of *so* as a focus marker, and leave these points exploratory.

2.3. Focus categories in the domain of *so*

If we have a look at the distribution of *so* from the point of view of information structure, we find that it can mark both information focus and contrastive focus. In our examples, information focus appears in (1), where a new aspect, the *Schwung* ‘impetus’ to carry the team through the tournament, is introduced (with some contrastive aspects, as this is described as differing from the 2006 cup), in (2), where the focus expression *coole Leute* ‘cool people’ provides the information asked for in the preceding question, and in (4), where the adjective *indisch* ‘Indian’ introduces a new assessment of a movie setting.

Evidence for contrastive focus comes from our examples in (3), where approaches by random people *auf der Straße* ‘in the street’ are distinguished from those by journalists; (5), where the fact that the Romanian team has *geil verteidigt* ‘awesomely defended’ is set against a nil-nil result for the game; and (6), where the recommended application time *tagsüber* ‘daytime’ for one cream is contrasted with nightly applications of the other cream.

I have not been able to attest a third focus category for the domain of *so* that would be interesting for our discussion, namely ‘metalinguistic focus’ (Jackendoff 1997: Ch. 12.5) or ‘expression focus’ (Krifka 2007: Ch. 3.2), which targets linguistic expressions qua expressions,²¹ but this might be a matter of more searches, rather than a systematic restriction.

An interesting case is the occurrence of *so* in contrastive topics introduced by ‘es gibt _’ or ‘gibt es _’, literally ‘it gives _’ (or ‘gives it _’), a construction used for existential statements similar to English ‘there is _’. In these constructions,

es is the syntactic subject, but as a semantically empty, weak pronominal subject, it tends to be realized not so much in a sentence-initial position before the verb, but rather in the so-called *Wackernagel's position* immediately after the finite verb *gibt*, where it gets cliticized in spoken (and sometimes also in written) language, yielding a form *gibt's* or *gibts*, or even *gibs* with a phonologically reduced coda.²² In contemporary German, this form seems to be undergoing reinterpretation as a monomorphemic element, rather than a complex form containing a pronominal subject *es*, which leads to constructions with an additional phonologically full pronoun *es* in sentence-initial position, as in (16).²³

- (16) *Es gibts nichts schöneres als*
 it *gibts* [lit.: gives-it_{CL}] nothing nicer than
[mit einem guten Buch mit einem Glas Rotwein vor einem offenen
Kamin zu sitzen]
 'There is nothing nicer than sitting in front of the fireplace with a good
 book and a glass of red wine.'

As an existential marker, *gib(t (e))s* is often used to introduce referents that serve as topics in the following discourse, and the emergence of a reinterpreted form *gibs* might actually support the development of a topic marker in German that would constitute a counterpart to the focus marker *so*. What is interesting for our present discussion of *so*, is a construction that combines both elements in a pattern "gib(t)s so NP [Mod]", where Mod is a modifier in the NP, usually a relative clause. (17) gives two examples.²⁴

- (17) a. *Es gibts so Wecker die einen nach*
 it *gibts so* alarm.clocks which one_{ACC} after
Schlafphasen wecken,
 sleep.phases wake.up
[also immer wenn man gerade nicht im Tiefschlaf ist]
 'There are alarm clocks that wake you up after sleep phases, that
 is, whenever you are not in deep sleep at that moment.'
- b. *[falls du den fleischgeschmack an sich vermisst (ging mir nie so),]*
gibts auch so bratwürstchen, die sid²⁵ vollkommen
gibts also so sausages which are completely
vegetarisch
 vegetarian
[aber schmecken total nach fleisch]
 'If you miss the taste of meat per se (which I never did), there are
 also sausages that are completely vegetarian, but taste absolutely
 of meat.'

These constructions contain contrastive topics: (17a) is a posting of somebody who talks about alarm clocks, but refers to a special kind here, namely one that takes into account different sleep patterns, in contrast to regular ones that ring at a particular time irrespective of the owner's sleep behavior, and (17b) is a posting in a discussion of diets without meat, introducing a special kind of sausage, namely vegetarian ones, in contrast to the regular ones containing meat.

Contrastive topics have been analyzed as a combination of topic and focus, namely as aboutness topics that contain a focus indicating alternatives to this topic (Krifka 2007: Ch. 5.2). This suggests that what we have in the *gibts so*-constructions here, is a division of labor where *gibts* introduces the topic while *so* serves as a focus marker. Semantically, the NP itself identifies the larger class that is the topic, e.g., alarm clocks in (17a) and sausages in (17b), while the relative clause identifies a special, exceptional subclass that is in the focus, e.g., those alarm clocks that wake their owner up after sleep phases in (17a), and those sausages that are completely vegetarian.²⁶ The existence of alternatives — i.e., the more regular members of the superclass — is often further emphasized by the occurrence of an additive particle *auch* 'also' / 'too' as in (17b).

After this discussion of different focus domains for *so*, let us now briefly take a look at the communicative contexts in which it is used, before section 3 presents some further, experimental, evidence for the status of *so* as a focus marker.

2.4. Communicative contexts for *so* as a focus marker

The examples in (1) through (6) above provide evidence for *so* as a focus marker from a range of different contexts, such as a newspaper interview (1), an internet chat (2), a TV talk show (3), informal conversations between young people ((4) and (5)), and between a doctor and her patient (6). What these contexts have in common, is that they are part of informal, spoken language — or, in the case of chats, that they are at least CONCEPTUALLY oral, even if they belong to writing, rather than speech.²⁷

Within this domain of informal, spoken or conceptually oral language, *so* as a focus marker seems to be particularly prominent in youth language, so much so that its use has been subject to stylistic criticism, e.g., by Bastian Sick, a language commentator in German popular media. In a contribution to a weekly magazine ('Quatsch mit so Soße', *SpiegelOnline* October 2008), Sick complains about the frequent use of *so* by young people, putting forward a number of cases in point which, even though they are constructed to show the

purportedly exaggerated and unnecessary use of *so*, all provide beautiful examples of *so* as a focus marker, and, while doing so, bring together the different options of linearization we discussed above (pre-, post- and circum-*so*).

- (18) a. *Was machen wir so heute so?*
 what make we so today so
 ‘What will we do today?’ [so-bracket around *today*]
- b. *Der Justin Timberlake, der ist echt süß so*
 the Justin Timberlake he/DEM is really sweet so
 ‘Justin Timberlake is really cute.’ [post-*so* after *really sweet*]
- c. *Mathe und Physik, das ist so überhaupt nicht so mein Ding so. Ich steh so eher auf Kunst und so Grafikdesign so*
 Maths and Physics that is so absolutely not so my thing so I stand so rather on art and so graphic.design so
 ‘Maths and Physics are absolutely not my cup of tea. I rather like art and graphic design.’
 [possible distribution of *so*:
so-bracket around *absolutely not*; post-*so* after *my cup of tea*;
so-bracket around *rather art and graphic design*; pre-*so* before *graphic design*]

Sick — quite mistakenly — criticizes *so* in such contexts as being superfluous and providing “zero information”, obviously misled by the desemanticization that *so* undergoes in its use as a focus marker.

There is further evidence for the development of focus markers in informal varieties in other Germanic languages, particularly in youth language. According to Toril Opsahl (p.c.), *sånn* ‘true’ in Norwegian Youth Language might be interpreted as a focus marker similar to *so*, and *bara/ba* ‘only/exclusively’ in Swedish Youth Language has been characterized as a conversational/discourse marker that can “highlight certain parts of the discourse” by preceding the focused element (cf. Erman and Kotsinas 1993: 83).

The prominence of *so* in German youth language is particularly pronounced in *Kiezdeutsch* ‘(neighbor-)hood German’, a multiethnolect spoken among adolescents in neighborhoods with a high migrant population.²⁸ For the understanding of this, it is important to note that this multiethnolect is not a form of second language acquisition, but rather constitutes a distinct variety, a new, multiethnic dialect of German: it is characterized by grammatical developments that point to a productive linguistic system in its own right, and is spoken by adolescents of different ethnic backgrounds, including the majority ethnicity, German (cf. Wiese et al. 2009).

A comparison of *so* usages in a corpus of spontaneous speech samples from adolescents in a multiethnic (Kiezdeutsch) vs. a monoethnic neighborhood²⁹ yielded about the same number of *so*-occurrences altogether (approximately 18 tokens per 1000 words), but a lead for the Kiezdeutsch sample with respect to focus marker *so*'s within this set (4.0 vs. 3.3 tokens per 1000 words, or 22% vs. 18% of all *so* tokens, in the multiethnic and the monoethnic neighborhood, respectively).

This suggests that in a variety like Kiezdeutsch, which gains a special dynamic from its multilingual environment (with a lot of speakers who have more than one first language), nonstandard constructions might be established more easily, and so we might find general phenomena of informal German more salient and further advanced here. Given the tendency towards stronger reliance on pragmatics in language contact, this might be particularly pronounced for developments at the interface of grammar and pragmatics. I will come back to these aspects when discussing our results from the point of view of linguistic architecture in the final section.

A qualitative comparison of the evidence for focus-marking *so* in the two corpora showed that, while overall *so* was used as a focus marker in a similar way in both kinds of youth language, one type of construction appeared only in Kiezdeutsch, namely combinations of *so* with bare singular count NPs, as illustrated in (19), which would be ungrammatical without an article in standard German (cf. also Wiese 2011):

- (19) Description of a friend's behavior during soccer cups:

[*Er is ENGLänder und er feiert MIT uns,*]

Er hat so TÜRKEITrikot und TÜRKEIFAHne um sich.

he has *so* Turkey.strip and Turkey.flag around himself

'He is English, and/but he celebrates with us. He has the Turkish team's strip, and wraps a Turkish flag around himself.'

A lesser reliance on (explicit) functional elements such as determiners is characteristic for Kiezdeutsch, and accordingly, the occurrence of bare singular count NPs is also attested outside of *so* combinations. The use of *so* in such constructions in Kiezdeutsch might be further supported by an additional function here, namely that of indicating nominal arguments in the absence of an article (cf. Wiese 2006).

Summing up our discussion so far, the evidence we have for *so* from spontaneous, naturally occurring language production, then, strongly points to an additional, hitherto undescribed, usage as a focus marker in contemporary German, where it occurs in informal contexts, is particularly prominent in youth language, and here especially in Kiezdeutsch, a multiethnic variety. In the following section, I am going to present evidence for the utilization of *so* as a focus marker under controlled experimental conditions.

3. *So* as a focus marker: evidence from a “Jeopardy!” study

3.1. *Background: focus identification via questions*

In order to test, under controlled conditions, the hypothesis that *so* is used as a focus marker, we conducted a study that made use of the answer-and-question format in the TV quiz show “Jeopardy!”, which has a well-known German version.

Our setup is based on the observation that the location of focus in a sentence can be revealed by a question, and this is what is usually done in discussions of focus. Focus characteristically involves alternatives or contrasts³⁰ and hence can be understood as an answer to an implicit or explicit question that requests a choice from a set of such alternatives, such that the content provided by the question identifies the presupposed information, and the open position, typically marked by a *wh*-word, identifies the requested, new information. Accordingly, different questions can mark different possible focus constituents in a sentence, and this is typically used as a test for information focus.³¹

Take, for instance, a sentence like (20) below. If the context in which (20) is uttered suggests that it answers to a question like (21a), then the focus constituent in (20) would be *on the subway*; if, however, (20) answers to (21b), then the focus constituent would be *peanut*'.

- (20) *Karen eats peanuts on the subway.*
 (21) a. *Where does Karen eat peanuts?* [→ ‘on the subway’ as focus constituent]
 b. *What does Karen eat on the subway?* [→ ‘peanuts’ as focus constituent]

Looking at this the other way round, if someone interprets ‘on the subway’ as the focus constituent in (20), this means that s/he understands (20) as an answer to something like (21a), and if s/he interprets ‘peanuts’ as the focus constituent, s/he understands (20) as answering to (21b). Hence, if one asks people “Jeopardy!” style, to provide suitable questions for given sentences like the one in (20), one targets their identification of the focus constituent in those sentences.

And so this is exactly what we did. We presented people with sentences where, like in (20), apart from the subject, there were two major constituents that might be interpreted as the focus expression, namely a direct object and an adverbial, and asked them to formulate suitable questions for these sentences. We compared answers to “bare” sentences like the one in (20) to those that contained an occurrence of *so* preceding either the object or the adverbial. Our hypothesis was that the occurrence (vs. nonoccurrence) of *so* would make it

more likely for a constituent to be interpreted as the focus expression of a sentence, hence, compared to bare sentences, we predicted an increase of adverbial questions for sentences with *so* in front of the adverbial, and an increase of object questions for sentences where *so* precedes the object.

Since focus is also marked intonationally in German, we had to present the sentences in written form, rather than auditorily. This was something we expected to somewhat reduce the possible effects of *so* as a focus marker, since, as our discussion from the previous section showed, focus marking by *so* is, at present, mostly restricted to informal spoken language, thus it would be less natural in written form. This was, however, unavoidable since intonation would pin down focus identification.

Given that word order also contributes to the identification of focus in German, we tested our hypothesis in two experiments using a different order of objects and adverbials. In Experiment 1, sentences appeared in the basic word order, which sets the adverbial in front of the object, as illustrated in (22), a German counterpart to (20).

- (22) *Sie isst in der U-Bahn Erdnüsse.*
 she eats in the sub-way peanuts

In this linearization, the object is in a position, which, all other things being equal, favors its interpretation as the focus expression: in German, the preferred position for information focus in VP is the deepest embedded position (cf. Drubig and Schaffar 2001), which in this case is the object's. In sentences such as (22), where the constituent in this position appears as the sentence-final element, this allocation of focus gets further support by a general cross-linguistic strategy 'Focus Last', as proposed by Klein and Perdue (1997).³² This is the case in verb-second sentences — where the verb appears outside VP and not in its basic position following its objects — when no additional, extraposed elements follow the VP, cf. the sketch in (22').³³

- (22') [Sie_j [isst_i [t_j [_{VP}[in der U-Bahn] [Erdnüsse t_i]]]]]

Given this partiality towards an interpretation of the object as the focus expression in this case, we conducted two experiments, one with the basic word order described here, and a second one that complemented this by testing sentences with a reverse order of adverbial and object, as illustrated in (23).

- (23) *Sie isst Erdnüsse in der U-Bahn.*
 she eats peanuts in the sub-way

This linearization is possible in German through a movement of the object to a position further to the left within the middle field, such that the adverbial ends up in final position, cf. the sketch of the VP structure in (23').

(23') [Sie_j [isst_i [t_j [VP [Erdnüsse]_k [[in der U-Bahn] [t_k t_i]]]]]]

This linearization provides us with a direct counterpart of the word order in the unmarked case (22). In this marked case, we expect the word order not to favor an interpretation of the direct object as the focus expression anymore. Instead, two outcomes are possible: (a) The word order might favor a focus interpretation for the adverbial instead, which is now in the sentence-final (= focus-preferred) position, or (b) Moving the object to the left might lead to a balanced outcome where both constituents are equally likely to be interpreted as focus expressions, given that the object's movement puts some emphasis on it and thus might counteract the advantage that the adverbial gains from its position.³⁴ In the first case, (a), we would have a mirror image of Experiment 1 that could serve as a basis for comparisons. In the second case, (b), the "bare" sentences in Experiment 2 would provide us with a balanced baseline for the calculation of *so* effects.

Finally, we conducted a third experiment with an English version of our questionnaire, where we tested US-American speakers of English and used sentences involving *like* as a focus marker. This allowed us to check our setup and the results we obtained in Experiments 1 and 2 against those for an element, namely *like*, whose usage as a focus marker in informal language is well established and will be less controversial than in the case of German *so*.

3.2. *Experiment 1*

3.2.1. *Participants.* Twenty-five native speakers of German, students from Potsdam University participated in this experiment. All had normal or corrected-to-normal vision.

3.2.2. *Materials.* Stimuli consisted of visually presented sentences of the kind illustrated in (22) above, that is, sentences containing, in this order, *subject* — *finite verb* — *adverbial* — *direct object*. In order to make sentences maximally comparable, we observed the following additional restrictions (also illustrated in (22) above): the subject was always a 3rd person pronoun, the verb was in present tense, the adverbial was a prepositional phrase, and the object was an indefinite plural or mass NP.³⁵ Semantically, the subject was always human, the adverbial was modal, identifying place ('on the subway'), time ('in the morning', 'during his lessons'), or companions ('with her parents'), and the object NP referred to food items ('cookies', 'wine'). We constructed 10 sentences from each of 5 categories:

- 1) BARE sentences containing no additional elements;
- 2) soADV sentences with *so* preceding the adverbial;
- 3) soOBJ sentences with *so* preceding the object;
- 4) FpADV sentences with a focus particle such as *nur* ‘only’, *auch* ‘also’, or *sogar* ‘even’ preceding the adverbial;
- 5) FpOBJ sentences with a focus particle preceding the object.

The last pair of categories, 4) and 5), allowed us to include particles into the analysis for which an interaction with focus has already been well established (see Section 1 above), and thus check whether the set-up is suitable to distinguish focus marking for different constituents.

Sentences were matched across categories such that in each case, a sentence with the same verb would appear in five different forms according to the five categories, and would contain the same focus particle in the case of categories (4) and (5) (that is, we had 10 basic sentences that were modified for five categories each).

The sentences were distributed in the final script in a pseudo-randomized manner such that no two sentences of the same category followed each other and that similar sentences (as matched across categories) were maximally apart.

3.2.3. *Procedure.* Participants were asked to provide a suitable question for each sentence. For practical reasons, this was conducted as a pen-and-paper task, where participants filled in their responses on a print-out of the script, which allowed us to test all participants at the same time. The questionnaire was headed by the following instructions (translated from German):

Do you know the TV show *Jeopardy*? In this game, you have to find an appropriate question for each given answer. This idea is the basis for our questionnaire.

In what follows, you will find 50 answers, taken from informal conversations. Your task is to tell what question might have been answered in each case.

Please keep close to the answers, and use, as far as possible, the same words!

The request in italics was included to avoid responses (= questions) that would not target a sentence constituent directly and thus unambiguously. This is something we encountered in a pilot run, where for a sentence like *He broils pork chops in the kitchen*, a participant offered the question *Is this kosher?*, which suggests that *pork chops* was interpreted as the focus expression, but does not pin this down unambiguously, as opposed to, say, *What does he broil in the kitchen?* (= object focus) vs. *Where does he broil pork chops?* (= adverbial focus).

The pointer to informal conversations was to decrease the risk that *so* would not be employed as a focus marker in the written stimuli sentences because of its affiliation with informal spoken language.

3.2.4. *Results and discussion.* Participants' responses (= the questions they provided) were classified according to the open constituent identified by the *wh*-phrase in the question: if the *wh*-phrase was the direct object, e.g., *What does she eat on the subway?*, the response was classified as indicating "object focus"; if the *wh*-phrase was an adverbial, e.g., *Where does she eat peanuts?*, the response was classified as indicating "adverbial focus"; all other responses (subject questions, questions pointing to verum focus, etc.) were classified as "others".

A comparison of means per subject yielded the following picture for the different focus classes in the different sentence categories.

In BARE sentences, we found a preference for object focus (mean for 10 sentences: 6.84) versus adverbial focus (1.8), as expected, given the position of the object (cf. the discussion above).

This preference was increased in FP_{OBJ} sentences where a conventional focus particle (*auch* 'also', *sogar* 'even', or *nur* 'only') preceded the direct object (mean: 7.68)³⁶, such that object focus was significantly more frequent than in BARE sentences ($\chi^2 = 4.15$; $p = 0.042$). In FP_{ADV} sentences with a focus particle preceding the adverbial, the preference got reversed (mean for adverbial focus: 5.48)³⁷, such that adverbial focus was significantly more frequent than in BARE sentences ($\chi^2 = 72.6$; $p = 0.000$). This pattern, then, supports our methodology; it suggests that the set-up is indeed sensitive to focus marking through focus-sensitive particles.

Against this basis, we got the following pattern for sentences with *so*. In *so*_{OBJ} sentences, where *so* precedes the direct object, we got similar results as for sentences with focus particles, namely an increased preference for object focus: object focus was even more frequent than in bare sentences (mean for object focus: 7.56)³⁸, although here, this difference in frequency did not quite reach significance ($\chi^2 = 3.83$; $p = 0.050$). In *so*_{ADV} sentences, where *so* preceded the adverbial, the preference for object focus observed in BARE sentences did not get reversed, as it did in FP_{ADV} sentences, but it got weaker (mean for object focus: 5.72)³⁹. Crucially, adverbial focus was significantly more frequent in *so*_{ADV} sentences than in BARE sentences ($\chi^2 = 4.77$; $p = 0.029$).

Figure 1 summarizes the means per sentence category.

These results suggest that *so* has a significant influence on the identification of focus, even though in the word order employed here, this influence is not strong enough to override the general preference for object focus. On this basis, Experiment 2, then, tested the effect of *so* on focus identification in

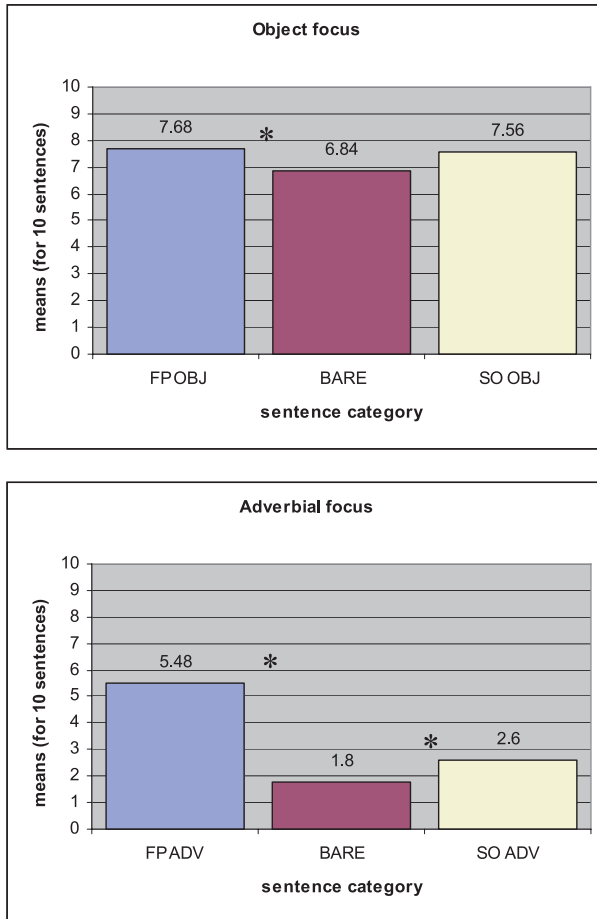


Figure 1. Choice of object focus and adverbial focus in Experiment 1

sentences with a word order that would not favor the object over the adverbial as the focus expression.

3.3. Experiment 2

3.3.1. *Participants.* Thirty-three native speakers of German, students from Potsdam University participated in this experiment. All had normal or corrected-to-normal vision.

3.3.2. *Materials.* Stimuli were the same as in Experiment 1, with one difference: the order of direct objects and adverbials was reversed, such that sentences were of the form *subject — finite verb — direct object — adverbial*. Apart from this modification, the script was identical to that used in Experiment 1.

3.3.3. *Procedure.* The procedure was identical to Experiment 1.

3.3.4. *Results and discussion.* Participants' responses were classified in the same way as in Experiment 1. A comparison of means per subject yielded the following results.

In BARE sentences, we did not observe a preference for either object focus or adverbial focus, but got a balanced distribution of responses, as envisioned in outcome (b) above (Section 3.1). — In fact, we got the rather unlikely result of participants choosing both constituents with *exactly* the same frequency, namely 146 times each, or 4.42 times per subject for the 10 sentences on average.

This, then, gives us a perfect baseline for our assessment of *so* effects. In addition, it also contributes to a more general point about the interaction of word order and information structure in German. It suggests that in the absence of intonational information, the kind of movement we employed here supports a focus interpretation for the object that balances the focus-favored sentence-final position of the adverbial, in a way that focus is assigned to both constituents with the same likelihood.⁴⁰

In this neutral, focus-impartial context, that is, against a (totally) balanced baseline, *so* had significant effects on the allocation of focus, and these effects did not differ from those of conventional focus particles. In FpOBJ sentences, where a conventional focus particle preceded the direct object, we observed a preference for object focus (mean: 5.52),⁴¹ such that the choice of object was significantly more frequent than in BARE sentences ($\chi^2 = 7.85$; $p = 0.005$). This held for soOBJ sentences as well, where *so* preceded the object: object focus was preferred (mean: 5.27),⁴² and it was significantly more frequent than in BARE sentences ($\chi^2 = 4.76$; $p = 0.029$).

The same pattern held for adverbial focus. In FpADV sentences, where a focus particle preceded the adverbial, adverbial focus was preferred (mean: 6.09),⁴³ and was chosen significantly more often than in BARE sentences ($\chi^2 = 18.4$; $p = 0.000$). In soADV sentences, where *so* preceded the adverbial, adverbial focus was preferred as well (mean: 6.03),⁴⁴ and it was also significantly more frequent than in BARE sentences ($\chi^2 = 17.1$; $p = 0.000$).

Figure 2 summarizes the means per sentence category.

These results suggest that, all things being equal, that is, without a biasing word order, *so* significantly influences the identification of focus, in a similar way as conventional focus particles.

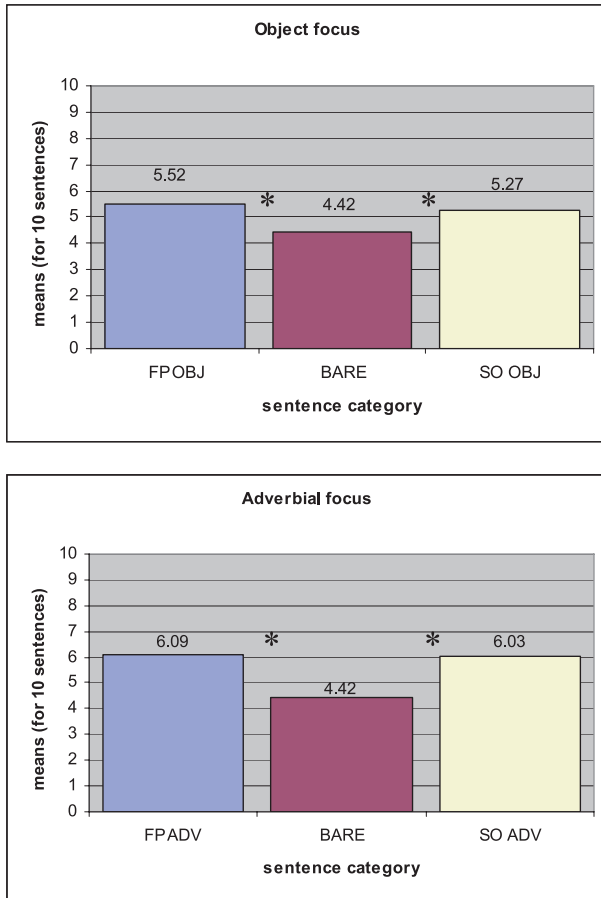


Figure 2. Choice of object focus and adverbial focus in Experiment 2

3.4. Experiment 3

3.4.1. *Participants.* Twenty-three native speakers of US-American English, recruited via the Psychological Research on the Net list, participated in this experiment.

3.4.2. *Materials.* Stimuli were English versions of the sentences in Experiments 1 and 2, presented visually in an online questionnaire. Sentences

contained, in this order, *subject* — *finite verb* — *direct object* — *adverbial*. Following the pattern from Experiments 1 and 2, we constructed 10 sentences from each of 5 categories:

- 1) BARE SENTENCES containing no additional elements;
- 2) LIKEADV sentences with *like* preceding the adverbial;
- 3) LIKEOBJ sentences with *like* preceding the object;
- 4) FPADV sentences with a focus particle such as *only*, *just*, *also*, or *even* preceding the adverbial;
- 5) FP OBJ sentences with a focus particle preceding the object.

3.4.3. *Procedure.* The procedure was the same as in the German experiments, with the only difference that the experiment was conducted online, such that the script was presented in electronic form and participants filled in their responses on their computer.

3.4.4. *Results and discussion.* Participants' responses were classified in the same way as in the German experiments. A comparison of means per subject yielded the following picture for the different focus classes in the different sentence categories.

In BARE sentences, we found a balanced distribution of responses, just as in the German Experiment 2 — in fact, we observed, again, a completely even outcome, with the means for our 10 sentences the same for object focus and adverbial focus, namely 4.83, which provided us with a balanced baseline for our comparisons.

Against this baseline, *like*, as well as conventional focus particles, had significant effects on the allocation of focus, as in Experiment 2. We observed a preference for object focus in both FP OBJ sentences (mean: 7.96)⁴⁵ and in LIKE OBJ sentences (mean: 7.52),⁴⁶ and in both cases the choice of object was significantly more frequent than in BARE sentences (BARE vs. FP OBJ sentences: $\chi^2 = 119$; $p = 0.000$; BARE vs. LIKE OBJ sentences: $\chi^2 = 35$; $p = 0.000$). Similarly, adverbial focus was preferred in FP ADV sentences (mean: 8.30),⁴⁷ as well as LIKE ADV sentences (mean: 7.52)⁴⁸ and again, in both cases was chosen significantly more often than in BARE sentences (BARE vs. FP ADV sentences: $\chi^2 = 62.2$; $p = 0.000$; BARE vs. LIKE ADV sentences: $\chi^2 = 36.5$; $p = 0.000$).

Figure 3 summarizes the means per sentence category.

These results indicate the same pattern for English *like* as for German *so*: they suggest that *like*, just as *so*, has a significant influence on the choice of focus in a sentence, similar to that of conventional focus particles like *only*, *just*, *also*, and *even*.

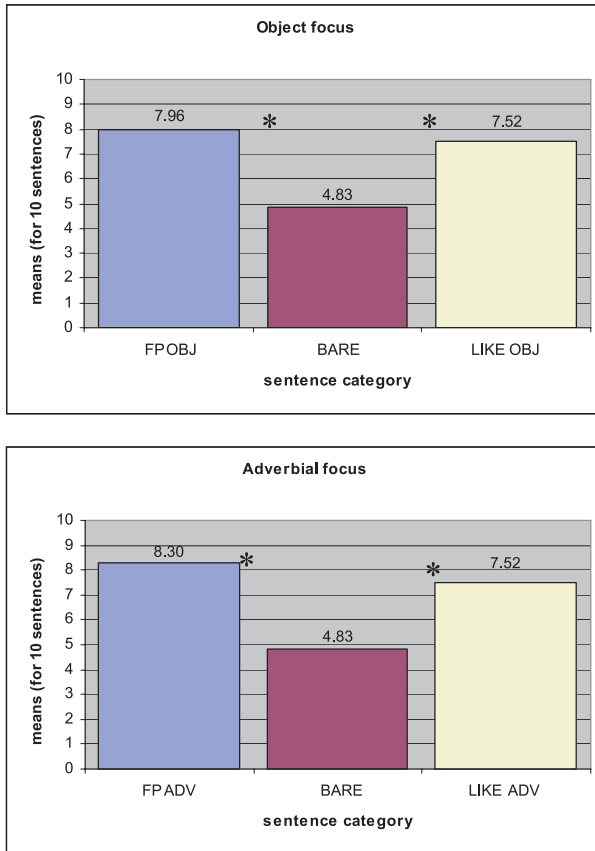


Figure 3. Choice of object focus and adverbial focus in Experiment 3

3.5. General discussion

In sum, we found that *so* has a significant effect on the identification of focus in a sentence: the occurrence of *so* in front of a constituent makes it more likely that this constituent is chosen as the focus expression of a sentence, than in the absence of *so*. The influence of *so* that we observed in our study was not as strong as that of conventional particles like *auch* ‘also’, *sogar* ‘even’, or *nur* ‘only’ in a context of biasing word order, that is, under conditions where the word order favors the interpretation of one constituent as focus expression over the other (Experiment 1). In the absence of such a bias, though, this difference disappeared (Experiment 2). What is more, the pattern we observed for *so* in

such a neutral context was the same as the one we found in an English version of the experiment for *like*, whose usage as a focus marker in informal language should be uncontroversial (Experiment 3).

I interpret this as further evidence for the status of *so* as a focus marker: it shows that the occurrence of *so* guided participants' selection of focus, that is, when interpreting sentences, speakers make use of *so* in order to identify the expressions that carried information focus. This, then, supports our claim for a novel, genuine focus marker *so* in German with evidence from controlled experimental conditions, complementing the analysis of naturally occurring language data above.

A reason why *so* could not override word-order advantages in the same way as conventional focus particles in Experiment 1, might be that we had to use written stimuli, whereas *so* as a focus marker, as noted above, occurs mostly in spoken language and has a nonstandard status as opposed to the well-established use of focus-sensitive particles like *auch* 'also', *sogar* 'even', or *nur* 'only'. The following section takes a closer look at where this usage of *so* might come from, that is, what might be a possible developmental path leading to it, and contrasts the outcome of such a development with these referential particles.

4. Where could this usage come from? A possible pragmaticalization path

In Section 2.1 I described the core meaning of *so* in conventional usages as that of an indexical modal element answering to *How?*, and argued that around this core meaning, we find a cluster of related meanings that can be derived from this, which captures the typical usages of *so* described in the literature. In comparison to these usages, *so* as a focus marker has lost its semantic content in favor of an information-structural function. What could be a possible path leading to focus-marking *so* and relating it to its other usages?

If we take as a feasible point of departure the main usages of *so* as a content word, then the semantic bleaching we find in focus-marking *so* suggests a process of grammaticalization or, more precisely, of pragmaticalization: what we observe here, is a usage as a function word that is presumably based on those as a content word, but is specialized for a pragmatic/discourse function, rather than a grammatical/ morphosyntactic function. What does this difference imply?

4.1. Some preliminary considerations: Grammaticalization and pragmaticalization

Language change leading from content words to elements with a primarily or exclusively pragmatic function is known, for instance, from the emergence

of discourse particles such as English *I think*, *you know*, or *and things / and stuff*, and German *weil* or *obwohl*.⁴⁹ The process leading to such elements has by some researchers been described as ‘pragmatization’,⁵⁰ while others include it under the concept of ‘grammaticalization’, often by employing a broader concept of grammar that includes pragmatic and discourse aspects.⁵¹

I take a more narrow view of grammar here, in line with a general architecture that encompasses syntax, morphology, phonology, and semantics as the basic domains involved in the formation of words and sentences,⁵² but does not include processes involved in the organization of discourse, conceptualization, and general information management. Such a distinction allows us to separate the grammatical system proper from more general domains involved in communication. The first, but not the latter, is specific to language and covers, in addition to possibly universal features of grammar, restrictions and rules, or constructions that are specific to a particular language or language family.

If we understand, in a general approach, information structure as a way of packaging information that takes into account communicative needs of speaker and hearer,⁵³ we can then tease apart (a) extralinguistic cognitive aspects that are relevant for the constitution of information-structural preferences, and (b) the realization of such preferences in linguistic expressions, which is subject to language-specific grammatical restrictions. Based on an architecture suggested by Jackendoff (2002), one can account for this by an information structure tier that is part of the conceptual system, and distinguish this from the syntactic counterparts of information-structural realizations.⁵⁴

In the extra-linguistic domain relevant for this information structure tier, we find general and language-independent cognitive aspects, such as the organization of conceptual representations and extra-linguistic conditions of the communicative situation, which bear on communicative needs and thus on preferences of information packaging. These preferences reflect communicative strategies of a speaker who must then conform to the grammatical and lexical conditions of a language in order to implement them.

Hence, *so* as a focus marker has as its domain a *pragmatic* category, namely *focus*, which is part of information structure, and as such interacts with grammar, but is not part of the grammatical system itself. In contrast to that, *grammaticalized* elements such as, for instance, *zu* ‘to’ as a semantically bleached infinitive marker, indicate categories (such as ‘infinitive’) that establish and organize morphosyntactic representations and as such are part of grammar proper.

If we compare the pragmatic usage of *so* with those as a content word, we find some, but not all, of the characteristics described for grammaticalization. Taking Lehmann’s (2000: Ch. 4) distinction of criteria, we can see that the usage of *so* as a focus marker fulfills those in two areas, namely the decrease in

semantic and phonological weight: as described above, focus marking *so* is semantically bleached compared to its referential usage and cannot be stressed anymore. In the syntactic domain, though, we miss a central characteristic of grammaticalization, namely that of decreased syntactic variability. If you recall our discussion in Section 2.2 above, *so* does not get more syntactically fixed as focus marker, but moves in the opposite direction, and is characterized by a higher positional variability, that is, an increase rather than a reduction of syntactic scope, compared to content word usages.

As mentioned above, this deviation is due to the particular pragmatic function of *so*: as a marker of focus, its position will vary with the location of focus in a sentence, that is, it will take as its co-constituent phrases of different syntactic categories. Hence, the development from a content word to a desemantized, bleached element in this case brings with it syntactic promiscuity, rather than fixation, due to a specific pragmatic function that requires this.

This points to an important difference between grammaticalization and pragmaticalization. While both are processes of language change where a content word becomes a function word, the domain of this function is not the same in both cases. In grammaticalization proper, the outcome is an element whose function targets the grammatical system. Accordingly, as Haspelmath (1998: 318) summarizes it, “grammaticalization is the gradual drift in all parts of the grammar toward tighter structures, toward less freedom in the use of linguistic expressions at all levels.”

However, when the outcome of the process is a word whose function does not contribute to grammatical structure as such, but rather to extragrammatical domains such as information structure or discourse organization, then there is no necessary connection to such tighter structures. As a result, unlike in grammaticalization, the reduction of syntactic scope is not part and parcel of the process in this case, but rather an optional add-on, a possible artifact of a particular outcome rather than an integral part of the process: discourse particles and pragmatic markers such as *so* might or might not get syntactically fixed depending on the requirements of their particular pragmatic function. In the case of focus marking *so*, this is a function that requires increased syntactic variability, and the same is true, for instance, for English *though* in its function as a discourse marker (Barth-Weingarten and Couper-Kuhlen 2002), while in other cases, we might find that the pragmatic function supports a loss of syntactic freedom similar to grammaticalization proper, but at the same time involves a decrease in bondedness, as for instance in the case of German *obwohl* and *wobei* as discourse markers (Günthner and Mutz 2004).

In view of this, I am going to use the concept *pragmaticalization* to identify a domain of language change processes that result in words with a *pragmatic* rather than a grammatical function. This domain is closely related to that of grammaticalization, and shares core features with it, in particular the decrease

in semantic weight that comes with the development of a function word, but differs from it in its functional outcome.⁵⁵

What would, then, be a possible path of pragmaticalization in the case of *so*? In order to answer this question, let us now have a closer look at the semantic and pragmatic relations between the different usages of *so*. In doing so, the development I suggest will show a *possible* route to focus-marking *so*, that is, a path that shows that we can relate focus-marking *so* to other *so* usages, based on semantic commonalities and processes of semantic bleaching, but which does not make specific diachronic claims.

4.2. *The basis: an indexical modal on sorts*

We can account for the core indexical modal meaning of *so* by describing it generally as an indexical on sorts, namely sorts of objects (“this kind”), sorts of activities (“this manner”), and sorts of properties or attributes (“this quality” / “this degree”). In its central usage as a comparison particle, the indexical content thus set up is specified by an explicit or implicit object of comparison (“of this / the same kind/manner/quality as . . .”).

As discussed in 2.1, this core semantics is accessed by a cluster of related usages, in particular a resultative meaning of *so*, where the comparison targets events (“of such a kind/manner/degree that . . .”); an intensifying meaning, with an implicit object of comparison whose quality is emphasized (“of *such* a kind/manner/quality”); and a relational usage, where the comparison establishes a relation between two attributes or a general phenomenon and an instance of it (“of such a quality/kind as . . .”).

4.3. *The status of quotative so*

The two other usages of *so* we discussed above, namely that of a quotative marker and of an approximation marker, are semantically further removed from the core meaning of *so* and can in fact be described as a first step of semantic reduction. As a quotative, *so* sets up an indexical relation between a speaker and her utterance and/or nonverbal behavior such as sound effects, gestures etc. Apart from establishing this relation, *so* does not contribute any additional content in this usage anymore. However, compared to the focus marker *so*, we still do have a semantic heritage from its core usage: as sketched above, we can account for the quotative usage of *so* via its indexical modal meaning ‘in this manner’. This aspect of its core meaning provides a natural basis to set up a quotative relation between a speaker S and the quoted material, targeted as the activity of S: ‘S acted (spoke/behaved/ . . .) in this manner: . . .’.

Given that the quoted material introduced by a quotative marker is normally in focus, there might be a diachronic relationship between quotative markers and focus marking.⁵⁶ In fact, some accounts on quotatives have suggested a developmental path connecting the two, going from focus marking to quotatives: Güldemann (2008), in his survey of quotative indexes in African languages, proposes focus operators as one source for the development of quotatives, and Eriksson (1995) analyzes Swedish *ba* ‘only / exclusively’ as a quotative whose basis is a foregrounding device.

Note though, that at least in the case of German *so*, we find a closer semantic relation of the quotative to the conventional meaning of *so*, compared to a further advanced semantic bleaching of the focus marker. In view of this, I do not endorse such a path for the case of German *so*, but regard both usages as part of a network of meanings, with the focus marker further along a pragmaticalization path than the quotative, but not necessarily derived from it.

A similar case has been made for English *like* by Romaine and Lange (1991), who maintain that *like* in its function as a quotative has retained some of its semantic meaning (based on its meaning in referential usages, such as comparison), in contrast to the semantically bleached usage as a focus marker. Based on Traugott’s (1982) distinction of propositional, textual, and expressive/interpersonal stages for semantic shifts underlying grammaticalization, they suggest a grammaticalization path that leads from a propositional usage of *like* as a preposition to its textual usage as a conjunction and from here, on the one hand, to its equally textual usage as a quotative, and, on the other hand, to its more semantically bleached interpersonal usage as a focus marker. As I am going to show in what follows, a similar (though not identical) path might hold for German *so*.

4.4. *Hedging as a basis of focus-marking so*

A suitable candidate for a derivational basis for German *so* as a focus marker, rather than quotative *so*, is the other remaining usage of *so* we illustrated in 2.1, namely that as an approximation or vagueness marker. This usage is also characterized by a reduction of semantic content compared to the core meaning: the indexicality that *so* has as a comparison marker is lost here, and only the element “sort” remains, such that *so* denotes a generalized kind/manner/quality, meaning “of a/some kind/manner/quality”. This thus gets us from *so* as an answer to *How?* to *so* meaning something similar to ‘*somehow*’. On the semantic level, this contributes vagueness, while pragmatically, *so* can be used for hedging, understood as a discourse-level process that reduces the speaker’s commitment to the truth of the utterance and thus allows him to distance himself from its content to some extent.⁵⁷ Since everything is of a certain sort (kind/

manner/quality), the generalization involved in this usage of *so* yields an element that fits into semantic representations without affecting them in a substantial way, and can thus be flexibly used for this kind of pragmatic hedging.

It is this usage for hedging that can provide a pertinent basis for a reinterpretation of *so* as a focus marker: given that the focus expression in a sentence is “the information center of the sentence” (Molnár and Winkler 2006: 3), hedging makes most sense for this domain, making focus expressions a prime target for hedging.⁵⁸ Hence, *so* as a hedging device will frequently co-occur with focus expressions, which makes it likely for it to be reinterpreted as a focus marker. And just like a focus marker, *so* is always unstressed when indicating vagueness and used as a hedging device in German.⁵⁹

The possibility of such a path from hedging to focus marking is further supported by the fact that both functions do not exclude each other, but can and do coexist, making a transition feasible. While hedging modulates the speaker’s *commitment* towards the information he conveys, focus marking affects the *structuring* of information and thus will not clash with such a modulation. Accordingly, we find many “gray” areas where *so* could be used for hedging as well as for focus marking — or, presumably, for both.⁶⁰

The examples in (24) illustrate such grey cases, where *so*, on the one hand, supports hedging, and on the other hand marks the respective focus expressions that provide the central and new information. (24a) to (24c) are taken from the DWDS corpus of spoken and written German,⁶¹ (24d) is from the Kiezdeutsch Korpus (see Note 29; MuP1MK), and (24e) is from Burkhardt (1987: 311), who quotes this as an example for hedging with *so*.⁶²

- (24) a. Conversation in TV show ‘Das literarische Quartett’, August 17th, 2001, where invited people discuss current literature; comment by literature journalist Iris Radisch:
Das mag originell sein und das mag irgendwie
 that might original be and that might somehow
so einen Kieztouch haben,
 so a ‘hood.touch have
 [wenn man Kreuzberg liebt, dann liest man das nicht völlig ohne Amüsement.]
 ‘That might be original and might have a neighborhood feel to it somehow. If one loves Kreuzberg, then reading it (i.e., the book under discussion) is not altogether unamusing.’
- b. Conversation in TV show *Das literarische Quartett*, April 23rd, 1999; comment by literary critic Marcel Reich-Ranicki:
 [Ich will Ihnen nur eines sagen: Wir empfehlen hier Bücher, aber wir sind nicht dazu da, irrezuführen. Wir machen also drauf aufmerksam:]

*Dies ist bedeutende, herrliche Literatur, und ich
this is significant wonderful literature and I
meine, nicht so ganz leicht lesbar.
think not so entirely easy readable*

‘Let me say just one thing about this: we recommend books here, but it is not our job to mislead people. Hence, we point out: This is significant, wonderful literature and, I think, not entirely easy to read.’

- c. From an opinion article in the weekly paper *Die Zeit* 33/2008 by Tanja Stelzer, criticizing over-ambitious urban middle-class parents in their attitude towards education for their children:

*Ich habe halt so ein romantisches Kindheitsideal,
I have after.all so a romantic childhood.ideal
[vielleicht ist das etwas altmodisch.]*

‘I have a somehow romantic ideal of childhood, maybe that’s a bit old-fashioned.’

- d. Discussion of an ongoing football on TV match among adolescents:

*Ich hab so das gefÜHL, die erste HALBzeit kein Tor,
I have so the feeling the first half.time no goal
JA?*

yes

‘I have the feeling there won’t be a goal in the first half, right?’

- e. Presenter Dieter Kürten commenting on a bike race, in a TV sports program, the *ZDF-Sportstudio*, April 19th, 1986 [stress marking after Burkhardt 1987: 311]:

*Die NiedersACHsenrundfahrt ist für die deutschen
the lower.Saxony.round.tour is for the German
RADrennfahrer so die erste PROBE für die neue
bike.race.drivers so the first test for the new
SAISON.*

season

‘For the German bike racers, the Lower Saxony Tour is the first test for the new season.’

In these utterances, *so* is part of hedging: by way of *so*, the speaker tones down the proposition, reducing his/her commitment. In (24a) and (24c), this is further indicated by the occurrence of vagueness elements like *irgendwie* ‘somehow’, and *halt* ‘just’/‘after all’. In the other examples, *so* stands by itself, but nevertheless reduces the force of the statement, without the additional support of vagueness markers, in particularly in (24b) and (24d) ((24e) might be a less clear case of hedging, cf. Note 62). In both cases, we find epistemic elements

like “I think” and “I have the feeling” that set up a suitable context for the hedging achieved by *so*. The same sentence without *so* would make the statement more forceful, and strengthen the speaker’s commitment to the proposition, for instance in (24b) to the proposition that this kind of literature is not easy to read. (25) illustrates this with a minimal pair extracted from (24b), together with English counterparts involving *like*, where we get a comparable effect.

- (25) a. *Diese Literatur ist nicht so ganz leicht lesbar;*
 this literature is not *so* entirely easy readable
 ‘This literature is not, like, entirely easy to read.’
 b. *Diese Literatur ist nicht ganz leicht lesbar;*
 this literature is not entirely easy readable
 ‘This literature is not entirely easy to read.’

While *so* is part of hedging in such examples, at the same time it combines with a focused co-constituent, the constituent that provides the new, highlighted information of the sentence, namely *einen Kieztouch* ‘a neighborhood feel’ in (24a), where a crucial, unanticipated impression of the book under discussion is described; *(nicht) ganz leicht lesbar* ‘(not) entirely easy to read’ in (24b), which might reduce the attraction of the literature in question and thus argue against the speaker’s recommendation; *ein romantisches Kindheitsideal* ‘a romantic childhood ideal’ in (24c), which introduces an alternative view of childhood to the one set up by overambitious parents; *ein Gefühl* ‘a feeling’ in (24c), which emphasizes the epistemic modulation for the statement; and, in (24d), *die erste Probe* ‘the first test’ as the central new information about the tour’s significance for the bike racers.

4.5. A possible developmental path for *so* as a focus marker

Such dual-purpose usages of *so* can constitute bridging contexts, providing a basis for a transition from semantic vagueness and pragmatic hedging to a usage of *so* as a focus marker, which then serves to identify focus expressions independently of hedging strategies. This transition would involve further semantic bleaching: as a focus marker, *so* makes no semantic contribution at all anymore — not even one as weak as vagueness — but is specialized for a purely pragmatic/discourse function.

This account of *so* is in accordance with that suggested by Meehan (1991) for English *like*, who regards usages where *like* indicates approximation as a basis for the development of *like* as focus marker. Similarly to focus-marking *so*, English *like* is also particularly prominent in youth language, and one reason for this might thus be that this is a context that is particularly open to hedging.⁶³

The possible path I sketched for *so* here, from comparison via vagueness/hedging to focus marking, implies a transition from propositional and textual to interpersonal, discourse usages of *so*, and is in this respect comparable to the one proposed by Romaine & Lange (1991) for English *like*, and as became clear above, I, too, assume an independent path from propositional to quotative and hence textual usages of *so* that does not lead to focus marking. In contrast to this, Eriksson's (1995) account of Swedish *ba*, as mentioned earlier, involves a path for *ba* from foregrounding to quotative, and hence from interpersonal to textual usage, which, as Eriksson (1995: 43) points out, constitutes a deviation from the tendency observed by Traugott (1982) that propositional, textual, and expressive/interpersonal stages succeed each other in grammaticalization paths (cf. also Hopper & Traugott 1993). According to the account for *so* developed here, German would then differ from Swedish in following the sequence captured by Traugott's account.

Figure 4 summarizes the possible path I propose for *so* on the semantic and discourse levels (the dotted horizontal line separates the two domains, distinguishing semantics as part of grammar from the extragrammatical domain of discourse). The path leads from the basic referential meaning of *so* as a modal indexical on sorts — with relational, resultative, and intensifying usages clustering around the core comparative meaning — via approximation (semantic vagueness and pragmatic hedging) to its usage as a focus marker, a path accompanied by an additional, separate sideline from basic (modal indexical) to quotative usages.

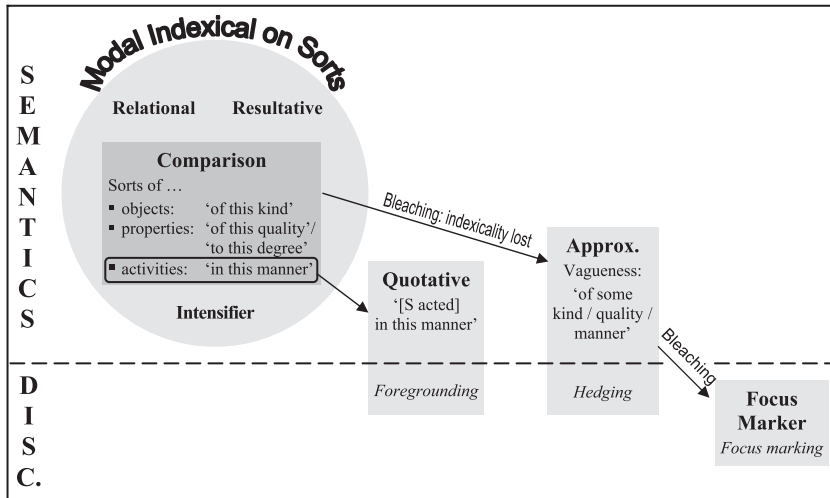


Figure 4. A possible pragmaticalization path for *so* as a focus marker in German

Note that this path does not involve a grammatical marker as the basis for focus-marking *so*: according to this scenario, *so* undergoes semantic bleaching that leads directly from a content word to a pragmatic marker without an intermediate stage characterized by a primarily grammatical function. This, then, would suggest a development similar to the one discussed by Erman and Kotsinas (1993) for the pragmaticalization of *bara/ba*'. There is one reservation to this, though: while this holds for *so* in most variants of informal spoken German, in Kiezdeutsch there is at least some evidence for a usage of *so* as a more — if not fully — grammaticalized element, namely as a marker of nominal arguments (see the discussion in Section 2.4 above). This, then, might provide an additional support for the spread of focus-marking *so* in this multiethnic dialect.

4.6. *The outcome of pragmaticalization: the meaning of focus-marking so*

The distinction between semantic and discourse/pragmatic aspects in the different usages of *so* makes evident that, if this path is correct, in focus-marking *so* we see the result of pragmaticalization: the development from (a) a referential item with the core meaning 'comparison' to (b) a semantically reduced, but not empty approximative element that, semantically, contributes vagueness, while, pragmatically, it can serve as a hedging device, and (c) via further semantic bleaching on to a focus marker that does not contribute any semantic content anymore, but acts solely in the pragmatic/discourse domain of information structure.

As mentioned above, such a focus marker is different from conventional German focus particles such as *auch* 'also'/'too' or *nur* 'only', which interact with focus, but at the same time also contribute their own lexical meaning to the semantic representation of a sentence. This contribution has been accounted for by operators that take as their argument a sentence denotation that identifies a set of alternatives for the focused element.

Within the *Alternative Semantics* framework of Rooth (1985, 1992), this is achieved through an alternative meaning that involves a function that is applied to the ordinary meaning of the focused expression and yields a set of context-appropriate alternatives for it. In a *Structured Meanings* approach (Jacobs 1983; von Stechow 1991), the alternatives come in via lambda-abstraction for the denotation of the focused expression, which divides the sentence denotation into focus (\rightarrow focused expression) and background (\rightarrow lambda-abstracted proposition). (26) and (27) illustrate this for the sentence 'Sue eats peanuts on the subway', with (a) the focus on *peanuts* (*What does Sue eat on the subway?*) and (b) the focus on *subway* (*Where does Sue eat peanuts?*). (26) gives a (simplified) representation for this within Alternative Semantics, (27) with Structured Meanings.⁶⁴

- (26) a. $\{\text{EAT}(\text{sue}, x): \text{LOC}(\text{on_the_subway}) \mid x \in \text{ALT}(\text{peanuts})\}$
 b. $\{\text{EAT}(\text{sue}, \text{peanuts}): \text{LOC}(x) \mid x \in \text{ALT}(\text{on_the_subway})\}$
 (27) a. $\langle \text{peanuts}, \lambda x \text{EAT}(\text{sue}, x): \text{LOC}(\text{on_the_subway}) \rangle$
 b. $\langle \text{on_the_subway}, \lambda x \text{EAT}(\text{sue}, \text{peanuts}): \text{LOC}(x) \rangle$

Against this background, the meaning of conventional focus particles has been accounted for by semantic operators that are applied to such representations and add to them in specific ways, for instance, in the case of *only*, requiring that there are no other elements from the alternative set, apart from the focused element, that fulfill the proposition; in the case of *also*, stating that there exist such additional elements. (28a) and (28b) show how this works for our two representations in (27) for the case of *only*, with a semantic operator ONLY that captures the meaning of the focus particle (adapted from Krifka 2007).

- (28) a. $\text{ONLY} (\langle \text{peanuts}, \lambda x \text{EAT}(\text{sue}, x): \text{LOC}(\text{on_the_subway}) \rangle) =$
 $\forall y ([\text{EAT}(\text{sue}, y): \text{LOC}(\text{on_the_subway})] \rightarrow \text{peanuts} = y)$
 b. $\text{ONLY} (\langle \text{on_the_subway}, \lambda x \text{EAT}(\text{sue}, \text{peanuts}): \text{LOC}(x) \rangle) =$
 $\forall y ([\text{EAT}(\text{sue}, \text{peanuts}): \text{LOC}(y)] \rightarrow \text{on_the_subway} = y)$

What ONLY does here is, it takes a semantic representation as its input which is already structured in a way that marks one element as the focus constituent, namely “peanuts” in (28a) and “on_the_subway” in (28b), and it then adds further semantic content to this representation: the output of ONLY gives us a representation that does not only mark “peanuts” (or “on_the_subway” in (28b)) as focus element, but also states that all other entities that fulfill the lambda-marked proposition for “peanuts” are identical with “peanuts”, hence that everything that Sue eats on the subway is peanuts — or indeed that Sue eats only peanuts on the subway. This means that conventional focus markers such as *only* (or *also*, *even*, etc.) do not yield a representation that sets a particular element in focus, but rather presuppose this: they need this as their input, and they then add further semantic (restrictive, additive, or scalar) content to this input. They do not MARK focus, they INTERACT with focus in a way that their presence with a particular constituent indicates that this constituent must be the focus expression in order to receive the semantic modification they supply.

For *so*, we have no such semantic contribution: as a specialized focus marker, it does not add to the semantic content of the sentence, but instead helps to identify the focus element in the first place: it forces a particular focus structure onto the semantic representation, but does not add any content to it, and accordingly does not affect its truth value. In other words, it leaves where focus particles like *only* begin.

Figure 5 illustrates this graphically for sentences: the box on the left stands for the ordinary meaning of a sentence expressing an Agent (“Ag”; e.g., Sue),

a Patient (“Pat”; e.g., peanuts), and a Location (“Loc”, e.g., on the subway), the ones in the middle show two (of more) different options for organizing this representation to mark a focus element (rendered graphically higher than the others), which is supported by word order and intonation and can be indicated by *so* in different positions (*soObj*: *so* with the object as a co-constituent, *soAdv*: *so* with the adverbial as a co-constituent), and, finally, the boxes on the right stand for semantic representations enriched by the contributions of a focus particle like *only* (where REST stands for the restrictive meaning that *only* adds to the semantic representation).

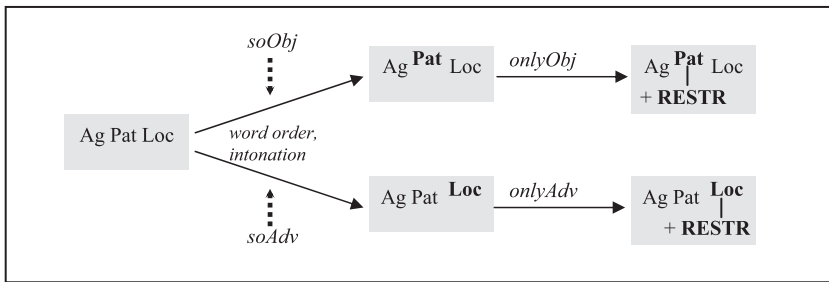


Figure 5. *The contribution of the focus marker so vs. conventional focus particles*

Against this background, we can now account for the contribution of *so* by an element, which, rather than operating on a sentence denotation that identifies a set of alternatives for the focused element, constitutes this alternative set to begin with: *so* generates a set of alternatives from the denotation of its co-constituent, thus supporting a particular representation of alternative or structured meanings, and this way identifies its co-constituent as the focus expression of the sentence.

I capture this contribution of *so* with a function SO_f that maps the denotation of a sentence, a proposition $P(b,a)$, onto one that characterizes one of the arguments as an element of a set of alternatives within Alternative Semantics (29a), or as a lambda-abstracted element in a Structured Meaning representation (29b).⁶⁵

- (29) a. $SO_a(P(b,a)) = \{P(b,x) \mid x \in ALT(a)\}$
 b. $SO_a(P(b,a)) = \langle a, \lambda x P(b,x) \rangle$

These representations, then, capture the fact that *so*, in its function as a focus marker, does not contribute to the meaning of a sentence on the semantic level proper, it does not add any new content to its denotation, but rather acts on the way the existing content is organized: it supports a certain way to organize the

information that the sentence conveys by invoking alternatives for a specific element and thus marking it as a focus expression.

5. Conclusions and outlook

In this article I have discussed a hitherto undescribed usage of the particle *so* in contemporary German, where *so* marks the focus expression in a sentence, in a similar way as *like* in colloquial English. I have argued that in this usage, *so* serves as a genuine, dedicated focus marker, that is, as a functional element that does not contribute any meaning to the sentence, but is specialized for its information-structural function.

In this usage, *so* shows some characteristic features. On the semantic and phonological levels, it is reduced: semantically, it is bleached of its content and does not affect the truth value of a sentence in which it appears; on the phonological level, it is systematically unaccented, while its co-constituent (= the focus expression of the sentence) carries the main sentence stress. On the syntactic level, the distribution of *so* expands, again as an effect of its pragmatic status: as a focus marker, *so* will tend to be adjacent to the focus expression of a sentence, and accordingly, in this usage *so* combines with all major lexical categories (or their immediate functional projections), including some combinations that would be ungrammatical in conventional, nonfocus marking usages of *so*. In its linearization with its co-constituent, *so* displayed an interesting range of options, preceding it, following, and even forming a bracket around it. Whether this reflects a functional distinction remains to be investigated. On the pragmatic/discourse level, I showed that *so* can mark both information focus and contrastive focus, and also marks the focus part in contrastive topics introduced by *gib(t)s* 'gives (it)', a form that might be on its way to a monomorphemic item in German.

As evidence for this usage of *so*, I have discussed examples from naturally occurring linguistic data, and complemented this by controlled experimental data from a study investigating speakers' identifications of focus expressions in sentences containing *so*. This study supports the analysis of *so* as a focus marker: it shows that *so* has a significant influence on speakers' understanding of what the focus expression in a sentence is, similar to traditional focus particles such as *auch* 'also'/'too', *nur* 'only', or *sogar* 'even'.

In contrast to such focus particles, *so*, according to the analysis I put forward in this paper, is a genuine focus marker that does not provide any additional content, but is reduced to its pragmatic/discourse function. Accordingly, I have sketched a possible pragmatization path for *so* that leads from the semantic contributions of conventional, referential usages via a semantically reduced, but not empty usage in hedging to a purely pragmatic role as a focus marker,

where *so* does not add any new content to a sentence denotation anymore, but supports a particular way to organize the information it conveys.

Taken together, the results of our discussion contribute to several domains:

- 1) With respect to contemporary German, this use of *so* indicates that focus marking in this language is not restricted to means of intonation and word order, but that the system is open to lexical markers as well, that is, to function words whose designated and only function it is to mark focus. Incidentally, evidence for this openness might also come from the first language acquisition of German, where this option might be explored as well. The following data from child language suggests that the particle *nur* ‘only’ is used here as a pure, semantically void focus marker, in contrast to its usage in adult language, where it carries a restrictive meaning.⁶⁶
- (30) a. Mother: *Willst du eine Apfelsine?*
 Y. (2;6): *Ich will nur ZWEI!*
 I want only two
 ‘[Mother: Do you want an orange?] — Y. (2;6): I want (*only*) two!’
- b. Mother: *Y, schäl das Ei!*
 Y. (2;7): *Ich SCHÄLE nur!*
 I peel only
 ‘[Mother: Y, peel the egg!] — Y. (2;7): I am (*only*) peeling it!’
- c. Y. has two sheets of colored paper, and is to give one to her sister.
 Y. (2;8): *Ich will nich Emen. Ich will den andern*
 I want not one I want the other
AUCH! Ich will nur ZWEI!
 also I want only two
 ‘Y. (2;8): I don’t want one, I want the other one as well. I want (*only*) two!’

This data needs to be further complemented by more systematic evidence, but even as it is, it shows that, at least in an individual case, there can be a stage in the acquisition of German where a lexical element is used as a pure focus marker, with the effect that only those of its aspects that target the organization of information structure are selected, while its lexical content is neglected.

- 2) From a typological point of view, the occurrence of focus-marking *so* in German, taken together with the evidence for English *like*, shows that the use of a genuine, purely pragmatic focus marker is not restricted to better known examples from typologically different languages, such as those from African languages and creoles, but is an option for Germanic languages as well.

- 3) From the perspective of grammaticalization and pragmatization, the use of *so* as a focus marker and its (synchronic) semantic/pragmatic relations to other usages of *so* together suggest a particular development of a function word from a content word that differs from typical grammaticalization in two ways: (i) the development does not seem to require an intermediate stage involving a dedicated grammatical, syntactic/morphological function, but might lead directly to a pragmatic marker; and (ii) it leads to an increase, rather than a decrease, in syntactic variability for the resulting function word. This motivates a view of grammaticalization and pragmatization as two closely related but distinct sister domains of diachronic processes that lead to function words, with commonalities, but also with differences resulting from whether this functional outcome has a grammatical or pragmatic status.
- 4) Finally, if we have a look at the communicative contexts we found for focus-marking *so*, this use of *so* can shed light on the interface between grammar and information structure from the point of view of linguistic architecture. As we saw above, focus-marking *so* is characteristically used in informal, spoken (or conceptually oral) language, and in particular in youth language. Within this domain, we found that the use of *so* as a focus marker is particularly pronounced in a new variety of German spoken among adolescents in multiethnic neighborhoods, namely *Kiezdeutsch* '(neighbor-)hood German'.

Due to its context of language contact and the multilingual competences of many of its speakers, this variety supports a special linguistic dynamic; it constitutes a more liberal linguistic system with softer grammatical constraints. This liberalization supports a number of grammatical reduction phenomena,⁶⁷ but also provides a basis for innovations.⁶⁸

Against this background, we can interpret the higher occurrence of focus-marking *so* in *Kiezdeutsch* as a sign of the stronger permeability of its grammatical system to the implementation of means that are motivated by extra-grammatical, information-structural, aspects; in this case, the use of lexical devices for focus marking which is not an option in standard German.⁶⁹ This points to a systematic interaction between weakened grammatical constraints and a more liberal realization of information-structural preferences, suggesting a scenario where extra-linguistic aspects of information structure put pressure on the linguistic system in a way that might — or might not, depending on the conditions offered by a particular system — lead to new lexical and grammatical developments, to the emergence of linguistic structures supporting particular extra-linguistic preferences in the packaging of information. Taken together, the emergence of *so* as a focus marker, then, is an option in the lin-

guistic system of German that is not only interesting as and by itself, but might also tell us something about the interplay between grammar and information structure.

Received 4 March 2010

University of Potsdam

Revised version received 30 September 2010

Appendix. Abbreviations

ACC	– accusative
CL	– clitic
DEM	– demonstrative
INF	– infinitive
VPRT	– verb particle
F	– focus

Notes

- * Work on this article was supported by a German Research Foundation (DFG) grant to the Collaborative Research Centre (SFB) “Information structure” 632 of Potsdam University and Humboldt University Berlin (project B6 “Kiezdeutsch”, PI: Heike Wiese). For the discussion of different parts of the account presented here, I would like to thank audiences at *Jugendsprache* 5, 2008; *ICLaVE* 5, 2009; *NWAV* 38, 2009; *DGfS* 2010, and at the Centre for Language, Discourse & Communication, King’s College London 2008. This paper has further benefited from valuable comments by Ray Jackendoff, Larry Horn, and my colleagues at the SFB “Information Structure”, in particular Manfred Krifka, Tom Güldemann, Malte Zimmermann, and my collaborators in project B6, Ulrike Freywald, Sören Schalowski, and Katharina Mayr. For their help with collecting the data in the ‘Jeopardy!’ study, I would like to thank Jennifer Jason, Jens Roeser, Kathleen Schumann, Eva Wittenberg, and, again, Katharina Mayr, who also thought of the name “Jeopardy!”, which proved so helpful for instructing participants! For constructive comments on an earlier version of this paper, I would like to thank two anonymous reviewers. Correspondence address: Institut für Germanistik, Universität Potsdam, Am Neuen Palais 10, 14469 Potsdam, Germany. E-mail: heike.wiese@uni-potsdam.de.
1. Small caps mark main stress; for readability, interlinear glosses are only provided for the immediate passages containing *so*, but not for the sentences in brackets, which provide further context.
 2. From DWDS (*Digitales Wörterbuch der Deutschen Sprache des 20. Jahrhunderts* [Digital Lexicon of German in the 20th century], Berlin-Brandenburg Academy of Science): *Die Zeit* corpus, encompassing all editions from 1996 to 2007, with 106 million word tokens in more than 200,000 articles.
 3. MuH1WD from the “Kiezdeutsch Korpus”, a corpus of spontaneous speech among adolescents in a multiethnic neighborhood of Berlin (transcriptions of self-recordings, approx. 127,000 words), cf. Wiese et al. (2009/10).
 4. Naturally occurring speech data, own observation.

5. IdS Mannheim, Archive "Spoken German", Pfeffer Corpus, 1961, Interaction PF041, Münster. Thanks to the anonymous reviewer for providing this example.
6. After the English translation by William Taylor.
7. Cf. also Freywald (2010), who gives several examples for nonstandard constructions in German that have been mischaracterized as newly emerging simply because they have only recently received linguistic attention.
8. Accordingly, König (1991: 29) points out that such focus particles "in contrast to 'pure' focus markers [. . .] have a lexical meaning." Rooth (1985, 1992), who calls these particles "focusing adverbs", analyzes their semantic contribution and its interaction with focus within an Alternative Semantics framework. Cf. Horn (1996) for a detailed discussion of the different meaning aspects involved in the case of *only*.
9. Cf. Boadi (1974); Heine and Reh (1983); Bearth (1999); Aboh et al. (2007).
10. Cf. Byrne and Winford (1993); Veenstra and den Besten (1994).
11. Cf. Zifonun et al. (1997); Heidolph et al. (1981); Helbig and Buscha (2001); Duden 4 (2009). Cf. also Hennig (2006) for a discussion of categorizations for *so*.
12. For discussions of common usages of *so* and their meaning cf., for instance, Ehlich (1987); Thurmair (2001); Weinrich (2003); Auer (2006).
13. *Solch-* initially goes back to a compound of *so* + Germanic **-līka-z*, meaning 'of a certain shape/design'.
14. In a similar vein, Ehlich (1987: Ch. 4.1) characterizes this usage of *so* as an expansion of its deictic domain. Note that there is a slight difference between the constructions in (9e) and (9f): in (9e), *so* is in the main clause as a particle in the VP, supporting an interpretation where the verb's referent is high on a scale, whereas in (9f), *so* is in the subordinate clause and syntactically forms a complex conjunction with *dass* (as in English 'such that'). In both cases, though, *so* sets up a comparison of events (denoted by the main clause and the subordinate clause) that leads to a resultative meaning.
15. Cf. Golato (2000); Auer (2006: Ch. 2.1) on quotative usages of *so* in German. On English *like* as a quotative marker cf., for instance, Romaine and Lange (1991); Jones and Schieffelin (2009).
16. Given this large range, Güldemann (2008) calls *so* in such contexts a "mimesis marker".
17. DWDS corpus, see Note 2. Cf. also Pittner (1993); Hennig (2006) on this usage of *so*.
18. Cf. Burkhardt (1987) and Sandig (1987) for *so* in hedging.
19. Note that this is not always true for occurrences of comparative *so* in constructions such as (9a) and (9b), where *so* can be left implicit. This option, which seems to be more easily available in some dialects than in others, is based on the combination of *so* with *wie* 'as' (or, in interrogative usage, 'how') in these examples: since *so* answers indexically to *How?*, it can remain implicit when *wie* is present.
20. A similar point has been made by Underhill (1988), who states that *like* as a focus marker, in contrast to its other usages, can never be stressed. Note that this is also something that distinguishes *so* from focus particles such as *auch/also, nur/only* etc., which can carry nuclear tone if they follow the focus constituent (cf. König 1991 and the discussion of (15) below).
21. E.g. in a sentence like 'You say to-MAY-to, and I say to-MAH-to.' (Jackendoff 1997: 411).
22. This is supported by corpus data. The DWDS (*Digitales Wörterbuch der Deutschen Sprache*, see Note 2) core corpus of over 100,000 word tokens from a range of mostly written texts, provides about 13,000 hits for *es gibt*, but about 19,000 hits, that is about 1.5 times as many, for *gibt es* and the cliticized form *gibt's/gibts*.
23. From a discussion forum on IT News, <http://forum.golem.de/read.php?27931,1463768,1464287> (last viewing Dec 8, 2009). A Google search for *es gibt's* and *es gibts* on German websites (Dec 8, 2009) yielded about 45,000 hits.

24. From internet sources: (17a) is a posting on gutefrage.net, an advice community (<http://www.gutefrage.net/frage/es-gibts-so-wecker-die-einen-nach-schlafphasen-wecken-also-immer-wenn-man-gerade-nicht-im-tiefschlaf-ist>, last viewing Dec 7, 2009), (17b) is a contribution on med1, a forum on health and healthy diets (<http://www.med1.de/Forum/Ernaehrung/388944/>, last viewing Dec. 7, 2009).
25. "Sid" is the original spelling. I assume that this is a misspelling of *sind* 'are'.
26. Relative clauses are particularly well suited for this function, given their semantics: cf., for instance, the seminal work by Keenan and Comrie (1977), who define relative clauses as constructions that specify a set of objects by restricting a larger set, the domain of relativization (identified by the NP), to a smaller subset, for which a certain sentence (identified by the relative clause) is true.
27. Cf. Koch and Oesterreicher (1994), who distinguish the graphic vs. phonic *medium* of language on the one hand, and its *conceptuality* on the other hand, which covers a continuum from conceptually written to spoken. Cf. Crystal (2001) on aspects of informal oral communication in internet chats (but see also Androutsopoulos (2003) on the sociolinguistic variability within chats and other internet communication forms).
28. Cf. Wiese (2006, 2009) on a grammatical/pragmatic description of Kiezdeutsch as a multiethnolect. For sociolinguistic and discourse-theoretical analyses of language use among adolescents in areas with a high migrant population cf., for instance, Kallmeyer and Keim (2003); Dirim and Auer (2004: Ch. 4); Selting and Kern (2009).
29. The corpora comprise transcriptions of self-recordings of spontaneous speech from conversations among adolescents in a multiethnic neighborhood of Berlin and from a monoethnic neighborhood of Berlin with comparable socio-economic indicators, cf. Wiese et al. (2009/2010) (see also Note 3).
30. See, for instance, Jackendoff (1972); Jacobs (1983); Rooth (1985); König (1991); Krifka (2007).
31. See, for instance, Heusinger (2004: 167), who states that "the traditional test for the focus-background structure of a sentence consists in a constituent question and its felicitous answer", and König (1993: 978), who points out that "a clear delimitation of the focus is only possible on the basis of the context and some appropriate tests", and follows this by identifying focus expressions through questions. Cf. also Erteschik-Shir's (2007: Ch. 2.2) discussion of this "question-answer test" (2007: 28) for foci. Rooth (1992: Ch. 2.4 and 3) discusses the relationship between a question and the position of focus in an appropriate answer from the point of view of Alternative Semantics. Underhill (1988: 240) discusses instances of English *like* as a focus marker that identifies the significant information in answers to questions.
32. Klein and Perdue (1997) propose "focus last" for contact varieties like the *Basic Variety* in untutored second language acquisition. Jackendoff (2002) characterizes this strategy, together with a counterpart "topic first", as protolinguistic "fossil principles" (2002: 249) that hold crosslinguistically and operate when there are no conflicting language-specific syntactic restrictions. Note, though, that the resulting order "Topic > Focus" might not be the only linearization option: as Givón (1988) and Gundel (1988) argue, there might also be an opposite principle "attend to the most urgent task first" (Givón 1988: 252) or "first things first" (Gundel 1988: 229), which would favor well-introduced topics to be allocated to positions further towards the end of the sentence.
33. This sketch follows syntactic approaches to (Dutch and) German that account for the V2 pattern by a unified analysis that takes as the basic word order V-final, or SOV, and assumes two topological changes: (1) *verb fronting*, where the finite verb moves to first position, yielding V1 (= verb first), and (2) *topicalization*, where a major constituent, in our example the subject, moves to the left of this fronted verb, yielding V2 (cf. Koster 1975; Thiersch 1978; den Besten 1981).

34. Cf. also Lenerz (1977), who considers the order “direct object > adverbial” as the unmarked serialization in German (despite a syntactic bracketing [adverbial [direct object V]]), because it supports a rhematic status for either constituent, whereas the serialization ‘adverbial > direct object’ favors the object (in accordance with our discussion above).
35. This allowed us to use bare NPs without determiners, thus avoiding difficulties with cliticized vs. full forms of indefinite articles following *so*, and potential problems for combinations of *so* with definite articles (the data so far suggests that such a combination is possible in principle, but it is not clear how well established this is).
36. The mean for adverbial focus was 0.48.
37. Object focus: 1.52.
38. Adverbial focus: 1.24.
39. Adverbial focus: 2.6.
40. Note that in contrast to this, scrambling of direct and indirect object in the German middle field involves defocusing of the moved constituent in favor of focus marking for the then sentence-final constituent. The outcome we find here for adverbials and objects hence is in accordance with accounts such as Haider and Rosengren (2003), who restrict the notion of “scrambling” to arguments and thus exclude changes of relative argument-adjunct order.
41. Adverbial focus: 2.48.
42. Adverbial focus: 3.88.
43. Object focus: 1.76.
44. Object focus: 3.24.
45. Adverbial focus: 0.04.
46. Adverbial focus: 2.22.
47. Object focus: 1.48.
48. Object focus: 2.22.
49. Cf., for instance, Erman and Kotsinas (1993); Brinton (1996); Aijmer (1997, 2002); Cheshire (2007); Gohl and Günthner (1999); Günthner (2000).
50. E.g., Erman and Kotsinas (1993); Aijmer (1997); Günthner and Mutz (2004).
51. E.g., Traugott (2003); Diewald (2011).
52. I take these general assumptions to hold across theoretical models; differences occur in particular with respect to the relative weight allotted to the syntactic system (cf. Jackendoff (2002: Part II) for an overview).
53. Chafe (1976). Cf. Krifka (2007) for an explication in terms of Common Ground management.
54. See also Steube (2000), who proposes a model where information structure involves conceptual features such as ‘contextually new vs. bound’ that are *per se* extralinguistic and enter grammar in the course of derivation, with language-specific grammatical consequences, e.g., for focused vs. nonfocused expressions.
55. For a similar proposal see Wischer (2000), who subsumes cases of pragmaticalization under the concept of “grammaticalization”, but distinguishes two kinds of grammaticalization, where one covers discourse-pragmatic processes. Barth-Weingarten and Couper-Kuhlen (2002), in an alternative approach, suggest reinterpreting “grammaticalization” as a prototype concept. While this can cover cases of pragmaticalization, it would assign them a more peripheral status than those of grammaticalization proper, unlike the view suggested here.
56. Cf. also Golato (2000), who discusses the usage of *so* as a German quotative and points out its relation to a foregrounding function, where *so* can mark new and noteworthy information.
57. This is in accordance with a notion of hedging used, e.g., by Hyland (1998: Ch. 1). Note that pragmatic hedging differs from epistemic modality, which also concerns a speaker’s relation to the truth of a proposition, in that with hedging, this relation is that of the speaker’s *com-*

- mitment* to the truth, which is reduced, whereas the relation involved in epistemic modality is the speaker's *evaluation* of the truth (e.g., an evaluation of the proposition as being certainly true, probably true, etc.).
58. Cf. also Wiese (2011). Note that what we find here, is a gradual development involving the *reinterpretation* of an element, in contrast to an abrupt change based on syntactic *reanalysis* (see Haspelmath 1998 on a distinction of reanalysis from grammaticalization).
 59. Cf. also Weinrich (2003).
 60. A similar coexistence of hedging and focus functions is also known from typologically different languages, cf. for instance Sankoff (1993) who discusses the Tok Pisin expression *em* that can be used as a hesitation form as well as a focus marker. Underhill (1988: 240–241) discusses the use of English *like* for hedging, in addition to its use as a focus marker.
 61. Corpus *Gesprochene Sprache* 'Spoken Language', encompassing transcripts from the 20th century with approx. 2.5 million tokens, and Die ZEIT corpus (see Note 2), respectively.
 62. As one reviewer pointed out, it is actually not quite clear whether hedging is involved in (24d) as well. I included it following Burkhardt's (1987) interpretation, who provided the example, analyzing it as a hedging usage of *so*. Given that the usage of *so* as a focus marker, unlike that as a hedging device, has not been in the center of attention before, this might also be a case of focus *so* that was not recognized as such.
 63. Cf. Henne (1986), who also points out the frequent occurrence of complex hedging expressions containing *so*, such as *und so* 'and so on' and *oder so* 'or so', in German youth language. Androutsopoulos (1998: Ch. 6.4) describes frequently occurring combinations of *so* with semantically weak, generalized nouns in youth language, such as *so Zeug* 'such stuff', or *so Sachen* 'such things'.
 64. ALT is an operator that yields a set of alternatives; LOC identifies the location of an event; for the present purpose, I simplified the representation of plural NPs and ignored the instantiation of events, using the connective “:” ('such that', cf. Bierwisch 1988) to combine a proposition describing an event, and its location.
 65. Within the framework of Foreground-Background Semantics, as proposed by Heusinger (2004), *so* would accordingly support a partition of discourse representations into background and foreground representations.
 66. Natural data, own observations, conversations between mother and child (Y.); age of child given in brackets as “year;months”.
 67. Cf., for instance Dirim and Auer (2004: 207), who list characteristic omissions and reductions in this language variety.
 68. Cf. Wiese (2009).
 69. Cf. also Wiese et al. (2009); Wiese (2011) on the interaction between grammar and information structure in Kiezdeutsch.

References

- Aboh, Enoch Oladé, Katharina Hartmann & Malte Zimmermann (eds.), 2007. *Focus strategies in African languages: The interaction of focus and grammar in Niger-Congo and Afro-Asiatic*. Berlin & New York: Mouton de Gruyter.
- Aijmer, Karin. 1997. “I think” — an English modal particle. In Toril Svan & Olaf Jansen Westvik (eds.), *Modality in Germanic languages: Historical and comparative perspectives*, 1–47. Berlin & New York: Mouton de Gruyter.
- Aijmer, Karin. 2002. *English discourse particles: Evidence from a corpus*. Amsterdam: John Benjamins.

- Androutsopoulos, Jannis K. 1998. *Deutsche Jugendsprache. Untersuchungen zu ihren Strukturen und Funktionen*. Frankfurt am Main: Peter Lang.
- Androutsopoulos, Jannis K. 2003. Online-Gemeinschaften und Sprachvariation: Soziolinguistische Perspektiven auf Sprache im Internet. *Zeitschrift für Germanistische Linguistik* 31. 173–197.
- Auer, Peter. 2006. Construction Grammar meets Conversation: Einige Überlegungen am Beispiel von „so“-Konstruktionen. In Susanne Günthner & Wolfgang Imo (eds.), *Konstruktion in der Interaktion*, 291–314. Berlin & New York: De Gruyter.
- Barth-Weingarten, Dagmar & Elizabeth Couper-Kuhlen. 2002. On the development of final *though*: A case of grammaticalization? In Ilse Wischer & Gabriele Diewald (eds.), *New reflections on grammaticalization*, 345–361. Amsterdam & Philadelphia: John Benjamins.
- Beard, Thomas. 1999. The contribution of African linguistics towards a general theory of focus: Update and critical review. *Journal of African Languages and Linguistics* 20(2). 121–156.
- den Besten, Hans. 1981. Government, syntaktische Struktur und Kasus. In Manfred Kohrt & Jürgen Lenerz (eds.), *Sprache: Formen und Strukturen: Akten des 15. Linguistischen Kolloquiums Münster 1980*, Vol. 1. (Linguistische Arbeiten 98), 97–108. Tübingen: Max Niemeyer Verlag.
- Bierwisch, Manfred. 1988. On the grammar of local prepositions. In Manfred Bierwisch, Wolfgang Motsch & Ilse Zimmermann (eds.), *Syntax, Semantik und Lexikon* (Studia Grammatica 29), 1–65. Berlin: Akademie Verlag.
- Boadi, Laurence A. 1974. Focus-marking in Akan. *Linguistics* 140. 5–57.
- Brinton, Laurel J. 1996. *Pragmatic markers in English: Grammaticalization and discourse functions*. Berlin & New York: Mouton de Gruyter.
- Burkhardt, Armin. 1987. SOSO? Kritik und weiterführende Überlegungen zu Konrad Ehlich's Aufsatz über die Funktionen des deutschen *so*. In Inger Rosengren (ed.), *Sprache und Pragmatik: Lunder Symposium 1986*, 299–313. Stockholm: Almqvist & Wiksell.
- Byrne, Francis & Donald Winford (eds.), 1993. *Focus and grammatical relations in Creole languages*. Amsterdam & Philadelphia: John Benjamins.
- Chafe, Wallace L. 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In Charles N. Li (ed.), *Subject and topic*, 27–55. New York: Academic Press.
- Cheshire, Jenny. 2007. Discourse variation, grammaticalisation and stuff like that. *Journal of Sociolinguistics* 11(2). 155–193.
- Crystal, David. 2001. *Language and the internet*. Cambridge: Cambridge University Press.
- Diewald, Gabriele. 2011. Pragmaticalization (defined) as grammaticalization of discourse functions. *Linguistics* 49(2). 365–390.
- Dirim, İnci & Peter Auer. 2004. *Türkisch sprechen nicht nur die Türken: Über die Unschärfebeziehung zwischen Sprache und Ethnie in Deutschland* (Linguistik — Impulse und Tendenzen 4). Berlin & New York: De Gruyter.
- Drubig, Hans Bernhard & Wolfram Schaffar. 2001. Focus constructions. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher & Wolfgang Raible (eds.), *Language typology and language universals: An international handbook* (HSK 20/2), 1079–1104. Berlin & New York: De Gruyter.
- Ehlich, Konrad. 1987. *So* — Überlegungen zum Verhältnis sprachlicher Formen und sprachlichen Handelns, allgemein und an einem widerspenstigen Beispiel. In Inger Rosengren (ed.), *Sprache und Pragmatik. Lunder Symposium 1986*, 279–298. Stockholm: Almqvist & Wiksell.
- Eisenberg, Peter, Jörg Peters, Peter Gallmann, Cathrine Fabricius-Hansen, Damaris Nübling, Irmhild Barz, Thomas Fritz & Reinhard Fiehler. 2009. *Duden: Band 4: Die Grammatik*. 8th edn. Mannheim: Duden Verlag.
- Eriksson, Mats. 1995. A case of grammaticalization in modern Swedish: The use of *ba* in adolescent speech. *Language Sciences* 17(1). 19–48.
- Erman, Britt & Ulla-Britt Kotsinas. 1993. Pragmaticalization: the case of *ba'* and *you know*. *Studier i modern språkvetenskap* 10. 76–93.

- Erteschik-Shir, Nomi. 2007. *Information structure. The syntax-discourse interface* (Oxford Surveys in Syntax and Morphology 3). Oxford: Oxford University Press.
- Freywald, Ulrike. 2010. „Obwohl vielleicht war es ganz anders.“ Vorüberlegungen zum Alter der Verbzweitstellung nach subordinierenden Konjunktionen. In Arne Ziegler (ed.), *Historische Textgrammatik und Historische Syntax des Deutschen*, 55–84. Berlin & New York: De Gruyter.
- Givón, Talmy. 1988. The pragmatics of word order: Predictability, importance and attention. In Michael T. Hammond, Edith A. Moravcsik & Jessica R. Wirth (eds.), *Studies in syntactic typology* (Typological Studies in Language 17), 243–284. Amsterdam & Philadelphia: John Benjamins.
- Gohl, Christine & Susanne Günthner. 1999. Grammatikalisierung von *weil* als Diskursmarker in der gesprochenen Sprache. *Zeitschrift für Sprachwissenschaft* 18(1). 39–75.
- Golato, Andrea. 2000. An innovative German quotative for reporting on embodied actions: *Und ich so | und er so* ‘and I’m like | and he’s like’. *Journal of Pragmatics* 32. 29–54.
- Grimm, Jakob & Wilhelm Grimm. (1965– [1854–1961]) *Das Deutsche Wörterbuch*. (DWB) [New Edition in progress edited by the Berlin-Brandenburg and Göttingen Academy of Sciences].
- Güldemann, Tom. 2008. *Quotative indexes in African languages: A synchronic and diachronic survey* (Empirical Approaches to Language Typology 34). Berlin & New York: Mouton de Gruyter.
- Gundel, Jeanette K. 1988. Universals of topic-comment structure. In Michael T. Hammond, Edith A. Moravcsik & Jessica R. Wirth (eds.), *Studies in syntactic typology* (Typological Studies in Language 17), 209–239. Amsterdam & Philadelphia: John Benjamins.
- Günthner, Susanne & Katrin Mutz. 2004. Grammaticalization vs. pragmaticalization? The development of pragmatic markers in German and Italian. In Walter Bisang, Nikolaus P. Himmelmann & Björn Wiemer (eds.), *What makes grammaticalization? A Look from its fringes and its components* (Trends in Linguistics Studies and Monographs 158), 77–107. Berlin & New York: Mouton de Gruyter.
- Günthner, Susanne. 2000. From concessive connector to discourse maker: The use of *obwohl* in everyday German interaction. In Elizabeth Couper-Kuhlen & Bernd Kortmann (eds.), *Cause, condition, concession, contrast: Cognitive and discourse perspectives*, 439–468. Berlin & New York: Mouton de Gruyter.
- Haider, Hubert & Inger Rosengren. 2003. Scrambling: nontriggered chain formation in OV languages. *Journal of Germanic Linguistics* 15(3). 203–267.
- Haspelmath, Martin. 1998. Does grammaticalization need reanalysis? *Studies in Language* 22(2). 315–351.
- Heidolph, Karl Erich, Walter Flämig & Wolfgang Motsch. 1981. *Grundzüge einer deutschen Grammatik*. Berlin: Akademie Verlag.
- Heine, Bernd & Mechthild Reh. 1983. Diachronic observations on completive focus marking in some African languages. *Sprache und Geschichte in Afrika* 5. 7–44.
- Helbig, Gerhard & Joachim Buscha. 2001. *Deutsche Grammatik*. Berlin & München: Langenscheidt.
- Henne, Helmut. 1986. *Jugend und ihre Sprache*. Berlin & New York: De Gruyter.
- Hennig, Mathilde. 2006. „So, und so, und so weiter“: Vom Sinn und Unsinn der Wortartklassifikation. *Zeitschrift für Germanistische Linguistik* 34. 409–431.
- von Heusinger, Klaus. 2004. Focus particles, sentence meaning, and discourse structure. In Alice ter Meulen & Werner Abraham (eds.), *The composition of meaning*, 167–194. Amsterdam & Philadelphia: John Benjamins.
- Hole, Daniel & Gerson Klumpp. 2000. Definite type and indefinite token: The article ‘son’ in colloquial German. *Linguistische Berichte* 182. 231–244.

- Hopper, Paul and Elizabeth Closs Traugott. 1993. *Grammaticalization*. Cambridge University Press.
- Horn, Laurence R. 1996. *Exclusive company: Only and the dynamics of vertical inference*. *Journal of Semantics* 13. 1–40.
- Hyland, Ken. 1998. *Hedging in scientific research articles*. Amsterdam & Philadelphia: John Benjamins.
- Jackendoff, Ray S. 1972. *Semantic interpretation in generative grammar*. Cambridge, MA: MIT Press.
- Jackendoff, Ray S. 1997. *The architecture of the language faculty*. Cambridge, MA: MIT Press.
- Jackendoff, Ray S. 2002. *Foundations of language. Brain, meaning, grammar, evolution*. Oxford: Oxford University Press.
- Jacobs, Joachim. 1983. *Fokus und Skalen*. Tübingen: Niemeyer.
- Jones, Graham B. & Bambi B. Schieffelin. 2009. Enquoting voices, accomplishing talk: uses of *be* + *like* in Instant Messaging. *Language and Communication* 29. 77–113.
- Kallmeyer, Werner & Inken Keim. 2003. Linguistic variation and the construction of social identity in a German-Turkish setting. In Jannis Androutsopoulos (ed.), *Discourse constructions of youth identities* (Pragmatics and Beyond 110), 29–46. Amsterdam & Philadelphia: John Benjamins.
- Keenan, Edward L. & Bernard Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8(1). 63–99.
- Klein, Wolfgang & Clive Perdue. 1997. The Basic Variety or: Couldn't languages be much simpler?. *Second Language Research* 13(4). 301–347.
- Koch, Peter & Wulf Oesterreicher. 1994. Schriftlichkeit und Sprache. In Helmut Günther & Ott Ludwig (eds.), *Writing and its uses. An interdisciplinary handbook of international research*, vol. 1, 587–604. Berlin & New York: Mouton de Gruyter.
- König, Ekkehart. 1991. *The meaning of focus particles: A comparative perspective*. London & New York: Routledge.
- König, Ekkehart. 1993. Focus particles. In Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld & Theo Vennemann (eds.), *Syntax: An international handbook of contemporary research*, vol. 1 (HSK 9.1), 978–987. Berlin & New York: Mouton de Gruyter.
- Koster, Jan. 1975. Dutch as an SOV language. *Linguistic Analysis* 1(2). 111–136.
- Krifka, Manfred. 2007. Basic notions of information structure. In Caroline Féry, Gisbert Fanselow & Manfred Krifka (eds.), *The notions of information structure* (ISIS 6), 13–55. Berlin & Potsdam: Working Papers of the SFB 632, Interdisciplinary Studies on Information Structure.
- Lehmann, Christian. 2000. *Thoughts on grammaticalization* (Arbeitspapiere des Seminars für Sprachwissenschaft der Universität Erfurt 9). 2nd revised edn. Erfurt: Seminar für Sprachwissenschaft.
- Lenerz, Jürgen. 1977. *Zur Abfolge nominaler Satzglieder im Deutschen*. Tübingen: Niemeyer.
- Meehan, Teresa. 1991. It's like, 'What's happening in the evolution of like?': A theory of grammaticalization. *Kansas Working Papers in Linguistics* 16. 37–51.
- Molnár, Valéria & Susanne Winkler (eds.), 2006. *The architecture of focus*. Berlin & New York: Mouton de Gruyter.
- Paul, Hermann. 1992. *Deutsches Wörterbuch*. 9th revised edn. by Helmut Henne & Georg Objartel with Heidrun Kämper-Jensen. Tübingen: Niemeyer.
- Pittner, Karin. 1993. *So und wie* in Redekommentaren. *Deutsche Sprache* 21. 306–325.
- Romaine, Suzanne & Deborah Lange. 1991. The use of *like* as a marker of reported speech and thought: A case of grammaticalization in progress. *American Speech* 66(3). 227–279.
- Rooth, Mats. 1985. *Association with Focus*. Amherst, MA: University of Massachusetts dissertation.

- Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1. 75–116.
- Sandig, Barbara. 1987. Kontextualisierungshinweise: Verwendungen von *so* im Prozeß sprachlichen Handelns. In Inger Rosengren (ed.), *Sprache und Pragmatik. Lunder Symposium 1986*, 327–333. Stockholm: Almqvist & Wiksell.
- Sankoff, Gillian. 1993. Focus in Tok Pisin. In Francis Byrne & Donald Winford (eds.), *Focus and grammatical relations in Creole languages*, 117–140. Amsterdam & Philadelphia: John Benjamins.
- Selting, Margret & Friederike Kern. 2009. On some syntactic and prosodic structures of Turkish German in talk-in-interaction. *Journal of Pragmatics* 41(12). 2496–2514.
- Stechow, Arnim von. 1991. Focusing and backgrounding operators. In Werner Abraham (ed.), *Discourse particles*, 37–84. Amsterdam & Philadelphia: John Benjamins.
- Steube, Anita. 2000. Ein kognitionswissenschaftlich basiertes Modell für Informationsstrukturierung. In Josef Bayer & Christine Römer (eds.), *Von der Philologie zur Grammatiktheorie: Peter Suchsland zum 65. Geburtstag*, 213–238. Tübingen: Niemeyer.
- Thiersch, Craig L. 1978. *Topics in German syntax*. Cambridge, MA: MIT dissertation.
- Thurmair, Maria. 2001. *Vergleiche und Vergleichen: Eine Studie zu Form und Funktion der Vergleichsstrukturen im Deutschen* (Linguistische Arbeiten 433). Tübingen: Niemeyer.
- Traugott, Elizabeth Closs. 1982. From propositional to textual and expressive meanings: some semantic-pragmatic aspects of grammaticalization. In Winfred P. Lehmann & Yakov Malkiel (eds.), *Perspectives on historical linguistics* (Current Issues in Linguistic Theory 24), 245–271. Amsterdam & Philadelphia: John Benjamins.
- Traugott, Elizabeth Closs. 2003. Constructions in grammaticalization. In Brian Joseph & Richard Janda (eds.), *The handbook of historical linguistics*, 624–647. Oxford: Blackwell.
- Underhill, Robert. 1988. *Like* is, like, focus. *American Speech* 63. 234–246.
- Veenstra, Tonjes & Hans den Besten. 1994. Fronting. In Jacques Arends, Pieter Muysken & Norval Smith (eds.), *Pidgins and Creoles: An introduction*, 303–315. Amsterdam & Philadelphia: John Benjamins.
- Weinrich, Harald. 2003. *Textgrammatik der deutschen Sprache*. Hildesheim: Olms.
- Wiese, Heike, Ulrike Freywald & Katharina Mayr. 2009. *Kiezdeutsch as a test case for the interaction between grammar and information structure* (Working Papers of the SFB 632, Interdisciplinary Studies on Information Structure [ISIS] 12). Potsdam: Universitätsverlag Potsdam.
- Wiese, Heike, Ulrike Freywald, Sören Schalowski & Katharina Mayr. 2009–2010. *KiezDeutsch Korpus*. A linguistic corpus of natural conversations among young people in a multiethnic urban neighbourhood [Main Corpus: transcriptions of self-recordings from peer-group communications in Berlin-Kreuzberg; Supplementary Corpus: comparative data from a monoethnic neighbourhood, Berlin-Hellersdorf]. University of Potsdam.
- Wiese, Heike. 2006. “Ich mach dich Messer”: Grammatische Produktivität in Kiez-Sprache. *Linguistische Berichte* 207. 245–273.
- Wiese, Heike. 2009. Grammatical innovation in multiethnic urban Europe: New linguistic practices among adolescents. *Lingua* 119. 782–806.
- Wiese, Heike. 2011. The role of information structure in linguistic variation: Evidence from a German multiethnolect. In Frans Gregersen, Jeffrey Parrott & Pia Quist (eds.), *Language Variation — European Perspectives III. Proceedings of ICLaVE 5*, 83–95. Amsterdam & Philadelphia: John Benjamins.
- Wischer, Ilse (2000). Grammaticalization versus lexicalization — ‘methinks’ there is some confusion. In Olga Fischer, Annette Rosenbach & Dieter Stein (eds.), *Pathways of change: Grammaticalization in English*, 355–370. Amsterdam & Philadelphia: John Benjamins.
- Zifonun, Gisela, Ludger Hoffmann & Bruno Strecker. 1997. *Grammatik der deutschen Sprache* (Schriften des Instituts für deutsche Sprache). 3 Volumes. Berlin & New York: De Gruyter.