

Reflecting the past

Mapping the development of the Indo-European SE-form*

Kees de Schepper

1. Introduction

It is often noted that many Indo-European languages possess a morpheme beginning with an *s* (French *se*, Italian *si*, German *sich*, Icelandic *sig*, Serbo-Croatian *se*), which has traditionally been classified as a reflexive pronoun (cf. Puddu 2003). Throughout the paper the term SE will be used to refer to this morpheme cross-linguistically: *se* is the French SE-form, *si* is the Italian SE-form, etc. A French example of this reflexive function of SE (Reflexive SE) can be found in (1).

- (1) Jean *se* voit à la télé
John SE sees at the TV
“John sees himself on TV”

Furthermore, it has been noted that marking reflexivity is not the only function of SE (Wehrli 1986, Hulk and Cornips 1996, Steinbach 1998, Dobrovie-Sorin 1998, and Reinhart and Siloni 2005, among others). Some other functions of SE that are quite common are Middle SE (2a), Ergative SE (2b), Inherent SE (2c), Reciprocal SE (2d) and Passive SE (2e).

- (2) a. le grec *se* traduit facilement
the Greek SE translates easily
“Greek translates easily”
b. la branche *se* casse
the branch SE breaks
“The branch breaks”
c. George *se* baigne
George SE bathe
“George takes a bath”

- d. Jean et Max se détestent
John and Max SE hate
“John and Max hate each other”
- e. les pommes se mangent en hiver
the apples SE eat in winter
“Apples are eaten during winter”

Typical of these functions is that SE combines with a verb that also has a SE-less transitive configuration. So for example, the transitive counterpart of (2b) is (3).

- (3) Jean casse la branche
Jean breaks the branch

Middle SE, as in (2a), is a construction used in some languages to express the so-called middle meaning (used here in its narrow sense; other terms are facilitative, medium, mediopassive or middle-passive). Characteristic of the middle meaning is the property-relation that seems to hold between subject and predicate: the predicate functions as a property of the subject. For example, (2a) could also be paraphrased as “Greek is easy to translate”, or “Greek is easily translatable”.

Ergative SE, as in (2b), is used by various languages to mark the ergative construction (also known as unaccusative, neutral, inchoative or anticausative) for some of its verbs. The ergative construction is a construction in which the grammatical subject is semantically a Theme, as e.g., *la branche* in (2b). Note that *la branche* has the same thematic role — Theme — in both (2b) and (3).

Inherent SE (also known as inherent reflexive, endoreflexive or idiomatic pronominal) is the name given to those predicates in which the combination of predicate and SE expresses a unified bodily action: the distinction between the Agent and the Theme role is collapsed, and there is effectively only a single participant role left. The example in (4) illustrates the situation with this distinction still intact. This sentence is the transitive counterpart of the Inherent SE predicate *se baigner* (“to bathe”) in (2c).

- (4) George baigne Max
George bathes Max

In (4) there is still a clear distinction between the Agent, *George*, and the Theme, *Max*: George is the one doing the washing while standing next to the tub and Max is the one inside the tub. In (2c), however, there is only George. He is both the Agent and the Theme in that sentence, but his actions are rather different from the actions he would perform if he was bathing someone else. This is exactly where Inherent SE differs from Reflexive SE: when a person sees herself on television, cf. (1a), the seeing action she performs is not different from the seeing action she performs when she sees someone else on television.

Reciprocal SE (2d) is a marker of reciprocity. Reciprocity means approximately that in a situation person A does something to person B and person B does the same thing to person A. For more information on reciprocity see Kemmer (1993) and Gast and Haas (forthc.).

The last SE-function discussed here is Passive SE (2e). In Passive SE-constructions, as in Ergative and Middle SE-constructions, the Theme is in subject position. However, Passive SE implies the presence of an Agent. So (2e) can also be translated as “People eat apples during winter”, but (2a) cannot be translated as “People translate Greek easily”, nor can (2b) be translated as “Somebody breaks a branch”.

There are not only differences between the six SE-functions but also similarities. For example, as illustrated for French in (2a-e), in many languages SE is used for more than one of the functions, which suggests a relationship between the functions. There have been attempts at explicating this relationship in the literature (e.g. Wehrli 1986, Steinbach 1998, Dobrovie-Sorin 1998, Reinhart and Siloni 2005), but these attempts fail to account coherently for the cross-linguistic data of the six SE-functions. In Section 2, the next section, I will describe the problems faced by these previous accounts. Then, in Section 3, I will propose a framework that is able to account for the cross-linguistic SE data. In Section 4, finally, I will conclude with the main points of this paper. In this section I will also give an overview of some issues for further research.

2. Previous accounts

The question that I will address in this paper is: “How are the different SE-functions related?” This “relational” question can be split into sub-questions. One such question that has received a lot of attention in the literature is the locus question: “In which language module does the SE-form get its appropriate SE-function?” There have been three kinds of approaches to this question: pre-syntactic, syntactic and post-syntactic. Pre-syntactic approaches argue that the SE-function is selected in the lexicon, whereas syntactic approaches posit that it happens in the syntax itself. Post-syntactic approaches, finally, stipulate that sentence semantics determine the SE-function of the SE-morpheme.

Although the locus problem is interesting in itself, any results in this area will not solve the “relational” question, as information on the locus of the SE-functions does not reveal much about the relationship between the SE-functions. A more helpful question is the “unification” question: “What is the underlying characteristic that all the SE-functions have in common?” Every approach assumes such a characteristic because of the similarities between the types, and the fact that these

types are occasionally expressed by the same linguistic element. This “unification question” is not easy to answer either, however. In some cases SE behaves as an anaphor (i.e. a subject-oriented pronoun, see Dobrovie-Sorin 1998 and Steinbach 1998) and therefore as an argument of its predicate, while in other cases it is more of a marker of argument reduction (see Wehrli 1986, Reinhart and Siloni 2005) and hence not an argument of its predicate. Yet, if this dual status of SE is correct, then what is the unifying characteristic of SE? And if there is no such characteristic, how can we maintain the view that there is really one SE?

Another useful question is the classification question: “How can the SE-functions be grouped together?” The SE-functions of a language usually pattern in such a way that one or more SE-functions have a particular characteristic that is lacking in the other types. These patterns suggest that the SE-functions allow systematic grouping, which will help to find an answer to the “relational” question above. Classifying SE-functions is not an easy matter, as any two SE-functions differ from each other in quite a number of ways. Grouping two types together may be misleading as it suggests that these SE-functions are more similar to each other than to any other SE-function, which is often not the case. Treating each individual SE-function as an autonomous category, however, which is roughly Reinhart and Siloni’s (2005) position, suggests that there are no patterns among the SE-functions at all, or, at best, that those patterns are distributed randomly, which is not the case. As we will see below, there really are systematic principles that underlie the patterns found among SE-functions, even though these principles are not as clear-cut as has been suggested in the literature (e.g. Wehrli 1984, Dobrovie-Sorin 1998 or Steinbach 1998).

3. A different look at SE

The previous section presented the question how a unifying approach to SE can be maintained if there is no unifying characteristic of SE-functions. But perhaps the question should be why we would want such an approach in the first place. The answer to this lies in the history of SE-functions.

3.1 The history of SE-functions

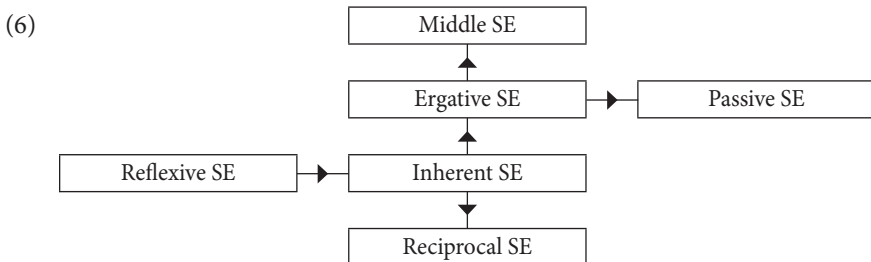
It is a well-known fact that all SE-forms go back to one Indo-European SE-morpheme. This suggests that SE has existed for a long time. Therefore it might be interesting to look at the history of the SE-functions too. Have all SE-functions been present from the start, or has there been a gradual increase in SE-functions? The answer to this may provide us information on the nature of the SE-functions.

Table (5) presents data on the presence of SE in some Indo-European languages old and new. It is possible to perceive a gradual increase in SE-functions: all languages developed Reflexive SE, but the same is not true of types like Passive SE. The data on the Romance languages suggest that this extension is indeed a historical process. Latin only had Reflexive SE, Old French developed an Inherent SE-function, and Modern French has all six types (cf. Kemmer 1993). Research on older stages of the Germanic languages (cf. Kemmer 1993 and references cited there) and Slavic languages (cf. Steinbach 1998:35–36 and references cited there) suggests a similar process. Furthermore, in the reconstruction of the Proto-Indo-European language that the function SE started out with was Reflexive SE (cf. Puddu 2003).

(5)	Reflexive	Inherent	Reciprocal	Ergative	Middle	Passive
Latin	+					
Old French	+	+				
Dutch	+	+		+		
German	+	+	+	+	+	
Danish	+	+	+	+		+
French	+	+	+	+	+	+
Russian	+	+	+	+	+	+

Note that languages may differ in the way they have implemented each function. The group of verbs that allow Inherent SE varies across languages, for example. It is however possible to compare languages in a more abstract fashion, as the table in (5) shows.

Another way of representing the data in (5) is by a semantic map, since the six SE-functions can be seen as different uses of one linguistic unit SE. A semantic map representation has the advantage of showing which types are closely related. Using the data in (5) I constructed a preliminary semantic map of SE-functions, which can be seen in (6).



The map in (6) clearly shows the diachronic development of the SE-functions. In the first stage a language only has Reflexive SE. The second type to develop is

Inherent SE, and the third Ergative and/or Reciprocal SE. Finally, once a language develops Ergative SE, it can also develop Middle and/or Passive SE. All this follows from the data in (5), with one possible exception. It is possible to argue that Reciprocal SE develops out of Ergative SE, because the set of languages in (5) that have Reciprocal SE is a subset of the languages that have Ergative SE. However, Reciprocal SE is closely related in meaning to Inherent and Reflexive SE, because only in these three types does the subject have to be animate. Therefore I choose to organize the SE-functions as in (6), with Reciprocal SE branching out of Inherent SE (see Gast and Haas *forthc.* and the next subsection for more discussion).

The map in (6) is not only useful for describing SE's diachronic development but it also solves the problem of unifying SE. In the first stage of SE's development, when its only type was the reflexive, SE was probably really an anaphor, as is claimed by Dobrovie-Sorin and Steinbach. Then, when SE developed other types it became a marker of argument reduction. The advantage of the map in (6) is that it posits a common ancestor of all SE types, and that it offers a diachronic explanation to account for the fact that each SE-function has some characteristics in common with the types to which it is closely related. The map also solves the problem of classifying SE, as it predicts that the only patterns possible for the SE-functions are patterns that occupy contingent areas of the semantic map. The patterns exhibited by the various languages offer some support for this prediction as a workable hypothesis.

Haspelmath (2003) presents a semantic map that is roughly similar to the one presented in (6). It should be noted, however, that Haspelmath's map also covers elements comparable to SE in all the languages of the world. Moreover, his (2003) map does not reflect the diachrony of SE, although this aspect was present in an earlier version (Haspelmath 1987). Although the central idea behind his map is correct, some of its details are problematic. One of his assumptions is that Passive SE branches out of Middle SE (Haspelmath's 'potential passive') instead of out of Ergative SE as in (6). From his assumption the prediction follows that a language that has Ergative SE (Haspelmath's 'anticausative') and Passive SE should also have Middle SE. There is, however, a clear counterexample to this prediction in Danish. Danish does not have Middle SE, see Bergeton (2004:293–294). This means that Passive SE and Middle SE cannot be combined into one medio-passive SE, *contra* much that has been said in the literature. The Danish data in (5) lead me to hypothesize that Passive SE develops directly out of Ergative SE.

Moreover, it is argued by some that Passive SE can be split into two functions, eventive Passive SE and habitual Passive SE. Dobrovie-Sorin (1998) shows that in French and European Portuguese Passive SE has a habitual meaning if its subject is in a preverbal position, and an eventive meaning if its subject is in postverbal position. However, in French there is no pro-drop, so the only situation in which the

subject is postverbal is when there is an impersonal dummy subject *il* occupying the preverbal subject positions. This is why Passive SE with impersonal *il* is evenitive in French, while Passive SE without impersonal *il* is habitual. Nevertheless, because eventive and habitual Passive SE coexist in both French and European Portuguese there is no compelling reason to include this distinction in the map in (6). However, there may be languages in which only one of the two Passive SE-types is available.

3.2 Morphological categories

In the previous subsection it was argued that there is a diachronic relationship between the SE-functions which explains their patterning. A similar diachronic ordering is also possible for the forms of SE. A SE-morpheme can take any of the following morphological forms: strong pronoun, weak pronoun, clitic or affix. The research on the grammaticalization of pronouns suggests a grammaticalization hierarchy for these four forms (cf. Bresnan 1998):

- (7) strong pronoun > weak pronoun > clitic > affix

If pronouns undergo grammaticalization, they tend to do it in the order described in (7), and the same order can be argued to hold for SE's forms, as stages in a process of grammaticalization.

The first stage in such a process is the strong pronoun stage. A strong pronoun can do anything a normal DP can, which means that it can be fronted. An example of a strong pronoun is Reflexive *sich* in German. What is interesting is that there are no strong pronoun forms for any of the other available SE-functions in German. In fact, other languages show similarly that if a language has a strong pronoun SE-form, this form is always limited to Reflexive SE, which must surely be significant.

A weak pronoun (a *simple clitic* in Zwicky's (1977) terminology) cannot be stressed, which explains why they cannot be fronted or undergo any other operation that requires stress, but is similar to the strong pronoun in other respects. Examples are Dutch *zich* and German non-reflexive *sich*. A clitic (a *special clitic* in Zwicky's (1977) terminology) is an element that occupies a position in syntax that is different from the normal DP position. An example of a SE-clitic is French *se*. An affix, finally, is a morpheme that is part of the inflectional paradigm of the verb, unlike clitics, which are always outside of the verb. An example of an affix is Russian *-sja/-s'*.

There are no restrictions on the SE-functions that clitics and affixes can be associated with. The only function that a weak pronoun cannot have is Passive SE. German, for example, in which there is only pronominal SE, has no Passive SE, but

it does have the other five functions of SE. This may be connected to the assumption that a weak pronoun like German *sich* still possesses accusative case, while for example a clitic like French *se* does not (cf. Reinhart and Siloni 2005). Passives in general cannot have accusative case and need additional verbal morphology to separate the Agent role from the subject position of the verb (example taken from Steinbach 1998):

- (8) a. *das Brot schneidet sich absichtlich
 the bread cuts SE on.purpose
 b. das Brot wurde absichtlich geschnitten
 the bread became on.purpose cut
 “The bread was cut on purpose”

In (8a), a sentence with SE and regular active morphology, the adverb *absichtlich* (“on purpose”) will be interpreted as referring to the subject *das Brot* (“the bread”) which makes the sentence nonsensical. In (8b), a sentence with passive morphology, the adverb *absichtlich* will be interpreted as referring to an entity other than the subject. In French and other Romance languages SE is a clitic and therefore part of the verbal morphology. This may be why Passive SE is possible in Romance but not in German (see also Dobrovie-Sorin 1998).

To sum up, these four morphological categories — strong pronouns, weak pronouns, clitics and affixes — correlate with the six SE-functions in the following way:

(9)

	Reflexive	Inherent	Reciprocal	Ergative	Middle	Passive
Strong pronoun	+	–	–	–	–	–
Weak pronoun	+	+	+	+	+	–
Clitic	+	+	+	+	+	+
Affix	+	+	+	+	+	+

Now because SE-functions seem to correlate with morphological form, and because both SE-function and morphological form express a diachronic dimension, (9) can be used to formulate claims about the development of SE and how formal and functional aspects tie into this development. The next subsection gives an example of such an application of the matrix in (9).

3.3 A scenario of the development of SE

In the light of our discussion so far, it seems likely that reflexive SE represents the initial stage of SE in the Indo-European languages, and that SE was a strong pronoun. The reflexive meaning of such a pronominal construction arises because of its anaphoric nature: the pronoun refers to the same person as the subject DP.

Then, at a later stage (for the Romance languages, for example, this is the stage in which the present Romance languages developed from Latin, see Kemmer 1993) the other SE-functions developed from this Reflexive SE. Inherent SE probably developed directly from Reflexive SE, as noted earlier. However, the emerging Inherent SE-verbs could no longer be considered transitive verbs, because these verbs cannot take other objects than SE under an Inherent SE meaning, and because the original two thematic roles effectively collapsed into a single role. As a consequence, SE could no longer be considered a strong pronoun, and was instead reanalyzed as a grammatical marker of some sort. This is how the SE-forms associated with Inherent SE lost their ability to be stressed.

The next stage is more problematic: the SE-form of Reflexive SE is associated with an argument — let us call it *argument SE* — whereas the SE-form of Inherent SE is not. These two forms behave differently, for example with respect to the ability to be stressed, and the resulting conflict could be solved by splitting its SE-form into two SE-forms, an argument and a non-argument one. The downside to this is that the language now has two nearly identical SE-forms instead of one. This option can be found in German, which has two *sich*'s, as illustrated in (10), an example from Gast and Haas (forthc.).

- (10) a. die Spieler konnten sich nicht leiden, aber sie mochten den Trainer
 the player could SE not bear but they liked the coach
 “The players couldn’t bear themselves, but they liked the coach”
 “The players couldn’t bear each other, but they liked the coach”
 b. sich konnten die Spieler nicht leiden, aber sie mochten den Trainer
 SE could the players not bear but they liked the coach
 “The players couldn’t bear themselves, but they liked the coach.”
 NOT:
 “The players couldn’t bear each other, but they liked the coach.”

In a sentence like (10a) *sich* is ambiguous between Reflexive SE and Reciprocal SE. However, if *sich* is fronted as in (10b), the Reciprocal SE reading is lost. When an element is fronted in German, it automatically receives stress, so Reciprocal *sich* is not possible in (10b), because it can only be a weak pronoun just like Inherent *sich*, while Reflexive *sich* is a strong pronoun.

Note that Gast and Haas (forthc.) have observed that there are languages with a strong pronoun Reciprocal SE. They claim that these forms developed directly out of Reflexive SE. This strong pronoun Reciprocal SE is similar to true reciprocal pronouns like English *each other*. Following Dimitriades (2004), among others, this suggests that strong pronoun Reciprocal SE belongs to the class of argument reciprocals and weak pronoun Reciprocal SE to the class of verbal reciprocals. There are some semantic differences between the two classes, so perhaps the map

in (6) should have two Reciprocal SE functions accordingly. However, because strong pronoun Reciprocal SE is only attested in Czech and some varieties of Polish, and because in both cases there is also a weak pronoun Reciprocal SE present, I will consider strong pronoun Reciprocal SE a marginal phenomenon. Future research may provide more insight into the phenomenon.

A second option to solve the problem of multiple SE's is to drop the strong pronoun SE-form. This option is found in Danish and Dutch. It means that the language has to come up with some other expression for an argument SE-form. In Dutch this led to the creation of *zichzelf*, a complex form of the reflexive:

- (11) a. *Jan wast zich en Marie
 John washes SE and Mary
 b. Jan wast zichzelf en Marie
 John wash SE.self and Mary
 "John washes himself"

In Dutch Reflexive *zich* can only be used when it does not bear stress, because of its status as weak pronoun. In a coordination structure as in (10a) both elements are automatically stressed, and therefore *zich* is not possible here. Instead Dutch must rely on the complex form *zichzelf*, see (11b), which is able to bear stress.

Another development introduced by Inherent SE is a change in the thematic role of the subject of an Inherent SE-verb, which was noted earlier. With Inherent SE, see (2d), the subject is still both the one initiating and undergoing the action, just like with Reflexive SE, but effectively there is one role: only one person is needed for taking a bath. When this is taken one step further we get the situation in (12).

- (12) a. George endort Jean
 George in.sleeps John
 "George puts John to sleep"
 b. Jean s' endort
 John SE in.sleeps
 "John puts himself to sleep" [Reflexive SE]
 "John falls asleep" [ErgativeSE]

What does this example show? As (12a) indicates the verb *endormir* means originally "to put someone to sleep". When combined with SE the verb obtains the predictable reflexive meaning "putting oneself to sleep", see (12b). Putting oneself to sleep is rather different from putting someone else to sleep, so *s'endormir* could be considered a one-participant predicate too. However, as putting oneself to sleep usually does not take that much effort, the focus of the predicate shifted to the "falling asleep" part rather than the causation part. This offers a plausible scenario

of how the Reflexive SE-function developed into an Ergative SE-function. The subject of an Ergative SE-verb cannot be considered an Agent, however, unlike the subject of an Inherent SE-verb. The subject of an Ergative SE-verb like *s'endormir* is really just a Theme.

So once the development illustrated by (12b) has begun, it became possible for SE-verbs to be combined with a Theme subject. For a verb like *s'endormir* this meant that their subject is obligatory animate, because the Theme of its transitive counterpart *endormir* is also animate. However, for many other Ergative SE-verbs this meant that inanimate subjects became possible. This development paved the way for the appearance of Middle and Passive SE.

To sum up, it is possible to construct a direct path from Reflexive SE via Inherent SE and Ergative SE towards Passive SE. As the examples above show, the table in (9) can be applied to specific cases in the development of SE. The validity of the table as a scenario of the development of SE must be left to future research.

4. Concluding remarks

In this paper I have dealt with some problems concerning SE, in particular the problem of unifying SE-functions, and the problem of classifying SE-functions. I showed that synchronic approaches are not sufficient for dealing with these problems. This is why I presented a diachronic semantic map approach, as in (6), based on Haspelmath (2003), which offers a solution to both the unification and the classification problems. I also showed that combining the semantic map with a diachronic take on the morphological form of SE can help us to construct a scenario for the historical development of the SE-morpheme.

This has only been a preliminary paper, so there are many aspects of SE that have not been addressed. Some of the aspects that future research may shed some more light on are the following. First of all there are SE-functions other than the six that are dealt with here, that need to be integrated in the theory, for example logophoric/long distance SE, middle distance/ECM SE, the Slavic object arbitrarization SE, nominative SE, and psych verb/subject experiencer verb SE. Secondly, it may be interesting to look at the meaning expressed by each SE-function to see if it can be expressed without SE, and look at the differences between these two modes of expression. Finally it might be useful to look at the actual development of SE in languages and language varieties in order to test the approach presented here.

Note

* I would like to thank Bettelou Los, the Optimal Communication research group of the Radboud University Nijmegen, the audience of the TIN-dag 2006 and two anonymous reviewers for their valuable comments.

References

- Bergeton, Uffe. 2004. The independence of binding and intensification. PhD diss., University of Southern California.
- Bresnan, Joan. 2001b. "The Emergence of the Unmarked Pronoun". *Optimality-Theoretic Syntax* ed. By Geraldine Legendre, Jane Grimshaw & Sten Vikner, 113–142. Cambridge, Mass.: The MIT Press.
- Dimitriadis, Alexis. 2004. Discontinuous reciprocals. Ms., UiL-OTS. University of Utrecht.
- Dobrovie-Sorin, Carmen. 1998. "Impersonal Se Constructions in Romance and the Passivization of unergatives". *Linguistic Inquiry* 29.399–437.
- Gast, Volker & Florian Haas. Forthcoming. "Reflexive and Reciprocal Readings of Anaphors in German and other European Languages". *Reciprocals and reflexives — Cross-linguistics and theoretical explorations (Trends in Linguistics)* ed. by Ekkehard König & Volker Gast, Berlin: Mouton de Gruyter.
- Haspelmath, Martin. 1987. *Transitivity alternations of the anticausative type*. (= *Arbeitspapiere N. F. 5*) Cologne, Germany: University of Cologne.
- Haspelmath, Martin. 2003. "The geometry of grammatical meaning: semantic maps and cross-linguistic comparison". *The new psychology of language: cognitive and functional approaches to language structure*, vol. 2, ed. by Michael Tomasello, 211–42. Mahwah, NJ: Lawrence Erlbaum.
- Hulk, Aafke & Leonie Cornips. 1996. Ergative reflexives in Heerlen Dutch and French. *Studia Linguistica* 50.1–21.
- Kemmer, Suzanne. 1993. *The middle voice*. Amsterdam: John Benjamins.
- Puddu, Nicoletta. 2003. "Reflecting on *se-/swe-: From typology to Indo-European and back". *Proceedings of the seventeenth annual UCLA Indo-European Conference (November 8–9, 2002)*. (= *The Journal of Indo-European Studies, Monograph 47*) ed. by Karlene Jones-Bley, Angela della Volpe, Martin Huld & Miriam Robbins Dexter, 63–82. Washington D.C.: Institute for the Study of Man.
- Reinhart, Tanya & Tal Siloni. 2005. "The Lexicon-Syntax parameter: Reflexivization and other arity operations". *Linguistic Inquiry* 36.389–436.
- Steinbach, Markus. 1998. "Middles in German: The syntax and semantics of transitive reflexive sentences". PhD Diss., Humboldt University, Berlin.
- Wehrli, Eric. 1986. On some properties of French clitic se. *The Syntax of Pronominal Clitics*. (= *Syntax and Semantics, 19*), ed. by Hagit Borer, 263–283. Orlando & London: Academic Press.
- Zwicky, Arnold M. 1977. *On clitics*. Bloomington: Indiana University Linguistics Club.