AN LFG ANALYSIS OF THE SO-CALLED REFLEXIVE MARKER IN POLISH

Agnieszka Patejuk and Adam Przepiórkowski
Institute of Computer Science, Polish Academy of Sciences

Proceedings of the LFG15 Conference
Miriam Butt and Tracy Holloway King (Editors)
2015
CSLI Publications
http://csli-publications.stanford.edu/
Abstract

This paper proposes an LFG analysis of SIĘ – a Polish word which is usually referred to as the reflexive pronoun even though it is not always the case that it is used in this function. Moreover, SIĘ can be used in more than one function at the same time; there is also the issue of haplology of SIĘ in verb chains and of sharing it under coordination – the proposed analysis captures these phenomena. The paper discusses the status of SIĘ, whether it is an argument or a marker (a co-head with no PRED), and treats it as a marker, though some counterevidence against a unified analysis is presented.

1 Introduction

This paper provides an LFG analysis of SIĘ, a Polish word which is typically referred to as a “reflexive” item (e.g. Kupść 2000, Przepiórkowski et al. 2002, Kibort 2004), even though this is only one of its possible functions. It aims to provide a comprehensive treatment (within the available space limits) instead of focusing on a selected aspect of this phenomenon. The analysis presented here has been implemented as a part of a large XLE grammar of Polish (Patejuk and Przepiórkowski 2012).

The initial sections of this paper present the linguistic data on the basis of attested examples taken from the National Corpus of Polish (Narodowy Korpus Języka Polskiego, NKJP; Przepiórkowski et al. 2012, http://nkjp.pl/) or retrieved using Google search engine. § 2 presents selected basic types of SIĘ in Polish and briefly discusses their properties. § 3 shows that SIĘ may bear more than one function in relation to the same predicate. § 4 presents the phenomenon of haplology of SIĘ in verb chains, while § 5 discusses sharing of SIĘ under coordination of verbal predicates.

§ 6 is devoted to the issue of the status of SIĘ: whether it should be represented as an argument of the relevant predicate or as a marker, i.e. a co-head that does not have its own PRED value, but only contributes certain features to the f-structure of the predicate it depends on. It presents the proposed analysis and formalisation of phenomena discussed earlier based on the latter representation of SIĘ – as a marker. However, this issue is discussed further in § 7, which shows potential problems with such an analysis of SIĘ. Finally, § 8 concludes the paper.

It is worth noting that Polish SIĘ has been described earlier in Kupść 2000, which provides a comprehensive discussion of rich data and an HPSG analysis for phenomena such as multifunctionality of SIĘ and haplology of SIĘ in verb chains. Sharing of SIĘ, though discussed there, was not formalised.  

1 All links provided in this paper were accessed on 20 December 2015.  

2 Sharing of SIĘ, though discussed there, was not formalised.
2 Basic uses of SIE

Examples below show selected basic uses of SIE: reflexive (REFL) in (1), reciprocal (RECIP) in (2), inherent (INH) in (3), impersonal (IMPRS) in (4).3,4

(1) Jacek golił się.
   Jacek.SG.M1 shaved.SG.M1 REFL
   ‘Jacek was shaving.’ (NKJP)

(2) Przez chwilę całowali się.
   for moment kissed.PL.M1 RECIP
   ‘For a moment, they were kissing each other.’ (NKJP)

(3) Ty też się śmiejesz.
   you.SG also INH laugh.SG.2
   ‘You’re also laughing.’ (NKJP)

(4) Po Edenie chodziło się nago.
   along Eden walked.3.SG.N IMPRS nude.ADV
   ‘One would walk nude in Eden.’ (NKJP)

Other uses of SIE, not discussed in this paper, include middle (which was, however, discussed in Kupść 2000), shown below (glossed as MID):

(5) Takie powieści zawsze dobrze się sprzedawały.
   such.PL.F novel.PL.F always well MID sold.PL.F
   ‘Such novels have always sold well.’ (NKJP)

While all types of SIE listed in (1)–(4) above have the same shape, the following subsections discuss their properties with respect to criteria such as the possibility of being replaced with SIEBIE (§ 2.1), being required lexically or being constructional (§ 2.2), making a contribution to the semantics (§ 2.3) and finally imposing extra constraints, for instance on verbal agreement (§ 2.4).

2.1 SIE vs. SIEBIE

Though SIE is typically described as a weak form of the pronoun SIEBIE ‘self’ (e.g. Kupść 2000, but also in more traditional linguistic literature), it can only be replaced with appropriate forms of SIEBIE in (1)–(2), which feature reflexive and reciprocal SIE, respectively:

(6) Jacek golił siebie (samego).
   Jacek.SG.M1 shaved.SG.M1 self.ACC own.ACC
   ‘Jacek was shaving himself.’

(7) Przez chwilę całowali siebie (nawzajem).
   for moment kissed.PL.M1 self.ACC reciprocally
   ‘For a moment, they were kissing each other.’

3Though the word nago is translated as ‘nude’ in (4), it is an adverb – as indicated in glosses.
4Morphosyntactic abbreviations mostly adhere to those recommended in Leipzig Glossing Rules (http://www.eva.mpg.de/lingua/resources/glossing-rules.php), with the exception of M1, which refers to the human-masculine gender (assuming the repertoire of five genders in Polish proposed in Mańczak 1956).
Though, as shown above, certain types of SIE can be replaced with SIEBIE, it does not follow that these forms are equivalent and that they have the same syntactic properties – this issue is discussed in more detail in § 6.

2.2 Required lexically vs. constructional

Some types of SIE are a lexical requirement of the given predicate: these include reflexive (cf. (1)), reciprocal (cf. (2)) and inherent (cf. (3)) SIE.

By contrast, impersonal SIE shown in (4) is not required lexically – it is a productive construction, which may be used with almost any predicate that allows for a human subject. This constraint is illustrated below: since the subject of BOLEĆ ‘hurt’ in (10) is not human, the use of impersonal SIE is ungrammatical.

(10) Bolało mnie (kolano).
    hurt.SG.N I.ACC knee.NOM.SG.N
    ‘My knee would (often) hurt.’

(11) *Bolało się
    hurt.SG.N I.ACC IMPRS
    ‘It would (often) hurt.’ (intended)

2.3 Contribution to semantics

Another property to be taken into account is whether the given type of SIE contributes to the semantics – it is the case with all types of SIE discussed above except for inherent SIE, which is required syntactically but it does not make any contribution to the semantics. By contrast, the remaining types of SIE make such a contribution: the reflexive SIE marks that the respective predicate is reflexive, the reciprocal SIE marks that it is reciprocal, influencing the semantic interpretation, while the impersonal SIE marks that the predicate is impersonal.

2.4 Agreement constraints

When the impersonal SIE is used, it triggers default agreement on the verb (Dziwirek 1990) – it must be third person, singular, neuter, though it must be noted that neuter gender specification is only visible with past tense forms such as in (12), but not with a present tense form such as in (13).

(12) Po Edenie chodziło/*chodziły się nago.
    along Eden walked.3.SG/PL.N IMPRS nude.ADV
    ‘One would walk nude in Eden.’ (NKJP)

(13) Po Edenie chodzi/*chodzi się nago.
    along Eden walk.3.SG/PL IMPRS nude.ADV
    ‘One walks nude in Eden.’
Furthermore, as shown in (14), the impersonal SIE blocks the use of a lexical subject. It is the case even when the lexical subject would normally be compatible with such a form (third person, singular, neuter), compare (15).

(14) Po Edenie (*każe dziecko) chodziło się nago.  
    along Eden (*every child.NOM.SG.N walked.3.SG.N IMPRS nude.ADV

(15) Po Edenie każde dziecko chodziło nago.  
    along Eden every child.NOM.SG.N walked.3.SG.N nude.ADV

    ‘Every child walked nude in Eden.’

3 Multifunctional use of SIE

It is possible for SIE to have more than one function with respect to the same predicate at the same time: in both examples below SIE is impersonal (as in (4)), apart from being inherent in (16) (as in (3)) and reflexive in (17) (as in (1)). Multifunctionality of SIE is represented in glosses by joining appropriate basic types of SIE using the + symbol.

(16) Kiedyś się śmiało z czerwonych.  
    sometime laughed.SG.N INH+IMPRS from red.PL

    ‘One used to laugh at the communists.’ (Google)

(17) Kiedyś goliło się zyletkami.  
    sometime shaved.SG.N REFL+IMPRS razor blade.PL.INST

    ‘One used to shave with razor blades.’ (Google)

Furthermore, as noted in Kupść 2000, § 3.5.1, it is ungrammatical to use more than one instance of SIE as dependents of the same predicate, so the use of multifunctional SIE is the only way to obtain the readings presented in (16)–(17) above:

(18) *Kiedyś się śmiało się z czerwonych.  
    sometime IMPRŠ laughed.SG.N INH from red.PL

(19) *Kiedyś się goliło się zyletkami.  
    sometime IMPRS shaved.SG.N REFL razor blade.PL.INST

    It is perhaps worth mentioning that although the interpretations provided in (16)–(17) are the intended readings, there are alternative readings of these sentences which depend on the interpretation of SIE. (16) has a reading where SIE is not multifunctional but only inherent – under this reading the subject is implicit and, because of the agreement properties of the verb, whose form is third person singular neuter, it could, for instance, refer to a child, as shown in (20):

(20) Kiedyś (dziecko) się śmiało z czerwonych.  
    sometime child.NOM.SG.N laughed.SG.N INH from red.PL

    ‘Once a child laughed at the communists.’

5http://forum.wirtualnemedia.pl/teologowie-moralisci-apeluja-do-mediow-t9375.html
(17) in turn has two alternative readings: one where SIĘ is only reflexive and
the subject is implicit (as discussed for (20)), shown in (21), and the other where
SIĘ is only impersonal (with an implicit subject that blocks the use of a lexical one,
see § 2.4) and the predicate GOLIĆ takes an implicit object, see (22):

(21) Kiedyś (dziecko) gośilo się żyletkami.
sometime child.NOM.SG.N shaved.SG.N REFL razor blade.PL.INST
‘A child used to shave with razor blades.’

(22) Kiedyś gośilo się (ludzi) żyletkami.
sometime shaved.SG.N IMPRS people.ACC razor blade.PL.INST
‘One used to shave people with razor blades.’

All readings presented here are accounted for by the proposed analysis.

4 Haplology of SIĘ in verb chains

When two (or more) predicates in a verb chain consisting of the main verb and any
nonempty sequence of subordinate infinitival clauses require SIĘ, the one belong-
ing to the structurally higher predicate can at the same time satisfy the requirements
of the lower one – as a result, it is possible to use only one instance of SIĘ, as in
(23).§ Such sharing of SIĘ in verb chains is represented in glosses by joining the
types of SIĘ required by respective verbs using the / symbol.

(23) A czy Tobie zdarzyło się śmiać z dowcipu który
and Q you.DAT happened.SG.N INH/INH laugh.INF from joke which
nie był [. . . ] śmieszny?
NEG was funny
‘Have you happened to laugh at a joke that was not funny?’ (Google)§

In (23) both predicates – ZDARZYĆ ‘happen’ and ŚMIAĆ ‘laugh’ – require the
inherent SIĘ (as shown in (24)), but there is only one textual instance of SIĘ, which
structurally belongs to the higher predicate, namely ZDARZYĆ, as demonstrated in
(25), which is ungrammatical with SIĘ placed in the lower clause.

(24) A czy Tobie zdarzyło się śmiać się z dowcipu który
and Q you.DAT happened.SG.N INH laugh.INF INH from joke which
nie był [. . . ] śmieszny?
NEG was funny

(25) *A czy Tobie zdarzyło się śmiać się z dowcipu który
and Q you.DAT happened.SG.N laugh.INF INH/INH from joke which
nie był [. . . ] śmieszny?
NEG was funny

The following examples show that haplology of SIĘ is also possible when,
unlike in (23), the respective verbs in the verb chain require different types of SIĘ:

§czy, glossed as Q in (23)–(25), is the Polish yes/no question particle.
§http://www.druga-strona.pl/oobe-podroz-poza-cialem/czy-warto-
zdobyc-kazda-wiedze-na-temat-oobe
(26) Sporo osób boi się golić pod włos.
many people fear.SG.N INH/REFL shave.INF under hair
‘Many people are afraid of shaving against the grain.’  (Google)
(27) Sporo osób boi się golić się pod włos.
many people fear.SG.N INH shave.INF REFL under hair
(28) Kiedy chodziło się do kina gapić na wielki mrugający sometime walked.SG.N IMPRS/INH to cinema stare.INF at great blinking ekran.
screen
‘One would go to a cinema to stare at the great blinking screen.’  (NKJP)
(29) Kiedy chodziło się do kina gapić się na wielki mrugający sometime walked.SG.N IMPRS to cinema stare.INF INH at great blinking ekran.
screen
In (26) the main verb, BAČ ‘fear’, requires the inherent SIĘ, while its infinitival complement, namely GOLIĆ ‘shave’, takes the reflexive SIĘ – as shown in (27). Similarly, in (28), the main verb CHODZIĆ ‘walk’ forms a construction with the impersonal SIĘ, while the embedded predicate GAPIĆ ‘stare’ requires the inherent SIĘ – as demonstrated in (29).
Finally, as mentioned when discussing multifunctionality of SIĘ, sentences with SIĘ may be ambiguous depending on the interpretation of SIĘ. This is the case in (26), which has has an alternative reading presented in (30) where haplogogy is not involved – SIĘ is only a dependent of the main verb (which takes inherent SIĘ) and SHAVE has an implicit object (as in (22)):
(30) Sporo osób boi się golić (ludzi) pod włos.
many people fear.SG.N INH shave.INF people.ACC under hair
‘Many people are afraid to shave people against the grain.’
This reading is also taken into account under the proposed analysis.

5 Sharing SIĘ under coordination

Like other dependents, SIĘ can be shared by coordinated predicates: not only in the typical situation when it is one of the edge constituents, as in (31), but also when it is “intertwined” inside one of the conjuncts, as shown in (32), where it would normally belong to the phrase headed by CAŁOWAĆ. Such sharing is represented in glosses by joining types of SIĘ required by the coordinated predicates using |.
(31) Całowali i przytulali się w metrze.
kissed.PL.M1 and hugged.PL.M1 RECIP|RECIP in subway
‘They were kissing and hugging each other in subway.’  (Google)
(32) Całowali | przytulali się w metrze.
kissed.PL.M1 | and hugged.PL.M1 RECIP|RECIP in subway
‘They were kissing and hugging each other in subway.’
(32) Całowali się i przytulali!
   kissed.PL.M1 RECIP | RECIP and hugged.PL.M1
   ‘They were kissing and hugging each other!’
   (Google)

(33) Całowali się i przytulali się!
   kissed.PL.M1 RECIP and hugged.PL.M1 RECIP
As it is the case in verb chains, while SIĘ may have the same function with both co-
ordinated verbs – see (31)–(32), where CAŁOWAĆ ‘kiss (each other)’ and PRZYTU-
LAĆ ‘hug (each other)’ both require the reciprocal SIĘ (as shown in (33)) – it may
also be different for each conjunct – compare (34), where SIĘ required by ŚMIĄĆ
‘laugh’ is inherent, while with the figurative PUKAĆ SIĘ W GŁOWĘ, lit. ‘knock
yourself in head’, i.e., ‘imply that somebody is nuts’, it is reflexive (see (35)).

(34) Śmiali się i pukali w głowy.
   laughed.PL.M1 INH | REFL and knocked.PL.M1 in heads
   ‘They were laughing and asking if somebody is nuts.’
   (NKJP)

(35) Śmiali się i pukali się w głowy.
   laughed.PL.M1 INH and knocked.PL.M1 REFL in heads

6 Representation of SIĘ: a marker

Though, as mentioned in § 2.1, some types of SIĘ can be replaced with SIEBIE, even
there the referential properties of the two forms differ in the sense of Jackendoff
1992 – the following examples, whose context is a photograph of Zlatan Ibramović
next to the wax figure of himself in the Musée Grévin, show differences in meaning
transfer between these two forms:

(36) Zlatan podziwiał się w muzeum.
   Zlatan admired REF in museum
   ‘Zlatan admired himself in the museum’.

(37) Zlatan podziwiał siebie w muzeum.
   Zlatan admired self.ACC in museum
   ‘Zlatan admired himself/the wax figure of himself in the museum’.

In (36) the use of SIĘ seems to force strict identity: it can only mean that Zlatan
admired himself (rather than the wax figure of himself). By contrast, (37) has two
readings: one is the reading with identity by co-reference where Zlatan admires
himself (same as in (36)), while the other is the sloppy reading where Zlatan ad-
mires the wax figure of himself (unavailable in (36)). This difference suggests that
while SIEBIE is an argument and there are 2 syntactic participants in sentences with
SIEBIE (subject and object, see (39)), SIĘ is rather a marker, a co-head that does
not have a PRED value of its own, so there is only one participant (the subject, as
in (38)), which would explain the requirement of the strict identity reading in (36).

http://www.pudelek.pl/artykul/63056/rozenek_i_majdan_razem_w_
mcdonaldzie_calowali_sie_i_przytulali/
The formalisation of the proposed analysis of SI\text{\`e} is inspired by the analysis of case offered in Dalrymple et al. 2009, which involves the use of a complex CASE attribute containing subattributes corresponding to particular values of case, each of which takes a boolean value, thereby making it possible to account for multiple case marking of a dependent in languages such as Polish, where different predicates can impose different case requirements on the shared dependent. This idea can be adapted to account for the multifunctionality of SI\text{\`e} discussed in § 3.

6.1 Templates used in lexical entries

The template called in the lexical entry of SI\text{\`e} is provided in (40):

(40)  \text{SIE} \equiv (\uparrow \text{SIE \text{\`e} PRESENT}) = + \land [\text{SIE-LEX-IMPRS} \lor \text{SIE-IMPRS-ONLY}]

(41)  \text{SIE-LEX-IMPRS} \equiv \text{SIE-LEX} \land (\text{SIE-IMPRS})

(42)  \text{SIE-IMPRS-ONLY} \equiv \text{SIE-IMPRS} \land \neg \text{SIE-LEX}

(43)  \text{SIE-LEX} \equiv (\uparrow \text{SIE \{INH|REFL|RECIP\}}) = c + \land (\uparrow \text{SUBJ}) = %S \land (\%S \text{ PRED}) = \text{‘PRO’} \land (\%S \text{ PERS}) = 3 \land (\%S \text{ NUM}) = \text{SG} \land (\%S \text{ GEND}) = \text{N}

The definition of the SIE template consists of two conjoined statements. The first one introduces the attribute SIE, which contains the attribute PRESENT, whose value is set to +: it marks the presence of SI\text{\`e}, which is analysed as a co-head of the relevant verb, so it contributes to its f-structure. In order to account for the fact that only one instance of SI\text{\`e} may be used with the given predicate, as discussed in § 3, the attribute PRESENT is an instantiated feature whose value may be set only once.

The second conjunct of (40) is a disjunction of two template calls; let us start with the discussion of its first disjunct. SIE-LEX-IMPRS, whose definition is given in (41), rewrites to a conjunction of two template calls: the first one, SIE-LEX defined in (43), checks whether the attribute SIE has the positive value for any of the relevant types of SI\text{\`e} (reflexive, reciprocal or inherent), which accounts for lexically required SI\text{\`e}, while the second one is the optional call (put in brackets) to the template SIE-IMPRS defined in (44). If this call is used, it adds the attribute IMPRS to the value of the SIE attribute, making SI\text{\`e} multifunctional: it is both impersonal and of one of the other three types listed above (see § 3). Additionally, SIE-IMPRS imposes constraints on the subject of the verb to which it attaches: it introduces a pro subject to block the occurrence of a lexical one and it imposes appropriate agreement constraints (third person, singular number, neuter gender).
Let us now discuss the template SIE-IMPRS-ONLY called in the second disjunct of the second conjunct of (40), defined in (42) as a conjunction of two template calls: the first conjunct is an obligatory (not in brackets) call to the template SIE-IMPRS defined in (44) and discussed above, which introduces the constructional impersonal SIE, while the second conjunct is a negated (¬) call to the SIE-LEX template – its result is precluding SIE from being of any of the types defined in (43), which includes inherent, reflective and reciprocal SIE. The effect of these constraints is that SIE introduced by SIE-IMPRS-ONLY cannot be multifunctional – it can only be impersonal.

The last part of the analysis is placed in the lexical entries of verbs requiring SIE – they contain calls to the template defined in (45), where the value of the parameter TYPE corresponds to the type of required SIE (INH, REFL or RECIP; impersonal SIE of IMPRS type is not included here as it is constructional, as discussed in § 2.2):

\[(45)\quad \text{SIE-TYPE}(\text{TYPE}) \equiv (\uparrow \text{SIE TYPE}) = + \land (\uparrow \text{XCOMP}^* \uparrow \text{SIE PRESENT}) = c +\]

The first conjunct of (45) specifies the type of required SIE, while the second one ensures that SIE is present in the relevant domain – it looks for the positive value of the PRESENT attribute inside the SIE attribute locally to the verb (when XCOMP* in (45) corresponds to zero instances of XCOMP, the path then points to the local verb) or higher in the verb chain (when the local verb is nested in a structure containing one or more successive XCOMPS), which accounts for the haplology of SIE discussed in § 4. Sample lexical entries for verbs used in examples (1)–(3) are provided below:

\[(46)\quad \text{golić} \quad \text{V} \quad (\uparrow \text{PRED}) = ‘\text{SHAVE}<\uparrow \text{SUBJ}>’ \land @ (\text{SIE-TYPE REFL})\]
\[(47)\quad \text{calować} \quad \text{V} \quad (\uparrow \text{PRED}) = ‘\text{KISS}<\uparrow \text{SUBJ}>’ \land @ (\text{SIE-TYPE RECIP})\]
\[(48)\quad \text{śmiać} \quad \text{V} \quad (\uparrow \text{PRED}) = ‘\text{LAUGH}<\uparrow \text{SUBJ}>’ \land @ (\text{SIE-TYPE INH})\]

6.2 Simple structures

The following examples provide simplified f-structures for sentences presented in § 2: (49) and (50) correspond to (1) and (2), which feature reflexive (REFL) and reciprocal (RECIP) SIE, respectively. Due to the adopted analysis of SIE as a marker that does not have its own PRED attribute (see § 6), these f-structures feature one-argument (rather than two-argument) predicates SHAVE and KISS.

\[(49)\quad \begin{array}{c}
\text{PRED ‘SHAVE']} \\
\text{SUBJ} \square \\
\text{CASE NOM} \\
\text{SIE} \quad \text{REFL} + \\
\text{PRESENT} + \\
\end{array} \quad \text{SUBJ} \quad \text{PRED ‘JACEK’} \\
\text{NUM SG} \quad \text{NUM PL} \quad \text{NUM SG} \quad \text{NUM PL} \quad \text{NUM SG} \quad \text{NUM PL}\]

\[(50)\quad \begin{array}{c}
\text{PRED ‘KISS']} \\
\text{SUBJ} \square \\
\text{CASE NOM} \\
\text{SIE} \quad \text{RECIP} + \\
\text{PRESENT} + \\
\end{array} \quad \text{SUBJ} \quad \text{PRED ‘PRO’} \\
\text{NUM PL} \quad \text{NUM PL} \quad \text{NUM PL} \quad \text{NUM PL}\]

The representation of the remaining two types of SIE is not controversial: (51) corresponds to (3), which features inherent (INH) SIE, while (52) provides a sim-
plified representation of (4), which contains an impersonal (IMPRS) SI. 

(51)

\[
\begin{array}{c}
\text{PRED} \quad \text{LAUGH} \langle 1 \rangle \\
\text{SUBJ} \quad \text{OBL} \quad \text{SIE}
\end{array}
\]

(52)

\[
\begin{array}{c}
\text{PRED} \quad \text{WALK} \langle 1 \rangle \\
\text{SUBJ} \quad \text{ADJ} \quad \text{SIE}
\end{array}
\]

6.3 Multifunctional SI

Let us now consider some structures produced by this analysis for sentences that involve multifunctional SI discussed in § 3 – the f-structures in (53)–(54) provide a representation of simplified versions of (16) (śmiało się z czerwonych) and (17) (goliło się zyletkami).

(53)

\[
\begin{array}{c}
\text{PRED} \quad \text{LAUGH} \langle 1 \rangle \\
\text{SUBJ} \quad \text{OBL} \quad \text{SIE}
\end{array}
\]

(54)

\[
\begin{array}{c}
\text{PRED} \quad \text{SHAVE} \langle 1 \rangle \\
\text{SUBJ} \quad \text{ADJ} \quad \text{SIE}
\end{array}
\]

In (16) SI is at the same time inherent (specified lexically by LAUGH) and impersonal (constructionally), while in (17) SI is both reflexive (as specified by SHAVE) and impersonal.

Such multifunctional specifications are the result of the interaction of lexical entries of verbs requiring SI lexically and the lexical entry of SI (see (40)), which contains a call to the template SIE-LEX-IMPRS defined in (41) that can optionally introduce the specification of the impersonal SI (given in (44)) if it satisfies the condition formalised in (43), namely that the given instance of SI is already specified lexically by the appropriate verb (using (45)) as belonging to one of the following types: INH (as in (53)), REFL (as in (54)) or RECIP.

6.4 Haplology of SI

Under the current analysis, SI contributes the attribute PRESENT inside the SIE attribute of the relevant verb (as a result of calling the template defined in (40)), while verbs whose lexical requirement of SI is satisfied as a result of haplology do not have this attribute, which makes it possible to identify instances of SI haplology.
in f-structures. Let us discuss the f-structure provided in (55), which corresponds to a simplified version of (26) (*Sporo boi si€ goli´c*) and uses this mechanism:

(55) 

\[
\begin{align*}
\text{PRED} &\rightarrow \text{FEAR}
\text{SUBJ} &\rightarrow \text{MANY}
\text{XCOMP} &\rightarrow \text{SHAVE}
\end{align*}
\]

The higher predicate, FEAR, to which SI€ belongs structurally, contains two attributes inside SI€: PRESENT, contributed by the lexical instance of SI€, and INH, which corresponds to the type of SI€ required lexically by FEAR (with the help of the template defined in (45)). By contrast, the SI€ attribute of the infinitival complement (XCOMP) only contains the REFL specification of SI€ which is required by SHAVE – since this predicate does not have a lexical SI€ as its dependent but it uses haplology to satisfy this requirement, there is no PRESENT attribute inside its SI€ attribute. It is worth noting that, since it is the verb that determines the type of SI€ (using a defining equation in (45)), it is possible to handle haplology where different types of SI€ are involved, as in (26) discussed above, where the main verb takes the inherent SI€, while its infinitival complement takes the reflexive SI€. Similarly, in (28) SI€ belonging to the main verb is impersonal (constructional, not required lexically as in (26)), while the infinitival complement takes the inherent SI€.

Finally, such an analysis of haplology ensures that the requirement of an appropriate type of SI€ of relevant predicates is satisfied without placing the SI€ attribute in the f-structures of all subsequent predicates in the verb chain which may have no such requirement, as in (56), whose f-structure is provided in (57) – as shown in (58), unlike BAĆ ‘fear’ and ŠMIAĆ ‘laugh’, PRÓBOWAĆ ‘try’ does not require SI€ (this is represented using . . . in glosses in (56)) and putting the SI€ attribute in the corresponding partial f-structure would distort its semantics.

(56) 

\[
\text{Bal si€ próbować uśmiechnąć.}
\]

‘He was afraid to try to smile.’

(57) 

\[
\begin{align*}
\text{PRED} &\rightarrow \text{FEAR}
\text{SUBJ} &\rightarrow \text{PRO}
\text{XCOMP} &\rightarrow \text{SMILE}
\end{align*}
\]
6.5 Sharing $\text{SIE}$ under coordination

The last remaining issue is sharing of $\text{SIE}$ under coordination. While examples where $\text{SIE}$ is one of the edge constituents, as in (31), can be handled using standard mechanisms for dependent sharing under coordination, such a solution is not sufficient to account for more difficult cases such as the ones provided in (32) and (34), where $\text{SIE}$ is placed after the first conjunct. It would therefore belong to the first conjunct and, under standard coordination rules, could not be shared with the other conjunct.

However, such examples can be accounted for by introducing a general analysis of “intertwined” shared dependents under coordination using the following rules, where DEP corresponds to the shared dependent and RM corresponds to $\text{SIE}$:

\begin{align}
\text{(59)} & \quad \text{IPtop} \rightarrow \text{IP} \quad \text{DEP} \quad \text{Conj} \quad \text{IP} \\
& \quad \downarrow \in \uparrow \quad \uparrow = \downarrow \quad \downarrow = \uparrow \quad \downarrow \in \uparrow \\
\text{(60)} & \quad \text{DEP} \rightarrow \{ \operatorname{ARG} \mid \operatorname{MOD} \mid \operatorname{RM} \} \\
& \quad (\uparrow \operatorname{GF}) = \downarrow \quad \downarrow \in (\uparrow \operatorname{ADJUNCT}) \quad \uparrow = \downarrow \\
\end{align}

Such rules provide an account of examples such as the following ones, which feature plain argument sharing, as well as (32) and (34), which feature sharing of $\text{SIE}$:

(61) Przyjechali żandarmi i chodzili od domu
domu.

to house

‘Military policemen came and they walked from house to house.’ (NKJP)

(62) Zakleiła kopertę i wepchnęła do torebki.

‘She sealed the envelope and she pushed it into the handbag.’ (NKJP)

In (61) the shared dependent is the subject żandarmi, while in (62) it is the object kopertę. In (32) verbs share $\text{SIE}$ of the same type (reciprocal), while in (34) they share $\text{SIE}$ of different types (inherent vs. reflexive). The analysis proposed above makes it possible to account for all these examples.

While Frank 2002 discusses German examples similar to (61) and offers an analysis of such coordination in terms of Grammaticalised Discourse Functions (GDF) and structure sharing of relevant f-structure fragments between conjuncts, her analysis does not seem to be applicable in the context of Polish. First, in German this phenomenon is strictly limited to the SUBJ grammatical function, which is not the case in Polish, where other dependents can also be shared, which includes the object, as shown in (62), adjuncts (not presented here for reasons of space) and, more importantly, $\text{SIE}$. Secondly, since $\text{SIE}$, a weak form that cannot be stressed (Kupść 2000), is under the current analysis a marker rather than an argument (this

\begin{footnote} (62) also features a shared implicit subject. \end{footnote}
is uncontroversial for inherent and impersonal \( \text{SE} \), it seems inappropriate to assign it discourse functions such as \text{TOPIC} or \text{FOCUS}. The last option discussed in Frank 2002, namely the \text{SUBJ} grammatical function, is not taken into account since \( \text{SE} \) is not the subject in Polish. Finally, the strongest argument against adopting such an analysis for Polish is that structure-sharing of the f-structure fragment corresponding to \( \text{SE} \) would make it impossible to account for sentences such as (34), where conjuncts require different basic types of \( \text{SE} \) – since particular verbs call (45) to specify the type of required \( \text{SE} \), two different calls would result in a multifunctional specification of \( \text{SE} \), which is an unwanted effect (changing the meaning).

Let us now proceed to the discussion of structures produced by the proposed analysis for Polish: (63) is the f-structure corresponding to (32),\(^{13}\) where both conjuncts require the same type of \( \text{SE} \), namely reciprocal \( \text{SE} \):

\[
(63) \begin{align*}
&\{[\text{PRED 'KISS]\langle 1 \rangle}'] \quad [\text{PRED 'PRO']}'] \\
&\{[\text{SUBJ } \text{CASE NOM} ] \quad [\text{NUM PL}]
\}
\end{align*}
\begin{align*}
&\{[\text{SIE } \text{RECIP} + ] \quad [\text{PRESENT} + ]
\}
\end{align*}
\begin{align*}
&\{[\text{COORD-FORM } \text{AND}]
\}
\end{align*}
\]

The f-structure in (64) corresponds to (34), where coordinated verbs take different types of \( \text{SE} \) (inherent vs. reflexive). Accounting for the fact that the coordinated predicates in (34) require two different basic types of \( \text{SE} \) is possible due to the fact that the call to the template \( \text{SE} \) defined in (40) (placed in the lexical entry of \( \text{SE} \)) introduces the defining equation \( \uparrow \text{SE PRESENT} = + \), which distributes over all conjuncts, but the constraining equation defined in (43) uses functional uncertainty, which makes it possible to evaluate it separately for each conjunct, which in turn makes it possible for particular conjoined verbs to feature different types of \( \text{SE} \) – as discussed in § 6.1, it is the verb that specifies the type of \( \text{SE} \) using the defining equation in (45).

\[
(64) \begin{align*}
&\{[\text{PRED 'LAUGH}\langle 1 \rangle}'] \quad [\text{PRED 'PRO')}'] \\
&\{[\text{SUBJ } \text{CASE NOM} ] \quad [\text{GEND M} 1] \\
&\{[\text{NUM PL} ] \quad [\text{PERS 3}]
\}
\end{align*}
\begin{align*}
&\{[\text{SIE } \text{INH} + ] \quad [\text{PRESENT} + ]
\}
\end{align*}
\begin{align*}
&\{[\text{COORD-FORM } \text{AND}]
\}
\end{align*}
\]

The proposed analysis also accounts for the impossibility of sharing \( \text{SE} \) under

\(^{13}\) As shown in (63) and (64), the implicit subject is shared by the coordinated verbs.
coordination where the predicate in one conjunct would lexically require SIĘ (inherent, reflexive or reciprocal), while with the other SIĘ would be constructional (impersonal), as in the following putative example from Kupšć 2000:14

(65) Dziecko potykało się i trudno mu *(się)
child.NOM.SG.N stumbled.3.SG.N INH and difficult it.DAT IMPRS
wchodziło po schodach.
climbed.3.SG.N after stairs
‘The child stumbled and had difficulty climbing the stairs.’
(Kupšć 2000, p. 150, ex. (3.127))

Under the current analysis (65) is ungrammatical not because Dziecko is the shared subject (this is because it is not obligatorily shared), which would conflict with the impersonal SIĘ introducing its own implicit subject (as discussed in § 2.4 and formalised in (44)). It is ungrammatical due to the way (40), the template called in the lexical entry of SIĘ, is defined – its second conjunct is a disjunction of two template calls: (41) and (42). The latter, (42), introduces the exclusively impersonal SIĘ, which fails to satisfy the requirements of POTYKAĆ ‘stumble’ in (65). By contrast, the former, (41), introduces lexically required SIĘ (inherent, reflexive or reciprocal), which can also optionally be impersonal (as a result of multifunctionality allowed in (41)), but it still fails to satisfy the requirements of WCHODZIĆ ‘climb’ as it does not take any type of lexically required SIĘ – it is exclusively impersonal in (65).

One possible problem for the proposed analysis is, however, how to restrict structure-sharing of the shared dependent such as SIĘ to relevant predicates in examples such as the following one, where only some conjuncts require SIĘ:

(66) Bał się, płakał i śmiali.
feared.SG.M1 INH cried.SG.M1 and laughed.SG.M1
‘He was afraid, he cried and he laughed.’

In (66) the first (BAĆ) and the last (ŚMIAĆ) conjunct require the inherent SIĘ, while PŁAKAĆ ‘cry’ has no such requirement – it cannot therefore satisfy the constraining equations defined in (43). Currently the only alternative is (42), the template introducing impersonal SIĘ, but this would not only distort the semantics of the relevant predicates as SIĘ would be impersonal in all conjuncts, but it would also conflict with the inherent SIĘ specification introduced lexically using (45) by the edge conjuncts (BAĆ and ŚMIAĆ), as (42) requires SIĘ to be only impersonal and blocks multifunctionality (by negating SIĘ-LEX).

7 Problematic issues: split analysis

The following data concerning the licensing of secondary predicates, in particular the acceptability of (69), is problematic for the uniform analysis of SIĘ as a marker:

---

14Two examples from Kupšć 2000, (3.127a)–(3.127b) were merged into one in (65). The glosses and the free translation were modified.
(67) Ja jeszcze nigdy nie widziałem siebie pijaneGO.
   I yet never NEG saw self.GEN drunk.GEN
   ‘I have never seen myself drunk.’ (Google)

(68) Ja jeszcze nigdy nie widziałem siebie pijanyM.
   I yet never NEG saw self.GEN drunk.INST

(69) Ja jeszcze nigdy nie widziałem się pijaneGO/pijanyM.
   I yet never NEG saw REFL drunk.GEN/INST

In (67)–(68) SIEBIE, corresponding to the object marked for structural genitive case (because sentential negation is present), licenses the secondary predicate PIJANY ‘drunk’, which may either agree in case (genitive) or appear in the non-agreeing instrumental. On the analysis of SIE as a marker, this should not be possible when SIE is used instead of SIEBIE, as in (69). However, this example is – perhaps somewhat marginally – acceptable, and multiple acceptable examples of this kind, with the instrumental secondary predicate, may be found in corpora, e.g.:

(70) Widział się gołym.
    saw.M1 REFL naked.INST.M1
    ‘He saw himself naked.’ (NKJP)

If only the instrumental version were acceptable, then perhaps such examples could be explained away by claiming that the secondary predicate refers to the nominative subject. However, attested examples of secondary predicates apparently agreeing in case with SIE may also be found, even if they are much rarer and a little marginal:

(71) Nie widzę się zmęczonej, w nowej sytuacji, wśród tabunów babć, cioć i tP.
    NEG see REFL tired.GEN in new situation among hordes grandmas aunts etc
    ‘I can’t see myself tired, in a new situation, among hordes of grandmas and aunts, etc.’ (NKJP)

Such cases are easy to analyse if SIE is treated as a case-bearing argument, but not when it is always claimed to be a marker.

Further evidence against treating SIE uniformly as a marker comes from referential properties of this element in verb chains:

(72) Piotrowicz, który z wykształcenia jest mechanikiem,
    Piotrowicz.NOM.SG.M1 who from education is mechanic
    Kazał się tytułować profesorem.
    ordered.SG.M1 REFL address.INF professor.INST
    ‘Piotrowicz, who is a car mechanic by profession, ordered people to address him as a professor.’ (NKJP)

In (72) SIE clearly refers to Piotrowicz, which is the subject of the main verb, KAZAĆ ‘order’. Because KAZAĆ is an object control verb, its dative object (implicit in (72)) is, as a result of structure sharing, at the same time the subject of the

embedded infinitival clause headed by TYTUŁOWAĆ ‘address’. If, as proposed in § 6, SIĘ is treated as a marker and not as an argument, the predicate TYTUŁOWAĆ can only refer to its own subject, which would block the intended reading of (72). In this example SIĘ behaves in the same way as SIEBIE, which, according to Polish binding rules, typically\(^{16}\) refers to the subject, but it does not have to be the subject of the predicate local to SIEBIE – it may also be the subject from a higher verb in the verb chain. This ambiguity is illustrated below:

(73) Jani kazał Piotrowi, kupić sobie \(\ldots \) książkę.

Jan.NOM ordered Piotr.DAT buy.INF self.DAT book.ACC

‘Jan ordered Piotr to buy himself a book.’

(Przepiórkowski et al. 2002, p. 178, ex. (6.11a))

The f-structure corresponding to (73)\(^{17}\) is provided in (74). As indicated in glosses, (73) has two readings: one where SIEBIE is bound by the local subject, structure shared with the matrix dative object (Piotrowi), and another one, where it is bound by the matrix subject (Jan).

(75) is similar to (73) in that it also has two readings, depending on which argument binds SIĘ: the intended one where it is the father (the matrix subject) who is to be shaved and another one, somewhat unnatural because the hairdresser is specified for feminine gender, where the hairdresser is ordered to shave herself (the subject of the embedded predicate, local to SIĘ as a result of structure sharing with the matrix object marked for dative case).

(75) Ojciec kazał fryzjerce się ogolić.

father.NOM.SG.M1 ordered hairdresser.DAT.F REFL shave.INF

‘Father ordered the hairdresser to shave himself/herself.’

Since the proposed analysis adopts the marker analysis of SIĘ (see § 6), it produces the f-structure in (76) as the representation of (75), which makes the intended reading where it is the father who is to be shaved unavailable, leaving the alternative unintended reading where SIĘ is bound by the local subject, the hairdresser.

---

\(^{16}\)SIEBIE is bound by the subject unless it is used with a reciprocal predicate such as in the example below where it is bound by the object:

(i) Przedstawił sobie (nawzajem) sąsiadów.

introduced.SG.M1 self.DAT reciprocallly neighbour.ACC.PL

‘He introduced the neighbours to each other.’

\(^{17}\)The glosses and free translation in (73) were added for the purposes of this paper.
This problem cannot be solved by adopting a clause union analysis of verb chains in Polish, as this would cause problems for the treatment of negation-related phenomena in this environment, which includes handling of genitive of negation and licensing of n-words discussed in Patejuk and Przepiórkowski 2014 in the context of passive. It seems that the only alternative would be to treat reflective (and possibly reciprocal) SIE as also having an incarnation as an argument, which would lead to a split analysis of SIE. This, however, requires further research.

8 Conclusion

This paper offered an LFG analysis of SIE – the Polish so-called reflexive marker – showing its different uses, providing an account of a number of related phenomena, which include multifunctionality of SIE, where it has more than one use with respect to the same predicate, haplology of SIE in verb chains and finally structure sharing of SIE under coordination. The proposed analysis assumes that SIE is a marker rather than an argument, but binding facts in verb chains suggest that its status requires further research, which might lead to a split analysis of SIE.

Acknowledgements

We are grateful to both reviewers for their comments, which led to various improvements in the form and content of this paper.

Work reported here has been partially financed by the Polish Ministry of Science and Higher Education within the CLARIN ERIC programme 2015–2016 (http://clarin.eu/).

References


