Position of Polish Clitics: An HPSG Approach

Anna Kupść

Polish Academy of Sciences and Eberhard–Karls–Universität Tübingen

Abstract

The aim of this paper is to provide an HPSG analysis of linear positions of Polish pronominal clitics.[1] The distribution of Polish clitics, though rather free, is more constrained than that of other NPs. Polish clitics either occur preverbally or must closely follow the verb. We treat all Polish pronominal clitics as syntactic items and propose an LP constraint to account for their distribution. This constraint uniformly accounts both for preverbal and postverbal positions of clitics. Moreover, it captures the adjacency contrast between preverbal and postverbal positions of clitics. We briefly compare our approach with two possible alternatives: an analysis based on the lexical weight, cf. Abeillé and Godard (1998), and on the topological fields of Kathol (1995). The former might turn out to be too general if *lite* items in Polish are not restricted to clitics only. The latter, on the other hand, seems too restrictive for a 'free' word order language like Polish.

1 Polish Pronominal Clitics as Syntactic Items

Polish pronouns may occur in two forms: weak and strong. The two forms can be morphologically distinct, cf. (1), but most weak pronominals in Polish double as strong forms, e.g., *nas* 'us' or *im* 'them', cf. Rappaport (1988) and Spencer (1991, ch.9). In order to avoid confusion, we exemplify the behaviour of clitics only with morphologically distinct forms throughout the paper.

(1) a. *Jan go / jego zobaczył.*
        John him_{weak} him_{strong} saw
        'John saw him.'

    b. *Chętnie mu / jemu pomagamy.*
        willingly him_{weak} him_{strong} help-we
        'We help him willingly.'

Weak pronouns are clitics, i.e., they are prosodically deficient items. Unlike strong forms, they do not bear heavy stress and cannot be used in isolation.

(2) a. *Jan zobaczył JEGO / *GO, a nie Piotra.*
        John saw him_{strong} him_{weak} and not Peter
        'John saw HIM, not Peter.'

b. *Komu pomogliście? Jemu   / *Mu.*
    whom helped-you   him$_{strong}$ him$_{weak}$
    'Whom did you help? Him.'

They share also syntactic properties with weak forms discussed in Cardinaletti and Starke (1994). Unlike strong forms, weak pronouns cannot be coordinated, (3a), and can never be modified, (3b).

(3) a. *Zaprosiliśmy Piotra i   jego   / *go.*
         invited-we   Peter and him$_{strong}$ him$_{weak}$
         'We have invited Peter and him.'

     b. *Zaprosiliśmy tylko jego   / *go.*
         invited-we   only him$_{strong}$ him$_{weak}$
         'We have invited only him.'

The grammatical status of clitics is usually difficult to establish since they have both affix-like (morphological) and word-like (syntactic) properties, cf. e.g., Zwicky (1985), Zwicky and Pullum (1983), Miller (1992) for various tests to distinguish properties of morphological and syntactic items. We treat all Polish pronominal clitics as syntactic items[2] (postlexical clitics (PLC) in the terminology of Miller (1992)).

Unlike affixes, Polish pronominal clitics do not form a prosodic lexical unit with the host, cf. Rappaport (1988).[3]

(4) a. *POdaj Piotrowi książkę!*   'Give Peter the book!'
         give Peter   book

     b. *POdaj mu   ją!*   'Give it to him!'
         give him$_{weak}$ her$_{weak}$

     c. *PoDAję Piotrowi książkę.*   'I give Peter the book.'
         give-I Peter   book

If (a sequence of) clitics follow(s) the word, (4b), the lexical (penultimate) stress of this word (a host) is not affected, similarly to (4a). However, if an inflectional affix is attached to a word, the lexical stress can be shifted, see (4c).

Polish pronominals can be elided, cf. (5).

(5) A: *Przedstawisz   mi   go?   B: Przedstawię.*
        will-introduce-you me$_{weak}$ him$_{weak}$ will-introduce-I

        'Will you introduce him to me? (I) will introduce.'

---

[2]The reflexive clitic *się* 'self' exhibits so-called haplology, e.g., Fowler (1993), Rappaport (1997), Kupść (n.d.). A single occurrence of *się* can satisfy several requirements for this item. Cross-linguistically, haplology is usually conditioned morphologically and is explained on morphological grounds. We argue, however, in Kupść (n.d.) that *się* is a syntactic item and propose a syntactic analysis of its haplology.

[3]Capital letters mark lexical stress.

Such a property is characteristic for syntactic rather than morphological items, cf. Bresnan and Mchombo (1995).

Polish pronominal clitics do not have a fixed position with respect to a (lexical) verb, cf. positions of *go* in (1a) vs. (5) and *mu* in (1b) vs. (4c), in contrast to inflectional affixes which have a rigid position with respect to the stem.

In §2, we discuss in detail the distribution of Polish weak pronominals, which shows that their position in a phrase is not fixed, either.

2 Distribution of Polish Pronominal Clitics

The distribution of Polish pronominal clitics differs significantly from that in other languages, e.g. Spencer (1991, ch.9), Halpern (1995), Franks (1998). Unlike e.g., in Romance and Bulgarian, Polish pronominal clitics need not be adjacent to the verb, cf. (6a–b).

(6) a. *Dziewczyna mu   cierpliwie wszystko wytłumaczyła.*
         girl him_{weak} patiently   everything explained
         'The girl patiently explained him everything.'

   b. *Dziewczyna cierpliwie mu wszystko wytłumaczyła.*

   c. *Dziewczyna cierpliwie wszystko mu wytłumaczyła.*

In contrast to, e.g., Serbian/Croatian or Czech, cf. Halpern (1995), Polish clitics do not have a rigid position in a sentence. As (6) shows, a clitic may 'float' and be attached to a preverbal item[4] (Rappaport 1988). Note that phrase internal positions are also available, (7). In Polish, this is not a unique property of clitics. As observed in Borsley and Rivero (1994), Polish NPs and PPs can be split by other external elements as well. We come back to this issue in §3.

(7) *[Żadnego mu   zadania]$_{NP}$ nie wytłumaczyłaś.*
     none him_{weak} exercise   not explained-you

   'You didn't explain him any exercise.'

As repeatedly noticed, e.g., Kubiński (1982), Spencer (1991, ch.9), a pronominal clitic cannot occur in the sentence initial position, (8).

(8) * *Go   / się   Jan zobaczył w lustrze.*
       him_{weak} / $self_{weak}$ John saw   in mirror

   'John has seen him/himself in the mirror.'

Moreover, a clitic must remain in a clause, i.e., a clause boundary constitutes a barrier for clitic placement, e.g. Dyła (1983), Dziwirek (1994), Rappaport (1997).

---

[4]Clitics cannot occur between the negative marker *nie* 'not' and the verb in Polish, cf. Spencer (1991, ch.9). This fact can be immediately explained if *nie*, unlike pronominal clitics, is a verbal prefix, see Kupść and Przepiórkowski (1999). Generally, a clitic cannot be a complement of a preposition, either. This, however, seems to be a more general property of weak forms, see Cardinaletti and Starke (1994).

(9) a. *Powiedział nam, że   go   / się   widzi   w lustrze.*
         told-he   us   COMP him$_{weak}$ / self$_{weak}$ sees-he in mirror
         'He told us that he saw him/himself in the mirror.'

     b. * *Powiedział nam go   / się,   że   widzi   w lustrze.*
         told-he   us   him$_{weak}$ / self$_{weak}$ COMP sees-he in mirror

The contrast between the ungrammatical (8) and the correct (9a) shows that Polish pronominal clitics are enclitics, i.e., they require a preceding phonological host to attach to.

Certain postverbal positions of clitics are possible as well, e.g., clitics may occur immediately after the verb, (10).

(10) *Dziewczyna cierpliwie wytłumaczyła mu   wszystko.*
      girl   patiently   explained   him$_{weak}$ everything

      'The girl patiently explained him everything.'

If postverbal, a clitic (unlike other NPs) can be separated from the verb only by another clitic, (11a). Neither a strong pronoun, cf. (11b), nor a phrase, (11c), can be used instead.

(11) a. *Jan przyjrzał mu   się   dokładnie.*
         John looked   him$_{weak}$ self$_{weak}$ carefully
         'John watched him carefully.'

     b. * *Jan przyjrzał jemu   się   dokładnie.*
         John looked   him$_{strong}$ self$_{weak}$ carefully

     c. * *Jan przyjrzał [nowemu koledze]$_{NP}$ się dokładnie.*
         John looked   new   colleague   self carefully
         'John watched a new colleague carefully.'

As observed in Rappaport (1988), if there are several pronominal clitics, either all clitics float, cf. (12), or none, (11a).[5]

(12) a. *Brat mi   się   podoba.*
         brother me$_{weak}$ self$_{weak}$ likes
         'I like the brother.'

     b. * *Brat mi podoba się.*

Unlike in many other languages, Polish weak pronominals need not form a cluster (Spencer 1991, ch.9; Franks 1998).

(13) *Teraz mu   takie książki się   nie podobają.*
      now him$_{weak}$ such books self$_{weak}$ not like

      'He doesn't like such books nowadays.'

---

[5]In fact, certain native speakers do not find (12b) fully ungrammatical, in contrast to our judgement. All informants, however, consider (12b) degraded with respect to (12a).

The distribution of Polish pronominal clitics has hardly ever been formalised. One exception is Rappaport (1997) (but see also Franks (1998) for a tentative proposal). Rappaport (1997) assumes that pronominal clitics are a spell-out of an abstract (morphological) feature and have a fixed position within some phrasal domain. He discusses in detail the behaviour of the reflexive clitic *się* 'self' and associates it with a post-onset position in some verbal projection, i.e., after (post) the first element (onset) in the syntactic verbal domain. For example, if the syntactic structure of (6) is represented as in (14), the clitic *mu* in (6a) is in the post-onset position of IP (a clause), while in (6b) and (6c) it is in the post-onset position of a VP.

(14) $[_{IP}$ *Dziewczyna mu* $[_{VP}$ *cierpliwie* $[_{VP}$ *wszystko*   *wytłumaczyła]]].*
      girl   him$_{weak}$ patiently   everything explained

     'The girl patiently explained him everything.'

As discussed in Kupść (n.d.), however, such a parametrization seems too rigid for the placement of the reflexive clitic. In particular, it is problematic if several clitics are present, e.g., (11a), (12a). The reflexive clitic does not occur immediately after the first element (onset) of the clause in (11a) or a lexical verb (head of VP) in (12a) as would be required by the 'onset' parameter. Moreover, Rappaport (1988) notes that there is a phonological word barrier between a pronominal clitic and its host. His observation is strongly supported by evidence from southern and western dialects of Polish. In these dialects, the last consonant of the word *brat* in (12a) is pronounced as [d], which is characteristic for inter-word rather than word-internal voicing. Hence, if the sequence *brat mi* does not form a phonological word, the first element of the domain (onset) in (12a) cannot be formulated in phonological terms, either.

The analysis of (4b), (11a) and (12a) could be rescued if we assumed that if there are several adjacent clitics, they form a cluster. Such a clitic cluster could be represented as a single feature. In (5) and (11a), this feature would be spelled out in the post-head of VP, as in the case of a single clitic, e.g., (10). Such a representation, however, seems redundant. Not only would each clitic be represented by an individual feature but, additionally, there would have to be a feature representing adjacent clitics.

Another problem is that adjacent clitics do not have typical properties of a cluster, e.g., they do not seem to be arbitrarily ordered. Although Rappaport (1988) claims that pronominal clitics occur in a fixed *dat<się<acc<gen* order, other orders are also acceptable, e.g., (15) vs. (12a).

(15) *Brat*   *się*   *mi*   *bardzo podoba.*
     brother self$_{weak}$ me$_{weak,dat}$ very   likes

     'I like the brother very much.'

Moreover, Rappaport (1988) does not treat adjacent clitics as a phonological lexical unit.

Finally, it is not clear how Rappaport's (1997) analysis could exclude examples such as (12b). Note that each of ('post-onset') positions in (12b) is licit if only

one clitic is present, cf. (6a), (6c) and (10). We cannot exclude (12b) simply by assuming that several clitics must form a cluster since this is not the case in Polish, cf. (13).

In our account, §4, clitics are not assigned any specific position in a syntactic tree and their distribution is regulated via a linear precedence constraint.

3   Polish Word Order

Before proceeding to the analysis of the placement of clitics, some general remarks on Polish word order are necessary.

Word order in Polish is relatively free. Unlike in Germanic and many Romance languages, the position of the subject is not fixed. In Polish, the subject need not be sentence initial, cf. (16b–e), and can occur inside a VP, cf. (16b) and (16e).

(16) a.   *[$_S$ Jan   [$_{VP}$ lubi Marysię]].*
          John$_{nom}$ likes Mary$_{acc}$
          'John likes Mary.'

    b.   *Lubi Jan Marysię.*

    c.   *Lubi Marysię Jan.*

    d.   *Marysię lubi Jan.*

    e.   *Marysię Jan lubi.*

    f.   *Jan Marysię lubi.*

What is relevant here, if a preverbal clitic is phonologically licensed, i.e., if there is a phonological host it can be attached to, the clitic may precede the subject, e.g., (17).

(17) *Codziennie mu   Marysia czyta książki.*
     everyday   him$_{weak}$ Mary$_{nom}$ reads books$_{acc}$

     'Mary reads him books every day.'

Hence, the ungrammaticality of (8) is not due to the fact that clitics cannot outrank the subject.

The order of VP elements is not fixed. Not only are dependents of a verb unordered among themselves but the position of the verb is not fixed, either. All 24 possible permutations of (18) are grammatical in Polish.

(18) *często Marysi   daje   kwiaty*   '(He) often gives flowers to Mary.'
     often Mary$_{dat}$ gives-he flowers$_{acc}$

Again, phonological licensing and general constraints on word order in Polish are the only restrictions for preverbal clitic positions:

(19) *kwiaty mu   często daje*   '(He) often gives him flowers.'
     flowers him$_{weak}$ often gives

As illustrated in (7), a clitic can occur inside an NP. This, however, is a more general property of Polish: NPs can be split by external elements, cf. Borsley and Rivero (1994).

(20) *Ewy   czytałeś książkę.*   'You've been reading Eve's book.'
     Eve's read-you book

Also PPs can be interleaved with external elements but this is more constrained. Borsley and Rivero (1994) show that a preposition forms an unseparable unit with the first element of its complement.[6] Hence, such PPs cannot be split right after a preposition, cf. (21).

(21) a.   *[$_{PP}$ Z   tym umówiła   mężczyzną] Piotra.*
                 with this set-she a date man   Peter
           'She set Peter a date with this man.'

     b. * *[$_{PP}$ Z   umówiła   [$_{NP}$ tym mężczyzną]] Piotra.*
                 with set-she a date   this man   Peter

As mentioned in fn. 4, prepositions generally cannot have a clitic complement[7] but possibilities of clitic insertion into a PP are analogous to those of other items, e.g., (22).
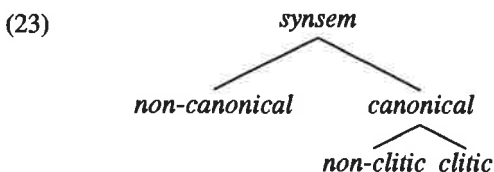
(22) *[$_{PP}$ Z   tym go   mężczyzną] umówiła.*
          with this him$_{weak}$ man   set-she a date

     'She set him a date with this man.'

In order to account for Polish data, we need a mechanism which orders not only syntactic constituents but also allows them to interleave. HPSG offers such an account, cf. Reape (1992) and Kathol (1995). In §4.2, we apply it to Polish.

4   HPSG Account

4.1   Representation of Polish Pronominal Clitics

As argued in §1, we treat Polish pronominal clitics as syntactic items rather than as affixes as in Romance, cf. Miller (1992), Miller and Sag (1997) for French and Monachesi (1995) for Italian. We distinguish clitics from other NPs by splitting *synsem* into subtypes. Since Polish clitics correspond to *signs*, we represent them via the type *clitic*, a subtype of *canonical (synsem)*, unlike in the analysis of French by Miller and Sag (1997), see (23).

(23)   *synsem*
                        /   \
              non-canonical   *canonical*
                                         /   \
                              non-clitic   clitic

---

[6] As shown in Rubach and Booij (1985), some prepositions form a prosodic word with the complement.
[7] The fixed expression *(samo) przez się* '(it) by itself' seems to be the only exception to this rule in Polish.

As in Romance, see Monachesi (1995, 1997) and Abeillé, Godard and Sag (1998), we assume that the reflexive clitic *się* is always introduced lexically, cf. Kupść (n.d.). We treat *się* as a syntactic item always present in the subcategorization frame of (lexical) verbs it occurs with. We use lexical rules to derive impersonal and middle verb forms, whereas inherent reflexive verbs have the reflexive clitic introduced directly in the lexical entry. Unlike in Romance, no lexical mechanisms are necessary to ensure the correct representation of other pronominal clitics. We represent all Polish weak pronominals (including *się*) on a par with syntactic complements, i.e., as elements of COMPS. We do not introduce a separate valence attribute for the reflexive clitic, unlike in our previous account in Kupść (n.d.). We constrain, however, the COMPS value to contain at most a single reflexive clitic.[8] (As argued in Kupść (n.d.), each verb may have at most one (syntactic) 'slot' for the reflexive clitic *się*.) Since the value of COMPS is specified as a list of *synsems*, both *clitics* and *non-clitics* can occur here. Being syntactic items, Polish pronominal clitics are removed from the COMPS list via general syntactic principles, i.e., the Valence Principle and ID schemata.

4.2   Ordering of Clitics

Facts presented in §3 show that in order to capture Polish data we need a more powerful mechanism than ordering of sisters in a local tree, i.e., LP statements of Pollard and Sag (1987). We use order domains of Reape (1992) and Kathol (1995) to account for Polish linear order facts.

Order domains allow one to encode linear order independently of the syntactic structure. We follow Kathol (1995) and assume that *domain_objects* (elements of the DOMAIN (DOM) list) have the following architecture:

$$(24) \quad \begin{bmatrix} dom\_obj \\ \text{PHON } phon \\ \text{SYNSEM } synsem \\ \cdots \end{bmatrix}$$

There are two basic operations which allow one to form order domains: shuffle (domain union), cf. Reape (1992), and (partial or total) compaction, cf. Kathol (1995). Roughly speaking, the former allows one to interleave elements from distinct order domains ('shuffle' them into a domain) while the latter (partly or totally) 'compacts' domains, i.e., treats them as unseparable units, see Reape (1992) and Kathol (1995) for details.

We assume that NPs and PPs[9] as well as VPs are not compacted phrases in a higher domain, i.e., they are not represented as a single domain object. Pronominal clitics are individual domain objects in a VP order domain. Clauses, however, are compacted phrases, i.e., they are represented as a single *dom_obj* in a higher order domain. The domain representation of (9a) is schematically illustrated in fig. 1.

---

[8] The reflexive clitic is represented as a *clitic:ana* element on the COMPS list.
[9] See restrictions mentioned in fn. 4 and §3.

$$\begin{bmatrix} \text{S} \\ \text{DOM } \langle \text{powiedział} \rangle, \langle \text{nam} \rangle, \langle \text{że się widzi} \rangle \end{bmatrix}$$

$$\begin{bmatrix} \text{V} \\ \text{DOM } \langle \text{powiedział} \rangle \end{bmatrix} \begin{bmatrix} \text{N} \\ \text{DOM } \langle \text{nam} \rangle \end{bmatrix} \begin{bmatrix} \text{S-BAR} \\ \text{DOM } \langle \text{że się widzi} \rangle \end{bmatrix}$$

$$\begin{bmatrix} \text{COMP} \\ \text{DOM } \langle \text{że} \rangle \end{bmatrix} \begin{bmatrix} \text{S} \\ \text{DOM } \langle \text{się} \rangle, \langle \text{widzi} \rangle \end{bmatrix}$$

$$\begin{bmatrix} \text{N} \\ \text{DOM } \langle \text{się} \rangle \end{bmatrix} \begin{bmatrix} \text{V} \\ \text{DOM } \langle \text{widzi} \rangle \end{bmatrix}$$
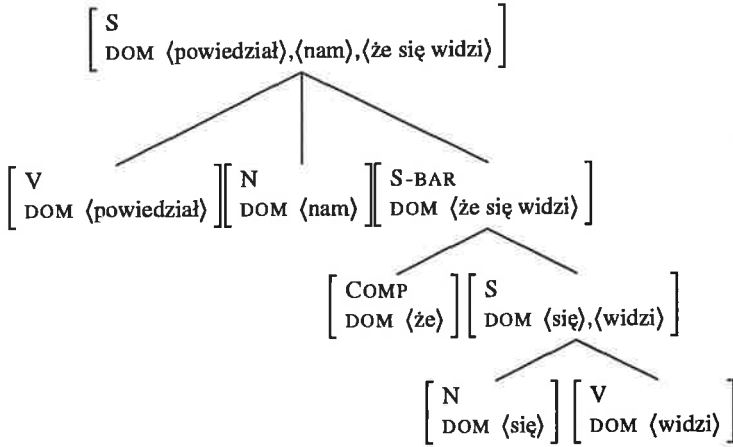
Figure 1: Domain representation of part of (9a)

Linear precedence principles are encoded as general well-formedness constraints on DOM values, cf. Kasper, Kathol and Pollard (1995):

$$(25) \quad \begin{bmatrix} \textit{sign} \\ \text{DOM } \boxed{1} \end{bmatrix} \rightarrow LP(\boxed{1})$$

If no constraints are given, all permutations of the DOM elements are possible.

The following implicational constraint on the DOM values uniformly accounts for preverbal and postverbal positions of clitics in a simple clause:[10]

$$(26) \quad \left(\begin{bmatrix} \textit{dom\_obj} \\ \text{SS|LOC|CAT} \begin{bmatrix} \text{HEAD } \textit{verb} \\ \text{VAL|COMPS } \boxed{1} \end{bmatrix} \end{bmatrix} \prec \begin{bmatrix} \textit{dom\_obj} \\ \text{SS } \boxed{2} \end{bmatrix} \prec \begin{bmatrix} \textit{dom\_obj} \\ \text{SS } \boxed{3} \textit{clitic} \end{bmatrix} \& \right.$$

$$\text{member}(\boxed{3}, \boxed{1}))$$

$$\rightarrow \left(\begin{bmatrix} \textit{dom\_obj} \\ \text{SS } \boxed{2} \textit{clitic} \end{bmatrix} \& \text{member}(\boxed{2}, \boxed{1}) \right)$$

This constraint should be understood as follows: for all DOM values (the antecedent of (25)), if it contains (at least) three elements x, y, z such that x is a verb, z is an argument clitic of this verb and x precedes y which precedes z, then y must also be the argument clitic of this verb (the consequent of (25)). Note that this constraint does not restrict the DOM list to contain three elements. The DOM list can be longer (or shorter) but no requirements on the order of other elements are specified here.

---

[10] A similar LP constraint has been previously proposed in Kupść (n.d.) to account for the distribution of the reflexive clitic.

(26) ensures that if postverbal, an argument clitic, i.e., $\boxed{3}$, can be separated from the verb only by another clitic, i.e., $\boxed{2}$ must be of type *clitic*. The 'member' relation ensures that clitics, i.e., $\boxed{2}$ and $\boxed{3}$, are arguments (elements of the COMPS list) of the verb. Since (26) is an implicational constraint, the antecedent may not be satisfied. This constraint does not force clitics to be postverbal. In particular, if there are no clitic arguments, the implication remains valid as well.

If we want to exclude (12b), we have to slightly modify (26). We must ensure that all clitics realised as arguments of a verb either occur preverbally or postverbally. We formalise this as follows:

$$(27) \quad \left(\begin{bmatrix} \textit{dom\_obj} \\ \text{SS } \boxed{1} \end{bmatrix} \prec \begin{bmatrix} \textit{dom\_obj} \\ \text{SS}|\text{L}|\text{C} \begin{bmatrix} \text{HEAD} & \text{verb} \\ \text{V}|\text{COMPS} & \boxed{4} \end{bmatrix} \end{bmatrix} \prec \begin{bmatrix} \textit{dom\_obj} \\ \text{SS } \boxed{2} \end{bmatrix} \prec \begin{bmatrix} \textit{dom\_obj} \\ \text{SS } \boxed{3} \textit{ clitic} \end{bmatrix}\right.$$

$$\left. \& \text{ member}(\boxed{3}, \boxed{4})\right)$$

$$\rightarrow \left(\begin{bmatrix} \textit{dom\_obj} \\ \text{SS } \boxed{1} \textit{ non-cl} \end{bmatrix} \& \begin{bmatrix} \textit{dom\_obj} \\ \text{SS } \boxed{2} \textit{ clitic} \end{bmatrix} \& \text{ member}(\boxed{2}, \boxed{4})\right)$$

According to (27), if one of the arguments, $\boxed{3}$, is realised as a postverbal clitic, clitic co-arguments, i.e., $\boxed{2}$, cannot occur preverbally. This is blocked since every element preceding the verb (with postverbal clitics) must be a *non-clitic*. Note that we do not require $\boxed{1}$ to be a complement of the verb. Hence, if we assume that the order domain of a clause is flat (all permutations are possible, §3), this constraint correctly rules out also (28) where the subject is outranked by an argument clitic. (In a clause, a clitic can precede the subject as well, cf. (17).)

(28)* *Codziennie mu   Marysia pożycza go   po   pracy.*
everyday   to-him$_{weak}$ Mary   lends   him$_{weak}$ after work

'Mary lends it to him everyday after office hours.'

As previously, a postverbal clitic can be separated from the verb only by another clitic, i.e., $\boxed{2}$ must be a *clitic*. Since (27) is an implicational constraint, it does not require all *non-clitics* to occur preverbally. In particular, (27) remains valid if there are no preverbal *non-clitics* at all, e.g., (5). Moreover, (27) does not require preverbal, unlike postverbal, clitics to cluster. Thus, (11a), (12a) and (13) are licensed.

Neither (26) nor (27) excludes (8). As the contrast with (9a) shows, the clause-initial position of clitics is not excluded in general. We assume here that (8) is ruled out by prosodic (or phonological) rather than linear or syntactic constraints. On the other hand, the ungrammaticality of (9b) is accounted for by pure syntax. The embedded clause in (9b) has an argument (clitic) unrealised whereas only fully saturated finite clauses are grammatical. There is no way to license the clitic as an argument of the main clause since in Polish clitics do not climb across a finite clause boundary.

4.3 Alternatives

4.3.1 Lexical Weight

Abeillé and Godard (1998) propose an account of word order in French based on the weight of syntactic items. Certain lexical syntactic items are 'lite' while others can be lexically specified as 'non-lite' or remain unspecified. 'Lite' items must precede 'non-lite' ones. Abeillé and Godard (1998) show that 'weight' is independent of lexical vs. phrasal distinction. They introduce a WEIGHT attribute appropriate for all *synsem*s. Although most phrases in French are *non-lite*, certain phrases can be *lite*.

How would this approach fit the ordering of Polish clitics? As pronominal clitics occur preverbally more freely, we could treat them as *lite* words, (lexical) verbs would be unspecified while other lexical items would be *non-lite*. Note, however, that we cannot retain the French *lite* < *non-lite* precedence rule for Polish. As (19) shows, *non-lite* complements may precede a *lite* clitic within a VP.

The easiest way of incorporating lexical weight into our analysis would be to reformulate (27) as follows:

$$
\begin{aligned}
(29) \quad (\begin{bmatrix} dom\_obj \\ \text{SS} | \text{WGT}\ \boxed{1} \end{bmatrix} &\prec \begin{bmatrix} dom\_obj \\ \text{SS} \begin{bmatrix} \text{L} | \text{C} \begin{bmatrix} \text{HEAD}\,verb \\ \text{V} | \text{COMPS}\ \boxed{4} \end{bmatrix} \end{bmatrix} \end{bmatrix} \prec \begin{bmatrix} dom\_obj \\ \text{SS}\ \boxed{2} \end{bmatrix} \prec \\
\begin{bmatrix} dom\_obj \\ \text{SS} | \text{WGT}\ \boxed{3}\ lite \end{bmatrix} &\ \&\ \text{member}(\begin{bmatrix} \text{WGT}\ \boxed{3} \end{bmatrix}, \boxed{4})) \\
\rightarrow (\begin{bmatrix} dom\_obj \\ \text{SS} | \text{WGT}\ \boxed{1}\ non\text{-}lite \end{bmatrix} &\ \&\ \begin{bmatrix} dom\_obj \\ \text{SS} | \text{WGT}\ \boxed{2}\ lite \end{bmatrix}\ \& \\
\text{member}(\begin{bmatrix} \text{WGT}\ \boxed{2} \end{bmatrix}, \boxed{4}))
\end{aligned}
$$

If clitics were the only *lite* elements in Polish, this constraint would be equivalent to (27). Since the assignment of the WEIGHT value in Polish is unclear to us at the moment, we retain the LP constraint for pronominal clitics in its more explicit formulation, i.e., (27).

4.3.2 Topological Fields

An appealing approach to linearization has been proposed in Penn (n.d.). His analysis derives from the idea of topological fields, cf. Kathol (1995), modified in order to include prosodic factors which constrain linear order in Serbian/Croatian.

It is clear that the distribution of Polish clitics cannot be purely syntactic. Prosodic factors involved in the placement of clitics in Serbian/Croatian and Polish seem to be quite similar. Polish clitics are enclitics, i.e., they require a phonological host to be licensed, (8). As in Serbian/Croatian, most Polish prepositions form a prosodic unit with the first element of its complement, cf. Rubach and Booij (1985). Clitics may occur PP-internally only after this unit. Due to such similarities, it would be desirable if the prosodic approach of Penn (n.d.) were carried

over to Polish. It seems, however, that a topological fields account is too rigid for Polish.

In order to account for the distribution of clitics in a clause, one may assume that the topology of a clause is organized around the verb.[11] Three basic fields could be distinguished: preverbal *pre-vf*, verbal *vf* and postverbal *post-vf*. Additionally, the embedded field *ef* can be used for verbal dependents (if any), cf. table 1.[12] On such an approach, *vf* could host a verb and clitics only. Addition-

	pre-vf	vf	post-vf	ef
(11a)	Jan	przyjrzał mu się	dokładnie	
(10)	Dz. cierpliwie	wytłumaczyła mu	wszystko	
(7)	żadnego mu zadania	nie wytłumaczyłaś		
(9a)		powiedział	nam	że go widzi w lustrze

Table 1: Assignment of topological fields in a clause

ally, clitics would be allowed in the preverbal field but would be excluded from the postverbal field. Nouns, adverbs and prepositions could appear in preverbal and postverbal fields.

As in Kathol (1995) and Penn (n.d.), the assignment of topological fields is indicated as a value of the additional TOPO attribute of a *dom_obj*. The assignment of topological fields can be lexical, e.g., as in the case of clitics or verbs, or syntactic, e.g., a *head-mark-ph* phrase would mark its singleton domain object as [TOPO *ef*], see the domain for S-BAR in fig. 1. Assuming that the order of topological fields in a clause is as indicated at the top of table 1, such an analysis would correctly account for all examples discussed in this paper but (12b) and (28). If we want to rule out sentences such as (12b) or (28), we must additionally assume that in a clause all pronominal clitics must have the same TOPO value, i.e., either *pre-vf* or *vf*.

A problem with this analysis is that not all verbal dependents can be treated on a par. In particular, we should distinguish verb clusters (VC, complementizerless infinitival complexes) from sentential dependents. The former, unlike the latter, allow much more flexible orderings. For example, an infinitive complement in VC can be preposed, (30a), whereas it is not possible if the complement is introduced by a complementizer, (30b).

(30) a. *Czytać Jan zaczyna książki.*   'John begins reading books.'
         read$_{inf}$ John starts   books

    b. * *Czytać Jan powiedział, żeby   książki.*
         read$_{inf}$ John said   COMP books
         'John told (us) to read books.'

---

[11] A similar idea has been expressed in Misz (1966) with respect to the position of the reflexive clitic *się*.
[12] As in Penn (n.d.), the internal topology of *ef* would be the same as that of the main clause.

Therefore, it seems that an infinitive and its complements in VC, should not be represented as a single domain object, unlike a subordinate clause in (9a) or (30b).

In fact, various permutations in VC are possible, cf. Przepiórkowski and Kupść (1997) for some examples. A natural assumption would be to represent all elements of VC as individual *dom_obj*s. We have to introduce additional topological markers in order to distinguish verbal fields in VC. As the ungrammaticality of (31) shows, a clitic cannot occur just in any verbal field.

(31)* *Boimy zacząć się   wybory.*
    fear-we star$_{inf}$ self$_{weak}$ elections

   'We are afraid to begin the elections.'

In order to account for (30a), one might introduce an additional topological field for preposed verbal dependents, e.g., *ppf*. In Polish, not only infinitives but also their complements can be preposed, cf. (32).

(32) *Czytać książki Jan   zaczyna.*
    read$_{inf}$ books John starts

It seems natural to assign *ppf* the internal topology analogous to that of the main clause. Then, where should the complement *książki* belong to? Should it occur in *post-vf* of the preposed field or rather in *pre-vf* of the main clause? The syntactic structure of Polish VC is unclear, cf. Przepiórkowski and Kupść (1997), and this issue cannot be resolved by syntax alone.

If possible at all, a topological fields account to Polish seems quite complex. On the other hand, as shown in Kupść (n.d.), a minor modification of an LP constraint for a reflexive clitic in a simple clause accounts for its positions in VC. Similarly, if we reformulate (27) to refer to realized clitics ([REAL+], cf. Przepiórkowski (n.d.)), such a LP constraint is able to account for the distribution of pronominal clitics in VC. Therefore, the analysis proposed in §4.2 seems both simpler and more accurate for Polish than a topological fields approach.

5   Conclusion

In this paper we presented an HPSG account of the placement of Polish pronominal clitics. The distribution of clitics in Polish is much less constrained than in other languages. Most of their positions follow from general principles of Polish linear order. We treat Polish pronominal clitics as syntactic items and use order domains to account for their distribution. We propose a single (implicational) LP constraint which uniformly accounts for all syntactic positions of clitics. In particular, this LP constraint allows us to capture the fact that postverbal, unlike preverbal, clitics must be adjacent to the verb. We briefly discuss also possible alternatives, i.e., an approach based on the lexical weight and topological fields. The former has to be at least temporarily dismissed since the weight assignment in Polish is not sufficiently investigated at present. The topological fields approach, on the other hand, appears unnecessarily complex and might impose too strong constraints on Polish word order.

This paper does not exhaust the subject. We only mentioned verb clusters and we did not deal at all with the distribution of clitics in other verbal environments, e.g., tense auxiliaries, participles or gerunds. Moreover, the relative order of clitics, in particular with respect to auxiliary clitics, has not been discussed. We leave these issues for future study.

References

Abeillé, A. and Godard, D.(1998), French Word Order and Lexicality, To appear in R. Borsley (ed.) *Syntactic Categories.*

Abeillé, A., Godard, D. and Sag, I. A.(1998), Two Kinds of Composition in French Complex Predicates, *in* E. Hinrichs, A. Kathol and T. Nakazawa (eds), *Complex Predicates in Nonderivational Syntax*, Vol. 30 of *Syntax and Semantics*, New York: Academic Press, pp. 1–41.

Borsley, R. D. and Rivero, M. L.(1994), Clitic Auxiliaries and Incorporation in Polish, **12**, 373–422.

Bresnan, J. and Mchombo, S.(1995), The Lexical Integrity Principle: Evidence from Bantu, **13**, 181–254.

Cardinaletti, A. and Starke, M.(1994), The Typology of Structural Deficiency on the Three Grammatical Classes, *University of Venice Working Papers in Linguistics* 4(2), 41–109.

Dyła, S.(1983), Evidence for S-deletion in Polish, *Folia Linguistica. Societas Linguistica Europaea* **XVII**, 327–337.

Dziwirek, K.(1994), *Polish Subjects*, Garland, New York & London.

Fowler, G.(1993), A syntactic account of derivational *-sja* in Russian, *in* R. Maguire and A. Timberlake (eds), *American Contributions to the Eleventh International Congress of Slavists*, Slavica Publishers, Columbus, Ohio, pp. 270–284.

Franks, S.(1998), Clitics in Slavic, Position paper at the Workshop "Comparative Slavic Morphosyntax", Spencer, Indiana, 5–7 June, 1998.

Halpern, A.(1995), *On the placement and morphology of clitics*, CSLI Publications, Stanford, California.

Kasper, R. T., Kathol, A. and Pollard, C.(1995), Linear Precedence Constraints and Reentrancy, *in* J. Kilbury and R. Wiese (eds), *Integrative Ansätze in der Computerlinguistik*, Düsseldorf, pp. 49–54.

Kathol, A.(1995), *Linearization-Based German Syntax.*, PhD dissertation, Ohio State University.

Kubiński, W.(1982), Polish *się* constructions and their English counterparts, **XV**, 55–65.

Kupść, A.(n.d.), Haplology of the Polish Reflexive Marker, pp. 91–124.

Kupść, A. and Przepiórkowski, A.(1999), Morphological Aspects of Verbal Negation in Polish, *Proceedings of the Second European Conference on Formal Description of Slavic Languages, Potsdam, Germany, November 20–22, 1997.* To appear.

Miller, P.(1992), *Clitics and Constituents in Phrase Structure Grammar*, Garland, New York.

Miller, P. H. and Sag, I. A.(1997), French Clitic Movement Without Clitics or Movement, **15**(3), 573–639.

Misz, H.(1966), Szyk *się* we współczesnej polszczyźnie pisanej, *Język Polski* **46**, 102–110.

Monachesi, P.(1995), *A Grammar of Italian Clitics*, PhD dissertation, Tilburg University, Tilburg. ITK Dissertation Series 1995-3 and TILDIL Dissertation Series 1995-3.

Monachesi, P.(1997), Italian Restructuring Verbs: A Lexical Analysis, *in* E. Hinrichs, T. Nakazawa and A. Kathol (eds), *Complex Predicates*. To appear.

Penn, G.(n.d.), Linearization and *wh*-Extraction in HPSG: Evidence from Serbo-Croatian, pp. 149–182.

Pollard, C. and Sag, I. A.(1987), *Information-Based Syntax and Semantics, Volume 1: Fundamentals*, CSLI Publications.

Przepiórkowski, A.(n.d.), On Case Assignment and "Adjuncts as Complements", pp. 231–245.

Przepiórkowski, A. and Kupść, A.(1997), Verbal Negation and Complex Predicate Formation in Polish, *Proceedings of the 1997 Texas Linguistics Society Conference on the Syntax and Semantics of Predication*, Vol. 38 of *Texas Linguistic Forum*, Austin, Texas, pp. 247–261.

Rappaport, G. C.(1988), On the relationship between prosodic and syntactic properties of pronouns in the Slavic languages, *in* A. M. Schenker (ed.), *American Contribution to the Tenth International Congress of Slavists*, Slavica Publishers, Columbus, Ohio, pp. 301–327.

Rappaport, G. C.(1997), Clitics as Features: A Non-semiotic Approach, To appear in: American Contributions to the International Congress of Slavists. Version of 7th March 1997.

Reape, M.(1992), *A formal theory of word order: A case study in West Germanic*, PhD dissertation.

Rubach, J. and Booij, G. E.(1985), A Grid Theory of Stress in Polish, *Lingua* **66**, 281–319.

Spencer, A.(1991), *Morphological Theory*, Oxford.

Zwicky, A. M.(1985), Clitics and Particles, *Language* **61**(2), 283–305.

Zwicky, A. M. and Pullum, G. K.(1983), Cliticization vs. Inflection: English *n't*, *Language* **59**(3), 502–513.