Proceedings of the 28th International Conference on Head-Driven Phrase Structure Grammar

Online (Frankfurt)

Stefan Müller, Nurit Melnik (Editors)

2021

Contents

Editor's note

The 27th International Conference on Head-Driven Phrase Structure Grammar (2020) was planned to take place in Leuven (organized by Frank Van Eynde and Liesbeth Augustinus), but due to the Corona pandemic it was organized online by Stefan Müller (Humboldt Universität zu Berlin) and Olga Zamaraeva (University of Washington, Seattle).

The conference featured 3 invited talks and 11 papers selected by the program committee (Anne Abeillé, Doug Arnold, Emily Bender, Felix Bildhauer, Hans Boas, Olivier Bonami, Francis Bond, Gosse Bouma, Antonio Branco, Rui Chaves, Philippa Cook, Berthold Crysmann, Dan Flickinger, Antske Fokkens, Petter Haugereid, Fabiola Henri, Thomas Hoffmann, Anke Holler (chair), Gianina Iordăchioaia, Paul Kay, Jong-Bok Kim, Jean-Pierre Koenig, David Lahm, Bob Levine, Nurit Melnik, Laura Michaelis, Philip Miller, Stefan Müller, Tsuneko Nakazawa, Petya Osenova, Rainer Osswald, Gerald Penn, Frank Richter, Louisa Sadler, Manfred Sailer, Pollet Samvellian, Jesse Tseng, Stephen Wechsler, Eun-Jung Yoo, Shûichi Yatabe).

We want to thank the program committee for putting this nice program together.

As in the past years the contributions to the conference proceedings are based on the five page abstract that was reviewed by the program committee, but there is no additional reviewing of the longer contribution to the proceedings. To ensure easy access and fast publication we have chosen an electronic format.

The proceedings include all the papers of the conference except the ones by Liesbeth Augustinus, Gosse Bouma, Frank Van Eynde & Jong-Bok Kim, Gert Webelhuth and Shûichi Yatabe.

An inside-out approach to French causatives

Gabriel Aguila-Multner
Université de Paris, Laboratoire de linguistique formelle, CNRS

Berthold Crysmann
Université de Paris, Laboratoire de linguistique formelle, CNRS

Proceedings of the 28th International Conference on
Head-Driven Phrase Structure Grammar

Online (Frankfurt)

Stefan Müller, Nurit Melnik (Editors)

2021

pages 5–25

Keywords: Clitic climbing, French, HPSG, causative, faire, periphrasis, VP structure

Aguila-Multner, Gabriel, & Crysmann, Berthold. 2021. An inside-out approach to French causatives. In Müller, Stefan, & Melnik, Nurit (Eds.), *Proceedings of the 28th International Conference on Head-Driven Phrase Structure Grammar, Online (Frankfurt)*, 5–25. DOI: 10.21248/hpsg.2021.1.

Abstract

In this paper[1], we provide a novel account of French causatives that crucially derives the core properties of the construction inside-out from the downstairs lexical verb to the causative verb, rather than outside-in, as is commonly assumed by argument composition (Miller & Sag, 1997; Abeillé et al., 1997; Abeillé et al., 1998). We shall argue on the basis of clitic trapping (Miller & Sag, 1997), as well as marking of the downstairs subject (Koenig, 1998) that the downstairs verb assumes a more active role than what is suggested by an argument composition approach and, conversely, we shall show that argument composition leads to problems with coordination and with en-cliticisation. The analysis we are going to propose combines an inversion analysis of the downstairs subject as a downstairs complement, accounting for scrambling and case marking, with an analysis of clitic climbing in terms of inflectional periphrasis (Aguila-Multner & Crysmann, 2020).

Clitic climbing, defined in Romance as the non-local realisation of clitic or affixal pronominal arguments, is limited in modern French to four cases: tense auxiliaries *avoir* and *être*, copular *être* and other predicative constructions, causative *faire* ("make") and *laisser* ("let"), and certain perception verbs like *voir* ("see"). Examples (1–4) illustrate the phenomenon in the four classes of constructions, respectively.

(1) a. Le chat l' a cassé.
the cat DO.SG have.PRS.3SG broken
'The cat broke it.'

b. Le chat y est allé.
the cat LOC be.PRS.3SG gone
'The cat went there.'

(2) a. Un chat leur sera donné.
a cat IO.PL be.FUT.3SG given
'A cat will be given to them.'

b. Le chat nous restera fidèle.
the cat 1PL remain.FUT.3SG loyal
'The cat will remain loyal to us.'

(3) a. Je le ferai manger au chat.
I DO.SG.M make.FUT.1SG eat to.the cat
'I will make the cat eat it.'

b. Je le laisserai manger au chat.
I DO.SG.M let.FUT.1SG eat to.the cat
'I will let the cat eat it.'

[1]We would like to thank the audience at HPSG 2020 for their comments and discussion, in particular Anne Abeillé, Olivier Bonami, Danièle Godard, Jean-Pierre Koenig, and Laura Michaelis. The research reported here has been supported by a doctoral grant from U Paris to Gabriel Aguila-Multner and also benefitted from a public grant overseen by the French National Research Agency (ANR) as part of the program "Investissements d'Avenir" (reference: ANR-10-LABX-0083). It contributes to the IdEx Université de Paris - ANR-18-IDEX-0001. Authors' names are listed in alphabetical order.

(4) Je l' ai vu casser par le chat.
I DO.SG have.PRS.3SG seen break by the cat
'I saw the cat break it.'

Within Head-driven Phrase Structure Grammar (HPSG, Pollard & Sag, 1994), French clitic climbing has been analysed (Abeillé & Godard, 2002; Abeillé et al., 1997; Abeillé et al., 1998) as a case of argument composition (Hinrichs & Nakazawa, 1990), i.e. generalised raising of the downstairs verb's arguments by the auxiliary. In a more recent proposal (Aguila-Multner & Crysmann, 2020), we suggested an alternative approach to clitic climbing, building on the model of inflectional periphrasis in HPSG by Bonami & Webelhuth (2013); Bonami (2015). However, this analysis so far only accounts for temporal, passive and predicative constructions. In this paper we extend the periphrasis approach to the causative construction.

We follow Abeillé et al. (1998) (after Hyman & Zimmer, 1976) in assuming two types of causative constructions in French, a generalisation summarised in the following section. The proposals based on argument composition by Abeillé & Godard (2002); Abeillé et al. (1998) are then presented in Section 2. Section 3 argues for an inside-out approach that gives more control to the downstairs verb, while a critical discussion of the argument composition approach is provided in Section 4. After an interim summary in Section 5, a new analysis based on periphrasis is laid out in Section 6.

1 Two types of French causatives

When *faire* is followed by an infinitive, it can give rise to two sorts of causative meanings. One takes the form of a three-place predicate assigning roles to a causer, a causee, and a caused event; the first corresponds to the subject of *faire*, the second to an object of *faire* co-indexed with the downstairs verb's subject in a control construction, and the third argument corresponds to the verb phrase. We call this type of *faire* "control *faire*". The other *faire* only assigns two roles: a causer and a caused event. This type of *faire*, which we call "non-control *faire*", contrasts with control *faire* in the semantic inferences it gives rise to: since a causee role is assigned by control *faire*, this kind of causation is generally interpreted as being direct, while non-control *faire* does not license such inferences (Abeillé et al., 1997, pp 66-67). This difference in semantics leads to verbs with experiencer subjects such as *aimer* (to like) being dispreferred in the control construction, as experiencers are not expected to have control over the caused event and are therefore incompatible with the causee role. This is illustrated in example (5), where the only compatible causative meaning is the non-control one (5b).

(5) a. # Faites-les aimer Proust !
make-DO.3PL love Proust
'Make them like Proust.'

b. Faites-leur aimer Proust !
make-IO.3PL love Proust
'Make them like Proust.'

Several syntactic properties correlate with this distinction. Control *faire* invariably realises the (controller of the) downstairs subject as an accusative pronominal affix, as opposed to a phrase:

(6) Je l' ai fait manger des épinards.
I DO.3SG have made eat INDEF.PL spinach
'I made him eat spinach.'

(7) * J'ai fait manger des épinards les enfants.

Furthermore, clitic climbing is impossible with control *faire*:

(8) Je l' ai fait en manger.
I DO.3SG have made DO.INDEF eat
'I have made him eat some.'

(9) * Je l'en ai fait manger.

In the non-control construction, however, the realisation of the downstairs subject varies according to the transitivity of the infinitive: transitives give rise to a dative pronominal or an NP[à], while intransitives lead to an accusative pronominal or a bare NP.

(10) J' ai fait manger des épinards aux enfants.
I have made eat INDEF.PL spinach to.the children
'I made the children eat spinach.'

(11) J' ai fait dormir les enfants.
I have made sleep the children
'I have made the children sleep.'

This construction does license clitic climbing to *faire* (subject to some constraints, cf. Section 3.1):

(12) Je lui en ai fait manger
I IO.3SG DO.INDEF have made eat
'I have made him eat some.'

Finally, the downstairs subject in the non-control construction displays a peculiar pattern of realisation: when realised pronominally, it is always attached to *faire*. In case of phrasal realisation, however, the downstairs subject may scramble with other downstairs complements (or adjuncts for that matter), as illustrated by the following example.

(13) a. J' ai fait manger aux enfants des épinards.
I have made eat to.the children INDEF.PL spinach
'I had the children eat spinach.'

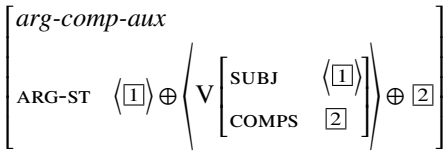

Figure 1: Argument composition

Thus, both marking and linearisation properties suggest that the logical subject of the downstairs verb enjoys the syntactic status of a non-subject complement.

2 Argument composition

Argument composition approaches to clitic climbing (Abeillé & Godard, 2002; Abeillé et al., 1997; Abeillé et al., 1998) rely on raising of the entirety of the downstairs verb's argument structure (and/or valence lists) by the upstairs verb. Arguments that are inherited in this way are naturally predicted to be hosted by the upstairs verb whenever they are affixal. In the case of causatives, non-control *faire* is then analysed as an argument composition verb. Figure 1 gives a schematic representation of such verbs.

Miller (1992) gives several arguments in defence of the flat structure of causatives. First, the free position of the downstairs subject relative to the complements of the infinitive is taken as evidence that the latter cannot form a VP with its complements alone. This does not rule out the possibility of a VP incorporating the downstairs subject, which we will explore in the analysis.

(14) Pierre a fait échanger les jouets aux enfants contre des livres.
Pierre has made exchange the toys to.the children against some books
'Pierre made the children swap the toys for books.' (Miller, 1992, 238)

Secondly, he draws an argument from the ungrammaticality of embedding of tense auxiliaries under a causative. He however admits that this ungrammaticality could be due to "some sort of independent semantico-pragmatic restriction" (p. 240 fn. 6), which is confirmed by the felicitous examples provided by Abeillé & Godard (1996, 38).

(15) a. Leur flair et leur ambition ont fait avoir fréquenté les
their intuition and their ambition have made have socialised.with the
gens qu' il fallait *(à) notre nouveau ministre et à sa femme.
people that EXPL had.to to our new minister and to his wife
'Their intuition and their ambition have made the new minister and his wife have been acquainted with the people that they needed to.'

b. La frugalité fait avoir vécu jusqu'à 110 ans (*à) notre fameuse
the frugality makes have lived until 110 years to our famous
concitoyenne, et la fera vivre encore longtemps.
copatriot and 3SG.ACC.F make live again a.long.time

'Frugality makes our famous copatriot have lived 110 years and will make her live an even longer time.'

(Abeillé & Godard, 1996, glossing and translation ours)

Thirdly, preverbal negation in the form of *ne pas* is impossible before the infinitive in the non-control case, which can easily be captured under the flat structure hypothesis, given that there is no infinitive VP for the negation to attach to. We return to this argument in the analysis.

(16) * Pierre fera ne pas rire Marie.
Pierre will make NEG not laugh Mary

(Miller, 1992, 240)

3 Restrictions imposed downstairs

As we have seen above, argument composition manages to reconcile climbing with a lexical perspective on pronominal affixation by means of giving the upstairs verb (*faire*) full control over the argument structure of the downstairs verb. In the *faire*-construction, however, there are still several cases where the downstairs verb maintains control over construction-specific aspects of realisation.

3.1 Trapping

With non-control *faire*, we typically observe climbing, i.e. upstairs realisation of all pronominal affixes of the downstairs verb. However, there are several exceptions: intrinsic arguments, medio-passive *se* and for most speakers even reflexive *se* resist climbing, as shown in (17).

(17) a. Le snobisme fait se vendre bien les classiques.
the snobism makes self sell well the classics
'Snobism makes the classics sell well.'

b. La chaleur a fait s'évanouir Paul.
the heat has made self.faint Paul
'The heat made Paul faint.'

c. (*) Marie a fait se laver les enfants.
Marie has made self wash the children
'Marie has made the children wash themselves.' (Abeillé et al., 1998, 24)

What is more, these intrinsic arguments also prevent any other pronominal affixes from being realised upstairs, with the exception of the downstairs subject.

(18) a. * Tout leur en fait vouloir à Paul.
everything IO.PL EN make angry to Paul
'Everything makes them/Paul angry at Paul/them.

b. Tout leur fait en vouloir à Paul.
everything IO.PL make EN angry to Paul
'Everything makes them angry at Paul.

c. Tout leur fait vous en vouloir.
everything IO.PL make 2.PL EN angry
'Everything makes them angry at you.'
(Miller & Sag, 1997, 609–610)

3.2 Subjects marked with *de/par*

Koenig (1998) notes another peculiarity of French *faire* construction which suggests that the downstairs verb plays a more active role with respect to argument realisation than what would be expected under an argument composition approach.

Agents of French passives can be expressed by either a *par*-phrase, or a *de*-phrase, the choice depending on the lexical aspect of the verb, i.e. whether it is dynamic (*par*) or stative *de*.

(19) Jean a été suivi *de / par Paul.
Jean has been followed of / by Paul
'Jean has been followed by Paul.'

(20) Le poisson a été suivi de / *par des rôtis.
the fish has been followed of / by INDEF.PL roasts
'The fish has been followed by a roast.'

In the *faire*-construction, realisation of the agent of the downstairs verb by an oblique by-phrase is equally possible, and we still observe sensitivity to the lexical aspect of the downstairs verb.

(21) Marc a fait suivre Jean *de / par Paul.
Marc has made follow Jean of / by Paul
'Marc had Jean followed by Paul.'

(22) Marc a fait suivre le poisson de / *par des rôtis.
Marc has made follow the fish of / by INDEF.PL roasts
'Marc had the fish be followed by a roast.'

With infinitives, however, realisation as a by-phrase is not a standard option. Koenig (1998) concludes that the grammatical function change must take effect on the downstairs verb, yet be conditioned inside-out on embedding in the causative construction.

4 Problems with argument composition

4.1 Controlling affixal realisation

The way argument composition is implemented in terms of structure sharing of ARG-ST lists, and therefore, structure sharing of the lists' elements, entails that any

constraint applied upstairs will also hold downstairs (and vice versa). If an upstairs *affixal-synsem* implies pronominal affixation, we would expect, *ceteris paribus*, that the same should hold downstairs. With auxiliary-participle constructions, this is a non-issue in French, since participles may not host pronominal affixes at all. French infinitives, however, can generally host pronominal affixes, so argument composition *per se* would predict affixal realisation to feature simultaneously on the upstairs and the downstairs verb. However, this expectation is not borne out, thereby weakening the appeal of argument composition.

Miller & Sag (1997, 609) work around the technical side of this problem by distinguishing the HEAD values of verbs into *bas(ic)-v(er)b* and *red(uced)-v(er)b*, where the former is the value for plain verbs without pronominal affixes, while the latter is the *default* value for verbs hosting pronominal affixes. This default is overridden with the value *bas-vb* in the case of verbs with intrinsic clitics (see section 3.1 on trapping), leading to the paradoxical situation that even the presence of regular, valence-reducing argument clitics does not imply the value *red-vb*. While the head types *bas-vb* and *red-vb* appear to be little more than diacritic features, their specific use in connection with trapping reveals their *ad hoc* nature.

4.2 Coordination

One key characteristic of argument composition is that gives it rise to a flat verb phrase structure that complicates the treatment of VP coordination: i.e. the lexical non-finite verb figures as a direct complement of *faire* and does not itself combine with its own complements to project a VP. Thus, what looks like a case of ordinary constituent coordination, as indicated by the bracketing in (23), must be analysed as a case of non-constituent coordination.

(23) a. Elle la leur a fait [apprendre par cœur] et [réciter le
she DO.SG.F IO.PL have.3SG.PRS made learn by heart and recite the
lendemain].
next.day
'She made them learn it by heart and recite it the next day.'

b. Elle a fait [lire Sartre par les garçons] et [réciter Prévert
she have.3SG.PRS made read Sartre by the boys and recite Prévert
aux filles].
to.the girls
'She made the boys read Sartre and the girls recite Prévert.'

Under a traditional layered VP structure non-finite VP coordination an analysis in terms of conventional VP coordination is possible, as has been pointed out for tense constructions already by Manning (1997) and Aguila-Multner & Crysmann (2020).

4.3 en-cliticisation

Another piece of evidence that challenges the argument composition approach is contributed by *en*-cliticisation in conjunction with trapping.

The relevant contrasts are given in (24) below: non-local *en*-cliticisation is subject to the same trapping effect as ordinary argument clitics.

(24) a. Je leur ai fait s' en rappeler la fin.
I IO.SG have.PRS.1SG make.PTCP REFL.3 EN remember.INF the end
'I have made them remember the end of it.'

b. * je leur en ai fait se rappeler la fin
I IO.SG EN have.PRS.1SG make.PTCP REFL.3 remember.INF the end

(25) Voici le roman dont je leur ai fait se rappeler la fin.
here's the novel OF.WHICH I IO.PL have made REFL.3 remember.INF the end.
'Here's the novel I made them remember the end of.'

With argument composition, the above contrast is actually quite surprising: as discussed by Miller & Sag (1997), *dont*-relativisation and *en*-cliticisation are non-local in that they refer to a *de*-NP that can be arbitrarily deeply embedded within a complement of the host. To capture this, they argue that *en*-cliticisation goes piggy-back on the unbounded dependency independently needed for *dont*-relativisation, and propose a lexical rule that inserts an affixal synsem to bind the *de*-NP SLASH value of the verb's *canonical* complement. Given argument composition, this lexical rule should be able to apply not only to the lexical verb, but also to *faire*, in which case upstairs realisation will be predicted where only downstairs trapping should be possible.

5 Summary

In the previous sections, we have observed that the downstairs verb plays a more prominent role in the French causative construction than an argument composition approach would suggest: most notably the realisation of the downstairs subject, i.e. whether it surfaces as a bare NP or an indirect object, is a property decided by the transitivity of the downstairs verb. Furthermore, as discussed by Koenig (1998), the choice between *par* and *de* as an alternate marking for the subject of a transitive is determined by the lexical aspect of the downstairs verb. As for clitic climbing, trapping also militates for a position that grants the downstairs verb more active control over the construction.

In the remainder of this paper, we shall present a novel approach to the grammar of French causatives that does away with argument composition and derives the core properties of the construction inside-out from the downstairs lexical verb. In essence we shall generalise the inside-out dependence of *par/de* marking on an embedding causative verb and suggest that realisation as a direct or indirect object is equally an instance of demotion of the downstairs subject valency to a complement.

This "inversion" approach shall prove capable of deriving a number of core facts of the construction without having to rely on argument composition: if the downstairs subject is demoted to COMPS, scrambling with other complements of the downstairs verb is expected, cf. (13). Similarly, indirect object marking with transitives can equally be motivated by a ban on double accusatives as a constraint on the COMPS list of the downstairs verb. Finally, the perspective of representing all arguments of the downstairs verb as its complements provides for a straightforward account of VP coordination, including mixed subject marking, as shown in (23b).

Turning to clitic climbing, we have suggested in recent work (Aguila-Multner & Crysmann, 2020) that climbing with tense auxiliaries is best understood as an instance of periphrastic inflection (Bonami, 2015), arguing more specifically that delegation of pronominal affixation to the auxiliary is parasitic on an existing morphological inside-out dependency, namely tense periphrasis. Here, we shall extend our approach and suggest that clitic climbing in causatives equally relies on an independently motivated inside-out dependency (Koenig, 1998).

The analysis we are going to propose improves over the argument composition approach also in the area of *en*-cliticisation: given that there is no argument composition, *en*, just as all other clitics, can only ever originate on the downstairs verb. With intrinsic clitics, *en* will then be trapped, while it can climb otherwise, the decision being ultimately made by the downstairs verb.

6 Analysis

We have seen in Section 3 that the downstairs verb in constructions with non-control *faire* exerts a significant amount of control on argument realisation, both in terms of the realisation of the downstairs subject and in terms of the possibility vs. impossibility of clitic climbing. Rather than using argument composition to make as much information as possible available to the causative verb, we shall build on the work on clitic climbing via periphrasis by Aguila-Multner & Crysmann (2020) and place the various constraints associated with this construction on the downstairs verb.

6.1 Clitic climbing as periphrastic morphology

In our analysis of clitic climbing in French tense constructions (Aguila-Multner & Crysmann, 2020), we built on Bonami (2015)'s theory of inflectional periphrasis to reduce clitic climbing to a case of periphrastic exponence. Bonami's theory relies on reverse selection, a form of inside-out constraint, to allow the lexical element in a periphrase to impose morphological constraints to the auxiliary that syntactically selects for it, effectively creating a dependency that can convey information output by the inflectional component, i.e. periphrastic exponence. Since pronominal clitics in French are best analysed as lexical affixes (Miller, 1992), their non-locality in tense periphrases with clitic climbing can be accounted for as a form of periphrastic exponence, reverse-selected for by the downstairs verb to the auxiliary; in other words,

realisation of pronominal arguments is just another property that is realised upstairs in a French tense auxiliary construction, along with TAM and subject agreement.

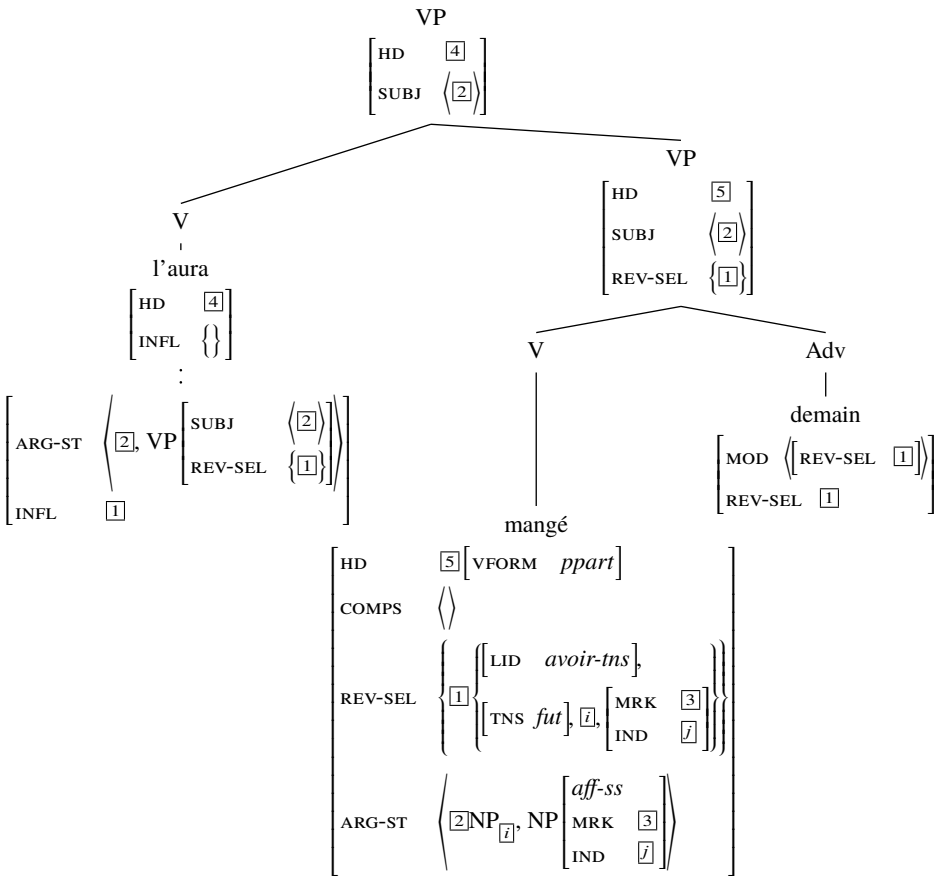

Figure 2: Percolation of periphrastic pronominal affixation

Figure 2 summarises the analysis of periphrastic realisation of pronominal affixes with a sample derivation of clitic climbing in periphrastic tenses. In essence, the lexical verb places its morphological requirements for the ancillary element in a feature REV(ERSE)-SEL(ECTION), which is passed up along the head projection path. The ancillary element, in this case a head governing the VP projection, lexically equates the requirements of its complement with its own INFL value, which defines the input for morphological realisation rules.

We define REV-SEL as a set-valued[2] feature located under CAT. The percolation mechanism of this feature is made explicit in Figure 3: ancillary lexemes subcategorise for an element with a non-empty REV-SEL set, one element of which they check against their INFL feature, while the rest is passed up to the ancillary lexeme's own REV-SEL. Non ancillary elements do not combine with elements carrying re-

[2]The REV-SEL feature was originally defined as list-valued by Aguila-Multner & Crysmann (2020), but we do not find any use for ordering of multiple reverse selection dependencies.

verse selection dependencies. Inheritance of REV-SEL by phrases proceeds simply from the head in head-valence phrases, and from both heads in coordinated phrases.

$$\textit{non-anc-hd-lex} \rightarrow \left[\begin{array}{ll} \textsc{ss} & \left[\begin{array}{ll} \textsc{loc} & \left[\textsc{cat.rev-sel} \quad \{\}\right] \\ \textsc{arg-st} & \left\langle \left[\textsc{loc.cat.rev-sel} \{\}\right] \ldots \left[\textsc{loc.cat.rev-sel} \{\}\right] \right\rangle \end{array}\right] \end{array}\right]$$

(a) Non-ancillary head

$$\textit{anc-hd-lex} \rightarrow \left[\begin{array}{ll} \textsc{ss} & \left[\begin{array}{ll} \textsc{loc} & \left[\textsc{cat.rev-sel} \quad \boxed{2}\right] \\ \textsc{arg-st} & \left\langle \ldots \left[\textsc{loc.cat.rev-sel} \{\boxed{1}\} \cup \boxed{2}\right] \ldots \right\rangle \end{array}\right] \\ \textsc{infl} & \boxed{1} \cup \textit{set} \end{array}\right]$$

(b) Ancillary head

Figure 3: Constraints on saturation of REV-SEL

6.2 Realisation of pronominal affixes

As stated in Aguila-Multner & Crysmann (2020), we assume argument mapping rules that type elements of ARG-ST with one of three synsem types (*canon-ss, gap-ss, praf-ss*) and insert them to the relevant features accordingly: canonical elements are left on valence lists, gap elements are tied to non-local features, and most relevantly here pronominal affixes are added to the inflectional agenda INFL as structures of type *praf*, containing case/marking and an index value. This is illustrated in Figure 4.

$$\left[\begin{array}{ll} \textsc{comps} & \boxed{2}\ \textit{list(canon)} \\ \textsc{infl} & \boxed{3} \cup \left\{\left[\begin{array}{lll} \textit{praf} & & \\ \textsc{mrk} & \boxed{m_1} & \\ \textsc{ind} & \boxed{i_1} & \end{array}\right] \ldots \left[\begin{array}{lll} \textit{praf} & & \\ \textsc{mrk} & \boxed{m_n} & \\ \textsc{ind} & \boxed{i_n} & \end{array}\right]\right\} \\ \textsc{dtr} & \left[\begin{array}{ll} \textsc{arg-st} & \left\langle \boxed{1}, \left[\begin{array}{lll} \textit{aff-ss} & & \\ \textsc{head}|\textsc{mrk} & \boxed{m_1} & \\ \textsc{cont}|\textsc{ind} & \boxed{i_1} & \end{array}\right] \ldots \left[\begin{array}{lll} \textit{aff-ss} & & \\ \textsc{head}|\textsc{mrk} & \boxed{m_n} & \\ \textsc{cont}|\textsc{ind} & \boxed{i_n} & \end{array}\right] \right\rangle \circ \textit{list(gap)} \circ \boxed{2} \\ \textsc{infl} & \boxed{3} \end{array}\right] \end{array}\right]$$

Figure 4: Mapping of pronominal arguments

Our implementation of the inflectional component is a set-valued feature INFL that acts as an agenda of morphosyntactic properties to be realised; realisation rules (synthetic and periphrastic) empty its contents and an empty INFL set is a requirement for entering syntax. A derivation for a simple tensed verb with local pronominal affixation is given in Figure 5 as an illustration of this morphology-syntax inter-

face: the verbal lexeme undergoes the mapping rule which adds a *praf* to its INFL, and inflectional rules symbolised by the dotted line realise it (along with TAM and agreement properties) accordingly with the form *les mangera*. Such rules can realise properties inherited by an ancillary element from their complement's REV-SEL, since their inheritance is mediated by INFL, as illustrated by the pronominalisation rule that applies to *l'aura* in Figure 2.

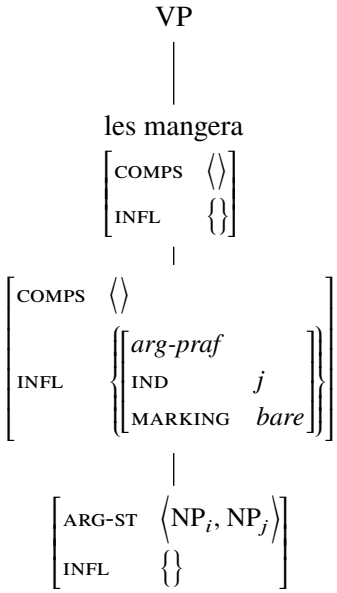

Figure 5: Synthetic pronominalisation

6.3 Realisation of the downstairs subject

All that is required now in order to model clitic climbing in causatives is a lexical rule to introduce the reverse-selection for *faire*. As argued by Koenig (1998), causativised infinitives independently need to carry an inside-out constraint for their licensor (causative *faire*) to properly constrain realisation of their subject as a *par* or *de*-phrase. I.e. the downstairs verb's subject is demoted to an oblique complement, contingent on the embedding under the causative verb. Our analysis goes piggyback on this independently required inside-out dependency (Koenig actually assumes argument composition together with a flat structure of VP): on the one side, we shall generalise realisation of the downstairs subject as an oblique complement to the case of realisation by an indirect object (transitives) or a direct object (intransitives). On the other hand, we shall argue that if there is already an inside-out dependency on a causative predicate, an analysis of clitic climbing as periphrasis will come at little extra cost. This is highly similar to the case of tense auxiliaries (Aguila-Multner & Crysmann, 2020) where periphrastic realisation of pronominal affixation depends on an already existing periphrastic relation between the participle and the auxiliary

for the expression of tense.

The relevant lexical rule is given in Figure 6, using the feature LID for identification of the causative verb. Generalising the case of oblique by-phrase realisation of the downstairs subject to direct and indirect objects, we suggest to extend the COMPS list of the downstairs verb with an NP co-indexed with the first element of ARG-ST. I.e. we essentially entertain an inversion analysis for downstairs NP subjects. Subtypes of this rule select the appropriate marking value on this inverted NP, sensitive to the argument structure and/or lexical semantics of the verb.

This rule only creates causative infinitive verbal lexemes, and given that French lacks a synthetic way of realising causative voice, these lexemes need a periphrasis rule to delegate their morphosyntactic properties to the relevant ancillary element (*faire*). As given in Figure 7, this rule not only delegates the realisation of causative voice, but also delegates expression of any *praf* specifications.

Finally, an entry for the causative verb is given in Figure 8. As was the case with *avoir*, *faire* inherits part of its inflection from its verbal complement's REV-SEL set, including any pronominal affixes delegated by the periphrasis rule.

The tree in Figure 9 summarises the analysis in the simple case of an intransitive verb (*dormir*) with an affixal subject. The one in Figure 10 features clitic climbing of the downstairs object.

$$\begin{bmatrix} \textsc{head} & \left[\textsc{vform} \quad \textit{nonfinite}\right] \\ \textsc{infl} & \left\{\left[\textsc{lid} \quad \textit{faire-lid}\right]\right\} \\ \textsc{subj} & \langle\rangle \\ \textsc{comps} & \boxed{1} \oplus \left\langle \mathrm{NP}_i \left[\textsc{marking} \quad \textit{bare} \vee \textit{à} \vee \textit{par} \vee \textit{de}\right]\right\rangle \\ \textsc{dtr} & \begin{bmatrix} \textsc{infl} & \{\} \\ \textsc{subj} & \langle \mathrm{NP}_i \rangle \\ \textsc{comps} & \boxed{1} \end{bmatrix} \end{bmatrix}$$

Figure 6: Lexical rule for causativised verbs

$$\begin{bmatrix} \textsc{head} & \left[\textsc{vform} \quad \textit{nonfinite}\right] \\ \textsc{rev-sel} & \left\{\boxed{1} \cup \boxed{2}\ \textit{set}\left(\left[\textit{praf}\right]\right)\right\} \cup \boxed{3} \\ \textsc{infl} & \{\} \\ \textsc{dtr} & \begin{bmatrix} \textsc{rev-sel} & \boxed{3} \\ \textsc{infl} & \boxed{1}\left\{\left[\textsc{lid} \quad \textit{faire-lid}\right]\right\} \cup \boxed{2} \end{bmatrix} \end{bmatrix}$$

Figure 7: Lexical rule for causative periphrasis

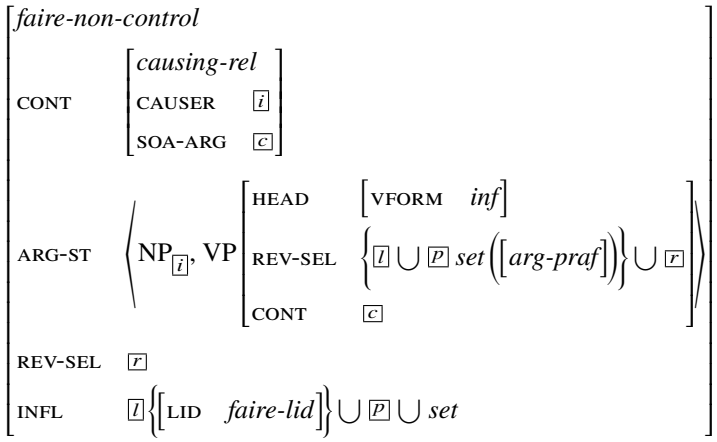

Figure 8: Non-control *faire*

6.4 Intrinsic arguments and trapping

Having laid out the basic line of analysis, we shall now turn to the treatment of trapping. As described in Section 3.1, trapping is triggered by lexically specified intrinsic arguments as well as inherent, medio-passive, and, for most speakers, reflexive *se*, so an important question is how such arguments are represented as part of the lexical entry of the verb. There are two basic observations regarding all these arguments: first, they are always realised affixally (cf. Abeillé et al., 1998), and second, intrinsic arguments, including inherent *se*, are not assigned a thematic role. This observation already carries over to medio-passive *se*, which is best understood as an exponent of grammatical function change (Grimshaw, 1982; Wehrli, 1986). Following Crysmann (2003), we shall therefore assume that intrinsic arguments and reflexives can be represented on ARG-ST as *aff-ss* objects whose CONT value is either *expl*, as is the case of intrinsic arguments, or else *refl*.

Given such an explicit representation of argument type, we shall always be able to detect the presence of intrinsic arguments and enforce their local realisation prior to the application of the causative lexical rule. This can be ensured by augmenting the description of non-control *faire* with a type constraint on the set of *praf* elements it may inherit as *arg(umental)-pr(onominal)af(fixes)*. This is exemplified in Figure 8. Figure 11 illustrates the derivation of a sentence with trapping of intrinsic *en* (*en vouloir* "to be angry with").

6.5 Interaction with tense auxiliaries

With at least two separate constructions (*faire* and *avoir/être*) entering a reverse selection dependency, the question arises what their possible combinations are and whether the analysis adequately generates them. A first combination is the possibility for *avoir* to embed a causative construction headed by *fait* (PTCP). In this case, any climbing from the downstairs infinitive to *faire* is simply further deferred to the

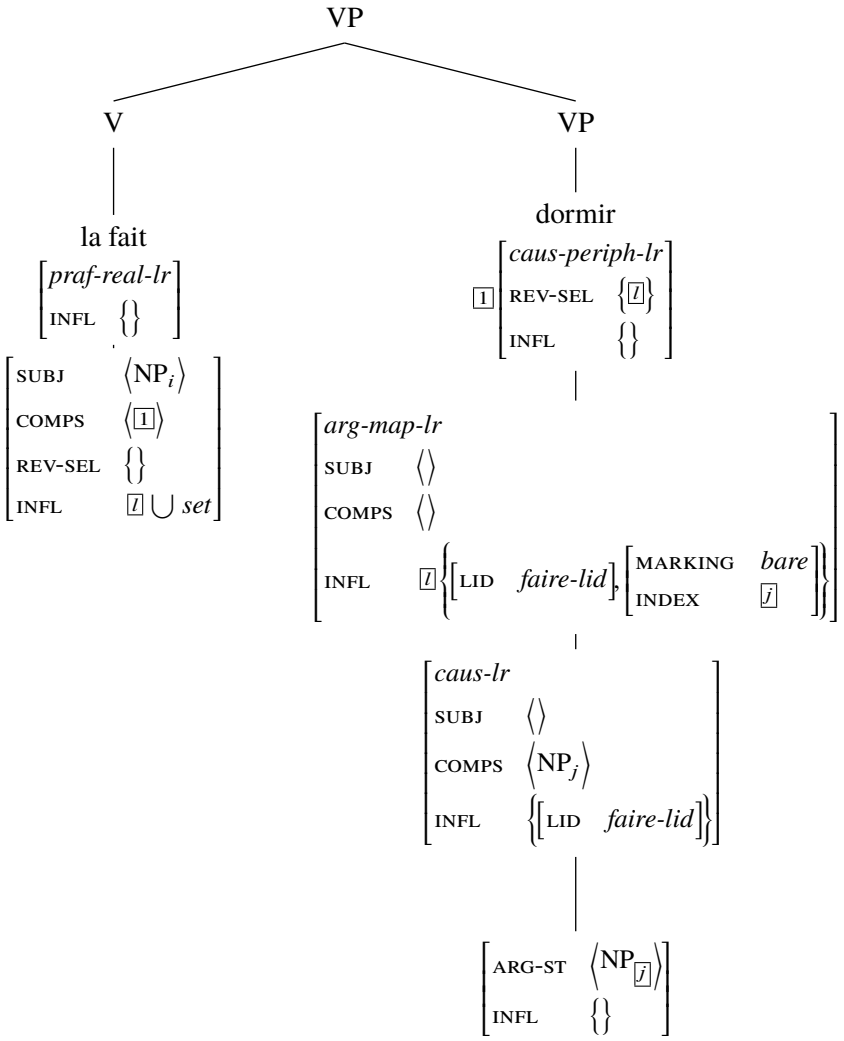

Figure 9: Sample derivation with affixal subject

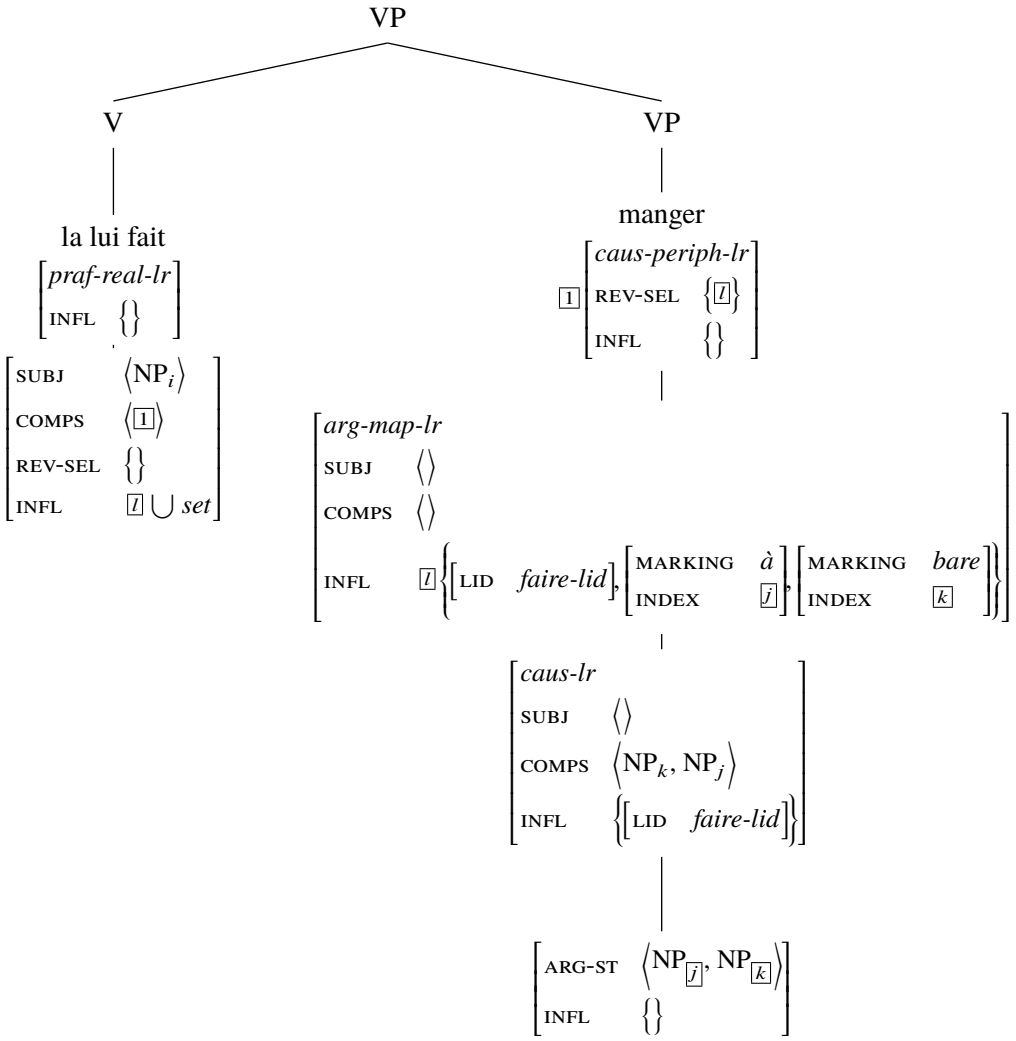

Figure 10: Sample derivation with affixal subject and climbing

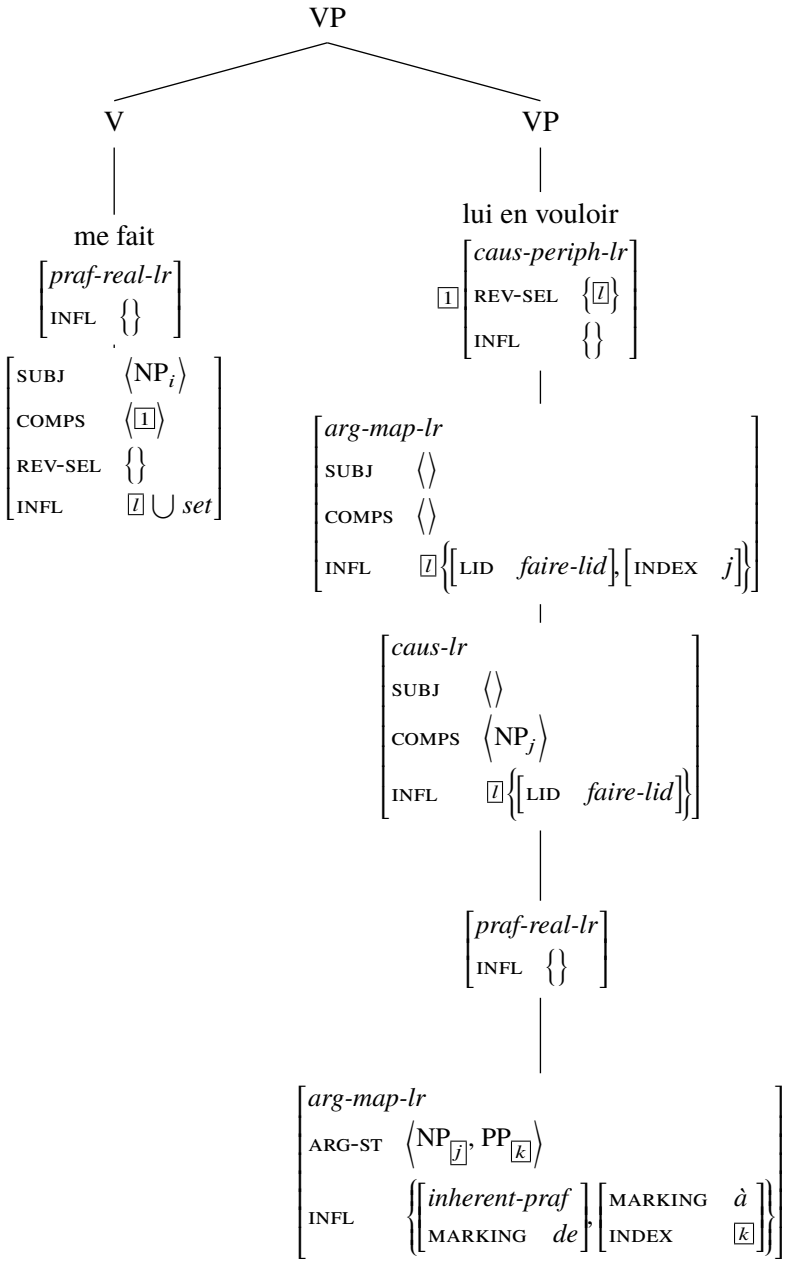

Figure 11: Sample derivation with affixal subject and trapping

tense auxiliaries, consistent with the rule of obligatory climbing from participles. Our approach readily captures this given the rules and lexical entries previously given: climbing from the infinitive is covered regardless of the particular form of *faire*, and the *praf* objects will be inherited by *fait*'s lexical entry on INFL. Participle periphrasis rules such as given in Aguila-Multner & Crysmann (2020) operate on a verb's INFL (and not e.g. directly on ARG-ST), and will appropriately create a new REV-SEL dependency with a tense auxiliary (in this case *avoir*) which contains all pronominal affixes inherited from the previous dependency.

More challenging is the second possible combination: a perfective periphrase can be embedded under *faire* (15). Again climbing of all affixes is obligatory from the participle to *avoir/être*, but further climbing to *faire* is limited to the downstairs subject in the presence of an intrinsic affix, following the trapping rule described in Section 3.1. Our approach as previously stated however suffices to produce the desired outcome, on the assumption that in the sequence of inflection rules aspect periphrasis precedes causativisation. This way the trapping case is covered by the early application of the mapping rule, after which all pronominal affixes will be inserted into the REV-SEL dependency by the aspect periphrasis rule, with the exception of the subject, which has not yet been inverted and is therefore not available to mapping. The affixal subject can only be mapped after the causativisation rule instead, and therefore after the aspect periphrasis rule; as a consequence its only possibility of realisation is to enter the REV-SEL dependency established by the causative periphrasis rule, which in the full climbing case will also contain all other pronominal affixes (Figure 7), effectively climbing from the participle to *faire* in one go.

Before closing, a remark is due concerning negation with non-control *faire*: as observed by Miller (1992), the downstairs infinitive cannot be modified by *ne pas*, unlike standard VPs. One way to capture this constraint is to ensure that negative modifiers cannot disrupt morphological periphrasis, e.g. by requiring that these modifiers select for a head whose REV-SEL value of the head is the empty set.

6.6 Control *faire*

To complement our analysis of French non-control *faire*, a brief remark is due to its counterpart, control *faire*: essentially, we shall follow Abeillé et al. (1998) in assuming that control *faire* is a standard object equi verb that assigns the thematic role of causee to its affixal direct object complement, the controller of the downstairs subject. Cf. Figure 12 for a sample lexical entry.

7 Conclusion

In this paper we have provided an analysis of clitic climbing in French causatives that is based on reverse selection from the downstairs infinitive to the causative verb. Building on Koenig (1998)'s argument for an inside-out view of such constructions

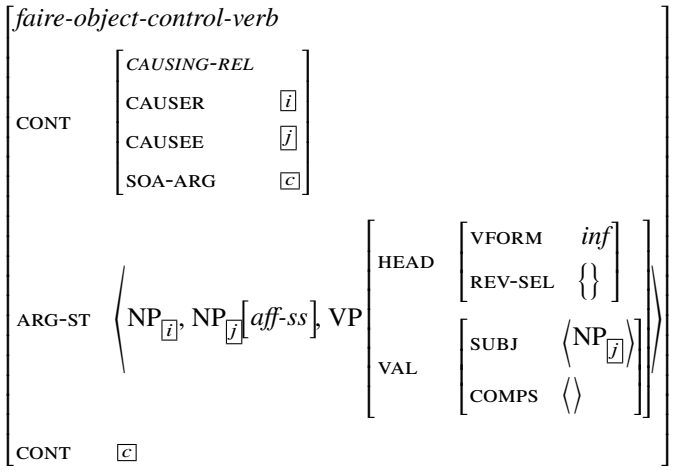

Figure 12: Control *faire*

and on Aguila-Multner & Crysmann (2020)'s proposal for clitic climbing by periphrasis with French tense auxiliaries and predicatives, this approach disposes with the need for argument composition and the concomitant flat structure of the VP. Instead, by giving the downstairs verb not just partial (Koenig, 1998) but full control over the realisation of its arguments, it covers the possibilities of climbing or trapping of arguments, the possible realisations of the subject including their dependence on lexical aspect, and the two possible realisations of the subject as either a climbing affix or a local phrasal complement. Moreover, the present approach to non-control *faire* is highly parallel to the periphrastic approach to climbing advanced by Aguila-Multner & Crysmann (2020): in both cases, morphological periphrasis goes piggyback on an independently required inside-out dependency, and in both cases, the syntax-semantics mismatch entailed by argument composition has been resolved in favour of syntax-semantics alignment. It is furthermore fully compatible with the approach to tense periphrasis in the interaction of the two phenomena. Finally, the present approach provides the missing piece towards a morphological theory of clitic climbing, showing that the periphrasis approach does scale up from auxiliary constructions to the full range of climbing phenomena.

References

Abeillé, Anne & Daniele Godard. 1996. La complémentation des auxiliaires français. *Langages* 30(122). 32–61.

Abeillé, Anne & Danièle Godard. 2002. The syntactic structure of French auxiliaries. *Language* 78(3). 404–452.

Abeillé, Anne, Danièle Godard & Philip Miller. 1997. Les causatives en français : un cas de compétition syntaxique. *Langue française* 62–74.

Abeillé, Anne, Danièle Godard & Ivan Sag. 1998. Two kinds of composition in French complex predicates. In Erhard Hinrichs, Andreas Kathol & Tsuneko Nakazawa (eds.), *Complex predicates in nonderivational syntax*, 1–41. New York: Academic Press.

Aguila-Multner, Gabriel & Berthold Crysmann. 2020. French clitic climbing as periphrasis. *Lingvisticæ Investigationes* 43(1). 23–61.

Bonami, Olivier. 2015. Periphrasis as collocation. *Morphology* 25(1). 63–110.

Bonami, Olivier & Gert Webelhuth. 2013. The phrase-structural diversity of periphrasis: a lexicalist account. In Marina Chumakina & Greville G. Corbett (eds.), *Periphrasis: The role of syntax and morphology in paradigms*, 141–167. Oxford: Oxford University Press.

Crysmann, Berthold. 2003. Clitic climbing revisited. In Jong-Bok Kim & Stephen Wechsler (eds.), *Proceedings of the 9th International Conference on Head-driven Phrase Structure Grammar, Kyung Hee University, Seoul, 5–7 August, 2002*, 67–89. Stanford: CSLI Publications.

Grimshaw, Jane. 1982. On the lexical representation of Romance reflexive clitics. In Joan Bresnan (ed.), *The mental representation of grammatical relations*, 87–148. Cambridge, Massachusetts: The MIT Press.

Hinrichs, Erhard & Tsuneko Nakazawa. 1990. Subcategorization and VP structure in German. In Shaun Hughes & Joe Salmons (eds.), *Proceedings of the third symposium on Germanic linguistics*, Amsterdam: Benjamins.

Hyman, Larry M. & Karl E. Zimmer. 1976. Embedded topic in French. In Charles N. Li (ed.), *Subject and Topic*, 191–211. New York: Academic Press.

Koenig, Jean-Pierre. 1998. Inside-out constraints and description languages for HPSG. In Andreas Kathol, Jean-Pierre Koenig & Gert Webelhuth (eds.), *Lexical and constructional aspects of linguistic explanation* Studies in Constraint-based Lexicalism, 265–279. Stanford: CSLI publications.

Manning, Christopher D. 1997. Romance complex predicates: In defence of the right-branching structure. *Ms., University of Sydney* .

Miller, Philip H. 1992. *Clitics and constituents in phrase structure grammar*. Garland, New York.

Miller, Philip H. & Ivan A. Sag. 1997. French clitic movement without clitics or movement. *Natural Language & Linguistic Theory* 15(3). 573–639.

Pollard, Carl & Ivan A Sag. 1994. *Head-driven phrase structure grammar*. University of Chicago Press.

Wehrli, Eric. 1986. On some properties of French clitic *se*. In Hagit Borer (ed.), *The syntax of pronominal clitics*, vol. 19 SYNTAX and SEMANTICS, 263–284. New York: Academic Press.

A Smurf-based Approach to Placeholder Expressions

Manfred Sailer
Goethe-University Frankfurt a.M.

Annika Dörner
University of Erfurt

Proceedings of the 28th International Conference on
Head-Driven Phrase Structure Grammar

Online (Frankfurt)

Stefan Müller, Nurit Melnik (Editors)

2021

pages 26–46

Keywords: German, HPSG, morphology, placeholder expression, smurfing

Sailer, Manfred, & Dörner, Annika. 2021. A Smurf-based Approach to Placeholder Expressions. In Müller, Stefan, & Melnik, Nurit (Eds.), *Proceedings of the 28th International Conference on Head-Driven Phrase Structure Grammar, Online (Frankfurt)*, 26–46. DOI: 10.21248/hpsg.2021.2.

Abstract

The *Smurf* comics series is, among others, famous for the so-called *smurf language*, in which words or parts of words can be replaced by *smurf*. We will argue that this *smurfing* has the properties of placeholding. Based on data from German translations of Smurf comics, we will provide a formalization of smurfing in German which can be generalized to a theory of placeholder expressions.

1 Introduction

Placeholder expressions are item such as those in (1).

(1) whatsit, whatchamacallit, thingamajig, what's-his/her-name, what-d'you-call-her/him, you-know-who, so-and-so

Cheung (2015) assignes them the following properties: First, they substitute a *target*. Second, the target can be a word, a phrase, but also a syllable. Third, there must be some pragmatic reasons for why the speaker utters the placeholder instead of the target.

In this paper, we will connect placeholder expressions and the *Smurf language*. The Smurfs are comics and cartoon figures which were invented by the Belgian cartoonist Peyo (Pierre Culliford) and which have been published since 1958. While originally in French, the Smurf comics have been translated into more than 25 languages. The "language" of the Smurfs, or rather their way of speaking is characterized by the use of the expression *smurf*, or its equivalent in the language of the publication, to replace words or parts of words. This is shown in (2) with the same example, taken from the English and German translations of the French original in (2c).

(2) *The Hungry Smurfs*. p. 7

a. What a disaster! It makes you want to smurf (= tear) your hair out! (en)

b. Welch eine Kataschlumpfe (= Katastrophe 'catastrophe')! Es ist zum Schlümpfe (= Haare 'hair') ausraufen! (de)

c. Quel désastre! C'est à s'arracher les schtroupfs (= cheveux 'hairs')! (fr)

The connection between the Smurf language and placeholder expressions that we want to explore in this paper can be traced back to the creation myth of the Smurf language. It has been repeatedly reported that the origin of the Smurf language goes back to the following conversation between Peyo and André Franquin, quoted from (Dayez, 2013, 9), our underlining and translation.

[†]We thank the reviewers and the audience for their comments, in particular Ash Asudeh, Berthold Crysmann, Marianne Desmets, and Adam Przepiórkowski. All errors are ours.

Peyo: André, passe-moi la... allez, le... le schtroumpf, là, près de toi!
'André, pass me the... smurf there, next to you!'
André: Tiens, voilà ton schtroumpf! 'Here comes your smurf'
Peyo: Merci de me l'avoir schtroumpfé, ...
'Thanks for smurfing it to me.'
quand je n'en aurai plus besoin, je te le reschtroumpferai.
'when I don't need it anymore, I'll smurf it back to you.'

In its first occurrence, the nonce word *schtroumpf* is used as an ad hoc placeholder as Peyo could not think of the word for salt. Franquin picks this word up and the two continue using it in an ever more playful way.

We will argue that smurfing has properties that make it a placeholder phenomenon: *smurf* replaces another expression and has special pragmatics. At the same time, *smurf* differs from other placeholders in that it is not used for filling a cognitive or lexical gap. Nonetheless, we claim that the analysis of smurfing will be a step towards an analysis of placeholder expressions in general.

We present two previous approaches to placeholder expressions in Section 2. In Section 3, we discuss the properties of German smurfing. An HPSG modelling of smurfing as a placeholder phenomenon is given in Section 4. We summarize our main results and point to future directions of research in Section 5.

2 Previous approaches to placeholder expressions

In this section, we will review the approaches to placeholder expressions in Enfield (2003) and Cheung (2015). Enfield (2003) argues that a placeholder makes a rather general truth-conditional semantic contribution, but its range of meaning can be constrained, and there are additional conditions on whether the speaker and/or the addressee can recover the target expression. We illustrate this with Enfield's analysis of the English placeholder *you-know-WHAT* in (3).

(3) *you-know-WHAT* (Enfield, 2003, 107)
   a. Something
   b. I don't want to say the word for this thing now
   c. I don't say it now because I know I don't have to
   d. By saying *you-know-WHAT* I think you'll know what I'm thinking of.

The variation in the range of meaning can be seen by constrasting *you-know-WHAT* with *you-know-WHO*, where only the latter is restricted to persons. Variation with respect to the accessibility of the replaced expression can be seen by contrasting *you-know-WHAT* with *WHAT-d'you-call-it*, which points to a (temporary) unavailability of the expression to the speaker.

The speaker-addressee related aspects of the characterization of placeholder expressions in Enfield (2003) have the properties of what is discussed under

the label of *use-conditional meaning* in Gutzmann (2013) and Gutzmann & McCready (2016), among others. Whereas we can ask whether the truth-conditional meaning of a sentence is true or false, the appropriate question for an expression with use-conditional meaning is whether it is used felicitously (Gutzmann & McCready, 2016). One type of expressions with use-conditional meaning are *slurs*, such as the word *kraut*. The word (truth-conditionally) refers to Germans, it is used felicitously if the speaker has a negative attitude of the speaker towards Germans in general. Use-conditional meaning has the projective properties of conventional implicature (Karttunen & Peters, 1979; Potts, 2005) and usually comes with speaker-addressee attitudes.

While we agree with most of Enfield's characterization of placeholder expressions, we do not think that they necessarily come with a vague, general truth-conditional semantics. We can show this with cases of replacement of parts of idioms by placeholder expressions. As we will concentrate on German data in the main part of this paper, we will use German examples here, too.

The idiom in (4) contains the word *Barthel*, a so-called *phraseologically bound word* or *cranberry word*. Such words are usually not found outside a particular expression (Aronoff, 1976; Dobrovol'skij, 1989; Richter & Sailer, 2003).[1] As the word *Barthel* is restricted to this expression, we cannot indicate a translation for it. Example (4) shows that the bound word cannot be replaced within the idiom with a general term such as *jemand* 'someone' or *etwas* 'something'.

(4) zeigen, wo Barthel/ #jemand/ #etwas den Most holt
show where ??/ someone/ something the cider gets
'show s.o. what's going on'

Contrary to this, idiom parts can be replaced with placeholder expressions. This is shown for the word *Barthel* in (5), which is replaced by the German placeholder expression *Dingsbums* 'thingamajig' (glossed as PHE for *placeholder expression*). This contrast between general terms and placeholder expressions is not compatible with Enfield's analysis of the truth-conditional meaning of placeholders.

(5) [sie] waren so motiviert, uns zu zeigen, wo Dingsbums
they were so motivated to.us to show where PHE
(= Barthel) den Most holt, daß...
?? the cider gets that
'they were to motivated to show us what's going on that ...'[2]

Cheung (2015) is not very detailed with respect to what we consider the use-conditional meaning of placeholder expressions, but provides a different

[1] See https://www.english-linguistics.de/codii/ for English and German bound words.
[2] https://tinyurl.com/y5f72cdz, accessed 14.10.2020.

approach to their referential meaning. He considers placeholders *metalinguistic demonstratives*. This means that they denote the expression that they replace. If *what's-her-name* is used instead of the name *Robin*, for example, the placeholder refers to the word *Robin*, i.e. $[\![$*what's-her-name*$]\!] =$ *Robin*. Whenever a placeholder is used, there is also an operator **SHIFT** in the structure. This operator, then, maps the expression denoted by the placeholder to its denotation, i.e. [[**SHIFT**(*Robin*)]] = [[*Robin*]] = **robin**.

This approach presents an elegant answer to the question of how a placeholder and its target are connected semantically. However, it is not clear if it captures all aspects of this connection. We will address this point in more detail in our discussion of the smurfing data.

There is, however, one problem. Cheung (2015, 276) shows with examples such as (6) that placeholding can also affect phonological units. In this example, the target of the placeholder is the second syllable of the Chinese form of the French name *Hollande*.

(6) Ao-shenme-de (= Ao-lang-de) shi xianren Faguo zongtong.
Ho-PHE-de (Hollande) be current France president
'Ao-something-de (= Hollande) is the current President of France.'

The problem of the analysis is that Cheung (2015, 302) assumes that the placeholder, together with the **SHIFT** operator, forms a syntactic constituent that is inserted between the two intact syllables. While it is conceivable that the placeholder denotes the syllabel *lang*, it is not clear what kind of semantic object **SHIFT**(*lang*) would denote.

Our brief discussion of two formal approaches to placeholding shows (i) that placeholders come with use-conditional semantic aspects, (ii) that they can substitute concrete lexical items, and (iii) that they can even substitute meaningless parts of lexical items. We will show in the next section that smurfing has the same properties.

3 Properties of smurfing

In this section, we will turn to smurfing. We mainly use the data collected in Dörner (2012), which consist of the 536 instances of smurfing found in 6 German Smurf stories. This will be expanded by browsing through early English, French, and German Smurf comics (1958–1988). In addition, we will use data and observations from the literature on smurfing (Bollig, 2016; Bourcier & Martin, 1996; Chatzopoulos, 2008). Chatzopoulos (2008) will be an important source as she provides the first formal analysis of smurfing.

Chatzopoulos (2008) distinguishes a literal and a "semantically unspecified" use of the morpheme *smurf* – or its equivalent in other languages. In its literal use, as in (7a), *smurf* refers to Smurfs or anything Smurf-related. This use is

also found in the speech balloons attributed to non-Smurfs, such as Gargamel. The second use, illustrated in (7b), constitutes the special use that Chatzopoulos refers to as *smurfing*.

(7) a. Gargamel: Smurf-prints! (= footprints left by Smurfs) (Chatzopoulos, 2008)
b. Are you making smurf (= fun) of me? (*The Fake Smurf*, 4)

In this section, we will look at two general properties of smurfing: its pragmatics and recoverability. Then, we will argue that there are two types of smurfing: phonological and morphological smurfing.

3.1 Pragmatics of smurfing

According to Chatzopoulos (2008), smurfing is a strong marker of Smurf identity. This is confirmed in an exemplary look at the German translations of two Smurf stories – *Der fliegende Schlumpf* (English title: *A Smurf in the Air*), and *Die Schlümpfe und die Zauberflöte* (English title: *The Smurfs and the Magic Flute*). In the first story, there are only Smurfs, i.e., all conversations are among Smurfs. We find smurfing in 33% of the panels in which there is a speech balloon. In contrast to this, the second story contains both Smurfs and non-Smurfs. In it, there is smurfing in 68% of the speech balloons showing Smurfs talking to each other. Smurfs use smurfing when talking to a non-Smurf in 23% of the panels. Non-Smurfs never use smurfing when talking to each other. There is no (successful) use of smurfing by a non-Smurf towards a Smurf.[3]

This indicates that non-Smurfs do not smurf. Furthermore, the smurfing rate is higher when Smurfs are depicted in contact with non-Smurfs than when they are among themselves. As smurfing is also used towards non-Smurfs – though to a lesser extent – it marks the Smurf identity of the speaker. This shows that smurfing is used to contrast Smurfs and non-Smurfs also linguistically. In other words, Smurfs are depicted in the comics as a special sociolinguistic group.

We will include this pragmatic property of smurfing in its use-conditional meaning, abbreviated as **Smurf-UC**.

(8) Identity use-conditional meaning of smurfing (**Smurf-UC**):
By smurfing, a speaker is marked as Smurf and signals their Smurf-ness.

Such a kind of use-conditional meaning is not uncommon. It is, in fact, at the heart of classical variationist research in sociolinguistics – though rephrased in terms of formal pragmatics.

[3] In this particular story, one non-Smurf tries to use smurfing to communicate with a Smurf, but never does so successfully.

3.2 Recoverability of smurfing

A central point in the research on smurfing is the question of the recoverability of the target. Chatzopoulos (2008) proposes an Optimality Theoretical approach to smurfing. She postulates a constraint SMURF: "Smurf all lexical morphemes." This constraint is outranked by a recoverability constraint which restricts smurfing to contexts in which the target can be recovered. She lists a number of factors that enable recoverability. One of them is phonological similarity, as in her example in (9).

(9) smurfday (= birthday) (Chatzopoulos, 2008)

Smurfing of parts of multiword expressions enhances recoverability as well. Chatzopoulos (2008) explicitly mentions "proverbs, idioms and phrases with some degree of fossilization." We saw an example of this in (2) above, where part of the idiom *tear one's hair out* is smurfed. The third factor named in Chatzopoulos (2008) is a pragmatically rich context. Such a context is often provided by the pictures or the general situation.

Bollig (2016, 55) challenges this common opinion that smurfing is always recoverable providing the example (10) from the story *Schtroumpf vert et vert schtroumpf* (English title: *Smurf vs. Smurf*).

(10) A: Qu'est-ce qui est schtroumpf, qui a un schtroumpf vert et qui schtroumpfe quand on le schtroumpfe?
'What is smurf, has a green smurf and smurfs when you smurf it?'
B: Je ne sais pas ... un schtroumpf?
'I don't know ... a smurf?'
A: Mais non, voyons! DEUX schtroumpfs!
'But no, look! TWO smurfs!'
(*Schtroumpf vert et vert schtroumpf*, p. 2; our translation)

While it is clear that the reader cannot recover the smurfing, the two Smurfs in conversation are depicted as being able to do so. Consequently, the speaker is depicted as assuming recoverability for the addressee. This is the same condition we find for the English placeholder *you-know-WHAT*, see (3d) above. We can formulate it as a second use-condition on smurfing in (11).

(11) Recoverability use-conditional meaning of smurfing (**Recov-UC**):
By smurfing the speaker thinks the addressee knows what target the speaker is thinking of.

After these two subsections on the pragmatics of smurfing, we will take a closer look at its grammatical properties in the following two sections.

3.3 Phonological smurfing

According to Chatzopoulos (2008), English *-smurf-* can replace one syllable, but keeps the overall metrical properties of the smurfed word. We find smurfing of individual (underlying) syllables in German as well, see (12). In the first line, we state the underlying syllable structure of the word containing the smurfing, together with its non-smurfed target. In the second line, we show the syllabification of the result.

(12) a. Ka.ta.schlumpf.e (= Ka.ta.stroph.e 'catastrophe')
→ Ka.ta.schlum.pfe

b. schlumpf.est.ier.en (= pro.test.ier.en 'protest')
→ schlum.pfes.tie.ren

Note, however, that the first segment of the syllable *test* is deleted in (12b). This might be due to some simplification of the syllable structure, even though *schlumpf.tes-tie-ren* is in line with the phonotactic constraints of German. In any case, this example indicates that the phonological replacement mechanism does not seem to be fully deterministic.

We saw with the Chinese example from Cheung (2015) in (6) that such a syllable replacement is possible with genuine placeholders as well. Example (13) shows the replacement of a syllable by a placeholder in German.

(13) soll an irgendeinen support eine analy...dingens datei
must.1.SG to some support an analy-PHE file
(= Analyse-datei) senden.
analysis-file send
'I must send an analysis file to some support.'[4]

As this type of smurfing is phonologically conditioned, we refer to it as *phonological (p-)smurfing*.

3.4 Morphological smurfing

In this subsection, we will show that, at least for German, a phonological account of smurfing is not sufficient. Instead, there is a second type of smurfing, which we will call *morphological (m-)smurfing*.

In (14a) the form *ver-schlumpf-t* occurs as the smurfed version of the word *verstanden* 'understood'. If we had a pure phonological smurfing, we would expect to find *verschlumpfen*, see (14b) instead. This shows that the smurfed verb *ver-schlumpf-en* has a different inflectional paradigm than the target *ver-steh-en* 'understand'.[5]

[4] https://tinyurl.com/y3kghypm, accessed 05.10.2020.

[5] We use the following abbreviations in the glosses of German examples: DER derivational affix, F feminine, GE prefix for past participle formation, INF infinitive, M masculine, N neuter, PTCP past participle, PL plural, PRS present tense, PST past tense, SG singular.

(14) a. Hast du ver-schlumpf-t (= ver-stand-en)?
have you DER-smurf-PTCP (= DER-stand-PTCP 'understood')
'Do you understand?'
b. *Hast du ver-schlumpf-en (= ver-stand-en)?

We also find cases in which there occurs a derivational affix that is not present in the target.

(15) Eine schlumpf-ig-e (= gut-e) Idee!
a smurf-DER-F.SG good-F.SG idea
'a good idea'

In (15), the target adjective consists just of a simple morpheme as the stem to which an inflectional suffix attaches. The smurfed version, however, consists of the root *schlumpf*, the derivational affix *-ig* (which marks denominal adjectives), and the inflectional marking. A purely phonological smurfing process would have let to the simpler form *schlumpf-e*.

These two examples suggest that the target of m-smurfing is a morphological unit rather than a syllable. This can be further substantiated by looking at the possible size of the target. In (14a) above, the target was a root morpheme, *steh* 'stand'. We also find cases in which the target is a combination of a root and a derivational affixes, such as (16).

(16) Um das Nützliche mit dem Angenehmen zu schlumpf-en
to the useful with the pleasant to smurf-INF
(= ver-bind-en), ...
(= DER-bind-INF 'connect')
'to mix business with pleasure, ...'

Finally, even entire compounds can be smurfed. The German expression of congratulation contains the compound *Glück-wunsch* 'luck-wish'. In our data, we find two possible smurfed forms of this word. In one, only the last component of the compound is replaced with *schlumpf*. In the other, the entire compound is realized as *Schlumpf*.

(17) Herzlichen Glück-schlumpf/ Schlumpf (= Glück-wunsch)!
hearly luck-smurf/ smurf! (= luck-wish)
'Congratulations!'

This shows that, in fact, the target of smurfing can be any morphological unit that contains at least one root and excludes inflection.

Interestingly, we find the same for general placeholder expressions in German as well. In (18), there are two occurrences of the placeholder *Dings*. The first one replaces the compound *Glück-wunsch* 'luck-wish', the second the compound *Geburts-tag* 'birth-day'.

(18) Herzlichen Dings (= Glück-wunsch) zum Dings (= Geburts-tag)!
heartly PHE (= luck-wish) to.the PHE (= birth-day)
'Congratulations to your birthday!'[6]

We have seen in (14a) that m-smurfing of verbs does not conserve the inflectional class of the target. The same is true for nouns. The masculine and neuter noun form *Schlumpf* 'smurf' forms its plural with an umlaut, *Schlümpf-e* 'smurf-PL'. This umlaut also occurs if the target does not have an umlaut. This is the case in example (2b) above. The neuter noun *Haar* 'hair' has an umlaut-less plural *Haar-e*. Nonetheless, its smurfed form, *Schlümpf-e*, shows the umlaut.

On the other hand, there is no umlaut or additional plural affix its feminine form, *Schlumpfe*, even if the target has an umlaut or an additional affix. This is shown in (19). The feminine word *Hand* 'hand' forms its plural with umlaut and *-e*, *Händ-e*. Nonetheless, the smurfed form remains *Schlumpfe*.

(19) ... bin ich dabei, mir die Schlumpf-e (= Händ-e) zu waschen
am I busy me the smurf-DER (= hand-PL) to wash
'I am busy washing my hands'

While a smurfed word does not inherit the inflectional properties and the internal morphological structure of the target, it inherits a number of morpho-syntactic properties, such as gender (for nouns) and auxiliary selection (for verbs), as well as argument selection in general.

We saw instances of gender inheritance in the examples above, such as (19) (feminine noun) and (17) (masculine noun). Auxiliary selection is illustrated in (20). The smurfed verb combines with the perfect auxiliary *haben* 'have' in (20a) and with the auxiliary *sein* 'be' in (20b), just as its target.

(20) a. Du hast schon ge-schlumpf-t (= ge-wähl-t)!
you have already GE-smurf-PTCP (= voted)
'You have already voted!'

b. Hast du aber Glück gehabt, dass ich gerade vorbei-ge-schlumpf-t
have you but luck had that I just along-GE-smurf-PTCP
(= vorbei-ge-komm-en) bin!
(= come along) am
'How lucky you are that I have just come along!'

In all examples, the argument structure of the smurfed expression is the same as that of the target. We add example (21), where the smurfed verb occurs with a reflexive pronoun *sich* 'himself'. The target of the smurfed verb is inherently reflexive. This shows that the argument structure does not follow from the target's semantics but is inherited from the target. Consequently, a

[6] https://tinyurl.com/y5unq8yb, accessed 30.09.2020.

m-smurfing	p-smurfing
depends on morphological structure	depends on syllable structure
target is (simple or complex) morphological unit	target is single underlying syllable
can combine with derivational affixes not in the target [*schlumpf-ig*]-*e* (= *gut-e*)	no internal complexity
form is determined by paradigm *schlumpf/schlümpf*	form is constant

Table 1: Differences between m- and p-smurfing

purely denotational relation between the placeholder and its target as in Cheung (2015) seems potentially problematic.

(21) Schnell, Gargamel be-schlumpf-t (= be-weg-t) sich etwas!
fast, Gargamel DER-smurf-3.SG.PRS (= moves) himself a bit
'Hurry up, Gargamel is moving a bit!'

The differences between m-smurfing and p-smurfing are summarized in Table 1. We showed that the distinction between phonological and morphological replacement can be found in general placeholding as well.

The data in this section illustrate the following properties of smurfing: There is a single inflectional word *Schlumpf* 'smurf' with umlaut, which has a literal and a placeholder use. The placeholder use comes with use conditions of Smurfness of the speaker and recoverablility of the target for the hearer. There are two types of smurfing: one that targets a phonological unit (p-smurfing), and one that targets a morphological unit (m-smurfing). We observed that these properties are also found in placeholder expressions. We take this as support for treating smurfing as an instance of placholding.

4 HPSG modelling

In the present section, we will develop our formal analysis of smurfing as part of an HPSG analysis of placeholding in general. We will use the morphological component of Bonami & Boyé (2006), extended with the modelling of compounding from Desmets & Villoing (2009). The inheritance of properties of the target to the overall smurfed expression will be expressed as *transparent heads*, which is a standard technique of HPSG, employed in Pollard & Sag (1994) and more explicitly in Levine (2010). We will assume that there are two lexemes *Schlumpf* – a literal and a non-literal lexeme. The two lexemes share the same

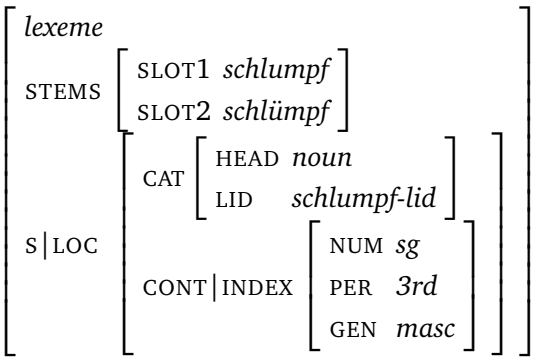

Figure 1: Lexical entry of the lexeme *Schlumpf* 'smurf'

inflectional paradigm. Smurfing is modelled as by allowing the use of the non-literal lexeme inside existing placeholder constructions of German.

4.1 Lexemes

Our formal modelling follows the empirical observations from Section 3 very closely. In this subsection, we will specify two related lexemes for the word *Schlumpf*, one for the literal and one for the non-literal use. We follow Sag (2012) and others in using a feature LEXICAL-IDENTIFIER (LID). The value of this feature is shared between a phrase and its head daughter. Nonetheless, contrary to other HPSG publications, we do not treat LID as a head feature but assume that it is defined on *category* objects.[7]

To account for the two uses of *Schlumpf* we assume that its LID value is of sort *schlumpf-lid*. This sort has two subsorts, *schlumpf-lit* and *schlumpf-phe*, for the literal and the non-literal use respectively.

Bonami & Boyé (2006) introduce a feature STEMS on lexemes, whose value for provides the stem allophones needed for inflection. Combining the features STEM and LID, we can specify the lexical entry for *Schlumpf* as in Figure 1.

This combines with constraints on the two different uses given in (22). The constraint in (22a) specifies that the literal use of *Schlumpf* refers to a Smurf. The non-literal, placeholder, use of *Schlumpf* does not have a semantics of its own, but contributes the two use conditions discussed in Sections 3.1 and 3.2: Smurf-ness of the speaker (**Smurf-UC**) and recoverability of the target for the addressee (**Recov-UC**). These two use conditions are included in the BACKGROUND feature, which is the place for projective semantic contributions in HPSG (Pollard & Sag, 1994; Green, 1994).

[7]This makes it possible to share head features among signs that do not belong to the same lexeme – see Figure 5 below and Soehn (2006) and Richter & Sailer (2009).

(22) a. Constraint on the literal use of *Schlumpf*:
$\left[\text{S}|\text{L}|\text{CAT}|\text{LID}\ \textit{schlumpf-lit} \right] \Rightarrow$

$$\left[\text{S}|\text{L}|\text{CONT} \left[\begin{array}{l} \text{INDEX}\ \boxed{1} \\ \text{RESTR}\ \left\{ \left[\begin{array}{l} \textit{smurf-rel} \\ \text{INST}\ \boxed{1} \end{array} \right] \right\} \end{array} \right] \right]$$

b. Constraint on the non-literal use of *Schlumpf*:
$\left[\text{S}|\text{L}|\text{CAT}|\text{LID}\ \textit{schlumpf-phe} \right] \Rightarrow$

$$\left[\text{S}|\text{L}|\text{CTXT}|\text{BACKGROUND}\ \left\{ \textbf{SMURF-UC}, \textbf{RECOV} \right\} \right]$$

The sort *schlumpf-phe* is the LID value of the placeholder use of the word *Schlumpf*. We assume that this sort is not only a subsort of *schlumpf-lid*, but also of a sort *placeholder-lid* (*phe-lid*). The same is true for other lexemes that have both a placeholder and a non-placeholder use, such as *Ding* 'thing'. If a word has only a placeholder use, such as German *Ding(en)s* or *Dingsbums* (both 'thingamajig'), the sort of its LID value is a subsort of *phe-lid* exclusively. We saw in our brief summary of Enfield (2003) that placeholder expressions in general come with a use-conditional meaning. They can, but need not, impose further restrictions on their syntax or semantics. In the case of *Schlumpf*, there is no semantic restriction, but it is specified as a masculine noun with a particular inflectional paradigm.

4.2 Placeholder constructions

In this subsection, we will present a general formalization of placeholding (in German). We assume that smurfing is regular placeholding, where the chosen placeholder is *Schlumpf*. We argued above that we need to distinguish two types of smurfing – and placeholding in general: phonological and morphological smurfing/placeholding. In all cases of smurfing/placeholding, the only semantic and pragmatic constraints contributed by placeholder are its use-conditional meaning. The semantics of the resulting expression as well as many of its morpho-syntactic properties are inherited from the target. To model this, we will assume that placeholding is a combination of two lexemes: a placeholder and its target. This combination is a generalization of compounding.[8] We think that phonological placeholding is an instance of *blending* (Fradin, 2015), whereas morphological placeholding is some other type of *subtractive lexeme combination* for which we have not found a fully parallel process outside placeholding yet.

As there is no explicit generalization of compounding in HPSG, we will take the architecture of Desmets & Villoing (2009) as starting point for our analysis.

[8]We are grateful to Berthold Crysmann and Marianne Desmets (p.c.) for stressing that placeholding is not an ordinary form of compounding.

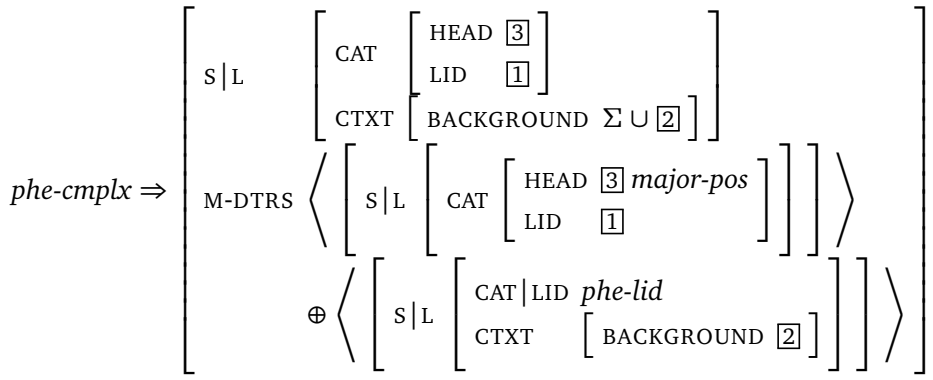

Figure 2: Constraint on the sort *phe-cmplx*

In this approach, lexemes can be simple or complex. Compounds are cases of complex lexemes. The components of a compound are its *morphological daughters*, for which Desmets & Villoing introduce a list-valued feature M-DAUGHTERS. Using this feature architecture, we assume that there is a sort *phe-complex* (*phe-cmplx*), which has two morphological daughters: the placeholder and the target.

In Figure 2, we provide the general constraint on all placeholder complexes: There are two morphological daughters. One is a placeholder, i.e., its LID value is of sort *phe-lid*. The other one is the target. The target must be of a major part of speech. The overall expression inherits from the target (at least) its LID value (1) and its HEAD value (3). The use-conditional meaning of the placeholder is integrated into the use conditions of the complex – specified as the BACKGROUND set 2.

This constraint on placeholding complexes captures a number of observations we have made in earlier sections. First, the target must be of a major part of speech and cannot be a functional morpheme. Second, placeholding is a process on lexemes, which means that inflection will be added to the placeholding complex. Third, the LID value of the complex is identical with that of the target.

This last property allows for the placeholding of parts of idioms and other fixed expressions, as observed in (4) and found massively in smurfing. Theories of idioms in HPSG have made heavy use of some way of lexeme-specific selection to guarantee that the idiom-specific words combine. Kay et al. (2015) show how this approach can be used even for syntactically fixed and semantically non-decomposable idioms such as *kick the bucket* 'die'. As a placeholding complex shares the LID value with the target, placeholding of idiom parts is immediately accounted for.

We can now turn to phonological placeholding, which we consider a type of blending. Fradin (2015) provides an overview of blending. In the classical cases, it is a very flexible type of combining lexemes in which truncated forms of the components are combined, as in English *brunch* (*br(eakfast) + (l)unch*)

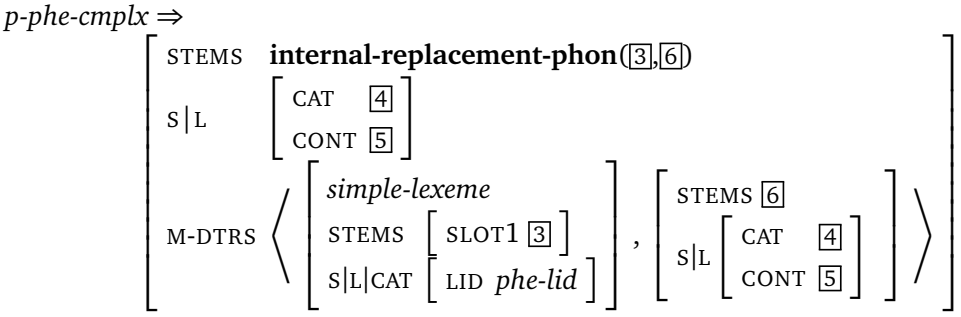

Figure 3: Constraint on p-placeholding

or German *Kripo* (*Kri(minal)* + *Po(lizei)* 'Criminal Investigation Department'). There are also blends in which one component is inserted into the other internally. Fradin (2015, 391) provides the French example *mét<amour>phose*, in which the word *amour* 'love' replaces (parts of) the second and the third syllable of *métamorphose* 'metamorphosis'. This is immediately reminiscent of the cases of p-smurfing discussed in Section 3.3, such as *Kata<schlumpf>e* (= Katastrophe 'catastrophe'), to use Fradin's notation.

We use the sort *p-phe-cmplx* to model p-placeholding. This sort is a subsort of *phe-cmplx* and, thus, inherits all properties from Figure 2. The sort *p-phe-cmplx* should ultimately also be treated as a subsort of the sort used for constraining (internal) blending. We treat the first component as the placeholder, the second component as the target. This is motivated by the fact that the second component determines most properties of complex words in German. The placeholder must be morphologically simple, whereas the target can be complex.

In this complex, the resulting word inherits its category and semantic properties from the target – given here as the CAT and CONT values (4 and 5). This is more information than the minimal information inherited from the target specified in Figure 2. The constraint on the supertype *phe-cmplx* ensures that the use-conditional meaning of the placeholder daughter projects to the overall complex. The special phonological effect of internal blending is encoded in the function **internal-replacement-phon**. This function takes as its arguments, the relevant stem form of the placeholder and the STEM value of the target. The output is a replacement of part of the target's phonology by the placeholder's phonology. Fradin (2015) names general principles governing the way in which the phonology of blends is determined.

In Figure 4, we illustrate how the constraints in Figures 2 and 3 are at work to derive the word *Katschlumpfe* as a p-placeholder combination of *Schlumpf* and *Katastrophe*. The highest local tree is the projection from an (unflected) lexeme to its inflected word form. We simplify the feature geometry in the tree.

Next, we can look at m-placeholding, which we will model using a sort *morphological-phe-complex* (*m-phe-cmplx*). The constraint on this type is given in Figure 5. We saw in Section 3.4 that the inflectional properties of the com-

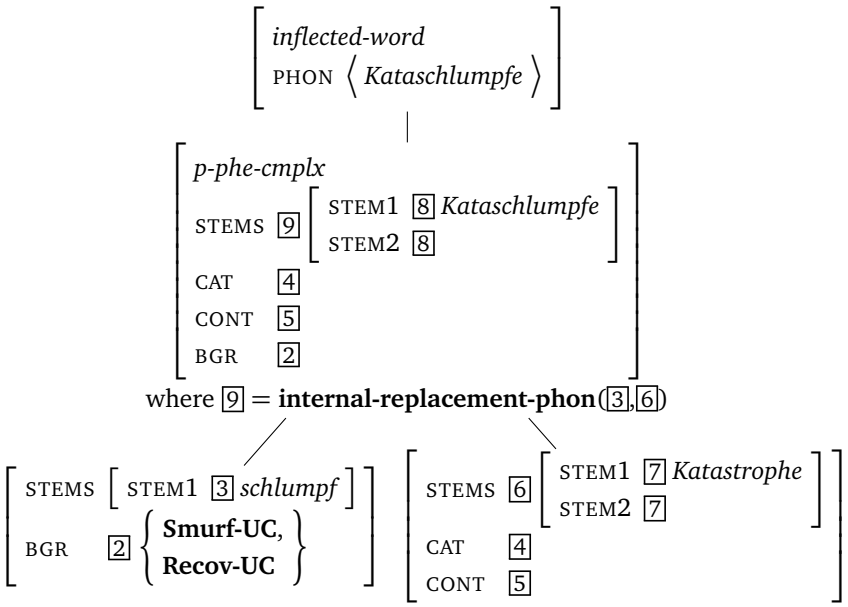

Figure 4: Derivation of the word *Kataschlumpfe*

bination are determined by the placeholder, not the target. For this reason, we assume that the placeholder is the second component in m-placeholding. At the same time, the complex shares its HEAD, VAL, and CONT values with the target. This is modelled through the technique of *transparent heads*, i.e., the morphological head is specified as inheriting these properties from the non-head. In Figure 6, we provide the derivation of the complex placeholder *schlumpf-ig-e*.

The constraint on the sort *m-phe-cmplx* not only allows for potentially complex targets, but the placeholder may be complex as well. This captures the fact that we find deriviational affixes on the placeholder that are absent from the target, such as *schlumpf-ig-e* 'smurf-DER-F.SG' for *gut-e* 'good-F.SG' in (15). We assume that the derivational affixes used with the basic placeholder expression only cause a conversion, i.e., a change of part of speech or of grammatical gender, but they do not change the LID value.

In our analysis, the placeholder and the target agree in part of speech, valence, and content. This explains the use of derivational affixes on the placeholder: if the target is an adjective, such as *gut* 'good', it cannot directly combine with the nominal placeholder *Schlumpf*. Consequently, the placeholder first needs to combine with a derivational affixes like *-ig*. Similarly, the placeholder can impose constraints on the kinds of targets it combines with. We saw this with *you-know-WHAT* in (3a). This placeholder constraints its target to objects and is not compatible with persons.

We should briefly turn to the question whether we can relate m-placeholding to a more general morphological process. There is a similarity to the replacement of lexemes by others in taboos or euphemisms (Allan & Burridge, 1991).

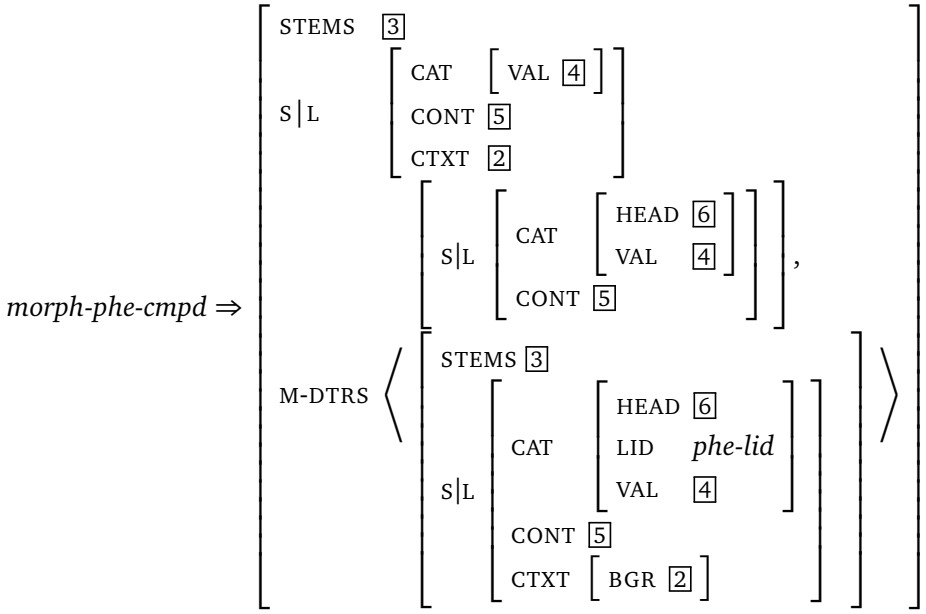

Figure 5: Constraint on m-placeholding

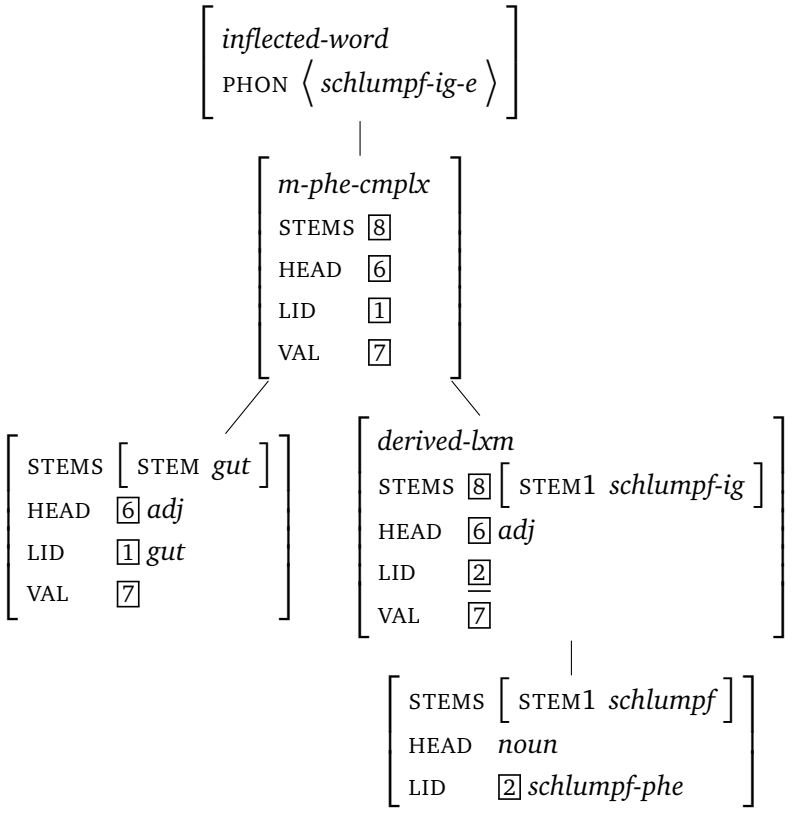

Figure 6: Derivation of the word *schlumpf-ig-e*

In German, the word *Scheiße* 'shit' is marked as vulgar. Replacing it with the phonologically similar word *Scheibe* 'disc' attenuates this marking. Allan & Burridge (1991, 15) refer to this type of lexeme replacement as *remodelling*. However, Allan & Burridge (1991) do not give examples in which such remodellings combine productively with derivational affixes.

Our analysis shows a certain parallel to the one in Cheung (2015): we assume a placeholder lexeme which does not have a lexically specified semantics. Cheung's operator **SHIFT** corresponds to our placeholder complex. Cheung (2015) needs to exclude free uses of the placeholder, i.e. uses of the placeholder not in the scope of **SHIFT**. Similarly, we need to exclude occurrences of placeholder lexemes outside placeholder complexes. This can be done with the constraint in (23). It determines that words cannot have an LID value of sort *phe-lid*. This constraint is modelled in analogy to the ban on the occurrence of non-canonical *synsem* objects in syntax.

(23) $word \Rightarrow \neg \left[\text{S}|\text{L}|\text{CAT}|\text{LID} \;\; phe\text{-}lid \right]$

All placeholder combinations inherit their LID value from the target – see Figure 2. Consequently, the constraint in (23) exclude the occurrence of targetless placeholders in sentences.[9]

A second distributional question that we would like to address is whether there may be recursive placeholding. At first sight, it seems reasonable to exclude spurious ambiguity. In the given modelling, it is possible to combine the noun *Wunsch* 'wish' with *schlumpf* in an m-placeholder complex. The result has a non-placeholder LID value. Consequently, it could combine recursively with yet another placeholder, even *Schlumpf*. To exclude this, we can require that a placeholder needs to make a recognizable contribution in a placeholder complex. As the main function of placeholders is use conditional, we can require that the use-conditional meaning of a placeholder, as collected in its BACKGROUND value, must not be included in the BACKGROUND value of the target.

This correctly excludes a redundant application of the same placeholder within one word. At the same time, it allows, in principle, the stacking of different placeholders.[10] It is conceivable that a placeholder word could be smurfed – though our database does not include such an example. This is illustrated in the constructed example in (24).

(24) Hast du den Dings-schlumpf gesehen?
have you the $\overline{\text{PHE-smurf}}$ seen?

'Did you see the WHAT-d'you-smurf-it?' (constructed)

The target of the smurfing is *Dingsbums*, which, itself is a placeholder expression. It could, for example, stand for the monster bird *Krakakass* (*Howli-*

[9]Depending on the available types of derivation and compounding, other morphological processes may also be restricted to non-placeholder lexemes.

[10]We are grateful to Ash Asudeh (p.c.) for bringing up this point.

bird in English). Consequently, it seems to be well motivated to require a non-redundant use-conditional contribution of placeholders.

5 Conclusion

In this paper, we provided a detailed look at smurfing in German Smurf comics. We classified smurfing as an instance of placeholding and developed a formal analysis. To our knowledge, such a placeholder analysis has not been proposed previously. This treatment can give an answer to the justified question of why smurfing should be of interest for formal linguistics, given that smurfing is an artistic invention. If our approach is on the right track, smurfing relies on the placeholding possibilities of a particular language. The only "invention" is in the choice of the placeholder lexeme and its use conditions. The smurfing complexes themselves fully rely on existing placeholder formation rules of the language. Support for this assumption comes from cross-linguistic differences in smurfing, as observed, for example, for English and Modern Greek in Chatzopoulos (2008). Smurfing can provide us with a rich database for placeholding as it is much more frequent than natural occurrences of placeholding.

Our approach combines the insights of previous work on placeholding and smurfing. The integration into a sign-based framework like HPSG makes it possible to combine semantic, pragmatic, and syntactic aspects. It is eventually this formal perspective that helped us identify two types of smurfing and placeholding. This separation helps solve problems of Chatzopoulos (2008), who only looked at phonological smurfing, and Cheung (2015), whose theory is suited for morphological placeholding, but less so for phonological placeholding.

There are at least two possible future directions to explore. First, we should explore smurfing and placeholding in other languages to test the hypothesis that smurfing builds on existing placeholding processes. Related to this, it is possible that there are other types of placeholding than phonological and morphological placeholding, which we identified for German.

Second, we can go beyond placeholding. We pointed out that phonological placeholding can be considered an instance of blending. Fradin (2015) describes blending as a systematic but extra-grammatical process. We reinterpreted this in treating blending as a grammatical process, which is, however, less constrained than compounding and has a potentially non-deterministic phonology. Similarly, we would like to explore the parallelism of morphological placeholding and other processes, such as remodelling in euphemisms.

References

Allan, Keith & Kate Burridge. 1991. *Euphemism and dysphemism: Language used as shield and weapon*. New York: Oxford University Press.

Aronoff, Mark. 1976. *Word formation in Generative Grammar*. Cambridge MA: MIT Press.

Bollig, Olivier. 2016. Análisis de la lengua de los Pitufos en la obra de Peyo. Comparativa de la versión original francófona con la traducción al español. doi:10.13140/RG.2.1.2463.3847. Master thesis. Brussels.

Bonami, Olivier & Gilles Boyé. 2006. Deriving inflectional irregularity. In Stefan Müller (ed.), *Proceedings of the 13th international conference on HPSG*, 361–380. https://cslipublications.stanford.edu/HPSG/2006/bonami-boye.pdf.

Bourcier, André & Brigitte Martin. 1996. Schtroumpfez-nous une explication grammaticale. In Julie Laberge & Robert Vézina (eds.), *Actes des 10e journées de linguistique (1996)*, 43–49. Quebec: Université Laval. https://files.eric.ed.gov/fulltext/ED403750.pdf.

Chatzopoulos, Catherine. 2008. Optimal smurfing in English and Greek. In Nikolaos Lavidas, Elissavet Nouchoutidou & Marietta Sionti (eds.), *New perspectives in Greek linguistics*, 189–202. Cambridge: Cambridge Scholars Publishing. www.academia.edu/1135038/Optimal_smurfing_in_English_and_Greek.

Cheung, Lawrence Y.-L. 2015. Uttering the unutterable with *wh*-placeholders. *Journal of East Asian Linguistics* 24. 271–308. doi:10.1007/s10831-014-9130-x.

Dayez, Hugues. 2013. Introduction. In *Les schtroumpfs. L'intégrale 1, 1958–1966*, 9–39. Marcinelle: Dupuis.

Desmets, Marianne & Florence Villoing. 2009. French VN lexemes: Morphological compounding in HPSG. In Stefan Müller (ed.), *Proceedings of the 16th international conference on HPSG*, 89–109. Stanford: CSLI Publications. https://cslipublications.stanford.edu/HPSG/2009/desmets-villoing.pdf.

Dobrovol'skij, Dmitrij. 1989. Formal gebundene phraseologische konstituenten: Klassifikationsgrundlagen und theoretische analyse. In Wolfgang Fleischer, Rudolf Große & Gotthard Lerchner (eds.), *Beiträge zur Erforschung der deutschen Sprache*, vol. 9, 57–78. Leipzig: Bibliographisches Institut.

Dörner, Annika. 2012. Korpusstudie zur Schlumpfsprache. Term paper, Göttingen University.

Enfield, Nick J. 2003. The definition of WHAT-*d'you-call-it*: Semantics and pragmatics of 'recognitional deixis'. *Journal of Pragmatics* 35(1). 101–117.

Fradin, Bernhard. 2015. Blends. In Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (eds.), *Word-formation. An international handbook of the languages of Europe*, vol. 1, 386–413. Berlin: Mouton de Gruyter.

Green, Georgia M. 1994. The structure of CONTEXT: The representation of pragmatic restrictions in HPSG. *Studies in the Linguistic Sciences* 24(1/2). 215–232.

Gutzmann, Daniel. 2013. Expressives and beyond. In Daniel Gutzmann & Hans-Martin Gärtner (eds.), *Beyond expressives* 28, 1–58. Leiden: Brill. doi:10.1163/9789004183988_002.

Gutzmann, Daniel & Eric McCready. 2016. Quantification with pejoratives. In Rita Finkbeiner, Jörg Meibauer & Heike Wiese (eds.), *Pejoration* (Linguistik Aktuell/Linguistics Today 2016), 75–102. Amsterdam and Philadelphia: John Benjamins.

Karttunen, Lauri & Stanley Peters. 1979. Conventional implicature. In Choon-Kyo Oh & David A. Dinneen (eds.), *Presupposition* (Syntax and Semantics 11), 1–56. New York: Academic Press.

Kay, Paul, Ivan A. Sag & Dan Flickinger. 2015. A lexical theory of phrasal idioms. Manuscript. www1.icsi.berkeley.edu/~kay/idiom-pdflatex.11-13-15.pdf.

Levine, Robert D. 2010. The *Ass* camouflage construction: Masks as parasitic heads. *Language* 86(2). 265–301.

Pollard, Carl & Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar*. Chicago and London: University of Chicago Press.

Potts, Christopher. 2005. *The logic of conventional implicatures* (Oxford Studies in Theoretical Linguistics 7). Oxford: Oxford University Press.

Richter, Frank & Manfred Sailer. 2003. Cranberry words in formal grammar. In Claire Beyssade, Olivier Bonami, Patricia Cabredo Hofherr & Francis Corblin (eds.), *Empirical issues in formal syntax and semantics*, vol. 4, 155–171. Paris: Presses Universitaires de Paris-Sorbonne.

Richter, Frank & Manfred Sailer. 2009. Phraseological clauses in constructional HPSG. In Stefan Müller (ed.), *Proceedings of the 16th international conference on HPSG*, 297–317. Stanford: CSLI Publications. https://cslipublications.stanford.edu/HPSG/2009/richter-sailer.pdf.

Sag, Ivan A. 2012. Sign-Based Construction Grammar: An informal synopsis. In Hans C. Boas & Ivan A. Sag (eds.), *Sign-Based Construction Grammar*, 69–202. Stanford: CSLI Publications.

Soehn, Jan-Philipp. 2006. *Über Bärendienste und erstaunte Bauklötze. Idiome ohne freie Lesart in der HPSG*. Frankfurt am Main: Peter Lang. Ph.D. thesis, Friedrich-Schiller-Universität Jena.

Verbal reduplication in Mandarin Chinese: An HPSG account

Yanru Lu
Humboldt-Universität zu Berlin

Stefan Müller
Humboldt-Universität zu Berlin

Proceedings of the 28th International Conference on
Head-Driven Phrase Structure Grammar

Online (Frankfurt)

Stefan Müller, Nurit Melnik (Editors)

2021

pages 47–67

Keywords: Chinese, HPSG, reduplication

Lu, Yanru, & Müller, Stefan. 2021. Verbal reduplication in Mandarin Chinese: An HPSG account. In Müller, Stefan, & Melnik, Nurit (Eds.), *Proceedings of the 28th International Conference on Head-Driven Phrase Structure Grammar, Online (Frankfurt)*, 47–67. DOI: 10.21248/hpsg.2021.3.

Abstract

The current study presents an HPSG analysis for verbal reduplication in Mandarin Chinese. After discussing its interaction with *Aktionsarten* and aspect markers, we argue that it is a morphological rather than syntactic process. We put forward a lexical rule for verbal reduplication in Mandarin Chinese and the different forms of reduplication are captured in an inheritance hierarchy. The interaction between verbal reduplication and aspect marking is handled by multiple inheritance. This analysis covers all forms of verbal reduplication in Mandarin Chinese and has none of the shortcomings of the previous analyses.

1 Introduction

In Mandarin Chinese, verbs can be reduplicated to express a delimitative aspectual meaning (e.g. Chao, 1968; Q. Chen, 2001; Dai, 1997; Li, 1996; Li & Thompson, 1981; Tsao, 2001; Xiao & McEnery, 2004; Yang, 2003; Zhu, 1998). It means that the event or state denoted by the verb happens in a short duration and/or a low frequency (Xiao & McEnery, 2004, 155), such as illustrated in (1).[1] Thus, verbal reduplication in Mandarin Chinese is often translated as doing something "a little bit/for a little while".

(1) a. qing ni chang zhe dao cai.
please you taste this CLF dish
'Please taste this dish.'

b. qing ni **chang-chang** zhe dao cai.
please you taste-taste this CLF dish
'Please taste this dish a little bit.'

The current study tries to determine a suitable formal and unified analysis for the structure of verbal reduplication in Mandarin Chinese. It provides a novel HPSG analysis to this phenomenon and avoids the problems of previous approaches.

After this introduction, we will present in Section 2 the forms and syntactic distribution as well as the semantics of verbal reduplication in Mandarin Chinese. Importantly, we restrict the object of this study to the AA, A-*yi*-A, A-*le*-A, A-*le*-*yi*-A, ABAB and AB-*le*-AB forms of verbal reduplication in Mandarin Chinese. We will also discuss in this section the question of whether the reduplication is a morphological or syntactic process with the help of empirical data. In Section 3, we will discuss the advantages and drawbacks of previous approaches. Later on, we will present a new HPSG account for verbal reduplication in Mandarin Chinese in Section 4.

In addition to introspection, the Modern Chinese subcorpus of the corpus of the *Center for Chinese Linguistics of Peking University* (CCL) (Zhan et al., 2003, 2019) was also consulted. Other examples from novels and plays written by native speakers were also considered.

[1] In the following paper, reduplications in the example sentences will be boldfaced.

2 The phenomenon

This section introduces the fundamental grammatical behaviors of verbal reduplication in Mandarin Chinese. After illustrating its forms, syntactic distribution and semantics, we discuss its interaction with *Aktionsarten* and other aspect markers in Mandarin Chinese, and the question of whether it is better analyzed as a morphological or a syntactic phenomenon.

Verbal reduplication in Mandarin Chinese takes the forms listed in (2).

(2) a. for monosyllabic verbs: *shuo* 'say'

i.	shuo-shuo say-say	AA
ii.	shuo-yi-shuo say-one-say	A-*yi*-A
iii.	shuo-le-shuo say-PFV-say	A-*le*-A
iv.	shuo-le-yi-shuo say-PFV-one-say	A-*le*-*yi*-A
v.	shuo-shuo-kan say-say-look	AA-*kan*
vi.	shuo-kan-kan say-look-look	A-*kan*-*kan*[2]

b. for disyllabic verbs: *lai-wang* come-go 'come and go/communicate'

i.	lai-wang-lai-wang come-go-come-go	ABAB
ii.	lai-wang-le-lai-wang come-go-PFV-come-go	AB-*le*-AB
iii.	lai-lai-wang-wang come-come-go-go	AABB

c. for V-O compounds: *chang-ge* sing-song 'sing'

i.	chang-chang-ge sing-sing-song	AAB
ii.	chang-yi-chang-ge sing-one-sing-song	A-*yi*-AB
iii.	chang-le-chang-ge sing-PFV-sing-song	A-*le*-AB

Arcodia et al. (2014), Fan (1964), Melloni & Basciano (2018) and Xie (2020) compared the AA, ABAB and AABB forms of reduplication and found a number of differences between the AA, ABAB forms compared to the AABB form in terms

[2]This form is more common in Taiwan than in Mainland China.

of their semantics, productivity, syntactic distribution and origin. The current study will only focus on the AA, A-*yi*-A, A-*le*-A, A-*le*-*yi*-A, ABAB and AB-*le*-AB forms of verbal reduplication in Mandarin Chinese.[3] AA-*kan*, A-*kan*-*kan*, AAB, A-*yi*-AB, A-*le*-AB will also be mentioned occasionally to provide further arguments.

The reduplication has a similar syntactic distribution as an unreduplicated verb. The reduplication cannot be aspect marked, though, except with the perfective aspect marker *le* (for further discussions see Section 2.2). The reduplication is incompatible with an expression that quantifies the duration or the extent of the event expressed in the sentence, as in (3) (L. Chen, 2005, 114–115; Li, 1998, 83–84). This is probably because the reduplication already contains a quantity meaning (L. Chen, 2005, 114–115; Li, 1998, 84), namely a short duration or a small extent.

(3) a. ta yi tian pao shi li. (Li, 1998, 83)
he one day run ten mile
'He runs ten miles a day.'

b. * ta yi tian **pao**-**pao** shi li.
he one day run-run ten mile

The reduplication has a *delimitativeness* meaning (e.g. Chao, 1968; Q. Chen, 2001; Dai, 1997; Li, 1996; Li & Thompson, 1981; Tsao, 2001; Xiao & McEnery, 2004; Yang, 2003; Zhu, 1998). The semantics of the reduplication has the properties of transitoriness, holisticity and dynamicity (Dai, 1997, 70–79; Xiao & McEnery, 2004, 155–159). It presents the situation as a transitory and non-decomposable whole, which involves not only changes in the initiation and termination of an event, but also changes in the transitory process itself. Compared to (4a), which could mean that the protagonist kept staring at the the footprint, (4b) indicates that the protagonist took a brief look or several brief looks at the footprint and looked away in the end, which is a process full of changes.

(4) a. Wu Xumang kan-le zuo-an shi liuxia de jiaoyin …
Wu Xumang look-PFV commit-crime when leave DE footprint
(Xiao & McEnery, 2004, 158)
'Wu Xumang looked at the footprint left when the crime was committed.'

b. Wu Xumang **kan-le-kan** zuo-an shi liuxia de jiaoyin …
Wu Xumang look-PFV-look commit-crime when leave DE footprint
(Xiao & McEnery, 2004, 158)
'Wu Xumang looked a little bit at the footprint left when the crime was committed.'

As for the other forms of the reduplication, A-*yi*-A is considered to have the same core semantics as AA, despite being described to tend to have different pragmatic

[3]For sake of simplicity, the term *reduplication* will be used in the following text to refer specifically to the AA, A-*yi*-A, A-*le*-A, A-*le*-*yi*-A, ABAB and AB-*le*-AB forms of verbal reduplication in Mandarin Chinese, if not specified otherwise.

uses (Yang, 2003). The semantics of A-*le*-A can be deduced compositionally from its structure. It is a hierarchical combination of the perfective aspect and delimitativeness, "conveying a transitory event which has been actualized" (Xiao & McEnery, 2004, 151). AA-*kan* and A-*kan-kan* are described to express a "try …and find out" meaning (Cheng, 2012, 63). Tsao (2001, 290) also observed that the tentative meaning is particularly prominent when the reduplication is followed by *kan* 'look'. We consider the tentativeness implied by these two forms to be a pragmatic extension of delimitativeness. The tentative meaning is made prominent by the verb *kan* 'look', and the whole structure can be understood as "do A a little bit and see".

2.1 Interaction with *Aktionsarten*

Previous research often claimed that the reduplication can only be used for verb classes of certain *Aktionsarten*, while it is infelicitous for other ones. Hong (1999, 277–278) and Li & Thompson (1981, 234–235) suggested that reduplication is only possible for volitional activity verbs. Dai (1997, 70–71) and Tsao (2001, 290) both considered that reduplication can only be used in dynamic situations. The former further claimed that achievement verbs cannot be reduplicated. Arcodia et al. (2014, 20), Basciano & Melloni (2017) and Xiao & McEnery (2004, 155) proposed that only [+dynamic] and [−result] verbs can be reduplicated. This means that the reduplication can only interact with activities and semelfactives, but not with states and achievements.

Q. Chen (2001, 53) and Yang (2003, 10–11) acknowledged that the reduplication of non-volitional verbs is more restricted than that of volitional ones. But Zhu (1998, 381–382) listed a number of non-volitional predicates that can be reduplicated. We found the examples shown in (5) in CCL where non-volitional verbs *weiqu* 'feel wronged', *ren-xing* 'be willful' and *diao* 'drop' are reduplicated.

(5) a. keshi xian mujin, dajia ye zhihao **weiqu-weiqu**
but now current everybody also can.only feel.wronged-feel.wronged
le. (CCL)
PTC
'But now, everybody can only feel wronged a little bit.'

b. ta-men neng zuo de buguo shi **ren-ren-xing** shua
she-PL can do DE just be be.willful-be.willful-temperament play
dian'er xiao piqi **diao-diao** yanlei shenme de. (CCL)
a.little small temper drop-drop tear what DE
'What they can do is just to be a little bit willful, to lose their temper a little bit and to drop a little bit of tears or something.'

It is true that the reduplication of stative and achievement verbs is not as easily acceptable as that of activities and semelfactives. Compared to the questionable reduplication of the stative verb *bing* 'be sick' in (6a) and that of the achievement verb

ying 'win' in (6b), the reduplication of the activity verb *kan* 'watch' in (6c) and that of the semelfactive verb *kesou* 'cough' in (6d) is readily acceptable.

(6) a. ? ta **bing-bing** jiu hao le. (Xiao & McEnery, 2004, 155)
he be.sick-be.sick then well PTC
Intended: 'He was sick for a little while and then got well.'

b. ? ta **ying-le-ying** na chang bisai. (Xiao & McEnery, 2004, 155)
he win-PFV-win that CLF competition
Intended: 'He won that competition a little bit.'

c. ta **kan-le-kan** na chang bisai.
he watch-PFV-watch that CLF competition
'He watched that competition for a little while.'

d. ta **kesou-kesou** jiu hao le.
he cough-cough then well PTC
'He coughed a little bit and then got well.'

However, examples such as those in (7a) – (7b) were found in novels and plays written by native speakers and example sentences like (7c) and (7d) constructed by native speaker linguists. Here, achievement verbs like *wang* 'forget' and *sheng* 'give birth to' and stative verbs like *shutan* 'be comfortable' and *bing* 'be sick' are reduplicated.

(7) a. deng ren-men ba zhe jian shi **wang-wang** zai shuo ba.[4]
wait people-PL BA this CLF incident forget-forget then talk PTC
'Let's wait until people forget this incident a little bit and then talk about it.'

b. huitou mo ge zao **shutan-shutan** ba.[5]
later wipe CLF bath be.comfortable-be.comfortable PTC
'Let's take a bath later and be comfortable for a little while.'

c. wo zhen xiang **bing-yi-bing**, xie ta ge shi tian ban yue.
I really want be.sick-one-be.sick rest it CLF ten day half month
(Q. Chen, 2001, 54)
'I really want to be sick for a little while and rest for ten days or half a month.'

d. jiao ta **sheng-sheng** xiaohai, jiu zhidao zuo muqin de
let she give.birth.to-give.birth.to child then know COP mother DE
gan-ku le. (L. Chen, 2005, 112)
sweet-bitter PTC
'Let her try to give birth to a child and then she will know the bittersweetness of being a mother.'

[4]Liu, Zhen. 1963. *Chang chang de liushui [Long long water]*, 72. Beijing: The Writers Publishing House.

[5]Tian, Han. 1959. *Tianhan xuanji [Selected works of Tianhan]*, 122. Beijing: People's Literature Publishing House.

This shows that although the reduplication does have a tendency to interact with volitional verbs and with activities and semelfactives due to its dynamic meaning, this is by no means a rigid constraint, and non-volitional verbs, states and achievements can be reduplicated in certain contexts as well.

2.2 Interaction with aspect markers

As mentioned above, the reduplication can only be marked by the perfective aspect marker *le* but not other aspect markers.[6] We consider this incompatibility be due to semantic reasons.

Xiao & McEnery (2004, Ch. 4) considered *le*, *guo* and reduplication to indicate perfective aspects. The perfective aspect marker *le* is compatible with reduplication while the experiential aspect marker *guo* is not. Xiao & McEnery (2004, 128–131) described that *le* "can focus on both heterogeneous internal structures and changing points" (Xiao & McEnery, 2004, 129). *Le* is compatible with the reduplication, because its dynamicity can relate to not only the termination or instantiation of an event (a point of change), but also the process of the situation, just like that of the reduplication.

On the other hand, the experiential aspect marker *guo* cannot co-occur with a reduplicated verb, because its dynamicity attributes to an "experiential change" (Xiao & McEnery, 2004, 148), namely that a situation has been experienced historically and that "the final state of the situation no longer obtains" at the reference time (Xiao & McEnery, 2004, 144). It is clear that *guo* only indicates a change at the termination of a situation and cannot express the dynamicity within a situation. Hence, it is incompatible with the semantics of the reduplication.

Due to the holistic semantics of the reduplication, it is incompatible with imperfective aspect markers: the durative aspect marker *zhe* and the progressive aspect marker *zai*, as both only focus on a part of the situation and do not view the situation as a whole (Xiao & McEnery, 2004, Ch. 5).

From the illustration above, it seems that due to its semantics, reduplication can only be marked by *le* but not the other aspect markers.

2.3 Word vs. phrase

The literature on reduplication makes different assumptions on whether it is a morphological or syntactic phenomenon. Chao (1968) and Li & Thompson (1981) listed reduplication under morphological processes. Arcodia et al. (2014), Basciano & Melloni (2017), Melloni & Basciano (2018), Xie (2020), Xiong (2016), Yang & Wei (2017), on the other hand, claimed it to be syntactic. This section reviews the arguments in Xie (2020), applies the tests proposed by Duanmu (1998) and Schäfer (2009) to distinguish words from phrases in Mandarin Chinese. The results argue for a morphological status of reduplication.

[6]There is no consensus on which elements exactly are considered aspect markers in Mandarin Chinese. We only discuss the most commonly recognized ones here.

Xie (2020) compared the AA and the ABAB forms of reduplication with the AABB form and claimed that AA and ABAB are syntactic processes while AABB is morphological. She pointed out that AA and ABAB behave differently from AABB in their productivity, possibility of *le* insertion, categorial stability, transitivity, and input/output constraints. While AA and ABAB are highly productive, AABB shows low productivity. *Le* can be inserted freely into AA and ABAB but not into AABB. The output of AA and ABAB does not change the grammatical category of the input (verb), but the output of AABB could have other categories such as adverb or adjective. AA and ABAB do not change the valency of the input verb, but AABB makes a transitive verb intransitive. The two groups also have different input and output constraints. Xie (2020) claimed that only dynamic and volitional verbs can undergo AA or ABAB reduplication (but see Section 2.1). On the other hand, AABB requires its input to be a complex verb whose constituents are either synonymous, antonymous or logically coordinated. Moreover, the output of AABB has an increasing meaning, i.e. an event happens repeatedly or continuously, as opposed to the delimitative meaning of AA and ABAB.

However, a morphological process can be productive, and it does not necessarily change the category or valency of the input. Further, if *le* is considered to be a morphological element (e.g. Huang et al., 2009; Müller & Lipenkova, 2013), the insertion of *le* does not have to be viewed as a syntactic process either. It seems that Xie (2020) only showed that AA and ABAB are different processes than AABB, but not necessarily that the former is syntactic while the latter morphological.

It is, therefore, necessary to resort to other tests that are intended to distinguish words from phrases. Duanmu (1998) and Schäfer (2009) proposed the following four tests to distinguish words from phrases in Mandarin Chinese: semantic compositionality, phrasal extension, phrasal substitution and conjunction reduction.[7]

The semantic criterion is that the meaning of a phrase is usually built up in a compositional way while that of a word usually not (Duanmu, 1998, 140; Schäfer, 2009, 275). The meaning of the reduplication is not compositional, as it does not mean that the event denoted by the verb happens twice or multiple times, but rather that the event happens for a short duration and/or a low frequency.

The first syntactic test is phrasal extension, namely the addition of optional elements (Duanmu, 1998, 150; Schäfer, 2009, 280). Optional elements that can possibly appear in a phrase should be able to be added into it. Subparts of a phrase should be able to be modified separately. And these should not be possible for a word. As illustrated in (2) in Section 1, the reduplication can only be separated by *le* and *yi*. As mentioned above, whether aspect markers are considered to be morphological or syntactic elements depends on the theoretical framework (and possibly the target language). And the status of *yi* is unclear. Also, each element in the reduplication cannot be modified by itself. Compared to (8a), where the adverbial *qingsheng de*

[7]It is important to note that none of these criteria are sufficient or necessary to determine the word or phrase status of an expression. Nevertheless, they together might suggest which of the two statuses is more likely.

'quietly' modifies the whole reduplication, (8b) is ungrammatical, as the adverbial cannot modify the second element in the reduplication alone.

(8) a. ta qingsheng de **xiao-le-xiao**.
he quietly DE laugh-LE-laugh
'He quietly laughed a little bit.'

b. * ta **xiao-le** qingsheng de **xiao**.
he laugh-PFV quietly DE laugh

The second syntactic test is phrasal substitution, namely the substitution of smaller exemplars of a specific category with a full blown XP (Duanmu, 1998, 152; Schäfer, 2009, 280). If a part of an expression is actually an XP that only contains one element, a full realization of this XP should be possible as well. Otherwise, this expression is considered to be a word. In a reduplication structure, it is ungrammatical to substitute each element with a full VP (9).

(9) a. ta **chang-le-chang** tang.
he taste-PFV-taste soup
'He tasted the soup a little bit.'

b. * ta **chang tang le chang tang**.
he taste soup PFV taste soup

The third syntactic criterion is conjunction reduction. It should only be possible for coordinated phrases and not for coordinated words (Duanmu, 1998, 137; Schäfer, 2009, 283). For the reduplication, conjunction reduction does not seem to be possible. In (10a), the reduplication *jiao-jiao* 'chew a little bit' is coordinated with a simple verb *mo* 'apply' together with the adverbial *yidian* 'a little bit'. Without the adverbial *yidian* 'a little bit', *mo* 'apply' by itself cannot express the additional 'a little bit' meaning even when it is coordinated with a reduplicated verb. Similarly, in (10b), the reduplication *kan-le-kan* 'look a little bit' is coordinated with the predicate *zou-le chulai* 'walked out'. The verb in the latter case is not reduplicated and it cannot express the delimitative meaning either.

(10) a. wujian gong-xiu mo dian bohe-gao huo **jiao-jiao**
midday work-break apply a.little mint-cream or chew-chew
kouxiangtang. (CCL)
chewing.gum
'During the working break in the midday, apply a little bit of mint cream or chew some chewing gum a little bit.'

b. Songailing **kan-le-kan** yupen you zou-le chulai. (CCL)
Songailing look-PFV-look bath.tub again walk-PFV out
'Songailing looked at the bath tub a little bit and walked out again.'

Following the analyses above, it is clear that the reduplication failed all of the tests. Therefore, it seems more likely to assume the reduplication to be a morphological process rather than a syntactic one.

3 Previous analyses

Previous analyses on the reduplication in Mandarin Chinese and in other languages can be classified into three groups: the reduplicant as a verbal classifier, the reduplicant as an aspect marker, and the postulation of a special reduplication structure.

Chao (1968), Fan (1964) and Xiong (2016) analyzed the reduplicant in Mandarin Chinese as a verbal classifier. A verbal classifier is "a measure for verbs of action expresses the number of times an action takes place" (Chao, 1968, 615). In this analysis, the first element in the reduplication is the actual verb, the second element is a verbal classifier borrowed from the verb, and *yi* 'one' is an optional pseudo-numeral that only has an abstract 'a little bit' meaning. Although the reduplication and the verbal classifier both serve to quantify the extent of an event and can often be used interchangeably, they behave differently in the following three aspects. First, the verb and the verbal classifier can be separated, while the reduplication cannot (Paris, 2013, 269). Second, unlike verbal classifiers, the *yi* 'one' in A-*yi*-A cannot be replaced by other numerals (Yang & Wei, 2017, 299–230). Third, idioms lose their idiomatic meaning when used with verbal classifiers, but maintain their idiomatic meaning with reduplications (Yang & Wei, 2017, 230–231). Based on these observations, it seems inappropriate to view the reduplicant as a kind of verbal classifier.

A number of studies consider the reduplicant to be a delimitative aspect marker (Arcodia et al., 2014; Basciano & Melloni, 2017; Yang & Wei, 2017) due to the delimitative meaning of the reduplication. Travis (1999, 2000) also analyzed the reduplication in Tagalog as an imperfective aspect marker. In Arcodia et al. (2014) and Basciano & Melloni (2017)'s analysis, the reduplication of stative and achievement verbs is structurally ruled out, which does not fit the empirical observations we presented in Section 2.1. The other analyses along the line of aspect marker all have problems with the A-*yi*-A form, as the addition of *yi* in the reduplication does not lead to further syntactic or semantic functions. Moreover, although the reduplicant is postulated as a special affix that copies the phonology of the base morpheme, the exact nature of this copying process is not formalized.

Ghomeshi et al. (2004) gave an analysis for Contrastive Reduplications (CRs) in English like (11) based on the Parallel Architecture proposed by Jackendoff (1997, 2002) as shown in Figure 1.[8]

(11) I make the tuna salad, and you make the **SALAD-salad**.

Applying this to the reduplication in Mandarin Chinese, the structure should be something like Figure 2.[9,10] Further, A-*le*-A can be handled as two compositional

[8] P = phonological unit, P/E/S CTR = prototypical/extreme/salient contrast, XP^{min} = XP without its specifier

[9] DELIM = delimitative

[10] Although the reduplication in Mandarin Chinese does not have a contrastive meaning, we preserved the notation of CR^{syn} in Ghomeshi et al. (2004) here to simply refer to the reduplicant. In English, it makes sense to assume CR^{syn} to be a syntactic unit, because the base can be XP^{min}. But for Mandarin Chinese, the base can only be V. As Ghomeshi et al. (2004, 353) wrote: "when applying to its

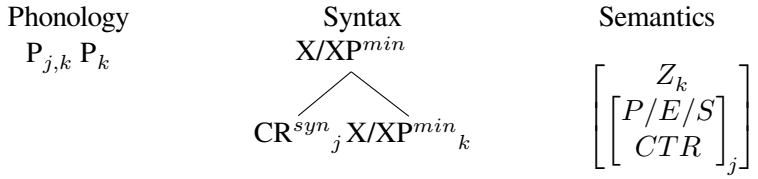

Figure 1: Analysis for CRs in English according to Ghomeshi et al. (2004, 344)

processes [[[A]-*le*] -A]. Moreover, the *yi* in A-*yi*-A and A-*le*-*yi*-A can simply be viewed as a dangling phonological unit. In this case, the phonological unit ⟨ *yi* ⟩ is neither coindexed with a syntactic unit nor with a semantic one.

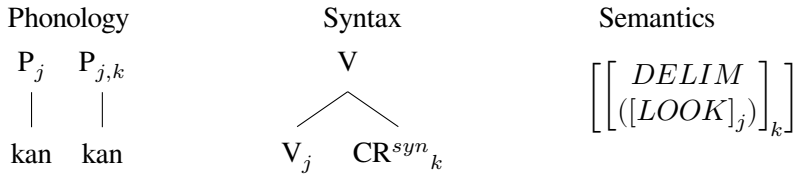

Figure 2: Analysis for AA following Ghomeshi et al. (2004)

This analysis correctly captures the fact that the addition of *yi* does not change the syntactic and semantic behaviors of the reduplication. It also provides a formal account for the phonology of the reduplication. On the other hand, by assuming a construction specially for the reduplication, Ghomeshi et al. (2004)'s approach lost the connection between the reduplication and other aspect markers in Mandarin Chinese, unlike the affixation analysis.

Finally, Fan et al. (2015) provided a unified HPSG analysis for the reduplication of both verbs and adjectives in Mandarin Chinese. They considered reduplication to be a morphological process and modeled the reduplication via lexical rules. They regarded the reduplication to function as an intensifier predicate, which has the subtypes *redup_up_x_rel* and *redup_down_x_rel*. They provided the lexical rule (12) for reduplication in general, and further proposed *redup-a-lr* and *redup-v-lr* as subtypes of *redup-type*, as illustrated in (13) and (14) respectively. The orthography is handled separately. The AABB form for adjectives and the ABAB form for verbs, as well as the AAB form for V-O compounds are handled as irregular derivation forms.

smallest scope, X inside of a word, it has the feel of other things that attach there, i.e., morphological affixes". It seems that it suffices to assume the reduplication in Mandarin Chinese to be a morphological phenomenon (cf. Section 2.3). We continue to call the second column "syntax" to preserve the consistency of the notations.

(12)
$$\left[\begin{array}{ll} \textit{redup-type} & \\ \textsc{cat}|\textsc{head} & \boxed{1} \\ \textsc{val} & \boxed{2} \\ \textsc{cont} & \boxed{3}\ \textsc{hook}\ \left[\begin{array}{ll}\textsc{ltop} & \boxed{4}\\ \textsc{ind} & \boxed{5}\end{array}\right] \\ \textsc{c-cont} & \left\langle \left[\begin{array}{ll}\textit{event-rel} & \\ \textsc{pred} & \textit{intensifier\_x\_rel}\\ \textsc{lbl} & \boxed{4}\\ \textsc{arg}1 & \boxed{5}\end{array}\right] \right\rangle \end{array}\right] \mapsto \left[\begin{array}{ll}\textsc{cat}|\textsc{head} & \boxed{1}\\ \textsc{val} & \boxed{2}\\ \textsc{cont} & \boxed{3}\end{array}\right]$$

(13)
$$\left[\begin{array}{ll} \textit{redup-a-lr} \subset \textit{redup-type} & \\ \textsc{cat}|\textsc{head} & \textit{adjective} \\ \textsc{val} & \left[\textsc{spr}\ \langle\rangle\right] \\ \textsc{c-cont} & \left\langle \left[\textsc{pred}\ \textit{redup\_up\_x\_rel}\right] \right\rangle \end{array}\right]$$
ORTHOGRAPHY: A → AA; (irregular AB → AABB)

(14)
$$\left[\begin{array}{ll} \textit{redup-v-lr} \subset \textit{redup-type} & \\ \textsc{cat}|\textsc{head} & \textit{verb} \\ \textsc{cont}|\textsc{hook} & \left[\textsc{aspect}\ \textit{non-aspect}\right] \\ \textsc{c-cont} & \left\langle \left[\textsc{pred}\ \textit{redup\_down\_x\_rel}\right] \right\rangle \end{array}\right]$$
ORTHOGRAPHY: A → AA; A → A-*yi*-A; (irregular AB → ABAB)

This approach provided a unified account for adjectival and verbal reduplication. Their commonalities are captured by inheritance hierarchies of the intensifier predicates and the lexical rules. In the case of verbal reduplication, A-*yi*-A is analyzed as an alternative orthographical form of AA. This correctly captured the intuition that AA and A-*yi*-A express the same meaning and only differ from each other phonologically/orthographically.

Nevertheless, this analysis has some shortcomings. To begin with, since the combination with aspect markers is completely forbidden, it is impossible for this approach to account for A-*le*-A. Moreover, as verbal reduplication is considered to express a delimitative aspectual meaning, it seems unconvincing to assume that there is no aspect information in its semantics. We consider a semantic explanation as described in Section 2.2 to be more reasonable for ruling out aspect markers other than *le*. Furthermore, this account can only deal with monosyllabic reduplication and handles ABAB and AAB as irregular forms, for the reason that ABAB and AAB reduplication of AB verbs "are not very productive in Chinese" (Fan et al., 2015, 102). This is not true. Basciano & Melloni (2017), Melloni & Basciano (2018), Xie (2020) and H. Xing (2000) all considered both AA and ABAB to be productive, and H. Xing (2000) concluded that AAB is productive as well. Therefore, ABAB and AAB should not be handled as irregular forms, but should be derivable from lexical rules.

The shortcomings of previous analyses lead us to propose a new analysis on verbal reduplication with HPSG, that formalizes its phonology, resolves the problem of *yi* and preserves the generalization on aspect marking, as we will elaborate in Section 4.

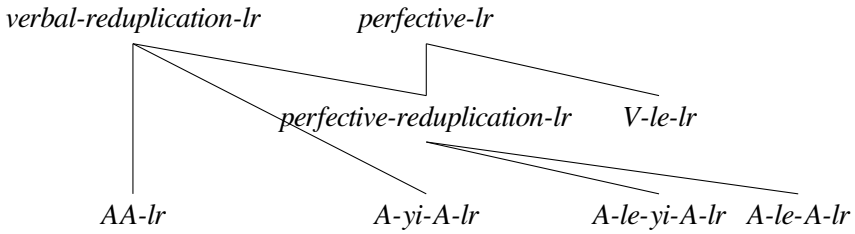

Figure 3: Type hierarchy for lexical rules of verbal reduplication and *le*

4 A new HPSG analysis

The lexical rule in (15) shows the general type of *verbal-reduplication-lr* for Mandarin Chinese. It takes a verb as LEX-DTR. The output reduplicates the phonology of the input verb with the possibility to have further phonological materials in between. □ indicates an underspecified list which could be empty or not. A delimitative relation is appended to the RELS value of the input verb and it takes the event index of the input verb as argument. Further relations can be added at the beginning of the RELS list to allow for the additional perfective meaning in A-*le*-A and A-*le*-*yi*-A. The combination with perfective will be elaborated in the following paragraphs.

(15)
$$\left[\begin{array}{ll} \textit{verbal-reduplication-lr} & \\ \textsc{phon} & \boxed{1} \oplus \square \oplus \boxed{1} \\ \textsc{rels} & \square \oplus \boxed{2} \oplus \left\langle \left[\begin{array}{l} \textit{delimitative-rel} \\ \textsc{arg}\ \boxed{3} \end{array}\right] \right\rangle \\ \textsc{lex-dtr} & \left[\begin{array}{ll} \textsc{phon}\ \boxed{1} & \\ \textsc{synsem}|\textsc{loc} & \left[\begin{array}{ll} \textsc{cat}|\textsc{head} & \textit{verb} \\ \textsc{cont}|\textsc{ind} & \boxed{3} \end{array}\right] \\ \textsc{rels}\ \boxed{2} & \end{array}\right] \end{array}\right]$$

To account for the variations in the phonology of the reduplication as well as the combination with the phonology and semantics of the perfective aspect marker *le*, the type hierarchy of lexical rules in Figure 3 is put forward. *verbal-reduplication-lr* has *AA-lr* and *A-yi-A-lr* as direct subtypes. *V-le-lr* is a direct subtype of the *perfective-lr*. *perfective-reduplication-lr* inherits from both *verbal-reduplication-lr* and *perfective-lr*, and has two subtypes *A-le-yi-A-lr* and *A-le-A-lr* itself. *verbal-redupliaction-lr* is already presented in (15). The other lexical rules in this type hierarchy will be illustrated in detail below.

(16) and (17) show *AA-lr* and *A-yi-A-lr*, respectively. As subtypes of *verbal-reduplication-lr* illustrated in (15), both inherit the constraints on the LEX-DTR and on the semantics of the output, and further specify that in the RELS list of the output, nothing else comes before the *delimitative-rel* and the RELS value of the input verb. In addition, they specify the phonology of the output differently. *AA-lr* determines that the □ between the two phonological copies is empty, whereas *A-yi-A-lr* specifies this possible list of phonological material as ⟨ *yi* ⟩.

(16)
$$\begin{bmatrix} \textit{AA-lr} \\ \textsc{phon} & \boxed{1} \oplus \boxed{1} \\ \textsc{rels} & \boxed{2} \oplus \left\langle \begin{bmatrix} \textit{delimitative-rel} \\ \textsc{arg} \ \boxed{3} \end{bmatrix} \right\rangle \\ \textsc{lex-dtr} & \begin{bmatrix} \textsc{phon} \ \boxed{1} \\ \textsc{synsem|loc} & \begin{bmatrix} \textsc{cat|head} & \textit{verb} \\ \textsc{cont|ind} & \boxed{3} \end{bmatrix} \\ \textsc{rels} \ \boxed{2} \end{bmatrix} \end{bmatrix}$$

(17)
$$\begin{bmatrix} \textit{A-yi-A-lr} \\ \textsc{phon} & \boxed{1} \oplus \langle\ \textit{yi}\ \rangle \oplus \boxed{1} \\ \textsc{rels} & \boxed{2} \oplus \left\langle \begin{bmatrix} \textit{delimitative-rel} \\ \textsc{arg} \ \boxed{3} \end{bmatrix} \right\rangle \\ \textsc{lex-dtr} & \begin{bmatrix} \textsc{phon} \ \boxed{1} \\ \textsc{synsem|loc} & \begin{bmatrix} \textsc{cat|head} & \textit{verb} \\ \textsc{cont|ind} & \boxed{3} \end{bmatrix} \\ \textsc{rels} \ \boxed{2} \end{bmatrix} \end{bmatrix}$$

Müller & Lipenkova (2013, 246) proposed the perfective lexical rule demonstrated in (18), adapted to the formalization adopted in the current paper. It takes a verb as LEX-DTR and appends ⟨ *le* ⟩ to its phonology. Further, it accounts for the change in semantics by appending the RELS value of the input verb to a *prefective-rel*.

(18) Perfective lexical rule adapted from Müller & Lipenkova (2013, 246):
$$\begin{bmatrix} \textit{perfective-lr} \\ \textsc{phon} & \boxed{1} \oplus \langle\ \textit{le}\ \rangle \\ \textsc{rels} & \left\langle \begin{bmatrix} \textit{perfective-rel} \\ \textsc{arg} \ \boxed{3} \end{bmatrix} \right\rangle \oplus \boxed{2} \\ \textsc{lex-dtr} & \begin{bmatrix} \textsc{phon} \ \boxed{1} \\ \textsc{synsem|loc} & \begin{bmatrix} \textsc{cat|head} & \textit{verb} \\ \textsc{cont|ind} & \boxed{3} \end{bmatrix} \\ \textsc{rels} \ \boxed{2} \end{bmatrix} \end{bmatrix}$$

The lexical rule suggested in (18) only explains simple perfective aspect marking with *le*, where *le* immediately follows the verb. But it cannot account for the perfective aspect marking of reduplicated verb, as *le* does not occur after the reduplication, nor can *le* be reduplicated together with the verb. It can only appear between the verb and the reduplicant. In order to accommodate *le* marking for both simple and reduplicated verbs, a general perfective lexical rule as in (19) and a subtype *V-le-lr* as in (20) are posited here. Besides adding a *perfective-rel* in the RELS list of the output as in (18), the *perfective-lr* in (19) allows an underspecified list to be append at the end of the RELS list. The PHON value of the output makes it possible for further phonological material to occur both before and after ⟨ *le* ⟩.

(19) Perfective lexical rule improved:

$$\begin{bmatrix} \textit{perfective-lr} \\ \textsc{phon} \quad \square \oplus \langle\, \textit{le} \,\rangle \oplus \square \\ \textsc{rels} \quad \left\langle \begin{bmatrix} \textit{perfective-rel} \\ \textsc{arg} \ \boxed{3} \end{bmatrix} \right\rangle \oplus \boxed{2} \oplus \square \\ \textsc{lex-dtr} \quad \begin{bmatrix} \textsc{phon} \ \boxed{1} \\ \textsc{synsem}|\textsc{loc} \begin{bmatrix} \textsc{cat}|\textsc{head} \ \textit{verb} \\ \textsc{cont}|\textsc{ind} \ \boxed{3} \end{bmatrix} \\ \textsc{rels} \ \boxed{2} \end{bmatrix} \end{bmatrix}$$

V-le-lr as in (20) inherits from *perfective-lr* and specifies that the first element in the output PHON list is identified with the PHON value of the input verb and nothing else comes after ⟨ *le* ⟩. Furthermore, no other list can be appended at the end of the RELS list of the output anymore. This corresponds to the proposal of Müller & Lipenkova (2013, 246) shown in (18), which accounts for the simple perfective marking of verbs.

(20) *V-le-lr* as a subtype of *perfective-lr*:

$$\begin{bmatrix} \textit{V-le-lr} \\ \textsc{phon} \quad \boxed{1} \oplus \langle\, \textit{le} \,\rangle \\ \textsc{rels} \quad \left\langle \begin{bmatrix} \textit{perfective-rel} \\ \textsc{arg} \ \boxed{3} \end{bmatrix} \right\rangle \oplus \boxed{2} \\ \textsc{lex-dtr} \quad \begin{bmatrix} \textsc{phon} \ \boxed{1} \\ \textsc{synsem}|\textsc{loc} \begin{bmatrix} \textsc{cat}|\textsc{head} \ \textit{verb} \\ \textsc{cont}|\textsc{ind} \ \boxed{3} \end{bmatrix} \\ \textsc{rels} \ \boxed{2} \end{bmatrix} \end{bmatrix}$$

perfective-reduplication-lr inherits from both *verbal-reduplication-lr* and *perfective-lr*. The PHON value of the output reduplicates the phonology of the input verb and states that there is ⟨ *le* ⟩ in between and potentially further phonological material. The RELS list of the output appends the *delimitative-rel* to the *perfective-rel* and the RELS value of the input verb. The arguments of both *perfective-rel* and *delimitative-rel* share the event index of the input verb to ensure that they apply to the same event denoted by the input verb.

(21)
$$\begin{bmatrix} \textit{perfective-reduplication-lr} \\ \textsc{phon} \quad \boxed{1} \oplus \langle\, \textit{le} \,\rangle \oplus \square \oplus \boxed{1} \\ \textsc{rels} \quad \left\langle \begin{bmatrix} \textit{perfective-rel} \\ \textsc{arg} \ \boxed{3} \end{bmatrix} \right\rangle \oplus \boxed{2} \oplus \left\langle \begin{bmatrix} \textit{delimitative-rel} \\ \textsc{arg} \ \boxed{3} \end{bmatrix} \right\rangle \\ \textsc{lex-dtr} \quad \begin{bmatrix} \textsc{phon} \ \boxed{1} \\ \textsc{synsem}|\textsc{loc} \begin{bmatrix} \textsc{cat}|\textsc{head} \ \textit{verb} \\ \textsc{cont}|\textsc{ind} \ \boxed{3} \end{bmatrix} \\ \textsc{rels} \ \boxed{2} \end{bmatrix} \end{bmatrix}$$

Two subtypes of *perfective-reduplication-lr* are posited: *A-le-yi-A-lr* as in (22) and *A-le-A-lr* as in (23). They take over the semantic change to the input from *perfective-reduplication-lr*, but specifies the PHON value differently. Specifically, *A-le-yi-A-lr* specifies the middle phonological material with ⟨ *le, yi* ⟩, while *A-le-A* only with ⟨ *le* ⟩.

(22)
$$\begin{bmatrix} \textit{A-le-yi-A-lr} \\ \textsc{phon} \quad \boxed{1} \oplus \langle\, \textit{le, yi} \,\rangle \oplus \boxed{1} \\ \textsc{lex-dtr} \quad \left[\textsc{phon} \ \boxed{1}\right] \end{bmatrix}$$

(23)
$$\begin{bmatrix} \textit{A-le-A-lr} \\ \textsc{phon} \quad \boxed{1} \oplus \langle\, \textit{le} \,\rangle \oplus \boxed{1} \\ \textsc{lex-dtr} \quad \left[\textsc{phon} \ \boxed{1}\right] \end{bmatrix}$$

Since the above described lexical rules do not constrain the number of syllables of the input verb, but simply reduplicate its phonology as a whole, they can also account for the ABAB and the AB-*le*-AB forms of reduplication, as long as the input verb is disyllabic. Notice that the lexical rules above also produce AB-*yi*-AB and AB-*le-yi*-AB for disyllabic input verbs. These forms, however, are generally considered unacceptable (Basciano & Melloni, 2017, 160, Hong, 1999, 275–276, Li & Thompson, 1981, 30, Yang & Wei, 2017, 239). On the other hand, Fan (1964, 269) and Sui (2018, 143) considered AB-*yi*-AB and AB-*le-yi*-AB to be possible, even though they both recognized that these two forms are rare. Indeed, a few examples of AB-*yi*-AB and AB-*le-yi*-AB in Early Mandarin (24a)–(24b) and Modern Mandarin (**??**)–(24e) were found.

(24) a. ni yu wo **zhengli-yi-zhengli**.[11]
you let me arrange-one-arrange
'Let me arrange it a little bit!"

b. ni **dating-yi-dating**.[12]
you inquire-one-inquire
'Inquire about it a little bit!"

c. ta **weixiao-le-yi-weixiao**, you **mingxiang-le-yi-mingxiang**.[13]
he smile-PFV-one-smile and meditate-PFV-one-meditate
'He smiled a little bit and meditated a little bit.'

d. feichang yansu de ba jinshi yanjing
very seriously DE BA nearsighted glasses
duanzheng-le-yi-duanzheng.[14]
straighten-PFV-one-straighten
'Very seriously put the nearsighted glasses straight quickly.'

e. er di-tou xiang jiang-xin yi kan, jiduo sanluan-zhe de
and bow-head towards river-middle one look many scattered-DUR DE
chuan li de dengguang, ye huyinhumie de **bianhuan-le-yi-bianhuan**
boat in DE light also flicker DE change-PFV-one-chang

[11] *Yuanqu xuan: Luzhailang [Selected Yuanqu: Luzhailang]*, as cited in Zhang (2000, 15)

[12] *Yuan Ming juan: Piaotongshi [Yuan and Ming volume: Piaotongshi]*, 308, as cited in Zhang (2000, 15)

[13] Rou, Shi. 1975. *Roushi xiaoshuo xuanji [Selected novels of Roushi]*, 31. Beijing: People's Literature Publishing House.

[14] Li, Jieren. 1962. *Da bo [Great wave]*, 3rd band, 171. Beijing: The Writers Publishing House.

weizhi. (CCL)
position

'And (I) bowed my head and took a look at the middle of the river. Many scattered lights in the boats also changed their positions a little bit, flickering.'

This suggests that even though AB-*yi*-AB and AB-*le*-*yi*-AB might be degraded, they are not ungrammatical *per se*. The reason for this degradedness is probably phonological, namely AB-*yi*-AB and AB-*le*-*yi*-AB contain too many syllables (Fan, 1964, 274, Sui, 2018, 143, Yang & Wei, 2017, 239, Zhang, 2000, 15). We argue that the reduced acceptability of AB-*yi*-AB and AB-*le*-*yi*-AB in Modern Mandarin Chinese is due to their phonological length and not their grammaticality. Thus, they can still be produced via the lexical rules posited above, but are ruled out or degraded due to a general phonological constraint.

AAB, A-*yi*-AB, A-*le*-AB, AA-*kan* and A-*kan*-*kan* can also be accounted for by the lexical rules proposed in this section. They can be analyzed as a compounds out of a reduplicated monosyllabic verb and another element. Specifically, AAB, A-*yi*-AB and A-*le*-AB can be considered as the compound of a reduplicated monosyllabic verb (A) and a noun (B).[15] AA-*kan* can be regarded as the compound of a reduplicated monosyllabic verb (A) and the verb *kan* 'look', whereas A-*kan*-*kan* is the compound of a monosyllabic verb (A) and the reduplication of *kan* 'look'. A-*yi*-A-*kan* is also possible, although rare, presumably due to its length as well. An inquiry in CCL found 55 hits of A-*yi*-A-*kan*. A sample is listed in (25).

(25) a. danshi dui fa mei fa-guo hege-zheng, yijing shuo bu
but about issue not issue-EXP conformity-certificate already say not
qing le, xuyao **cha-yi-cha-kan**. (CCL)
clealy PTC need check-one-chek-look

'But one already cannot say it clearly anymore, whether a certificate of conformity is issued or not. One needs to have a check and see.'

b. da-laoban-men yao **deng-yi-deng-kan** (CCL)
big-boss-PL need wait-one-wait-look

'Big bosses need to wait a little bit and see.'

c. furen ni dao **shu-yi-shu-kan**, zhe zhu hua de huaduo
madam you just count-one-count-look this CLF flower DE blossom
gong you ji zhong yanse. (CCL)
in.total have how.many CLF color

'Madam, just try to count and see how many color does the blossom of this flower have in total.'

[15] Huang (1984) and Her (1996, 2010) argued that some of this kind of structures are compounds, some are phrases, and some have dual status (both compounds and phrases). Following this approach, AAB, A-*yi*-AB and A-*le*-AB can (also) be considered as the phrasal combination of a reduplicated verb and its object.

Due to the prominent tentative, trying meaning of AA-*kan* and A-*kan*-*kan*, they are not compatible with the perfective aspect marker *le* semantically, as one usually cannot try something that is already realized. Unacceptable structures such as A-*le*-A-*kan* and A-*kan*-*le*-*kan* are thus semantically ruled out.

The current analysis provides a unified account for all forms of delimitative verbal reduplication in Mandarin Chinese. Like in Fan et al. (2015), *yi* is handled as a phonological element which does not make any contribution to the semantics, and inheritance hierarchy is used to capture the commonalities among different forms of reduplication. But the present proposal also reflects the connection between the reduplication and aspect marking via multiple inheritance. This account makes use of a semantic mechanism, which correctly rules out the aspect marking other than *le*. By providing a semantic explanation, this mechanism seems less *ad hoc* than the one used in Fan et al. (2015), which simply assumed that the reduplication cannot combine with aspect information. The present approach also has a broader coverage of the forms of verbal reduplication than Fan et al. (2015). Furthermore, all the forms are derivable from the lexical rules proposed here, so that there is no need to resort to irregular lexicon entries, and the productivity of these forms are correctly captured. In sum, the analysis proposed in this paper seems to possess greater explanatory power and resolves the problems in previous studies.

5 Conclusion

The current study provided an HPSG account for verbal reduplication in Mandarin Chinese. We presented empirical evidence that the reduplication is possible with all *Aktionsarten*. We gave a semantic explanation for the incompatibility of the reduplication with aspect markers other than *le*. We argued that the reduplication is a morphological rather than syntactic process. We modeled the reduplication as a lexical rule and the different forms of reduplication are captured in an inheritance hierarchy using underspecified lists. The interaction between verbal reduplication and aspect marking is handled by multiple inheritance. This analysis is compatible with both mono- and disyllabic verbs, so that all productive forms of reduplication are derivable from lexical rules.

References

Arcodia, Giorgio Francesco, Bianca Basciano & Chiara Melloni. 2014. Verbal reduplication in Sinitic. In Sandra Augendre, Graziella Couasnon-Torlois, Déborah Lebon, Michard Clément, Gilles Boyé & Fabio Montermini (eds.), *Proceedings of the Décembrettes 8th international conference on morphology*, 15–45. Bordeaux: CLLE-ERSS. https://iris.unive.it/retrieve/handle/10278/39972/85072/carnGram22.23-53.pdf.

Basciano, Bianca & Chiara Melloni. 2017. Event delimitation in Mandarin: The

case of diminishing reduplication. *Italian Journal of Linguistics* 29(1). 143–166. doi:10.26346/1120-2726-106.

Chao, Yuen Ren. 1968. *A grammar of spoken Chinese*. Berkeley, CA: University of California Press.

Cheng, Yi-Yang. 2012. Verbal reduplication and grammaticalization: A study of Mandarin VV-kan and V-kankan constructions. *Annual Meeting of the Berkeley Linguistics Society* 38. 63–77. doi:10.3765/bls.v38i0.3321.

Chen, Limin. 2005. Lun dongci chongdie de yufa yiyi [On the grammatical meaning of verbal reduplication]. *Zhongguo yuwen [Studies of the Chinese Language]* 305(2). 110–191.

Chen, Qianrui. 2001. Dongci chongdie de qingzhuang tezheng ji qi ti de diwei [The situational features of Chinese verbal reduplication and its aspectual status]. *Yuyan jiaoxue yu yanjiu [Language Teaching and Linguistic Studies]* (4). 48–56.

Dai, Yaojing. 1997. *Xiandai hanyu shiti xitong yanjiu [Studies on the tense and aspect system of Modern Chinese]*. Hangzhou: Zhejiang Education Publishing Group.

Duanmu, San. 1998. Wordhood in Chinese. In Jerome L. Packard (ed.), *New approaches to Chinese word formations*, 135–196. Berlin: De Gruyter Mouton. doi:10.1515/9783110809084.135.

Fan, Fanglian. 1964. Shi lun shuowei "dongci chongdie" [A tentative analysis of the so-called "verbal reduplication"]. *Zhongguo Yuwen [Studies of the Chinese Language]* 4. 264–278.

Fan, Zhenzhen, Sanghoun Song & Francis Bond. 2015. Building Zhong, a Chinese HPSG Shared-Grammar. In Stefan Müller (ed.), *Proceedings of the 22nd International Conference on Head-Driven Phrase Structure Grammar, Nanyang Technological University (NTU), Singapore*, 96–109. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2015/fsb.pdf.

Ghomeshi, Jila, Ray Jackendoff, Nicole Rosen & Kevin Russell. 2004. Contrastive focus reduplication in English (the salad-salad paper). *Natural Language & Linguistic Theory* 22(2). 307–357. doi:10.1023/B:NALA.0000015789.98638.f9.

Her, One-Soon. 1996. Variation of the VO construction in Chinese: A synchronic account. *Linguistics* 34(4). 733–751. doi:10.1515/ling.1996.34.4.733.

Her, One-Soon. 2010. *Interaction and variation in the Chinese VO construction*. Taipei: Crane Publishing revised edn. edn. http://www3.nccu.edu.tw/$\sim$osh/publication/D07-Interaction-REV.pdf.

Hong, Zhou. 1999. Cognate objects in Chinese. *Toronto Working Papers in Linguistics* 17. 263–284. https://twpl.library.utoronto.ca/index.php/twpl/article/view/6265.

Huang, C.-T. James. 1984. Phrase structure, lexical integrity, and Chinese compounds. *Journal of Chinese Language Teachers Association* 12(2). 53–78. https://scholar.harvard.edu/files/ctjhuang/files/1984.lexical.integrity.pdf?m=1443058311.

Huang, C.-T. James, Y.-H. Audrey Li & Yafei Li. 2009. *The syntax of Chinese* Cambridge Syntax Guides. Cambridge: Cambridge University Press. doi:10.1017/CBO9781139166935.

Jackendoff, Ray. 1997. *The architecture of the language faculty* Linguistic Inquiry Monographs. Cambridge, MA: MIT Press.

Jackendoff, Ray. 2002. *Foundations of language: Brain, meaning, grammar, evolution*. Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780198270126.001.0001.

Li, Charles N. & Sandra A. Thompson. 1981. *Mandarin Chinese: A functional reference grammar*. Berkeley, CA: University of California Press.

Li, Yuming. 1996. Lun ciyu chongdie de yiyi [On the meaning of reduplication]. *Shijie hanyu jiaoxue [Chinese Teaching in the World]* 35(1). 10–19.

Li, Yuming. 1998. Dongci chongdie de ruogan jufa wenti [Some syntactic issues on verbal reduplication]. *Zhongguo Yuwen [Studies of the Chinese Language]* 2. 83–92.

Melloni, Chiara & Bianca Basciano. 2018. Reduplication across boundaries: The case of Mandarin. In Olivier Bonami, Gilles Boyé, Georgette Dal, Hélène Giraudo & Fiammetta Namer (eds.), *The lexeme in descriptive and theoretical morphology*, 325–363. Berlin: Language Science Press. doi:10.5281/zenodo.1407013.

Müller, Stefan & Janna Lipenkova. 2013. ChinGram: A TRALE implementation of an HPSG fragment of Mandarin Chinese. In *Proceedings of the 27th Pacific Asia Conference on Language, Information, and Computation (PACLIC 27)*, 240–249. Taipei: Department of English, National Chengchi University. https://www.aclweb.org/anthology/Y13-1023.pdf.

Paris, Marie-Claude. 2013. Verbal reduplication and verbal classifiers in Chinese. In Guangshun Cao, Hilary Chappell, Redouane Djamouri & Thekla Wiebusch (eds.), *Breaking down the barriers: Interdisciplinary studies in Chinese linguistics and beyond*, vol. 1, 257–278. Taipei, Taiwan: Institute of Linguistics, Academia Sinica. http://www.ling.sinica.edu.tw/Files/LL/Docments/Monographs/Breaking%20Down%20the%20Barriers/Volume%201/12-Paris.pdf.

Schäfer, Martin. 2009. A+N constructions in Mandarin and the 'compound vs. phrase' debate. *Word Structure* 2(2). 272–293. doi:10.3366/E1750124509000452.

Sui, Yanyan. 2018. Affixation or compounding? Reduplication in Standard Chinese. In Rita Finkbeiner & Ulrike Freywald (eds.), *Exact repetition in grammar and discourse* (Trends in Linguistics, Studies and Monographs (TiLSM) 323), 127–157. De Gruyter Mouton. doi:10.1515/9783110592498-006.

Travis, Lisa deMena. 1999. A syntactician's view of reduplication. In Carolyn Smallwood & Catherine Kitto (eds.), *Proceedings of AFLA VI (the sixth meeting of the Austronesian Formal Linguistics Association)*, 313–331. Toronto: Department of Linguistics, University of Toronto. https://twpl.library.utoronto.ca/index.php/twpl/article/view/6295.

Travis, Lisa deMena. 2000. Event structure in syntax. In Carol L. Tenny & James Pustejovsky (eds.), *Events as grammatical objects: The converging perspectives of lexical semantics and syntax* (CSLI Lecture Notes 100), 145–185. Stanford, CA: CSLI Publications.

Tsao, Feng-fu. 2001. Semantics and syntax of verbal and adjectival reduplication in

Mandarin and Taiwanese Southern Min. In Hilary Chappell (ed.), *Sinitic grammar: Synchronic and diachronic perspectives*, 285–308. Oxford University Press.

Xiao, Richard & Tony McEnery. 2004. *Aspect in Mandarin Chinese: A corpus-based study* (Studies in Language Companion Series 73). Amsterdam: John Benjamins. doi:10.1075/slcs.73.

Xie, Zhu. 2020. Two types of verb reduplications in Mandarin Chinese. *Studies in Chinese Linguistics* 41(1). 73–108. doi:10.2478/scl-2020-0003.

Xing, Hongbing. 2000. Hanyu ciyu chongdie jiegou tongji fenxi [Statistical analysis on reduplication Modern Chinese words]. *Yuyan Jiaoxue yu Yanjiu [Language Teaching and Linguistic Studies]* (1). 32–37.

Xiong, Zhongru. 2016. Dongci chongdie de jufa fenxi [A syntactic analysis of verbal reduplication]. *Shijie Hanyu Jiaoxue [Chinese Teaching in the World]* 30(2). 156–169. doi:10.13724/j.cnki.ctiw.2016.02.002.

Yang, Ping. 2003. Dongci chongdie de jiben yiyi [The basic function of verb reduplication in Chinese Language]. *Yuyan Jiaoxue yu Yanjiu [Language Teaching and Linguistic Studies]* (5). 8–16.

Yang, Yifan & Wei Wei. 2017. Verbal reduplication in Mandarin Chinese: An analysis at the syntax-phonology interface. In Michael Yoshitaka Erlewine (ed.), *Proceedings of GLOW in Asia XI*, vol. 1, 227–242. Cambridge, MA: MIT Working Papers in Linguistics. https://glowlinguistics.org/asia11/wp-content/uploads/sites/3/2017/11/v1-yang-wei.pdf.

Zhan, Weidong, Rui Guo, Baobao Chang, Yirong Chen & Chen Long. 2019. Beijing Daxue CCL yuliaoku de yanzhi [the building of the CCL corpus: Its design and implementation]. *Yuliaoku Yuyanxue [Corpus Linguistics]* 6(1). 71–86.

Zhan, Weidong, Rui Guo & Yirong Chen. 2003. Beijing Daxue Zhongguo Yuyanxue Yanjiu Zhongxin CCL yuliaoku (Guimo: 7 yi zi; shijian: gongyuanqian 11 shiji - dangdai)[the CCL corpus of Chinese texts: 700 million Chinese characters, the 11th Century B.C. - present]. http://ccl.pku.edu.cn:8080/ccl_corpus.

Zhang, Cheng. 2000. Xiandai hanyu "VyiV" shi he "VV" shi de laiyuan [The origin of "V-yi-V" and "VV" form in modern Chinese]. *Yuyan Jiaoxue yu Yanjiu [Language Teaching and Linguistic Studies]* 4. 10–17.

Zhu, Jingsong. 1998. Dongci chongdieshi de yufa yiyi [The semantics of verbal reduplication]. *Zhongguo Yuwen [Studies of the Chinese Language]* 266(5). 378–386.