

A smurf-based analysis of placeholder expressions

Manfred Sailer & Annika Dörner

Languages have expressions which are used in place of other expressions. In English, these include *whatsit*, *whatchamacallit*, or *thingamajig* for common nouns and *What's-his/her-name* or *you-know-who* for proper nouns. Such *placeholder expressions* (PHEs) are used under various circumstances (Cheung, 2015): (i) when the speaker cannot think of the correct expression right away, (ii) to fill a lexical gap temporarily, (iii) as euphemism when the speaker wants to avoid a certain term, (iv) as an in-group term to mark solidarity among the communication partners. All of these uses share the following properties: (PHE-Ref) The PHE refers to a linguistic sign, i.e., some form-meaning pairing. (PHE-Truth) The truth-conditional meaning of a PHE is not fixed but is identical to that of the expression it replaces. (PHE-Constr) There can be morpho-syntactic and semantico-pragmatic constraints on which expressions a given PHE can replace. (PHE-Use) PHEs have a use-conditional meaning, which can be more or less flexible.

In this paper, we will look at *smurf*-expressions, or *smurfemes*, as they are used in the German translation of the Smurf comics series. The Smurf comics were created by the Belgian cartoonist Peyo and have been published since 1958, originally in French but translated into 25 languages (wikipedia). While we will concentrate on the German Smurf translations, we will also relate our research to studies on the English, Modern Greek, French, and Spanish versions (Bollig, 2016; Bourcier & Martin, 1996; Chatzopoulos, 2008).

The main motivation for our research is that we consider *smurfemes* as a particularly productive type of PHEs. Peyo allegedly used *schtroumpf* 'smurf' the first time during a dinner with a friend when he forgot the word *sel* 'salt', asking his friend to pass him the *schtroumpf*. The friend took this up saying "There's the Schtroumpf – when you are done schtroumpfing, schtroumpf it back ..." (wikipedia). This creation myth of the Smurf language contains two essential aspects of PHEs: first, the use of a *smurfeme* as a placeholder for a temporarily unavailable expression, and, second, the solidarity creating function. No attempt has been made in the literature to relate the origin of *smurfing* as a PHE to its linguistic analysis. This paper aspires to do exactly this.

We will first discuss properties and analyses of PHEs and then turn to *smurfing*. We will propose a constraint-based modelling of PHEs (including *smurfing*) as compound formation. We will show how this captures the observed properties of PHEs in general and *smurfing* in German particular.

Previous approaches to PHEs

Enfield (2003) analyses English PHEs as having a very general truth-conditional semantics, like 'something.' In addition, each PHE comes with constraints on the conditions under which a speaker can use it. Enfield's analysis is given in (1).

- (1) *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.

In this approach, PHEs may differ with respect to their truth-conditional semantics in (1-a) (thing, person, ...), but also with respect to their use-conditions, (1-b)–(1-d).

We agree with Enfield's usage characterization of PHEs and think that they show all the properties of typical *use-conditional meanings* or *speaker meanings* (Gutzmann, 2013). Our major criticism of Enfield's analysis concerns his truth-conditional meaning of PHEs: we cannot replace PHEs with a general expression in all cases. In particular, PHEs can be used to replace parts of idioms and collocations. In (2), the German PHE *Dingsbums* replaces the word *Barthel*, which is a bound word restricted to the idiom *zeigen/ wissen, wo Barthel den Most holt* 'show/ know what's going on' (lit.: know/ show where ?? gets the cider).

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

Whatever one's analysis of such idioms looks like, the semantics of the PHE in (2) cannot be 'something.' Rather, the PHE must share enough properties with the bound word *Barthel* to legitimately occur in (2).

Cheung (2015) discusses Chinese *wh*-PHEs. He is less precise than Enfield (2003) about the use conditions of PHE, but he overcomes the just-mentioned problem. Cheung analyzes this group of PHEs as *metalinguistic demonstratives*. This means that they refer to an intended linguistic expression – for which Cheung uses the semantic type *u* (for *utterance*), borrowed from the literature on quotation. He requires a shifting operation that maps the denoted linguistic expression to its meaning. Consequently, the PHE is analyzed as a real placeholder of the substitute linguistic expression.

Cheung’s analysis is, however, less convincing for cases of pure phonological replacement as in (3). Here, the PHE *shenme* replaces the second syllable of the Mandarin rendering of the French name *Hollande*, which is *Ao-lang-de* (Cheung, 2015, 276).

- (3) Ao-shenme-de shi xianren Faguo zongtong.
Ho-llan-de be current France president
‘Ao-something-de is the current President of F.’

Cheung assumes that the PHE (3) is a DP that refers to the syllable *lang*. As this syllable has no semantic interpretation, it is just inserted into the further combination of the structure of the name *Ao-lang-de* ‘Hollande’. However, the name does not plausibly have an internal syntactic structure. It is, thus, far from clear how a PHE-DP can occur structurally inside the name.

Properties of smurfing

The present paper is based on an inspection of 6 German smurf comics with a total of 536 instances of smurfing. In addition, we used examples from the literature on smurfing. Consequently, our study is not intended as a quantitative study on smurfing, but rather as a data-based theoretical modelling. Chatzopoulos (2008) distinguishes between a *literal* use of *smurf*, referring to Smurfs or things related to them (*smurf-soufflé*), and a *semantically unspecified* use – in which, according to her, the smurfeme “can take various meanings or remain empty.” *Smurfing* refers to the non-literal use of smurfemes, which we will consider a PHE-use.

Smurfing has a number of aspects: (i) its morpho-syntactic properties, (ii) its phonological properties, (iii) its conversational or social function, and (iv) the recoverability of the smurfed expression,

Morpho-syntax of smurfing: In all languages discussed in the literature, smurfemes are used as morphemes of major parts of speech (POS), where the POS is usually unambiguously recoverable – see Bourcier & Martin (1996) for French. For words which contain a smurfeme as their only (visible) root, our German data shows a tendency that verbal uses of *schlumpf* are usually directly inflected, nominal uses are more likely combined with derivational affixes that indicate noun-hood, and adjectival uses are often marked as *schlumpf-ig* ‘smurf-y’, though the plain form, *schlumpf*, occurs as well. The same holds for English *smurf*-adjectives, which often occur with the affix *-y*.

In German, the smurfeme not only inherits its POS from the expression it replaces, but also its gender (for nouns) and its choice of the perfect auxiliary *haben/sein* ‘have/be’ (for verbs). It does not, however, inherit the inflection class. The verbal form shows weak (or regular) inflection. The masculine and neuter nominal form forms its plural with an *umlaut* (*Schlumpf*.M/N.SG – *Schlümpfe*.M/N.PL), independently of the plural formation of the re-

placed noun. The feminine noun is *Schlumpf-e*, even if the replaced noun does not show a gender-specific affix, as in (4). The feminine plural is *Schlumpfe*, without *umlaut*, see (5), where the replaced noun has an *umlaut* plural.

- (4) Vorsicht ist die Schlumpf-e (=Mutter.F.SG) der Porzellan-kiste.
‘Caution is the smurf (=mother) of wisdom’
- (5) ... bin ich dabei, mir die Schlumpf-e (= Hände.F.PL) zu waschen
‘I am busy washing my smurfs (=hands)’

The morpho-syntactic behavior of the German smurfeme suggests that it is a masculine, *umlaut*-forming noun that can be converted into a verb and turned into a female noun or an adjective by (optionally) adding an appropriate affix (*-e* or *-ig*). This addition of POS-singalling derivational affixes is not expected in previous approaches, as those affixes are not necessarily part of the replaced word.

There is a straightforward explanation for this observation. We can start from a lexeme (or rather *flexeme*, Bonami & Crysmann (2018)) *Schlumpf* ‘smurf’ with its plural form *Schlümpfe*. The mentioned POS-marked forms all can be formed from this lexeme, such as the adjective forms *schlumpf* and *schlumpf-ig*, and the verb *schlumpf-en*. With the literal use of *Schlumpf*, these forms mean ‘Smurf-related’ or ‘act in a Smurf way’. With the PHE use of *Schlumpf*, they are just POS-marked forms.

German smurfing is not restricted to replacing a single morpheme but can cover bigger morphological units. The congratulation formula *Herzlichen Glückwunsch!* ‘heartly luck-wish’ is alternatively smurfed as *Herzlichen Glückschlumpf!* ‘heartly luck-smurf’ or *Herzlichen Schlumpf!*. In the second case, an entire compound is smurfed rather than just a morpheme. Thus, any morphological complex containing at least one root can undergo smurfing.

Phonology of smurfing: According to Chatzopoulos (2008), English *-smurf-* can replace one syllable, but may not interfere with the overall metrical properties of the smurfed word.

Our data rather suggest that we need to distinguish two types of smurfing: M(orphological) and p(honological) smurfing. In m-smurfing, a morphological unit is replaced with the POS-adjusted smurfeme. In p-smurfing, the phonology of the smurfeme replaces as syllable in a word. Greek forms like *strumfonía* (*sinfonía* ‘symphony’) and *strumfonáftis* (*astronáftis* ‘astronaut’) (Chatzopoulos, 2008) are p-smurfings. *Kataschlumpfe* (*Katas-trophe* ‘catastrophe’) and *schlumpfestieren* (*protes-tieren* ‘to protest’) are German examples.

P-smurfing differs from m-smurfing in three respects: M-smurfing replaces a morpheme (complex) and relies on the morphological structure (including the POS) of the word in which smurfing

occurs. P-smurfing depends on the phonological structure and replaces one syllable. M-smurfing can contain POS-marking affixes, p-smurfing is restricted to the string *schlumpf*. M-smurfing may change the inflectional paradigm, p-smurfing does not. This last point is illustrated with *Isoschlumpfe* (*Isotope.N.PL* ‘isotopes’). Note that we do not get the *umlaut*, **Isoschlümpfe*, which we would expect for the neuter plural in m-smurfing.

Note that this distinction applies to PHEs beyond smurfemes. We can treat the syllable replacement illustrated in (3), as p-replacement, whereas the other cases discussed in Cheung (2015) are instances of m-replacement.

Pragmatics of smurfing: Chatzopoulos (2008) considers smurfing a strong marker of Smurf identity. Our data confirm this. We computed the smurfing rate in the German translations of *A Smurf in the Air* and *The Smurfs and the Magic Flute*. The first story only has dialogues among Smurfs and contains smurfing in 33% of the frames. In the second story, there are conversations among Smurfs, among non-Smurfs, and of Smurfs talking with non-Smurfs. There is a smurfing in 68% of the frames when Smurfs talk among themselves, in 0% for conversations among non-Smurfs, and in 23% when Smurfs and non-Smurfs talk to each other. This confirms that smurfing is used as an in-group marker and that Smurfs are depicted in the comics as a special sociolinguistic group.

Consequently, we assume that smurfing contributes a use-conditional meaning to a clause, marking it as assuming a Smurf-specific cultural and pragmatic background, which we state in (6).

- (6) Conversational function of smurfing:
The speaker identifies themselves as Smurf and signals their Smurf-ness. The speaker presupposes a shared Smurf background with the addressee.

Recoverability of smurfing: Chatzopoulos (2008) assumes an OT constraint SMURF: “smurf all lexical morphemes.” This constraint is outranked by a recoverability constraint, which restricts smurfing to recoverable contexts. She names three factors that facilitate recoverability and, thus, smurfing: First, phonological similarity, which she illustrates with examples like *smurfdag* ‘birthday’ or Greek *strumfonía* (*simfonía* ‘symphony’); second, multiword expressions such as “proverbs, idioms and phrases with some degree of fossilization;” third, pragmatically rich context. Bollig (2016) names additional factors for smurfing in French and Spanish. He also stresses that recoverability of the smurfed expression (for the reader) is not always given: There are instances in which there is no uniquely identifiable smurfed expression – as in *schtroumpf de schtroumpf* ‘smurf of a smurf’, which could replace *filz de putain* ‘son of a bitch’, but also

similar expletive expressions. In other cases, there is too much smurfing and too little context for recovering what is being said. Bollig illustrates this with the riddle in (7).

- (7) 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 schtroumpf!
‘But no, look! TWO smurfs!’

While the reader cannot recover the smurfed expressions, the Smurf asked to solve the riddle, can, as the continuation of the dialogue shows. We, thus, assume that there is a recoverability constraint very much like (1-d), which needs to be added to the use conditions of smurfing in (6).

HPSG modelling

We will follow the model of HPSG morphology as sketched in Bonami & Boyé (2006) and applied to compounding in Desmets & Villoing (2009), though other approaches would be equally feasible. In this approach, the sorts *phrase* and *word* are subsorts of *syn(tactic)-sign*, and *word* and *lexeme* are treated subsorts of *lex(ical)-sign*. Syntactic signs have a PHON feature. Lexical signs have a feature M-DTRS, whose value is a list of lexemes. A lexeme’s inflectional forms are encoded in a feature STEMS, whose value contains all stem forms.

The lexical description of the lexeme *Schlumpf* is given in Fig. 1a. The lexeme has two stem forms, *schlumpf* and *schlümpf*, the latter of which is used for plural and diminutive, for example.

We assume that all PHE lexemes have a L(EXICAL-)ID(ENTIFIER) value which is a subsort of *phe*. Some lexemes are pure PHEs, such as German *Dings* ‘thingamajig.’ Their LID value would be a subsort of *phe*, *dings-phe*. *Schlumpf* has a both literal and a placeholder use. We, thus, need two distinct LID values, *schlumpf-lit* and *schlumpf-phe*. There is an additional constraint for the literal use, see Fig. 1b, which specifies its meaning as referring to a Smurf. The constraint in Fig. 1c restricts the PHE use. It does not constrain the CONTENT value, but contributes its idiosyncratic use conditions, sketched in (6), abbreviated as SMURF-UC.

Desmets & Villoing (2009) treat compounds as morphologically complex lexemes with two lexemes on their M-DTRS list. We follow Cheung (2015) in assuming that the PHE combines with the sign that it stands for. However, we treat this as a compound. We argued above that there are two types of replacement by a PHE: morphological placeholder compounds (m-PHC) and phonological placeholder compounds (p-PHC).

The constraint on **morphological replacement** is given in Fig. 2. The PHE is the right-most daughter (which means that it is the head of the compound since German compounds are right-headed). The head can be morphologically complex as long as it is a PHE according to its LID value. The compound shares its inflectional information with the PHE, i.e., both have the same STEM value, [1]. The PHE will also contribute its use-conditional meaning, indicated by the identity of the CONTEXT values of the PHE and the compound, [6]. The CAT value of the compound and its CONT value are the same as those of the non-head, i.e. the replaced lexeme. Note, however, that the PHE and the replaced lexeme also agree in these values, except for the LID value. This makes it possible to restrict the syntactic or semantic properties of the replaced lexeme – as the restriction to female proper nouns for English *what’s-her-name*. It also allows us to require that we use a gender- or POS-marked version of the smurfeme (i.e., *Schlumpf-e.F* or *schlumpf-ig.ADJ*). To avoid circular application of replacement, we require that the replaced lexeme have a LID value that is not of sort *phe*.

Phonological replacement is licensed by the constraint in Fig. 3. In this case, we find the replacement of a phonological unit – a syllable in smurfing – but all other properties of the compound are identical with those of the non-PHE, which we treat as the morphological head for this reason. There are a number of similarities between the two placeholder compounds: Only one component may be a PHE, the CAT value and the CONT value of the compound are determined by the non-PHE component, and the use conditions are inherited from the PHE-component.

Our analysis also captures the differences between m- and p-replacement: In a p-PHC, the PHE must be a simple lexeme, i.e. of sort *simple-lexeme*, which excludes any derivational affixes. We also need to change the STEM value. This is done by the relation **phe-phon**. This relation will incorporate the basic stem form of the PHE, [1], into all stem forms of the non-PHE component, [2].

Our analysis captures the four properties of PHEs mentioned in the introduction. First, PHE-Ref (the PHE refers to a linguistic sign) is satisfied in the sense that the PHE and the sign it replaces are combined in a compound structure and the result shares all semantic and categorial properties with the sign it replaces. Second, PHE-Truth (no fixed truth-conditional meaning of the PHE) is equally satisfied as the PHE has no lexically fixed CONT value. In a m-PHC, it inherits its CONT value from the non-PHE component. In a p-PHC, no truth-conditional contribution of the PHE enters into the overall semantic combinatorics. Third, the gist of PHE-Constr (possible morpho-

syntactic requirements on the replaced item) is encoded in m-PHCs as both components must have identical CAT values. Consequently, they can constrain each other. For smurfing, we saw that the POS of the smurf-lexeme and the replaced lexeme need to be the same, for example. In p-PHCs, we did not specify such a mutual constraining as the replaced unit is not morpho-syntactic but rather phonological. Fourth, we express PHE-Use (the PHE has its own use conditions) in terms of the CONTEXT value. The CONTEXT value of the PHE always percolates to the PHE compound and can override that of the non-PHE component.

Conclusion

We developed the first analysis of placeholder morphemes in HPSG. We showed that smurfing is an instance of a placeholder construction. Since smurfing occurs with high frequency in the Smurf comics, it provides a good data base for our analysis. Our analysis incorporates the insights of previous approaches to both PHEs and smurfing, but solves some of their problems. We adopted the use-conditional aspect of Enfield (2003), but solved his problem of a too general semantics for PHEs. Similarly, our distinction of m-PHCs and p-PHCs solves the problems in Cheung (2015) when a PHE replaces a single syllable rather than a full linguistic sign. In German, smurfing often occurs with a derived form of the smurfeme, such as *schlumpf-ig* ‘smurf-y’. This is not expected in Chatzopoulos’ analysis, but is accounted for by our analysis of m-PHCs, which allows for a morphologically complex PHE component.

The domain of PHEs combines morphological, syntactic, semantic, and pragmatic aspects. HPSG allows us to include all of these aspects in a single analysis. Nonetheless, we could not discuss them all in detail. An additional open question is whether the two types of PHE-constructions identified in this paper are the only ones. Further research on PHEs in different languages – including smurfing – may shed a light on this question.

Figure 1: The lexeme *Schlumpf* ‘smurf’

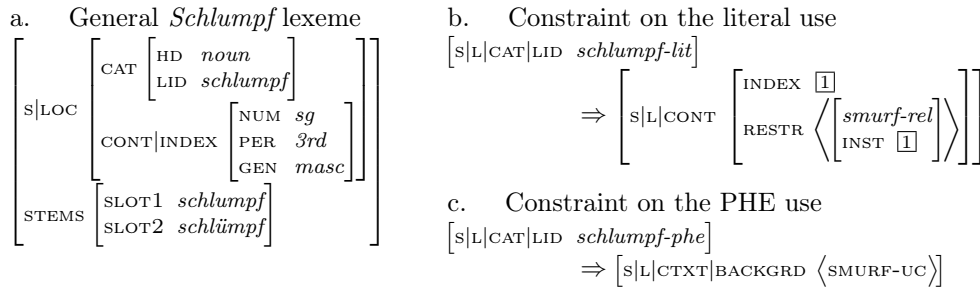


Figure 2: Morphological Placeholder Compound (m-PHC)

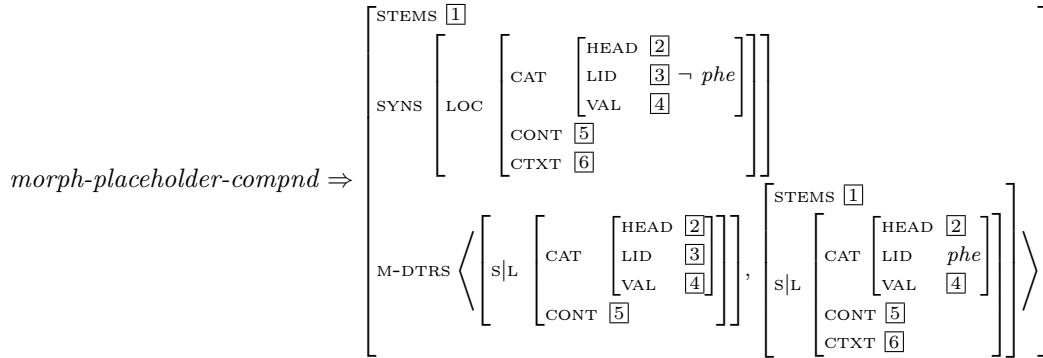
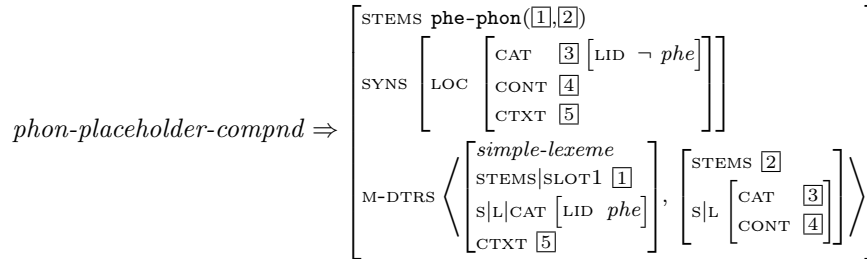


Figure 3: Phonological Placeholder Compound (p-PHC)



- References** ◇ 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. Université Libre de Bruxelles. ◇ Bonami, Olivier & Gilles Boyé. 2006. Deriving inflectional irregularity. In Stefan Müller (ed.), *Proceedings of the 13th international conference on head-driven phrase structure grammar, varna*, 361–380. <http://csli-publications.stanford.edu/HPSG/2006/bonami-boyé.pdf>. ◇ Bonami, Olivier & Berthold Crysmann. 2018. Lexeme and flexeme in a formal theory of grammar. In Olivier Bonami, Gilles Boyé, Georgette Dal, Hélène Giraudo & Fiammetta Namer (eds.), *The lexeme in descriptive and theoretical morphology*, 175–202. Berlin: Language Science Press. <https://langsci-press.org/catalog/book/165>. ◇ 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. https://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. ◇ Desmets, Marianne & Florence Villoing. 2009. French VN lexemes: Morphological compounding in HPSG. In *Proceedings of the 16th International Conference on Head-Driven Phrase Structure Grammar, University of Göttingen, Germany*, 89–109. Stanford: CSLI Publications. <http://csli-publications.stanford.edu/HPSG/2009/>. ◇ 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. ◇ Gutzmann, Daniel. 2013. Expressives and beyond. In Daniel Gutzmann & Hans-Martin Gärtner (eds.), *Beyond expressives* 28, 1–58. Brill.