Constraining Expletives in English

Alex Alsina
Universitat Pompeu Fabra

Fengrong Yang
Universitat Pompeu Fabra

Abstract

This paper aims to present a theory of expletives[1] in English within LFG in which the distribution of expletives follows from general principles and from the lexical entries of the relevant expletives. Consequently, expletives are not subcategorized for, i.e. verbs do not lexically specify whether they take an expletive or what expletive they take, unlike what is assumed in current LFG approaches to expletives. In addition, there are no alternative lexical entries for verbs depending on whether they cooccur with an expletive or not. The proposed analysis diverges from standard mapping theories in the assumption that argument-to-function linking takes place in the syntax, instead of in the lexicon. The current analysis assumes that there are two kinds of expletives: non-thematic expletives, which do not fill an argument position and are licensed by satisfying the Subject Condition, and argumental expletives, which do fill an argument position, but make no semantic contribution.

1 Introduction

The goal of this paper is to propose an alternative theory to the current accounts of the English expletives *there* and *it*, a theory that does not account for the distribution of these expletives through the stipulation of the specific expletive required. This paper is organized as follows. In section 2, we review current analyses of English expletives *there* and *it* and point out the problems they face. We propose our theory about expletive *there* and propositional *it* in section 3 and 4, respectively. Main conclusions are summarized in section 5.

2 Current analyses of expletives

A commonly accepted assumption in LFG is that an expletive GF has to be included in the lexical entry of the predicate that cooccurs with this GF in the clause. For example, Bresnan (1982), in addition to positing the lexical entry of expletive *there*, as in (1a), assumes that the lexical entry of copular *be*, as used in *There is a pig running through the garden*, includes only one thematic argument in its PRED value (i.e., the XCOMP), and has to stipulate that **i)** it takes a non-thematic SUBJ, **ii)** this SUBJ is an expletive, and **iii)** the form of the expletive is *there*, represented as in (1b):

(1) a. *there*: NP$_{[PRO]}$, (FORM) = there
     b. V,   'there-be ((XCOMP))', X = P, V, A
              (OBJ) = (XCOMP SUBJ)
              (SUBJ FORM) =$_C$ there
              (SUBJ NUM) = (OBJ NUM)

(Bresnan 1982: 73)

---

[1] By an expletive we refer to a grammatical function (GF) with no semantic content. Weather "it" will not be considered as an instance of an expletive, given compelling arguments for its non-expletive status (see for example Krejci 2014, Levin 2017, etc.).

Falk (2001) takes a similar position, proposing that in constructions in which an expletive subject cooccurs with a clausal complement, the main predicate should be analyzed as subcategorizing for a COMP and an expletive subject, with a special requirement on the form of the expletive. Namely, the non-thematic subject must be realized as *it* in such constructions:

(2)   It seems that the geneticist clones dinosaurs.

<div align="right">(Falk 2001: 137)</div>

The lexical entry of *seem* in (2), with the requirement on the form of the non-thematic expletive subject, is suggested to be represented as follows:

(3)   *seem*:   V   (↑ PRED) = 'seem <(↑ COMP)> (↑ SUBJ)'
                       (↑ SUBJ FORM) = it

<div align="right">(Falk 2001: 138)</div>

As an alternative to stipulating the specific expletive required by means of the feature FORM, which is also present in the lexical entry of the expletive (as in (1a)), some LFG theories achieve the same result by stipulating in the lexical entry of a predicate the person, number, and gender features of one of the GFs required by the predicate, which can only match those of a particular expletive. For example, Kaplan and Zaenen (1995) propose that the predicate *likely* as used in example (4a) has the lexical entry in (4b) and the expletive *it* has the lexical entry in (4c).

(4) a.   It is likely that Susan will be late.
     b.   *likely*: (↑ PRED) = 'likely <(↑ COMP)> (↑ SUBJ)'
                      (↑ SUBJ PERS) = 3
                      (↑ SUBJ NUM) = SG
                      (↑ SUBJ GEND) = NEUT
     c.   *it*:   (↑ PERS) = 3
                      (↑ NUM) = SG
                      (↑ GEND) = NEUT
                      ¬ (↑ PRED)

<div align="right">(Kaplan and Zaenen 1995: 158)</div>

A consequence of stipulating the presence of an expletive in the lexical entry of a verb is that the distribution of expletives does not follow from general principles (such as the Subject Condition), unlike what happens in GB/MP, where it is a direct consequence of the Extended Projection Principle (EPP, Chomsky 1981, 1982, etc.). Moreover, it does not provide a way of explaining the distribution of the expletives *there* and *it*, namely, the fact that, in standard modern English, expletive *there* cooccurs with a postverbal NP, whereas expletive *it* cooccurs with phrases with propositional content, such as infinitival phrases or full clauses:[2]

(5) a.   There/*it are flowers in the yard.

---

[2] Expletive *it* was also used in existential constructions in earlier stages of English, and is still used in African American Vernacular English (Louise McNally, p.c.).

b. It/*there seems that a new idea is emerging.

c. It/*there surprised me that you won the lottery.

Another implication of the assumption in current LFG that expletives are lexically selected is that there are two lexical entries for predicates that may take an expletive as their subject. In other words, predicates must have two different lexical entries depending on whether they use an expletive or not. This is the case for Kaplan and Zaenen (1995) with "extraposable" predicates such as *likely*, *important*, and *advisable*, etc., which are proposed to have two different lexical entries, one with and one without the expletive subject *it*:

(6) a. That Susan will be late is likely.

     likely: (↑ PRED) = 'likely <(↑ SUBJ)>'

b. It is likely that Susan will be late.

     likely: (↑ PRED) = 'likely <(↑ COMP)> (↑ SUBJ)'

Generation of the lexical entry in (6b) is achieved by positing an extraposition rule for "extraposable" lexical entries, which licenses a second lexical entry with a non-thematic subject (for another proposal within LFG, see Falk 2001):

(7) Extraposition rule:

   a. Extraposable entry:

     (↑ PRED) = 'R <(↑ SUBJ) … >'

   b. Lexical entry added:

     (↑ PRED) = 'R <(↑ COMP) … > (↑ SUBJ)'

     (↑ SUBJ PERS) = 3

     (↑ SUBJ NUM) = SG

     (↑ SUBJ GEND) = NEUT

<div align="right">(Kaplan and Zaenen 1995: 158)</div>

However, the extraposition rule proposed by Kaplan and Zaenen (1995) can only account for constructions in which the original *subject* clause is extraposed. Such a rule gives no explanation, for example, for constructions in which an original object clause is extraposed, as in the following examples:

(8) a. I resent it greatly that you didn't call me.

b. I regret it very much that we could not hire Mosconi.

<div align="right">(Postal and Pullum 1988: 642)</div>

In summary, current LFG accounts of expletive GFs depend largely on stipulations in the lexical entries of the expletive-taking predicates. In addition, attempts to capture generalizations by means of lexical rules are partial (as they only address extraposition *it*) and incomplete (as they do not generalize to all instances of extraposition *it*).

3 Analysis of expletive *there*

In this section, we propose our analysis of expletive *there*. We assume that the distribution of expletive *there* is regulated by its lexical entry and by independently required constraints, such as the Subject Condition. For instance, verbs such as *appear* or *stand* allow their NP argument to be realized

alternatively as SUBJ or OBJ, and when the OBJ realization is chosen, the expletive *there* is the syntactic SUBJ. This is illustrated by examples in (9) and (10), respectively:

(9) a.   A bird appeared on the windowsill.
    b.   There appeared a bird on the windowsill.
(10) a.   A monument stood in the square.
    b.   There stood a monument in the square.

We propose that verbs like *appear* or *stand,* despite having two alternative realizations of their NP argument, have only one lexical entry, with an a-structure consisting of one core argument, as shown in (11) for *appear*.

(11)   '*appear* < I A >'[3]

The mapping principles allow for an internal argument to map onto either SUBJ or OBJ, as assumed in current versions of mapping theories such as LMT (see Bresnan 1994, Kibort 2001, or Findlay 2016) or FMT (see Alsina 1996a). Therefore, the a-structure in (11) is used in all the c-structure/f-structure pairs corresponding to the examples in (9). As for the association between arguments and GFs, we are in line with FMT in assuming it takes place in the f-structure and not in the lexicon.

Let us assume the following two well-formedness conditions on f-structures, as violable constraints, i.e., the Subject Condition (SUBJCON) and GF Faithfulness (GF-FAITH), stated respectively as:

(12)   Subject Condition (SUBJCON):
        Every verbal f-structure must include a subject.

(13)   GF Faithfulness (GF-FAITH):
        i) Every direct GF must be lexically required (i.e., required by the
           lexical information of some element of the clause, such as the a-
           structure of the predicate).
        ii) A GF has a PRED value iff it corresponds to an argument or has
            semantic content.

GF-FAITH is roughly equivalent to the commonly assumed Coherence Condition (Kaplan and Bresnan 1982) and replaces it in the present theory. Adopting an Optimality Theory (OT) approach to constraint interaction (see Kuhn 2003), two rankings of these constraints are possible, listed as follows:

(14) a.   Ranking 1: SUBJCON ≫ GF-FAITH
    b.   Ranking 2: GF-FAITH ≫ SUBJCON

With Ranking 1, we have languages in which every clause must include a subject, even if that incurs a violation of GF-FAITH, as in the case of a structure with an expletive subject that is not lexically required. English is an example of such a language. By contrast, with Ranking 2, we have languages in which expletive subjects are not possible, as every direct GF must be

---

[3] Here we use *I* to represent the internal argument and *A* to represent the non-core locative argument. *E* will be used to represent the external argument in (24).

lexically required (typically, an argument in the a-structure), even if that implies violating SUBJCON. Spanish or Catalan is an example of such a language.
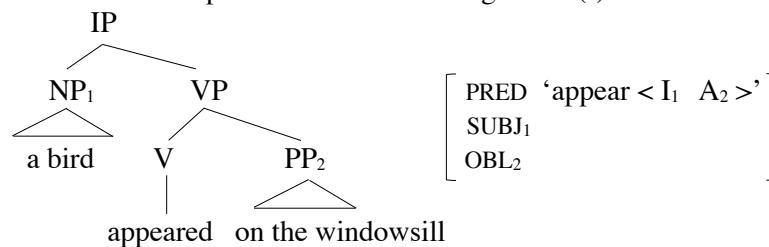
As for the lexical entry of expletive *there*,[4] we assume that it merely states that its category is NP and that it maps onto a GF with oblique case and that it cooccurs with an NP object in the same clause.[5]

(15)   *there*: NP_1 $\begin{bmatrix} GF \begin{bmatrix} CASE & OBL \end{bmatrix}_1 \\ OBJ \end{bmatrix}$

        where OBJ maps onto an NP

Let us consider the representation of the examples in (9). If we choose to map the internal argument of the predicate in (11) to SUBJ, we obtain the c-structure/f-structure pair in (16), corresponding to example (9a). Here and in what follows, the subscripted integers signal the correspondence between arguments, GFs, and c-structure nodes:

(16) C-structure/f-structure pair with the internal argument (I) as SUBJ:

IP
NP$_1$ VP
a bird   V   PP$_2$
appeared   on the windowsill

$\begin{bmatrix} PRED & \text{'appear} < I_1 \ A_2 >\text{'} \\ SUBJ_1 \\ OBL_2 \end{bmatrix}$

This is the optimal candidate: the f-structure in (16) satisfies both SUBJCON and GF-FAITH, because it contains a subject which is lexically required (it maps onto an argument).

If we choose the option in which the internal argument of the predicate in (11) maps to OBJ, there are two possible f-structures consistent with this mapping, as shown in the two c-structure/f-structure pairs in (17):

(17) Possible c-structure/f-structure pairs with internal argument (I) as OBJ:

a.   VP
V   NP$_1$ PP$_2$
appeared   a bird   on the windowsill

$\begin{bmatrix} PRED & \text{'appear} < I_1 \ A_2 >\text{'} \\ OBJ_1 \\ OBL_2 \end{bmatrix}$

---
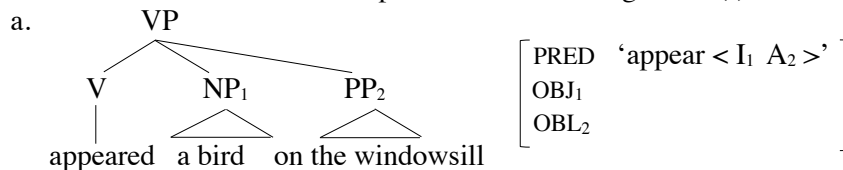
[4] We are not using annotations for lexical entries, but we give the f-structures that would be generated by the annotations.

[5] The reference to oblique case in the lexical entry of expletive *there* serves to block verb agreement with the expletive subject, allowing the direct case object to be the agreement trigger, as in Bresnan (1994).

b.
```
        IP
       /  \
     NP₁   VP
      |    / \
    there V   NP₂   PP₃
          |
      appeared  a bird  on the windowsill
```

$$\begin{bmatrix} \text{PRED} & \text{'appear} < \text{I}_1 \ \text{A}_2 > \text{'} \\ \text{SUBJ} & [\text{CASE} \ \text{OBL}]_1 \\ \text{OBJ}_2 & \\ \text{OBL}_3 & \end{bmatrix}$$
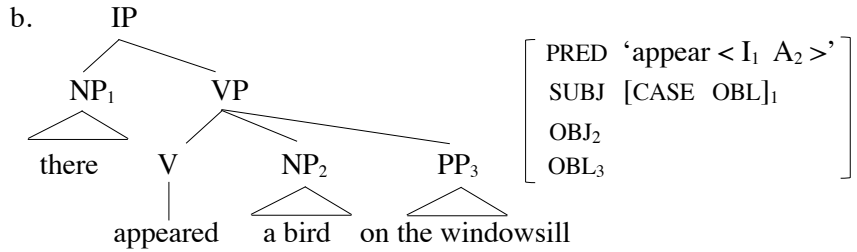
There is a difference in information structure between (9a) and (9b), as, in such constructions, the OBJ realization is discourse-new, unlike what we find with the SUBJ realization (Huddleston and Pullum 2002: 1396–1397). Given this difference in information structure, the two c-structure/f-structure pairs in (17) are not in competition with that in (16), but are in competition with each other. We assume that, in order for two structures to be candidates for optimization, they must be identical in meaning and that a difference in information structure entails a difference in meaning. In both structures in (17), the internal argument of *appear* corresponds to the OBJ; they differ in that there is no SUBJ in (17a) and there is a non-thematic SUBJ in (17b). The optimization for these two candidates is represented in the tableau in (18):

(18)   Optimization for (17a) and (17b):

	SUBJCON	GF-FAITH
(17a)	*!	
☞ (17b)		*

As the optimization above illustrates, (17b) satisfies SUBJCON and, although it violates GF-FAITH, it is chosen over (17a), resulting in the grammatical sentence with an expletive subject, i.e., ***There** appeared a bird on the windowsill*.

Within this theory, expletive *there* can only be used when it is required to satisfy SUBJCON, as it violates GF-FAITH because it is not lexically required. This includes sentences in which the expletive is in Spec of IP, as in (9b) or (10b), as well as raising to object constructions, where the expletive is in the object position, rather than in Spec of IP, but satisfies SUBJCON in the complement clause, as in (19a):

(19) a.   Kim believed there to be flowers in the yard.

    b. * Kim believed Ø to be flowers in the yard.

(19a) is chosen over (19b) for the following reason. We assume that raising-to-object verbs such as *believe* or *expect* specify in their lexical entry (but not in their a-structure) that they take a direct GF that is not an argument of the verb. This non-thematic GF is the one that is structure-shared with the subject of the clausal complement of the raising-to-object verb. The expletive subject in the embedded f-structure of (19a) satisfies SUBJCON but violates GF-FAITH, because it is not lexically required by *be*. With respect to the raising

object in the main clause, it does not violate the GF-FAITH, since it is lexically required by *believe*. By contrast, the f-structure of (19b) violates SUBJCON in the embedded clause and GF-FAITH in the main clause, as it does not have an expletive subject in the embedded clause and a raising object in the main clause, respectively. The two f-structures and the optimization are represented as in (20) and (21), respectively:

(20) a.   F-structure of (19a):

$$
\begin{bmatrix}
\text{PRED} & \text{'believe} < E_1, I_2 > \text{'} \\
\text{SUBJ} & [\text{PRED} \quad \text{'Kim'}]_1 \\
\text{OBJ}_3 \\
\text{OBJ} & \begin{bmatrix}
\text{PRED} & \text{'be} \quad < I_4, A_5 > \text{'} \\
\text{SUBJ} & [\text{CASE} \quad \text{OBL}]_6 \\
\text{OBJ} & [\text{PRED 'flower'}]_4 \\
\text{OBL} & \begin{bmatrix} \text{PRED 'in} < I > \text{'} \\ \text{OBJ} \quad [\text{PRED 'yard'}] \end{bmatrix}_5
\end{bmatrix}_2
\end{bmatrix}
$$

b.   F-structure of (19b):

$$
\begin{bmatrix}
\text{PRED} & \text{'believe} < E_1, I_2 > \text{'} \\
\text{SUBJ} & [\text{PRED} \quad \text{'Kim'}]_1 \\
\text{OBJ} & \begin{bmatrix}
\text{PRED} & \text{'be} \quad < I_3, A_4 > \text{'} \\
\text{OBJ} & [\text{PRED 'flower'}]_3 \\
\text{OBL} & \begin{bmatrix} \text{PRED 'in} < I > \text{'} \\ \text{OBJ} \quad [\text{PRED 'yard'}] \end{bmatrix}_4
\end{bmatrix}_2
\end{bmatrix}
$$
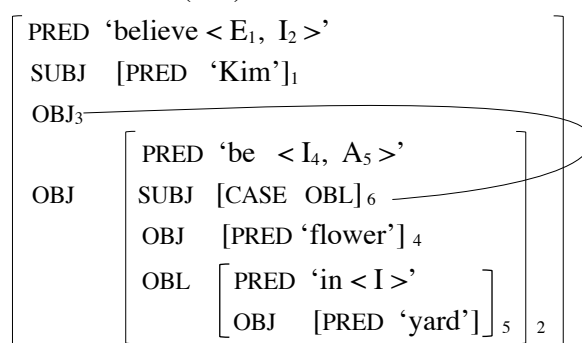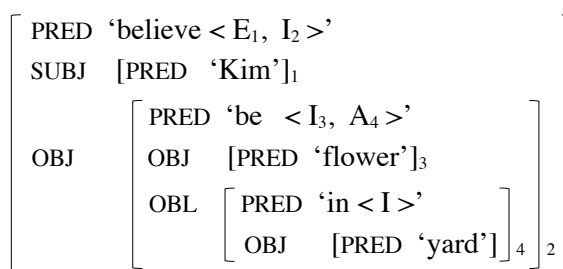
(21)   Optimization for (20):

	SUBJCON	GF-FAITH
☞ (20a)		*
(20b)	*!	*

In summary, the distribution of expletive *there* follows from the lexical entry of this expletive and independently motivated constraints. Its occurrence need not be stipulated in the argument structure (or in any other part of the lexical information) of the predicate that cooccurs with it.

4 Analysis of propositional expletive *it*

In this section, we give an account of propositional expletive *it*. We argue against the idea of a special argument-to-function mapping principle licensing a propositional argument as a complement of the verb, in 4.1. In 4.2, we discuss the grammatical function of the propositional constituent in structures with

propositional expletive *it*. In 4.3, we give our proposal about propositional expletive *it*. 4.4 highlights the main features of this proposal and compares it with that of Kim and Sag (2005).

4.1 Against a special propositional mapping principle

Propositional *it* can only occur in sentences with a propositional constituent. We propose that propositional *it* is not restricted to satisfy SUBJCON, unlike expletive *there*. One might be tempted to assume that there is a mapping constraint which allows propositional elements to map onto a non-subject function, thereby vacating a position that can be filled by this expletive. We argue that there are three reasons against this assumption. First, "proposition" is not an a-structure notion: it is not a thematic role, but a type of semantic content that an argument may have. An argument role can be semantically a "thing" or a "proposition", among other types (see Jackendoff 1990: 49 for more information about semantic types or conceptual categories), as we see with the subject of *surprise* in (22):
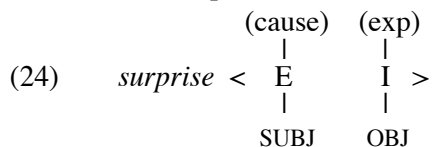
(22) a.   That you won the lottery surprised me.
     b.   The result surprised me.

Thus, there should not be two a-structures depending on whether the theme argument of *surprise* is a "thing" or a "proposition". Likewise, the theme argument of *expect* can also be a thing or a proposition, as in (23), and there should not be two a-structures of *expect* for these two cases. There is just one a-structure with an experiencer and a theme:

(23) a.   Nobody expected anything of me.
     b.   Nobody expected (it) of you that you could be so cruel.

(Kim and Sag 2005: 194)

Second, it would be very problematic to assume that there is a mapping principle that allows a propositional argument to map onto a non-subject GF (whether it is OBJ, as in a framework with a reduced inventory of GFs, such as the present one, or COMP, XCOMP, etc., if a larger inventory of GFs is available), overriding the more general mapping principle that maps the higher of two core arguments to SUBJ (which is implemented in different ways depending on the particular theory, from Bresnan and Kanerva 1989 to Kibort 2001 and other work). For example, *surprise* is a verb whose a-structure includes a cause, which can be a thing or a proposition, and an experiencer argument, expressed as an object, as (22) illustrates. The canonical mapping of arguments to GFs for this verb is represented in (24):

$$
\begin{array}{ccc}
 & \text{(cause)} & \text{(exp)} \\
 & | & | \\
(24) \quad \textit{surprise} \; < & \text{E} & \text{I} \quad > \\
 & | & | \\
 & \text{SUBJ} & \text{OBJ}
\end{array}
$$

If we employed the special propositional mapping rule mentioned, which would allow the propositional *cause* argument to map onto an OBJ/COMP, there

should be no reason to insert expletive *it* to fill the SUBJ function, (i.e., it would not be required to satisfy SUBJCON) because the *experiencer* argument would be forced to fill the SUBJ, resulting in the ungrammatical **I surprised that you won the lottery*, with the same meaning as (22a).

Third, as is well known since Postal and Pullum (1988), propositional *it* is not restricted to filling the SUBJ function: it may also be an OBJ, for instance. We repeat examples (6b) and (8) as (25) for ease of consultation:

(25) a.   **It** is likely that Susan will be late.

b.   I resent **it** greatly that you didn't call me.

c.   I regret **it** very much that we could not hire Mosconi.

We therefore conclude that it is inadequate to propose special mapping constraints to license the realization of a propositional argument as a complement in order to capture the distribution of propositional expletive *it*. We will assume that the mapping of arguments onto GFs is independent of their semantic type as things or propositions, while looking for other ways of explaining the distribution of this expletive.

4.2 Syntactic function of the extraposed clause

Some words need to be said about the grammatical function of the subordinate clause in examples such as (25). This subordinate clause has been analyzed as an ADJ (i.e., as an appositive clause, e.g. Vikner 1995: 241), a COMP (e.g. Kaplan and Zaenen 1995), or has been suggested to contribute to the same argument slot as the expletive *it*, thus unifying their f-structure information under the same but "discontinuous" function (either SUBJ or OBJ), as proposed by Berman et al. (1998) for German.

We will not assume the discontinuous analysis, for the following reason. In a structure like (25a), the propositional *it* and the subordinate clause would be analyzed as contributing together to the SUBJ function of the matrix clause according to the discontinuous approach. Note that grammatical functions in English are strictly constrained by their c-structure realization: the subject always occupies the specifier position of IP, whereas the canonical object position is VP-internal.[6] Now, while it is true that the propositional *it* occupies the subject position in (25a), and thus is uncontroversially the subject of the clause, the subordinate clause is clearly not occupying the specifier position of IP. Therefore, analyzing the clausal complement as (part of) the SUBJ function would not be appropriate. On the other hand, considering this element as a non-subject function will not cause such a problem.

Adopting the reduced inventory of grammatical functions proposed in Alsina (1996a) and Patejuk and Przepiórkowski (2016) consisting only of SUBJ, OBJ, and OBL (see also Alsina et. al. 2005 and Forst 2006 for arguments against the GF COMP), and taking into consideration the arguments from Kim

---

[6] The default object position is immediately following the verb, but certain conditions of the object phrase (heaviness, focus or discourse-newness, or a non-nominal category) allow or favor having other sister constituents linearized before it.

15

and Sag (2005: 197) against analyzing the clausal phrase in extraposition as an adjunct, we will not analyze it as ADJ or COMP. Instead, we assume that the syntactic function of the clausal phrase is always OBJ, which is a distinct function from the function of propositional *it*.

One might object to the claim that the clausal complement in sentences like (25b,c) is an OBJ with the argument that clausal adjuncts cannot appear before objects (pointed out to us by Bresnan p.c.), and therefore the clausal constituent in sentences such as (25b,c) should not be analyzed as bearing the OBJ function because it is preceded by a clausal adjunct such as *greatly* or *very much*. We admit that an NP object cannot be separated from its preceding verb by a clausal adjunct (excepting instances of heavy NP shift), as illustrated below:

(26) a. * She saw often Tom.
     b.   She saw Tom often.
     c.   She often saw Tom.

<div align="right">(Huddleston and Pullum 2002: 247)</div>

But when a construction has a clausal object, it is in fact perfectly acceptable for an adverb of the matrix clause to precede the object:

(27)   I regret deeply volunteering to take part.

<div align="right">(Huddleston and Pullum 2002: 781)</div>

It can be argued that *regret* takes an object, which can be an NP, as in *I regret my intolerance*, or a clause, as in (25c) or (27). The same happens with three-place predicates, such as *tell*, whose second object can be either an NP or a clause. With two NP objects, it is not possible for an adverb such as *yesterday* in the following example, which modifies the matrix clause, to precede either of the NPs:

(28) a. * He told me yesterday the story.
     b. * He told yesterday me the story.

By contrast, the adjunct *yesterday* can appear between the two objects, when the object following the adverbial adjunct is a clause, as opposed to an NP:

(29)   He told me yesterday you wanted it.

<div align="right">(Huddleston and Pullum 2002: 781)</div>

As we can see from the example above, *yesterday*, as an unambiguous adjunct of the matrix clause, appears before the clausal object of (29), a double object construction (for the object status of the clausal complement of *tell*, see Dalrymple and Lødrup 2000). Therefore, constructions with an adjunct separating two objects are not necessarily bad: they are only bad if the adjunct appears immediately before an NP object, but not if it precedes a clausal object.

Independent evidence for the assumption that the clausal complement is syntactically an OBJ comes from the impossibility of extraposing the second object of a predicate that takes two objects, as (30) illustrates:

(30)   * He told me it that he has tried.

The ungrammaticality of (30) can be explained by analyzing the extraposed complement as an OBJ. Unlike other languages, which allow more than two

objects, for example, some Bantu languages (see Bresnan and Moshi 1990, Alsina 1996b, among others), it is clear that English allows at most two objects. The principle that enforces this limitation explains the ungrammaticality of (30), as there would be three objects (i.e., *me*, *it*, and *that he has tried*) if we assume that the extraposed clause is an object. If we assumed the extraposed clause to bear a GF other than OBJ, there would only be two objects in (30), which would leave the ungrammaticality of (30) unexplained.

4.3 Propositional *it*

We assume that the distribution of propositional expletive *it* depends on a special lexical entry that allows it to be used in the presence of an OBJ with propositional semantics:

(31)   Lexical entry of propositional *it*:

$$\textit{it}: \text{NP}_1 \quad \left[\begin{array}{l} \text{GF} \left[\begin{array}{ll} (\text{PRED 'pro'}) \\ \text{PERS} \quad 3 \\ \text{NUM} \quad \text{SG} \end{array}\right]_1 \\ \text{OBJ}_2 \end{array}\right] \qquad \text{Semantic Structure} \quad \left[\begin{array}{ll} \text{TYPE} & \text{proposition} \end{array}\right]_2$$

The lexical entry in (31) does not restrict the c-structure realization of the propositional complement: it can be a *that*-CP, as in the examples given, a *for-to*-infinitive clause, or a *to*-infinitive clause, etc. For example:

(32) a.   It is advisable for students to prepare for the exam.
     b.   It is important to buy a lottery ticket.

This lexical entry allows propositional *it* to be used in two different situations: the non-thematic (or true) expletive *it* (as in (2)) and the argumental expletive *it*, found in extraposition, as in (25) and (32). The former violates GF-FAITH, as it is not lexically required, but satisfies SUBJCON, very much like expletive *there*; the latter satisfies GF-FAITH, as it is an argument of the predicate, like most NPs, and has the effect of licensing a clausal complement. Let us consider the two in turn.

Non-thematic expletive *it*

On the one hand, *it* can be used in constructions like (2), in which the verb has a single argument with propositional content that maps onto a non-subject function. This is termed the impersonal construction in Huddleston and Pullum (2002: 960) and it involves verbs such as *seem*, *appear*, *happen*, *turn out*, etc.[7] With these verbs, the subordinate clause cannot appear in subject position, but only in postverbal position, which distinguishes them from extraposable predicates such as *surprising*, which can have a clausal expression either preverbally or postverbally:

(33) a.   It seems that he was lying.
     b. * That he was lying seems.

---

[7] These verbs also have a use as raising verbs, taking a predicative complement, instead of a full clausal complement, but we will not be concerned with this use here.

(34) a.   It is surprising that he was lying.
   b.   That he was lying is surprising.

We therefore propose that predicates of the impersonal construction like *seem* lexically specify both the grammatical function and the grammatical category of their single argument: it is a CP and an OBJ:

(35)   *seem*   < Arg >
                         |
                        OBJ
                         |
                        CP

Given that the sole argument of *seem* maps onto an object, the clause in which this verb appears needs a subject. In this situation, propositional *it* fills the non-thematic SUBJ function, satisfying SUBJCON. Notice that this GF is not lexically required, and thus violates GF-FAITH; but since SUBJCON ranks higher than GF-FAITH in English, the structure with the expletive subject is chosen over the one that lacks it. Being a non-thematic subject, the option without PRED is chosen. In such cases, the subject *it* is a true, or non-thematic, expletive.

Extraposition *it*

On the other hand, the lexical entry in (31) allows any argument that can be expressed as an NP and is semantically consistent with a proposition to be expressed by means of *it*, which licenses a clausal object that provides the propositional content of the argument. This is the extraposition *it*, which we find in (25) and (32). In these examples, *it* is not a true expletive – if we take *true expletive* to be a GF that does not map onto a semantic participant – as it maps onto an argument of the predicate. In these cases, the option with PRED is chosen. The GF with propositional content does not correspond to an argument of the verb, but yet satisfies GF-FAITH, as it is required by the lexical entry of *it*.

Thus, a single lexical entry for expletive *it*, in (31), in combination with the general constraints GF-FAITH and SUBJCON, gives rise to both the non-thematic expletive *it* of impersonal constructions, as in (33), and the argumental expletive *it* of the extraposition construction, as in (25), (32), and (34). A consequence of this theory is that, in the extraposition construction, expletive *it* can only occur in the position that corresponds to the argument GF it realizes, whether SUBJ or OBJ, whereas the phrase that corresponds to the GF with propositional content appears in postverbal position, as a clausal object, which explains the ungrammaticality of (36), as expletive *it* is not in the position that corresponds to its GF, nor is the phrase with propositional content.

(36)   * That you won the lottery surprised me **it**.

(cf. (5c))

However, nothing that we have said so far explains the ungrammaticality of (37a), where expletive *it* is an OBJ, as corresponds to the argument it fills, and the phrase with propositional content is also an OBJ, as required by the lexical

entry (31). The grammatical example (37b) has the same elements as (37a), only in a different linear order.

(37) a. * I resent that you didn't call me it.

　　 b. 　I resent it that you didn't call me.

Following Kim and Sag (2005), we can assume the existence of a linear precedence rule that requires clausal phrases to linearly follow a sister GF:

(38) 　　Linear precedence rule 1:

　　　　A clausal phrase must follow a sister GF.

As the two postverbal constituents in (37) are OBJ and they are sister constituents, the clausal object must follow the NP OBJ *it*, which explains the contrast between (37a) and (37b).

　　Another fact that needs to be explained is that, as observed by Huddleston and Pullum (2002), clausal extraposition is normally required when there is an object predicative complement (such as *hardly surprising* in (39)), except if the clause is topicalized (as in (39c)):

(39) a. 　* I find that he tried to retract his statement hardly surprising.

　　 b. 　　I find **it** hardly surprising that he tried to retract his statement.

　　 c. 　　That he tried to retract his statement I find hardly surprising.

　　　　　　　　　　　　　　　　　　　(Huddleston and Pullum 2002: 963)

The ungrammaticality of (39a) follows from the linear precedence rule (38). However, reordering of the predicative phrase and the extraposed clause yields another ungrammatical, or marginal, sentence:

(40) 　　* I find hardly surprising that he tried to retract his statement.

　　We can explain this fact by adopting the explanation in Kim and Sag (2005), which involves another linear precedence statement that requires a subject to precede the phrase of which it is the subject

(41) 　　Linear precedence rule 2:

　　　　If XP is the subject of YP, XP linearly precedes YP.

Example (40) does not comply with this precedence rule, as the subject of *hardly surprising* – the *that*-clause – follows that phrase, instead of preceding it. In contrast, example (39b), with expletive *it*, meets both precedence requirements, as the *that*-clause follows its sister predicative complement, as well as *it*, in compliance with linear precedence rule 1, and *it* precedes the predicative phrase of which it is the subject, in compliance with linear precedence rule 2.

　　A last fact that needs to be explained regarding the extraposition construction is that the propositional constituent in this construction cannot be topicalized:

(42) 　　* That you won the lottery, **it** surprised me.

The ungrammaticality of topicalizing an extraposed complement clause does not mean that a complement clause cannot be topicalized at all. It can be preposed, without propositional *it*, as in (39c), or in the following example (43):

(43) For them to sack him we would regard as a gross miscarriage of justice.
   (*cf*. We would regard it as a gross miscarriage of justice for them to sack him.)

<div align="right">(Huddleston and Pullum 2002: 1255)</div>

The same observation is also reported by Kaplan and Zaenen (1995) and Kim and Sag (2005), among others:

(44) a.   That Susan would be late John didn't think was very likely.
     b. * That Susan would be late John didn't think **it** was very likely.

<div align="right">(Kaplan and Zaenen 1995: 158)</div>

(45) a.   That Kim would lose to Pat, nobody had expected.
     b. * That Sandy snores, **it** bothers Kim more and more.

<div align="right">(Kim and Sag 2005: 202)</div>

Note that informants providing judgement about (42) point out that the sentence is acceptable. They nevertheless also point out that the sentence would be more natural if *it* is replaced by *that*, i.e.:

(46)   That you won the lottery, **that** surprised me.

The possibility of using a demonstrative *that* shows that the acceptable utterance *That you won the lottery, it surprised me* is a case of left-dislocation, instead of topicalization of an extraposed clause; and the *it* is not an expletive *it*, but a pronominal *it* with explicit reference. In the unacceptable situation of (42), *it* is the expletive, which indicates that it is ungrammatical to topicalize an extraposed clause. The explanation that Kaplan and Zaenen (1995) propose for this resorts to a restriction on functional control in long-distance dependencies. They assume that the extraposed clause bears the function COMP and that the functional uncertainty equation that models long-distance dependencies cannot have the GF COMP as its bottom. In other words, they assume that a COMP cannot undergo topicalization.[8] But see Alsina et al. (2005) for an alternative analysis that does not involve COMP.

Our explanation of this last fact, i.e., the ungrammaticality of topicalizing an extraposed clause in English, takes into account the relation between clausal heaviness and extraposition. We are in line with Huddleston and Pullum (2002: 1403) in considering that, "the effect of extraposition is to place a heavy constituent at the end of the clause". Let us assume that expletive *it* marks the clausal object in its lexical entry as heavy, [H+], and that [H+] constituents are linearized as final in their clause. It follows from this that the clausal constituent in a clause with expletive *it* cannot be preposed (i.e., topicalized), as then it would not be final. In contrast with the clausal object of *it*-clauses, other clausal objects are not constrained to be [H+] and therefore are free to be preposed (i.e., topicalized), as in (39c) or (43).

---

[8] See Berman et al. (1998), who report a similar contrast in German between the ungrammaticality of topicalizing an extraposed clause and the possibility of topicalizing a clausal complement and propose an explanation that depends on the assumption that the expletive and the extraposed clause are coarguments.

To conclude this subsection, we would like to point out a consequence of our analysis of propositional expletive *it*. Even though the same lexical entry licenses both extraposition *it* and the non-thematic dummy *it*, these two uses of the expletive have a different representation, as has been noted: extraposition *it*, being an argument, has the [PRED 'pro'] feature, whereas the non-thematic expletive lacks this feature. This difference implies that expletive *it* cannot be the shared constituent in a coordination of an impersonal and an extraposition construction, as shown in (47):

(47) a. It seemed that he was trying to hide his true identity.
　　 b. It was later confirmed that he was trying to hide his true identity.
　　 c. * It seemed and was later confirmed that he was trying to hide his true identity.

<div align="right">(Huddleston and Pullum 2002: 961)</div>

According to Huddleston and Pullum (2002: 961), the ungrammaticality of (47c) suggests that the extraposed clause "does not have the same function in the two cases" (i.e., in (47a) and (47b)). However, according to our analysis, the *that*-clause does have the same function in both cases, namely, object, but the ungrammaticality of (47c) is attributed to the impossibility of expletive *it* being at the same time non-thematic in the *seem* case, as in (47a), where it lacks a PRED value, and thematic in the extraposition case, as in (47b), where it has a PRED value. The following example shows that expletive *it* cannot be the shared constituent in a coordination of an impersonal construction and an extraposition construction also when each conjunct contains its own clausal complement:

(48) 　 * It seemed that he didn't stand a chance and was hardly surprising that he didn't win.

The ungrammaticality of this example cannot be attributed to the putative difference in the grammatical function of the clausal complement in the two constructions involved, but can be attributed to the incompatible requirements imposed in each construction to the expletive *it*.

4.4 Comparison with Kim and Sag (2005)

Kim and Sag (2005), henceforth KS, develop an analysis of the English extraposition construction within HPSG that has some similarities with the present proposal. They propose a rule that creates new words out of words whose SUBCAT feature includes an S or CP argument such that in the new words this S or CP argument is not in the SUBCAT feature, but in the EXTRA feature. Also, an expletive NP (NP[*it*]) holds the place of the extraposed complement in the new word's SUBCAT list.

We will not analyze here the advantages or disadvantages of introducing the additional selection feature EXTRA, although there seems to be little independent evidence for it. The differences between the KS analysis and the present one that we will focus on are: (a) the difference in generality between the two analyses; (b) the difference regarding verbs that select a complement

that is necessarily clausal; and (c) the difference regarding verbs that select a complement that is necessarily an NP.

Whereas in the present theory the same lexical entry for expletive *it* accounts for the use of this expletive in both the impersonal construction and the extraposition construction, the KS analysis limits its scope to the extraposition construction. It is clear that the two constructions have common elements: in both cases, the same expletive is used and the structure includes a clausal complement. This is captured in the present theory, but no indication is given that the KS analysis of the extraposition construction can be extended to the impersonal construction. We will not speculate as to whether and how this can be done, but, while the lexical rule approach implies that the rule applies optionally, the impersonal construction has the subject expletive *it* and the complement clause as obligatory elements.

The two analyses make differing predictions with respect to the classes of verbs that can appear in the extraposition construction. For KS, the extraposition rule is only possible with verbs that select an argument that can be expressed as a clause and, therefore is not possible with verbs whose arguments are restricted to be of other categories, such as NP or PP. In contrast, the present analysis restricts extraposition to occur with verbs that select an argument that can be expressed as an NP: the expletive *it*, being an NP, can only appear in positions where NPs are possible; the argument must also allow the propositional semantics associated with the clausal object licensed by *it*.

With respect to verbs that take an argument that is constrained to be an NP, the present theory predicts that this argument should be expressible by means of extraposition *it*, whereas the extraposition rule of KS cannot be used with it. These verbs include *take*, *put, like*, or *dislike*, etc. With such verbs, a CP complement without the co-appearance of extraposition *it* will result in an ungrammatical construction. For example, the object of *put* is obligatorily an NP, and appearance of a CP object is acceptable with this verb, provided that expletive *it* also appears. On the contrary, without extraposition *it*, such sentences are judged to be unacceptable. The contrast is illustrated as follows:[9]

(49)   I put **it** to you that you know what the consequences would be.

   (*cf.* * I put to you that you know what the consequences would be.

        * I put that you know what the consequences would be to you.)

(Huddleston and Pullum 2002: 247)

As for verbs that only take CP – but not NP – complements, the present analysis predicts they do not allow extraposition *it*, whereas KS lead us to expect them in the extraposition construction. These verbs include *object*, *conclude*, *reason*, *reflect*, *reply*, *complain*, *decide*, etc. Some of these verbs cannot take any kind of complement except for a clause, as is the case of *reason*

---

[9] KS account for these cases by means of additional lexical entries of the verbs involved that directly stipulate that they take the expletive *it* as an object and a clausal structure in the EXTRA feature.

(but not *reason out*, which does take an NP object). Others have alternative lexical entries taking either an NP, an oblique PP, or a CP complement. For example, *conclude* can take either an NP object or a CP complement, with different meanings, as in *I concluded my work* (where *conclude* equals *finish*) and *I concluded that there was no satisfactory solution* (where *conclude* equals *reach the conclusion*), respectively. *Complain* can take either an oblique *about*-PP, which might also be analyzed as an adjunct, or a clause. In *I complained about the weather*, *the weather* is the object or target of the complaint; in *I complained that it was too hot*, the CP is the argument on which the complaint is based. With all of these verbs when taking an CP complement, extraposition *it* is bad:

(50) a. Local authorities complained (*it) that they lacked sufficient resources.

<div align="right">(based on Oxford Dictionary of English, p.356)</div>

  b. The boy's father objected (*it) that the police had arrested him unlawfully.

<div align="right">(based on Oxford Dictionary of English, p.1226)</div>

These facts confirm our prediction that verbs that take a complement that can only be expressed as a CP will not allow extraposition of this complement and constitute a crucial bit of evidence to distinguish our analysis from KS. In their discussion of different classes of verbs with respect to extraposition, KS only mention *think* as an example of Group III verbs, i.e., verbs that only take a clausal complement, but not an NP object. However, this verb turns out not to be a very good example of this type, because it also takes NP objects in restricted conditions, as in *What are you thinking?* or *What do you think?*, given that *what* is an NP, or *She thought something else entirely* or *She thought a few things I cannot explain*, etc. This shows that *think* can take an NP object. Consequently, just as the NP object of *think* is possible, with semantic limitations, it is not surprising that extraposition *it* is possible with this verb, although not at all frequent:

(51) I thought it that it would be nearly impossible for the filmmakers to sustain such a level of excitement through the rest of the movie.

<div align="right">(Kim and Sag 2005: 209)</div>

  To conclude this subsection, the present theory not only accounts for the distribution of expletive *it* in both the impersonal and the extraposition construction by means of a single lexical entry, but correctly predicts that, in the extraposition construction, this expletive can appear wherever we have a verb that takes an NP subject or object and cannot appear in place of the complement with verbs whose complement is constrained to be a CP. KS make the opposite prediction: they predict extraposition to be possible with the latter class of verbs and impossible when the expletive corresponds to an argument whose categorial expression is constrained to be NP.

5 Conclusions

In this paper, we have proposed a theory in which expletives are not stipulated in the lexical entry of the predicate that cooccurs with the expletive, but their distribution instead follows from general principles such as the Subject Condition and from the lexical entries of the relevant expletives. As a result, there are no expletive insertion rules or lexical rules to generate verbs that select expletives, no alternative lexical entries for verbs depending on whether they have an expletive or not, and no need to stipulate in any way the FORM feature of the expletives.

An idea that the present theory crucially depends on is that the assignment of GFs to arguments and the licensing of GFs in a clause take place in the syntax. Whereas standardly in LFG this process is assumed to take place in the lexicon, so that words exit the lexicon with the list of GFs that they take, we assume that the lexical entries of predicates do not fully specify the GFs that they take, but in general only specify the argument structure of the predicate, which constrains, but does not determine, the GFs associated with the predicate. In other words, argument structure in our paper replaces the lexical form in previous analyses, and there is no list of GFs in the PRED value. This is necessary for two reasons. First, the Subject Condition plays an important role in licensing the expletives *there* and non-thematic *it*: the Subject Condition is a constraint on f-structures and it helps choose f-structures with a subject over f-structures without a subject, even if that subject is not an argument of the predicate. Second, the complement clause in extraposition is licensed by expletive *it*; the predicate that cooccurs with that clause should not foresee in its lexical entry that it takes a complement clause; rather, if one of its arguments can be an NP and is realized as the expletive NP *it*, it is this word that licenses the complement clause, thanks to GF-FAITH, a reinterpretation of Coherence.[10]

The proposed analysis indicates that there are two kinds of expletives: true, or non-thematic expletives, as the case of *there* and non-thematic *it*, and argumental expletives, as the case of extraposition *it*. Our proposal about the two expletives, especially expletive *it*, makes use of a reduced inventory of grammatical functions: the three strongly motivated SUBJ, OBJ and OBL, as argued by Alsina (1996a), Alsina et al. (2005), Forst (2006), and Patejuk and Przepiórkowski (2016); and we do not need to enrich the inventory with other grammatical functions, such as COMP or XCOMP, as many LFG analyses such as Kaplan and Zaenen (1995) or Bresnan (1982), etc. do, or other theoretical constructs, such as EXTRA, as in Kim and Sag (2005), which do not have strong independent motivation.

---

[10] The effects of Completeness are captured by the assumption that the argument-to-GF mapping principles apply whenever possible and that lexical requirements must be satisfied in the syntax.

References

Alsina, Alex. 1996a. *The Role of Argument Structure in Grammar: Evidence from Romance*. Stanford: CSLI Publications.

Alsina, Alex. 1996b. Passive types and the theory of object asymmetries. *Natural Language & Linguistic Theory* 14(4), 673–723.

Alsina, Alex, K. P. Mohanan, and Tara Mohanan. 2005. How to get rid of the COMP. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG05 Conference*, 21–24. Stanford: CSLI Publications.

Berman, Judith, Stefanie Dipper, Christian Fortmann, and Jonas Kuhn. 1998. Argument clauses and correlative *es* in German – deriving discourse properties in a unification analysis. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG98 Conference*. Stanford: CSLI Publications.

Bresnan, Joan. 1982. The passive in lexical theory. In Joan Bresnan (ed.), *The Mental Representation of Grammatical Relations*, 3–86. Cambridge MA: The MIT Press.

Bresnan, Joan. 1994. Locative inversion and the architecture of Universal Grammar. *Language* 70(1), 72–131.

Bresnan, Joan and Jonni Kanerva. 1989. Locative inversion in Chicheŵa: a case study of factorization in grammar. *Linguistic Inquiry* 20(1), 1–50.

Bresnan, Joan and Lioba Moshi. 1990. Object asymmetries in comparative Bantu syntax. *Linguistic Inquiry* 21(2), 147–186.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, Noam. 1982. *Some Concepts and Consequences of the Theory of Government and Binding*. Cambridge MA: The MIT Press.

Dalrymple, Mary and Helge Lødrup. 2000. The grammatical function of complement clauses. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG00 Conference*, 104–121. Stanford: CSLI Publications.

Falk, Yehuda. 2001. *Lexical-Functional Grammar: an Introduction to Parallel Constraint-Based Syntax*. Stanford: CSLI Publications.

Findlay, Jamie. 2016. Mapping theory without argument structure. *Journal of Language Modelling* 4(2), 293–338.

Forst, Martin. 2006. COMP in (parallel) grammar writing. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG06 Conference*, 222–239. Stanford: CSLI Publications.

Huddleston, Rodney and Geoffrey Pullum. 2002. *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.

Jackendoff, Ray. 1990. *Semantic Structures*. Cambridge MA: The MIT Press.

Kaplan, Ronald and Joan Bresnan. 1982. Lexical-functional grammar: a formal system for grammatical representation. In Joan Bresnan (ed.), *The Mental Representation of Grammatical Relations*, 173–281. Cambridge MA: The MIT Press.

Kaplan, Ronald and Annie Zaenen. 1995. Long-distance dependencies, constituent structure, and functional uncertainty. In Mary Dalrymple, Ronald Kaplan, John Maxwell, and Annie Zaenen (eds.), *Formal Issues in Lexical-Functional Grammar*, 137–166. Stanford: CSLI Publications. First printed in Mark R. Baltin and Anthony S. Kroch, (eds.). 1989. *Alternative Conceptions of Phrase Structure*, 17–42. Chicago: University of Chicago Press.

Kibort, Anna. 2001. The Polish passive and impersonal in Lexical Mapping Theory. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG01 Conference*, 163–183. Stanford: CSLI Publications.

Kim, Jong-Bok and Ivan Sag. 2005. English object extraposition: a constraint-based approach. In Stefan Müller (ed.), *Proceedings of the HPSG05 Conference*, 192–212. Stanford: CSLI Publications.

Krejci, Bonnie. 2014. What is raining? English weather *it* revisited. *Linguistic Society of America* 88, Minneapolis, MN. 3 January.

Kuhn, Jonas. 2003. *Optimality-Theoretic Syntax: a Declarative Approach*. Stanford: CSLI Publications.

Levin, Beth. 2017. Talking about the weather: a case study of precipitation verbs. Linguistics Colloquium. University at Buffalo.

Oxford Dictionary of English. 2010. Ed. by Angus Stevenson. Third Edition. Oxford: Oxford University Press.

Patejuk, Agnieszka and Adam Przepiórkowski. 2016. Reducing grammatical functions in LFG. In Doug Arnold, Miriam Butt, Berthold Crysmann, Tracy Holloway King and Stefan Müller (eds.), *Proceedings of the Joint 2016 Conference on Head-driven Phrase Structure Grammar and Lexical Functional Grammar*, 541–559. Stanford: CSLI Publications.

Postal, Paul and Geoffrey Pullum. 1988. Expletive noun phrases in subcategorized positions. *Linguistic Inquiry* 19(4), 635–670.

Vikner, Sten. 1995. *Verb Movement and Expletive Subjects in the Germanic Languages*. Oxford University Press.