Abstract

The paper examines borrowed instances of what we call *emphatic superlative ever* (ES-*ever*) into two Germanic languages (Dutch and German) and two Romance languages (French and Spanish). We base our study on extensive corpus data. We model the data in three stages ranging from constructional borrowing (Stage-1: *el coolest job ever* 'the coolest job ever'), via diaconstructions (Stage-2: *la mejor canción ever* 'the best song ever'), up to lexical borrowing (Stage-3: *las portadas más photoshopeadas ever* 'the most photoshoped portals ever'). We extend an earlier approach to social meaning in HPSG to borrowing.
The data extracted for this study is available at: https://osf.io/juewa/?view_only=215970c573d34b148815cc5653965697

1 Introduction

The paper at hand examines borrowed instances of what we call emphatic superlative *ever* (ES-*ever*) into two Germanic languages (Dutch and German) and two Romance languages (French and Spanish), see (1), in order to deepen our understanding of borrowing, which in turn will feed a theoretic implementation into a HPSG model for borrowing.

(1) a. nl: de beste opmerking *ever* 'the best comment ever'
matrix language alternative: *ooit* 'ever'

b. de: bestes Bild *ever* 'best picture ever'
matrix language alternative: *aller Zeiten* 'of all times'

c. fr: la meilleure idée *ever* 'the best idea ever'
matrix language alternative: *de tous les temps* 'of all times'

d. es: la mejor foto *ever* 'the best picture ever'
matrix language alternative: *de siempre* 'of always'

Based on data from an extensive corpus research via Sketch Engine (Kilgarriff et al., 2014), we will show that although the languages at hand exhibit ES-*ever* with varying frequencies, they do so in a homogenous way:

1. Only ES-*ever* is borrowed, other well-formed and well-attested uses in English (such as with negation or in questions) do not occur or only in all-English passages.

2. ES-*ever* respects the rules of the source language (English): (i) it requires licensing by either a morphological superlative form or a semantic/pragmatic superlative; (ii) it can only occur in an extraposed position.

[†]Aspects of this paper were presented at the *45th Annual Meeting of the German Linguistic Society*, March 2023. We would like to thank the reviewers and the audience of that event and of *HPSG 2023* for their comments, in particular Emily M. Bender, Nurit Melnik, and Elodie Winckel. All errors are ours.

3. The usage of ES-*ever* can be modelled in stages.

4. ES-*ever* has blended in with the grammatical rules of the respective matrix languages allowing to be combined with constructions not permissive in English (what we will take as indicative of the last stage)

These findings will serve as a starting point to an HPSG approach to borrowing, which we envision as a three stage model ranging from constructional borrowing (Stage-1) up to lexical borrowing (Stage-3).

We will proceed in the following fashion: We will present our corpus study in Section 2. Section 3 will give deeper insights into the syntactic structure of ES-*ever* in the respective borrowing languages. Section 4 will present our HPSG approach to borrowing. We end with a conclusion.

2 Corpus study

The main goal of our corpus study was to collect relevant data in order to determine the licensing conditions of ES-*ever* in the different matrix languages. In the following, we will outline our methodology and present some of our results.

2.1 Methodology

In order to acquire comparable results in all the languages scurtinized we used the Timestamped corpora which are available for all four languages. In order to optimize the quantity of results we opted for the largest version of the corpus for each language, i.e. Timestamped JSI web corpus 2014-2021 French, Spanish, German and Dutch respectively. Since the Timestamped corpora contain texts extracted from RSS feeds of News-websites, we hoped to limit the extraction of irrelevant hits such as e.g. fragments from all-English texts, gibberish and machine created texts. Creating appropriate queries turned out to be a challenging task: ES-*ever* constructions allow for the (recursive) embedding of constituents, e.g. PPs, relative clauses and so forth. However, such elementary properties of human languages can only be indirectly operationlized in CQL (Corpus Query Language, the language which Sketch Engine provides for conducting more advanced searches), as it is generally restricted to describing "flat" linear order of strings and is generally unaware of hierarchical structures. Consequently, any query can only be an approximation; ours are no exception. The expressions in Q1, Q2, and Q3 show our queries for French, which – except for references to language specific lexical items (i.e. *et* 'and') – are the same for all languages.

Q1
```
[tag="ADJ.*"][]{0,5}
[tag="NOM|NAM"&word!="than|for|4"]
[word="ever|EVER"][word!="after|closer"]within<s/>
```

Q2
```
[tag="ADJ.*"] []{0,5}
[tag="NOM|NAM" & word!="than|for|4"]
[]{0,5} [word!="et|ever|for|than|4"]
[word="ever|EVER"] [word!="after|closer"] within <s/>
```

Q3
```
[tag="NOM|NAM" & word!="than|for|4"] []{0,5}
[tag="ADJ.*"] [word="ever|EVER"] within <s/>
```

Q1 and Q2 represent the queries for pre-nominal adjectives (used in all four languages) and Q3 the query for post-nominal adjectives (used only in the Romance part of the sample). Q1 will find "simple" examples of ES-*ever* as in (2), which may contain additional pre-nominal material.

(2) de: der höflichste, freundlichste, teamwilligste Camper ever
the kindest, friendliest, most teamwilling camper ever
'the kindest, friendliest, most teamwilling camper ever'

Q1 translates to: "Find some adjectival form; followed by a span of zero to five unspecified word-forms; followed by a common or proper noun which may not have the form "than", "for", or "4"; followed by the word-form "ever" or "EVER"; not followed by the word-forms "after" or "closer". All within one sentence."

Q2 will find examples with material between a noun and *ever*, see (3).

(3) de: die wohl beste Werbung für das Hotel ihrer Eltern ever
the probably best advertisment for the hotel of.her parents ever
'the apparently best advertisement for her parents' hotel'

Q2 translates to: "Find a string which begins by some form of an adjective; followed by zero to five occurrences of some undefined wordform; followed by a noun which is either tagged as a noun or a proper name but has not the form "than", "for", or "4"; followed by zero to five occurrences of some undefined word; followed by the wordforms "ever" or "EVER", which may not be succeeded by the wordforms "after" or "closer". All within one sentence."

Q3 will find examples such as the following, featuring a post-nominal adjective and an optional post-nominal span of arbitrary words:

(4) fr: l'un des joueurs les plus fragiles ever
the one of.the players the most fragile ever
'one of the most fragile players ever'

Q3 translates to: "Find a string which begins by some common or proper noun, which may not have the wordform "than", "for", or "4"; followed by followed by zero to five occurrences of some undefined wordform; followed by some adjective; followed by the wordforms "ever" or "EVER". All within one sentence."

The reasoning behind the queries is the following:

- **Reference to an unspecified adjectival form**: We decided to not restrict the query to only include morphological superlatives, which would be supported by the tagging system, as not all languages attest synthetic superlatives for all adjectives and we wanted to evaluate if indeed all instances of ES-*ever* included a superlative form.

- **Reference to "than" "for" and "4", "after", "closer"**: in prior versions of this query we identified quite a lot of false hits, i.e. fixed expression, which included irrelevant instances of *for/4 ever*, *than ever*, *ever closer union* etc.

- **Reference to zero to five occurrences of some undefined word forms**: These two wild cards of varying length allow us to account for recursion, like additional pre- or post-nominal material such as PPs, additional adjectives, adverbs, relative clauses etc. We restricted the span to maximally five items, because enlarging the span (i) leads strings matching the query multiple times, and (ii) increases the amount of false hits.

After extracting all the hits found by the queries, we evaluated them by hand in order to sort out any false hits. Subsequently we collected all results for each language into one set via the identifiers of each hit. This allowed us to further analyze the results according to their geographical origin. Finally, we exported them to .csv files and tagged them according to the following criteria:[1]

- Pattern inside the noun phrase in terms of POS tags

- Type of superlative, i.e. analytic (*most beautiful*), synthetic (*best*) or inherent (*absolute*)

- positioning of the adjective, i.e. pre- or post-nominal

- Stage of nativization (see Section 2.2)

2.2 Results and discussion

After merging the results of each query by language, eliminating false hits by hand and removing duplicates, we received 369 instances of ES-*ever* for Dutch (0.23 hits per million tokens), 2,230 for German (0.26 hits per million tokens, hmt), 159 for French (0.02 hmt) and 120 for Spanish (0.01 hmt).[2]

Generally speaking, ES-*ever* is strikingly more frequent in the Germanic than in the Romance languages of our sample. We can further correlate the occurrences of ES-*ever* with their source country by looking at the top-level domain of the site an utterance is taken from. We show the variation for some countries in Table 1 for Dutch and German, and in Table 2 for French and Spanish.

[1] An exception in this approach was German: Since we were left with 2,230 hits in German after excluding false positives, we decided to create a random sample of 300 items for evaluation.

[2] The data extracted for this study is available at: https://osf.io/juewa/?view_only=215970c573d34b148815cc5653965697

Country	*N*	hmt
Netherlands	266	0.25
Belgium	79	0.18
All countries	369	0.23

Country	*N*	hmt
Germany	1,661	0.26
Austria	142	0.27
Switzerland	187	0.29
All countries	2,230	0.26

Table 1: Frequencies in Dutch and German

Country	*N*	hmt
France	77	0.02
Canada	27	0.05
Belgium	3	< 0.01
Switzerland	1	< 0.01
All countries:	159	0.02

Country	*N*	hmt
Spain	44	< 0.01
Mexico	13	< 0.01
Chile	13	0.02
Peru	8	0.01
All countries:	120	0.01

Table 2: Frequencies in French and Spanish

Canada is English-French bilingual and shows the highest relative frequency of ES-*ever* in the French data, but this value is still much lower than for the Germanic languages. For Belgium, the relative frequency for both Dutch and French is lower than the overall results for these languages. For Switzerland, the relative frequency for German is the highest in the table, but for French, it is among the lowest. This shows that societal bilingualism does not explain the variation among the French-speaking countries. We tentatively conclude that the Germanic-Romance contrast is the prominent, consistent, determining factor of the frequency of ES-*ever* in our data.

As mentioned in the previous section we classified the extracted and hand-sorted findings according in four structurally distinct stages of nativization: Stage-0 contains fully English expressions, like (5). As there is no interaction between the grammar of the noun phrase in the matrix and the source language in Stage-0, we will ignore this stage in rest of this paper. Stage-1 refers to expressions which contain an uninflected English adjective followed by the noun, see (6). An expression that features a matrix language determiner, noun and adjective was sorted under Stage-2 (cf. (7)). Finally, expressions that are additionally incompatible with the English source language grammar were collected under Stage-3, see (8).

(5) de: Best party ever

(6) es: el worst deal ever
the worst deal ever 'the worst deal ever'

Language	Stage-1		Stage-2		Stage-3	
Dutch	38	(10%)	276	(75%)	1	(< 1%)
German	3	(1%)	287	(96%)	7	(2%)
French	14	(9%)	123	(59%)	15	(9%)
Spanish	11	(1%)	71	(59%)	33	(28%)

Table 3: Distribution of the nativization stages

(7) fr: la plus belle fin de chanson ever
the most beautiful end of song ever
'the most beautiful song ending ever'

(8) nl: Gisteren de allermooiste babyshower gehad ever ...
yesterday the utmost beautiful babyshower had.PTCP ever
'Yesterday (we) had the utmost beautiful babyshower ever ...'

For the Romance languages, which allow for pre- and post-nominal adjectives we found that pre-nominal adjectives are more common, yet post-nominal adjectives are clearly possible: Our French sample contained 144 ES-*ever* instances featuring a pre-nominal adjective and 14 featuring a post-nominal adjective, as in (4) above. Our Spanish data revealed 88 ES-*ever* instances with a pre-nominal adjective and 32 with a post-nominal adjective. We classified all instances of post-nominal adjectives as belonging to Stage-3, as post-nominal uses of adjectives in English are rather rare, but see Section 3.

For every language we found examples from every stage. The numbers and percentages are given in Table 3. It shows that Stage-2 is the most frequently represented stage in all four languages. The higher percentage of Stage-3 cases in the Romance languages is primarily due to the use of ES-*ever* with post-nominal adjectives.

3 Syntax of emphatic superlative expressions

In this section, we will first look at the general syntax of adjectival modification in the languages under discussion. Then, we will show how the matrix language emphatic superlative expressions are integrated.

Pollard & Sag (1994) analyzed adjectival modification in English as an AP combining with a nominal category that is saturated for complements. Sadler & Arnold (1994) show that this is not adequate since pre-nominal adjectives are rather restricted in their complexity. For example, they allow degree particles but no complements (*the very proud (*of their kids) parents*). In contrast to this, post-nominal adjectives can show full complexity (*the parents very proud of their kids*). Consequently, Sadler & Arnold analyze only post-nominal adjectives as full APs that combine with a complement-saturated nominal projection.

Combinations with pre-nomial adjectives are treated as "small constructions," for which they propose that an A^0 category combines with an N^0 head.

Abeillé & Godard (2000) argue that the structures of French adjectival modification are analogous to those found in English. French also has severe restrictions on the syntactic complexity of pre-nominal adjectives, but none on post-nominal adjectives. However, only very few adjectives can occur pre-nominally at all. Abeillé & Godard use a head feature WEIGHT whose value is *light* for pre-nominal adjectives, i.e. for what Sadler & Arnold call "small constructions." Post-nominal adjectives have the WEIGHT value *non-light*. Machicao y Priemer & Winckel (2015) propose that Spanish can be analyzed like French, though with an even smaller set of pre-nominal adjectives.

In Dutch and German, all adjectives occur pre-nominally, allowing full APs in prenominal position, i.e., an analysis like that proposed in Pollard & Sag (1994) is unproblematic for Dutch and German, see (9).

(9) nl: Nederland is een met zichzelf tevreden natie, ...
The Netherlands is a with itself content nation
'The Netherlands is a nation content with itself.'

To sum up, the grammar of English adjective placement is like that of French and Spanish, even though, most adjectives occur in pre-nominal position which, therefore, make English look more like Dutch and German from the point of frequency. This means that in all five languages under discussion we find the same word order for the NP *the best book*. However, in English, French, and Spanish, *best book* is a "small construction," i.e. an $[A^0\ N^0]$ combination, whereas it is an $[AP\ N']$ combination in Dutch and German.

We can now turn to the syntax of emphatic superlative expressions from (1) in the languages looked at in this paper. Just like *ever* in English, they all occur post-nominally. Therefore, we can assume the same syntactic position for these expressions in all of our languages. In English, ES-*ever* has the same syntax as other NP-internally extraposed degree phrases. We will follow the extraposition analysis developed in Kay & Sag (2012) for the obligatorily extraposed clauses introduced by degree particles such as *so*. Their analysis is sketched in (10).

(10) en: [[so willing to help out][that they called early]]

The degree particle *so* selects the *that* clause via a list-valued feature EXTRA. In a phrase, the EXTRA values of all daughters are concatenated, unless it is an extraposition structure. There, the non-head daughter corresponds to the first element of the head daughter's EXTRA list, and the mother's EXTRA value is the rest of that list.

Just as degree particles can select extraposed clauses, we assume that elements with superlative semantics can select an emphatic superlative expression via the EXTRA list. These come in two groups. The first group consists of elements that introduce a morpho-syntactic superlative form: a superlative particle

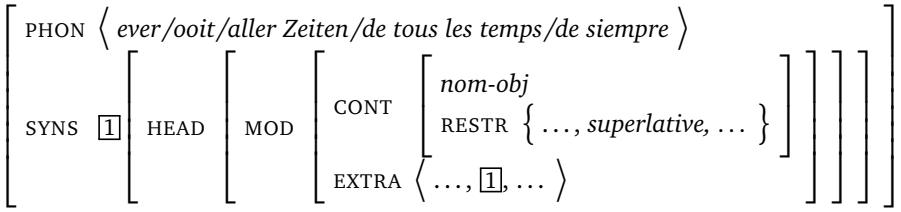

Figure 1: Sketch of an emphatic superlative expression

(like English *most* or French *plus* 'most'), a superlative morpheme attached to the positive form of an adjective by some lexical rule (like the morpheme *-st* in Dutch, English, and German), or a suppletive superlative form (like English *best* or Spanish *peor* 'worst'). Such morpho-syntactic superlatives can alternatively select a clausal comparative class expression, see (11).

(11) en: the best book ever/ of all times/ [that I have read in a long time]

The second group of expressions that can introduce an emphatic superlative expression are purely semantic/pragmatic superlatives. These include adjectives in their positive form if they have a superlative-like semantics such as *favorite* or Spanish *único* 'only', but also top-degree nouns such as *highlight* or the combination of a noun with a top-degree prefixoid like *top-* or German *Lieblings-N* 'favorite N'. For items of the second group, the occurrence of a clausal comparative class expression is less typical, though not ungrammatical, see (12)

(12) en: In the gallery, you'll find 15 of our favorite essays that we published this year. (Timestamped JSI English 2014–2021)

We will largely ignore the semantics and pragmatics of emphatic superlative expressions in this paper. Clearly, they indicate the comparison class of the superlative operator having the effect of a domain widening. Pragmatically, this has the effect of stressing the extraordinary degree to which the property in the scope of the superlative operator holds. This pragmatic effect may vary between different expressions.

We sketch a description of an emphatic superlative expression in Figure 1. Note that it modifies an element whose semantics is of the sort *nominal-object*, i.e., a noun or an adjective. The modified element must have a superlative operator on its RESTR list. Finally, the SYNSEM value of the emphatic superlative expression, 1, must be on the EXTRA list of the modified element.[3]

In Figure 2 we sketch the resulting syntactic structure for the English NP *the best book ever* and its French and Spanish translations. The structure for the Dutch and German equivalents (*het beste boek ooit/das beste Buch* [*aller Zeiten*]) looks the same, but the adjective *beste* would be an AP and the noun *boek/Buch* would be required to have an empty COMPS list.

[3]From here on, we will drastically simplify the AVMs, ignoring, for example the SYNSEM feature. Even when not displaying it, we commit to a full, standard HPSG feature geometry.

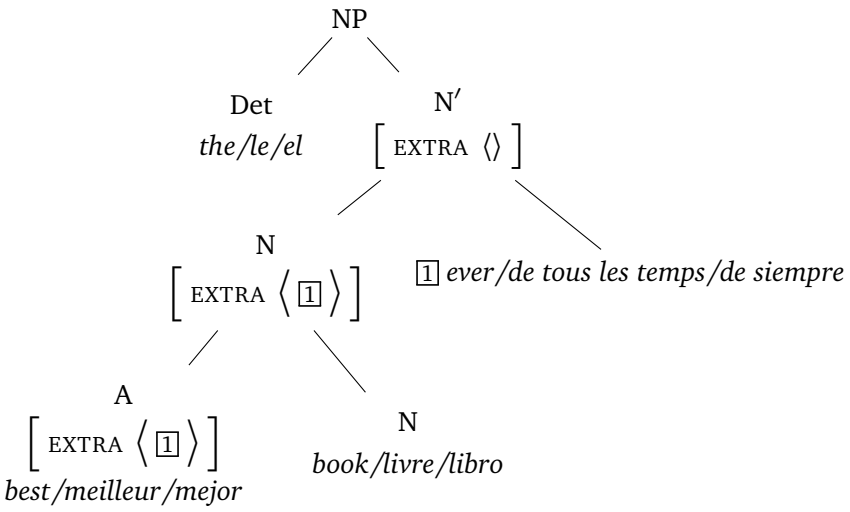

Figure 2: Sketch of the structure of an NP with emphatic superlative (en/fr/es)

4 A model of borrowing

In this section, we will turn to the main theoretical contribution of this paper, a development of a modelling of borrowing in HPSG. We will model the data presented in Section 2 as a three-stage process, which reflects the stages described in Section 2. We provide a Spanish example NP for each of these stages in (13).

(13) a. es: el coolest job *ever* 'the coolest job ever'
b. es: la mejor canción *ever* 'the best song ever'
c. es: las portadas más photoshopeadas *ever*

For the first stage, we postulate that a structure of the form "A N *ever*" is borrowed from English (constructional borrowing). In the second stage, we still only find the combination with pre-nominal adjectives, but matrix language lexical material is used, leaving *ever* as the only English item. Finally, *ever* is turned into a lexical borrowing and can occupy all positions found for emphatic superlative expressions in the matrix language. This means that ES-*ever* can occur with post-nominal adjectives in French and Spanish from that stage on.

In Section 4.1, we will sketch the two existing approaches whose ideas we (partly) incorporate in our own approach in Section 4.2. Finally, we will show how the three steps sketched in (13) are expressed in our model (Section 4.3).

4.1 Background

We will embed in our approach some fundamental concepts of *Diasystematic Construction Grammar* (DCxG, Höder 2012, 2018) and of the *Communicative*

Situations approach CSA, Wiese 2021). Common to them is the assumption that the linguistic knowledge of multilectal and multilingual language users consists of a single repertoire, containing all elements of the apparently different linguistic system available to them. Each of these elements is, however, marked for the varieties to which they belong, or the communicative contexts in which they are appropriate. However, while DCxG is *usage based*, CSA is *competence based*. We will side with the latter.

Clearly, a single item can be appropriately used in a number of communicative contexts. In this case, DCxG assumes that items can be specific or unspecific for a particular variety, with unspecific items (so-called *diaconstructions*) being compatible with more than one. In this approach, the borrowing of ES-*ever* could be modelled in the following way: ES-*ever* starts off as specific for English and is, then, turned into a diaconstruction and can appear in structures which are similar between English and the matrix language. However, we saw that in French and Spanish, ES-*ever* occurs with post-nominal adjectives, a position barely found in English. Consequently, there must, eventually, be a language-specific variant of borrowed ES-*ever*. Nonetheless, we will assume something along the lines of a diaconstruction for Stage-2 as in (13b).

Wiese's (2021) CSA shares many basic assumptions with DCxG, but takes a competence-based stance and is less devoted to a particular framework for expressing linguistic generalizations, though she uses Jackendoff's *Parallal Architecture* framework, summarized for example in Jackendoff (2007). Wiese annotates each unit of linguistic knowledge for the *communicative situation* (ComSit) in which this unit is usually, and recurrently encountered. These ComSits can be of any degree of concreteness or abstractness. Wiese argues that the notion of a "named language" or a particular "named dialect" can be understood as very abstract socially constructed entities, and, consequently, as ComSits.

Wiese (2021, Section 2.4) provides an example of a lexical borrowing from English to German. She argues that the English word *chicken* occurs as a borrowing in German only as chicken meat and in the context of diners. I.e., the English word is integrated into German with a subpart of its English meanings and in special types of communicative situations. At the same time, the German word *chicken* is still explicitly marked as a borrowing from the (semantically and situationally more general) English word *chicken* (Wiese, 2021, 14).

If we want to adopt this approach to the borrowing of ES-*ever*, we could say that English, more general *ever* is borrowed only in its emphatic superlative particle use into the discussed matrix languages. It is not clear to us if the borrowing of a particular meaning of *ever* can be as plausibly attributed to the communicative situations in which the word is used as in the case of *chicken*, where a particular food item is connected to the places in which it is usually consumed. Nonetheless, we will, in fact, model the final borrowing stage, i.e. ES-*ever* as in (13c), in a way that is an HPSG-rendering of the CSA analysis just sketched. We will also make the link between the borrowing from "English" situations to matrix language situations even more explicit than Wiese (2021).

4.2 An HPSG-approach to borrowing in HPSG

In this section, we will provide the basic architecture for our modelling of the borrowing of ES-*ever*. So far, there has been no work on borrowing in HPSG to our knowledge. However, there has been work on social meaning and/or register.[4] Under the assumption that language users have just one all-comprising grammar with marking for communicative situations, borrowing from one language to another is nothing else but, as in Wiese's (2021) *chicken* example, making an element from ComSits typical for one language available in ComSits typical for another language. Within the CSA, there is, thus, no principled difference between borrowing among languages and among registers, or from an item acquiring a new social meaning.

In this paper, we will follow the pragmatic tradition of the modelling of sociolinguistic aspects of language, based on the architecture of CONTEXT proposed in Green (1994), and applied to diglossia in Paolillo (2000). Asadpour et al. (2022) provide a recent incarnation of this approach, modelling regional and register variation in the realization of relative clauses in English and Kurdish.[5] They assume that social meaning takes the form of statements as in (14).

(14) (X believes that) X and Y mutually believe that community Z normally believes that expression E signals ϕ. (Asadpour et al., 2022, 18)

Such statements have the formal status of *conventional implicatures* (Grice, 1975; Potts, 2005), or rather of *expressive/use-conditional meaning* (Potts, 2007; Gutzmann, 2013), i.e., we take it that social meaning has the following properties: its truth conditions are independent of the at-issue content; it relates to the current utterance situation (non-displaceability), usually it expresses something about the speaker (perspective dependency); it is hard to paraphrase explicitly (descriptive ineffability); it performs its meaning simply by being uttered (immediacy); and using several items with the same social meaning reinforces their effect rather than being perceived as redundant (repeatability).

In the pragmatic approach, statements as (14) are introduced as elements of the projective, non-at-issue content. Green (1994) follows Pollard & Sag (1994) in using the set BACKGROUND for this. Asadpour et al. (2022) assume that there can be different types of projective meaning, each of which having its own set- or list-valued attribute. Therefore, they use attributes PRESUPPOSITION and CONVENTIONAL-IMPLICATURE (CI) instead of a single BACKGROUND feature.[6] Social meaning statements are, then, treated as elements in a sign's CI value.

[4] We adopt the view that the *social meaning* of a linguistic entity is the knowledge of its typical/-conventionalized association with particular communicative situations; a *register* is the subset of a language user's linguistic repertoire consistent with a particular communicative situation.

[5] For reasons of space, we cannot elaborate on the differences to other HPSG approaches such as Wilcock (1999), Bender (2007), and Machicao y Priemer et al. (2022).

[6] See Sailer & Am-David (2016) and Rizea & Sailer (2020) for other applications of this refined structure of the CONTEXT feature.

In the context of the present paper, we look at particular types of social meaning statements, namely at statements of the form in (15), i.e., that the social meaning associated with an expression is marked for a particular named language (the matrix language or the source language, here English).

(15) (X believes that) X and Y mutually believe that the speech community of the matrix language normally believes that expression *E* signals that X and Y are in a matrix-language communicative situation/ in an English communicative situation.

We will use abbreviated forms of social meaning statements in this paper which ignore the various embeddings of attitude predicates. An example of such an abbreviated form is given in (16). The AVM in (16) is part of the lexical entry of English *ever*. It expresses that the word *ever* is perceived as signaling communicative situations in which English is typically used. The relevant element in the CI set specifies the C(OMM)-SIT value as *en*. In an UTT(ERANCE) value it indicates which expression is marked for this communicative situation.[7]

(16)
$$\left[\begin{array}{ll} \text{PHON} & \boxed{1}\ \textit{ever} \\ \text{CTXT} & \left[\text{CI}\ \left\langle \ldots, \left[\begin{array}{ll}\text{C-SIT} & \textit{en}\\ \text{UTT} & \boxed{1}\end{array}\right], \ldots \right\rangle\right] \end{array}\right]$$

Such a marking for C-SIT is part of each linguistic expression. An utterance is fully English, for example, when all linguistic expressions in it have an element of the form $\left[\ \text{C-SIT}\ \ \textit{en}\ \right]$ in their CI set. If we find elements with different C-SIT specifications, this does not lead to an ungrammatical utterance, but simply to one that is not purely monolingual.[8]

We will ignore the problematic difference between borrowing and code-switching here, and, instead, assume that what we find in ES-*ever* is a development from what Muysken (2000, 72) calls *conventionalized code mixing* to an *establish loan*. Muysken (2000) uses the term *borrowing* exclusively for lexical material that is, in clear cases, morphologically integrated into the matrix language. We will, instead, use the term *borrowing* for conventionalized or "listed" elements of mixing of any level of linguistic complexity – here, at the word and the phrasal level. This reflects the rather constructional approach taken in our paper. We would use the term *codeswitching*, then, for the non-listed cases, i.e., for cases in which word- or phrase-level elements of different languages co-occur spontaneously – though not randomly, but rather in well-defined patterns, as elaborated in Muysken (2000).

[7] We use the PHON value of an expression as value for UTT. Green (1994) even includes the entire sign, not just its PHON value. Similarly, a semantic type *u* for *utterance* is often assumed in the literature on quotes and meta-linguistic language use, such as in Potts (2005), for example.

[8] We assume a supertype *lang* for named languages, with subtypes *de*, *en*, *es*, *fr*, and *nl* for the languages discussed here.

We assume that a borrowed element is considered part of the matrix language, but that it is still connected to the corresponding element from the source language, just as in Wiese's (2021) analysis of the German use of *chicken*. We treat the information on borrowing as a kind of social meaning, i.e., communication participants can have mutual beliefs on what a borrowing is. This means that we will have additional elements in the CI values of borrowed expressions that note the borrowing property.

Borrowings need not satisfy the principles of grammar of the matrix language. To achieve this, they are marked as idiosyncratic signs by means of the specification $\left[\text{ COLL } \textit{irregular} \right]$, introduced in Richter & Sailer (2009, 307). This exempts them from the regular principles of grammar (such as the Head Feature Principle or the Immediate Dominance Principle).

We introduce a sort *borrowing* with at least two subsorts *lexical-borrowing* (*l-borrow*), and *constructional-borrowing* (*cx-borrow*). Each *borrowing* object has an attribute SOURCE encoding information on the source element, and an attribute TARGET with information on the target element.

In the case of lexical borrowing, these will be the lexical identifiers of the related items, i.e., the commonly used LID value or the LISTEME value of Soehn (2006).[9] For the *chicken* example, we would assume a LID value *chicken_en* for the English chicken and a LID value *chicken_de* for its borrowed version. The latter would contain a C-SIT specification as German, but also a *l-borrow* object indicating that *chicken_en* is the source, and *chicken_de* is the target.

For constructional borrowing, the source and the target are more complex. We assume that the SOURCE value is a list of signs. This captures the observation that borrowed constructions need not be structurally analyzable by the borrower. The TARGET value is the (possibly phonologically adapted) concatenation of the PHON values of the elements in the SOURCE list. The TARGET value is identical with the PHON of the overall phrase.

4.3 Modelling the borrowing steps

In this subsection, we will use the machinery introduced in Section 4.2 to model the borrowing stages for ES-*ever*, as illustrated in (13).

Stage-1: Constructional borrowing (es: *el coolest job ever* 'the coolest job ever') In Stage-1, we find an English adjective with synthetic superlative combining with a noun and ES-*ever*, as in (13a). The noun is typically an English word that can also occur freely in Spanish utterances, whereas the adjectives can't. We model this stage by the borrowing construction in Figure 3.

[9] Whereas LID is assumed to be a head feature, LISTEME is not. This allows Soehn (2006) to maintain the Head-Feature Principle within a default-free grammar even for idiomatic expressions, where the lexical identifier changes, but morpho-syntactic properties percolate regularly.

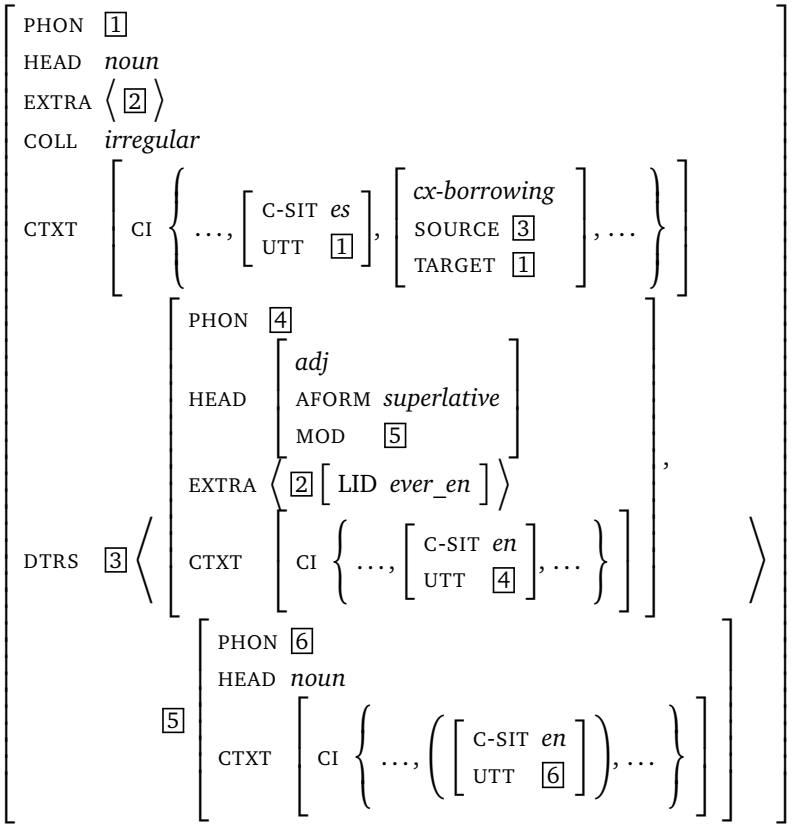

Figure 3: Description of the borrowing construction for Stage-1

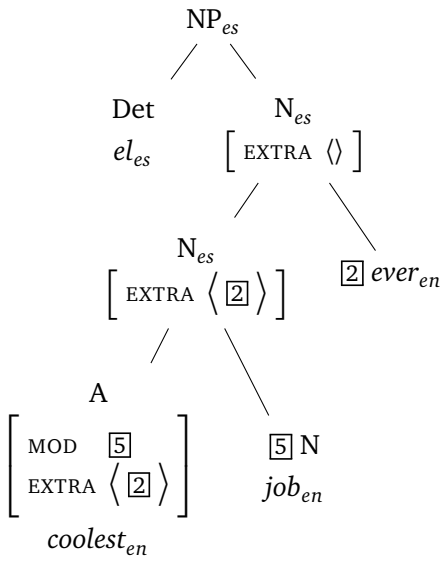

Figure 4: Structure of the Spanish NP *el coolest job ever* 'the coolest job ever'

The source, 3, consists of two daughters. The first daughter is a superlative adjective that is marked for English communicative situations and that modifies the second daughter (tag 5), and has the English ES-*ever* (2) on its EXTRA list. The second daughter is a noun that is also marked as English. The construction does not specify whether any standard construction of the matrix language grammar is used to combine the daughters. In fact, there need not be any, as the overall phrase is marked as *irregular*.

English ES-*ever* (2) occurs on the EXTRA list of the embedded adjective. It is explicitly inherited by the overall construction. Since Spanish has an NP-internal extraposition in its grammar, ES-*ever* can combine with the construction in Figure 3 by an ordinary construction of Spanish. Note, however, that within that construction, the extraposed element is the English word *ever*.

We sketch the resulting tree in Figure 4. The subscript on words and phrases indicates the C-SIT value associated with the phonology of those signs. The node dominating *coolest job* is licensed by the borrowing construction in Figure 3.

Stage-2: Diaconstruction (es: *la mejor canción ever* 'the best song ever') In this step, we will adopt the a DCxG-style modelling. While the borrowing construction in Figure 3 does not specify any immediate dominance schema, the word order A N is compatible with the matrix grammar: An A-N combination in French and Spanish, and AP-N combination in Dutch and German. In line with the ideas of DCxG, we assume that a more abstract, less idiosyncratic construction emerges, which is syntactically regular, allows for matrix language lexical elements, but has as its only idiosyncratic property the requirement that

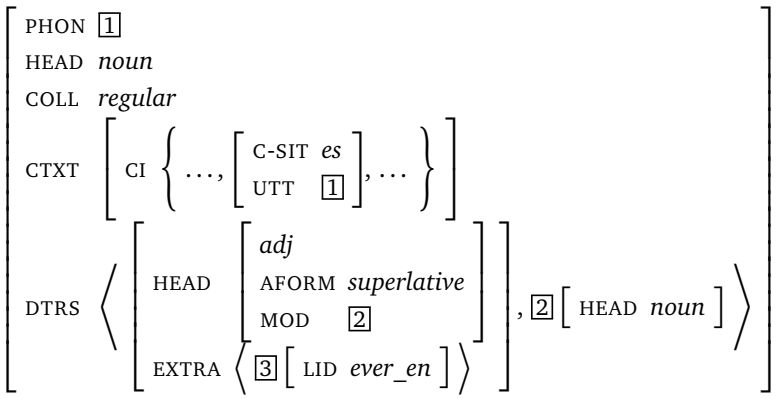

Figure 5: Description of the diaconstruction for Stage-2

the adjective introduces the English word *ever*. We provide the constraint on this construction in Figure 5.

The construction in Figure 5 is marked as *regular* and as adequate for Spanish communicative situations. Consequently, all Spanish principles of grammar apply. It specifies two daughters: a superlative adjective, followed by a noun modified by that adjective. Within the Spanish grammar this restricts the set of possible adjectives. The adjective is, furthermore, required to have the English word *ever* on its EXTRA list. Since the phrase is regular, we need not specify that the EXTRA value percolates from the adjective to the mother.

Stage-3: Lexical borrowing (es: *las portadas más photoshopeadas ever* 'the most photoshoped portals ever') In Stage-2, the English word *ever* was the only idiosyncratic item in the construction. In Stage-3, this lexical item gets borrowed into the matrix language, but only with the particular meaning that it has in this construction. Figure 6 shows the lexical entry of borrowed *ever*.

Borrowed ES-*ever* is specified as modifying an item that has a nominal object in its content with a superlative semantics. It must be on the EXTRA list of the modified element (indicated with tag 1). Borrowed ES-*ever* has its proper LID value *ever_es* and is marked as appropriate for Spanish communicative situations in the CI value. The CI value also expresses the pragmatic effect of ES-*ever*, skteched with the *excite(ment)* object as some high excitement of the speaker. The word is marked as a borrowing whose source is the English word *ever*, and whose target is the LID value of the matrix language ES-*ever* (3).

At this stage, ES-*ever* can occur in any position in which matrix language emphatic superlative expressions are possible. As a consequence of our analysis, all words and phrases in an NP like the one in (13c) are marked for matrix language communicative situations. The lexical item *ever* is specified as being related to English *ever*, but it is a matrix language word and its semantics and pragmatics is much more specific than that of its English source item.

The three-stage approach developed here allows us to capture a number of

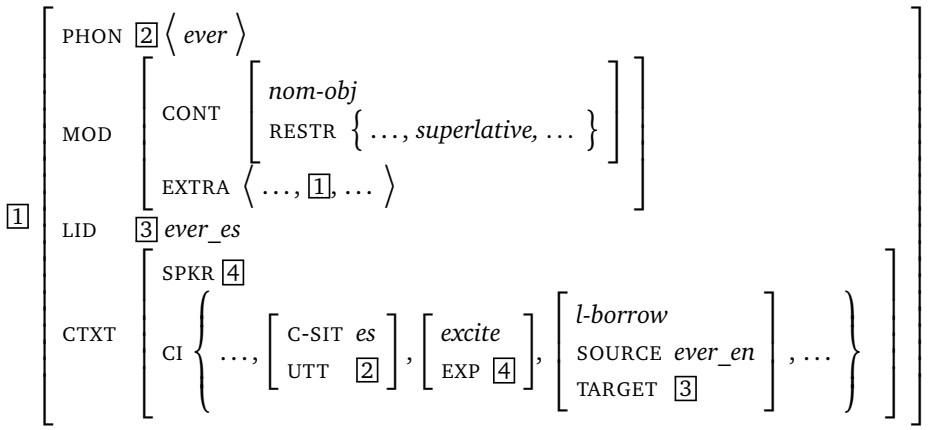

Figure 6: Lexical entry of matrix language ES-*ever* as a lexical borrowing

observations. First, the fact that some English adjectives can occur in the matrix languages exclusively when accompanied by *ever*. These cases are licensed by constructional borrowing. Second, the tendency in French and Spanish for using *ever* with pre-nominal adjectives, contrary to the overall dominance of post-nominal adjectives. This is achieved by the diaconstruction in Stage-2. Third, even though the syntactic rules of adjectival modification of English are more similar to those of French and Spanish than to those of Dutch and German, the surface combinations of English look more like those of the other Germanic languages than those of the Romance languages. The diaconstructional perspective in Stage-2 captures this, as the surface sequence adjective-noun is interpreted differently depending on the matrix language. The Dutch/German interpretation as AP N makes the structure immediately available for all adjectives of the matrix language, whereas the French/Spanish interpretation as A N opens it only for a restricted set of adjecitves.

Finally, and perhaps most importantly, the lexical borrowing is a generalization step resulting from the earlier two stages. In those stages, the ordinary English word *ever* is used, independently of whether it has an emphatic superlative semantics by itself or gets this interpretation through the use as modifying a superlative. In Stage-3, the new, matrix language word is "extracted" from the diaconstruction and, consequently, has the meaning that is specific to an emphatic superlative marker. Thus the question of why *ever* is borrowed in its emphatic superlative reading receives a natural answer as the lexical borrowing is just a final stage in a sequence starting with a constructional borrowing.

5 Conclusion

We discussed the occurrence of emphatic superlative *ever* in Dutch, French, German, and Spanish. Besides strong similarities, our corpus data revealed qual-

itative and quantitative differences, primarily between the Romance and the Germanic languages. We provided an analysis of the data within the pragmatic approach to social meaning in HPSG outlined in Asadpour et al. (2022). Adopting insights and techniques from Diasystematic Construction Grammar, and the Communicative Situation Approach, we proposed a three-stage modelling of the borrowing process that accounts for the observation that while, by now, ES-*ever* should be considered a lexical borrowing, it is only in that specific use that English *ever* occurs in all four languages we considered.

It is important that often, a structure that is compatible with a certain stage can also be analyzed according to the next stage. In particular, the Dutch and German data from Stage-2 are all also compatible with a Stage-3 analysis, i.e., we cannot tell whether ES-*ever* is introduced constructionally or lexically. We consider this a strength of our approach, as it allows us to model the transition from one step to the next in terms of a re-analysis of existing data.

This paper shows that Asadpour et al.'s (2022) approach to dialect and register variation can be extended to model phenomena of code-mixing, in particular conventionalized code-mixing such as constructional and lexical borrowing. Placing our analysis in the broader picture of code-mixing, we can distinguish between structures in which all expressions belong to the same language (our Stage-3) – even if some are marked as borrowings – and structures in which expressions belong to different languages (our Stage-1 and Stage-2). Within our particular technical implementation, the latter type of structures might be called instances of *codeswitching*. Though far beyond the scope of this paper, we consider this a first and promising step towards a formal integration of different phenomena of language (and register) mixing and shifting within HPSG.

References

Abeillé, Anne & Danièle Godard. 2000. French word order and lexical weight. In Robert Borsley (ed.), *The nature and function of syntactic categories* Syntax and Semantics, 325–360. San Diego: Academic Press.

Asadpour, Hiwa, Shene Hassan & Manfred Sailer. 2022. Non-*wh* relatives in English and Kurdish: Constraints on grammar and use. In Stefan Müller & Elodie Winckel (eds.), *Proceedings of the 29th International Conference on Head-Driven Phrase Structure Grammar, Online (Nagoya/Tokyo)*, 6–26. Frankfurt/Main: University Library. doi:10.21248/hpsg.2022.1.

Bender, Emily M. 2007. Socially meaningful syntactic variation in sign-based grammar. *English Language and Linguistics* 2(11). 347–381. doi:10.1017/S1360674307002286.

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.

Grice, Paul. 1975. Logic and conversation. In Peter Cole & Jerry L. Morgan (eds.), *Syntax and semantics*, vol. 3, 41–58. New York: Academic Press.

Gutzmann, Daniel. 2013. Expressives and beyond. In Daniel Gutzmann & Hans-Martin Gärtner (eds.), *Beyond expressives* (Current Research in the Semantics Pragmatics-Interface (CRiSPI) 28), 1–58. Brill. doi:10.1163/9789004183988_002. danielgutzmann.com/work/expressives-and-beyond.

Höder, Steffen. 2012. Multilingual constructions: A diasystematic approach to common structures. In Kurt Braunmüller & Christoph Gabriel (eds.), *Multilingual individuals and multilingual societies*, 241–257. Amsterdam, Philadelphia: Benjamins. doi:https://doi.org/10.1075/hsm.13.17hod.

Höder, Steffen. 2018. Grammar is community-specific: Background and basic concepts of Diasystematic Construction Grammar. In Hans C. Boas & Steffen Höder (eds.), *Constructions in contact: Constructional perspectives on contact phenomena in Germanic languages*, 37–70. Amsterdam: Benjamins. doi:https://doi.org/10.1075/cal.24.02hod.

Jackendoff, Ray. 2007. A Parallel Architecture perspective on language processing. *Brain Research* 1146. 2–22. doi:10.1016/j.brainres.2006.08.111.

Kay, Paul & Ivan A. Sag. 2012. Cleaning up the Big Mess: Discontinuous dependencies and complex determiners. In Hans C. Boas & Ivan A. Sag (eds.), *Sign-based construction grammar*, 229–256. Stanford: CSLI Publications.

Kilgarriff, Adam, VÃt Baisa, Jan Bušta, Miloš JakubÃček, Vojtěch Kovář, Jan Michelfeit, Pavel Rychlý & VÃt Suchomel. 2014. The Sketch Engine: Ten years on. *Lexicography* 1. 7–36. http://www.sketchengine.eu.

Muysken, Pieter. 2000. *Bilingual speech: A typology of code-mixing*. Cambridge: Cambridge University Press.

Paolillo, John C. 2000. Formalizing formality: An analysis of register variation in Sinhala. *Journal of Linguistics* 36(2). 215–259. doi:10.1017/S0022226700008148.

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.

Potts, Christopher. 2007. Into the conventional-implicature dimension. *Philosophy Compass* 4(2). 665–679.

Machicao y Priemer, Antonio, Stefan Müller, Roland Schäfer & Felix Bildhauer. 2022. Towards a treatment of register phenomena in HPSG. In Stefan

Müller & Elodie Winckel (eds.), *Proceedings of the 29th International Conference on Head-Driven Phrase Structure Grammar, Online (Nagoya/Tokyo)*, 86–101. Frankfurt/Main: University Library. doi:10.21248/hpsg.2022.5.

Machicao y Priemer, Antonio & Elodie Winckel. 2015. Adjectives within Spanish and French NPs: Position and interpretation. https://bond-lab.github.io/events/2015-hpsg/pdf/MachicaoyPrimer_Winckel.pdf.

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

Rizea, Monica-Mihaela & Manfred Sailer. 2020. A constraint-based modeling of negative polarity items in result clause constructions in Romanian. *Lingvisticae Investigationes* 43(1). 129–168. doi:https://doi-org.proxy.ub.uni-frankfurt.de/10.1075/li.00042.riz.

Sadler, Louise & Doug Arnold. 1994. Prenominal adjectives and the phrasal/lexical distinction. *Journal of Linguistics* 30(1). 187–226. doi:https://doi.org/10.1017/S0022226700016224.

Sailer, Manfred & Assif Am-David. 2016. Definite meaning and definite marking. In Doug Arnold, Miriam Butt, Berthold Crysmann, Tracy Holloway King & Stefan Müller (eds.), *Proceedings of the HeadLex 2016*, 641–661. Stanford: CSLI Publications. https://proceedings.hpsg.xyz/article/view/552.

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.

Wiese, Heike. 2021. Communicative situations as a basis for linguistic systems. *Working Papers in Urban Language & Literacies (WPULL)* 287.

Wilcock, Graham. 1999. Lexicalization of context. In Gert Webelhuth, Jean-Pierre Koenig & Andreas Kathol (eds.), *Lexical and constructional aspects of linguistic explanation*, 373–387. Stanford: CSLI Publications. http://www.ling.helsinki.fi/~gwilcock/Pubs/CSLI-98.ps.