Towards Interlingual Constructicography
On correspondence between constructicon resources for English and Swedish

Linnéa Bäckström, Benjamin Lyngfelt & Emma Sköldberg
Dept. of Swedish, University of Gothenburg

Abstract: This article addresses possibilities to link constructicon resources for different languages, in particular English and Swedish. The entries in Berkeley’s English constructicon have been compared to Swedish, with a focus on potential correspondents in a Swedish constructicon. In most cases, approximately corresponding Swedish constructions could be established, although typically with minor differences, often concerning grammatical markers. The closest equivalents are typically relatively general grammatical constructions, whereas those containing specific lexical elements tend to differ more. In order to link all correspondents between the two resources, a combination of strategies seems to be required. Constructions with a referential meaning may be linked via FrameNet frames, while those with a more abstract grammatical function may be related in terms of their grammatical properties.

Keywords: constructicon, construction, constructicography, English, Swedish, equivalence, correspondence, FrameNet, construction grammar, lexicography

1. Introduction

One of the central ideas in construction grammar (CxG) is the conception of language, or at least lexicogrammar, as a structured inventory of constructions: a constructicon. Although this notion has been around since the early days of CxG (e.g. Fillmore 1988), actual efforts to develop constructicons as descriptive resources have only recently been started. The first constructicon endeavor was initiated for English in Berkeley (Fillmore 2008, Fillmore et al. 2012), and there are now constructicon projects under way for Brazilian Portuguese (Torrent et al., this volume), Japanese (Ohara 2013), and Swedish (Bäckström et al. 2013, Lyngfelt et al. 2012, Sköldberg et al. 2013) as well. There are also plans for a German constructicon (Boas in press, Ziem et al. in press). Each constructicon is designed as a complementary resource to a FrameNet of the language in question.

The notion of constructicon thus displays the same kind of systematic polysemy as ‘grammar’ and ‘lexicon’. On the one hand, a constructicon is the network of constructions presumably present in the mind of a language user; on the other hand, it may be construed as a systematic description of the corresponding linguistic generalizations. In the first, more theoretical sense of the term, the constructicon is the ultimate object of study for CxG; whereas the second, more practical application is the domain of what we may call constructicography (in analogy with the distinction sometimes made between ‘lexicology’ and ‘lexicography’). Both deal with construction descriptions, although with somewhat different goals and, hence, methods. For instance, in aiming for large-scale coverage, a constructicon resource requires a simpler description format and less detail than the typical CxG analysis. In the following, we will be concerned with constructicons as resources, and, unless mentioned otherwise, this will be the intended sense of the term.

While each constructicon is primarily intended to be a monolingual resource, the existence of related projects for several languages also opens possibilities for cross-linguistic
application. The close relation to FrameNet and shared use of FrameNet methodology should be facilitating in this regard. A complicating factor, however, is that constructions are more or less language-specific. They are typically defined as “conventional, learned form-function pairings at varying levels of complexity and abstraction” (Goldberg 2013: 17); and, being conventions, such pairings are to some degree arbitrary. Although different languages often host strikingly similar constructions – due to mutual heritage, borrowing, or analogous developments – these are rarely entirely equivalent. However similar they may appear, there are typically some differences in form, meaning, function, or at least distribution (the latter possibly an indication of functional differences) to take into account. This is one of the main problems for anyone attempting to relate constructional resources across languages. For a discussion, see Boas 2010a (cf. also Croft 2001, chapter 1).

The same basic problem also applies to words, without stopping lexicographers from creating cross-linguistically useful lexical resources. However, grammatical constructions offer added complexity by containing constituent structure, dependency relations, and grammatical markers of various kinds. Hence, interlingual comparisons between constructions have to take internal structure into consideration; in the present context, any useful notion of equivalence or correspondence between constructions presupposes not only semantic/functional similarity, but also a degree of structural similarity. If, for instance, givenness is indicated by word order in one language and definiteness morphology in another, there is clearly a functional similarity between the two forms of expression. However, to assume those word order constructions and definiteness markers to be corresponding constructions would be stretching the notion of correspondence beyond the bounds of practicality. In fact, a degree of structural similarity is typically presupposed in interlingual lexicography as well: translational equivalents tend to be of the same grammatical category, or at least of categories with similar syntactic functions. This is roughly the degree of structural similarity required of corresponding constructions in the present work. Note that this restriction concerns the external function of the constructions, while correspondents may still differ greatly with respect to their internal structure.

As a first step towards connecting constructicon resources for different languages, we have worked through the entries in Berkeley’s English constructicon (henceforth, BCcn) from a Swedish perspective and, where applicable, developed corresponding entries in our own constructicon for Swedish (SweCcn). Based on this work, the present paper addresses some possibilities and difficulties for future interlingual constructicon development. A comparison between the constructions in BCcn and SweCcn is presented in section 2. In section 3, we turn to the issue of relating constructicon resources across languages and outline some preliminaries for interlingual constructicography.

2. Comparing constructicon entries for English and Swedish

The English constructicon at Berkeley is essentially a pilot resource, consisting of 73 construction (cxn) entries, 50 of them complete and 23 incomplete. For each of these, we considered possible Swedish correspondents, both in relation to the cxn descriptions themselves and based on how one would translate the examples presented. For satisfactory correspondence, at least partial equivalence in both form and function, and near equivalence in one of these respects, was required.

It should be noted that the comparisons concern the English and Swedish resources, not the entire languages. While the Swedish entries are all based on corpus studies conducted by ourselves, we do not make any claims about English, but simply assume the descriptions in BCcn to be adequate. Furthermore, due to the need for a relatively simple constructicon
format, the cxn descriptions are idealizations and do not account for the complexity of the
cxns in full detail. Thus, the comparisons do not aspire to match the level of specificity
typical of contrastive work in CxG (see e.g. the papers in Boas 2010b) or the typological
approach of Radical Construction Grammar (Croft 2001). On the other hand, we are able to
cover a larger number of constructions.

Corresponding Swedish constructions were established for a bit over 80% of the BCcn
entries. Most of the correspondences are one-to-one relations, but in a few cases a BCcn
entry rendered two SweCcns, or several BCcn entries were collapsed into one (see
section 2.4.1). As for the English cxns for which no satisfactory matches were found in
Swedish, most of these are either cases of partial correspondence, or specific subtypes within
a group of cxns where the other types could be matched (section 2.4.2).

In the following we will demonstrate some examples of both highly and less corresponding
cxns. Unsurprisingly, the closest equivalents are relatively abstract and general cxns, such as
COORDINATION and ADJECTIVE AS NOMINAL (section 2.1), whereas idioms and other cxns
containing lexically specific elements tend to differ more (section 2.3). Full equivalence is
rare, and most correspondence pairs display some minor differences, often concerning
grammatical markers for definiteness, gender etc. or relational expressions such as
prepositions and conjunctions (section 2.2). Minor functional and distributional differences
are also common, usually without seriously challenging the approximate equivalence.

Throughout, the constructions will be referred to by their names in BCcn, indicated in
small caps as above. The English examples sentences are from BCcn as well.

2.1. Relatively general grammatical constructions are often similar

English and Swedish are closely related languages, where many cxns are quite similar. The
most closely equivalent cxns typically represent relatively general grammatical structures. A
fairly straight-forward example of this tendency is the ellipsis construction GAPPING (1).

(1) a. Again, targets were set, times [ ] carefully recorded and fitness [ ] improved.
   b. Återigen blev mål uppsatta, tider [ ] noggrant nedtecknade och hälsan [ ] förbättrat.

The GAPPING cxn contains two or more conjuncts. Each conjunct consists of a Before
(targets, times, fitness) and an After (set, recorded, improved), which are separated by a
Gapped portion (were). The Gapped portion can be omitted in every conjunct but the first,
and has to be omitted in the final conjunct. Three minor differences can be noted in the
example: first, the placement of the finite verb (blev) directly after the initial adverbial in the
Swedish variant, due to the Swedish V2 pattern; second, the definite form hälsan in place of
the English indefinite fitness; and third, number agreement on the Swedish participles
uppsatta and nedtecknade. Neither is directly related to GAPPING as such, but rather following
from other cxns. V2 in particular will be addressed later in section 2.4.2.

Another construction with a fairly general function is the WITH_ABSOLUTE cxn, illustrated
in (2).

(2) a. With these events in mind, Naipaul wrote a novel.
   b. Med dessa händelser i åtanke skrev Naipaul en roman.
This is a type of absolute construction starting with the preposition with, or its Swedish counterpart med, respectively. The English and Swedish cxns are similar in both form and function. Although one might have expected differences due to the presence of a lexically specific element (cf. below), in this case this element is a general function word with roughly synonymous realizations in both languages. Admittedly, prepositions are notoriously multifunctional; if the full range of variation is taken into account, there are both distributional and functional differences between English with and Swedish med. Nevertheless, that is of no concern for the presently relevant senses of the words. In the present context, they are basically equivalent.

A lexically specific element is also present in RATHER THAN_COORDINATION. This cxn implicates a scenario where one conjunct is favored at the expense of another, as in (3).

(3)  a. Reliance has been placed on actuaries’ judgement rather than on constricting regulations.
    b. Tilltron har satts till aktuariers bedömning snarare än till begränsande regelverk.

The multi-word expression rather than corresponds to two different lexical units in Swedish: snarare än and hellre än. They differ slightly in that snarare än is used to express meta-comparison and factual tendencies, and hellre än is preferred when it comes to subjective preference. In English, rather than covers both functions. The difference is illustrated in (4), where (4a) could be interpreted either way, but the different readings render different translations in Swedish (4b-c).

(4)  a. But even then, she had listened rather than talked back.
    b. Men till och med då lyssnade hon snarare än att svara emot. (behavioral tendency)
    c. Men till och med då lyssnade hon hellre än att svara emot. (subjective preference)

Syntactically, however, the constructions are basically equivalent (although clausal complements are more often non-finite in Swedish, as illustrated by the contrast in (4)). For instance, both English and Swedish display the same word order variation, where the rather than expression – as well as either of its Swedish counterparts – can either occur between the conjuncts, as in (3–4), or split, as in (5).

(5)  a. Americans would rather have beer than wine.
    b. Amerikaner vill hellre ha öl än vin.

Since the difference between hellre än and snarare än has no grammatical consequences beyond lexical meaning, the coordination constructions as such have been treated as equivalent in English and Swedish. In cases where lexical differences affect the construction as a whole, we have rather assumed different constructions (cf. section 2.4).

It should be noted that the tendency for more general cxns to correspond more closely between English and Swedish than more lexically specific cxns do, is just that: a tendency. It is well known that there exist some quite general grammatical differences between the two languages, however closely related they are. Two striking exceptions in BCcn are the SUBJECT-AUXILIARY INVERSION and BE_PRESENT-PARTICIPLE, which will be addressed in section 2.4.2.
2.2. Differences typically concern grammatical markers

A large portion of the cxns, although often quite similar in other respects, differ somewhat in how grammatical properties such as definiteness, gender, and number are marked, and in the use of relational elements like prepositions and conjunctions. For example, consider the two time adverbials in (6), illustrating the LOCATION_INCALENDAR_UNIT cxn.

(6) a Mr Gorbatchev [...] called for a Helsinki 2 summit next year
b Gorbatchov begärde ett toppmöte nästa år

in Rome last week.
i Rom förra veckan.

In English, both temporal noun phrases are indefinite. In Swedish, however, while the noun phrase concerning future events is indefinite, the one indicating past time is definite. The determiners next and last are otherwise equivalent. Such minor but systematic differences between English and Swedish regarding definiteness marking are quite common.

In the case of ADJECTIVE_AS_NOMINAL, the English and Swedish versions differ with respect to the expression of number and gender, as well as definiteness. This is a group of cxns functioning as noun phrases despite lacking a head noun. They consist of a definite determiner and an adjective phrase, and refer to entities bearing the property denoted by the adjective. In BCen, as well as in SweCen, three subtypes are recognized: PEOPLE (7), ABSTRACT (8), and ANAPHORIC (9).

(7) a The rich live as fearful princes: the poor
b Det richa lever som rikda prinsar: de fattiga

live as angry beggars.
liver som arga tiggare.

(8) a. For Albert Edward the inevitable happened.
b. För Albert Edward skedde det omdviklig-a.

(9) a Conversely, the economies of the New World benefited immeasurably
b Omvänt gynnade ekonomi-er-na i den Nya Världen omätligt

corversely benefited-PASS economy-PL-DEF in Det new-DEF world-DEF immeasurably
by the exodus from the Old.

gennom utträdet ur den gamla.

The PEOPLE type always has plural reference to a (generic) set of people with the property expressed by the adjective, whereas the ABSTRACT type refers to an abstract property or entities characterized by this property. In the ANAPHORIC subtype, the interpretation of the noun phrase is linked to a correlate in the preceding context, for instance World/Världen in (9). Functionally, the cxns are virtually identical in English and Swedish, although Swedish seems to be more hospitable to the anaphoric type, whereas a resumptive pronoun (’one’) is often preferred in English.
Formally, all three types are identical in English, except for lexical differences between the adjectives. In Swedish, however, the gender and number of the noun phrase are marked on the determiner. The **people** cxn has plural marking (7), the **abstract** cxn is marked both for singular and for neuter gender (8), and the **anaphoric** cxn agrees with its correlate in both number and gender (9). Thus, unlike in English, the three types are formally distinct in Swedish. In addition, the adjective is marked as definite in the Swedish cxns, with no formal difference between the types.\(^5\) All of these formal differences between English and Swedish variants follow from general agreement patterns in Swedish; they are not directly related to **adjective** _as_ **nominal**.

Other differences in grammatical marking concern relational expressions. A typical example of this is **dimension_conjunction**, which represents measurements of the dimensions of an object. In English, the conjuncts are separated by the preposition *by*, as in (10a). In Swedish, however, the same function is filled by the noun *gånger* ‘times’, as in (10b). Also notice the genitive form *en femtedels tum* (‘a fifth of an inch’).

\begin{itemize}
\item (10) a A package approximately two inches long  
width  by a fifth of an inch deep.
\item b Ett paket ungefär två tim långt *gånger* en halv tum
width-Ø long-NEUT times a half inch
\end{itemize}

In Swedish dimension expressions with *gånger*, the unit of measurement is often omitted except in the last conjunct, as illustrated in (11).

\begin{itemize}
\item (11) Ett paket på två *gånger* en halv *gånger* en femtedel-s tum.
\end{itemize}

All of the above constructions are quite similar in both form and function. Hence, minor differences in grammatical marking of the kinds illustrated here might arguably be considered negligible. On the other hand, there are many situations where they really matter. In language technology applications, for instance, the exact form of expression is often of prime importance; in language pedagogy, idiomatic usage is of course a highly desirable goal, etc.

Consequently, while these cxns are clearly to be considered satisfactory equivalents for interlingual constructicon application, there are good reasons to also account for the minor formal differences between them.

### 2.3. Major differences typically occur in lexically specific, idiom-like constructions

If the closest equivalents typically are relatively general grammatical cxns, major idiomatic differences tend to correlate with specific lexical elements. Even for the more idiom-like cxns, however, there is usually a roughly corresponding expression in the other language, albeit differing to varying degrees in form and/or function. A case where the differences are relatively minor is the **ones_very_eyes** cxn.

\begin{itemize}
\item (12) a Gardeners had seen their life’s work destroyed
\item b Trädgårdsmästare hade sett sina livsverk förstöra-s destroy-PASS
\end{itemize}

\begin{itemize}
\item *mitt framför* ög-on-en på dem. 
\item right before eyes-DEF.PL on them
\end{itemize}
The English and Swedish cxns are functionally quite similar, both conveying a certain amount of knockdown expressivity. In Swedish the adverb mitt ‘right’ is crucial to obtain that effect, somewhat corresponding to the English use of very. Formally, although both versions display a few particular idiosyncrasies (such as the possessive pronoun and the modifier very in English, and the PP in Swedish), these are arranged around the central elements ‘before’ and ‘eyes’ in both languages.

A perhaps more striking difference concerns the often cited WAY cxn (cf. e.g. Goldberg 1995), of which there are three variants in BCcn. Presently, we are concerned with WAY_MEANS, where the verb expresses the means of motion.

(13) a But mum just **elbowed** her way past the two women.
    b Men mamma bara **armbågade sig** förbi de två kvinnor-or-na.
    c The Police **forced** **their way** into the pub.
    d Polis-en **trängde sig** in i puben.

The possessive pronoun + way in the English cxn (13a,c) correspond to a plain reflexive in Swedish (13b,d). Due to its famous ‘way’ feature, the English cxn is quite conspicuous. By contrast, since its Swedish counterpart has no such feature, it is simply a reflexive cxn among many others. Put differently, the Swedish cxn is perceived as part of a larger system, whereas the English cxn stands out on its own. Still, they both express the same basic meaning.

Another famous idiom (at least in the CxG literature) is the LET_ALONE cxn (cf. Filmore, Kay & O’Connor 1988), for which it is harder to establish a general Swedish correspondent:

(14) a Most wives are too bloody old, **let alone** mothers.
    b De flesta fruar är för jäska gamla, **för att inte tala om** mödrar.
    c Emigration to Australia or New Zealand, **let alone** the United States, which had its own extreme problems […] was little more than a trickle
    d Emigrationen till Australien eller Nya Zeeland, **för att inte tala om** USA, var knappt mer än en rännil.

The closest Swedish correspondent seems to be **för att inte tala om** (lit. ‘for to not speak of’, approx. ‘not to mention’); at least it is the most appropriate Swedish paraphrase in the majority of LET_ALONE examples in BCcn (14). This expression is, however, more restricted than **let alone**, and in many cases other alternatives are more appropriate, as in (15). Occasionally, no corresponding paraphrase is available, requiring the use of a wholly different type of cxn (15c–d).
From the viewpoint of BCcn, för att inte tala om is apparently the most closely corresponding Swedish paraphrase to let alone, but the converse relation does not hold. From a Swedish viewpoint, the closest English equivalent to för att inte tala om seems to be the almost cognate ‘not to mention’ rather than ‘let alone’. Hence, linking it to ‘let alone’ is simply a consequence of what happens to be present in BCcn. Nonetheless, we have established FÖR_ATT_INTE_TALA_OM as a cxn entry in SweCcn and will treat it as a correspondent to LET_ALONE for interlingual purposes. It is, after all, the closest equivalent there is. That it also happens to be the closest Swedish equivalent to ‘not to mention’ does not alter this fact.

A perhaps less idiom-like cxn, at least in English, is TAGGED_SENTENCE.CANONICAL. Tagged sentences occur in both English and Swedish, albeit in different ways. In English, the tag consists of an auxiliary verb + a pronoun referring to the subject. The tag also includes a negation, either cliticized to the auxiliary or following the subject, unless there is a negation in the matrix clause. In Swedish, by contrast, the tag is a fixed expression, eller hur ‘or how’. Some examples of TAGGED_SENTENCE.CANONICAL are presented in (16).

(16) a You are homeless, aren’t you?
   b Du är hemlös, eller hur?
   c He never learns, does he?
   d Han lär sig aldrig, eller hur?
   e A little vino loosens the conversation, does it not?
   d Lite vin lätter upp samtalet, eller hur?

Although the English and Swedish tags are quite different in form (except for their syntactic position), they are functionally equivalent. Therefore, it is reasonable to treat them as corresponding cxns. Likewise, satisfactory Swedish correspondents could be established for most of the BCcn entries, even the ones including lexically specific elements and other idiosyncracies.

2.4. Non one-to-one matches

There are, however, also quite a few cases which are not to be treated as more or less straightforward correspondence pairs. On the one hand, this concerns one-to-many and many-to-one mappings (section 2.4.1); on the other hand, some BCcn entries simply lack close enough correspondents in Swedish (section 2.4.2).

2.4.1. Different groupings in BCcn and SweCcn

In some cases, we have established Swedish correspondents to BCcn entries but not in a one-to-one relation. A striking example of this is the group of RATE cxns, where there are four entries in BCcn: FREQUENCY, MILEAGE, SPEED and COST_TIME, distinguished in terms of the domains they apply to. All four of them express a ratio relation and consist of two noun phrases, functioning as numerator and denominator (cf. Fillmore et al. 2012), joined either by an indefinite article or by the preposition per. RATE.FREQUENCY is illustrated in (17–18):
In both languages, there are two formal variants. Starting with English, there is one version where the relational expression is an indefinite article, as in (17), and one where it is the preposition per, as in (18). In both variants, the noun following the multiplier is indefinite. For the per variant, the Swedish correspondent is basically identical, even down to the preposition per, which both languages got from Latin.

The other Swedish variant, however, differs from both English and the per variant in three respects: the second noun phrase is definite, the relational expression is one of the prepositions i (‘in’) and om (‘about, around’), and the distribution is somewhat different. The choice of preposition depends on the following noun; i being the standard choice, but om occurring with dagen (‘the day’) and året (‘the year’). As regards distribution, i/om does not occur with MILAGE, while the version with per, as well as both English variants, apply to all four RATE domains.

In summary, where the formal variants in English are identical except for the lexical realization of the relational element, Swedish clearly employs two formally distinct RATE cnxs (called PROPORTION_I/OM and PROPORTION_PER in SweCcn). On the other hand, we do not assume different Swedish cnxs based on domain alone, as in BCcn. The latter choice is more of an editorial decision than due to any differences between English and Swedish. There are a few other cases where minor functional differences have led to cnx distinctions in BCcn but not in SweCcn, but this does probably not reflect any general difference in policy. One might venture a guess that the editors of BCcn would not have assumed eight RATE cnxs if they had also made a formal distinction.

One domain that requires special consideration is MILAGE. First, the default perspective of comparison in English and Swedish is reversed (although both variants are at least possible in both languages). In English, milage is typically expressed in terms of ‘distance per fuel unit’; where in Swedish the normal way of expression concerns ‘fuel consumption per distance’; compare (19a) and (b). Second, as already mentioned, it is incompatible with the Swedish I/OM cnx. Hence, there may be good reason to assume a third Swedish rate cnx: PROPORTION.FUEL.

(19) a. 100 miles per gallon
    b. 0.5 liter per mil
        0.5 litres per Sw. mile (=10 km)

In other cases one might either split an English cnx entry into two Swedish or treat one of them as a restricted partial correspondent and handle the other in a different context. One example of this dilemma is the BE_RECIP cnx:
The reciprocal relation expressed by this cxn is syntactically symmetrical, i.e. holding between two subject noun phrases, in (20) – but asymmetrical in (21), where one of them is oblique. In both cases, the reciprocal noun (friends) is plural in English. In Swedish, however, it is plural in the symmetrical case but singular in the asymmetrical.

If BCcn had assumed two distinct cxns for the two variants, the symmetrical cxn would clearly have a matching cxn in Swedish. However, without connection to the asymmetrical variant, would it even be considered a cxn, that is, a conventional pattern of its own? Its properties follow (a) from general properties of predicative complements, e.g. agreement with the (plural) subject, and (b) from general properties of reciprocals. Perhaps the combination of the two is common enough to be considered entrenched in the minds of the language users, but it is the asymmetrical pattern that displays a striking idiosyncrasy. If that had been considered a cxn of its own in BCcn, one could either assume that a corresponding cxn is lacking in Swedish or that there is a corresponding cxn in the singular. In the latter case, we again encounter the question of whether it should be considered a cxn of its own. The example in (21b) patterns perfectly with, say, Jag är god vän med presidenten (‘I’m a good friend of the president’), which suggests that it should not.

However, (20a) and (21a) do not represent distinct cxns in BCcn but variants of the same plural reciprocal cxn. In building a strictly monolingual constructicon we would not assume a corresponding cxn for Swedish. Nonetheless, for interlingual purposes it is clear what the Swedish correspondents for both the symmetrical and the asymmetrical cxns would be. That would amount to a one-to-two mapping between the relevant BCcn and SweCcn entries.

2.4.2. Deficient correspondence

Regarding English cxns that lack a satisfactory Swedish correspondent, the two most general ones in BCcn are SUBJECT-AUXILIARY_INVERSION (SAI) and BE_PRESENT-PARTICIPLE. SAI is a quite atypical case that will be addressed by the end of this section. As for BE_PRESENT-PARTICIPLE, there is a corresponding form in Swedish, but it lacks the general imperfectivity marking function of its English counterpart. In fact, Swedish has no such general imperfectivity cxn at all.9 More typical examples of deficient correspondence are cxns such as OWN_RIGHT (22a) and WHAT-WITH_ABSOLUTE (22b).

(22) a. Linguistics, like psychology, has grown up, and flown the nest of philosophy, to become a science in its own right.
   b. What with health budgets being pruned and cut back I'm asking the health board if staff shortages perhaps were a contributory factor here.
WITH_ABSOLUTE (cf. section 2.1). The latter does have a Swedish equivalent, and WHAT-WITH may be considered a special variant lacking in Swedish. There are, however, other Swedish expressions with a similar function, e.g. med tanke på (approx ‘considering’, lit. ‘with thought on’). With a more generous treatment of partial correspondence, med tanke på would be a likely candidate. At present, however, we do not assume any Swedish correspondent to the WHAT-WITH cxn.

Other examples of partial correspondence are ABSOLUTE_CLAUSE and the group of WAY cxns. Both are fairly general cxns with several variants, and their Swedish correspondents cover most but not all of them. For instance, the use of absolute -ing clauses is more restricted in Swedish than in English. This, however, has more to do with the -ing-forms than with absolute clauses.

A more critical difference, with far-reaching consequences, concerns subject-auxiliary inversion (SAI). In addition to the general SAI cxn, there are no less than six more specific inversion cxns in BCcn, four of which conform to Swedish. Note, however, that SAI is not a productive cxn in itself in English, but merely a generalization over those particular cxns that share this formal feature. In Swedish, however, there is a general V2 (verb second) feature, i.e. the finite verb occupies the second position of the clause. This means that Swedish gets “inversion” whenever the subject is not in first position, as illustrated in (23). One can of course construe this as a general SAI cxn in Swedish, but the notion of inversion presupposes a norm that the subject precedes the verb. There is better support for such a notion in English, which is more consistently an SVO language, than in Swedish.

(23) a. Igår spelade vi golf.
   yesterday played we golf
   ‘Yesterday, we played golf’

b. Henne känner jag inte.
   her know I not
   ‘Her, I don’t know / I don’t know her’

V2 is a characteristic feature of Germanic languages. English used to be a V2 language as well, but lost this property during the Middle English period (see Fischer et al. 2000, ch. 4). What remains in present-day English is the group of SAI cxns, arguably with little in common except this word order feature. Goldberg & Del Giudice (2005; see also Goldberg 2006, ch. 8) suggest that they form a category by sharing the property of non-prototypicality; but, given that they are remnants of a former pattern, we find it more probable that their non-prototypicality is simply the reason why they were not affected by the general word-order change, especially since there are a lot of non-prototypical clause patterns in English that do not display SAI.

In terms of cxn correspondence, English SAI and Swedish V2 – as general patterns – are formally similar, but differ too much in distribution to be considered corresponding cxns. On a more specific level, however, the majority of the English SAI cxns correspond to Swedish clause types of not only similar word order but also similar functions, as exemplified in (24–25).\(^\text{10}\) Those that do not, such as the EMPHATIC_NEGATIVE_IMPERATIVE cxn (ex: Don’t you dare move, boy!), depend on English do-support.

(24) a. Where could she be? (OPEN_INTERROGATIVE.NON-SUBJECT)

b. Var kan hon vara?

(25) a. Had it not been for human kindness he would have ended up in a pork pie. (SAI.CONDITIONAL)

b. Hade det inte varit för mänsklig vänlighet hade han slutat som köttpaj.
2.5. Summary

The above comparison illustrates how Swedish correspondents could be established for most of the BCcn entries. Most, if not all, of these correspondence pairs, however, are approximately rather than fully equivalent (see discussion in the following section). With a few notable exceptions, the most closely equivalent pairs are relatively general cxns, whereas cxns with lexically specific elements tend to differ more. The most common formal differences between corresponding English and Swedish cxns concern grammatical markers.

BCcn entries lacking Swedish correspondents are typically cases of partial correspondence or special variants of corresponding cxns. A few general differences between English and Swedish affect several different cxns, often indirectly. Only when they directly concern a cxn as such do they matter for the equivalence judgement.

3. Relating constructions across languages

In light of the above comparison, we will now address some preliminaries for interlingual application of constructicon resources. While finding corresponding Swedish cxns for over 80% of the BCcn entries may seem promising, one should not attach too much importance to this number. First, BCcn is small and cannot be regarded as totally representative of the English cxn repository in general. Second, what counts as a satisfactory correspondent is a somewhat subjective decision, dependent on the notion of correspondence employed (cf. section 3.1). Third, it is likely, if not inevitable, that distinctions in the source language influence the analyses of the target language, which are based on the search for similarities (cf. Svensén 2009: 253) In fact, this bias is further narrowed down to the source resource, in our case BCcn, including editorial decisions. On the other hand, the contrastive perspective has revealed special characteristics of the Swedish cxns that might have gone unnoticed using another method (cf. e.g. Colson 2008; Hannesdóttir 2012).

Finally, even assuming that 80% is an accurate measure, there still remain 20% unaccounted for by correspondence pairs. Furthermore, that gap will increase when languages less closely related than English and Swedish are also brought into the picture. To conclude, while establishing Swedish correspondents to the BCcn entries has been a both productive and revealing endeavor, direct correspondence pairs must be considered an insufficient method for connecting constructicon resources for different languages in a useful way. Alternative approaches will be addressed in the concluding section 3.2, after a discussion of the notions of equivalence and correspondence.

3.1. Equivalence and correspondence

The concept of equivalence is of central importance both in translation theory and in bilingual lexicography. In translation theory, the modern discussion starts with Nida’s (1964: 165ff.) well-known distinction between formal and dynamic equivalence (cf. e.g. Bassnett 2002: 32–38). According to Nida, a formal equivalence translation is basically source oriented; it is designed to reveal as much as possible of the form and content of the original message. A translation oriented toward dynamic equivalence, on the other hand, focuses on the receptor response. It is based on the principle of equivalent effect, i.e. that the relationship between receiver and message of the translation should aim at being the same as the one between the original receivers and the source language message.
As we all know, translation involves far more than mere replacement of words and phrases between languages. The lexical units in the source text cannot be regarded in isolation; the translators always have to take not only linguistic, but also cultural context into consideration (cf. Nida 1997). As a consequence, in some contexts, translators might find it suitable to translate an idiom in the source text with an idiom in the target text. But in other contexts, they choose a single word or a paraphrase, if that captures the original message in a better way. In other words, the concept of equivalence in translation theories covers relationships even on text and discourse level.

For our purposes, however, the lexicological/lexicographical concept of equivalence is more relevant. In bilingual lexicography, equivalence can, according to Svensén (2009: 255–256), be defined as the relationship between a source-language expression and a target-language expression with regard to meaning and usage (cf. semantic equivalence, involving both denotative and connotative meanings, and pragmatic equivalence) (see also Atkins & Rundell 2008: 467ff.; Adamska-Sałaciak 2010).

Farø (2004), discussing idioms, goes further than Svensén in his classification of equivalence types. Comparing pairs of idioms from different languages, a number of different aspects can be taken into consideration. In addition to semantics and pragmatics, aspects like diachrony, register, frequency, iconography, etc. can be regarded (Farø 2004: 92, see also e.g. Korhonen 2007). According to Farø, in a purely lexicological analysis of idiom equivalence, all of these aspects are in principle equivalent, i.e. none of them is more essential (and relevant) than the others. However, from a lexicographical point of view, equivalence of idioms is rather different as the lexicographer, when compiling a dictionary entry, always has to bear the intended user and dictionary use in mind (Farø 2004: 92f., 104f.).

Svensén (2009: 257ff.) and others distinguish three degrees of equivalence: full, partial and zero equivalence, although this classification is rough and the boundaries fluid, essentially making up a scale of more or less conceptual correspondence. Full equivalence implies complete agreement between two expressions in different languages as regard to content and usage. Svensén indicates that full equivalence between expressions in two languages is quite uncommon. Considering all the possible aspects mentioned by Farø (2004) above, it is probably extremely unusual.

However, not even monolingual dictionaries account for all these nuances. Lexicographic descriptions are necessarily idealizations, covering the conventional meanings and typical usage of lexical units but not all aspects of interpretation or the full range of variation. Accordingly, the notion of equivalence in interlingual lexicography is by practical necessity relativized to similar kinds of approximations. Therefore we have mostly used the somewhat weaker term correspondents rather than equivalents.

In the case of constructions, issues of correspondence concern not only content and usage but also form, making full equivalence even less likely. Hence, both the parameters of comparison and the need for approximations include an extra dimension.

Furthermore, there is – at least potentially – a fundamental difference between traditional lexicography and electronic resources regarding the relations between the languages involved. In traditional interlingual lexicography, compiling print dictionaries, the purpose is to provide the selected linguistic items in the source language with equivalents in the target language(s). In this way, the starting point of the work is fixed. In electronic dictionaries (with more or less advanced search functions), the conditions are quite different. Instead of there being just one source language, which serves as the natural point of entry, all languages in the resource can serve both as source languages and target languages.

The development of such resources – both in theory and practice – is still in its infancy, and current e-dictionaries still resemble traditional dictionaries in most respects. Nevertheless, the possibility to establish other than unidirectional links poses new questions about the
treatment of partial correspondence in particular. Instead of just considering whether an expression in the target language is an acceptable translational equivalent to the entry in the source language, one also faces the question whether the relation also holds in the opposite direction, etc. The treatment of such matters in future constructicon development is beyond the scope of the present paper. It is also an issue that has to be approached in collaboration between all the constructicon projects involved.

3.2 Towards interlingual constructicography

All constructicon resources presently under development are designed in relation to a FrameNet (FN) of the language in question, mostly following FN methodology. Hence, an obvious approach is to consider FN methods for cross-linguistic application as well. The core units in FN are frames, “a script-like conceptual structure that describes a particular type of situation, object, or event along with its participants and props.” (Ruppenhofer et al. 2010: 5). Frames are essentially semantic categories, instantiated by various lexical units (LU$s$, pairings of words and meanings). They have been fairly successfully applied across languages, since the same (semantic) frames may be assumed in different languages, and cross-linguistic differences can usually be ascribed to the LUs instantiating them (although not entirely without problems, cf. Boas 2005, Padó 2007, Friberg Heppin & Toporowska Gronostaj 2012).

In a constructicon, the central units are constructions (cxns). They differ from frames in having both form and meaning, a property they share with LUs (from a Construction Grammar point of view, words are simply lexical cxns). Consequently, cxns are also language specific to a higher degree than frames are. Nevertheless, also like LUs, quite a few cxns have a meaning corresponding to frames. For example, the (general) RATE cxn instantiate the Ratio frame, the WAY cxn instantiate the Motion frame, and BE_RECIP instantiate the Reciprocity frame (Fillmore et al. 2012: 325). To the extent that such relations hold, cxns may be related between languages via frames in basically the same way as LUs.

However, some cxns correspond to frames and some do not. It is hard to imagine a frame whose meaning would be instantiated by e.g. GAPPING, ADJECTIVE_AS_NOMINAL, or SAI. What these cxns have in common is that their meaning is grammatical rather than referential; put differently, they are better characterized in terms of function than in terms of meaning. Hence, what is needed to account for them is some kind of abstract grammatical representation. That, on the other hand, would be less suited to handle the more framelike cxns. The most promising approach therefore seems to be a combination of frames and a grammatical representation, where some cxns are linked via frames and others based on grammatical properties.

How such a grammatical metalanguage should be designed depends on the purpose of the resource. Like FN, the constructicon resources are primarily designed for use by experts, in linguistics and/or language technology, not by the general public. Hence, the grammatical representations may and should be both technical and formal. Still, relative simplicity is a high priority. Because of the aim for large-scale coverage in a constructicon, a full-fledged Construction Grammar formalism would simply be too time consuming. The basic format of SweCcn presently consists of (a) a free-text definition, similar to FN- and dictionary-type definitions, combined with (b) a simple structure sketch. Any representation more complex than that would be undesirable. In any case, whatever method is chosen should be compatible with all the constructicon resources concerned, which requires collaboration between the various project groups.

While each constructicon is related to FN, and shared methodology between the two types of resources is therefore desirable, they are also different kinds of resources, with somewhat different properties, uses, and needs. FN is a lexicographic resource, using frames to account
for words, their meanings, and their usage affordances. A constructicon, however, deals with constructions, which makes it a product of constructicography. One crucial difference, except partly different objects of study, is that a constructicon is less dependent on lexical entries, another that it has to account for constituent structure and grammatical relations.

In light of the last point, interlingual constructicon applications should be able to handle structural differences such as the ones demonstrated in this article. Judging from the comparison between BCcn and SweCcn presented here, our working hypothesis is that combining a free-text definition and a structure sketch will provide a rich enough format to account for such distinguishing features, and a simple enough format to do so in a relatively convenient way.

To conclude: Constructicography combines principles of Construction Grammar and lexicography, especially FrameNet lexicography, while also differing from both. Arguably, it is a natural development of the former and the logical complement to the latter (cf. Fillmore 2008, Fillmore et al. 2012). Constructicons are, however, a new kind of resource, and the methodological groundwork that was pioneered in Berkeley and is now continued by all the ongoing constructicon projects has so far mainly focused on monolingual uses. We hope that the present study can serve as a useful point of departure for interlingual constructicon development.

* The work presented here was inspired by Chuck Fillmore, who recently passed away, and we gratefully dedicate this article to his memory. We also wish to thank Hans Boas and two anonymous reviewers for valuable comments to an earlier version of this paper. The Swedish Constructicon project is supported by the Bank of Sweden Tercentenary Foundation (grant agreement P12-0076:1).

1 To what extent these criteria are applicable when more constructions and more languages are taken into account remains to be tested.

2 For two of the incomplete BCcn entries, the information available was insufficient for us to be able to make a reliable comparison. Of the remaining 71, we established Swedish correspondents for 59 of the English entries. 59 out of 71 makes 83%.

3 All italics in the examples have been added by the authors.

4 Note that the contrast in definiteness only concerns the morphosyntax. In terms of information structure, the English and Swedish NPs are no different, both having anchored / accessible referents.

5 The -a suffix on Swedish adjectives is used as a plural marker as well as a definiteness marker. Hence, it may be perceived as marking either or both of these properties in (7), as indicated by the slash in the glosses.

6 Formally, *eller hur* resembles the English tag construction *or what*; functionally, however, the latter expression is much more restricted.

7 Note that the alternation between *i* and *om* in this construction has no systematic relation to the distribution of English *in/on* in expressions such as *in the following hour* and *on this day*.

8 There are also additional formal variants to consider. Swedish has a special type without a relational multiplier, as in (i); whereas English has a type with the preposition *to* and a definite denominator, as in (ii).

(i) en halv liter milen
    a half litre mile-DEF
    ‘half a litre per Swedish mile’

(ii) 764 miles to the gallon

9 There are a few Swedish cxns expressing imperfectivity, for instance pseudo-coordinations and the imperfective auxiliary *hålla på*, but neither of them is used as generally as the English present participle.

10 Notice the contrast between English SVO and Swedish V2 in (25), where the English main clause in (a) displays SV word order (*he would*) following the initial conditional clause but the Swedish translation in (b), adhering to V2, has VS (*hade han*).
For constructicon application to e.g. language pedagogy, other adaptations will have to be made, though not necessarily in the general database format.

The cnx entries in SweCcn also contain annotated examples, links to frames where applicable, particular information about the construction elements etc. Such features, however, are not part of the basic description format but rather provide illustrations and additional information.

References


