Translating imperfective Scandinavian light verb constructions

Karen Brøcker
Uppsala University
Aarhus University
20082334@post.au.dk

Abstract

This paper deals with the Swedish and Danish versions of a specific light verb construction consisting of a posture verb, a coordinating conjunction, and another verb. It is shown how this specific multi word expression is a challenge to Natural Language Processing applications like Google Translate, and a detailed linguistic description of the construction is given. It is concluded that the construction, which is formally similar to the regular coordinating construction, can be differentiated from this on a number of syntactic criteria. The semantics of the construction is characterised as the situation described by the second verb of the construction in the imperfective aspect.

1 Introduction

Multi word expressions (MWEs) are known to challenge various NLP applications such as machine translation and syntactic parsing. This study focuses on one specific MWE, the posture verb and verb construction (PV and V construction for short), which is found in the three Scandinavian languages, Norwegian Swedish and Danish, as well as in a number of other languages (see e.g. Kuteva 1999). The construction is made up of one of the posture verbs corresponding to ‘sit’, ‘lie’, or ‘stand’, a conjunction meaning ‘and’, and another verb. It seems to be generally agreed that the semantics of the construction involves imperfective aspect (Kuteva 1999; Hilpert and Koops 2008; Bjerre and Bjerre 2007a among others). An example of the construction in Swedish can be seen below in (1) (constructed example):

(1) Jag står och fiskar. (S)
I stand and fish.
‘I am fishing.’

That the aspectual implications of the construction can pose difficulties for a NLP application like machine translation can be witnessed by even a superficial analysis of the output of a Google Translate translation of examples like (1) above. The Swedish sentence Jag står och fiskar is translated literally into the English sentence ‘I stand and fish’, and while some alternative translations are given, the one that comes closest to conveying the aspectual semantics of the Swedish original is clumsy and grammatically incorrect: ‘I am standing and fish’.

Baldwin and Kim (2010, p. 268) cite a number of studies that conclude that the “number of MWEs is estimated to be of the same order of magnitude as the number of simplex words in a speaker’s lexicon”, which means that on the type level at least, MWEs are approximately as frequent as simplex words. In addition, new MWEs emerge regularly. This makes modelling MWEs an interesting and important task for language technology applications. The problem of MWEs for language technology can be divided into the tasks of identifying and extracting MWEs from corpus data, and interpreting the semantics of MWEs.

The aim of the present study is to provide an in-depth linguistic description of the construction that can shed light on both its formal properties, for use in identification and extraction tasks, as well as it semantic properties, for use in interpretation tasks. This study only considers the construction in Swedish and Danish (for the Norwegian construc-
tion see Lødrup 2002, for Bulgarian and a small selection of other languages see Kuteva 1999). The analysis is divided into a part on form and one on semantics (Section 5.1 and 5.2). Before the main analysis, some notes on MWEs in relation to language technology are given, and related research on the PV and V construction is summarised.

The PV and V construction has been described from different points of view as either auxiliation (Kuteva 1999), pseudocoordination (Bjerre and Bjerre 2007a; Bjerre and Bjerre 2007b; Hilpert and Koops 2008; Lødrup 2002), or more generally as a subtype of complex predication (Hilpert and Koops 2008). In this paper, I will characterise the construction as a light verb construction (LVC) which is a type of complex predicate. Butt (2010) note that light verbs do not fully predicate, and their contribution to the joint predication is not necessarily transparent, but neither are they completely devoid of semantic content. Together with a complement, light verbs form a joint predication within a complex predicate. In English LVCs, the complements are nouns (e.g. take a nap) but complements with other parts-of-speech exist as well (Butt 2010). In English, a LCV can often be paraphrased with just the verbal form of the noun complement (e.g. walk for take a walk, Baldwin and Kim 2010, p. 281). As we will see in Section 5.2.1, the PV and V construction can similarly be paraphrased with its verbal complement.

In Section 5, I will argue that the PV and V construction is a monopredicative complex predicate, and I will analyse the semantically light contribution of the posture verb to the joint predication of the PV and V construction. On this basis, I analyse the PV and V construction as a LVC. Butt 2010 claims that light verbs can not contribute aspectual meaning to a LVC. She does not give any detailed arguments, however, and as the posture verbs in the PV and V construction in all other ways fit her description of light verbs, I maintain this analysis. Although there is a fair amount of research on how to deal with MWEs for the purposes of language technology, this specific LVC has (to my knowledge) not been investigated in this perspective. And while constructions of this kind are usually highly language specific, this construction exists in the three Scandinavian languages and has what appears to be closely related counterparts in a number of other languages. This makes a description of this construction interesting as it builds a platform for descriptions of the related constructions in other languages to build on.

2 Multi word expressions

The construction under investigation consists of multiple individual words yet seem to have one, unified semantic interpretation. It thus belongs to the group of expressions called multiword expressions (MWEs) which are characterised precisely by a) being decomposable into multiple lexemes, and b) having a component of idiomaticity, giving them lexical status in the mental lexicon (Baldwin and Kim 2010, pp. 267-269). Examples of types of MWEs include noun-noun compounds (e.g. home rule), verb-particle constructions (e.g. take off) and complex prepositions (e.g. in addition to). Light-verb constructions (LVCs) like the PV and V construction is another type of MWE. The PV and V construction clearly consists of multiple lexical items. Its meaning is semantically idiomatic as the imperfective aspect (which is part of the meaning of the construction as will be argued in later sections) is not derivable from the meaning of any of the individual parts of the construction.

MWEs occur with all parts of speech and in many different syntactic constructions. Some have a fixed form (e.g. by and large), other allow for some modification (e.g. in her/his/their shoes) while others are even more syntactically flexible (e.g. hand in the paper vs. hand the paper in) (Baldwin and Kim 2010). As is evident from the examples of semi-flexible and flexible MWEs, these are not necessarily contiguous. The semantics of the MWE might be partially derivable from the semantics of the component parts (e.g. tennis shoe), or there might be no obvious connection (e.g. kick the bucket) (Baldwin and Kim 2010).

In identification and extraction tasks, the main problems that MWEs pose for NLP are a) differentiating between MWE usage and literal usage, (e.g. distinguishing between She’s dead, she kicked the bucket vs. She kicked the old bucket lying in the grass), and b) syntactic disambiguation (e.g. between verb-particle constructions and constructions...
with an intransitive verb and a preposition). Approaches to solving these problems have tended to be language and MWE type specific (Baldwin and Kim 2010, p. 286).

One method for differentiating MWE usage from literal usage is to assume that the MWE will have less syntactic variability than the literal usages (Baldwin and Kim 2010, p. 288). If we have some specification of the syntactic variability of the MWE, we can use that to filter out any occurrences that are examples of literal usage. This specification of syntactic variability can be performed manually for the MWE in question, or it can be done using unsupervised machine learning to predict what the standard forms of the MWE in question look like (Baldwin and Kim 2010, p. 288). In this paper, I give such a specification for the \textit{PV and V} construction done on the basis of manual analysis of examples of the construction.

In semantic interpretation tasks, it is often assumed that the semantics of a specific MWE can be arrived at by synthesising the individual semantics of the component words of the MWE (Baldwin and Kim 2010, p. 293). According to Baldwin and Kim (2010, p. 292), the interpretation of MWEs are mostly done in one of two ways: either as paraphrasing into some “generalised semantic inventory (compatible with both simplex words and MWEs, such as WordNet)” or based on “a set of semantic relations capturing semantic interplay between component words” (see e.g. Megerdoomian 2004). I will return to the compositionality assumption in Section 5.2 and discuss its applicability to the \textit{PV and V} construction.

3 Related work

The \textit{PV and V} construction has been treated from a number of different perspectives, a small selection of which will be presented here. Kuteva (1999) describes the construction as a case of auxiliation. She describes the construction as a way to express continuative, durative or progressive aspect in Bulgarian, the Scandinavian languages, and a number of non-indoeuropean languages. In her view, this use of the construction may have grammaticalised from an earlier, more literal use.

In their article on complex predicates, Hilpert and Koops (2008) present a diachronic perspective on the version of the Swedish construction with the posture verb \textit{sitta} (‘to sit’). In their analysis, the construction is an example of complex predication, more specifically a LVC. In the paper, they argue for a monoclausal interpretation of the construction calling it a “monoclausal construction with two verbal heads” (Hilpert and Koops 2008, p. 244). They characterise the semantic function of the \textit{PV and V} construction as marking the event described as having duration or being in progress. They also note that the literal bodily meaning of ‘sitting’ is not necessarily a prominent part of the resulting constructions semantics, although the postural semantics may be emphasised in some contexts.

In contrast to this analysis is Lødrup (2002) who argues for a biclausal interpretation of the corresponding (and very similar) Norwegian construction. His analysis is done within the framework of Lexical-Functional Grammar, and he treats a wider range of constructions than that considered in this paper.

Bjerre and Bjerre (2007a); Bjerre and Bjerre (2007b) describe the syntactic properties of the Danish version of the \textit{PV and V} construction with the posture verb \textit{sidde} (‘to sit’). They analyse the construction as a hybrid phrase expressing both coordinating and subordination. According to them, the semantic function of the construction is to express imperfective aspect.

Kahr (2013) analyses the Danish version of the construction with the posture verb \textit{står} (‘to stand’) from a Semiotic Grammar perspective. She gives a detailed description of the form and semantics of the construction, arguing how they form the signifier and signified of a grammatical sign. She characterises the semantic function of the construction as marking the aspect of the content verb as imperfective.

While the papers presented above differ in their focuses, theoretical backgrounds and to some extent in their conclusions, they seem to agree on a superficial characterisation of the construction in question. There is some disagreement as to the degree of semantic bleaching of the literal sense of the posture verb, and it is possible that the different posture verbs in different languages have retained their literal semantic meaning to different degrees. The
different papers do seem to agree, however, that the meaning of the construction includes aspectual meaning of imperfective type. The present project will aim to refine the description of both the form and semantics of the construction in Swedish and Danish.

Finally, a note on two seemingly related constructions in Danish and Swedish. In Danish you find a very similar construction with motion verbs like gå (‘to walk’), løbe (‘to run’) and køre (‘to drive’). The particular version with ‘walk’ has been described in Brøcker (2013), where its semantic function is characterised as expressing different subtypes of imperfective aspect according to the aktivsart of the content verb. Interestingly, there is a Swedish construction which formally resembles the Danish motion verb construction, yet its semantic function seems to be to indicate that actor is about to begin the action denoted by the content verb (Wiklund 2009).

4 On data and methodology

In this study, I present my analysis of the PV and V construction. In forming hypotheses about the construction, I have been guided by my intuition as a native speaker of Danish and my knowledge of Swedish as a non-native speaker. These initial hypotheses were then checked against corpus data and refined to match the new insights from the data, then checked and refined again iteratively until the results were constant across a number of iterations.

The corpora used for this study are the Danish language corpus KorpusDK and a subset of the Swedish language corpus Språkbanken. KorpusDK contains 56 million words (at the token level) from Danish texts collected between 1998 and 2002 from news media, magazines, prose, technical writings, and web pages. Språkbanken contains a much wider range of material than KorpusDK, so to ensure a comparable result, a subset of Språkbanken was chosen to try and match the types of text and time period of KorpusDK. The subset of Språkbanken selected for this study is the complete Bloggmix subcorpus, the subcorpus containing Norstedtsromaner, the GP, Webnyheter and 8 sidor subcorpora, the part of the Press subcorpora from 1995 to 1998, and the Tidskrifter subcorpus. In total, this collected subcorpus contains slightly less than 100 million tokens. As the number of examples in the two languages are not directly compared, the difference in size between the two is not problematic for the study.

Examples of the PV and V construction were extracted from the corpora by specifying the pattern of the construction: one of the three posture verbs, sitta/ligga/stå in Swedish and sidde/ligge/stå in Danish, the word och/og, and a verb. In both languages, there exist fixed expressions on this form, like står och/og faller/falder (S/D) and står och stampar (S). These were disregarded in the analysis since they have idiomatic semantics.

Unless otherwise indicated, all examples presented in the study are from one of these corpora. In a few cases, examples have been altered to make a point about the range of variety of the construction, and a few examples are constructed by me (as cited). For the sake of readability, the corpus examples are sometimes simplified with respect to issues that are not relevant to the form or meaning of the PV and V construction. Each example has been provided with a simple word-by-word gloss and a translation into English (translations done by me). Morpheme level glossing has been deemed unnecessary for the present purpose as the PV and V construction is morphologically simple and constant. Ungrammatical examples are glossed with an asterisk *, and examples that are either grammatically questionable or possibly examples of something other than the PV and V construction have been glossed with a question mark ?. All examples are marked with either S or D to indicate the language of the example.

5 The construction

In a small scale pilot study, I manually and semi-randomly picked five examples for each of the posture verbs in each of the two languages, one example with each of the five content verbs that most frequently follow each posture verbs in the two corpora used. I then translated the examples into English using Google Translate. Each of the thirty translations was then analysed with respect to how the PV and V construction part of the sentences was translated (other errors were not considered). The examples can be found in Appendix A.

The analysis revealed two major types of error: translations including a posture verb where none
was necessary, and translations where the verb was not marked correctly for aspect. Furthermore, the two types of error often combine (as can be seen from (3) below; the part of the translation corresponding to the PV and V construction is underlined).

(2) Det är ganska många människor som står och skriker.  
‘There are quite a few people standing and screaming.’

(3) De sidder og venter ved grillen derhjemme.  
‘They sit and wait by the grill at home.’

(4) Rutschkanan bara ligger och skräpar på marken.  
‘The slide just lying around on the ground.’

Of the 30 examples, 19 included a literal translation of the posture verb, and for 17 of these, this was classified as an error. For the remaining two examples, the posture verb is part of the translation of the content verb as in (4) where ligger och skräpar is translated as ‘lying around’. In 9 of these 17 cases, the content verb was marked for the right aspect as in (2). In the other 8 of the 17 cases, the content verb was not marked for the right aspect as in (3).

In 8 of the 9 cases that included the posture verb and were marked correctly for aspect, the correct translation includes the English progressive form (ending in -ing), yet in one case, the correct translation included the simple present form of the verb. I will return to this point in Section 5.2.1.

This simple test shows that the performance of Google Translate for this specific MWE is varied. In cases of error, the problem is including a literal translation for the posture verb and/or not having the content verb in the right aspect. These problems could stem from two sources: 1) not being able to recognise a particular occurrence of the PV and V construction as such, i.e. interpreting it as a normal coordinating construction, and 2) not knowing the correct semantic interpretation of the construction. In this section, I first give a description of the form and syntactic behaviour of the construction, and then I give an analysis of the semantic interpretation and the role of the posture verb.

5.1 Form
As already mentioned, the PV and V construction consists of a posture verb, the coordinating conjunction, and a content verb. The content verb can be a simple verb or a verbal construction such as the copula er (‘is’) and its predicate or har (‘have’) and its argument as in (5).

(5) Han sat just and was so happy.  
‘He was just so happy.’

At first glance, the PV and V construction looks almost identical to the regular coordinating construction. There are, however, some syntactic differences between the two, so in the following, I will compare the PV and V construction to the regular coordinating construction to specify the specific syntactic variability of the PV and V construction.

Like for the regular coordinating construction, the posture verb and the content verb always have the same value for tense and finiteness, and the construction appears in all tenses (Present, Perfect, Preterit, Pluperfect, and Future tense).

Contrary to the coordinating construction, the PV and V construction has precisely one overt subject that functions as the subject of both verbs as seen in (6). If there is more than one overt subject, as in (7), the meaning changes and it is no longer an example of the PV and V construction. In coordinating constructions, we often find ellipsis of the second subject, as in (8) and (9), but in contrast to what is the case for the PV and V construction, the variants with one subject and the one with two subjects are semantically interchangeable ((7) and (9) are constructed examples).

(6) Hun står og fryser.  
She stands and freezes.  
‘She is freezing.’

(7) Hun står og hun fryser.  
She stands and she freezes.  
‘She is standing and she is freezing.’

(8) Rektorn lyckades fly och var sedan tvingen att låsa in sig på sitt  
The principal succeeded escape and was then forced to lock in himself in his
‘The principal was able to escape and was then forced to lock himself into his room.’

Another characteristic of the PV and V construction is, as Kahr (2013) notes about the står version of the construction in Danish, that only one negation is required to negate the full construction. This is another thing that sets the PV and V construction apart from the normal coordinating construction. The negation is placed after the posture verb and has scope over both verbs as can be seen with the negation ikke (D) in (10). If the negation is placed after the content verb, the result is slightly unnatural and the negation then only has scope over the content verb. In coordinating constructions, two negations can be used as in (11) while this is not possible in the PV and V construction. It is also possible for one negation to have scope over both verbs in a coordinating construction, but in that case the negation is placed after the second verb in the construction, as can be seen with the negation inte (S) in (12).

(10) Jeg sidder ikke og griner. (D)  
I sit not and laugh.  
‘I am not laughing’

(11) Læg mærke til, hvordan liljerne gro;  
Place notice to how the lilies grow;  
de arbejder ikke og spinder ikke. (D)  
they work not and spin not.  
‘Consider the lilies, how they grow: they neither toil nor spin.’

(12) Han klarade och vågade inte vänta  
He managed and dared not wait  
längre. (S)  
longer.  
‘He neither managed nor dared to wait any longer.’

In both the PV and V construction and the coordinating construction, time, place, and manner adjuncts can occur both after the posture verb or after the content verb, and if just one adjunct is used, the adjunct has scope over both verbs. For the coordinating construction, two different adjuncts can be used, and then the reading entails two situations as in (13) (constructed example). This is not possible for the PV and V construction. In PV and V constructions, manner adjuncts mostly follow the content verb, and when they follow the posture verb, the posture verb gets a more literal interpretation as in (14).

(13) ? En restaurang där man sitter inne  
A restaurant where you sit inside  
och äter i en lum mig trädgård. (S)  
and eat in a shady garden.  
‘A restaurant where you sit inside and eat in a shady garden.’

(14) Värdparet sitter lugnt och tittar på  
The host couple sit quietly and look at  
spektaklet. (S)  
the scene.  
‘The hosts sit quietly and watch the scene.’

That one negation or adjunct has scope over both verbs in PV and V construction suggests that, in contrast to the normal coordinating construction, the posture verb construction is monopredicative. As can be seen from (15), the two verbs in a regular coordinating construction refer to two distinct situations, here fishing and harvesting:

(15) De var så ugodelige, at de fiskede  
They were so ungodly that they fished  
og høstede selv i kirketiden. (D)  
and harvested even during church hours.  
‘They were so ungodly that they would fish and harvest even during church hours.’

Kahr (2013) mentions two other syntactic characteristics that set the PV and V construction apart from the regular coordinating construction. The first one is that the word order in questions containing the PV and V construction is different from that for questions containing the coordinating construction. As can be seen from (16) and (17), the subject follows the posture verb in the PV and V construction.
whereas the subject follows the second verb in a normal coordinating construction. The second is that for the current construction, topicalisation is possible when the content verbs are transitive whereas this is not the case for the coordinating construction ((17), (18), and (19) are constructed examples).

(16) *Sidder De og belurer Deres
Sit you and spy on your svigerfar? (D)
father-in-law?
‘Are you spying on your father-in-law?’

(17) Synger og danser de allesammen? (D)
Sing and dance they all together?
‘Do they all sing and dance together?’

(18) Bogen han sidder og læser...
The book he sits and reads...
‘The book that he is reading...’

(19) *Bogen han skriver og læser...
The book he writes and reads...

To sum up, the construction has the form posture verb + och/og + content verb. The posture verb and content verb have the same value for tense and finiteness, and they share one subject that functions as subject for both verbs. When modified, one negation or adjunct has scope over both verbs. In the case of manner adjuncts, the construction gets a literal reading when this is placed immediately after the posture verb. In contrast to coordinating construction, the subject follows the posture verb directly in interrogatives, and topicalisation is possible with transitive content verbs.

This means that in certain situations, purely syntactic criteria can be used to tell the two types of constructions apart. If a candidate construction has two subjects, negations or adjuncts, it is a coordinating construction. If a question contains a candidate construction which has the subject after the first verb, the construction is a PV and V construction. Finally, if a candidate construction is topicalised, it is a PV and V construction.

5.1.1 Content verbs

In both Swedish and Danish, the majority of the verbs that follow the group of posture verbs are atelic in their natural aspect (aktionsart), and there tends to be more examples of content verbs that describe situations that are durative in nature compared to punctual situations (I here follow Comrie 1976 in using the term ‘situation’ to refer to both events, processes, and states). Verbs with telic aktionsart and verbs describing punctual situations do occur as well, however. In Table 1 and 2, the five most frequent content verbs following posture verbs in both the languages are presented.

<table>
<thead>
<tr>
<th>Swedish</th>
<th>Danish</th>
<th>Norwegian</th>
</tr>
</thead>
<tbody>
<tr>
<td>står och...</td>
<td>sitter och...</td>
<td>ligger och...</td>
</tr>
<tr>
<td>väntar</td>
<td>väntar</td>
<td>sover</td>
</tr>
<tr>
<td>‘wait’</td>
<td>‘wait’</td>
<td>‘sleep’</td>
</tr>
<tr>
<td>tittar</td>
<td>kollar</td>
<td>väntar</td>
</tr>
<tr>
<td>‘look’</td>
<td>‘check’</td>
<td>‘wait’</td>
</tr>
<tr>
<td>pratar</td>
<td>tittar</td>
<td>skräpar</td>
</tr>
<tr>
<td>‘talk’</td>
<td>‘look’</td>
<td>‘clutter’</td>
</tr>
<tr>
<td>skriker</td>
<td>åter</td>
<td>kollar</td>
</tr>
<tr>
<td>‘cry/shout’</td>
<td>‘eat’</td>
<td>‘check’</td>
</tr>
<tr>
<td>lagar</td>
<td>pratar</td>
<td>vilar</td>
</tr>
<tr>
<td>‘make’</td>
<td>‘talk’</td>
<td>‘rest’</td>
</tr>
</tbody>
</table>

Table 1: The five most frequent content verbs after posture verbs in the used subset of Språkbanken (Swedish)

<table>
<thead>
<tr>
<th>Swedish</th>
<th>Danish</th>
<th>Norwegian</th>
</tr>
</thead>
<tbody>
<tr>
<td>står og...</td>
<td>sidder og...</td>
<td>ligger og...</td>
</tr>
<tr>
<td>mangler</td>
<td>venter</td>
<td>venter</td>
</tr>
<tr>
<td>‘need’</td>
<td>‘wait’</td>
<td>‘wait’</td>
</tr>
<tr>
<td>venter</td>
<td>snakker</td>
<td>flyder</td>
</tr>
<tr>
<td>‘wait’</td>
<td>‘talk’</td>
<td>‘float’</td>
</tr>
<tr>
<td>skal</td>
<td>skriver</td>
<td>sover</td>
</tr>
<tr>
<td>‘going to do’</td>
<td>‘write’</td>
<td>‘sleep’</td>
</tr>
<tr>
<td>kigger</td>
<td>kigger</td>
<td>lurer</td>
</tr>
<tr>
<td>‘look’</td>
<td>‘look’</td>
<td>‘lurk’</td>
</tr>
<tr>
<td>siger</td>
<td>tænker</td>
<td>roder</td>
</tr>
<tr>
<td>‘say’</td>
<td>‘think’</td>
<td>‘clutter’</td>
</tr>
</tbody>
</table>

Table 2: The five most frequent content verbs after posture verbs in KorpusDK (Danish)

The construction is productive and appears with a wide range of content verbs. However, there seems to be a semantic constraint on which verbs appear as content verbs: the situation that the content verb describes has to be compatible with the posture that the posture verb describes. Some motion verbs appear as content verbs, as can be seen in (20). In the case of ‘standing and jumping’, this action is compatible with a person being in the upright position...
and staying in the same place.

(20) Præsten stod og hoppede. (D)
    The priest stood and jumped.
    ‘The priest was jumping (up and down).’

5.2 Semantics

As we saw earlier in this section, the two major problems with translating the $PV$ and $V$ construction is getting the aspectual marking right and deciding whether or not to leave out the posture verb. These two issues are addressed in turn below.

5.2.1 Aspect

As described in previous studies, the construction generally contributes an imperfective reading to the situation described by the content verb (Kuteva 1999; Hilpert and Koops 2008; Bjerre and Bjerre 2007a; Kahr 2013). We can see how the focus in (21) (constructed example) is on the situation as a closed unit viewed from the outside. In (22), on the other hand, the focus is on the internal temporal structure of the situation.

(21) Dom avtalar en gemensam frukost. (S)
    They plan a common breakfast.
    ‘They plan a common breakfast.’

(22) Dom står och avtalar en gemensam
    They stand and plan a common
    frukost. (S)
    breakfast.
    ‘They are planning a common breakfast.’

Comrie (1976) and others make a distinction within the category of imperfective aspect between habitual and continuous aspect. In languages that have a special habitual form, this form is used to mark situations that can be said to be characteristic for a longer period of time. In English, habitual can be expressed in the past tense with ‘used to’ as in ‘John used to work here’ (Comrie 1976, p. 25).

Comrie (1976, p. 26) defines continuous aspect as “imperfectivity that is not habituality”. Progressive is a subcategory of continuous that is used to describe situations that are in progress at the time in question. In English, the progressive (formed with the coupla to be and the verbal suffix -ing) can not be applied to verbs describing states (He knows the answer but *He’s knowing the answer).

The specific type of imperfective aspect that the construction contributes to the joint predication of the construction seems to be dependent on the nature of the content verb. In some cases, the construction gets a progressive reading (as in (23)), but in a small number of cases it gets a habitual reading (24). In all the cases I have analysed where the construction gets a habitual rather than a progressive reading, the construction is accompanied by a time reference that marks the situation in question as something that occurs over a longer period of time, like på dagarna (‘by day’) in (24). The habitual aspect is therefore possibly a result of the time specification combined with the imperfective aspect of the construction. A systematic analysis of the aspectual meaning of different temporal adjuncts would be relevant here but is outside the scope of this paper.

(23) Han står och flöjter, mens han
    He stands and whistles while he
    barberer sig. (D)
    shaves himself.
    ‘He whistles while shaving himself.’

(24) På dagarna är hamstern tråkig och
    On the_days is the_hamser boring and
    ligger och sover. (S)
    lies and sleeps.
    ‘By day, the hamster is boring and does nothing but sleep.’

Apart from the progressive meaning, the construction additionally gets an iterative interpretation when the content verb describes a punctual situation (25). With state verbs (as in (26)), the construction gets a continuous reading, but as state verbs in English can not take the progressive form, the translation into English makes use of the simple present form.

(25) Jim sitter och nyser i soffan. (S)
    Jim sits and sneezes in the_couch.
    ‘Jim is on the couch, sneezing repeatedly.’

(26) Den reservedel, han står og
    The spare_part he stands and
    mangler.(D)
    needs.
    ‘The spare part that he needs.’
When translating the PV and V construction into English, the following guidelines regarding aspect thus applies: if the content verb is a state verb or if there is a time specification in the immediate context that suggests habituality, the verb in the translation should be in the simple tense, otherwise the verb should be in the progressive tense.

5.2.2 Posture

In her treatment of constructions like the PV and V construction in a number of other languages, Kuteva (1999, p. 194) notes that in Bulgarian, “it is freely used in contexts where the agent does not have to necessarily ‘lie’ or ‘sit’ while performing the action denoted by the main verb” although the construction is often used “according to the bodily position of the subject” (Kuteva 1999, p. 193). On this basis, Kuteva (1999, p. 208) suggests that the specific “‘human body’ semantics” of the construction has been lost.

This does not appear to be equally true for the corresponding construction in Swedish and Danish. As described in Section 5.1.1, the range of possible content words seem to be restricted to those that are conceptually compatible with the posture verb of the construction. In some cases, however, the posture verb can only be interpreted figuratively, as in (27) where the bodily interpretation of ‘sitting’ has little relevance for the subject Turkey.

(27) Tyrkiet sidder og venter på at kunne
    Turkey sits and waits on to be able to
    spille en rolle igen i Mellemøsten.(D)
    play a part again in the Middle East.
    ‘Turkey is waiting to be able to play a part
    in the Middle East again.’

In some cases, like in (27) above, the semantics of the full construction seems to be coloured by the connotations of the posture verbs like being stationary and not being active. In her article on aspect in Swedish, Greek and Polish, Lindvall (1997, p. 209) notes that posture verbs express temporary states and have a low degree of dynamism and that they are used in PV and V constructions to express extended duration of the content verb.

This suggests that the posture verb is at least in some cases semantically bleached. This is perhaps not surprising as all three posture verbs in both Swedish and Danish are polysemous, their meanings ranging from the bodily posture to more abstract meanings like being/existing (Svensk Ordbok, Den Danske Ordbog).

It is, however, not the case that the physical, bodily meaning of the posture verbs has disappeared completely. In the majority of the examples analysed in this study, the posture verb corresponds to the most common bodily posture of a agent in the situation described by the content verb. If we substitute one posture verb for another, we get results that are pragmatically strange, as in (28) which is a version of (25) where ‘sit’ has been exchanged for ‘stand’.

(28) Jim står och nyser i soffan. (S)
    Jim stands and sneezes in the couch.
    ‘Jim is standing on the couch, sneezing.’

There are, however, some exceptions, as in (29), where ‘lying’ and ‘driving’ are used together, although the actors involved in the driving are more likely to be sitting up than lying down. In this case, it seems that the car is perceived as an extension of the body, and the car can be said to ‘lie’ horizontally on the road.

(29) Många föräldrar ligger och kör fram och
    Many parents lie and drive forth and
    tillbaka... (S)
    back...
    ‘Many parents are driving back and forth...’

As the posture verb almost always corresponds to the canonical posture of the agent engaged in the situation described by the content verb, most of the examples do not change their meaning if removing the posture verb and coordinating conjunction (disregarding the aspectual semantics for the moment), as seen in (30) and (31) (constructed example).

(30) Jag sitter och skriver rekommendationer.(S)
    I sit and write recommendations.
    ‘I am writing recommendations.’

(31) Jag skriver rekommendationer. (S)
    I write recommendations.
    ‘I write recommendations.’

Another indication that the posture verbs have become semantically light is that they are hardly ever
relevant when translating the construction into English. As can be seen in (30) and (31) above, the only difference between the translations for the two examples is the aspectual marking on the content verb. As we saw in the case of (4), one of the the examples translated with Google Translate, the translation of the PV and V construction can sometimes contain the posture verb. This seems only to be the case, however, when it is relevant as part of the translation of the content verb.

Summing up, the posture verb can be ignored when translating the PV and V construction from Swedish or Danish into English, and the content verb should be marked as imperfective as described above. In this paper, I have not dealt with translating from English into Swedish or Danish. However, the analysis above suggests that when translating expressions in the progressive aspect from English into Swedish or Danish, the PV and V construction might often offer the closest translation, and in that case it would be important to select the posture verb that naturally corresponds to the unmarked or canonical posture of the actor engaged in the situation described by the content verb.

Finally, a note on compositional semantics. As described in Section 2, a basic assumption when in most approaches to automatic semantic interpretation of MWEs is that the semantics of the full MWE can be arrived at by somehow synthesising the semantics of the individual parts of the MWE. In the case of the PV and V construction, I propose that the semantics of this particular MWE is a composition of the semantics of the content verb and grammatical, more specifically aspectual, meaning. In this case, the compositionality view can be maintained if one assumes that the posture verb contributes not its literal semantics but instead imperfective aspectual semantics. Whether or not this means that the PV and V construction is undergoing grammaticalisation is a discussion that I is outside the scope of this investigation.

6 Conclusion

In this paper, I have given a detailed linguistic description of the PV and V construction for Swedish and Danish. Syntactically, it differs on a number of points from the otherwise very similar regular coordinating construction. Thus, syntactic information will in some cases be useful when deciding whether a specific expression is an example of the PV and V construction or not and thus might be useful for identification or extraction purposes. Semantically, the construction renders an imperfective reading to the situation referred to by the content verb. Generally, the PV and V construction translates into the English progressive form of the content verb, yet some more detailed considerations which might be useful for interpretation purposes are outlined as well.

References


**List of resources**

*Svensk Ordbok*. Last checked: 04.12.13. URL: [https://www.ne.se/om/svensk-ordbok](https://www.ne.se/om/svensk-ordbok).

**Appendix A. Collection of Google Translate Examples**

The examples presented here all come from Språkbanken (Swedish) or KorpusDK (Danish). For each posture verb (‘sit’/’stand’/’lie’) there are five examples in each language, one with each of the five most frequent content verbs for that posture verb. Each example is presented with its unedited Google Translate translation. For each example, the part of the translation corresponding to the *PV and V* construction has been underlined. The examples are grouped according to the quality of the translation: correct translations, translations that include the posture though it is irrelevant but have the right aspectual marking, and translations that include the posture though it is irrelevant and have the wrong aspectual marking.

**Correct translation**

*De sitter och väntar nere i bilen.*
They are waiting down there in the car.

*Jeg sidder og skriver på vej hjem i bilen.*
I am writing on the way home in the car.

*Några är på väg ner till tunnelbanan, andra står och väntar på bussen.*
Some are heading down to the subway, others are waiting for the bus.

*Vi taler lidt om det, de står och mangler.*
We talk a bit about what they are missing.

*De står och venter på mig nede i kælderen.*
They are waiting for me in the basement.

*Hvad fanden er det, du står og siger?*  
What the hell is it you are saying?

*En man ligger och sover på järnvägsspår.*  
A man is sleeping on the train tracks.

*Hun ligger og venter på at blive fundet.*  
She is waiting to be found.

*De har andra ärenden som ligger och väntar på att behandlas.*  
They have other cases that are waiting to be treated.

*Rutschkanan bara ligger och skräpar på marken.*  
The slide just lying around on the ground.

*Papirerne ligger og flyder, hvor de så længe har ligget.*  
The papers lying around where they have so long lain.

*Han ligger og sover ligesom pigen och den sorte killing.*  
He is sleeping like the girl and the black kitten.

*En beruset herre ligger og roder efter noget i rendestenen.*  
A drunk men is fumbling for something in the gutter.
Posture irrelevant + right aspect

*En katt sitter och tittar på en tom vägg.*
A cat sitting and looking at a blank wall.

*Vi sitter och pratar i en skinnsoffa i vardagsrummet.*
We are sitting and talking on a leather couch in the living room.

*Camilla sidder og snakker med en jævnaldrende pige.*
Camilla sitting and talking with peers girl.

*Det är ganska många människor som står och skriker.*
There are quite a few people standing and screaming.

*Ingen går ut, säger en kille som står och tittar på hålet efter raketen.*
Nobody goes out, says a guy standing and looking at the hole left by the rocket.

*En gammel kone står och kigger på en skotte, der spiller sækkepibe.*
An old woman standing and looking at a Scotsman playing the bagpipes.

*Sälar som ligger och vilar på land får rinniga ögon.*
Seals lying and resting on land may runny eyes.

*Vildskaben ligger och lurder under overfladen.*
The wildness lies lurking beneath the surface.

*Jag åbner en dåse öl og sidder og kigger på lysene i vinduet.*
I open a can of beer and sit and look at the lights in the window.

Posture irrelevant + wrong aspect

*Jag sitter och kollar på bilderna från New York.*
I sit and watch the pictures from New York.

*Nu är jag uppe och sitter och åter min gröt.*
Now I’m up there and sit and eat my porridge.

*De sidder og venter ved grillen derhjemme.*
They sit and wait by the grill at home.

*Jeg sidder og tænker på de døde dyr, jeg har set.*
I sit and think of the dead animals I have seen.

*Vi står och pratar en stund i solskenet framfor bostadshuset.*
We stand and talk for a while in the sunshine in front of the apartment building.

*Min far står och lagar något gott i köket.*
My father stands and cook something good in the kitchen.

*De står og skal afsted på ferie.*
They stand and have to go on vacation.

*Jag ligger och kollar på film.*
I lie and watch movies.