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 multiword expression is a challenge to Google Translate, and based on a detailed linguistic description of the construction, guidelines for the identification and interpretation of the construction are given.

1 Introduction

Multiword 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 (in this paper called the content verb following Jørgensen 2001). 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) but a more detailed semantic analysis of the general construction has not, to my knowledge, been done.

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 an 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, 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”. 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 a more detailed linguistic description of the construction than what has previously been done focusing on both its formal properties, for use in identification and extraction tasks, as well as its semantic properties, for use in interpretation tasks. This study only considers the construction in Swedish and Danish (for the Norwegian construction see Lødrup 2002, for Bulgarian and a small selection of other languages see...
Kuteva 1999). The analysis is divided into one part on how to identify the construction and one on how to interpret it (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) notes 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, 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 (i.e. that the two verbs in the construction describe just one situation), 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 an LVC. Butt (2010) claims that light verbs can not contribute aspectual meaning to an 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.

2 Automatic Translation of the PV and V Construction

To test the quality of automatic translations of the PV and V construction, I did a small exploratory study, in which 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 (see Section 5.1.1). I then translated the examples into English using Google Translate (GT). 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 translations given by GT were classified as either acceptable or non-acceptable on an case-by-case basis. The examples and their translations 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 (4) below (the part of the translation corresponding to the PV and V construction is underlined).

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 (2) where ligger och skräpar is translated as ‘lying around’.

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

In 9 of the 17 cases, the content verb was marked for the right aspect as in (3).

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

In the remaining 8 of the 17 cases, the content verb was not marked for the right aspect as in (4).

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

1The GT translation models are continually updated. The translations presented in this paper were done in the first half of December 2013.
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 examples like this 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, a literal translation for the posture verb is included in the translation and/or the content verb is not in the right aspect. These problems could stem from two sources: 1) GT is not able to recognise a particular occurrence of the *PV and V* construction as such, i.e. interprets it as a normal coordinating construction, and 2) GT does not know the correct semantic interpretation of the construction. In Section 5, 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. First, however, I describe the background for the current paper.

3 Background

3.1 Multiword Expressions

The construction under investigation consists of multiple individual words yet seem to have one, unified semantic interpretation. It is thus a MWE, and these 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, 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 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 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 (Baldwin and Kim 2010 call this semantic idiomaticity).

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 that was 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, 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, 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, 288). In this paper, I give such a specification for the *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, 293). According to Baldwin and Kim (2010, 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).

The *PV and V* construction combines with a wide
range of different content verbs (see Section 5.1.1), and it makes little sense to analyse each instance of the construction with each new content verb as a separate MWE. Rather, the PV and V construction consists of a posture verb, a conjunction, and a slot for a verb. The PV and V construction, as we will see in Section 5.2, gives an imperfective reading of the situation referred to by the content verb. Thus the semantics of this particular MWE is a composition of the aspectual meaning of the construction itself and the semantics of the content verb that fills the open slot in the construction. Whether the grammatical nature of the meaning of the PV and V construction means that it is undergoing grammaticalisation is a discussion that is outside the scope of this investigation.

3.2 Related Work
The 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 sitta (‘to sit’). In their analysis, the construction is an example of complex predication, more specifically an LVC. In the paper, they argue that the two verbs in the construction are describing one rather than two situations (as is the case with coordinating constructions), and they call it a “monoclusal construction with two verbal heads” (Hilpert and Koops 2008, 244). They characterise the semantic function of the 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 construction’s 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; 2007b) describe the syntactic properties of the Danish version of the PV and V construction with the posture verb sidde (‘to sit’). Theyanalyse the construction as a hybrid phrase expressing both coordination 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 stå (‘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 the 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 aktionsart 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 the actor is about to begin the
action denoted by the content verb (Wiklund 2009).

4 Data and Methodology

In this study, I present my analysis of the \textit{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 analysis stabilised.

The corpora used for this study are the Danish language corpus KorpusDK and a subset of the Swedish language corpora that are accessible through the website of Språkbanken, a research unit at the University of Gothenburg. 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 the corpora found at Språkbanken was chosen to try and match the types of text and time period of KorpusDK. The subset of corpora from Språkbanken selected for this study is the complete Bloggmix corpus, the corpus containing Norstedtsromaner, the GP, Webnyheter and 8 sidor corpora, the part of the Press corpora from 1995 to 1998, and the Tidskrifter corpus. In total, this collected corpus 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 \textit{PV and V} construction were extracted from the corpora by specifying the pattern of the construction: one of the three posture verbs, \texttt{sitta/ligga/stå} in Swedish and \texttt{sidde/ligge/stå} in Danish, the word \texttt{och/og}, and a verb. In both languages, there exist fixed expressions on this form, like \texttt{står och/og faller/falder} (S/D) and \texttt{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. 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 \textit{PV and V} construction. Each example has been provided with a simple word-by-word gloss and a translation into English (except from the translations in Section 2, all the translations are done by me). Morpheme level glossing has been deemed unnecessary for the present purpose as the \textit{PV and V} construction is morphologically simple and constant. Ungrammatical examples are marked with an asterisk *, and examples that are either grammatically questionable or possibly examples of something other than the \textit{PV and V} construction have been marked with a question mark ?. All examples are marked with either (S) or (D) to indicate the language of the example.

5 The Construction

5.1 Identification

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

\begin{center}
\begin{tabular}{ll}
(5) & \texttt{Han sad bare og var så glad.} (D) \\
& \texttt{He sat just and was so happy.} \\
& \texttt{‘He was just so happy.’}
\end{tabular}
\end{center}

At first glance, the \textit{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 these two constructions to characterise the specific syntactic variability of the \textit{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 tense and with auxiliary future tense marking).

Contrary to the coordinating construction, the \textit{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 ((7) and (9) below are constructed examples).

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

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

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 variant with one subject and the one with two subjects are semantically interchangeable.

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

(9) *Rektorn lyckades fly och han var sedan tvungen att låsa in sig på sitt rum.* (S)
   The principal succeeded escape and he was then forced to lock in himself in his room.
   ‘The principal was able to escape and he 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. 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).

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

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.

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

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).

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

In both the PV and V construction and the coordinating construction, time, place, and manner adjuncts can occur 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.

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

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).

(14) *Värdparet sitter lugnt och tittar på spektaklet.* (S)
    The hosts sit quietly and watch the scene.'
That one negation or adjunct has scope over both verbs in the *PV and V* construction suggests that, in contrast to the normal coordinating construction, the *PV and V* 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 og høstede selv i kirketiden.* (D)

‘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 always follows the posture verb in the *PV and V* construction whereas the subject normally follows the second verb in a normal coordinating construction. It might be possible to have the subject after the posture verb with normal coordination as well, but this would be a marked construction.

(16) *Sидder De og belurer Deres svigerfar?*

Sit you and spy:on your 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?’

The second is that for the *PV and V* 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).

(18) *Bogen sidder han og læser.* (D)

The:book sits he and reads.

‘He is reading the book.’ or ‘It is the book that he is reading.’

(19) *Bogen skriver han og læser.* (D)

The:book writes he and reads.

To sum up, 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

The construction is productive and appears with a wide range of content verbs. In Table 1 and 2, the five most frequent content verbs following posture verbs in both the languages are presented.

<table>
<thead>
<tr>
<th>står och...</th>
<th>sitter och...</th>
<th>ligger och...</th>
</tr>
</thead>
<tbody>
<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>äter</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)

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 (I here follow Comrie 1976 in using the term ‘situation’ to refer to both events, processes, and states). Some motion verbs appear as content verbs, as can
Table 2: The five most frequent content verbs after posture verbs in KorpusDK (Danish)

<table>
<thead>
<tr>
<th>står og…</th>
<th>sidder og…</th>
<th>ligger og…</th>
</tr>
</thead>
<tbody>
<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>tanken</td>
<td>roder</td>
</tr>
<tr>
<td>‘say’</td>
<td>‘think’</td>
<td>‘clutter’</td>
</tr>
</tbody>
</table>

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 Interpretation

As we saw in Section 2, 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 (see Section 3.2). 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* frukost. (S)
They stand and plan a common 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, 25). Comrie (1976, 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.

The examples analysed for this study show that the specific type of imperfective aspect that the construction contributes to the joint predication of the construction varies. In most cases, the construction gets a progressive reading, as in (23), but in a small number of cases it gets a habitual reading (24).

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

(24) *På dagarna är hamstern tråkig och ligger och sover.* (S)
By day, the hamster is boring and does nothing but sleep.

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. This would mean that the meaning of the construction is still progressive but that the superimposed habitual structure dictates the simple present form when the construction is translated into English.

Apart from the imperfective meaning, the construction additionally gets an iterative interpretation when the content verb describes a punctual situation as in (25).
(25)  *Jim sitter och nyser i soffan.*  (S)
     Jim sits and sneezes in the: couch.
     ‘Jim is on the couch, sneezing repeatedly.’

With state verbs (as in (26)), the construction gets
a progressive reading, but as state verbs in English
can not take the progressive form, the translation
into English makes use of the simple present form.

(26)  *Den reservedel, han står og mangler.*  (D)
     The spare: part he stands and 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 con-
text that suggests habituality, the verb in the trans-
lation 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 languages, Kuteva (1999, 194) notes that in Bulgarian, ‘it is freely used in contexts where the agent does not have to neces-
sarily ‘lie’ or ‘sit’ while performing the action de-
noted by the main verb’ although the construction is
often used “according to the bodily position of the
subject” (Kuteva 1999, 193). On this basis, Kuteva
(1999, 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 constructions 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 station-
ary and not being active. In her article on aspect
in Swedish, Greek and Polish, Lindvall (1997, 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 ex-
tended duration of the content verb.

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

It is, however, not the case that the physical, bod-
ily meaning of the posture verbs has disappeared
completely. In the majority of the examples anal-
ysed 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 sub-
stitute 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, al-
though the actors involved in the driving are more
likely to be sitting up than lying down. In this case,
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 (dis-
regarding the aspectual semantics for the moment),
as seen in (30) and (31) (constructed example).
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 (2), one of 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.

An interesting future investigation would be to look into how human translators translate the PV and V construction into English. This could shed light on to which degree human translations follow guidelines like the ones given in this section as well as on to which degree paraphrasing, a strategy that has not been discussed in this paper, is used.

6 Conclusion

In this paper, I have given a detailed linguistic description of the PV and V construction for Swedish and Danish. It was shown that Google Translate often though not always makes errors when translating the construction, specifically errors relating to aspect and the posture verb of the construction. Errors like this could either come from a lack of identification (i.e. this construction was not identified correctly) or from a lack of knowledge of how to interpret the construction correctly.

The study showed that the PV and V construction differs syntactically from the otherwise very similar regular coordinating construction on a number of points. 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.

Information like this can hopefully contribute to improving translations, either as guidelines in human post-editing or by updating automatically generated phrase translation tables with the information on the PV and V construction, yet further research is needed to test the viability and efficiency of these possible solutions to the current problems with automatic translation of the PV and V construction.

Acknowledgments

Many thanks to Joakim Nivre, Aaron Smith, and Karin Ahlén for their valuable comments on the first draft of this paper.

References


Karen Brøcker. 2013. The Danish går og V construction in a Semiotic Grammar perspective. Unpublished pa-
Appendix A. Collection of Google Translate Examples

The examples presented here all come from corpora in Språkbanken (Swedish) or from 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: translations that are correct with to the PV and V construction (but not necessarily the rest of the sentence), 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 järnvägsspåren.
A man is sleeping on the train tracks.

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

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).

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 og den sorte killing.
He is sleeping like the girl and the black kitten.

En beruset herre ligger och 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 sидder og snakker med en jævnaldrende pig.
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 raket.
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 søkekibbe.
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 og lurer under overfladen. The wildness lies lurking beneath the surface.

Jeg åbner en døse øl 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 sitter och 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 framför 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 cooks 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.