

## Using translation corpora to explore synonymy and polysemy [1]

Thomas Egan

Hedmark University College

### Abstract

This paper investigates whether the Norwegian translation equivalents of the two English verbs *begin* and *start* and of the multi-polysemous preposition *at* can aid us in ascertaining the extent to which the former pair may be said to be synonymous and in tracing the polysemous semantic network of the preposition. The corpus data consist of tokens of the three forms in the English language original texts of the English-Norwegian Parallel Corpus. The study shows that *begin* and *start* are to all intents and purposes synonymous in some, but not all, syntactic frames. It also transpires that, with one significant exception (the Perception sense, instantiated by *look at*) the various senses of *at* cluster into two main semantic sub-networks.

### 1. Introduction

Polysemy and synonymy are sometimes viewed as two sides of the same coin; in both cases we are faced with one-to-many form-meaning relationships. The purpose of the paper is to show how translation corpora can be mined to shed light on such one-to-many relationships. With respect to synonymy, the greater the degree of semantic overlap there is between two lexemes or constructions in language A, the more difficult it should be to predict the original forms given their translations into language B. As for polysemy, it is hypothesised that putatively different senses of a lexeme or construction are more likely to be translated differently than similar senses (see Egan 2012).

The data for this paper are taken from the English-Norwegian Parallel Corpus (ENPC: see Johansson 2007). Compiled under the direction of Stig Johansson at the University of Oslo, the ENPC consists of 50 extracts from English texts of some 12,000 words in length, together with their translations into Norwegian, and 50 extracts from Norwegian texts of similar length, together with their translations into English. Both fictional and non-fictional texts are represented in the corpus. For the present paper only original English texts and their translations into Norwegian were consulted. The data comprise all tokens of the two near-synonymous verbs *begin* and *start* in both the fiction and non-fiction components of the corpus, and all tokens of the multi-polysemous preposition *at* in the fiction component alone. The reason for restricting the study of *at* to the tokens in the fiction component is its frequency. There are more than 2,500 tokens of the preposition in the fiction texts.

Section 2 contains a brief discussion of the notions of synonymy and polysemy and of the possible utility of translation corpus data for their study. Section 3 contains the first case study, of *begin* and *start*. The various senses of the preposition *at* and their translation equivalents, i.e. items in the translations corresponding semantically to items in original texts, are introduced and discussed in section 4. The results for *at* are compared to the results of previous studies of *through* (Egan 2012) and *between* (Egan forth.). Finally, section 5 contains a summary and conclusion.

### 2. Synonymy, polysemy and translation

As mentioned above, synonymy and polysemy both comprise one-to-many form-meaning relationships. This has led to their being viewed as two sides of the same coin, as it were. A moment's reflection, however, will suffice to conclude that this metaphor is flawed, at best. Let us imagine that there are two lexemes, *x* and *y*, both of which are polysemous. For *x* and *y* to be defined as completely synonymous, they would have to have the exact same number of senses and each of these senses would have to be synonymous. It is unlikely, though not of course impossible, that two lexemes in a language would overlap to this extent. Moreover, if a language were to contain two such items, it is likely that over time one of them would fall into disuse. Given the improbability, then, of two polysemous lexemes overlapping completely in all of their meaning extensions, the reverse of the polysemy coin is much more likely to be partial, rather than complete, synonymy. The other side of the complete synonymy coin is monosemy. It is, in fact, doubtful if such a thing as complete synonymy actually exists, at least with respect to two expressions used in the same speech register over a period of time. As Murphy (2003) puts it:

One reason that synonymy is often defined as a *sense relation* is that synonyms usually involve a match between some, but not all, of a word's senses. The *sense relation* description is convenient, then, because it only considers one sense of a word at a time. However, what is related here is not two senses (because in absolute synonymy a single sense is shared), but two lexical units, that is, instantiations of lexical items, associated with a particular sense. (Murphy 2003: 145)

Most studies of synonymy are studies of what Murphy calls *sense synonyms*, polysemous words that overlap in one, or possibly several, but not all of their senses. The challenge facing the researcher into putative synonymy is to distinguish between overlapping and non-overlapping senses of the lexemes being compared. One corpus-based method of establishing the extent of sense synonymy is the behavioural profile approach pioneered by Gries & Divjak (2009), whereby all tokens of the putative synonyms in a corpus are tagged for a wide range of morphological, syntactic and semantic features and then analysed using a hierarchical agglomerative cluster analysis. The results of this analysis show which lexemes are most similar to one another, both syntactically and semantically.

The present study makes use of what Krzeszowski (1990: 25) calls a 2-text translation corpus. The justification for the

employment of cross-linguistic data in a study of synonyms in one language rests on the realisations that, on the one hand “Semantic relations are **universal** at both general and particular levels” (Murphy 2003: 40), while, on the other “all languages have different vocabularies, and so a map of relations in one language is not the same as that in another” (2003: 42). My hypothesis is that we can exploit the differences in coding of similar semantic relations in two languages to throw light on these relations in either language. For instance, it is axiomatic that if two lexemes or constructions in language A are completely synonymous, it should not be possible to predict the original form on the basis of translations into language B. In addition, the greater the degree of semantic overlap there is between two constructions in language A, the more difficult it should be to predict the original forms given their translations into language B. If there turn out to be significant differences between the translations of the two forms of putative synonymous lexemes or constructions, then we are forced to conclude that the translators are picking up on semantic differences between the two originals that were not apparent at first sight to advocates of their synonymy.

While a difference in translation equivalents should always alert us to the possibility of a difference in the semantics of the original items, the opposite is not necessarily the case. It may, for instance, turn out that a semantic distinction in the original language is not evidenced in the language of translation. In this sort of situation, one must take care not to be misled into postulating a greater degree of synonymy in the original language than is in fact the case. The opposite situation, where a semantic distinction in the language of translation is not evidenced in the original language, should not, however, carry the danger of leading to false conclusions.

Of the two phenomena explored in this paper, it is fair to say that polysemy has received a lot more attention in the literature in recent decades than has synonymy. The growth of interest in polysemy over the last thirty years or so was triggered by the exploration of the nature of linguistic categorisation that followed in the wake of linguists’ discovery of Rosch’s work on prototype theory (see, for instance, Geeraerts 2009: 183, Lakoff 1987: 15, Taylor 2003: 45). Of particular interest to linguists has been the interrelation of the various senses of a polysemous lexeme, as pointed out by Cuyckens, Sandra & Rice.

One of the major issues in cognitive lexical semantics over the past two decades has been the analysis of polysemous lexical items in terms of a family-resemblance network of multiple, interrelated senses of usage types. [...] The links between the different senses in a lexical network are manifold (conceptual/semantic overlap, metaphor, metonymy, image-schema transformation) and are supposed to represent the cognitive principles behind the processes of meaning extension. (Cuyckens, Sandra & Rice 1999)

One well-known example of the structure of a polysemous network is Langacker’s (1990) network for the verb *run*, reproduced here as Figure 1. The structure of the network in Figure 1 is informed by Langacker’s knowledge of his own mother tongue and his long training and experience as a linguist. The data on which it is based, however, are the product of his introspection. The introspective approach to the construction of networks may result in various scholars not only positing different nodes in a network, but also different numbers of nodes. In this connection, Tyler & Evans (2001) argue against what they term ‘rampant polysemy’. They maintain that the number of senses in a network should be limited to those spatial senses that cannot be understood given what they term the protoscene and contextual information as well as non-spatial senses. Nevertheless, their approach, like Langacker’s, is based on their own linguistically-informed analysis of data of their own devising.

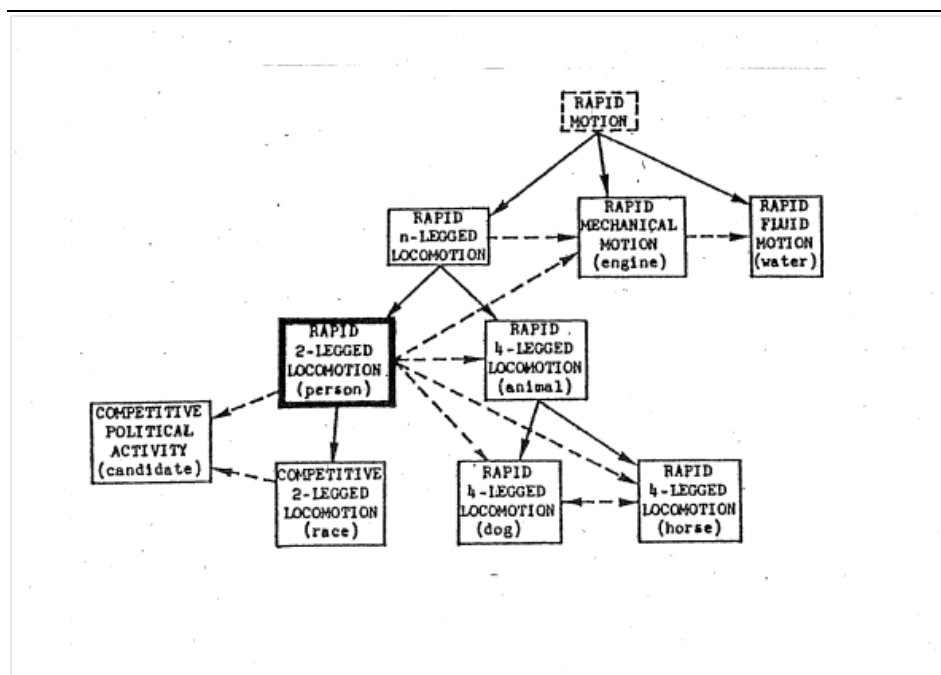


Figure 1. Langacker’s (1990: 267) semantic network for the polysemous verb *run*.

If one wishes to avoid the dangers inherent in introspection as a source for the data to be analysed, one can make use of corpora. For instance, Gries (2006) makes use of a behavioural profile approach to investigate the polysemous network

of the verb *run*. A hierarchical agglomerative cluster analysis showed that “the senses that amalgamated earliest are exactly those that are most strongly branching in the network-like representation of [a figure]” (Gries 2006: 83). The figure to which Gries refers here was drawn on the basis of semantic similarities between the various senses. The sense distinguished as the most central in the ‘run’ network is the ‘fast pedestrian motion’ sense which Langacker intuitively placed at the centre of the network reproduced above. Addressing the question of the links between the senses in the network, Gries computed correlations between the various senses and showed that some of them, including the ‘fast pedestrian motion’ sense and the ‘escape’ senses, are much more similar to one another than are other pairs of senses (Gries 2006: 80).

In my study of *at* in section 4 I again employ a 2-text translation corpus. The justification for the employment of translation equivalents in a study of polysemy rests on two common assumptions about linguistic networks. In the words of Tyler and Evans:

The lexicon constitutes an elaborate network of form-meaning associations (Langacker 1987, 1991a, 1991b), in which each form is paired with a semantic network or continuum (Brisard 1997). This follows from two basic assumptions, widely demonstrated within the framework of cognitive linguistics. First, semantic structure derives from and mirrors conceptual structure (see, for example, Fauconnier 1994, 1997; Heine 1997; Jackendoff 1983; Lakoff 1987). Second, the kinds of bodies and neural architecture human beings have – how we experience – and the nature of the spatio-physical world we happen to live in – what we experience – determine the conceptual structure we have. (Tyler & Evans 2001: 95)

If, as Tyler and Evans maintain, our conceptual structure is determined by the kinds of bodies and neural architecture human beings have and the nature of the spatio-physical world we happen to live in; and if, moreover, semantic structure derives from and mirrors conceptual structure: then one would expect that people with broadly similar cultural backgrounds might, given the similarity between their bodies and backgrounds, speak languages which contain similar (though, of course, not identical) semantic structures. In other words, one can hypothesise that:

If in language E the lexeme (or morpheme or construction) *x* is used with senses:

[a], [b], [c], [d], [e], [f], [g], [h], [i] ....

and in language N the lexeme *y* is used with senses:

[a], [b], [c], [e], [g], [j], [k], [l], [m] ....

and in language F the lexeme *z* is used with senses:

[a], [b], [d], [e], [i], [k], [l], [n], [o] ....

Then one can formulate a hypothesis that one or more of the senses [a], [b] and [e] represent(s) the prototype of each of the lexemes. This hypothesis presumes that one can identify senses cross-linguistically. But how can we be sure that the three lexemes, *x*, *y* and *z* code the same senses [a], [b] and [e]? In other words what evidence do we have for the (partial) equivalence of the semantic spaces in the three languages? In theory there are at least three sorts of evidence for cross-linguistic similarity. The first consists of introspection on the part of someone with a good knowledge of both languages. The second source of evidence might come from laboratory experiments, such as describing pictures or video snippets. The third consists of data from translation corpora. Translation corpora reveal which lexemes or constructions in one language are felt to correspond most closely to a given lexeme in another (see, for example, Dyvik 1998, 2004, Johansson 2007, Noël 2003).

Translation corpora have been mined extensively by Viberg (1998, 1999, 2002, 2003) in studies of the polysymies of certain very common verbs in various languages, including Swedish, English, Finnish and French. He demonstrates that translations can provide evidence of the internal structure of the polysemous network of a single lexeme. Moreover, different senses of a lexeme in language E which are usually translated into language N by one and the same lexeme (or construction) may be hypothesised, according to Garretson (2004), to be more closely related within the semantic network of the lexeme in language E than those translated by different lexemes. Should translations into other languages display a similar patterning, this would serve to strengthen the hypothesis. It goes without saying that any such hypothesis grounded on the testimony of translations should be evaluated in the light of other forms of (mono-linguistic) evidence.

### 3. Synonyms: *begin* and *start*?

Previous studies of *begin* and *start*, such as Freed (1979), Dixon (1991, 2005), Duffley (1999), Mair (2003) and Egan (2008), all agree that, while the two verbs both code ingressive aspect, in the sense that they both code a situation as (re)commencing, they are not completely synonymous. They also agree that it is difficult to tease out the difference between them. As Dixon puts it: “In many sentences *start* and *begin* may be substituted one for the other with little or no change in meaning [...]. But there do appear to be some semantic preferences for each verb, which motivates their use to a considerable extent” (Dixon 1991: 176 & 2005: 181). There is certainly a difference between the incidence of the two in spoken and written English, with *start* being significantly more common in spoken English (Egan 2008: 257).

One of the few clear semantic differences between the two matrix verbs posited in the literature is Freed's (1979: 71) assertion that “[...] only from a sentence with *begin* does it necessarily follow that the nucleus (or characteristic activity) of the event has been initiated; a sentence with *start* followed by a *to V* complement can have as a consequence that only the onset of the event named in the complement has been initiated”. I will return to this assertion below.

There are 433 tokens of the lemma *begin* and 232 tokens of the lemma *start* in the English original texts in the ENPC.

They both occur in four main construction types, listed here as (a)–(d). Raw figures for the distribution of each construction type are illustrated in Figure 2.

- (a) *to*-infinitive complement constructions : i.e. *begin/start to do*  
 (b) *-ing* complement constructions: i.e. *begin/start doing*  
 (c) Intransitive constructions: i.e. *begin/start #*  
 (d) Transitive constructions with nominal objects: i.e. *begin/start something*

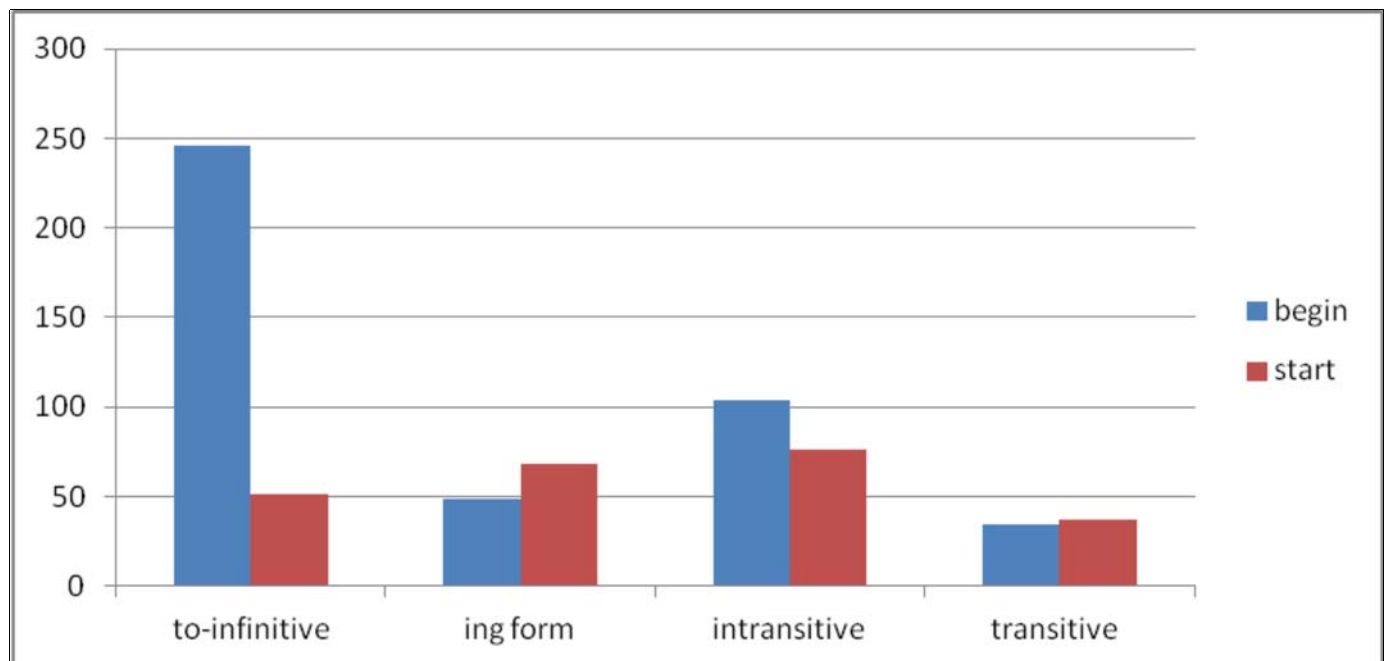


Figure 2. All tokens of *begin* & *start* in the ENPC: four types of complementation

There are five main options for translators of expressions containing *begin* and *start*. These differ as to whether the ingressive aspect is encoded in the predication, and if this is the case, in the manner of its encoding. The five options and the number of times each is used are listed as (e)–(i):

- (e) The Norwegian verb *begynne*: 481 tokens  
 (f) Other ingressive verbs (including *starte* 'start'): 88 tokens  
 (g) Divergent forms (encoding ingression): 47 tokens [2]  
 (h) Ingressive aspect not translated: 34 tokens  
 (i) Whole phrase not translated: 15 tokens

There are thus four alternative ways of translating predications containing *begin* and *start*, given that option (i) consists of avoiding translating the predication altogether. In three of the remaining four options, (e) to (g), the ingressive aspect is encoded, albeit by different forms. In examples (1) and (2), both *begin* and *start* are translated by the Norwegian matrix verb *begynne*, which is cognate with English *begin*.

- (1) After a pause, Dorothy controlled herself and *began* consoling them. (DL1) ... *begynte å trøste*... = began to console.  
 (2) He *started* breathing through his mouth. (JC1) ... *begynte å puste* ... = began to breathe

This is by far the most common strategy for translating both *begin* and *start*. In (3) and (4) the translators opt for alternative (f), coding the ingressive aspect by means of another verb.

- (3) Starvation *began*. (MAW1) *Sulten satte inn*... = ... set in  
 (4) Her problems *started* the day she married him. (SG1) *De oppstod* ... lit. They stood up

Option (g), in which the ingressive aspect is coded by a non-verbal form is illustrated in (5) and (6).

- (5) It is dark now and I stand at the end of a street, where the desert *begins*, and I weep like a fool. (RF1) ... *ved overgangen til ørkenen*... = at the transition to the desert  
 (6) As the story *starts* (ROB1) *I åpningen av denne historien* ... = At the beginning of ...

Finally, the translator may choose to omit the ingressive aspect, in the case of the constructions with non-finite complements, (a) and (b), while still translating the predication in the complement clause. This option has been chosen in (7) and (8).

- (7) Is your scalp *beginning* to burn, dear?“(RD1) *Svir det i hårbunnen?* = Is your scalp burning?  
 (8) The phone *started* to make gravelly noises. (PM1) *Telefonen gryntet...* = The telephone grunted

The extent to which the translators availed of the five strategies (e) to (f) is indicated in Figure 3.

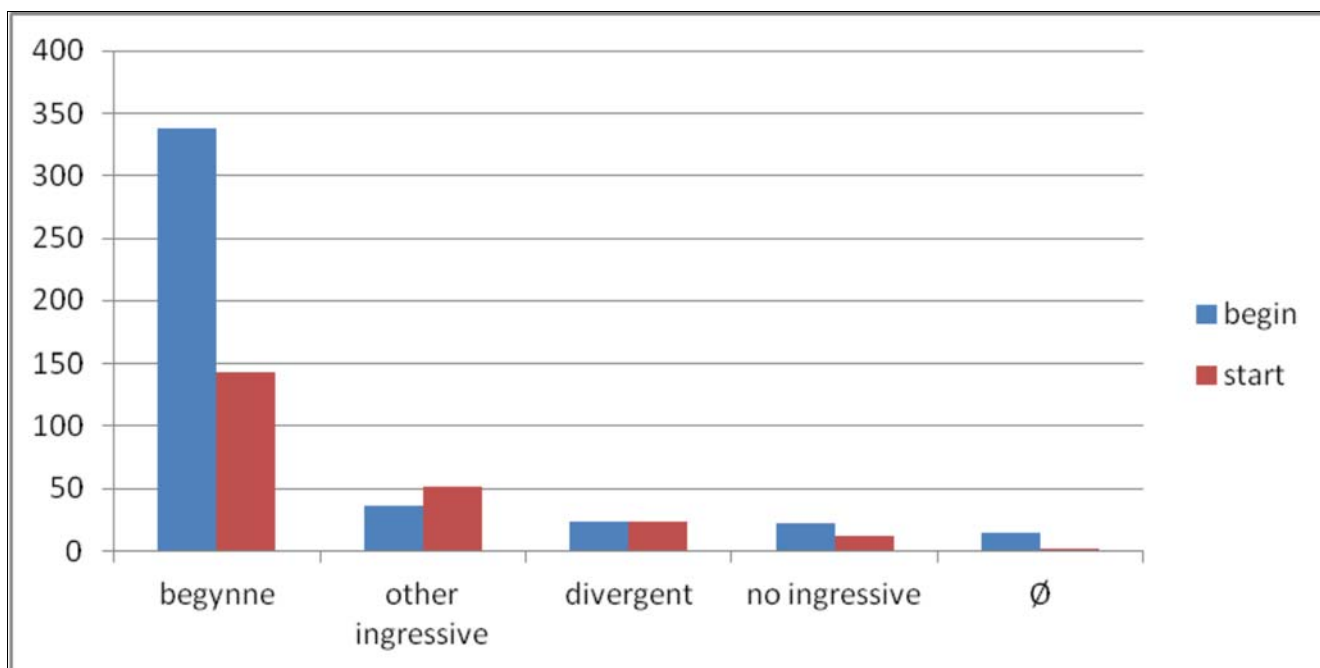


Figure 3. All tokens of *begin* and *start* in ENPC: five translation strategies

	begynne	other ingressive	divergent	no ingressive	Ø	totals
<b>Begin to</b>	200	18	4	17	7	246
<b>Start to</b>	39	5	3	4	0	51
<b>Begin -ing</b>	40	3	0	5	1	49
<b>Start -ing</b>	48	7	4	8	1	68
<b>Begin intrans.</b>	71	11	17	-	5	104
<b>Start intrans.</b>	41	17	11	-	0	69
<b>Begin + NP</b>	27	4	2	-	1	34
<b>Start + NP</b>	15	12	6	-	0	33
<b>All begin</b>	338	36	23	22	14	433
<b>All start</b>	143	52	24	12	1	232
<b>All tokens</b>	481	88	47	34	15	665

Table 1. Translation strategies according to four types of predication

The difference between the translations of *begin* and *start* shown in Figure 3 and Table 1 is significant at the  $p=0.00001$  level according to the Fisher Exact test with 4 df. However, the difference between the two forms in the construction with *to*-infinitive complements, type (a), is not significant at the  $p=0.05$  level ( $p=0.264019$ ). Nor is there any significant difference between the forms with *-ing* complements, type (b) ( $p=0.418827$ ). It follows that the overall difference must be due to differences in either the translations of intransitive and transitive nominal constructions, types (c) and (d), or to both of these. This is indeed the case. The difference between the two forms in the intransitive construction is significant ( $p=0.029692$  according to the Exact test with 3 df) and the difference between the two forms in the transitive construction is also significant ( $p=0.009382$ ). [3]

Before proceeding to look in greater detail at the transitive nominal and intransitive constructions, I return to Freed's (1979) assertion that by employing the *begin to* construction, as opposed to the *start to* construction, the speaker always guarantees that the action in the complement clause was actually carried out. Consider the evidence of examples (9)–(11).

- (9) I was so afraid that I got down from the barrel and *started* to move away when the girl pointed and cried:

(BO1) ... *skulle til å gå* ... = was about to go

(10) I *began* to move away when my legs brushed against something hairy. (BO1) ... *hadde så vidt begynt å gå*... = had just about begun to move...

(11) Before he could add, as he had *begun* to, suppressing a tone of irony, “Only the people”, she exclaimed, “Thank God for that!”(RR1) *Før han rakk å tilføye*... = Before he managed to add...

The evidence of (9) confirms Freed’s conclusion that the complement of ‘start to’ may occasionally be interrupted before the subject has started to realise it. This element of non-realisation is brought out by the translator who employs the Norwegian equivalent of the ‘was about to’ construction. (10) is written by the same author and translated by the same translator as (9). It also contains the same complement predicate ‘move away’. However, the translation in (10) differs from that in (9) in that the subject is encoded as having taken the very first step towards the realisation of the complement event. The differences in the two translations may thus be seen as providing support for Freed’s contention. This cannot be said of example (11), in which the translation only serves to bring out the element of non-realisation already coded in the initial adverbial clause in the original English text. The translator has here availed of option (i), non-translation of the ingressive aspect. One way in which Freed’s thesis regarding ‘begin to’ could be saved is by considering the action of speaking as containing a prior act of thinking. Thus the (non-)speaker in (11) had certainly thought of what he was going to say. He was, nevertheless, interrupted before he actually did so. At the very least we can conclude that (11) constitutes a possible counter-example to the hypothesis that ‘begin to’ always guarantees the initiation of the complement event.

Turning now to the two types of construction that exhibit significant differences in translation equivalents, the transitive nominal object and intransitive constructions, these were analysed with respect to various parameters, including animacy and specificity of the subject (and object), TAM (Tense, Aspect, Mood) features of the verb and adverbial modification. Many of these parameters proved not to be significant, and no further mention will be made of them. Figure 4 contains raw figures for the translation options in the case of the transitive nominal constructions.

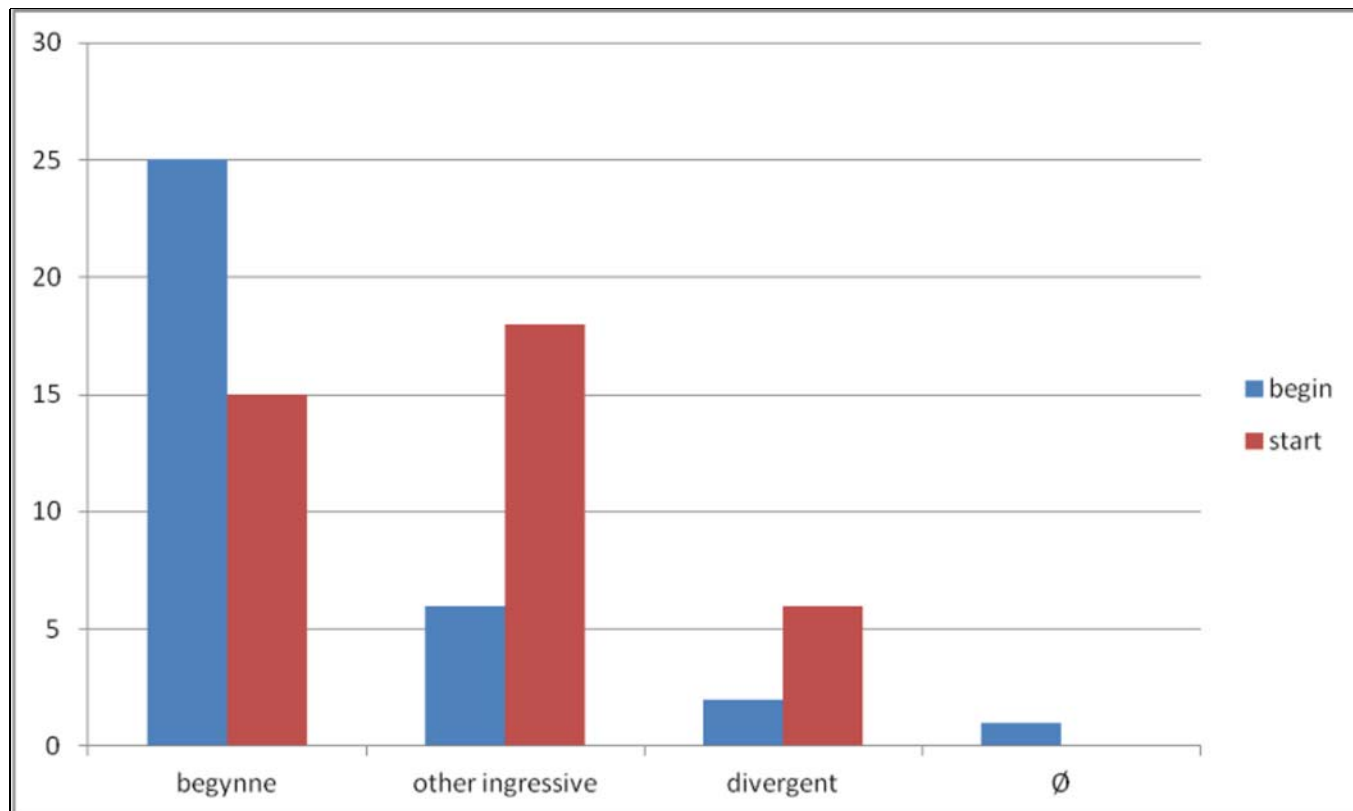


Figure 4. Raw figures for translations of transitive tokens of *begin* and *start* with nominal objects

To begin with one area in which the translations of the two verbs resemble one another closely, those translated by *begynne*, we may consider the semantic type of the objects in these tokens. Table 2 contains details of the categorisation of these objects in five broad semantic fields.

Semantic field of object	<i>Start</i>	<i>Begin</i>
Life	2 (13%)	4 (16%)
Work	2 (13%)	6 (24%)
Education	4 (26%)	4 (16%)
Physical action/product	6 (40%)	8 (32%)

<b>Abstract</b>	1 (7%)	3 (12%)
<b>Total</b>	15	25

Table 2. Transitive tokens translated by *begynne*

As an illustration of the similarity between the translations of the two English constructions, consider (12) and (13), both of which contain an object from the semantic field 'life', and (14) and (15), both of which contain an object from the semantic field 'work'.

- (12) *Begin* a new life abroad? (BC1) *Begynne et nytt liv* = Begin a new life
- (13) He can work hard to convince his readers that the reason he went to California in 1913 was to try to *start* a new life for himself and Pauline ... (MD1) ... *å begynne et nytt liv* ... = to begin a new life
- (14) Instead he *began* work as a filing clerk... (RF1) *I stedet begynte han å arbeide som arkivar...* = ... began to work
- (15) When she had first *started* work ... (AB1) *Da hun begynte å arbeide...* = ... began to work

One would not need to consult a translation corpus in order to discover that *start* and *begin* occur with the same semantic types of object in the transitive nominal construction. A monolingual English language corpus could serve the same purpose. On the other hand, the fact that different translators employ the exact same forms in (12) and (13) and in (14) and (15) does serve to underline the extent to which the two constructions are felt to be synonymous.

The raw figures in Figure 4 show that *start* transitive constructions with nominal objects are more likely to be translated into Norwegian by ingressives other than *begynne*. A total of 18 tokens of *start* are translated by 10 forms, while 6 tokens of *begin* are translated by 5 forms. Two forms are common to translations of both verbs, *grunnlegge* ('found') and *påbegynne* ('begin on'). Both tokens of *grunnlegge* are from the same text and the object in both cases is the same: colony (of bees).

Of the ingressives exclusive to translations of *start*, two stand out in terms of number. Four tokens of *start* are translated by its Norwegian cognate *starte*, as in (16), and three by *sette i gang* ('get going', lit. set in motion), as in (17).

- (16) They're not interested in harming the earth or *starting* wars. (ROB1) ... *i å starte kriger...* = in to start wars
- (17) As a result, it may cost more in foreign exchange to *start* domestic arms production. (CS1) ... *å sette i gang egen våpenproduksjon* = to get one's own arms production going

The occurrence of *starte* as a translation equivalent of *start*, but never of *begin*, may possibly be attributed, at least in part, to translation effects, the cognate form being the first to spring to the mind of the translator. On the other hand, it may also be the case that the English and Norwegian lemmas share a semantic component absent from the semantic make-up of both English *begin* and Norwegian *begynne*. Such a component might consist of the lack of any implication of the prior existence of a related situation, and would also motivate the employment of verbs like *sette i gang* in (17) and *oppstå* ('merge', lit. stand up) in (28).

There are three ingressives exclusive to *begin*. Of these, only one occurs more than once, *innlede*, which means open (lit. lead in), as in (18).

- (18) The banks, bursting with dollars, *began* a hard sell to encourage developing countries to borrow them. (LTLT1)... *innledet en beinhard salgsprosess* = opened

The other two forms just used for *begin* are *legge ut på* ('set out on') and *ta fatt på* ('get a grip on').

The second construction that exhibits significant differences in translation equivalents of *begin* and *start* is the intransitive construction. Figure 5 contains raw figures for both verbs.

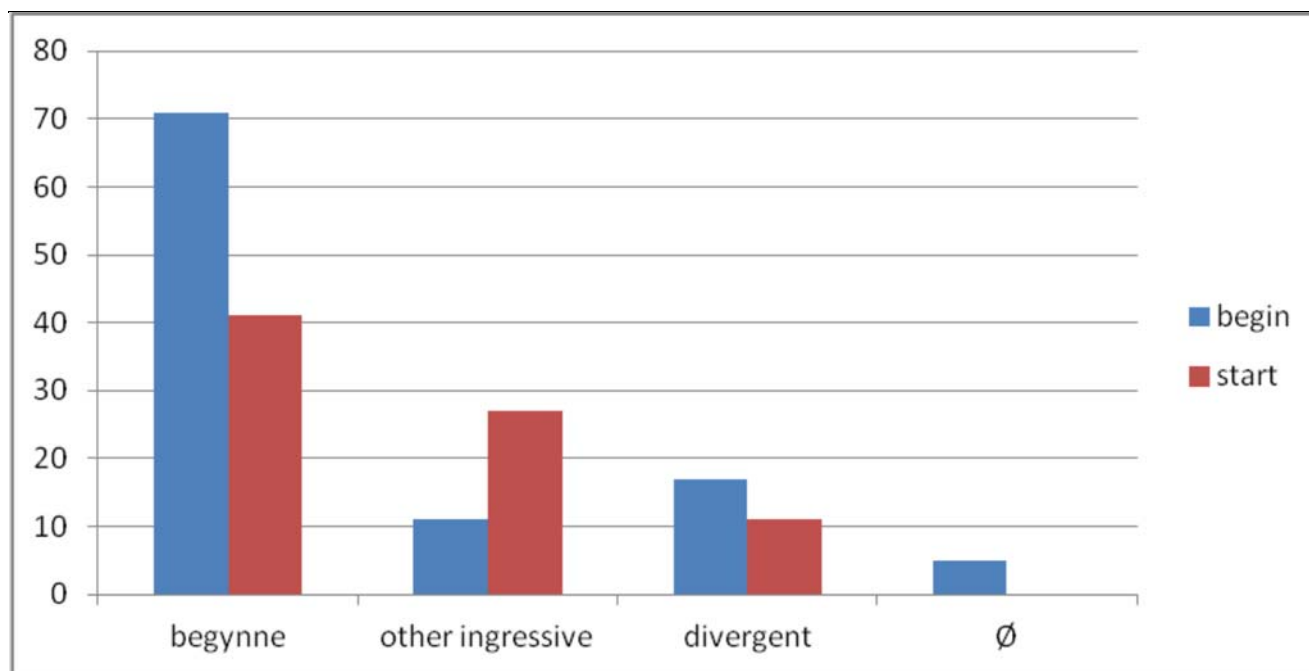


Figure 5. Raw figures for intransitive tokens of *begin* and *start*

Figure 5 resembles Figure 4 closely, at least with respect to the tokens translated by options (e) and (f), i.e. translation by *begynne* and by other ingressives. As was the case with the transitive nominal constructions there is little to distinguish the tokens of *start* and *begin* translated by *begynne*. There is one difference between them, nonetheless, that is perhaps worth noting. This concerns the tokens in which they occur with time adverbials. There are fourteen such tokens in the case of both verbs. The difference between them is related to the definiteness of the adverbial. 11 tokens of *start* (79%) occur with an indefinite time adverbial, as in (19), while 10 tokens of *begin* (71%) occur with a definite time adverbial, as in (20).

- (19) Most boys *started* earlier, but Macon had kept delaying it. (AT1) ... *De fleste guttene begynte tidligere...* = ... began earlier
- (20) The scientific study of the relationship between brain and mind *began* in 1861...(OS1) ... *begynte i Frankrike 1861...* = ... began in France 1861

As was the case with the semantic types of object in the transitive nominal construction, one would not have needed to consult a translation corpus in order to distinguish between the two verbs when used intransitively in terms of their collocating adverbials. On the other hand, sometimes the translations can reveal differences not immediately apparent in data from a monolingual corpus. For instance, the Norwegian adverbial *på ny(tt)* (anew) is used in the translation of six of the *start* tokens and none of the *begin* tokens. Two of these are cited as (21) and (22).

- (21) I hated the discouraging task of *starting over* (TH1) ... *begynne på nytt...* = ... begin anew
- (22) The planet becomes barren *again*, and so hot that there is no way for daisy life to *start* again. (JL1) ... *å begynne livet på ny* ... to begin life anew

Of the remaining four tokens translated by *på ny(tt)*, three contain 'start again' in the original texts and one 'start all over again'. There are also three tokens where 'start again' is rendered by *begynne igjen* (= begin again) as in (23), as well as three in which this translation is employed for 'begin again', as in (24).

- (23) And as soon as he came out of prison it *started* again ... (NG1) ... *begynte igjen* ... = began again
- (24) Once she was installed in her bed her talking stopped; soon the moans *began* again ... (AB1) ... *begynte igjen* ... = began again

(24) is typical of the three 'begin again' tokens insofar as they all encode the resumption of a situation after a comparatively short time interval, as indicated here by the adverbial *soon*. When something is said to start again, on the other hand, the interval in question is much longer, both in the cases where the adverbial *igjen* is used by the translator, as in (23), or where the translator prefers *på ny(tt)*, as in (22). It would seem that when an interrupted situation is re-begun, the trajector can continue where he or she had left off, as it were. When an interrupted situation is restarted, one is more likely to begin at the very start.

As is the case with the constructions with transitive nominal objects, there are significant differences between the two verbs when it comes to translations using other ingressive verbs. 27 tokens of *start* are translated by 12 ingressives, while 11 tokens of *begin* are translated by eight ingressives. There are three forms in common, *ta til* (lit. take to), used by



the same translator in (25) and (26), *komme i gang* (get going) and *sette i* (set in).

- (25) The full exercise of their powers shall *start* from the first day of the third stage.(MAAS1) *skal ta til fra første dag...* = will get going (lit. take to)
- (26) 1. The second stage for achieving economic and monetary union shall *begin* on 1 January 1994.(MAAS1) ... *skal ta til 1. januar 1994...* = will get going (lit. take to)

Of the ingressives exclusive to *start*, there are seven tokens translated by its Norwegian cognate *starte* as in (27) and 4 by *oppstå* ('emerge', lit. stand up), as in (28).

- (27) We *start* from the Embankment. (PDJ3) *Vi starter ...* = We start
- (28) Your mother's problems didn't *start* yesterday.(SG1) ... *oppstod ikke i går.* = lit. didn't stand up

We have already seen that *starte* is also used to translate *start* in transitive nominal constructions, like (16). It is never used to translate *begin*. Two other translation equivalents used more than once for *start* are *skulle til* ('be about to') and *ta fatt på* ('get going on', lit. get a grip on).

Of five ingressives exclusive to *begin*, two occur more than once, *åpne* ('open'), as in (29) and *innlede*, as in (30), which also means open (lit. lead in).

- (29) The Senator's letter began "Dear Donald".(RDA1) ... *åpnet med "Kjære Donald"* ... = opened with ...
- (30) The year began with lunch.(PM1) ... *ble innledet med ...* = ... was opened with

To conclude this section on *begin* and *start*, we have seen that if we look at the various constructions containing *begin* and *start* through the prism of Norwegian translations, we are unable to predict originals in the case of *begin to* and *start to*, *begin -ing* and *start -ing*. These two pairs would appear to be, to all intents and purposes, sense synonyms. This degree of synonymy is presumably also a prerequisite for the current expansion of *start* at the expense of *begin*, a development noted by Mair (2002) and Skandera (2003). In the case of the constructions with nominal objects and the intransitive constructions, differences in favoured translation options may reflect semantic differences in the English originals. For instance, in the case of interrupted situations in intransitive constructions we have seen that the longer the interruption, the more likely its recommencement is to be coded by *start* rather than *begin*. In addition, only tokens of *begin* are translated by Norwegian equivalents of 'open', while only tokens of *start* are translated by *oppstå* (= stand up). The use of a verb meaning 'open' implies the pre-existence of a participant (trajector or landmark) or process, since one cannot open something that does not already exist. On the other hand, the use of a verb like *oppstå* (= stand up), implies the creation of a new participant or process. The implications of these two translation equivalents may be taken as support for Freed's contention that "*start* refers to the onset of an event while *begin* refers to the initial temporal segment of the nucleus of an event" (Freed 1979: 71).

#### 4. The polysemous network structure of *at*

According to Herskovits (1986: 127) *at* is one of three basic topological prepositions in English (the other two are *in* and *on*). She states that: "The use types of *at* center around one ideal meaning: *at*: for a point to coincide with another" (Herskovits 1986: 128). There are many possible types of spatial coincidence, i.e. many possible locations of the trajector vis-à-vis the landmark, as pointed out by Lindstromberg (who uses the term 'Subject' for the more common 'trajector').

*AT* itself is often quite vague about whether the Subject is near the Landmark but not touching it; whether it is right by the Landmark *and* touching it; or whether indeed the Subject is on, in or among the Landmark. For this reason, even the spatial meaning of *AT* is barely depictable. (Lindstromberg 2010: 173)

This variety with respect to the exact locational relation of the trajector to the landmark in the spatial meaning of *at* is reflected in the translations in the ENPC. While there are close equivalents in Norwegian to the other two main topographical prepositions mentioned by Herskovits, *in* and *on*, this is not the case for *at*. In fact we often find tokens of locational *at* translated by the Norwegian equivalents of the other two prepositions, as in (31) and (32).

- (31) Once she had taken all three of them to hear Myra Hess play Bach *at* the National Gallery.(AB1) ... *i The National Gallery (Nasjonalgalleriet)*. = ... *in the National Gallery*
- (32) The foghorn *at* Mouille Point at night, if the wind is right.(ABR1) *Tåkeluren på Mouille Point* = ... *on Mouille Point*

(31) and (32) may be said to be underspecified as to the exact location of the trajector with respect to the landmark. That is, in (31) it is of no importance exactly where *at* (or *in*) the National Gallery the listeners were actually located. In other cases, such as in (33)–(35), the area in question is more circumscribed.

- (33) Ma and Pa were *at* the front door of a dirty old house, trying to get a key in the lock.(ST1)... *Mammi og pappi sto foran døren...* = ... *in front of the door*
- (34) It put her down *at* the Crown and Anchor *at* Cobb's Marsh, only fifty yards from her cottage. (PDJ3) *Den*

*satte henne av utenfor Crown and Anchor* = ... outside the Crown and Anchor...

- (35) He'd stand *at* the bedroom window looking over the neighborhood ... (AT1)... *Han stod ved soveværelsesvinduet* = ... by the bedroom window

Even in cases where the coincidence of the trajector and landmark do not admit of any great variation in terms of actual location, their relationship may be construed in quite different ways, as evidenced by (36) and (37).

- (36) ...when hip to haunch *at* Anna's desk, they checked the agenda of Victor's day. (JC1) ...*med hofte mot lår bøyd over Annas skrivebord...* = ... over Anna's desk
- (37) "We'll see," she said, not quite aloud, her fingers church-and-steepled *at* her chins ... (JC1)... *med fingrene under hakene sine...* = ... under her chins

The examples in (31)–(37) do not exhaust the list of Norwegian prepositions used to translate spatial *at*. Other prepositions used by at least two translators include *langs* ('along'), *rundt* ('around') and *hos* ('at the home of', 'chez'). The very wealth of translational options serves to underline the underspecification of mutual spatial orientation coded by English *at*, at least when it codes location. When used with a motion predicate to encode the end-point of a path, there is much less variation in prepositional use, the majority of translators employing *til* ('to'), as in (38), although we also find multiple uses of *mot* ('towards'), *på* ('on') and *ved* ('by'). Goal predications closely resemble motion predications, differing from them in that no physical motion in the direction of the target actually takes place. They are also most commonly translated by *til* ('to') as in (39), followed in frequency by *på* ('on') and *mot* ('towards').

- (38) We arrived *at* the island... (BO1) *Vi kom frem til øya...* = ... to the island
- (39) She winked *at* me. (TH1) *Hun blunket til meg.* = ... winked to

Temporal predications coded by *at* are translated into Norwegian by a broad range of prepositions, the most common of which are *i* ('in'), as in (40), *om* ('about'/'around'), as in (41), *på* ('on'), as in (42), *til* ('to'), as in (43), and *ved* ('by'), as in (44).

- (40) She had a knack for finding the right thing *at* the right moment. (MD1) ... *i det rette øyeblikk...* = ... in the right moment
- (41) Nor, except *at* night, did you see any sign of Cabot. (JSM1) ... *unntatt om kvelden.* = ... except about the night
- (42) It was *at* this point that she gave up hopes for her own life... (AB1) *Det var på dette tidspunkt ...* = ... on this point
- (43) We had about seven hundred refugees *at* any one time... (ABR1) ... *til enhver tid* = ... to every time
- (44) *At* dawn he accepted a cup of coffee ... (AT1) *Ved daggry* = By dawn

In addition to spatial and temporal predications, I have divided the predications coded in English by *at* into five more main types. These include predications of perception, as in (45), of degree, as in (46), of reaction, as in (47), of event, as in (48), as well as tokens in which the *at* phrase functions as the complement of an adjective, as in (49). These I have labelled 'specifying' since they serve to specify the exact nature of the quality coded by the adjective. [4] There is no space here to illustrate the variety of prepositional and divergent constructions used to encode all five. Instead, each will be exemplified with the most common prepositional translation.

- (45) He glared *at* Willie. (MM1) *Han glodde på Willie.* = ... glared on
- (46) Arthur could *at* least make his wife weep... (FW1) ... *Arthur kunne i det minste...* = ... in the least
- (47) Young Ben was horrified *at* his mother's act of kindness. (FW1) ... *over sin mors vennlige handling* = over his mother's friendly act
- (48) She had once eaten a mouthful of rat *at* a banquet in Belize. (ST1) ... *ved en bankett* = by a banquet
- (49) I was good *at* swimming. (RDO1) ... *til å svømme* = to to swim

Of all types of predications containing *at*, predications of perception are the most homogeneous with respect to their translations. Indeed perception verbs are homogenous in their predilection for collocating with *at* in English (see Perek & Lemmens 2010: 11). Just over 84% of the tokens of 'look at', 'stare at' etc. are translated into Norwegian by a perception verb followed by the preposition *på* ('on'). It may be worth noting here that it was only in the Late Modern period that *at* replaced (*up*)*on* in English as the default preposition with these verbs. As for predications of degree, *at* occurs in several very common fixed expressions, such as 'at least' and 'at all'. These are most often translated by fixed expressions, as in (46) for 'at least'. Two prepositions account for most of the translations of predications of reaction. The most common of these is *over* ('over'), as in (47). The other is *av* ('of'). Event expressions containing *at*, as in (48), are most often translated by expressions containing the Norwegian equivalent of 'by'. Finally, in all but two of the specifying tokens translated by prepositions, the preposition of choice is either *til* ('to'), as in (49), or *i* ('in').

The nine broad categories of Location, Motion, Goal, Time, Perception, Degree, Reaction, Event and Specifying do not

suffice to categorise all tokens of *at* in the ENPC. There remains a residue of over a hundred tokens of idiomatic expressions, three of which are exemplified by (50)–(52).

- (50) “Oh, Simon, you don’t have to tell me that you are an artist *at heart*.” (RDA1) ... *på bunnen* = on the bottom
- (51) Bert stood upright, slightly bent forward, arms *at ease*, looking at her. (DL2) ... *armene var avslappet* = (his) arms were relaxed
- (52) David disliked this trait of Harriet’s, a fatalism that seemed so *at odds* with the rest of her. (DL1) ... *var så ulikt henne* = was so unlike her

As was the case with perception predications, *på* (‘on’), as in (50), is the most frequent preposition found in translations of idiomatic expressions containing *at*. However, the majority of idioms are not translated using a preposition at all, but by a divergent construction, as in (51) and (52).

There are 2,570 tokens in all of *at* in the English original fictional texts in the ENPC. In the case of 142 of these the translator has omitted to translate the phrase containing the preposition. Of the remaining 2,428 tokens, 583 are translated by a divergent construction. The remaining 1,845 are translated by a congruent construction, i.e. one containing a preposition. By far the most common of these is *på* (‘on’), which accounts for some 31% the congruent translations. Since it is the most common translation equivalent, the tokens translated by *på* were separated from the remaining congruent translations. All of the 2,428 translated tokens were labelled according to whether they were translated by a prepositional construction containing *på*, a construction with another preposition, or a divergent construction. The translations of the various senses were then compared to one another, yielding the statistical results contained in Table 3. In the table comparisons which yielded non-significant results at the  $p = 0.05$  level are in red.

	Time	Degree	Event	Goal	Idiom	Motion	Perception	Reaction	Specifying
Location	<0.0001	<0.0001	0.1531	0.5814	<0.0001	0.0061	<0.0001	0.0001	0.0007
Time		0.0046	<0.0001	<0.0001	0.0035	<0.0001	<0.0001	<0.0001	<0.0001
Degree			<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	0.0002
Event				0.5283	<0.0001	0.2433	<0.0001	0.0769	0.0222
Goal					<0.0001	0.0341	<0.0001	0.0022	0.0034
Idiom						<0.0001	<0.0001	<0.0001	<0.0001
Motion							<0.0001	0.3952	0.1111
Perception								<0.0001	<0.0001
Reaction									0.3965

Table 3. Chi-square probabilities with two degrees of freedom to four decimal points for three sorts of translation of nine subtypes of *at*

Table 3 shows that there are significant differences in mode of translation between the perception, time, degree and idiom senses and all the other four senses. The location sense is translated in a similar fashion to both goal and event, being significantly different to the remaining seven senses. The specifying sense is translated similarly to motion and reaction and the goal sense similarly to location and event. Two of the senses, motion and reaction, are similar to three other senses, while the event sense is the most central in the network, as it is translated similarly to four other senses. Figure 6 is an attempt to convey in a network the resemblance between senses in Table 3.

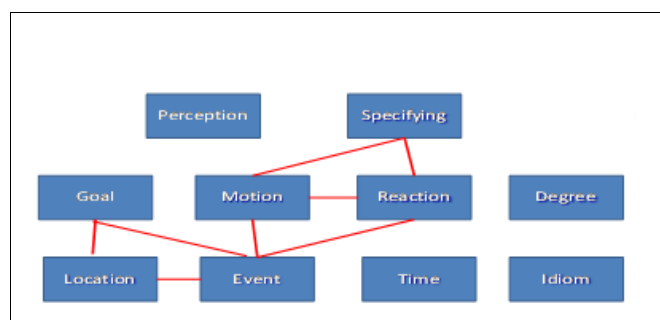


Figure 6. Network of *at* as reflected in Norwegian translations

Figure 6 illustrates a network of the various senses of English *at* as reflected in Norwegian translations. More accurately, it represents a network of similarities and differences between the way ‘at-ness’ is encoded in English and Norwegian. The degree, time and idiom senses are placed in proximity to one another because, although the translation options chosen in their cases are significantly different at  $p = 0.05$ , they are still less different to one another, as indicated by the statistical calculations, than they are to the other senses.

The structure of the network for the senses of *at* in Figure 6 resembles in one important respect similar networks drawn for *through* (in Egan 2012) and *between* (in Egan forth.) on the basis of their Norwegian translation equivalents. In the case of all three English prepositions, spatial meanings, which are not only the most frequent, but also often taken to represent the core or basic sense of the prepositions, do not assume a central position in the network of cross-linguistic correspondences. The most central meanings, in the sense of the meanings which are most similar to other meanings, including the spatial ones, are metaphorical extensions of the spatial senses, the channel sense in the case of *through* and the scalar sense in the case of *between*. These two senses occupy the same position in their respective networks as does the event sense in the case of *at*. This suggests that, at least for the prepositions in question, we should distinguish between the basic sense and the central sense, despite the fact that both basicness and centrality are often taken to be characteristics of a category prototype. The semantic network of a preposition may instead resemble the sort of structure portrayed in Figure 7.

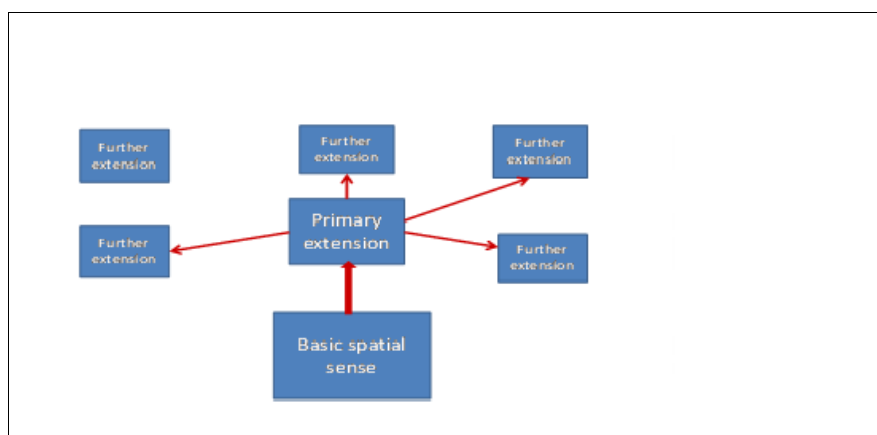


Figure 7. A model showing the links between the senses in a prepositional network

Figure 7 is an extremely simplified sketch of a possible prepositional network. It does not attempt to illustrate any possible links between the various extended senses. What it does serve to emphasise, however, is the fact that the basic spatial sense, while it may be both the most frequent and the most psychologically salient, and thus a prime candidate for status as category prototype may well not be the most central sense in the network. (See Arppe et al. 2010 for a discussion of some of the problems related to operationalising the notions of frequency and salience in positing prototypicality.) Centrality is often also taken to a characteristic of the prototype (see Gries 2006: 76). Should further research result in our positing similar networks for other prepositions, we may need to further nuance our conception of prototypicality.

## 5. Summary and conclusion

In this paper I have explored how translation corpora may contribute to the exploration of synonymy and polysemy. In section 2 I mentioned some previous approaches to the study of synonymy and polysemy and proffered some arguments as to how translation corpora may be exploited in the study of these sense relations. Section 3 was devoted to a case study of the putatively synonymous verbs *begin* and *start*. The evidence of their Norwegian translation equivalents points to the two verbs being sense synonyms when they occur with *to*-infinitive and gerund complements. When used intransitively, or with a nominal object, on the other hand, there are significant differences in the translation options chosen by translators. These differences may point to differences in the semantics of the expressions in the original language.

Would the use of a translation corpus prove a useful addition to the methodological toolkit of the linguist studying synonymy? Insofar as translation corpora allow us access to the intuitive insights of a cross-section of competent speakers, they are obviously to be preferred to the intuitions of a single analyst. More particularly, translation corpora can contribute to studies of synonymy in two ways. Firstly, similarities/differences in translation equivalents can alert us to similarities/differences in the original constructions of which we may not have been aware. Secondly, the actual forms chosen in the case of different translation equivalents may highlight aspects of the semantics of the original forms (in the present case *åpne/innlede* for 'begin' and *oppstå* for 'start'). It goes without saying that translation corpora containing more than one language may be even more useful as diagnostics for differences and similarities.

Section 4 was devoted to a study of the Norwegian translation equivalents of the preposition *at*. Nine senses of the preposition were distinguished and the translations of these various senses were analysed as either congruent or divergent in form. The congruent tokens were further divided into those containing the Norwegian preposition *på* (on) and those containing another preposition. The translation equivalents of the nine senses were compared to one another and a network was drawn on the basis of these comparisons. It transpired that the event sense is the most central in the network, insofar as it is the sense with the greatest number of links to the other senses.

I asked above whether the use of a translation corpus would prove a useful addition to the methodological toolkit of the linguist studying synonymy. The same question may also be raised with respect to polysemy. A similar answer may also be given. The point has been forcefully made in recent years (by, among other, Arppe et al. 2010 and Gilquin & Gries

2009) that we need to supplement corpus evidence with experimental evidence. Translation corpora partake of some of the advantages of elicitation experiments. They contain the intuitive linguistic responses of competent language users to a series of linguistic prompts. Moreover, since translators are not aware that their products will be subject to linguistic analysis, the dangers inherent in the observer's paradox do not apply. Just as systematic differences in translation equivalents may serve to highlight differences in meaning between putative synonyms, they may also highlight differences in meaning between the various senses of polysemous items. Translation corpora provide a rich vein for future research into these semantic relations.

## Notes

[1] I would like to thank the editors and two anonymous referees for insightful comments and suggestions for improving this paper. I would also like to thank Susan Nacey for her cooperation on a previous paper on *begin* and *start*, "The (near-)synonyms *begin* and *start*: evidence from translation corpora", which we co-presented at the conference "Rethinking Synonymy" at the University of Helsinki in October 2010.

[2] By 'divergent forms' I mean forms that differ syntactically from the original item, while at the same time conserving the meaning of the original.

[3] The reason why there are only three and not four degrees of freedom in the case of types (c) and (d) is that translation option (h), which involves dropping the ingressive element while retaining the complement predicate is not possible in constructions which lack a complement predicate.

[4] One reviewer makes the point that "it would seem that tokens "of reaction", as in "horrified at" are simply a special case of the "specifying" class, as in "good at". Certainly there is a generalization that can be made about the fact that we may say "horrified/surprised/astonished/shocked at", but I see the real connection here as between the adjective and the preposition, not between the preposition and its own complement." While I agree that the relationship between the adjective and the preposition is similar in the two types, I would still prefer to distinguish between them on the basis of the larger constructions in which they occur. Thus in the reaction type the trajector is an Experiencer and the landmark of the preposition codes the Stimulus. In the specifying type, on the other hand, the trajector is a Theme and the landmark of the preposition merely further specifies the quality predicated of this Theme.

## References:

- Arppe, Antti, Gaëtanelle Gilquin, Dylan Glynn, Martin Hilpert & Arne Zeschel. 2010. "Cognitive Corpus Linguistics: five points of debate on current theory and methodology". *Corpora* 2010 Vol. 5(1): 1–27
- Cuyckens, Hubert, Dominiek Sandra & Sally Rice. 1999. "Towards an empirical lexical semantics". *Human Contact Through Language and Linguistics*, ed. by Birgit Smieja & Meike Tasch, 35–54. Wiesbaden: Peter Lang. Reprinted in *The Cognitive Linguistics Reader*. 2007, ed. by Vyvyan Evans, Benjamin K. Bergen, & Jörg Zinken. 57–74. London: Equinox.
- Dixon, Robert M. W. 1991. *A New Approach to English Grammar, on Semantic Principles*. Oxford: Clarendon Press.
- Dixon, Robert M. W. 2005. *A Semantic Approach to English Grammar (2nd. ed.)*. Oxford: Oxford University Press.
- Duffley, Patrick J. 1999. "The use of the infinitive and the -ing after verbs denoting the beginning, middle and end of an event". *Folia Linguistica*, 33, 295–331.
- Dyvik, Helge. 1998. "A translational basis for semantics". *Corpora and Cross-linguistic Research: Theory, Method, and Case Studies*, ed. by Stig Johansson & Signe Oksefjell, 51–86. Amsterdam: Rodopi.
- Dyvik, Helge. 2004. "Translations as semantic mirrors: from parallel corpus to wordnet." *Advances in Corpus Linguistics: Papers from the 23rd International Conference on English Language Research on Computerized Corpora (ICAME 23)*, Göteborg 22–26 May 2002, ed. by Karin Aijmer & Bengt Altenberg, 313–326. Amsterdam: Rodopi.
- Egan, Thomas. 2008. *Non-finite Complementation: A Usage-based Study of Infinitive and -ing Clauses in English*. Amsterdam: Rodopi.
- Egan, Thomas. 2012. "Through seen through the looking glass of translation equivalence: a proposed method for determining closeness of word senses". *Corpus Linguistics: Looking Back - Moving Forward*, ed. by Sebastian Hoffman, Paul Rayson & Geoffrey N. Leech, 41–56. Amsterdam: Rodopi.
- Egan, Thomas. forthcoming. "Between and through revisited". *Proceedings of ICAME31*, ed. by Joybrato Mukherjee & Magnus Huber. (Studies in Variation, Contacts and Change in English). Helsinki: Research Unit for Variation, Contacts, and Change in English.
- Freed, Alice F. 1979. *The Semantics of English aspectual complementation*. Dordrecht: Reidel.
- Garretson, Gregory. 2004. *The Meanings of English "of": Uncovering Semantic Distinctions Via a Translation Corpus*. Unpublished MA thesis, Boston University.
- Geeraerts, Dirk. 2009. *Theories of Lexical Semantics*. Oxford: Oxford University Press.

- Gilquin, Gaëtanelle & Stefan Th. Gries. 2009. "Corpora and experimental methods: a state-of-the-art review". *Corpus Linguistics and Linguistic Theory* 5(1): 1–26.
- Gries, Stefan Th. 2006. "Corpus-based methods and cognitive semantics: the many meanings of *to run*". *Corpora in Cognitive Linguistics: Corpus-based Approaches to Syntax and Lexis*, ed. by Stefan Th. Gries & Anatol Stefanowitsch, 57–99. Berlin: Mouton
- Gries, Stefan Th. & Dagmar S. Divjak. 2009. "Behavioral profiles: a corpus-based approach towards cognitive semantic analysis". *New Directions in Cognitive Linguistics*, ed. by Vyvyan Evans & Stephanie S. Pourcel, 57&75. Amsterdam: Benjamins.
- Johansson, Stig. 2007. *Seeing through Multilingual Corpora: On the Use of Corpora in Contrastive Studies*. Amsterdam: John Benjamins.
- Herskovits, Annette. 1986. *Language and Apatial Cognition: An Interdisciplinary Study of the Prepositions in English*. Cambridge: Cambridge University Press
- Krzyszowski, Tomasz P. 1990. *Contrasting Languages: The Scope of Contrastive Linguistics*. Berlin: Mouton.
- Langacker, Ronald. W. 1990. *Concept, Image, and Symbol: The Cognitive Basis of Grammar*. Berlin: Mouton.
- Lakoff, George. 1987. *Fire, Women and Dangerous Things*. Chicago, IL: University of Chicago Press.
- Lindstromberg, Seth. 2010. *English Prepositions Explained: Revised Edition*. Amsterdam: Benjamins.
- Mair, Christian. 2002. "The changing forms of complementation in late Modern English: a real-time study based on matching text corpora." *English Language and Linguistics*, 6(1): 105–131.
- Mair, Christian. 2003. "Gerundial complements after begin and start: Grammatical and sociolinguistic factors, and how they work against each other". *Determinants of Grammatical Variation in English*, ed. by Britta Mondorf & Günter Rohdenburg, 329–343. Berlin: Mouton.
- Murphy, M. Lynne. 2003. *Semantic Relations and the Lexicon*. Cambridge: Cambridge University Press
- Noël, Dirk. 2003. "Translations as evidence for semantics: an illustration". *Linguistics* 41(4):757–785.
- Perek, Florent & Maarten Lemmens. 2010. "Getting at the meaning of the English at-constructions: the case of a constructional split". *CogniTextes [en ligne]* 5: 1–23. <http://cognitextes.revues.org/331>
- Skandera, Paul. 2003. "Start doing or start to do: Is the gerund spreading in American English?" *English Core Linguistics: Essays in honour of D. J. Allerton*, ed. by Cornelia Tschichold, 343–352. Bern: Peter Lang.
- Taylor, John R. 2003. *Linguistic Categorization* [3rd edition]. Oxford: Oxford University Press
- Tyler, Andrea & Vyvyan Evans. 2001. "Reconsidering propositional Polysemy Networks: The Case of *Over*". *Language* 77(4): 724–765
- Viberg, Åke. 1998. "Contrasts in polysemy and differentiation: Running and putting in English and Swedish." *Corpora and Cross-linguistic Research*, ed. by Stig Johansson & Signe Oksefjell, 343–376. Amsterdam: Rodopi.
- Viberg, Åke. 1999. "Polysemy and differentiation in the lexicon. Verbs of physical contact in Swedish." *Cognitive Semantics. Meaning and Cognition*, ed. by Jens Allwood & Peter Gärdenfors, 87–129. Amsterdam: Benjamins.
- Viberg, Åke. 2002. "Polysemy and disambiguation cues across languages. The case of Swedish *få* and English *get*." *Lexis in Contrast*, ed. By Bengt Altenberg & Sylviane Granger, 119–150. Amsterdam: Benjamins.
- Viberg, Åke. 2003. "The polysemy of the Swedish verb *komma* 'come': A view from translation corpora." *Meaning through Language Contrast. Vol. 2*, ed. by Kasia M. Jaszczolt & Ken Turner, 75–105. Amsterdam: Benjamins.