ROLL UP / VENEZ: AN INVITATION TO CORPUS-BASED RESEARCH IN MOTION TYPOLOGY

ROLL UP / VENEZ: UM CONVITE À PESQUISA BASEADA NO USO EM TIPOLOGIA DE MOVIMENTO

Bert Cappelle

Roll up (and that’s an invitation)
Roll up for the Mystery Tour
Roll up (to make a reservation)
Roll up for the Mystery Tour
The Magical Mystery Tour is waiting to
Take you away, waiting to take you away

Venez (c’est une invitation)
Venez à la Tournée mystérieuse
Venez (et réservez vos places)
Venez à la Tournée mystérieuse
La Tournée Magique et Mystérieuse attend de
Vous amener, attend de vous amener

Part of the lyrics of The Magical Mystery Tour by The Beatles, and the translation into French

---

1  Associate Professor in English Linguistics, Université de Lille (Faculté Langues, Littératures et Civilisation Étrangères, Département des Études Anglophones), bert.cappelle@univ-lille.fr.
ABSTRACT:

This paper highlights some facets of motion typology, applied here to mainly English and French. These two languages are not perfect examples of satellite-framed and verb-framed languages, in Leonard Talmy’s well-known typology, but they can nonetheless be shown to differ in a number of related respects: compared to English (and other Germanic languages), French (like other Romance languages) is quite constrained in its use of Manner-of-motion verbs. French also lacks true particles – Path satellites without a Ground that can be syntactically detached from the verb. Drawing on some of my previous research, I briefly discuss two simple but apparently sufficiently efficient corpus-based translation studies that reveal that these differences show up when we compare English texts originally written in English with English texts translated from French vs. English texts translated from German (or other Germanic languages). A third, more recent, study contrasts a single English novel with its French and Dutch translations, focusing on expressions of visual motion. Here, too, some of the basic encoding preferences (satellite-framed vs. verb-framed) that these languages exhibit for actual motion appear to apply, by and large, for visual motion. This paper also lists some precursors of Talmy, one of whom is famously linked with the linguistic relative hypothesis. It is suggested that French, because of its typological nature, may not urge its speakers to convey much detail (neither of Manner nor of Path) in the encoding of motion. It remains an open question, one that goes beyond the purview of corpus linguistics, whether this stylistic difference is matched with a deeper cognitive one.

KEYWORDS: Motion events; lexicalization patterns; Manner; Path; satellite; visual motion; Romance language; Germanic language; stylistics; corpus-based translation studies.

RESUMO:

Este artigo destaca algumas facetas da tipologia de movimento, aplicadas aqui principalmente ao Inglês e ao Francês. Essas duas línguas não são exemplos perfeitos de línguas de perfil satellite-framed e verb-framed, na conhecida tipologia de Leonard Talmy, mas, não obstante, pode-se mostrar que elas diferem em vários aspectos relacionados: em comparação com o Inglês (e outras línguas germânicas), o Francês (como outras línguas românicas) é bastante restrito no uso dos verbos de Modo de movimento. O Francês também carece de partículas verdadeiras – satélites do tipo Path/Percurso sem um Ground/Plano que podem ser sintaticamente desconectados do verbo. Com base em algumas das minhas pesquisas anteriores, discuto brevemente dois estudos de tradução baseados em corpus que, simples, mas aparentemente suficientemente eficientes, revelam que essas diferenças aparecem quando comparamos textos originalmente escritos em Inglês com textos em Inglês traduzidos do Francês vs. textos em Inglês traduzidos do Alemão (ou outras línguas germânicas). Um terceiro estudo, mais recente, compara um romance Inglês com suas traduções em Francês e Holandês, com foco em expressões de movimento visual. Aqui, também, algumas das preferências básicas de codificação (satellite-framed vs. verb-framed) que essas línguas exibem para o movimento real parecem aplicar-se, geralmente, ao movimento visual. Este artigo também lista alguns precursores de Talmy, um dos quais é famoso por estar ligado à hipótese do relativismo linguístico. Sugere-se que o Francês, por causa de sua natureza tipológica, não pode incitar seus falantes a transmitir muitos detalhes (nem de Manner/Modo nem de Path/Percurso) na codificação do movimento. Permanece uma questão em aberto, que vai além do alcance da Linguística de Corpus, se essa diferença estilística é combinada com uma cognitiva mais profunda.

PALAVRAS-CHAVE: Eventos de movimento; padrões de lexicalização; Modo; Percurso; satélite; movimento visual; língua românica; língua germânica; estilística; estudos de tradução baseados em corpus.
1. Purpose and plan of the trip

All animal species are familiar with the sensation or experience – without necessarily being highly conscious of it – of moving through their spatial environment. Only humans, as far as we know, have the ability to talk about that basic kind of event. Not all humans speak the same language, and so we can expect speakers of different languages to talk differently about directed motion. My main aim is to show how corpus research can provide a fragment of an answer to the following question: If you speak a particular language, how does that language make you talk about directed motion?

A more interesting question, still, is this: If you speak a particular language, does that language make you think in a particular way about directed motion? This is really one of the all-time favourite questions in Cognitive Linguistics. Corpus linguists should admit they can’t say much about that. However, they can say something about it.

This paper therefore invites the reader interested in corpus-based translation studies and motion typology on a short tour, taking brief stops at three small-scale corpus studies that jointly form a partial answer to that question about cognition. These studies were rather simple in their setup and were easy to perform. Yet, for all their simplicity, they provided some quite clear results. The studies in question are ones that appeared in the course of roughly a decade, the first in 2012, the second (done with my colleague Rudy Loock) in 2017 and the third one in 2020. Other research, by other authors, obtained some similar results like the ones presented here. I would therefore like to make it clear from the outset that this trip is a bit like showing people some of the highlights of one’s own town; I hope the reader will understand that no implication is intended that there are no other tourist-worthy towns.

I will present the three studies very succinctly, as we barely have enough time in the span of this single paper to get off and on the coach, so to speak. So here’s the plan of our trip. In Section 2, I will explain why corpus linguistics can be useful, albeit only to a limited extent, to answer questions about how we think. Section 3 deals with some basic concepts used in motion event typology, in the Talmyan tradition, where a distinction between satellite-framed and verb-framed languages is made. I will, however, also discuss some forgotten scholars in this field and give credit to one very famous scholar whose work on motion events you never see quoted (Section 4). Section 5 will zoom in on the concept of ‘satellite’ and Section 6 will be devoted to showing that English and French, the two languages focused on in this paper, are good but not perfect examples of a satellite-framed and verb-framed language, respectively. After all

---

2 This article is a written-out and slightly revised version of a plenary talk given at the 8thCogLing Days, the biennial conference of the Belgium Netherlands Cognitive Linguistics Association, held on 13-14 December 2018 at the Université Catholique de Louvain in Louvain-la-Neuve, Belgium.

3 The waggle dance of bees, which could be adduced as a counter-example, is not a way of talking about an event of directed motion: bees don’t talk, or communicate in any other way, about who moves where in what way. Bees’ waggle dance is merely a complex way of pointing.

these preliminaries, we will finally get to the aforementioned case studies (Sections 7-9). I will then present my earlier study on manner-of-motion verbs (Section 7), particles (Section 8), and a recent one on visual motion expressions (Section 9). The tour ends with some concluding remarks (Section 10).

2. What corpus linguistics can and cannot do

The start of the tour is at a nice coffee house called Elizabeth’s in Vieux-Lille, the old town quarter of Lille in northern France. Before its recent extension, there was some lettering on the window that was meant to lure customers in. It said “Come in!”, followed by an asterisk. In smaller print, right under it, there was another asterisk, followed by “Entrez” (see Figure 1).

![Figure 1. Come in! and its translation into French](image)

Of course, providing a translation was a tongue-in-cheek gesture. Even in France, few local passers-by will be really puzzled by what the two English words *come in* together mean. Nonetheless, the English welcoming invitation and its unnecessary French translation provides us with an interesting opportunity to contrast the two languages. Indeed, not often does one see a juxtaposition of two such short expressions that capture a crucial difference between two languages – a difference that runs between Germanic and Romance languages more generally.

English uses two words to express an idea that can be expressed in one word in French. At first blush, you could conclude that French is more efficient than English. However, two elements can in principle achieve twice as much as one element. So, perhaps, French is missing something.

In this case, there is a small portion of extra meaning conveyed by the English version: by allowing the change of location to be expressed by a particle, speakers of English have not ‘used up’ their options of what to put in the verb position. As a result, they can toggle between *come* and *go*, not to mention many other possibilities. “Come in!” states explicitly that the invited movement is towards the sender of the message. “Entrez!” leaves this information to be deduced from the situational context. Hence the question: Is there any linguistic evidence to suggest that speakers of French tend to omit some information compared to speakers of
English? Or, from the point of view of hearers or readers: Do you miss something when you hear or read French, something that you do not miss when you hear or read English?

To determine this, we should collect more data than just what we find on a single windowpane. An obvious way is to ask a number of speakers of English and speakers of French to describe the same motion events. This is done in a recent study by Hickmann et al. (2017), among many other similar studies (e.g., Berman and Slobin 1994; Strömqvist and Verhoeven 2004; Fortis et al. 2011; Fagard et al. 2013). However, I will show that running a quite simple corpus ‘experiment’ can also go a long way in answering that question.

Using corpora may even answer the more fundamental question mentioned in Section 1: Do speakers of English and French think about motion along different lines? Sure, the best way to address this question is by psycholinguistic means. The study by Hickmann et al. (2017), mentioned above, does exactly that. It ran categorization tasks whereby participants were first presented with a stimulus involving an animal moving in a particular way in a particular direction. They then had to judge as most similar to such a target either a cartoon in which the animal moves in the same manner but in a different direction, or a cartoon in which it moves in the same direction but in a different manner. So why use corpora instead? Corpora have the advantage that they allow us to observe language use ‘in the wild’ rather than in test conditions that can sometimes be artificial, like deciding which kind of event is more similar than which other event. Large portions of naturally produced text can reveal cross-linguistic differences in words and constructions people do and do not use, or more precisely, cross-linguistic differences in how often speakers use them. Therefore, corpora can also give us a glimpse into what is at the centre of speakers’ minds and what is less so.

 Obviously, when we, as corpus linguists, want to give an answer to the question of whether different languages make their speakers think differently about directed motion, we have to tread with care. Corpus findings cannot directly demonstrate any deep cognitive differences among speakers of structurally divergent languages. Different usage patterns may not show that speakers of French and speakers of English differ in their not-specifically-linguistic thinking about motion events. I would love to find a corpus-linguistic method that we can use to prove that speakers of different languages really conceptualize, categorize, remember or even perceive directed motion events differently. I’m afraid, though, that such a method is not within reach.

Can’t corpus linguists say anything about a hypothetical effect of language on thought? In fact, they can, even though their claims will have to remain modest. Corpus findings can reveal differences between languages in which and how many aspects of directed motion events language users feel compelled to put into words. Different usage patterns across languages can only indicate differences in the cognitive processes people tap into when they use language. I am not entirely sure at this moment whether such claims actually avoid the trap of triviality. It’s quite obvious that speakers of French take recourse to French words, expression and constructions when they use French! These French-language units don’t need to be taken as directly reflecting these speakers’ non-linguistic abilities to think about motion. Still, I sometimes like to think of a large corpus
of literary texts as one massive production experiment, in which language users, without really
knowing they participate in an experiment, are being observed. Novelists across different languages
have to describe how their characters walk, how they talk, what they see, and so on. This is true for
writers of French novels just as it is true for writers of English novels.

When we look at a corpus as the output of a large-scale production task, then differences
across corpora may serve to suggest the existence of cross-linguistic differences in what Dan
Slobin calls ‘thinking for speaking’. By this term Slobin refers to the hypothesis that there is “a
special form of thought that is mobilized for communication” and that “involves picking those
characteristics of objects and events that (a) fit some conceptualization of the event, and (b) are
readily encodable in the language” (Slobin 1996: 76).

3. Talmyan lexicalization patterns: satellite framing vs. verb framing

The language we speak nudges us into certain encoding choices. If we are a speaker of
English or Dutch and we want to talk about someone or something moving along a trajectory,
or ‘Path’, then we often rely on a so-called ‘satellite’ to the verb to express that Path. If we are
a speaker of French or Spanish, we typically express Path in the root of the main verb itself.
The notion of ‘satellite’ was famously launched by Leonard Talmy and deployed by him in a
series of highly influential publications leading up to and including his two-volume book on
may not be fully accurate: Talmy is often credited with the term ‘satellite’ used in a linguistic
sense, but he may have borrowed it from other authors (cf. Fortis 2010: 2, fn. 3). The term is
not understood and defined consistently in the literature, not even by Talmy himself. Talmy has
more recently defined a satellite as

“a constituent in construction with the main verb (root) and syntactically
subordinate to it as a dependent to a head.” (Talmy 2009: 390)

Although Talmy writes “syntactically subordinate” here, it is clear from his earlier work
that he also allows satellites to be morphologically dependent elements, such as prefixes in
German and Dutch. Thus, a satellite is an entity that is somehow combined with the main verb,
or its root (the part without any affixes), as a kind of complement. For example, the English
sentence we see here expresses the Path completely outside the verb, by using the particle in or
the prepositional phrase (PP) in the room or into the room:

(1) The man walked in((to) the room).
    [Figure] [motion + Manner] [Path [Ground]]
Satellite

We can just use in to express the Path, but the Path may also contain an explicit mention
of the so-called ‘Ground’, that is, the reference object with respect to which the Subject referent
is displaced. ‘Figure’ is the term used for a moving entity, in this case, the man. In Section 5, I
will make some further comments about the term ‘satellite’, justify the nested bracketing under the PP in (1) and touch on the variation between in the room and into the room.

Now, the equivalent French sentence encodes the Path in the main verb (entra), either exclusively so or in conjunction with the PP dans la pièce:

(2) L’homme entra (dans la pièce (‘en marchant)).

Main verb

the.man entered in the room by walking

English and French could be said to exhibit different ‘lexicalization patterns’, which is the term by which Talmy refers to broad regularities in how the conceptual components that make up the description of a situation characteristically find expression in the lexical units of a given language (Talmy 1985). These lexical units can be words, morphemes, or multi-word expressions, such as à pied (‘on foot’), which could also have been used in our French example (but which, like en marchant, might more likely not be used, as we’ll see shortly). In a recent publication on lexicalization patterns, Beth Levin and Malka Rappaport Hovav (2019) state that

[4]cross languages, clauses containing descriptions of similar events are likely to include the same conceptual components, but these may be distributed differently across the constituents of the clause. (Levin and Rappaport Hovav 2019: 1)

“Likely to include the same conceptual components”: really? Well, Levin and Rappaport Hovav also write that “manner is only optionally expressed and often omitted in Romance languages” (Levin and Rappaport Hovav 2019: 7).

If we compare the English sentence and the corresponding French sentence, we see that we can leave some components unexpressed and that leaving out Manner certainly does not harm the French sentence in any way.

(3) a. English:
The man walked in.

b. French:
L’homme entra.

Indeed, since walking is the default way of entering a room, it would be odd to add this information at the end of the clause, where it would probably receive focal emphasis, rather unnecessarily so. The use of such a manner component, therefore, while not exactly ungrammatical, is a marked choice in French.

4 The page number is to the online document mentioned in the references.
5 Cf. ibid. for further references on manner omission. In the same publication, the authors write: “If the verb lexicalizes the path, (…) the manner of motion may be expressed outside the verb in a PP or adverbial phrase [or a gerund, as in the French example shown – B.C.], but may also be omitted since the verb itself does not require its expression” (Levin and Rappaport Hovav 2019: 5).
In the typological literature based on Talmy’s work, it is common to refer to English and other Germanic languages as *Satellite-framed languages*, because in these languages, the Path is lexicalized in a satellite. Alternatively, they’re called *Manner languages*, because the main verb can lexicalize Manner. French and other Romance languages are called *Verb-framed languages* or *Path languages*, as the Path is lexicalized in the main verb, and Manner is expressed outside of the main verb, or not at all.6

Dividing all of the world’s languages in just two families is too simplistic. To account for the full range of lexicalization patterns found across the world, linguists have meanwhile argued we need at least a third type, so-called ‘equipollently-framed languages’ (Slobin 2004). Talmyan and even ‘post-Talmyan’ motion event typology is, now perhaps more than ever, a very lively field of research. For recent collections of studies, see Ibarretxe-Antuñano 2017, Matsumoto and Kawachi 2020.

4. Precursors to Talmy

Typologists may not always agree with Talmy’s work, but they cannot *not* mention it. Talmy’s work is just too groundbreaking for that. The thing is, though, that Talmy was not the first to observe that languages can diverge in how (or how readily) they encode Path and Manner of motion. A Danish linguist, Casper Woldersgaard, recently dug up a passage from a book on the style of the French language, written by one Fritz Strohmeyer, published over a hundred years ago, in 1910. This is my translation of the original German:

> The French aversion to specialization of concepts is particularly evident in its expressions of movement. As is well known, *aller* means everything: ‘go’, ‘drive’, ‘ride’, ‘fly’, ‘climb’, ‘sink’, and so on. (...) This is what has now more or less happened to the other verbs of movement, especially to *entrer*, (...), *sortir*, (...), and so on. Whether there is talk of a train pulling in and pulling out, a ship floating in and floating out, a bee flying into the hive and flying back out, a miner climbing into the shaft and climbing back out, a spring-loaded machine part snapping into its bearing and snapping back out, the Frenchman settles in all cases for a simple *entrer* and *sortir*. (Strohmeyer 1910: 191–192, quoted in Woldersgaard 2017: 14; my own translation from German)

This citation by Strohmeyer may (so far) be the earliest known instance antedating Talmy’s work. But it is not the only known precursor to Talmy. Other scholars made similar comments about differences between French and German or between French and English. Below is an overview:7

---

6 We also find shortenings and abbreviations in the literature: S-framed languages, satellite languages or just S-languages; likewise for verb-framed languages. The terminology goes back to Talmy (1975, 1985, 1991, 2000).

7 This list has been compiled from partial mentions in Berthele 2007: 66; Levin, Beavers and Tham 2009: 341; Fortis 2010: 6; Fagard, Stosic and Cerruti 2017: 638, 640; Woldersgaard 2017: 13–15; Levin
- Bally, Charles. 1932. *Linguistique générale et linguistique française*.
- Bergh, Lars. 1939. L’idée de direction exprimée par un adverbe ou par une préposition en suédois, par un verbe et une préposition en français; cf. also his 1948 PhD thesis.
- Tesnière, Lucien. 1959. *Éléments de syntaxe structurale*.
- Wandruszka, Mario. 1969. *Sprachen: Vergleichbar und unvergleichlich* (and specifically his discussion on pp. 460–469)

Among the pre-Talmyan works, we also find a publication that our generation has apparently almost forgotten (exceptions being Oliveira 2012: 37 and Levin 2015: 1). This oversight is all the more surprising because the publication was written by a linguist whose name has become inextricably linked with a hypothesis that is still hotly debated, also in the field of motion event typology. That linguist is none other than Edward Sapir, of the much-discussed Sapir-Whorf hypothesis. It is as if biologists studying variation in beaks of birds would be wondering what that variation might tell us about evolution while being blissfully unaware that Darwin himself had in fact written some quite interesting things about beaks of finches.

In 1932, Sapir (co-)wrote a study titled ‘The expression of the ending-point relation in English, French, and German’. Google Scholar has indexed a meagre 23 citations to this publication.8 For comparison, Talmy’s 1985 paper on lexicalization patterns has some 5,000 citations to it. I think it’s only fair to cite Sapir (and his co-authors):

The expressions of action and the direction or ending-point of action are frequently crowded into the French verb, while in English and German the tendency is rather to use a verb with a broad meaning and a preposition or adverb to complete the idea of direction or ending-point. Thus, Fr. ‘*accoster le navire*’ = Eng. ‘*to come alongside the ship*’, Ger. ‘*sich langseit dem Dampfer legen*’; Fr. *descendre* = Eng. *to come down*, Ger. *Herunterkommen*. (Sapir et al. 1932: 21–22)

The image of elements being “crowded into the French verb” foreshadows Talmy’s concept of *conflation* – the simultaneous lexicalization of two conceptual components in a single lexical item (Talmy 1972: 257).

Unfortunately, Sapir does not seem to reflect in this study on the question of whether

---

8 At the time of writing the present paper. As Sapir’s writings are also available in the form of collected works, the actual number of publications referring to Sapir et al. (1932) is probably quite a bit higher. The point remains that work in motion typology does not typically refer to this publication, as far as I can see, and yet often discusses findings in the light of the ‘Sapir-Whorf’ hypothesis of linguistic relativity.
Frenchmen, because of this frequent conflation imposed by their grammar, live in a mental world which differs from that of Englishmen or Germans. It would have been nice if we had found a passage in which Sapir wondered whether the French language limits its speakers’ ability to get their minds around the fact that a directed motion event consists of both a direction or goal and the action of motion itself. Or, for that matter, it’s a pity Sapir can’t be caught wondering whether the English or German language prevents its speakers from grasping the fact that coming and goal form inseparable aspects of one single event in reality. Then again, Sapir was never really clearly in favor of the Sapir-Whorf hypothesis partly named after him (which, therefore, is more correctly referred to as the Whorf hypothesis).

The attentive reader may have noticed how Sapir in this fragment does not say anything about Manner of motion. Sapir here claims that both English and German make frequent use of a general verb, like *come* and *kommen*. This is in sharp contrast with what we are accustomed to find in other works before and after Talmy’s main writings, for instance in Malblanc’s (1944) book on comparative stylistics of French and German:

> Whether poverty or miracle, a single small sign ‘passer’ can correspond to some fifty German verbs. That is because ‘passer’ nuances or changes its meaning depending on the other words with which it is combined. Very often, in French, language elements only have their precise value through their context, while in German, words much more often carry their meaning by themselves. (Malblanc 1944: 14; my own translation from French)

Perhaps the best-known forerunner of Talmy’s work is Vinay and Darbelnet’s book on French-English translation ‘stylistics’, published in 1958. These authors are the first to demonstrate a common translation trick whose name must ring a familiar bell with many students of English-French translation: the *chassé-croisé* or ‘interchange’ (to use the term in the English edition of their book). Here’s an example of theirs:

(4) blown away

![Diagram](image)

emporté par le vent
taken away by the wind

---

9 The term *chassé-croisé* has been adopted in other popular handbooks of translation used in France (Ballard 1980: 151–153; Chuquet and Paillard 1989: 13–14; cf. also Chuquet and Paillard 2017: 30). When you apply a *chassé-croisé*, you translate two elements in the source language clause in such a way that they change grammatical categories and swap positions – it is a double application of a technique known as transposition (rendering a source language element by a target language element that belongs to a different category), and a permutation at that.

Discourse seems to matter a lot. If we know from the surrounding discourse that the subject referent got into car, then a sentence such as

(5) They drove onto the scene of the accident.

can simply be rendered as

(6) Ils arrivèrent sur les lieux de l’accident.

‘They arrived at the scene of the accident.’

The authors comment:

The context tells us that “they” were in a car. French does not feel the need to restate it. This leads to a loss of information (…). Taken on its own, the French sentence says less than the English sentence. But it would be against the nature of the French language to go into such detail because French prefers the level of abstract expression. (Vinay and Darbelnet 1995 [1958]: 51)

Do the authors imply a value judgement here? It is hard to say. On the one hand, they objectively state a difference in syntactic structure:

the English sentence is organised around a concrete word whereas the French sentence is organised around an abstract word. (Vinay and Darbelnet 1995 [1958]: 51)

On the other hand, Vinay and Darbelnet don’t use such neutral phrasing when they compare the entire lexicons of French and English:

Generally, it can be said that French words function at a higher degree of abstraction than the corresponding English words. They tend to be less cluttered with details of reality. (Vinay and Darbelnet 1995 [1958]: 52)\(^\text{10}\)

This sounds as if the authors do not really value manner information. According to them, French words capture the essence of a situation and don’t force language users to fuss about aspects that are clearly, as far as they are concerned, of secondary importance. This attitude contrasts to what French stylists typically pride themselves on, namely the ability to always find le mot juste, that is, select the exactly right word or expression. Curiously, providing manner information isn’t a matter of using the perfect phrasing in French! On the contrary, it would distract from the essence of the message.

\(^\text{10}\) In the original French, it was formulated as follows: “Ils s’embarrassent moins des détails de la réalité” (Vinay and Darbelnet 1958: 59), which means something like ‘They [i.e. the French words] are less worried about the details of reality’.
5. Some notes on Path, satellite, and location understood as direction

Some brief comments about the nested bracketing under the PP in (1) and (2), about the term ‘satellite’, and about the variation between in the room and into the room may be in order.

My bracketed notation in (1) and (2) deviates from the practice of treating just the particle or the preposition as encoding Path. As I see it, the conceptual component Path denotes a trajectory, not just a relation between the Figure and the Ground. One cannot think of a Path independently of a Ground. Therefore, if a Ground is explicitly mentioned in the clause, this must necessarily be considered as part of the Path, not something that exists alongside it.

In earlier publications, Talmy excluded PPs from what can be comprised under the label ‘satellite’, which he defined as “the grammatical category of any constituent other than a nominal or prepositional-phrase complement that is in a sister relation to the verb root” (Talmy 2000: 102). The exclusion of PPs met with opposition, as these fulfill much the same function in a motion expression as verbal particles (e.g., Beavers 2008: 286, fn. 3; Beavers, Levin, and Tham 2010: 337–339; Croft et al. 2010: 205–206, Imbert 2012: 240–241). I see no principled reason to treat particles but not PPs as satellites: they can both express the motion’s Path. However, particles should not be seen as ‘reduced’ prepositional phrases, nor do particles and their full-PP counterparts necessarily license the same interpretation, even if the unexpressed Ground when a particle is used is understood to be the same as the complement of the preposition (Cappelle 2005; Cappelle and Declerck 2005). For instance, He jumped in can only be understood as an achievement (i.e., an essentially non-durative event), describing a change of state from being, say, out of the water of a swimming pool to being in the water, if the context makes it clear that a swimming pool is being talked about. He jumped in the pool, with a full PP, can describe the same situation but could also refer to an activity, namely that of jumping up and down (for some reason) while in the water. On that reading (which is admittedly much less likely to arrive at), the situation is similar, in aspectual terms, to He jumped on the trampoline for an hour, in that it does not involve a change of state. This difference between particles and full PPs may tie in with the fact that some full PPs are adjuncts rather than complements (Mateu 2012).

The PP in the room, just like its French counterpart dans la pièce, is strictly speaking an expression of location, not direction. It is felt to encode Path by virtue of a pragmatic inferencing process: in a context in which we know or suspect that directed motion is involved, we treat such a location as the goal of motion which metonymically stands for the whole Path (as it is the most salient portion of that Path). That in the room nonetheless semantically only expresses a location in He walked in the room is clear from the fact that a non-boundary-crossing reading is still (marginally) possible. We more easily get such an interpretation in, for instance, We walked in the woods for an hour, which describes an activity. The PP into the room, by contrast, makes explicit the idea of a transition – it really means something like ‘to [Location in the room]’ (cf., among others, Jespersen 1961: 8–9; Declerck 1977: 312) – and thus rules out an interpretation in which the movement is not a boundary-crossing one. For discussion of the distribution of directional in and into, see Nikitina (2008), who lists a variety of factors that help hearers or readers to give a directional interpretation to a locative PP.
6. English and French: not such perfect examples of satellite- and verb-framed languages

Does French really more often use abstract words in descriptions of directed motion than English? That is, to some extent, still an empirical question. In any case, there are reasons to say that French and English are not that different in nature.

In spite of what was stated in Section 3, Romance languages are not all perfectly verb-framed. Some of them may even use phrasal verbs, as should be clear to everyone who realizes that the Italian loanword tiramisu literally means ‘pull me up’ or, more freely, ‘cheer me up’. Romance languages also have many prefixed verbs. Thus, in French we find such verbs as ac-courir (‘hasten (somewhere)’), dé-crocher (‘unhook’), re-venir (‘come back’), s’en-dormir (‘fall asleep’), sur-voler (‘fly over (something)’), é-crémer (‘skim’), and so on. Such prefixed verbs have an incorporated Path satellite (Pourcel and Kopecka 2006).

Furthermore, French does have many Manner(-of-motion) verbs and it is often perfectly fine to use them even as main verb (rather than, as in (2), in a gerund added to a Manner-neutral main verb). There are some conditions, though:

1. The motion event should be non-boundary-crossing or non-boundary-reaching
2. The Ground should be expressed explicitly

Here’s an example:

(7) Des hommes titubaient autour de l’ancien hôpital de campagne, hébétés, couverts de poussière grise.11
   ‘Men staggered around the old country hospital, dazed, covered with gray dust.’

Even in boundary-reaching events and occasionally boundary-crossing events, it is possible to use a Manner-of-motion verb, when the expression of Manner is somehow highly relevant in the discourse:

(8) Il était en retard et il a couru jusqu’à l’école.12
   ‘He was late and he ran all the way to school.’

(9) Il a sauté dans mes bras du 3ième étage.13
   ‘He jumped in(to) my arms from the third floor.’

With regard to (9), it would even be odd to use the typical lexicalization pattern for French here, shown in (10), perhaps in part also because one’s arms don’t form a real container that one can go into:

Il est entré dans mes bras en sautant du 3ième étage.

‘He entered my arms jumping from the third floor.’

In the ‘constructicon’ of French speakers, sauter dans les bras de quelqu’un ‘jump in(to) someone’s arms’ may well be stored as a semi-preformed expression (allowing some modification for tense and with respect to whose arms are involved).

Finally, French has many Manner-expressing reflexive motion verbs that can be found in (sometimes figurative) boundary-crossing contexts: s’enfiler ‘to slip, to sneak’, s’enfoncer ‘to sink’, s’infiltrer ‘to infiltrate’, se couler ‘to slip, to glide’, se faufiler ‘to sneak, to edge one’s way (into…), to thread one’s way (among… / between…’), se glisser ‘to slip’, se jeter ‘to throw oneself, to jump’, se lancer ‘to rush, to launch, to embark’, se plonger ‘to plunger, to get absorbed (e.g. in a book)’, se précipiter ‘to rush’, etc. For example:

Il se faufila sous les draps.

‘He snuck under the sheets.’

Elle se lança dans la politique.

‘She went into politics.’

Apart from the above points, it’s true what typologists claim about French, namely that it tends to use Path verbs, which do not express Manner (entrer ‘come/go in’, sortir ‘come/go out’, monter ‘come/go up’, descendre ‘come/go down’, traverser ‘cross’, etc.).

Turning to English, just like French is not perfectly verb-framed, as we’ve just seen, this Germanic language is, for its part, not perfectly satellite-framed. It’s not just that there are also path verbs available, such as ascend and descend, but some path verbs in English are actually default choices and not (as could be wrongly assumed) lofty alternatives to Germanic structures. Consider for example this sentence, which sounds perfectly colloquial:

The girls arrived late and left early.

To sum up, French and English are definitely not perfect examples of the types of languages they are generally said to belong to – although, in truth, some authors have also provided a more nuanced picture (cf., inter alia, Martinez-Vasguez 2013; Fagard et al. 2013, 2017).

Still, I wouldn’t want to go as far as to say that French is not verb-framed or that English is not satellite-framed. To get a clearer picture about how these languages behave typologically, let’s see if we can detect traces of these languages’ typological nature in corpora. This is the topic of the next three sections, in which I will represent three small corpus studies. We’ll begin with looking at how often we find Manner-of-motion verbs in French compared to English.
7. A corpus study on Manner-of-motion verbs

Remember that one of my aims in this paper is to demonstrate how corpus research can be easy and still produce convincing results. In particular, I want to show that, if you don’t know a language very well there, there is a trick to ‘avoid’ studying that language from close-by. That is what I did for French in a paper on Manner-of-motion verbs (Cappelle 2012), for which I used a method that is well known in corpus-based translation studies. I compared originally produced English and English translated from French. Translations inevitably suffer from shining-through effects. No matter how hard translators do their best to make their products conform to target-language norms, there will always be traces of source text properties. Researchers can capitalize on that. If French is really verb-framed, it should use fewer Manner-of-motion verbs. We should therefore also expect translations from French to contain fewer Manner-of-motion verbs.

As a reference corpus of English, I used the fiction and biography components of the British National Corpus. This reference corpus consisted of about 20 million words. I then used the Translational English Corpus and also selected the subcorpora of fiction and biography. I took a corpus of English texts translated from French and a similar corpus with English texts translated from German. This way, we can compare the impact of having a Romance versus another Germanic language as source language. The size of both translational corpora was in the order of half a million words.

I then used Beth Levin’s (1993) book of English verb classes to select Manner-of-motion verbs and Path verbs. From that very useful resource, I selected 237 Manner-of-motion verbs. This set included such verbs as crawl, drift, leap and trot; and also transitive ones like dump, ladle, squeeze and toss. There are of course fewer motion verbs that just express some sort of Path, but still I managed to select 36 of them. This smaller list included, obviously, the verbs come and go, but also verbs such as cross and leave, and transitive ones such as lower, place, situate and take.

The results produced by looking for these verbs in the distinct corpora were very clear. As shown in Figure 2, in all three corpora, Manner-less motion verbs were in the majority. So, maybe Sapir was right in claiming that English uses a lot of general motion verbs. What’s more interesting was the difference in ratio of Manner-of-motion verbs to Manner-less motion verbs. Originally produced English (the material in the BNC, shown at the top) and English translated from German (the bar below it) displayed about 30% of Manner-of-motion verbs. There was hardly any difference between what we found in these two corpora. The lowest bar looks different: English translated from French only has about half that percentage of many Manner-of-motion verbs.
Figure 2. Manner-of-motion verbs and Manner-neutral (‘manner-less’) motion verbs in three varieties of English: originally produced English (top), English translated from German (middle) and English translated from French (bottom)

We can only attribute this difference to what was apparently present, or ‘less present’, in the source text. The French source texts must have had fewer Manner-of-motion verbs. This confirms our expectations. Again, French does have Manner-of-Motion verbs, but it cannot use them in the same range of conditions as in English. In resultative (namely, boundary-crossing/boundary-reaching) constructions, we don’t often find them. That is why they are also underrepresented in the translations, even though the translator probably added some Manner information here and there. This, in any case, is what Dan Slobin found for translations from Spanish to English:

English translators … add manner to the Spanish original in almost a quarter of their translations. (Slobin 1996:112)

If such additions of Manner also happened from French to English, they clearly weren’t sufficient.

8. A corpus study on particles

Moving on to the satellite in motion expressions, we can focus look at how often particles are used in original English and English translated from typologically different languages. This was the idea of a study carried out with a colleague of mine (Cappelle and Loock 2017). Before going into that study, a brief discussion about the grammatical status of this class of words may be useful.
8.1 How we should analyse English particles: a quick roundup of facts

I just referred to particles as a “class of words”, but the reality is that linguists disagree on whether they are really words. There is huge debate on that topic, and the following elements seem to be crucial in this.

First, in some constructions, the verb particle seems to behave as a morphological part of the verb. Take the ‘quotative inversion construction’ (e.g. McIntyre 2003):

(13) “Get lost,” shouted out Gertrude.

In this construction, what can appear between the quote and the Subject is normally just one word, namely the verb. For example:

(14) “Get lost,” exclaimed Gertrude.

In the position wedged between quote and Subject, we cannot find combinations of a verb and anything else, such as a verb and a prepositional phrase. What we cannot say, for instance, is (15):

(15) *“Get lost,” shouted at him Gertrude.

That is why we can conclude that shouted out in (13) forms a word, not a combination of two words, despite the inflection not appearing at the end of that word.

Second, in the closely related language Dutch, we also find particle verbs in positions where you can only have verbs without anything else (Cappelle 2012: 279):

(16) We zijn een plaat aan het maken/op-nemen.
    we are an album at the make/up-take
    ‘We are making/recording an album.’

After aan het, we cannot find phrasal material next to the verb, not even if that material forms a lexical unit with the verb. For instance, a speaker of Dutch cannot say this:

(17) *We zijn een plaat aan het op de markt brengen.
    we are an album at the on the market bringing
    ‘We’re releasing an album.’

This suggests again that particle verbs are morphologically complex units, not structures that are assembled by phrasal syntax.

On the other hand, there is also irrefutable evidence that particles can be words separate from the verb. For instance, they can be the head of a phrase, which is something only words, not morphemes can be:

(18) The balloon rose [all the way up to the sky].

Moreover, particles can be separated from the verb by a direct object:

(19) He cleaned his nose out.
When we find a particle in post-object position, following a full (non-pronominal) NP, it clearly doesn’t behave like an affix to the verb.

The most sensible conclusion we can make is that particles sometimes look like just a part of the verb and sometimes like words in their own right. Besides, their grammatical analysis has nothing to do with their meaning, as both idiomatic and semantically compositional verb-particle combinations can occur in both the joined (V Prt NP) and the split (V NP Prt) construction. Of course, a particle verb can’t at the same time be a word and a multi-word unit. That’s why it may be useful to see its variable manifestations as ‘allostructions’ of an underspecified construction (Cappelle 2006). That more general construction doesn’t specify about itself whether it’s a word or a combination of words.

Whatever their grammatical status, frequently used particle verbs are processed by the brain as lexical units (Cappelle et al. 2010, Hanna et al. 2017). For instance, when subjects hear up in an existing combination like heat up, then that particle gives a stronger spark in their brain than when they hear it in a pseudo-phrasal verb like cool up. This is the typical activity shown by a speaker’s brain when it comes across a word it ‘knows’. Since phrasal verbs are not necessarily words, what the brain activity shows is that known combinations trigger the response people get when they hear a lexical unit they can retrieve from their lexicon rather than a combination they still have to put together as they form a sentence. This heightened activation can also be found with combinations that are transparent, like rise up or fall down. And we also find it with particles that are separated from the verb.

Another indication that particle verbs are lexical units is that they usually have one-word translation equivalents in French:

(20) a. come in = entrer
b. get out = sortir
c. give up = abandonner

These translations are obviously stored as lexical units; by consequence, the English combinations are probably also stored as lexical units.

8.2 Particles in translation

We’ve just seen that particle verbs can often be translated by a one-word French counterpart. Some combinations, though, as we already know, call for a full chassé-croisé:

(21) a. Don’t worry, it’ll wash out
    = … ça partira au lavage
    ‘… that will leave in the washing’
    (Chuquet and Paillard 1989: 14)

b. He talked on.
    = Il poursuivait son discours.
    ‘he pursued his discourse’
Talmy noted that particles encode the upshot of the event, what it all comes down to, no matter whether they express goal or direction, or something more abstract, like the aspectual ‘contouring’ of another event. They thoroughly fit a satellite-framed lexicalization pattern.

Does French have particles? In a sense, it does. French has exclamations like Dehors! (‘Out!’) and Dedans! (‘In!’). But these ‘particles’ are seldom used with verbs. If French has particles, it doesn’t have particle verbs.¹⁴

No surprise then, that if we look at translations into English, we find the impact of French not having much of a satellite-framed arsenal of expressions. With my colleague Rudy Loock, I used the same procedure as before: comparing originally produced English with English translated from Germanic languages and from Romance languages (Cappelle and Loock 2017). We used a somewhat larger set of languages for this study, but most source texts were still French. As expected, the number of times we found the particles up, down and out in originally produced English was not very different from the number of times we can find these particles in English translated from other Germanic languages, which have similar constructions (Figure 3). However, in English translated from Romance languages, particles were underrepresented, which would seem to mean that the translators had not introduced a sufficiently high number of particles.

Figure 3. Normalized frequency of up, out and down in three varieties of English: originally produced English (top), English translated from Germanic languages (middle) and English translated from Romance languages (bottom)

Clearly, the translations from Romance languages are not completely lacking in particles either. This begs the question: What was in the source text that was produced by particles?

¹⁴ We can safely ignore exceptions like tu viens avec? (‘Are you coming with us?’).
Remember that French has prefixed verbs, such as *revenir* and *survoler*. We suspected that these were source items that lent themselves well to being translated by a particle verb. After all, prefixed verbs are a bit like particle verbs. A master’s student looked up all the verbs in *Le Petit Prince*, as a source text. The verbs were grouped into morphologically simplex ones, like *sortir*, and verbs that have a recognizable prefix, such as *revenir*.

Much to our satisfaction, we found our hunch confirmed. Prefixed French verbs in the source text triggered a particle verb as translation more readily than did non-prefixed French source verbs (Figures 4-5).

**Figures 4-5.** Particle verbs (blue slices) used in English translated from French; top (Figure 4): particle verbs translated from morphologically simplex verbs; bottom (Figure 5): particle verbs translated from prefixed verbs

![Pie chart](image)

This supports the idea that language users may still think of French prefixes as satellite-like elements or at least, that translators do. For instance, *réfaire* was translated as *do over* and *s’enfermer* as *shut oneself up* (not as *shut oneself in*, though). The correspondence isn’t always perfect, but still, it looks as though verbs consisting of two parts were also translated by two
elements.

9. A corpus study on visual motion

A lot of work has been done lately on what is called fictive or non-actual motion. One kind of fictive motion is visual motion (Slobin 2009; see Matsumoto et al. 2021 for a recent study). Talmy has devoted quite some attention to fictive motion, but Sapir again preceded him by calling “directed sight, a type of movement with ending-point” (Sapir et al. 1932: 39). Visual motion is fictive because we imagine a gaze emanating from the eyes of the person who is looking at something. Of course, the gaze doesn’t really exist as such – there’s no laser-like beam shooting out of our eyes.

For a study on visual motion, I used a children’s novel by Roald Dahl, which was chockful of events of looking, glancing, glaring, peeking, peering, and so on (Cappelle 2020). The book in question is The Witches, and in fact, in that book, the gaze of one of the witches, the Grand High Witch, is something very real and tangible, as in this sentence:

(22) “A moment later, a stream of sparks that looked like tiny white-hot metal-filings came shooting out of The Grand High Witch’s eyes and flew straight towards the one who had dared to speak.”

(Roald Dahl, The Witches)

I looked up all the instances in the book of visual motion, in the original English book and in the French and the Dutch translation. For the Dutch translation, I found full confirmation of the hypothesis that for the expression of visual motion, Dutch draws on the same encoding patterns as for actual (physical) motion. Thus, for encoding actual motion, Dutch can make use, among other construction, of a pattern with a (so-called) postposition (cf. (23a)); it could be found that this construction also lends itself to encoding visual motion (cf. (23b)):

(23) a. Ik liep een lange brede gang door.
   I ran a long wide corridor through
   ‘I went down a long wide corridor.’

b. Ze keek de lange, lege hotelgang af.
   She looked the long, empty hotel-corridor off
   ‘She looked up and down the long empty hotel corridor.’

(Roald Dahl, The Witches)

Next, I looked at what happened in the French translation. Matsumoto (2001) and Slobin (2009) previously assumed that perhaps for the expression of visual motion, French would lose its essentially verb-framed nature. They had a good reason for thinking that. To quote Slobin:

[V]erb-framed languages do not provide specialized verbs for visual paths, on a par with ‘enter’, ‘ascend’ and the like; rather, both types of languages [satellite- and verb-framed ones] rely on all-purpose perception verbs such as ‘look’, combined with various sorts of adjuncts (...). It may be, therefore,
that the marked differences between the two language types in the domain of physical motion fall away in the domain of visual motion (Slobin 2009: 201)

What I found, though, was that French appears to retain some of its verb-framing characteristics in the domain of visual motion, by taking frequent recourse to the transitive caused-motion construction. The entity that moves (i.e., the Figure) is then the looker’s eyes or gaze. In the following translations, French use Path verbs, as it typically would for physical motion:

(24) look up (source text)
lever les yeux (translation)
‘raise the eyes’

(25) look away (source text)
détacher le regard (translation)
‘detach the gaze’

Somewhat surprisingly, the French translation also made frequent use of Manner-rich motion expression, such as jeter un regard ‘throw a glance’ and fixer quelqu’un de ses yeux ‘fixate someone from/with one’s eyes’. A nice manner-rich visual motion expression is balayer quelqu’un de son regard, which literally means ‘sweep someone from/with one’s gaze’ and which is used when one scans someone from top to toe. Compared to the source text, though, the French translation was, on the whole, very poor in Path information (see Figure 6).

**Figure 6.** Frequency of having path information (either in the main verb or in a satellite) in Roald Dahl’s The Witches (top), its Dutch translation (middle) and its French translation (bottom)

![Path Information Frequency Graph](image)

Clearly, even though French can rely on verb-framed patterns of the type shown in (24) and (25), Path information was very often lost in the French translation. Sometimes, the
translator stuck to just the generic transitive looking verb regarder (‘look-at’)

(26) I saw The Grand High Witch peering down at the floor and staring with obvious puzzlement at William and Mary. (source text)
La Grandissime Sorcière regardait William et Mary, visiblement déconcertée. (translation)
‘the Grand Witch looked (at) William and Mary, visibly bewildered’

10. Envoi

In this paper, I have discussed some subtle but pervasive differences across some languages, focusing mainly on French and English. We have seen that French is not a purely verb-framed language; neither is English a purely satellite-framed one. Yet, French is still so different from English and other Germanic languages that English texts translated from French contain fewer Manner-of-motion verbs than originally produced English or English texts translated from German (cf. the first corpus study discussed). Furthermore, French (together with other Romance languages) also contain fewer Path elements – we specifically looked at particles – than either original English or than English translated from another Germanic language (cf. the second corpus study discussed). When we turned our attention to visual motion, we saw that French and English, and also Dutch, try to stick to their well-trodden typological paths (cf. the third corpus study discussed), in that French frequently adopts a verb-framing strategy (e.g. lever ses yeux ‘raise one’s eyes’), while English and Dutch predominantly use a satellite-framing strategy (e.g. stare up at something). The picture that emerges for visual motion is a little more complicated, though, because French unexpectedly uses many rather vivid Manner-of-looking expressions (e.g. jeter un regard ‘throw a glance’). Yet, when it comes to expressing Path information, the translation into French from English lost a lot of Path information, due apparently, to French using the Manner- and Path-less verb regarder ‘look’. This confirms the view of stylistic experts who stated that if you speak the French language, you are apparently not that “bothered” about the finer “details of reality” (cf. Section 4). Perhaps this also leaves an imprint on the minds of speakers of French, compared to speakers of English and other Germanic languages, although the results may not be suggestive of speakers’ conceptualization of motion beyond speakers’ thinking about how to put motion into words.

The ‘corpus studies’ presented here, you might say, are hardly worthy of that name. Indeed, for a part of the second study, I used a single short work of fiction (Le Petit Prince) and for the third one, I again used just one children’s novel (The Witches). The results I obtained from such potentially unrepresentative samples are in need of substantiation from larger-scale studies, involving also many more languages. Alternatively, many different studies of the sort discussed here, using similarly modest-sized datasets, could, when brought together, also go some way towards getting a more complete view of how languages differ in how they express actual and fictive motion. If this tour has given you some inspiration to go out and explore this
issue further using your own corpus of choice, I am happy to send you on your way.

REFERENCES


Beavers, John, Beth Levin and Shiao Wei Tham. 2010. The typology of motion expressions revisited. *Journal of Linguistics* 46,

Bergh, Lars. 1939. L’idée de direction exprimée par un adverbe ou par une préposition en suédois, par un verbe et une préposition en français. *Studia Neophilologica* 12, 66–90.


Cappelle, Bert and Renaat Declerck. 2005. Spatial and temporal boundedness in English motion

Roll up / Venez: An invitation to corpus-based research in motion typology

Bert Cappelle


Cappelle, Bert, Yury Shtyrov et Friedemann Pulvermüller. 2010. Heating up or cooling up the brain? MEG evidence that phrasal verbs are lexical units. Brain and Language 115(3), 189–201.


Martínez-Vázquez, Montserrat. 2013. Intralinguistic variation in the expression of motion events in English and Spanish. *Lingue e Linguaggi Lingue Linguaggi* 9, 143-156


Nikitina, Tatiana. 2008. Pragmatic factors and variation in the expression of spatial goals: The case of into vs. in. In Anna Asbury, Jakub Dotlačil, Berit Gehrke and Rick Nouwen (eds.),
Roll up / Venez: An invitation to corpus-based research in motion typology
Bert Cappelle


