

Fabienne Martin, Ingrid Falk, Maximilian Köper,
Sylvia Springorum, Sabine Schulte im Walde & Achim Stein

2019

Cross-Linguistic Interplay
of Lexical Aspect and (Non-)Literalness



In

A. Gattnar, R. Hörnig, M. Störzer & S. Featherston (Eds.)

Proceedings of Linguistic Evidence 2018: Experimental Data Drives Linguistic Theory

Tübingen: University of Tübingen

<https://publikationen.uni-tuebingen.de/xmlui/handle/10900/87132>

Cross-Linguistic Interplay of Lexical Aspect and (Non-)Literalness

Fabienne Martin, Ingrid Falk, Maximilian Köper, Sylvia Springorum,
Sabine Schulte im Walde & Achim Stein

Universität Stuttgart

fabienne.martin@hu-berlin.de

1 Introduction

Verbs often exhibit non-literal meanings, besides their literal senses. Recent literature has suggested interesting correlations between (non)-literalness and key dimensions of verb syntax and semantics. One tendency observed for various languages is that linguistic material that is optional when a verb is used literally may become compulsory for non-literal senses. (By contrast, the literature reports no examples of the opposite case, where linguistic material used under the literal sense needs to be deleted for the non-literal sense to obtain.) For French, Boons (1971) and Lamiroy (1987) notice that adjuncts sometimes lose their optionality when the verb is used non-literally, and Ruppenhofer (2006) makes the same observation for English. For instance, while the dative argument is optional under the literal reading of the French verb *sourire* ‘smile’, see (1a), Lamiroy (1987) observes that the same argument has to be syntactically realised under the non-literal reading in (1b).

- (1) a. Luc **sourit** (à Marie). FRENCH
Luc smiles at Marie
‘Luc is smiling (at Marie).’ *literal*
- b. Le destin **sourit** #(à Marie).
the fate smiles at Marie
‘Destiny is smiling at Marie.’ *non-literal*

For English, McNally & Spalek (2017) notice that non-literal uses sometimes require a particle that is optional in the literal use; e.g., their (2b) requires *off* in order for *cut* to be used non-literally, as a paraphrase of *stop*. By contrast, the non-literal sense is readily available in Spanish, which has no direct counterpart of English particles, see (2c). We observe similar contrasts for German (which uses verb particles like English) and French. For instance, *steigen* ‘raise’ can be used without particle in its literal use (3a), but not in its non-literal use (3b). By contrast, the non-literal sense is readily available in French (which has no particle, like Spanish), see (3c).

- (2) a. They **cut (off)** the head of the snake. *literal*
- b. I can **cut #(off)** the circulation in your feet, then they get really cold. *non-literal*
(GloWbE, McNally & Spalek 2017)

- c. Francia **corta** la entrega de etarras. SPANISH
 France cut-PST the extradition of ETA members
 ‘France cuts the extradition of ETA members.’ *non-literal*
 (McNally & Spalek 2017)
- (3) a. Der Ballon ist **(auf)gestiegen**. GERMAN
 the balloon is *auf*-raised
 ‘The balloon rose.’ *literal*
- b. In ihm **ist** Wut **#(auf)gestiegen**.
 in him is anger *auf*-raised
 ‘He started to feel angry.’ *non-literal*
- c. La colère **est montée** en lui. FRENCH
 The anger raise-PFV in him
 ‘Anger raised in him.’ *non-literal*

We formulate the generalization suggested by the data (1)-(3) as in (4). The name we attribute to this generalization was suggested to us by Lamiroy (1987), who claims that the shift from the literal to the non-literal use makes the verbal syntax ‘more rigid’.

(4) *Grammar rigidification under non-literalness:*

The set of (morpho-)syntactic frames a verb may enter in under its non-literal uses is a subset of the (morpho-)syntactic frames the same verb can instantiate under its literal uses.

A second and similar tendency has been observed in a key dimension of verb *semantics*. The observation is that non-literal extensions of verbal meaning regularly go along with a reduction in the number of aspectual frames a VP instantiates. As Vendler (1957) already observed, many predicates are ‘aspectually flexible’ in that they can instantiate more than one aspectual class. A first example is provided by predicates which can either have activity uses (compatible with the progressive), or stative ones under a habitual or generic reading (typically not compatible with the progressive), see for instance (5).

- (5) a. Peter **is running** through the valley. *literal, ✓ activity use*
 b. Peter **runs** every day. *literal, ✓ stative use*

Among motion verbs that have both activity and stative uses, some have a (non-literal) fictive motion use, see (6), under which they give an indication of the location or topology of a stationary entity (Talmy, 1996; Matlock, 2004). And as the contrast in (6) shows, only the stative use survives under this non-literal extension of motion verbs. In other words, verbs such as *run* seem to lose some of its aspectual potentialities once they are used non-literally.

- (6) a. **#The road is running** through the valley. *non-literal, ✗ activity use*
 b. The road **runs** through the valley. *non-literal, ✓ stative use*

The tendency can be formulated as in (7). It is illustrated by two additional examples below.

(7) *Aspectual rigidification under non-literality:*

The set of aspectual frames a verb may exhibit under its non-literal uses is a subset of the aspectual frames the same verb can instantiate under its literal uses.

Our second example concerns VPs that can have both telic and atelic uses. Atelic predicates (states or activities) are compatible with durative adverbials when they measure the event duration,¹ but not with *in*-adverbials, see (8). Telic predicates (accomplishments or achievements) show the opposite pattern, see (9) (e.g., Krifka, 1992).

- | | | |
|--------|---|-------------------------------|
| (8) a. | Peter ran for five minutes. | <i>literal</i> , ✓ atelic use |
| b. | #Peter ran in five minutes. | <i>literal</i> , ✗ telic use |
| (9) a. | #Peter killed the mosquito for five minutes. | <i>literal</i> , ✗ atelic use |
| b. | Peter killed the mosquito in five minutes. | <i>literal</i> , ✓ telic use |

As many authors observe, however, a subset of accomplishment VPs can, in fact, easily have an atelic use, and are consequently quite acceptable with *for*-adverbials measuring the event duration (see Smollett, 2005; Piñón, 2008; Deo & Piñango, 2011; Kennedy, 2012; Rothstein, 2012 for English; Bott, 2010, Bott & Hamm, 2014 for German; Arche, 2014 for Spanish; Martin, 2014 for French). The French verb *fondre* ‘melt’ is an example, see (10).

- | | | |
|---------|--|-------------------------------|
| (10) a. | La glace a fondu en deux minutes.
the ice cream melted in two minutes
‘The ice cream melted in two minutes.’ | <i>literal</i> , ✓ telic use |
| b. | La glace a fondu pendant deux minutes.
the ice cream melted for two minutes
‘The ice cream melted for two minutes.’ | <i>literal</i> , ✓ atelic use |

The relevant observation is that these predicates with telic and atelic uses tend to lose their aspectual flexibility when the verb is used non-literally. We illustrate the point with the verb *fondre* ‘melt’:²

¹Durative adverbials can also be used to measure the result state (cf. *I opened the window for one hour*), see Piñón (1999). Under its ‘result state related’ reading, *for*-adverbials are in principle compatible with telic predicates. Compatibility with *for*-adverbials diagnoses atelicity under their event-related reading only.

²Sentence (11b) is felicitous on the result-state related reading of the durative adverbial (meaning that his anger disappeared for two minutes), but is less acceptable on the event-related reading.

- (11) a. Sa colère **a fondu en** deux minutes. *non-literal*, ✓ telic use
his anger melted in two minutes
‘His anger melted in two minutes.’
- b. #Sa colère **a fondu pendant** deux minutes. *non-literal*, ✗ atelic use
his anger melted for two minutes
‘His anger melted for two minutes.’

A third example of the tendency (7) is provided by telic VPs which lose the accomplishmenthood they have under their literal use when used non-literally. As a result, they show properties of achievement VPs in their non-literal use—e.g., are not compatible with adverbials of completion such as *completely* or *part(ial)ly*, see Piñón (1997).³ For Spanish, Spalek (2014) observes that Spanish change-of-state verbs such as *romper* ‘break’ lose their gradability in non-literal uses; see (12), where the adverbial of completion is acceptable with the literal use of *romper* but not with the non-literal use. For French, predicates such as *monter* ‘raise’ or *assouplir* ‘soften’ may be embedded under aspectual verbs when used literally, but not when used non-literally (thus similar to achievements in this respect), see (13).

- (12) a. Su ligamento se **rompió (parcialmente)**. SPANISH
his ligament REFL break-PFV partly
‘His ligament tore (partially).’ *literal*, ✓ gradable
(Spalek, 2014)
- b. El grito **rompió (#parcialmente)** el silencio.
the scream broke partially the silence
‘The scream (partially) broke the silence.’ *non-literal*, ✗ gradable
(Spalek, 2014)
- (13) a. J’ai **fini d’assouplir** mes chaussures. FRENCH
I have finished to soften my shoes
‘I’ve finished softening my shoes.’ *literal*, ✓ accomplishment
- b. #L’assemblée a **fini d’assouplir** la loi.
the assembly has finished to soften the law
‘The assembly finished softening the law.’ *non-literal*, ✗ accomplishment
- c. Le ballon/#sa colère **a arrêté de monter**.
the balloon/his anger has stopped to raise
‘The balloon/his anger stopped raising.’ *literal/non-literal*, ✓/✗ accomplishment

³Both accomplishments and achievements are telic predicates, but the latter differ from the former in that they describe atomic events, devoid of proper parts. As a result, achievements are generally unfelicitous with adverbials of completion or aspectual verbs (*finish, start...*), see Piñón (1997) and Gyarmathy (2015) a.o.

2 Hypotheses

We believe that the two tendencies (4) and (7) illustrated above are quite significant, and inter-related at a deeper level, in that verb particles losing their optionality under the non-literal meaning have been argued to contribute to changing the lexical aspect/event structure of the VP they are part of (Smollett, 2005; Roßdeutscher, 2011, 2015). For instance, *aufsteigen* ‘raise’ is easier to combine with a completion adverbial than *steigen* ‘raise’, see (14a); *ein Buch auslesen* ‘read a book to the end’ has no atelic use, while *ein Buch lesen* ‘read a book’ does, see (14b) from Roßdeutscher (2011); *cut* has a conative use that *cut off* does not have (Levin & Rappaport Hovav, 2013), see (14c), etc.

- (14) a. Der Ballon **ist vollständig #(auf)gestiegen**. GERMAN
 the balloon is completely *auf*-raised
 ‘The balloon rose completely.’
- b. Er hat das Buch **drei Tage lang (#aus)gelesen**.
 he has the book three days long *aus*-read
 ‘He read the book (to the end) for three days.’
- c. I **cut (#off) at** the rope around my neck, but the knife was too dull. ENGLISH

To our knowledge, however, these tendencies have not been supported by large-scale empirical experiments yet. In the next sections, we report a set of experiments on French verbs and German particle verbs (with particles *ab*, *an*, *auf*, *aus*) and their base verb counterparts. Our goal is to evaluate the empirical scope of Hypotheses H1-H3 below. To our knowledge, H1 and H3 are new; H2 has been suggested before but not yet been supported by empirical experiments.

- (H1)** In German, non-literal senses prefer particle verbs (PVs) rather than the corresponding base verbs (BVs); BVs prefer literal over non-literal senses.
- (H2)** The German particles under study (*ab*, *an*, *auf*, *aus*) contribute to the aspectual profile of the VP they enter in, and do so in a particle-specific way (cf. Roßdeutscher, 2011, 2015, a.o.).
- (H3)** *strong version*: In non-literal senses, verbs tend to be either strictly stative (with no dynamic felicitous use) or strictly telic (‘rigid’ accomplishments with no atelic use or achievements).
weak version: In non-literal senses, verbs are aspectually less flexible (i.e., they enter a narrower range of aspectual frames) in comparison to their literal senses.

The experiments reported in the following Sections 3, 4 and 5 provide support for H1-H2 (which focus on German) and for both versions of H3 for French, but not for German. The last section, Section 6, sketches a preliminary theoretical account for the empirical findings.

3 Particle vs. bare verbs and non-literal vs. literal senses in German

Our first experiment empirically tests hypothesis H1. Our dataset comprises 16 German base verbs (BVs), and a systematic combination of these BVs with the four particles under study. For all BVs and PVs, we extracted sentences from the German COW corpus (Schäfer & Bildhauer, 2012) and automatically assigned a degree of (non-)literalness by applying the classifier from Köper & Schulte im Walde (2016). From this set of sentences, we randomly selected ten literal

and ten non-literal sentences for the 16 BVs, as well as 20 literal and 20 non-literal ones for each PV, for a total of 860 sentences (325 for BVs and 535 for PVs), as shown in the left panel of Table 1.

Table 1. (Non-)literal language usage in BV/PV sentences

	Automatic classification			Manual annotation		
	literal	non-literal	total	literal	non-literal	total
BVs	162	163	325	177 (79.7 %)	45 (20.3 %)	222
PVs	245	290	535	225 (57.4 %)	167 (42.6 %)	392
	407	453	860	402	212	614

Three annotators annotated these 860 sentences (in random order, so that BVs and PVs as well as literal and non-literal sentences were scrambled) on their degree of literalness on a 6-point [0-5] scale (0 = clearly literal; 5 = clearly non-literal). Breaking the results down to a binary decision by dividing the scale into two disjunctive ranges ([0-2] for literal; [3-5] for non-literal), we obtained the results summarized in the right panel of Table 1, relying on all sentences but excluding those without full annotator agreement. As Table 1 shows, 79.7 % of the BV-sentences were judged literal (20.3 % were judged non-literal), while only 57.4 % of the PV-sentences were judged literal (42.6 % were judged non-literal). According to χ -square, the difference in literal/non-literal distribution is significant ($p < 0.01$). We conclude that for the German verbs under study, hypothesis H1 is confirmed: non-literal senses prefer particle verbs rather than the corresponding base verbs; base verbs favor literal senses over non-literal senses. Schulte im Walde et al. (2018) replicate the annotation of (non-)literalness on a larger set of BVs and PVs and report the same tendency.

4 Interplay between lexical aspect and (non)-literalness

4.1 A finer-grained aspectual classification

Our further experiments aim to test a correlation between a shift from literal to non-literal language and a shift in lexical aspect (hypothesis H3). For the design of our experiments, we went beyond the traditional Vendlerian distinction in four aspectual classes, often claimed to be inadequate for the investigation of lexical aspect (see, e.g., Rappaport Hovav, 2008). We adopted a classification into eight aspectual subclasses exemplified in Table 2, capitalizing on previous distinctions proposed by Singh (1994), Kearns (2007), Piñón (2006, 2008), Tatevosov & Ivanov (2009), Rothstein (2012), Altshuler (2014), Martin (2014) and Wright (2014) on the basis of data from a variety of languages.

Table 2. A finer-grained typology of aspectual classes. The most (resp. least) telic class corresponds to the darkest (resp. clearest) cell

Aspectual subclasses	Example	
strictly stative VPs	<i>own a house</i>	
flexible stative VPs	<i>sit</i>	
activity VPs	<i>play in the garden</i>	
VPs of variable telicity	<i>read the Bible</i>	
weak accomplishment VPs	<i>write the letter</i>	
strong accomplishment VPs	<i>open the door</i>	
quasi-achievement VPs	<i>kill the cat</i>	
achievement VPs	<i>find the key</i>	

Atelic predicates are sorted into three subclasses, namely *strictly stative predicates* with no dynamic use; *flexible stative predicates*, that have derived dynamic uses besides their stative meaning, and *activity predicates*.

Predicates with a telic use are sorted into five rather than two aspectual subclasses according to their degree of telicity. On one extreme we find the predicates which are the least lexically biased towards telicity (corresponding to the clearest grey cell in Table 2), namely *predicates of variable telicity*. The aspectual lability of these predicates has been studied in Hay et al. (1999) and Kennedy & Levin (2008) a.o. Degree achievement derived from open-scale adjectives like *widen a scarf* are other examples of such predicates. In absence of an *in*-adverbial, perfective (PFV) sentences built with these predicates do not trigger an inference of culmination in a default context. They are nevertheless fully compatible with an *in*-adverbial (with no resort to a reinterpretation process through coercion, as is the case with activity predicates giving rise to an acceptable reading with *in*-adverbials). On the other extreme of the ‘telicity scale’ we find predicates that exhibit the highest degree of telicity, namely *achievement verbs*. These predicates do not license any atelic use, are not gradable and entail culmination when used with the perfective in most languages (corresponding to the black cell in Table 2).

Between these extreme classes there are three subclasses of accomplishments. Many authors have noted that accomplishments differ from each other with regard to the strength of their telic property: some accomplishments resemble more activity verbs than others, in that they require less forcefully a complete event interpretation when used with a perfect(ive) (e.g., Tenny, 1994; Smollett, 2005; Piñón, 2006; Kratzer, 2004; Mittwoch, 2013; Wright, 2014). We adopt the terminology introduced in Piñón (2006) and distinguish accomplishments by their strength: ‘Weak’ accomplishments satisfy less defining properties of accomplishmenthood than ‘strong’ ones: they accept quite well the event-related use of *for*-adverbials, see (15a), and in perfective sentences they allow subsequent denials of event culmination, see (15b). Incremental theme verbs such as verbs of consumption and creation often form weak accomplishment predicates. In this class, we also find ‘incremental’ causative accomplishments such as *empty*, for which an incremental relation is typically satisfied between the activity *e* of the subject’s referent and the change of state *e’* of the object’s referent (Rothstein, 2004, § 4.2.3.).⁴

(15) *Weak accomplishments:*

- a. Kathleen ate an apple for a couple of minutes while talking on the phone. (Piñón, 2008)
- b. I emptied the tub but not completely. (Rappaport Hovav, 2008)

On the other hand, strong accomplishments show a higher degree of telicity than weak accomplishments. They are not used naturally as activity predicates and consequently are much less acceptable with durative adverbials under the event-related use, see (16a), and in perfective sentences they do not license subsequent denials of event culmination so easily, (16b).⁵ Among this

⁴Informally, such a relation takes place between *e* and *e’* if there is a contextually available one-one function between each subpart of *e* and *e’*, such that related subevents temporarily coincide (e.g., the more I am emptying the tub, the more the tub gets emptier).

⁵Sentence (16b) is arguably acceptable if *not completely* does not serve to deny the completion of the waking up event, but rather indicates that the encoded result state property (here *awake*) is not satisfied at a maximal degree. Under this reading, the acceptability of *but not completely* does not reflect the acceptability of an event completion denial (see Piñón (2005) and Martin (forthcoming) for discussion).

class, we find non-incremental causative accomplishments, i.e. verbs for which an incremental relation between the activity involving the subject's referent and the change of state endured by the object's referent is non-natural, difficult or impossible, as well as VPs including a particle or a result phrase (*eat up, drink the teapot empty*). The fact that denying event completion and/or use a durative adverbial is more difficult with these two classes of VPs than with weak accomplishments has been noted for many languages (see Smollett (2005) and Roßdeutscher (2015) on English and German particle verbs, Singh (1994) on Hindi and Tatevosov & Ivanov (2009) on Karachay Balkar).

(16) *Strong accomplishments:*

- a. #Kathleen woke Mary up for a couple of minutes while talking on the phone.
- b. #Kathleen woke Mary up, but not completely.

Finally, the class of *quasi-achievement* predicates, of which *kill* is an example, straddles the class of strong accomplishments and achievements. This class gathers predicates that show properties of achievement verbs, such as the incompatibility with adverbials of completion or with aspectual verbs, cf. (17a/b). But they behave as accomplishments for others, such as the compatibility with the two (*in/after*) readings of *in*-adverbials (whereas achievements only select the 'after'-interpretation, cf. Piñón (1997)), see (17c), or the acceptability with the standard progressive, see (17d). Achievements are either not acceptable with the progressive, or under some specific readings only, see Gyarmathy (2015) for a thorough discussion.

(17) *Quasi-achievements:*

- a. John (#completely) killed the mosquito.
- b. #John started killing the mosquito.
- c. John killed the mosquito in 10 minutes. in/after 10 minutes
- d. John was killing the mosquito. ✓ standard progressive

Table 3 summarizes some of the aspectual properties underlying the aspectual typology we adopted, as well as the gold standard values we assign to each aspectual frame for each subclass on a six-point [0-5] scale (0 = unacceptable; 5 = acceptable). Sometimes, we give a range of several potential values instead of a categorical one. This corresponds to cases where a variation in degrees of acceptability arises within the same class. For instance, among weak accomplishment predicates, verbs of consumption have been argued to be more acceptable with *for*-adverbials than verbs of creation (Rothstein, 2012), which motivates the range [3-5] we attribute to weak accomplishments with *for*-adverbials. Aspectual classes showing the highest flexibility by entering the highest number of aspectual frames (and totalizing the highest score across frames) are called *flexible classes* (in grey in Table 3), while those being the most rigid are called *rigid* (in white).

4.2 Distribution of literal vs. non-literal verb senses across aspectual classes in French

Our second experiment tests the strong version of H3 for French. The dataset is composed of 167 frequent French verbs, with a good balance among the four Vendlerian aspectual classes. For these verbs, we extracted each sense as defined in the lexical resource 'Les Verbes Français'

Table 3. Aspectual subclasses with gold standard values w.r.t. aspectual frames

	Rigid class	Flexible classes				Rigid classes		
	X1	X2	X3	X4	X5	X6	X7	X8
	STAT	STAT- ACT	ACT	VAR	W-ACC	S-ACC	Q-ACH	ACH
	<i>own</i>	<i>sit</i>	<i>play</i>	<i>widen</i>	<i>eat an apple</i>	<i>wake up Ana</i>	<i>kill the cat</i>	<i>find the key</i>
T1. progressive	0	5	5	5	5	5	5	0 or 5
T2. <i>for</i> -adverbial (any reading)	5	5	5	5	3–5	0–5	0	0 or 5
T3. partitive <i>for</i> -adverbial	0	0	0	4	3–5	0–2	0	0
T4. <i>in</i> -adverbial	0	0	0	5	5	5	5	0 or 5
T5. <i>completely P</i>	0	0	0	3	5	0–5	0	0
T6. <i>P, but not completely</i>	0	0	0	3	5	0–5	0	0
T7. aspectual verbs	0	5	5	5	5	0	0	0
TOTAL (MINIMUM)	5	15	15	30	31	10	10	0

(LVF, François et al., 2007), for a total of 1199 verb senses. We also extracted the exemplifying sentences provided by the LVF for each of these 1199 senses.

We normalized the sentences in order to control for independent factors influencing the VP's lexical aspect. For instance, the LVF's exemplifying sentences (18) for *manger 8* and *manger 2* have been modified as in (19): the indefinite pronoun *on* favoring a generic interpretation is replaced by a full pronoun, the non-quantized DP *de la viande* 'meat' is replaced by a quantized DP, the durative adverbial is deleted, and the present tense is replaced by a perfective.

- (18) a. Les moustiques ont mangé P toute la nuit. = *manger 8* in LVF
the mosquitos have eaten P all the night
Literally: 'Mosquitos ate P the whole night.'
- b. On mange de la viande. = *manger 2* in LVF
one eats of the meat
'We are eating/eat meat.'
- (19) a. Les moustiques ont mangé Pierre.
the mosquitos have eaten Pierre
Literally: 'Mosquitos ate Pierre.'
- b. Pierre a mangé la pizza.
Pierre has eaten the pizza
'Pierre ate the pizza.'

The first author of this article annotated the sentences according to the aspectual classification into 8 aspectual subclasses described in Section 4.1. We also extracted the LVF sense classification according to (non)-literalness for all sentences, and selected the subset of lemmas which (i) have literal and non-literal senses and (ii) have senses in both rigid aspectual classes X1, X6, X7 and X8 and flexible aspectual classes X2-X5. 31 % of the lemmas fulfilled these two criteria.

We then compared the percentage of non-literal vs. literal uses within the sentences instantiating rigid aspectual classes (X1, X6-X8).

Table 4. Distribution of literal vs. non-literal uses of *vider* across aspectual classes

Aspectual classes	non-literal	literal	total
flexible	1	7	8
rigid	3	2	5
	4	9	13

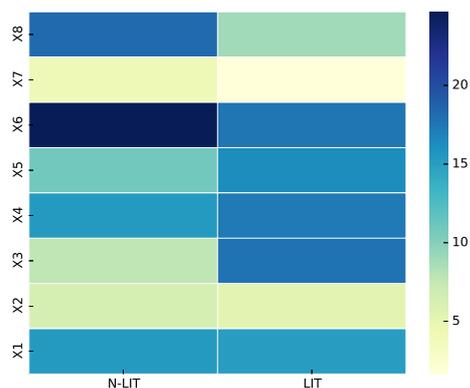


Figure 1. Literal vs. non-literal senses across eight aspectual classes

Table 4 gives the results for the lemma *vider* ‘empty’, for which the LVF lists 13 different senses. As the table shows, the percentage of non-literal senses in rigid aspectual classes (3 out of 4, i.e., 75 %) is higher than the percentage of literal senses in the same classes (2 out of 9, i.e. 22 %).

The same tendency was observed for most of the 53 relevant lemmas (60 %). For these lemmas, the percentage of non-literal verb senses in rigid aspectual classes was superior to the percentage of literal verb senses.

Figure 1 shows the distribution of the literal vs. non-literal senses across the 8 aspectual classes. We take these data to offer support for the strong version of H3.

4.3 Further cross-linguistic explorations

Further experiments aimed to bring together the cross-linguistic explorations in a parallel fashion for French and German, and going beyond a single-annotator setup and a rigid aspectual categorisation.

4.3.1 Dataset and annotation

For French, we selected out of the 167 lemmas used in the previous experiment the verbs instantiating the broadest range of aspectual classes, for a total of 57 lemmas. To this set, we added lemmas prefixed in *a-*, *en-* and *dé-* and their bare (unprefixed) counterparts so as to analyse the aspectual shape of these prefixes, for a total of 100 verb senses for each prefix. We arrived at a total of 1200 verb senses. For German, we extracted a total of 1905 example sentences from GermaNet version 9 for 1100 *ab/an/aus/auf*-PVs in GermaNet, with a good balance across particles (*ab*: 27 %, *aus*: 25 %, *auf*: 22 %, *an*: 25 %). We normalized the sentences in the same way as we did for French (the present perfect was used).

Three German and three French annotators classified the German/French sentences according to (non-)literalness on a six-point [0-5] scale. The agreement was moderate for both languages ($\kappa = 0.41$ and $\kappa = 0.43$). For aspect, the task was to evaluate the acceptability of eight key aspectual properties, using a [0-5] scale (0 = totally unacceptable; 5 = totally acceptable). For German, we assigned 7 aspectual tasks corresponding to the aspectual frames T2–T7 in Table 3 plus an additional task (T8).

- T2 acceptability of *X Minuten/Stunden ... lang* on any (non-iterative) use (resultant state related or event-related readings, cf. Piñón, 1999);
- T3 acceptability of *X Minuten/Stunden ... lang* on what Champollion (2013) calls the ‘parti-

tive’ (event-related) interpretation, where the sentence describes a part of a culminating event (*Er hat einen Apfel 2 Minuten lang gegessen* ‘He ate an apple for 2 minutes’);

T4 acceptability of *in X Minuten/Stunden*;

T5 acceptability of an adverbial of completion within the VP (*vollständig P*);

T6 acceptability of an extra-clausal completion denial (*P, aber nicht vollständig*);

T7 embeddability under *aufhören* (embeddability under *anfangen* ‘begin’ is tested in T1);

T8 an evaluation of whether the sentence triggers an inference of culmination (0 = absolutely not; 5 = yes with no doubt).

Note that task T2 does not target a specific (non-iterative) reading of the *for*-adverbial. Only T3 does, since it specifically tests the event-related reading. This partitive interpretation of *for*-adverbials is selected only by telic predicates (since atelic predicates do not describe events associated to an inherent culmination point).

Table 5. Inter-annotator agreement (Fleiss’ κ) for aspectual tests in French and German

	French tasks	Fleiss’ κ	German tasks	Fleiss’ κ
T1	<i>être en train de P</i>	0.33	<i>anfangen zu P</i>	0.37
T2	<i>pendant Xmin</i>	0.20	<i>XminLang</i>	0.32
T3	partitive <i>pendant Xmin</i>	0.06	partitive <i>XminLang</i>	0.36
T4	<i>en Xmin</i>	0.46	<i>in Xmin</i>	0.42
T5	<i>complètement P</i>	0.38	<i>vollständig P</i>	0.49
T6	<i>P, mais pas complètement</i>	0.30	<i>P, aber nicht vollständig</i>	0.32
T7	<i>arrêter de P</i>	0.34	<i>aufhören zu P</i>	0.28
T8	completion inference	0.25	completion inference	0.26

For French, the three annotators classified the 1200 sentences according to the same aspectual tasks, except that the *anfangen*-task was replaced by a progressive (*être en train de* ‘be V-ing’) task. Again, all test sentences were perfective (with the *passé composé*), except for the progressive test. In both languages, the agreement was fair to moderate depending on the tasks.⁶ Table 5 gives an overview of the inter-annotator agreement (Fleiss’ κ) for each test in French and German. As one can see, the agreement is roughly similar in both languages, except for the partitive *for*-adverbial (T3). This may be due to the fact that in French, but not in German, the annotators have been instructed to skip task T3 when the acceptability of the *for*-adverbial (T2) under any reading was below 2. For the computation of the kappa, we converted the absence of a value into a 0, which possibly distorted the ratings.

4.3.2 Spearman correlations across annotations

For German and French, we converted all annotations to the mean value of all annotators. If the mean value on the (non-)literalness dimension was below 2.5, we categorized the sentence as literal (or non-literal otherwise). For French, roughly 56 % (602) sentences were found to have a literal meaning and 44 % (464) sentences were labeled non-literal. For German, only 14.3 % of

⁶The annotation guidelines provided to the annotators can be found in the supplementary material found at github.com/ifalk/Interplay-Aspect-Literalness.git. An annotation file summarising all annotations will be added to the supplementary material in the nearest future. For French, the annotators also had to evaluate the resultant-state related reading of frame (*for*-adverbials as well as the paraphrasability of frame *in*-adverbials by *après* ‘after’), but these annotations have not been taken into account in the experiments reported in this paper.

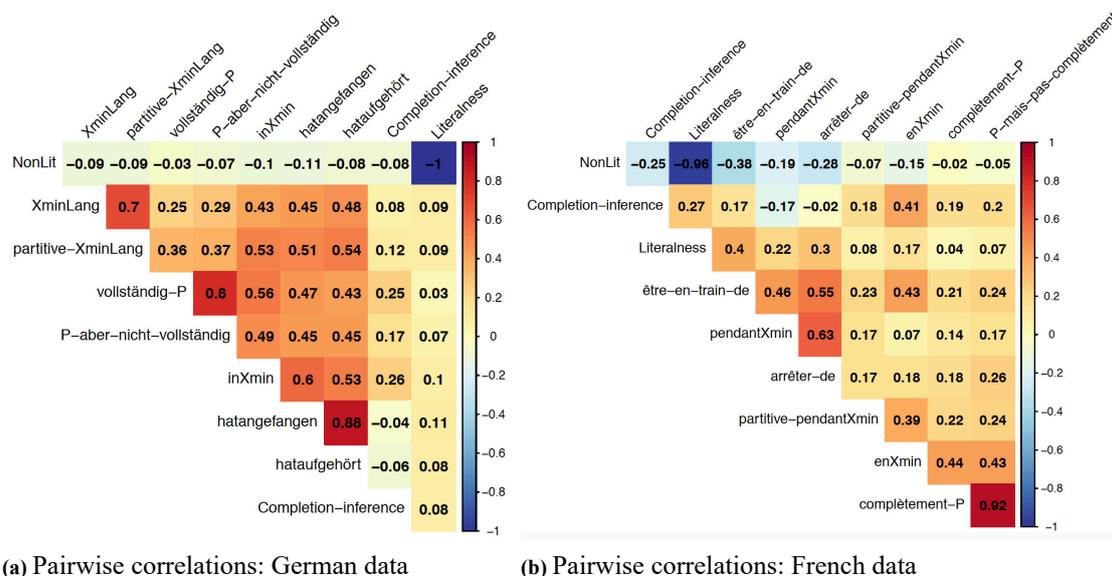


Figure 2. Pairwise Spearman correlations between German and French annotations of aspectual tests

the data was labeled non-literal, for a total of 272 sentences. Note that this difference is not due to the way the data were extracted nor to a difference in the architecture of LVF and GermaNet. For French, each verb sense delineated in the LVF lexicon corresponds to exactly one sentence in our dataset so that the percentage of literal vs. non-literal verb senses in the lexicon and in our dataset is the same. For German, the 1905 example sentences also correspond to one sentence per verb sense. This suggests that the *ab/an/aus/auf*-PVs in GermaNet overall have more literal than non-literal meanings.⁷

We computed the Spearman ρ correlations between all tasks for both languages. Figure 2 shows all pairwise correlations for (a) German and (b) French. The positive correlations are in orange and red (for the strongest ones), and the negative correlations in blue. For instance, we see that literalness and non-literalness show the most negative correlation (-1 and -0.96), which is expected given the binary classification adopted.

Correlations between (non)-literalness and aspect. Figure 2b shows that in French, non-literalness negatively correlates with progressivity (-0.38) and with the inference of completion (-0.25). This can be seen as a confirmation that non-literal verb senses tend to be more often stative (in accordance with H3), as the contrast in (6) illustrated. What is also interesting is that non-literalness negatively correlates with the embeddability under *arrêter* ‘stop’ in French (-0.28). This suggests that altogether, verbal predicates under their non-literal use more often tend to be either stative predicates, or to behave as achievement predicates (again supporting H3), as the contrasts in (13) illustrated. However, we do not observe a correlation between non-literalness and the acceptability of the partitive reading of durative adverbials (-0.07), although we should observe a negative correlation if non-literal verb senses tend to be either stative or highly telic. Note, however, that the inter-annotator agreement for the partitive *for*-adverbials was particularly low in French, which may explain why a negative correlation is not observed. As for German, Figure 2a shows hardly any correlation between non-literalness and any aspectual

⁷In cases where GermaNet provided more than one example sentence per verb sense, we only used the first one mentioned, but it is not the case that literal sentences are listed first, and non-literal sentences later (which means that we did not lose non-literal senses by proceeding this way). Also, when we transferred the example sentences to the present perfect, we lost 161 sentences (i.e., the original number of sentences was 2066) where the change of tense generated infelicity.

features. This cross-linguistic difference may be due to the fact that non-literal verb senses are much less represented in our data in German than in French (14 % vs. 44 %). But it may also reflect a structural difference between German, which has particle verbs, and French, which does not. We will explore this hypothesis in the last section.

Correlations between aspectual features. Many of the correlations between aspectual features shown in Figures 2b and 2a are relatively unsurprising given standard theoretical assumptions on lexical aspect. For instance, a positive correlation between the *in*-adverbial and the inference of completion is expected (since the value for these tests covaries for telic predicates as well as for atelic ones), a prediction confirmed for French (0.41), although (surprisingly) less so for German (0.26). Relatedly, the absence of positive or negative correlation between the values for the *for*-adverbial and the inference of completion is also expected (-0.17 for French and 0.08 in German), since predicates accepting *for*-adverbials can either be atelic (or weakly telic) when the adverbial has its event-related reading, or telic when it has its result-state related reading (recall that the *for*-adverbial test T2 does not disambiguate between readings). The relatively strong positive correlation between the embeddability under *stop* and the *for*-adverbial (0.63 in French, 0.48 in German) is also expected, given that *for X minutes* adverbials trigger the inference that the described event *stopped* after X minutes (Piñón, 1999; Bott, 2010; Bott & Hamm, 2014). The strong positive correlation between the value for the *anfangen* test and the *aufhören* one (0.88) is also unsurprising, given that these two aspectual frames basically test the same aspectual properties (dynamicity and non-atomicity).

For French, we also expect the observed positive correlation (0.55) between the acceptability of the progressive and the embeddability under *stop*, given that predicates that accept (respectively reject) the progressive are also those that can (respectively cannot) embed under *stop* (see the high correlation between the gold standard values for T1 and T7 in Table 3), except for one subclass only (the strong accomplishments). Also expected is the positive correlation between *completely P* on one hand and the other aspectual frames indicative of telicity, such as the *in*-adverbial (0.44 in French, 0.56 in German), as well as the partitive event-related reading of the durative adverbial (0.39 in French, 0.53 in German). That many expected correlations are in fact observed further confirms the reliability of the manual annotations.

What is from a traditional point of view more surprising is that no strongly negative correlation occurs between the values for the *for*-adverbial test and the *in*-adverbial test, since the standard assumption is that these adverbials are in complementary distribution across most predicates (predicates of variable telicity aside). No negative correlation is observed in French (0.07), which confirms our (and others') assumption that aspectual lability is a much broader phenomenon than generally supposed. Even more surprisingly, however, these two tests are *positively* correlated in German (0.43), for reasons unclear to us at this point.

Another intriguing fact from a theoretical perspective is the very high correlation (0.92 for French, 0.80 for German) between the acceptability of completion adverbials (*completely P*, T5) and extra-clausal completion denials with an adverbial of completion (*P, but not completely*, T6). This is surprising: remember that the sentences are tested with a perfective tense, which is standardly assumed to require the described event to be complete. Under this perspective, event completion denials of the '*P, but not completely*' type should trigger a contradiction after a perfective sentence, and thus give rise to weaker judgments of acceptability than the corresponding *completely P* sentences. The fact that, on the contrary, a strongly positive correlation arises between these two tests confirms recent observations that perfective telic sentences *x VP-PFV* fail to entail *completely VP-PFV*. The lack of entailment has been attributed to a variety of (in some cases compatible) sources, such as (i) the possibility to interpret the determiner used in the VP

as referring to a proper part of the entity rather than to the whole of it (*the pizza* used to describe a proper part of the pizza, see Križ (2016), Kennedy & Levin (2008), Martin & Gyarmathy (forthcoming)), (ii) the fact that some accomplishments may entail a standard telos rather than a maximal telos (Kearns, 2007; Fleischhauer, 2016; see McNally, 2017, for discussion), or (iii) a switch from a coarse-granularity level to a fine-granularity level triggered by *completely* (Sassoon & Zevakhina, 2012; Martin, forthcoming). In this perspective, the high correlation between the values for the tests T5 and T6 is less surprising.

An intriguing difference between French and German concerns the correlations between the acceptability of the partitive interpretation of *for*-adverbials (only possible with telic predicates), and the acceptability of other frames indicative of telicity, such as *completely P, P but not completely* and the *in*-adverbial. These correlations are (as expected) positive in German (0.36, 0.37 and 0.53), but consistently less so in French (0.22, 0.24 and 0.39). This difference may be due to the low inter-annotator agreement for the partitive *for*-adverbial task in French (see Table 5). But it may also reflect a semantic difference between the French perfective tense on one hand, a perfective morphology requiring event completion, and the German present perfect on the other, which is not a perfective, and thus at best triggers a *defeasible* inference of completion with a telic predicate. We could then explain the French vs. German difference as follows: under the partitive interpretation, the *for*-adverbial cancels the completion inference triggered by perfect(ive) telic predicates (in *He ate the sandwich for 2 minutes*, the phrase *for 2 minutes* cancels the inference of completion by default triggered by *He ate the sandwich*). Since this inference is stronger in French than in German, it is more difficult to defeat it with the adverbial. As a consequence, the partitive *for*-adverbial is less easily accepted in French than in German, hence the weaker correlation in French with other frames indicative of telicity.

4.3.3 Are verbs aspectually less flexible in their non-literal meaning?

Our next experiment was designed so as to test the weak version of H3, according to which predicates are aspectually more flexible—i.e., enter a broader range of aspectual frames—in their literal senses than in their non-literal senses. We approximated the aspectual flexibility of each verb sense by summing the annotators’ scores for five of the aspectual tasks (T2 and T4–T7). We only considered lemmas having both literal and non-literal senses according to the annotators’ scores. For each of these lemmas, we computed the average of the sums obtained for each of its sense. We found that for 62 % of the French and for 54 % of the German lemmas, the literal average was greater than the non-literal average. This experiment thus provides support to the weak version of H3 (a positive correlation between aspectual flexibility and literalness) for French, but less so for German, as shown by the pairwise Spearman correlations in Figure 2.

The next experiment tests the strong version of H3, according to which non-literal verb senses tend to gather at both extremes of the ‘telicity scale’ illustrated in Table 3, namely stative senses, or highly telic ones (strong accomplishments and achievements).

Table 6. Mapping of verb senses onto three aspectual superclasses

$0 < F-I$	ATE(LIC)
$-0.3 < F-I < 0$	VAR(IABLE)
$F-I < -0.3$	H-TEL(IC)

The aspectual profile for each verb sense was approximated by subtracting the value I for the *inXmin* task (T4) from the value F for the *forXmin* task (T2). Verb senses for which $F-I > 0$ were mapped to ATE(LIC), those for which $-0.3 < F-I \leq 0$ were mapped to VAR(IABLE), and those for which $F-I < -0.3$ were mapped to H(IGHLY)TE(LIC), see Table 6.

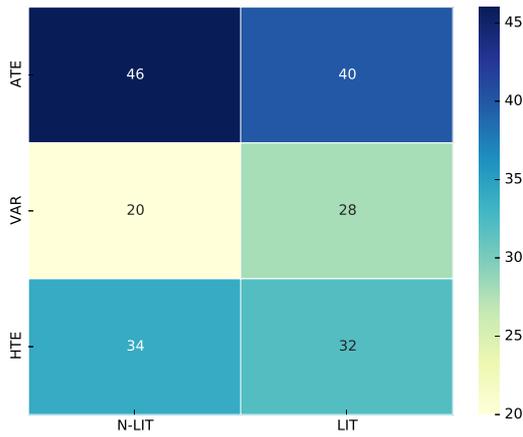


Figure 3. Aspectual profile of French verbs in their literal vs. non-literal senses

The heatmap in Figure 3 shows the distribution for literal vs. non-literal verb senses for French. We see that 46 % of all non-literal senses are atelic, 20 % of variable telicity and 34 % highly telic. On the other hand, 40 % of all literal senses are atelic, 28 % are variable and 32 % highly telic. In accordance with the strong version of H3, non-literal senses are less present in the variable class than literal senses. However, the results obtained from the German data do not show the same tendency: 34 % of the non-literal senses and only 27 % of the literal sense are of variable telicity. Whereas the distribution in aspectual classes of the literal senses is similar to that in

French, the distribution of the non-literal senses differs from French in that it is less skewed towards the “rigid” aspectual classes: 34 % are of variable telicity, and 24 % are highly telic.

5 Aspectual shape of the German particles *ab*, *an*, *auf*, *aus*

The goal of our last experiment is to provide support through large-scale empirical data to the hypothesis that the particles *ab*, *an*, *auf* and *aus* contribute to the aspectual profile of the VP they enter in, and do so in a particle-specific way (H2). We used the same method of approximation of the aspectual profile as in the previous experiment (by mapping the German PV sentences onto ATE, VAR, HTE depending on the value obtained for T2-T4, see Table 6).

The results in Figure 4 show the distribution of atelic, variable and highly telic senses across particles. For instance, as indicated in the darkest box on the left bottom, 59 % of the atelic senses are instantiated by sentences with a *an*-PV. All in all, the results suggest that on one hand, *an*-PVs instantiate the most atelic senses, *ab*-PVs the least atelic senses, while on the other, *ab*-PVs and *aus*-PVs exhibit most of the highly telic senses, and *an*-PVs the least of them. We then compared the annotator scores across tasks (divided into the two classes literal/non-literal), with additional splits into particles. The boxes in Figure 5 reflect the distribution. Again, most features show little correlation/differences with respect to (non)-literalness, but the distribution confirms that *ab* has the most telic profile, with the highest scores for *inXmin*, the inference of culmination and the adverbial of completion, while *an* has the most atelic profile, with the lowest scores for *inXmin* and adverbials of completion. Obviously, the aspectual shape we provide to particles can only be very coarse-grained, since we compute it on the basis of annotations across all verb types and verb senses, while particles arguably contribute quite differently to aspect from verb type to verb type, and even from verb sense to verb sense. However, given that each verb sense corresponds to one sentence

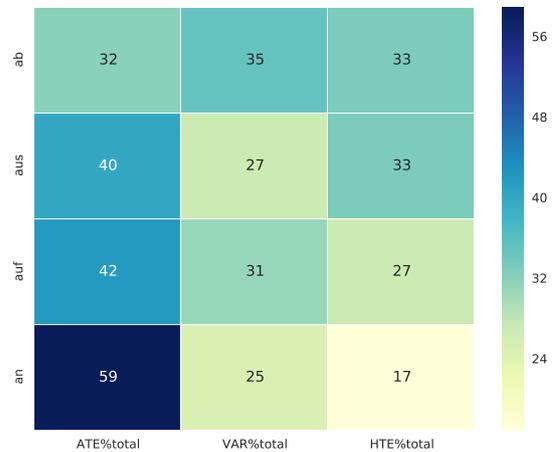
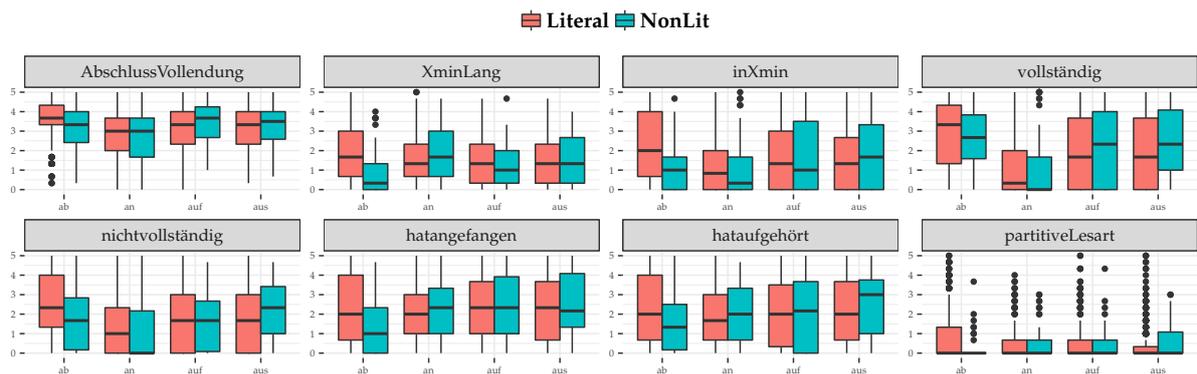


Figure 4. Aspectual classes across German particle verbs; split by particle type

Figure 5. Annotator scores for literal (red) vs. non-literal (green) sentences across tasks; split by particle



only in our German dataset, our computation at least has the merit to reflect the *average* or *most frequent* aspectual profile of each particle across verb types and verb senses.

6 The grammar of non-literality

The experiment reported in Section 3 show that particles which are optional under the non-literal meaning are often required for non-literal meaning (our hypothesis H1), as illustrated by the examples repeated below under (20a/b). We saw that the same is true of adjuncts in languages without particle verbs, see (20c).

- (20) a. I can **cut** #(off) the circulation in your feet, then they get really cold. *non-literal*
 (GloWbE, McNally & Spalek, 2017)
- b. In ihm **ist** Wut #(auf)gestiegen.
 in him is anger *auf*-raised
 ‘He started to feel angry.’ *non-literal*
- c. Le destin **sourit** #(à Marie).
 the destiny smiles at Marie
 ‘Destiny is smiling at Marie.’ *non-literal*

In this last section, we would like to provide some reflections on the potential basis of such a restriction, and along the way provide a reason for why we observe a positive correlation between literalness and aspectual flexibility in French, but not in German.

Let us assume that the event structure and the basic semantic features associated with a VP, such as the semantic types (manner or result predicates) assigned to subevents and the semantic relations defined between arguments and subevents, are kept constant under its literal and non-literal interpretations. Some key differences characterize non-literal interpretations, however:

- i. *ontological shift*:
 often, the VP predicates and quantifies over a domain of abstract entities (rather than the domain of physical entities as is often the case under the literal reading);

ii. *semantic bleaching*:

some semantic components may be ‘bleached’, weakened, in that they do not define the truth conditions of the sentence in the same way as under the literal sense;

iii. *semantic strengthening*:

some semantic components optional under the literal meaning of the VP may become compulsory under the non-literal sense.

- (21) a. Pierre **a lavé** la voiture, mais ce fut absolument sans aucun effet!
 Pierre has washed the car but it was absolutely without any effect
 ‘Pierre washed the car, but this had absolutely no effect!’
- b. Ces vacances **m’ont lavé** la tête, #mais ce fut absolument sans
 These holidays me have washed the head but it was absolutely without
 aucun effet!
 any effect
 ‘These holidays cleared my head, but this had absolutely no effect!’

Take for instance the contrast in (21). While *laver* ‘wash’ predicates over a physical entity under its literal sense (21a), it does so over a psychological one under its non-literal sense (21b). The manner component of *laver* (the use of fluid forces) is compulsory under the literal reading, and it is bleached under the non-literal one. On the contrary, however, the result component of the verb (the result conventionally associated with the process described by the verb, e.g., be clean(er) in the case of *wash*) is optional and thus cancellable under the literal reading (Talmy, 1991; Rappaport Hovav & Levin, 1998), but not under the non-literal one (Alexiadou et al., 2017), as witnessed by the fact that denying the result generates a contradiction in (21b). We propose that the result component becomes compulsory under the non-literal meaning precisely *because* the manner component is not satisfied. Otherwise, *no* event type at all associated to the verb would be satisfied under the non-literal meaning, which would empty the verb from the semantic component that differentiates it from any other transitive verb (with the same semantic roles).

We propose that if particles (or other optional elements such as adjuncts) are often required to express non-literal meaning, it is basically for the same reason: under the non-literal sense, some components (such as the manner component) are bleached, and the particle/adjunct, otherwise optional, becomes compulsory because it brings in the single semantic component justifying the applicability of the verb under its non-literal meaning. For instance, in (2a), the manner component of *cut* (the use of a sharp tool on the theme’s referent) is bleached. It is the result component that justifies the extension of *cut* to this abstract domain, and this component is brought about by the particle *off*, not by the base verb *cut* (McNally & Spalek, 2017). In (2b), the movement upwards is not brought in by the (manner) predicate *steigen* ‘raise’ (compatible with *runter* ‘down’, as in *runtersteigen*), but by the particle *auf* (Lechler & Roßdeutscher, 2011), therefore necessary for *steigen* to get its ‘coming up/into appearance’ sense. In (20c), the destiny does not smile (manner component) under the non-literal meaning, but it produces the result a smile usually triggers.

On the other hand, if non-literal senses make verbs aspectually less flexible, it is—among other factors—because the unfolding of abstract events is less easy to track down than the unfolding of concrete events, for abstract events are assigned a poorer, more opaque, mereological

structure (see also Spalek, 2014). For instance, the non-literal sense of (13c) is not felicitous because it is odd to assign several stop-overs to the trajectory of X's anger in an abstract space.

Finally, we propose that if H3 is less clearly supported in German than in French, it is perhaps because the aspectual shift triggered by non-literalness in French partly manifests itself in German through the loss of the particle's optionality. In this view, H1 and H3 represent two faces of the same coin.

Acknowledgments

This work was funded by Projects B5 & D12 of the SFB 732 granted by the DFG and hosted by the University of Stuttgart. We are grateful to the anonymous reviewers for their comments on a previous version as well as Nils Hirsch, Tillmann Pross and Florian Schäfer for discussions on our German data. We thank our annotators: Delia Alf, Anne-Lou Bitouzet, Katharina Fezer, Jessica Foc, Alexander Frey, Lea Gustedt, Stéphanie Günther, Manon Jeanningros, Ingrid Kasimir, Florian Lux, Monique Marquez and Daniela Naumann.

References

- Alexiadou, A., Martin, F., & Schäfer, F. (2017). Optionally causative manner verbs: When implied results get entailed. Talk presented at 'Roots V', Queen Mary University, London.
- Altshuler, D. (2014). A typology of partitive aspectual operators. *Natural Language & Linguistic Theory*, 32(3), 735–775.
- Arche, M. (2014). The construction of viewpoint aspect: The imperfective revisited. *Natural Language & Linguistic Theory*, 32(3), 791–831.
- Boons, J.-P. (1971). Métaphore et baisse de la redondance. *Langue Française*, 11, 15–16.
- Bott, O. (2010). *The Processing of Events*. Amsterdam: John Benjamins.
- Bott, O., & Hamm, F. (2014). Cross-linguistic variation in the processing of aspect. In B. Hemforth, B. Mertins & C. Fabricius-Hansen (Eds.), *Psycholinguistic Approaches to Meaning and Understanding Across Languages* (pp. 83–109). Berlin: Springer.
- Champollion, L. (2013). The scope and processing of *for*-adverbials: A reply to Deo and Piñango. In T. Snider (Ed.), *Proceedings of Semantics and Linguistic Theory (SALT) 23* (pp. 432–452).
- Deo, A. S., & Piñango, M. M. (2011). Quantification and context in measure adverbs. In N. Ashton, A. Chereches & D. Lutz (Eds.), *Proceedings of Semantics and Linguistic Theory (SALT) 21* (pp. 295–312).
- Fleischhauer, J. (2016). *Degree Gradation of Verbs*. Düsseldorf: Düsseldorf University Press.
- François, J., Le Pesant, D., & Leeman, D. (2007). Présentation de la classification des *Verbes Français* de Jean Dubois et Françoise Dubois-Charlier. *Langue Française*, 153, 3–19.
- Gyarmathy, Z. (2015). *Achievements, Durativity and Scales*. Berlin: Logos Verlag.
- Hay, J., Kennedy, C., & Levin, B. (1999). Scalar structure underlies telicity in degree achievements. In T. Matthews & D. Strolovitch (Eds.), *Proceedings of Semantics and Linguistics Theory (SALT) 9* (pp. 127–144).
- Kearns, K. (2007). Telic sense of deadjectival verbs. *Lingua*, 117, 26–66.
- Kennedy, C. (2012). The composition of incremental change. In V. Demonte & L. McNally

- (Eds.), *Telicity, Change, State: A Cross-Categorical View of Event Structure* (pp. 103–121). Oxford: Oxford University Press.
- Kennedy, C., & Levin, B. (2008). Measure of change: The adjectival core of verbs of variable telicity. In L. McNally & C. Kennedy (Eds.), *Adjectives and Adverbs: Syntax, Semantics, and Discourse* (pp. 156–182). Oxford: Oxford University Press.
- Köper, M., & Schulte im Walde, S. (2016). Distinguishing literal and non-literal usage of German particle verbs. In K. Knight, A. Nenkova & O. Rambow (Eds.), *Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (pp. 353–362). San Diego, CA.
- Kratzer, A. (2004). Telicity and the meaning of objective case. In J. Guéron & J. Lecarme (Eds.), *The Syntax of Time* (pp. 389–423). Cambridge, MA: MIT Press.
- Krifka, M. (1992). Thematic relations as links between nominal reference and temporal constitution. In I. Sag & A. Szabolsci (Eds.), *Lexical Matters* (pp. 29–53). Stanford, CA: CSLI Publications.
- Križ, M. (2016). Homogeneity, non-maximality and *all*. *Journal of Semantics*, 33, 493–539.
- Lamiroy, B. (1987). Les verbes de mouvement. Emplois figurés et extensions métaphoriques. *Langue Française*, 76, 41–58.
- Lechler, A., & Roßdeutscher, A. (2009). German particle verbs with ‘auf’. Reconstructing their composition in a DRT-based framework. *Linguistische Berichte*, 220, 439–478.
- Levin, B., & Rappaport Hovav, M. (2013). Lexicalized meaning and manner/result complementarity. In B. Arsenijević, B. Gehrke & R. Marín (Eds.), *Studies in the Composition and Decomposition of Event Predicates* (pp. 49–70). Amsterdam: Springer.
- Martin, F. (2014). Revisiting the distinction between achievements and accomplishments. In W. de Mulder, J. Mortelmans & T. Mortelmans (Eds.), *Cahiers Chronos 22. From Now to Eternity* (pp. 43–64). Amsterdam: Rodopi.
- Martin, F. (forthcoming). Non-culminating accomplishments.
- Martin, F., & Gyarmathy, Z. (forthcoming). A more fined-grained typology of perfective operators. In C. Piñón (Ed.), *Empirical Issues in Formal Syntax and Semantics*, 12.
- Matlock, T. (2004). Fictive motion as cognitive simulation. *Memory & Cognition*, 32(8), 1389–1400.
- McNally, L. (2017). On the scalar properties and telicity of degree achievements. In O. Fernández-Soriano, E. Castroviejo & I. Pérez-Jiménez (Eds.), *Boundary, Phases and Interfaces. Case studies in honor of Violeta Demonte* (pp. 173–192). Amsterdam: Benjamins.
- McNally, L., & Spalek, A. (2017). *‘Figurative’ uses of verbs and grammar*. Universitat Pompeu Fabra and Universitetet i Oslo.
- Mittwoch, A. (2013). On the criteria for distinguishing accomplishments from activities, and two types of aspectual misfits. In B. Arsenijevic, B. Gehrke & R. Marín (Eds.), *Studies in the Composition and Decomposition of Event Predicates* (pp. 27–48). Berlin: Springer.
- Piñón, C. (1997). Achievements in an event semantics. In A. Lawson (Ed.), *Semantics & Linguistic Theory (SALT) 7* (pp. 273–296).
- Piñón, C. (1999). Durative adverbials for result states. In S. Bird, A. Carnie, J. D. Haugen & Peter Norquest (Eds.), *Proceedings of the 18th West Coast Conference in Formal Linguistics*

- (WCCFL) (pp. 420–433). Somerville, MA: Cascadilla Press.
- Piñón, C. (2005). Adverbs of completion with events and degrees. In A. van Hout, H. de Swart & H. Verkuyl (Eds.), *Proceedings of Perspectives on Aspect* (pp. 149–166). Dordrecht: Kluwer.
- Piñón, C. (2006). Weak and strong accomplishments. In K. É. Kiss (Ed.), *Event Structure and the Left Periphery. Studies on Hungarian* (pp. 91–106). Dordrecht: Springer.
- Piñón, C. (2008). Aspectual composition with degrees. In L. McNally & C. Kennedy (Eds.), *Adjectives and Adverbs: Syntax, Semantics, and Discourse* (pp. 183–219). Oxford: Oxford University Press.
- Rappaport Hovav, M. (2008). Lexicalized meaning and the internal temporal structure of events. In S. Rothstein (Ed.), *Theoretical and Crosslinguistic Approaches to the Semantics of Aspect* (pp. 13–42). Amsterdam: Benjamins.
- Rappaport Hovav, M., & Levin, B. (1998). Building verb meanings. In M. Butt & W. Geuder (Eds.), *The Projection of Arguments: Lexical and Compositional Factors* (pp. 97–134). Chicago, IL: CSLI Publications.
- Roßdeutscher, A. (2011). Particle verbs and prefix verbs in German: Linking theory versus word-syntax. *Leuvense Bijdragen*, 97, 1–53.
- Roßdeutscher, A. (2015). Hidden universal quantification and change of argument structure in particle-verb constructions. In E. Csipak & H. Zeijlstra (Eds.), *Proceedings of Sinn und Bedeutung 19* (pp. 518–534).
- Rothstein, S. (2004). *Structuring Events: A Study in the Semantics of Lexical Aspect*. Malden/Oxford: Blackwell Publishing.
- Rothstein, S. (2012). Another look at accomplishments and incrementality. In V. Demonte & L. McNally (Eds.), *Telicity, Change and State* (pp. 60–120). Oxford: Oxford University Press.
- Ruppenhofer, J. (2006). Fictive motion: Construction or construal? In Z. Antić, C. B. Chang, E. Cibelli, J. Hong, M. J. Houser, C. S. Sandy, M. Toosarvandani & Y. Yao (Eds.), *Proceedings of the 32nd Annual Meeting of the Berkeley Linguistics Society* (pp. 307–319).
- Sassoon, G. W., & Zevakhina, N. (2012). Granularity shifting: Experimental evidence from degree modifiers. In A. Chereches (Ed.), *Proceedings of Semantics and Linguistics Theory (SALT) 22* (pp. 226–246).
- Schäfer, R., & Bildhauer, F. (2012). Building large corpora from the web using a new efficient tool chain. In N. Calzolari, K. Choukri, T. Declerck, M. U. Doğan, B. Maegaard, J. Mariani, A. Moreno, J. Odijk & S. Piperidis. (Eds.), *Proceedings of the 8th International Conference on Language Resources and Evaluation* (pp. 486–493). Istanbul.
- Schulte im Walde, S., Köper, M., & Springorum, S. (2018). Assessing meaning components in German complex verbs: A collection of source-target domains and directionality. In *Proceedings of the 7th Joint Conference on Lexical and Computational Semantics* (pp. 22–32).
- Singh, M. (1994). *Perfectivity, Definiteness, and Specificity: A Classification of Verbal Predicates in Hindi*. Dissertation, University of Texas at Austin.
- Smollett, R. (2005). Quantized direct object don't delimit after all. In H. Verkuyl & H. de Swart (Eds.), *Perspectives on Aspect* (pp. 41–60). Dordrecht: Kluwer.
- Spalek, A. (2014). *Verb Meaning and Combinatory Semantics: A Corpus Based Study of Spanish Change of State Verbs*. Dissertation, Universitat Pompeu Fabra.

- Talmy, L. (1991). Path to realization: A typology of event conflation. In *Annual Meeting of the Berkeley Linguistics Society, Vol. 7(1)* (pp. 480–519).
- Talmy, L. (1996). Fictive motion in language and ‘ception’. In P. Bloom, M. A. Peterson, L. Nadel & M. F. Garrett (Eds.), *Language and Space* (pp. 211–276). Cambridge, MA: MIT Press.
- Tatevosov, S., & Ivanov, M. (2009). Event structure of non-culminating accomplishments. In L. Hogeweg, H. de Hoop & A. Malchukov (Eds.), *Cross-linguistic Semantics of Tense, Aspect, and Modality* (pp. 83–130). Amsterdam: Benjamins.
- Tenny, C. (1994). *Aspectual Roles and the Syntax-Semantics Interface*. Dordrecht: Kluwer.
- Vendler, Z. (1957). Verbs and times. *Philosophical Review*, 66, 143–160.
- Wright, T. A. (2014). *Strict vs. Flexible Accomplishment Predicates*. Dissertation, University of Texas at Austin, TX.