Verbs of letting in Germanic and Romance languages

A quantitative investigation based on a parallel corpus of film subtitles

Natalia Levshina

F.R.S.–FNRS, Université catholique de Louvain (Belgium)

This study compares eleven verbs of letting in six Germanic and five Romance languages. The aim of this paper is to pinpoint the differences and similarities in the semasiological variation of these verbs, both across and within the two language groups they represent. The results of a Multidimensional Scaling analysis based on a parallel corpus of film subtitles show that the verbs differ along several semantic dimensions, such as letting versus leaving, factitive versus permissive causation, as well as modality and discourse function. Although the main differences between the verbs lend themselves to a genealogical interpretation (Germanic vs. Romance), a distributional analysis of constructional patterns in which the verbs occur reveals that these differences are in fact distributed areally, with a centre and a periphery.

Keywords: letting, parallel corpora, film subtitles, Multidimensional Scaling, Correspondence Analysis, distributional semantics, Romance/Germanic

1. Introduction

The aim of this paper is to compare eleven verbs of letting in six Germanic and five Romance languages. The Germanic verbs are the Danish *lade*, Dutch *laten*, English *let*, German *lassen*, Norwegian *la* and Swedish *låta*. The Romance verbs are the French *laisser*, Italian *lasciare*, Portuguese *deixar*, Romanian *a lăsa* and Spanish *dejar*. These verbs share the sense of 'let', which is illustrated in Example (1):

(1) a. EN Let my people go… (King James Version, Exodus 8:1).
     b. DE Lass mein Volk ziehen…
     c. DK Lad mit folk gå…

d.   ES   Deja ir á mi pueblo…
e.   FR   Laisse aller mon peuple…
f.   IT   Lascia andare il mio popolo…
g.   NL   Laat mijn volk trekken…
h.   NO La mitt folk fare…
i.   PT   Deixa ir o meu povo…
j.   RO   Lasă-l pe poporul meu să plece…
k.   SV   Låt mitt folk gå…

The Germanic verbs come from Proto-Germanic *lētan*, via the extended form *lēd-*, and are also related to Gothic *lats* ("sluggish, weary") and Latin *lassus* "weary". The primary sense of these verbs, according to the OED, is "to let go through weariness, to neglect" (OED, 1970: 211). The Romance cognates — the verbs *laisser*, *lasciare*, *dejar*, *deixar* (Old Portuguese *leixar*) and *a lăsa* — originate from the Latin etymon *laxare* "to loosen, slacken", which developed from *laxus* "slack, loose".[1] However, some contend that the Romance verbs have been partly derived or influenced by the Germanic verbs (Diez and Donkin, 1864: 268–269).

This paper aims to pinpoint the differences and similarities in the semasiological variation of these verbs, both across and within the two language groups they represent. The main research questions of this study are as follows:

1. What are the most important common dimensions of variation of these verbs, and which are language and language-group specific? What are the cross-linguistic differences in the form-meaning mapping with regard to these dimensions?
2. Are these differences and similarities determined by the genealogical relationships between the languages? Are there patterns that can be explained by geographic proximity and language contact?

The languages selected for this study represent a convenient sample for comparison of genealogical and areal patterns, representing two different language groups in the Indo-European family and spoken in geographically contingent areas with tight economic, political and cultural links that have existed for many centuries (e.g. Haspelmath, 2001).

The paper also proposes a methodological innovation. Since these verbs share many similar functions, as will be demonstrated below, the cross-linguistic differences are often a matter of degree. For example, most Germanic verbs of letting can

---

1.   The initial *d-* in the Spanish, Portuguese and some other cognates has been a vexing problem for etymologists for centuries. Although there are different accounts, e.g. that it originates from *de-laxare* or from the contamination of *laxare* by another verb, such as *dare* (Malkiel, 1968: 220), *laxare* seems to be the closest Latin etymon.

express the sense 'to leave', e.g. the German *Warum hast du mich allein gelassen?* "Why have you left me alone?"; however, this usage seems to be less frequent and productive than in Romance languages. This intuition, which was first based on dictionary information, is supported by a small quantitative analysis of the parallel corpus of subtitles used in this study (see Section 3 for more information). This analysis shows that the relative frequency of the 'leaving' sense in the Germanic verbs of letting is from 0% (English *let*) to 8.6% (German *lassen*), whereas in the Romance languages it fluctuates from 21.3 % (Portuguese *deixar*) to 32.1% (Italian *lasciare*). From this it follows that one needs corpus counts to pinpoint the quantitative differences, which is not easy for two main reasons. First, many relevant semantic distinctions, as will be shown below, are not discrete. Second, many senses contain sub-extensions, which makes it difficult to choose the optimal level of semantic granularity *a priori*. As a result, it is very difficult to come up with an etic grid based on discrete semantic features that would be similar to the one used in semantic typology of kinship terms (cf. Evans, 2010).

In light of these limitations and challenges, it seems more appropriate to perform a bottom-up quantitative analysis. This paper presents two analyses of this kind. The first one employs token-based probabilistic maps created with the help of Multidimensional Scaling based on the translations of letting situations in a multilingual parallel corpus (e.g. Wälchli and Cysouw, 2012). This approach allows for the identification of salient dimensions of variation and for the exploration of functional overlap between the verbs. The second method is based on the ideas of distributional semantics, which go back to Apresjan (1966). In this approach, the distribution of the constructional patterns in which a verb occurs serves as an operationalization of the verb's semantic functions. A comparison of the distributional profiles of the verbs provides information about the semantic similarities and differences between the words. In addition, Simple Correspondence Analysis is used to visualize and explore the associations between the constructional patterns and individual verbs of letting. To the best of my knowledge, these approaches have not been applied to contrastive semantics before.

Yet another innovative aspect of this paper is the type of data. The analyses are based on a multilingual parallel corpus of film subtitles. The paper also deals with the potential problems of 'translationese' by testing the results of the study at the quantitative and qualitative levels against a comparable English corpus of soap opera scripts (Davies, 2012–) and several other corpora.

The structure of the remaining part of the paper is as follows. Section 2 introduces the concept of letting, outlining the main dimensions of its variation. Section 3 introduces the multilingual corpus of subtitles and deals with the caveats related to this kind of evidence. The quantitative analyses are presented in

Section 4. Section 5 offers a discussion of the results and a hypothesis for future diachronic research.

2.   Multidimensional conceptual structure of letting

LET is a highly complex semantic category,[2] which varies along numerous dimensions, cross-cutting such grammatical and conceptual domains as causation, modality, voice and negation. In Talmy's (2000) theory of force dynamics, letting is a conceptual configuration that involves an Agonist — a participant with an intrinsic tendency towards motion or rest — and an Antagonist — who fails, deliberately or not, to override the intrinsic tendency of the Agonist. Consider Example (2):

(2)   The police **let** the protesters take the buildings.

In this context, *the protesters* are the Agonist with an intrinsic tendency towards motion (i.e. taking the buildings), whereas *the police* are the Antagonist who does not override the Agonist's intrinsic tendency.

This section presents a tentative list of ten main dimensions relevant for onomasiological and semasiological variation of the selected verbs of letting, as well for pragmatic differences between them. These dimensions were identified on the basis of previous research and various lexicographic sources. Some of the dimensions contain discrete values, similar to distinctive features in the compositional approach — e.g. *bachelor* [+male, +adult, -married] — but most represent a continuum. In addition, some dimensions contain subdimensions, revealing a highly complex LET structure.[3] One of the main questions of this study is which dimensions reveal the main cross-linguistic differences in the semantics of the verbs.

Dimension 1. The intrinsic tendency of the Agonist towards motion or rest

In Talmy's (2000) theory of force dynamics, the intrinsic tendency of the Agonist can, in a broad sense, be towards motion or rest. Example (3) contains an Agonist (the thief) with the intrinsic tendency towards motion (escaping):

(3)   The detective **let** the thief escape again.

---

2. Henceforth, capitalized words without italics refer to semantic concepts.

3. The dimensional approach, which is similar in spirit to the multidimensional approach in Cognitive Semantics (Geeraerts, 1998), differs from the latter in that it does not centre on the most experientially basic and simple schema (the prototype), which is elaborated by the extensions. In fact, the verbs under consideration do not share a common prototype, diachronically or synchronically (Soares da Silva, 2007).

In (4), the intrinsic tendency of the Agonist (the illegal immigrants) is towards rest (staying in the United States):

(4)   The new policy **allows** young illegal immigrants to stay in the United States.

It seems that all verbs under consideration can express these two force-dynamic configurations. The intrinsic tendency of the Agonist plays an important role in a few distinctions presented below.

Dimension 2. Onset or extended letting

Another distinction proposed by Talmy (2000) is that between onset and extended letting. Onset letting means that the Antagonist stops previous impingement, as in the earlier example from the Old Testament, repeated here as (5). The Antagonist (the Pharaoh) is told to release the Israelites (the Agonist) from their captivity. This example contrasts with (6), an example of extended letting, in which the Antagonist (the addressee) is not supposed to intervene at all.

(5)   **Let** my people go.

(6)   **Let** it be.

According to Egan (2008: Section 6.2.3), this distinction between onset and extended letting is relevant for explaining the differences in the use of the constructions *allow* + *to* V and *let* + V: The former tends to expresses a removal of impingement, whereas the latter more frequently involves extended letting.

A combination of Dimensions 1 and 2 is crucial for describing the difference between French *laisser* and Romanian *a lăsa*, as well as for the other Romance cognates. According to Soares da Silva (2003), these two verbs do not normally express the meaning 'let go', which involves onset letting (Dimension 2) with a dynamic Agonist (Dimension 1). Instead, French uses *lâcher*, whereas Romanian uses *a da drumul*.

Dimension 3. Domain of letting

This is yet another distinction made in Talmy's (2000) theory of force dynamics. Letting can occur in different domains, most importantly, physical, as in (7), and psychosocial, as in (8).

(7)   The farmer **let** his dog loose.

(8)   Mom won't **allow** me to go to the party.

In psychosocial letting, the Antagonist is a socially acknowledged authority to which the Agonist has to conform. As Soares da Silva puts it, "[i]n fact, what the Antagonist (the permitter) does, in exerting its authority, is not so much cause the action as legitimize it socially and morally" (Soares da Silva, 2007: 180–181). The meaning of permission is not particularly prominent in the verbs that are being compared in the present paper. Other verbs, such as *allow*, *permit* and *authorize*, seem to be more onomasiologically salient in the verbalization of this sense (cf. Viberg, 2009).

Dimension 4. Degree of semantic integration of the letting and effected event

Verbs of letting can express different degrees of semantic integration of the letting and effected, or permitted event. According to Givón's (1980) hierarchy of semantic binding, there is a continuum of semantic integration of two events, which is iconically related to the formal syntactic integration of the corresponding two clauses. The stronger the semantic binding, the more syntactically integrated the clauses are. In the Romance languages (e.g. Achard, 1996; Soares da Silva, 2004; Maldonado, 2007), there is a continuum of causative constructions, which can be represented in a very general form, as shown below (Figure 1).[4] The leftmost construction in Figure 1 implies the greatest independence of the letting and permitted events, while that on the extreme right involves the highest level of semantic binding. Not all elements of this continuum are present in all Romance languages. For instance, the VSV construction, where the second verb is an inflected infinitive that agrees with the Agonist (S), seems to exist only in Portuguese.
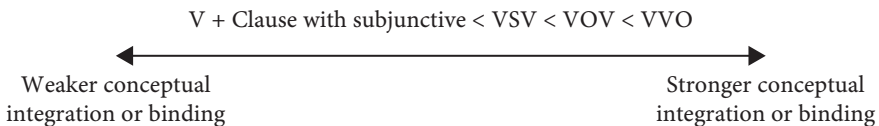
V + Clause with subjunctive < VSV < VOV < VVO

⟵─────────────────────────────────────⟶

Weaker conceptual   Stronger conceptual
integration or binding   integration or binding

Figure 1. Continuum of conceptual and syntactic integration in Romance languages. V: verb (auxiliary verb of letting or effected predicate), S: subject (Agonist), O: object (Agonist).

Compare (9a)–(9d), four Portuguese sentences (Soares da Silva, 2004: 297) that exemplify the four syntactic options, from left to right. All sentences can be translated as "Maria allowed the children to run."

(9) PT a. A   Maria deixou   que os miúdos corresem.
             the Maria let.PST.3SG that the children run.SBJV

---

4. An exception is Romanian, where the infinitive is demoted, and the subjunctive finite form is used instead. A demoted or non-existent infinitive is a typical feature of the Balkan Sprachbund (Friedman, 2006: 665).

b. A   Maria deixou   os miúdos correrem.
    the Maria let.PST.3SG the children run.INF.3pl

c. A   Maria deixou   os miúdos correr.
    the Maria let.PST.3SG the children run.INF

d. A   Maria deixou   correr os miúdos.
    the Maria let.PST.3SG run.INF the children
    "Maria allowed the children to run."

On the basis of corpus observations, which will be introduced below, one can conclude that the Germanic verbs of letting exhibit less variation in the degree of syntactic integration, occurring predominantly in the VOV construction in the main clause.[5]

Dimension 5. Permissive or factive causation
The difference between permissive and factive causation (Nedjalkov, 1976: Ch. 3) can be defined as the difference between letting and making X do Y.[6] In many Germanic languages the verbs of letting can express factive causation, as in (10):

(10) NL Ik **liet**   mijn huis   schilderen.
         I   let.PST.SG my   house paint.INF
         "I had my house painted."

In many instances, the interpretation of the construction in terms of making or letting can be made only on the basis of context (e.g. see a list of conceptual and syntactic cues for disambiguation of Swedish *låta* in Rawoens and Egan, 2013). For example, Example (11) in German can be interpreted as either an instance of non-interference or as an example of indirect causation (e.g. if the speaker makes someone roll the stone or exploits the force of gravity):

(11) DE Ich **lasse**   den   Stein rollen.
         I   let.PRS.1sg the.ACC   stone roll.INF
         "I let the stone roll/have the stone rolled."

---

5. The degree of semantic integration may be also expressed by the case form of the Agonist. For example, the preposition *door* "by" marks the agentive Causee in Dutch causative constructions with *laten*, whereas *aan* "to" marks the Causee in the role of an experiencer of mental events (Kemmer and Verhagen, 1994). These two prepositions indicate a lower degree of semantic binding than the bare NP (the default). Similar variation in case marking is also possible in Romance languages (e.g. Soares da Silva, 2004).

6. This dimension is in fact related to the previous dimension of semantic binding. According to Givón (1980), letting involves greater Agonist autonomy than does making. However, since this distinction involves systematic formal variation (e.g. *make* + V vs. *let* + V, *faire* + V vs. *laisser* + V*,* etc.), it is listed here as a separate dimension.

English *let* does not perform this function — at least, not anymore (see Lowrey, 2013). An exception is the construction *let X know Y*, which can also be interpreted as an intermediate case between letting and making. Such in-between uses with mental verbs are also observed with the Romance verbs of letting, which do not normally function as factitive causative auxiliaries, either.

An interesting question is whether the factitive causation function in Germanic languages developed as an extension of the permissive function, or whether the former emerged simultaneously (or even prior) in the auxiliaries (see an overview in Rawoens, 2013). For instance, the Swedish auxiliary *låta* has been used to express both factitive and permissive causation since the earliest available texts (*ibid.*). Unfortunately, there is insufficient historical evidence to answer this question conclusively.

Dimension 6. Polarity

The distinction between permissive and factitive causation is cross-cut by another dimension, the presence or absence of negation. While letting verbs in positive contexts (e.g. *She let him go*) express non-impingement, they express coercion in negative contexts (Nedjalkov, 1976:24). For instance, the sentence *She didn't let him go* means that she made him stay. Even though making and not letting describe the same referential situations, the construals of these situations are different. Unlike positive constructions with factitive causation verbs (e.g. *make*, *have*, *cause*), negative expressions with *let* presuppose some potential world where non-occurrence or cessation of impingement might or should take place.

Dimension 7. Letting vs. leaving

This dimension is based on a subtle difference in construal that introduces a whole new family of semantic extensions. In constructions of the type *X lets Y do Z*, the Antagonist *X* is the primary Trajector, and the Agonist *Y* is the primary Landmark. However, the Agonist *Y* is also the Trajector of the action or process *do Z*, which represents the permitted event (cf. Langacker, 1991:409–410).

The Agonist may also be construed as a Landmark only, a case of leaving, rather than letting. Typically, in such cases the Agonist's intrinsic tendency is towards rest (cf. Soares da Silva, 2007:176). The closely related categories LET and LEAVE can be distinguished with the help of syntactic cues: LET is represented by the periphrastic causative with a second verb, which specifies the event that is brought about by the Agonist, whereas LEAVE includes only one predicate. Compare Examples (12a) and (12b), which demonstrate letting and leaving, respectively:

(12) a. After months of abusive behaviour, he finally **let** his family live in peace.
      b. After months of abusive behaviour, he finally **left** his family in peace.

In contemporary English, these construals are represented by two different verbs, *let* and *leave*.[7] The Romance verbs of letting, in contrast, have a rich variety of extensions of this type. German and Dutch seem to be intermediary, with LEAVE expressed by prefixal verbs (e.g. German *verlassen* and Dutch *verlaten*) or by the verbs of letting in fixed expressions, such as the German *Lass mich in Ruhe!* "Leave me in peace!".

In instances of LEAVE, the focus is no longer on the action or state of the Agonist; instead, it shifts to the actions of the Antagonist (Soares da Silva, 2007). This reversal is especially evident in such subschemas as LEAVE_LOCATION and ABANDON, which contain a dynamic Antagonist. The subschema LEAVE_LOCATION typically involves physical motion, as in Example (13):

(13) PT O comboio **deixou**   a   estação
          the train   leave.PST.3SG the station
          às   12 h.
          at.ART 12 o'clock
          "The train left the station at 12 o'clock."

The subschema ABANDON is similar, but the principal domain is psychosocial. In Example (14), the focus is on social and psychological distance, rather than on the physical distance of the Antagonist from his ex-family.

(14) PT Ele **deixou**   a   mulher.
          he leave.PST.3SG the wife
          "He left his wife."

Notably, these extensions are not typically found in connection with the French *laisser* and Romanian *a lăsa*, in contrast with the other Romance cognates (Soares da Silva 2003). Instead, French typically uses *quitter* "to leave, quit, part" and *abandonner* "to give up, abandon", whereas the Romanian equivalent is *a părăsi* "to abandon, neglect".

The basic schema of LEAVE can be elaborated further. One of the examples is TRANSFER, when the Recipient is added to the scene. The Agonist then becomes the Theme, and the Antagonist becomes the Giver. An example is the concept of BEQUEATH:

---

7. The Old English verb *l'tan* had both the meaning of 'to let, allow' and 'to leave behind, leave undone' (The American Heritage Dictionary, 1970: 1526). This function seems to have survived only in the idiomatic expression *let alone*, which is equal to *leave alone*.

(15) FR Mais ce que nous **laisserons**   aux
        But   it that we   leave.FUT.1PL to.ART
        générations futures   depend   de
        generations future.PL   depend.PRS.3SG   of
        notre volonté   commune.
        our   will   common
        "But what we bequeath to future generations will depend on our
        common will." (from http://opus.lingfil.uu.se/MultiUN.php, last
        accessed 17 December 2014)

The Germanic verbs possess the subschema TRANSFER only in a few specific contexts.[8] The domains of giving and letting are very closely related, as can be seen in the polysemy patterns of verbs in different languages. For example, in Slavic languages the verbs of giving also play the role of causative auxiliaries; for example, the Russian *Daj mne pospat'!* "Let me sleep!" contains *daj*, the imperative form of *dat'* "to give (perf.)" (see Newman, 1996 and von Waldenfels, 2012 for more examples).

Yet another semantic modification of LEAVE is observed when the Agonist ceases to be an individuated participant and becomes an abstract entity, as in (16):

(16) IT Vuole   **lasciare** gli studi   universitari.
        want.PRS.3SG leave.INF the studies university.ADJ
        "He/she wants to leave college."

Moreover, the abandoned object can be an activity expressed by the verb. This extension — cessation of an activity — is observed in the Ibero-Romance languages (Soares da Silva, 2007), as in Example (17), from Portuguese:

(17) PT Ele **deixou**   de   fumar.
        he let.PST.3SG PREP smoke.INF
        "He gave up smoking." (Soares da Silva, 2007: 184)

Dimension 8. The voice (and coreferentiality)

In the present study, the voice denotes the construal of the Antagonist as a potentially affecting or affected entity. In Examples 1–17, the Antagonist was the entity that could potentially influence the actions or state of another entity (the Agonist). However, the Antagonist and Agonist may also be coreferential with each other or with other participants in the causation chain. Consider Example (18), in Dutch:

---

8. E.g. British English *let* has a specific sense 'lease', as in to *let a house*.

(18) NL **Laat**   je   niet misleiden.
   let.IMP REFL.2 not mislead.INF
   "Don't be misled."

In this sentence, the Antagonist is coreferential with the semantic object of the verb *misleiden* "mislead" (*je* "you"). Therefore, he/she is also a potentially affected entity. Such constructions demonstrate that the Antagonist is considered responsible for prevention of something bad that happens or may happen to him or her (cf. Loewenthal, 2003 for Dutch). Thus, although the Antagonist is in fact an affected entity, it still has some properties of an Agent, most importantly, the control over the effected event.

  Letting verbs in reflexive constructions can also be used to convey the semantics close to the middle voice. The Antagonist in such cases has some properties that facilitate the action specified by the effected predicate to be carried out (Davidse and Heyvaert, 2003). Consider Example (19), from German:

(19) DE Das **lässt**   sich   schon machen.
   that let.PRS.3SG REFL.3 indeed make.INF
   "That's manageable."

One can also find examples of letting constructions conveying completely passive meanings, as in Example (20), from Spanish:

(20) ES Además, se   pueden   dañar   si
   also,   REFL.3 can.PRS.3PL damage.INF if
   se **dejan**   caer.
   REFL.3 let.PRS.3pl fall.INF
   "They can also be damaged if dropped." (from http://help.vertu.com/
   downloads/user-guides/Constellation_C/RHV8/RHV8_es_Original.pdf
   [last accessed 17 December 2014])

The Antagonist has no agentive properties and has no influence on the situation. In such cases, the participant in fact ceases to be a true Antagonist.

Dimension 9. Modality

Letting is closely related to modality. If X lets Y do something, this means that Y can or may do something. Some letting verbs can be used in a purely modal function to express the speaker's desire, encouragement, surprise, etc. One example is the hortative use of *let* and its Germanic cognates, as in (21):

(21) **Let**'s pray!

Another use is optative, as in (22), whereby the speaker expresses a wish:

(22)   **Let** no one be dead today.

The hortative and optative uses are examples of deontic modality. Examples of epistemic modality can be found, as well. For instance, in (23), the Dutch *laten* is used to express the speaker's surprise and disbelief:

(23)   NL Hij zei dat hij tien Duvels zou   drinken,
            he said that he ten Duvels would.SG drink.INF
            en **laat**   hij het nu   doen   ook!
            and let.PRS.SG he it   now do.INF too
            "He said that he would drink ten Duvels, and imagine, he did that!"

Dimension 10. Discourse function

In some contexts, the original conceptual meaning of verbs of letting is bleached, and a verb of letting may be used as a part of a discourse marker, most commonly for making metalinguistic comments. Such discourse markers may introduce a minor break in discourse, as in (24), or clarification and elaboration, as in (25). Another function is signalling the speaker's intention to perform some action, as in (26):

(24)   **Let** me think…

(25)   **Let** me put it differently…

(26)   **Let** me help you.

Such imperative uses of *let* are labelled by Wierzbicka (2006: 183–202) as letting of 'cooperative dialogue' and 'cooperative thinking'. In these examples, the speaker formally asks the hearer for permission, but the hearer's permission is in fact not necessary. The purpose is to inform the hearer of the speaker's intentions in a cooperative and non-imposing way. The hortative *let's*, which was mentioned in Example (21), can be regarded as another instance of interpersonal use. Although possible in the other languages reviewed in the present paper (see Viberg, 2009 for Swedish), such cooperative uses seem to be particularly prominent in English (*ibid.*).

     Section 4 will investigate which of the dimensions of semantic and pragmatic variation described above play prominent roles in describing the cross-linguistic variation in the use of the eleven verbs of letting. Before this, a description of the data and methodology is presented in Section 3.

3.   Data and method

3.1 Parallel corpus

This study is based on a multilingual parallel corpus of film subtitles. Subtitles represent a new source of data in typology and contrastive linguistics. They have several advantages over traditional written parallel corpora. First, they are freely downloadable from numerous online repositories for a large variety of languages. Second, they are the closest to spoken language of all available multilingual parallel corpora, which mostly represent legal, political or religious texts, as well as works of fiction. In psycholinguistic studies, film subtitles in the original language have been shown to be a reliable source of lexical norms, sometimes outperforming other sources (Keuleers *et al.*, 2010). Similar to real life, the viewer of a film is immersed in a situation where multimodal (visual and audial) cues are present. As in everyday communication, the linguistic structures are often elliptical, relatively simple and compact. Consider (27), an example from *The Hobbit: An Unexpected Journey*.

(27)   …
646
00:51:27,880 — > 00:51:32,920
<i>For always evil will look to
find a foothold in this world.</i>
647
00:51:39,440 — > 00:51:42,603
Not good. Not good at all.
648
00:51:50,040 — > 00:51:51,326
Eww.
649
00:52:06,760 — > 00:52:09,081
Oh, no. Sebastian.
…

More complex structures can be represented, as well. A long sentence can be represented by several subsequent captions. The format of the files is SubRip (.srt), which includes information about the time when a caption should appear on screen and disappear, as well as optional formatting information.

However, there are two main concerns related to this kind of evidence. The first problem is medium-specific. Professional subtitle translators have rigid rules to follow with regard to the maximum length of a line, the time that a caption should

stay on screen, etc. (e.g. Deckert, 2013). This may influence the constructional choices that a creator of subtitles has at his or her disposal.

The second problem is more general, and concerns using translation corpora in linguistic research. Parallel corpora are very common in contrastive linguistics and, more recently, in typology, because they allow for the quantification of equivalence of expressions in two or more languages. However, there have been some concerns about possible deviations of 'translationese' from the original language. It seems that the distorting effect of translationese may differ from one linguistic phenomenon to another. For example, Johansson (2007: 32–33) demonstrates that Norwegian paraphrases of English tag questions (the equivalent construction does not exist in Norwegian), such as *ikke sant* "not true", are much more frequent in texts translated from English than in original Norwegian texts. In such cases, analysis of a translated text instead of an original one might produce a distorted picture. Some researchers, however, report evidence of structural similarity. For instance, Lefer and Cartoni (2013) show that the use of the English prefix *un-* and suffix *-less* does not vary significantly across original and translated English texts. This may happen because translators tend to adjust to the target language, as the evidence in Van Olmen's (2011) study of imperatives in English and Dutch suggests.

From this conflicting evidence it follows that the effect of translationese is an empirical question. To see whether these characteristics of the data represent a potential problem in the analysis of letting verbs, two experiments were conducted. First, the relative frequency of *let* was obtained from the Corpus of American Soap Operas scripts (Davies, 2012–),[9] which contains scripts from ten U.S. soap operas. In addition, relative frequencies were extracted from the spoken part of the Corpus of Contemporary American English (COCA) (Davies, 2008–), which contains transcripts of unscripted conversations from diverse TV and radio programmes. Finally, two relative frequencies of *let* were computed on the basis of subtitles of films of different genres. The first relative frequency is based on the author's own corpus of twenty films in original English downloaded from www.opensubtitles. org [last accessed 17 December 2014]. The second number is based on subtitles of twenty films translated into English from French, downloaded from the same repository. French was chosen because it is relatively easy to find French films translated into English, which is not the case for most other languages. The graph in Figure 2 displays the relative frequencies (per ten thousand words/tokens).

It seems that the frequencies of *let* in the soap operas and subtitles are comparable. One can also see that the subtitles and soap opera scripts have higher proportions of *let* than TV and radio transcripts from the spoken segment of the COCA. This may be due to the number of instances of *let's* and letting of cooperative

---

9. URL http://corpus2.byu.edu/soap/ [last accessed 15 December 2014].
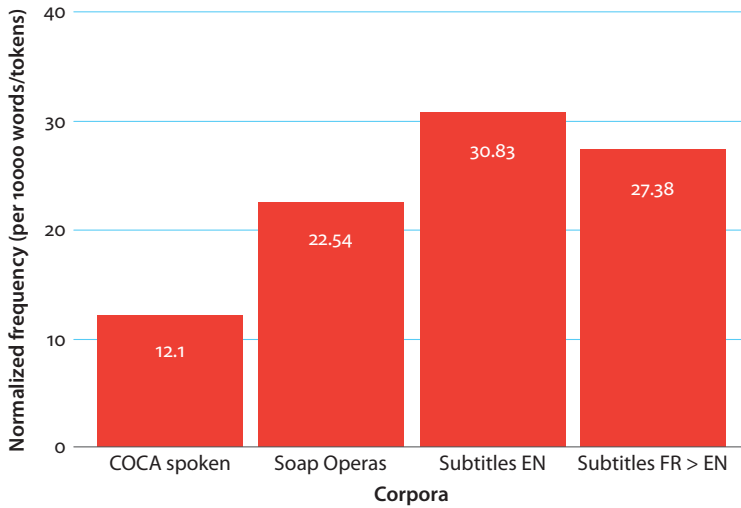
Figure 2. Relative frequencies (per 10,000 words) in different corpora.

dialogue, stemming from the fact that films and soap operas involve a great deal of action and interaction. The frequency of *let* in the translated subtitles (27.38 per 10,000) is somewhat lower than in the original subtitles (30.83 per 10,000). One may wonder whether this difference is due to the fact that *laisser* has relatively low frequencies in the original French subtitles (see Table 1). The difference between the proportion of *let* in the original English subtitles and that of *laisser* in the original French subtitles is statistically significant ($\chi^2 = 24.6491$, df = 1, $p < 0.001$). The difference between the original and translated English subtitles, however, is not statistically significant ($\chi^2 = 2.4348$, df = 1, $p = 0.119$). Thus, we do not find sufficient evidence of a potential influence of the frequency of *laisser* in the French originals on the use of *let* in the translations. This result supports Van Olmen's (2011) conclusion that translators make adjustments to the target language.

Table 1. Frequencies of *let* and *laisser* in different corpora of subtitles.

	let in original subtitles	*let* in translated English subtitles	*laisser* in original French subtitles
Frequency	493	286	202
Normalized per 10,000	30.83	27.38	20.37
Total corpus size (tokens)	159926	104445	99182

The second test concerned qualitative differences in the distribution of *let* in the subtitles and in the corpus of American soap operas. We constructed a profile vector with relative frequencies of the constructional patterns of English *let* in soap

opera scripts from the Corpus of American Soap Operas (Davies, 2012–). These patterns are introduced in Section 4.3. A random sample of 100 observations was collected, which were coded for their constructional features. The frequencies of the features are very similar to those in the English subtitles. There is a strong positive correlation between the vectors of distributional features of *let* in both corpora (Spearman's $\rho = 0.896$, $p < 0.001$). This is the highest correlation that can be found if one computes all correlations between all pairs of distributional vectors of the letting verbs (see Section 4.3). Although other languages need to be investigated as well, the working hypothesis of this paper is that subtitles reliably represent the general patterns in the use of verbs of letting.

3.2 The data set

The data come from a multilingual parallel corpus for typological comparisons (ParTy) compiled by the author.[10] For this case study, subtitles of five films of different genres and original languages were selected that represent the eleven languages given in Section 1. The films are listed in Table 2. The genres were determined according to the classification used by the International Movie Database.[11] All files were downloaded from the online databases at www.opensubtitles.org and www.subscene.com [last accessed 17 December 2014] and have the .srt format. These files were inter-aligned with the help of the subtitle alignment software *sub-align*, written by Jörg Tiedemann.[12]

The total size of the corpus is 294,000 tokens. All occurrences of the letting verbs in all eleven languages were extracted semi-automatically, by searching for all morphological forms of the verbs, and all occurrences were included, including auxiliary and phrasal verbs. Prefixed verbs, e.g. the Dutch *verlaten* "to leave, abandon" and German *nachlassen* "to decline, slack", were excluded. Homonymous uses of Swedish *låta*, with the meaning "to sound, appear", which may be historically related to the meaning of letting as releasing (Viberg, 2009), were not taken into account. For every detected exemplar of a letting verb, the corresponding sentences in the other languages were extracted with the help of the alignment information and checked manually. The final data set consisted of 280 multilingual contexts, for which at least one translation of the sentence contained a letting verb.

---

10. See more information at http://www.natalialevshina.com/corpus.

11. URL http://www.imdb.com [last accessed 19 June 2014].

12. The software (in Perl) is available at https://bitbucket.org/tiedemann/subalign [last accessed 17 December 2014]. The author of this paper is personally grateful to Jörg Tiedemann for his help.

Table 2. Films used in the study.

Original title	Title in English	Year	Original Language	Genre
Le fabuleux destin d'Amélie Poulain	*Amélie*	2001	French	Comedy, romance
Avatar	*Avatar*	2009	English	Action, adventure, fantasy
Das Leben der Anderen	*The Lives of Others*	2006	German	Drama, thriller
El laberinto del fauno	*Pan's Labyrinth*	2006	Spanish	Drama, fantasy, war
Wo hu cang long	*Crouching Tiger, Hidden Dragon*	2000	Mandarin Chinese	Action, drama, romance

For statistical analyses, a matrix was created, with 280 rows and eleven columns, each column corresponding to one of the verbs. If the letting verb was used in a given context, the cell contained 'Yes'. If no letting verb was used in the translation, the cell contained 'No'. If a translation was missing or was erroneous, the cell contained a missing value, coded as 'NA'. A snapshot of the matrix in Excel is shown in Figure 3. The corresponding contexts were stored in a separate file.

For example, the first context (Row 2) in Figure 3 corresponds to the sentences shown in (28):

(28)   DE   Entspannen Sie sich und denken Sie einfach an gar nichts.
   DK   Bare slap af og **lad** sindet gå blank.
   EN   Just relax and **let** your mind go blank.
   ES   Relájate y pon tu mente en blanco.
   FR   Détendez vous et faites le vide.
   IT   Rilassati e svuota la mente.
   NL   Ontspan en maak je hoofd leeg.
   NO   Bare slappe av og nullstill hodet.
   PT   É só relaxar e **deixar** tua mente vazia.
   RO   Relaxează-te şi goleşte-ţi mintea.
   SV   Slappna av och töm huvudet.

This matrix served as input for the statistical analyses presented in Sections 4.1 and 4.2. All analyses were performed with the help of R software (R Core Team, 2013) with add-on packages (cluster, ggplot2, MASS and smacof).

The languages exhibit a lot of variation in the expression of letting. In 142 contexts — representing more than a half of all observations — only one language used a verb of letting. Only five contexts out of 280 contained letting verbs in all languages. This variability is due to the high degree of creativity that is typical of this register (cf. Levshina, 2015), and provides indirect evidence against

Figure 3. The matrix in Excel with languages (columns) and multilingual contexts (rows).

translationese effects. Among the Germanic languages, the Danish *lade* has the highest absolute frequency (95) and the Swedish *låta* the lowest (only 45). As for the Romance languages, the Spanish *dejar* occurs the most frequently (75), whereas the Italian *lasciare* is the least frequent (53). Thus, the Germanic verbs have a greater dispersion of counts in comparison with the Romance ones.

4. Statistical analyses

4.1 Probabilistic map of letting

This subsection identifies the most important semantic dimensions of variation on the basis of quantitative analysis and compares the results with the dimensions described in Section 2. The main tool is a probabilistic map of letting exemplars, which was created with the help of Multidimensional Scaling. The map represents the distances between all pairs of multilingual exemplars in the data set. The distances were based on the Gower similarity scores for categorical data. If two exemplars (rows in Figure 3) had the same values 'Yes' or 'No' across all languages,

the distance was zero. The smaller the overlap of the values, the greater the distances between the exemplars. As a result, the contexts for which most sentences use a letting verb are located in the centre, and the exemplars for which only one language uses a letting verb form the periphery. According to the stress values, the optimal representation was a two-dimensional solution (see the scree plot in Figure 4). Adding more dimensions does not lead to a substantial improvement in representation.
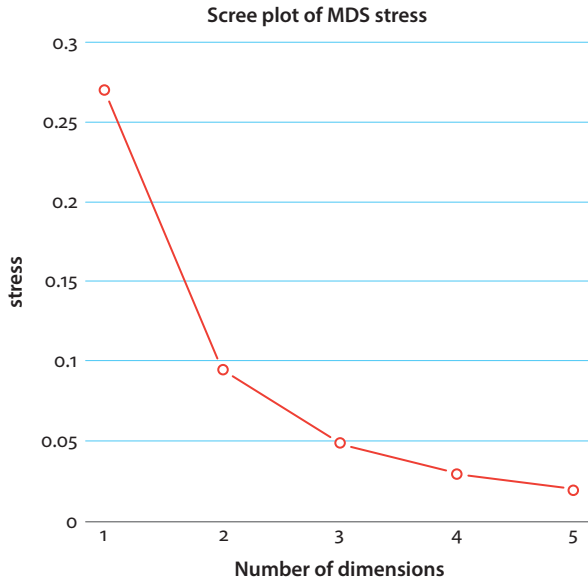
Scree plot of MDS stress

Figure 4. A scree plot of stress values in SMACOF MDS solutions depending on the number of dimensions. The line 'elbows' at two dimensions.

The resulting two-dimensional solution (stress=0.094) with some individual points (Figure 5) will be discussed below.

Note that the semantic dimensions that one expects to find need not correspond to two automatically obtained dimensions of the MDS; rather, they have to correspond to the patterns and clusters that emerge when one compares the semantics of the points on the map. An examination of the points in the middle of the map reveals that this area contains exemplars that express non-occurrence of impingement (i.e. extended letting); one such example is (29), which corresponds to Exemplar 223 in Figure 5. This area also contains exemplars that express removal of blockage (i.e. onset letting); one such example is (30), which corresponds to Exemplar 109 on the map.
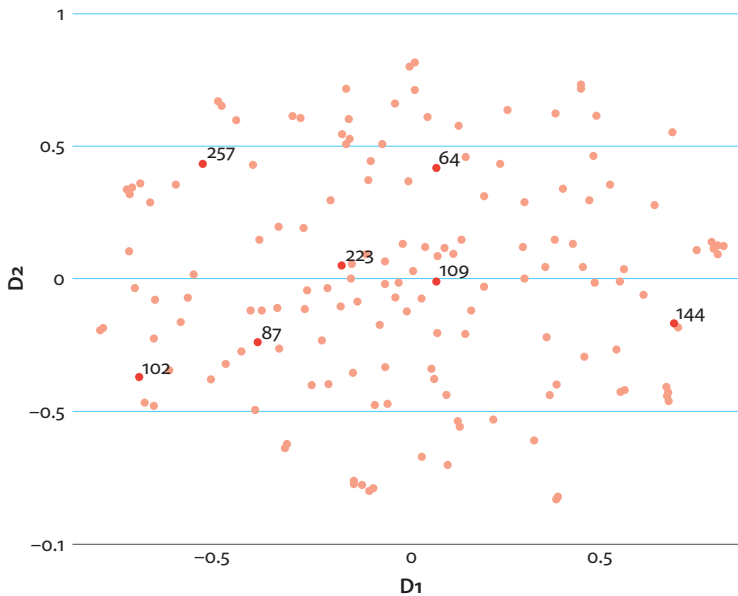
(29)   You can't **let** Li Mu Bai die! [Ex. 223]

Figure 5. The probabilistic space of letting exemplars.

(30)   Look: He **let** her go. [Ex. 109]

Notably, non-occurrence of impingement and removal of blockage are the senses that are considered prototypical instances of letting. For Talmy (2000), the proto-typical sense is the cessation of impingement, whereas Soares da Silva (2007) considers non-occurrence of impingement prototypical in contemporary Romance verbs. However, since both senses are central, it is difficult to conclusively determine which is more basic.

The area at the top of Figure 5 contains different instances of leaving; for example, leaving someone in a specific location, as in Exemplar 64, presented in (31):

(31)   FR   Et   moi, je vous **laisse**   ici… [Ex. 64]
           And me, I   you   leave.PRS.1sg here
           "I'll leave you here…"

The sense 'to quit, stop' is illustrated by Exemplar 257, presented in (32):

(32)   IT   Ho   **lasciato** l'addestramento   prima
           have.PRS. 1SG   leave.PTCP.PST.ART.training before
           del   tempo. [Ex. 257]
           of.ART time
           "I left the training early."

The lower portion of the map in Figure 5 contains examples for which verbs of letting are accompanied by another verb. Exemplars on the left-hand side include numerous instances of cooperative dialogue and thinking (e.g. Exemplar 87, presented in (33)), as well as hortative uses (e.g. Exemplar 102, presented in (34)):

(33) NL **Laat**   me zeggen… gemiddelde lengte.[Ex. 87]
          let.IMP me say.INF   average   height
          "Let's say… average height."

(34) **Let**'s both do one together. [Ex. 102]

Exemplars on the right-hand side include instances of factitive causation, as in (35):

(35) NL Ik **laat**   u   opsturen wat   wij nog aan
          I   let.PRS.SG you send.INF what we still at
          materiaal hebben. [Ex. 144]
          material   have.PRS.PL
          "I'll send you all the material we have."

To summarize, the main distinctions revealed by the map are as follows: letting vs. leaving, modal/cooperative vs. non-modal/non-cooperative uses, and factitive vs. permissive causation. These are therefore the most important semantic dimensions of the cross-linguistic variation of the verbs.

Let us now see how the verbs are mapped onto the semantic space. Figures 6 and 7 display the distributions of the Germanic and Romance verbs, respectively. Each black point represents one instance of a given verb of letting.

Several conclusions can be made on the basis of these figures. In general, the Germanic verbs tend to occupy the lower area, which corresponds to different types of letting, while the Romance verbs occupy the upper area, which corresponds to different types of leaving. The German and Dutch verbs are different from the other Germanic languages, also covering the region of factitive causation. The letting verbs in English, Danish and Norwegian cover the region containing cooperative and hortative uses. Notably, the Dutch *laten* looks to be an intermediary between the English *let* and German *lassen*, supporting the idea of the so-called 'Germanic sandwich', which puts Dutch between English and German with regard to a range of synchronic phenomena, from phonology to word order (cf. Hüning *et al.*, 2006). Another interesting detail is the similarity of the Romanian *a lăsa* and Portuguese *deixar*, the eastern-most and western-most Romance languages, respectively. Swedish occupies the most compact area, close to the centre, and without prominent peripheral functions.
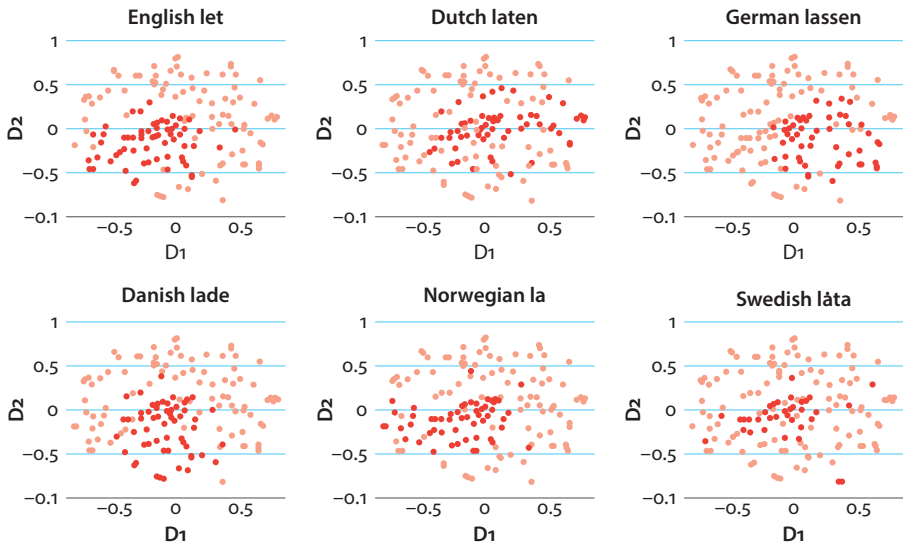
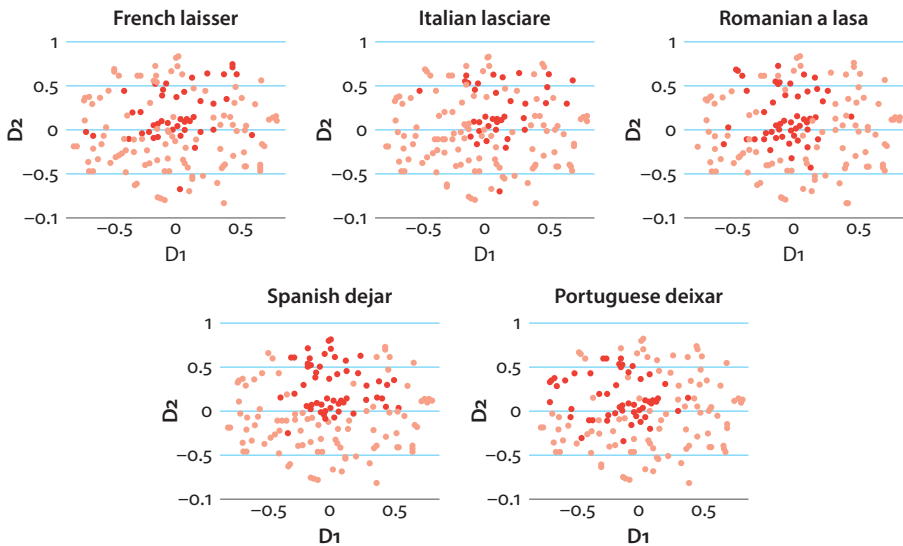**Figure 6.** Distribution of the Germanic verbs of letting on the semantic map.

Figure 7. Distribution of the Romance verbs of letting on the semantic map.

4.2 A hierarchical cluster analysis of the verbs

In Section 4.1, the verbs of letting were compared by looking at their semantic functions. This section offers a quantitative comparison based on a hierarchical cluster analysis. First, the Gower distances between all verbs were computed (see Section 4.1). These scores represent the proportion of overlapping values ('Yes' or

'No') between the verb columns in the matrix shown in Figure 3. The greater the proportion of overlapping values between a pair of verbs, the smaller the distance. Next, a hierarchical cluster analysis was performed. The resulting clustering tree is shown in Figure 8.
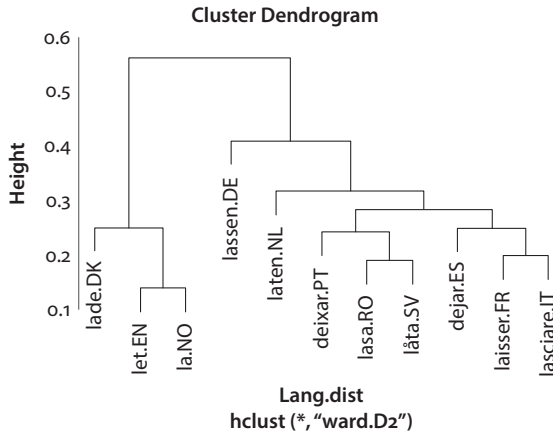

Figure 8. A hierarchical cluster analysis of the verbs of letting. Distance: Gower; Method: Ward.

Notably, the Romance and Germanic verbs of letting do not constitute separate clusters, as one might expect on the basis of the genealogical information. Instead, the German *lassen*, Dutch *laten* and Swedish *låta* belong to the large cluster on the right, which also contains the Romance verbs. Most of the lower-level genealogical subgroups are not supported, either. For example, Portuguese clusters together with Romanian and Swedish, whereas the other Ibero-Romance language, Spanish, is in the same cluster with French and Italian. English, a West Germanic language, is in a cluster with Danish and Norwegian, which are both North Germanic languages. These conclusions support the observations made in Section 4.1. Although the tree does display some genealogical patterns (e.g. French, Italian and Spanish cluster together), this is not systematic. This suggests that there could be other factors that explain these cross-linguistic differences, such as geographical proximity and language contact.

4.3 Distributional analysis based on constructional patterns

To examine the differences between the verbs in more detail, all examples in every language were coded for a number of constructional patterns. Constructional information serves here as an operationalization of semantic differences. This approach has been applied fruitfully in neo-structuralist approaches to semantics (cf.

Geeraerts, 2010). For example, Levin's (1993) verb classes are based on the syntactic alternations in which the verbs can participate. Alternatively, Levshina and Heylen (2014) use syntactic information, e.g. subcategorization frames and syntactic relationships, as well as information about the lexical neighbours, to identify relevant semantic classes of nouns and verbs for modelling near synonymy in Dutch. Although the constructions are not functionally identical in the languages under consideration, they are similar enough to enable a cross-linguistic comparison. The constructions are shown in Table 3. As one can see, the classification is quite coarse-grained. This is done for two reasons. First, there needs to be a sufficient number of tokens per type to draw robust conclusions. Second, more general patterns have higher chances of conveying similar semantics across this group of languages than more specific ones.

Table 3. Constructional patterns used for the distributional analysis.

Label	Explication	Example	Semantic correspondence
LET_Vintr	a letting verb as an auxiliary, followed by the infinitive form of an intransitive verb	*I let him go.*	permissive or factitive causation
LET_Vtr	the same as above, but the second verb is used transitively (i.e. has a direct object)	*I let him do it.*	more probable in facti-tive causation
LET_NP_ GOAL	a letting verb followed by an NP and an NP or adverb that specifies the goal	*He let her into the village.*	the Antagonist lets the Agonist (NP) into some place specified by a PP or adverb
LET_Refl	the Antagonist is coreferential with the Agonist or another argument of the second verb	DE *Lasst euch nicht täuschen!* "Don't be misled!"	affected Antagonist, often passive and middle voice
LET_ Clause	a verb of letting governs a clausal complement	ES *Deja que X hable.* "Let X do the talking."	non-interference, permission with low semantic binding between the letting and effected events
LET_NP	a verb of letting followed by an NP, without a second verb (optionally: another NP, adverb, adjective or PP)	RO *Las-o în pace.* "Leave her alone (in peace)."	leaving
LET'S	a letting verb in the imperative + 'we'/'us'	*Let's go!*	hortative function

Table 3.   (*continued*)

Label	Explication	Example	Semantic correspondence
LET_V	a verb of letting followed by another verb, no arguments	FR *Laisse tomber!* "Forget (it)!" DE *Lass sehen!* "Show (me)!"	miscellaneous
LET_ Prep_V	a verb of letting is followed by a preposition and an infinitive	PT *deixar de fumar* "stop smoking"	stopping or failing to perform the activity specified by the second verb
LET_V_ Prep_V	a verb of letting followed by another verb, preposition and a third verb (infinitive)	DK *Lad være med at opføre Dem sådan.* "Stop behaving like that."	stopping or failing to perform the activity specified by the verb after the preposition
LET	single letting verb without any arguments	NL *Laat maar zo.* "Never mind."	never mind, not interfere, forget
Other	other constructions	*Let go of him!*	miscellaneous

All occurrences of the letting verbs in the corpus were coded for the constructional patterns listed in Table 3. Note that, during the coding, the Romanian subjunctive form with the marker *să* was treated similarly to the infinitive in other languages. This was because the subjunctive in Romanian performs functions that overlap with those of the infinitive in the other languages. The relative and absolute frequencies of the patterns are displayed in Table 4.

A comparison of the frequencies of the distributional patterns leads to the following conclusions:

1. The Germanic verbs have a higher proportion of auxiliary uses (*LET_Vintr*, *LET_Vtr*, *LET_Refl*, *LET_V*, *LET_V_(Prep)_V*; associated with the semantics of letting) than lexical transitive uses (*LET_NP*; associated with the semantics of leaving); for the Romance verbs, the reverse holds. Among the Germanic verbs, the Swedish *låta* has the highest proportion of auxiliary uses (87%), whereas the German *lassen* occurs most frequently in the *LET_NP* pattern (19%), followed by the Dutch *laten* (14%). As for the Romance verbs, the French *laisser* has the highest proportion of auxiliary uses (59%), followed by the Romanian *a lăsa* (52%) and Italian *lasciare* (47%). The Spanish and Portuguese cognates have lower frequencies, 32% and 45%, respectively.
2. The English, Danish and Norwegian verbs have the highest proportions of hortative uses (*LET'S*), as one would logically expect based on the semantic map. The leader is the English *let*, with 40%. For the other Germanic verbs, the

Table 4. Absolute and relative (in italics) frequencies of constructional patterns.

Verb Pattern	lassen DE	lade DK	let EN	dejar ES	laisser FR	lasciare IT	laten NL	la NO	deixar PT	a lăsa RO	låta SV
LET_Vintr	27 *0.33*	30 *0.32*	23 *0.28*	18 *0.23*	19 *0.31*	17 *0.32*	27 *0.34*	31 *0.40*	23 *0.31*	20 *0.29*	21 *0.47*
LET_Vtr	11 *0.14*	20 *0.21*	17 *0.20*	3 *0.04*	9 *0.15*	2 *0.04*	27 *0.34*	15 *0.20*	8 *0.11*	15 *0.22*	17 *0.38*
LET_NP_GOAL	2 *0.03*	1 *0.01*	5 *0.06*	0 *0.00*	0 *0.00*	0 *0.00*	0 *0.00*	0 *0.00*	0 *0.00*	2 *0.03*	0 *0.00*
LET_REFL	13 *0.16*	2 *0.02*	3 *0.03*	4 *0.05*	5 *0.08*	4 *0.08*	3 *0.04*	4 *0.05*	1 *0.01*	1 *0.01*	2 *0.04*
LET_Clause	0 *0.00*	0 *0.00*	0 *0.00*	9 *0.12*	0 *0.00*	5 *0.09*	0 *0.00*	0 *0.00*	7 *0.09*	0 *0.00*	0 *0.00*
LET_NP	15 *0.19*	2 *0.02*	0 *0.00*	32 *0.42*	23 *0.38*	21 *0.40*	11 *0.14*	1 *0.01*	26 *0.35*	29 *0.42*	0 *0.00*
LET'S	7 *0.09*	24 *0.25*	35 *0.40*	0 *0.00*	0 *0.00*	0 *0.00*	8 *0.10*	25 *0.33*	0 *0.00*	0 *0.00*	4 *0.09*
LET_V(only)	5 *0.06*	4 *0.04*	1 *0.01*	0 *0.00*	3 *0.05*	2 *0.04*	1 *0.01*	0 *0.00*	2 *0.03*	0 *0.00*	1 *0.02*
LET_Prep_V	0 *0.00*	0 *0.00*	0 *0.00*	8 *0.10*	0 *0.00*	0 *0.00*	0 *0.00*	0 *0.00*	3 *0.04*	0 *0.00*	0 *0.00*
LET_V_(Prep)_V	0 *0.00*	11 *0.12*	0 *0.00*	0 *0.00*	0 *0.00*	0 *0.00*	0 *0.00*	0 *0.00*	0 *0.00*	0 *0.00*	0 *0.00*
LET	1 *0.01*	0 *0.00*	0 *0.00*	2 *0.03*	2 *0.03*	2 *0.04*	1 *0.01*	0 *0.00*	1 *0.01*	2 *0.03*	0 *0.00*
Other	0 *0.00*	1 *0.01*	3 *0.02*	1 *0.01*	0 *0.00*	0 *0.00*	1 *0.01*	1 *0.01*	4 *0.05*	0 *0.00*	0 *0.00*
Total	81	95	87	77	61	53	79	77	75	69	45

proportion of this function is about 10%. The Romance verbs are not used in this function.

3. The German *lassen* has the highest proportion of coreferential uses (16%). Among the Romance verbs, the leaders are *laisser* and *lasciare*, with 8% each.

4. Only the Spanish *dejar* (12%), Portuguese *deixar* (9%) and Italian *lasciare* (9%) are used with a complement clause (*LET_Clause*).

5. Only the Ibero-Romance verbs *dejar* (10%) and *deixar* (4%) can be followed by a preposition with an Infinitive. This construction designates "to cease, stop, fail to do something".

6. The Danish *lade* has a special expression with the infinitive, *lade være + med + at + V* "stop (with) doing something", as in Example (36):

(36) DK Så lad   være med at opføre   Dem sådan.
      so let.IMP.SG be.INF with PRT behaveINF REFL.2 like-that
      "So stop behaving like that."

Similar expressions without a comitative preposition exist in Norwegian, i.e. *la være + å* "to" + Vinf,[13] and Swedish, i.e. *låta bli* "let (the matter) rest" + *att* "to" + Vinf, (Viberg, 2009), but they did not occur in the data.

To visualize the relationships between the distributional features and verbs, Simple Correspondence Analysis was performed. The results are shown in Figure 9.[14] The relationships between the features and verbs should be interpreted dimensionally. Let us begin with the horizontal dimension, which seems to correspond most faithfully to the genealogical relationships between the languages. The left-hand side of the plot contains the Romance verbs and the distributional features for which the letting verb is not an auxiliary, with the exception of *LET_Refl*, which is very close to the vertical axis. The German *lassen* is also located in that sector, close to the *y*-axis. The right-hand side displays the remaining Germanic verbs, and the features associated with letting, as opposed to leaving. This supports the previously made conclusion about the LET : LEAVE ratio in Germanic and Romance languages.

In contrast, the vertical dimension seems to support an areal interpretation. The languages in the upper part of the plot, with the exception of Romanian and Swedish, are those that belong to the core of Standard Average European (van der Auwera, 1998; Haspelmath, 2001): Dutch, French, German and Italian. The letting verbs in these four languages seem to be associated with the most common features: *LET_Vintr*, *LET_Vtr*, *LET_Refl*, as well as with less frequent *LET* on

13. I thank the editors of *Languages in Contrast* for pointing this fact out.

14. Adding a third dimension leads to a total explained inertia of 81.9%. However, this dimension separates the Danish verb from all others, which is not particularly informative.
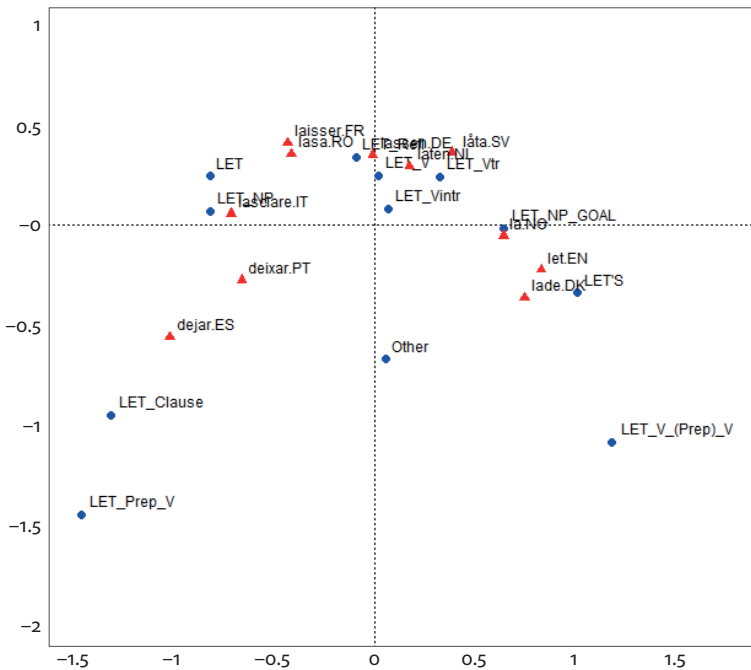
Figure 9. A Simple Correspondence Analysis map of letting verbs and their constructional features. Dimensions 1 and 2. Explained inertia: 69.4%.

the left-hand side and less frequent *LET_NP_Goal* on the right-hand side. French and German are particularly interesting, since they have the most balanced distributional vectors with regard to *LET_NP* (leaving) and auxiliary uses (letting or making). The German *lassen* has the highest proportion of *LET_NP* among all of the Germanic verbs in this study, while the French *laisser* has the highest proportion of auxiliary uses of all of the Romance verbs of letting. In other words, *laisser* is the most 'Germanic' among the Romance verbs, and *lassen* is the most 'Romance' among the Germanic verbs. In addition, *laisser* is not used with a clause complement in the data set, unlike the Italian, Spanish and Portuguese verbs. In Germanic, this use is impossible, as well. Moreover, German has the highest proportion of coreferential uses (*LET_Refl*) (16%), followed by French (8.2%) and Italian (7.5%). As for the Romanian and Swedish verbs, which are also located in the upper part of the plot, they were the most compact and central on the exemplar-based MDS map, with the least number of original extensions. This explains their position in the top area of the Correspondence Analysis map.

In the lower part of the plot, one can find the more peripheral languages. On the left are the Spanish *dejar* and Portuguese *deixar*, with their extensions *LET_Clause* and *LET_Prep_V*. On the right, one finds English, Norwegian and

Danish, with the hortative construction and the extension of the Danish *lade være + med + at + Vinf*. It is interesting that the Iberian and Scandinavian languages have a similar extension 'stop doing something' although the constructional make-up and origin is different. In the case of the Iberian languages, the action that is stopped is construed as an object abandoned by the Agent. In the Scandinavian languages, as was mentioned above, this expression is made from another frequent construction, which means 'let (something or someone) be', followed by an infinitival complement. This difference reflects the preferences for the lexical or auxiliary use of verbs of letting in the Romance and Germanic languages.

Finally, if one combines both dimensions and reads the plot from the bottom right to the top and then to the bottom left, one finds a pattern that roughly corresponds to the geographical continuum from North (Scandinavia and Great Britain) to South (Iberian Peninsula). Such a horseshoe pattern is quite common in multivariate analyses.

5.   Discussion

The aim of this study was to find similarities and differences in the semasiological variation of eleven verbs of letting in Germanic and Romance languages. Using a parallel corpus of film subtitles, I employed exploratory statistical techniques to visualize the common semantic space of letting and to establish the distinctive and overlapping functions of these verbs with regard to the most important semantic dimensions. The most important dimensions of cross-linguistic semantic variation for these verbs were letting vs. leaving, factitive vs. permissive causation and modality/discourse function.

We also observed some general genealogical patterns that relate to these dimensions of variation. Most importantly, the Germanic and Romance verbs differ with regard to the distinction between letting and leaving; more specifically, the Germanic verbs more often serve as auxiliaries with the Agonist as a secondary Trajector, while the Romance verbs more frequently express the sense of leaving, with the second participant as a Landmark. In addition, only the Germanic verbs can express factitive causation and perform different modal functions, in particular, the hortative one.

However, this difference between Germanic and Romance verbs is not clear-cut. With regard to this distinction, the German *lassen*, followed by the Dutch *laten*, seem to be more similar to the Romance verbs than to the other Germanic verbs; likewise, the French *laisser* is the most similar to the Germanic verbs among the Romance verbs. Regarding modal functions, the English *let* has the highest relative frequency of hortative uses, followed by the Danish and Norwegian verbs,

while the German, Dutch and Swedish cognates have only moderate frequencies of the hortative *let*. These results support the view of Europe as a linguistic area with a centre and a periphery (van der Auwera, 1998; Haspelmath, 2001).

The observed differences can be interpreted from a historical perspective. According to Soares da Silva (2007), the Germanic cognates *let*, *laten*, *lassen*, etc., as well as the Gothic *lētan* illustrate the evolutionary course from the sense of LEAVE to the semantics of LET/PERMIT.[15] The results of the quantitative study suggest that, in the contemporary Germanic languages, the semantics of leaving is marginal. Instead, prefixal derivatives are used as verbs of leaving, such as the German *verlassen* "to leave, abandon" or the Dutch *nalaten* "to leave, bequeath", as well as other verbs that may be etymologically related to the verbs of letting, such as the English *leave* or Swedish *lämna* "to leave".

The development of the Romance letting verbs can be traced in greater detail than that of the Germanic ones. The meaning 'to let go, release' of the Latin etymon *laxare* became prototypical around the second century CE. Later, the meaning 'to go away, abandon' developed (replacing the verb *relinquere*, which previously expressed this meaning), as well as the meaning 'to let, allow', a development that made *laxare* a semantic successor of *sinere* (Soares da Silva, 2007: 185). With time, the historical prototype of active letting with previous impingement 'to release, let go' has become less prominent in the Romance languages. As noted by Soares da Silva (2003), the French *laisser* and Romanian *a lăsa* have more salient verbs for expressing the sense of 'let go' (see Section 2). This sense implies an initial physical contact and a spatiotemporal contingency of the letting and permitted events, giving the latter a high degree of semantic integration. At the same time, the meaning 'let', as opposed to 'leave', has become overall less prominent in the Romance verbs under consideration.

Considering this historical information, one can formulate a hypothesis for diachronic research. It seems that the Germanic and Romance verbs of letting have been developing in opposite directions. The Germanic verbs previously denoted both 'let' and 'leave', but now specialize mostly in letting. They have also acquired a higher degree of semantic and syntactic integration with the caused event than have the Romance verbs, and overall seem to stand further on the grammaticalization cline. In contrast, although the historical prototype of the Romance verbs had a prominent component related to 'let', these verbs are now frequently used as lexical verbs of leaving. At the same time, when expressing letting, some Romance verbs have a relatively high proportion of finite clausal complements,

---

15. This assertion is not a unanimous one, however. For example, The Etymological Dictionary of Proto-Germanic (Kroonen, 2013: 332) says that the meaning of the proto-Germanic **lētan* was 'let, allow'.

which suggests a lower degree of syntactic and semantic integration of the letting and permitted events in comparison with the infinitival construction. Thus, it seems that the position of the Romance verbs (or at least, that of their particular meanings) has become lower on the semantic binding hierarchy. Importantly, the languages in the middle of the geographical continuum (especially German and French) tend to combine both tendencies, while the languages on the periphery tend to more strongly exemplify one of the tendencies. Of course, this hypothesis requires a thorough diachronic investigation based on solid corpus evidence.

Abbreviations

DE	German	NL	Dutch
DK	Danish	NO	Norwegian
EN	English	PT	Portuguese
ES	Spanish	RO	Romanian
FR	French	SV	Swedish
IT	Italian		

Acknowledgements

References

AHD: The American Heritage Dictionary of the English Language. (1970). New York: American Heritage.

Achard, M. 1998. Two Causation/Perception Constructions in French. *Cognitive Linguistics* 7: 315–357. doi:10.1515/cogl.1996.7.4.315

Apresjan, J. 1966. Analyse distributionelle des significations et champs sémantiques structurés. *Langages* 1: 44–74. doi:10.3406/lgge.1966.2865

Levshina, N. 2015. How Europeans GIVE: A two-layered semantic typology based on parallel corpora. In *Causation, Permission, and Transfer: Argument realization in GET, TAKE, PUT, GIVE and LET verbs*, B. Nolan, G. Rawoens and E. Diedrichsen (eds), 147–175. Amsterdam: John Benjamins. doi:10.1075/slcs.167.05lev

Levshina, N. and Heylen, K. 2014. A Radically Data-driven Construction Grammar: Experiments with Dutch causative constructions. In *Extending the Scope of Construction Grammar*, R. Boogaart, T. Colleman, and G. Rutten (eds), 17–46. Berlin: Mouton de Gruyter.

Givón, T. 1980. The Binding Hierarchy and the Typology of Complements. *Studies in Language* 4(3): 333–377. doi:10.1075/sl.4.3.03giv

Davidse, K. and Heyvaert, L. 2003. On the so-called 'middle' construction in English and Dutch. In *Corpus-based Approaches to Contrastive Linguistics and Translation Studies*, S. Granger, J. Lerot and S. Petch-Tyson (eds), 57–73. Amsterdam: Rodopi.

Davies, M. 2004–. BYU-BNC (Based on the British National Corpus from Oxford University Press). Retrieved from http://corpus.byu.edu/bnc/ [last accessed 15 December 2014].

Davies, M. 2008–. The Corpus of Contemporary American English: 450 million words, 1990-present. Retrieved from http://corpus.byu.edu/coca/ [last accessed 15 December 2014].

Davies, M. 2012–. The Corpus of American Soap Operas. Retrieved from http://corpus2.byu.edu/soap/ [last accessed 15 December 2014].

Deckert, M. 2013. *Meaning in Subtitling: Toward a Contrastive Cognitive Semantic Model*. Frankfurt am Main: Peter Lang. doi:10.3726/978-3-653-02868-3

Diez, F. Ch. and Donkin, T.C. 1864. *An Etymological Dictionary of the Romance Languages; Chiefly from the German of Friedrich Diez*. Edinburgh: Williams and Norgate. Retrieved from https://archive.org/details/anetymologicald00diezgoog [last accessed 15 December 2014].

Egan, T. 2008. *Non-finite Complementation: A Usage-Based Study of Infinitive and -ing Clauses in English*. Amsterdam: Rodopi.

Evans, N. (2010). Semantic typology. In *The Oxford Handbook of Linguistic Typology*, J. J. Song (ed.), 504–533. Oxford: Oxford University Press.

Friedman, V.A. 2006. Balkans as a linguistic area. In *Encyclopedia of Language and Linguistics*, Keith Brown (ed.), 2nd ed., 657–672. Oxford: Elsevier. doi:10.1016/B0-08-044854-2/00178-4

Geeraerts, D. 1998. The semantic structure of the indirect object in Dutch. In *The Dative II. Theoretical Approaches*, W. Van Langendonck and W. Van Belle (eds), 185–210. Amsterdam: John Benjamins. doi:10.1075/cagral.3.08gee

Geeraerts, D. 2010. *Theories of Lexical Semantics*. Oxford: Oxford University Press.

Johansson, S. 2007. *Seeing through Multilingual Corpora. On the Use of Corpora in Contrastive Studies*. Amsterdam: John Benjamins. doi:10.1075/scl.26

Haspelmath, M. 2001. The European linguistic area: Standard Average European. In *Language Typology and Language Universals*, M. Haspelmath, E. König, W. Oesterreicher and W. Raible (eds), 1492–1510. Berlin: de Gruyter. doi:10.1515/9783110171549.2

Hüning, M., Vogl, U., van der Wouden, T. and Verhagen, A. (eds). 2006. *Nederlands tussen Duits en Engels* [Dutch between German and English]. Leiden: Stichting Neerlandistiek Leiden.

Keuleers, E., Brysbaert, M. and New, B. 2010. SUBTLEX-NL: A New Frequency Measure for Dutch Words Based on Film Subtitles. *Behavior Research Methods* 42: 643–650. doi:10.3758/BRM.42.3.643

Kroonen, G. 2013. *Etymological Dictionary of Proto-Germanic*. Leiden: Brill.

Langacker, R.W. 1991. *Foundations of Cognitive Grammar, Volume II, Descriptive Application*. Stanford, CA: Stanford University Press.

Lefer, M-A. and Cartoni, B. 2013. Word-formation in original and translated English: Source language influence on the use of un- and -less. In *Theoretical and Empirical Advances in Word Formation*, J. Fernandez (ed.), 49–59. Special issue of the Quaderns de Filologia de la Universidad de Valencia 18.

Levin, B. 1993. *English Verb Classes and Alternations: A Preliminary Investigation*. Chicago: University of Chicago Press.

Loewenthal, J. 2003. Meaning and use of causeeless causative constructions with laten in Dutch. In *Usage-based Approaches to Dutch. Lexicon*, Grammar, Discourse, A. Verhagen and J. van de Wejer (eds), 97–129. Utrecht: LOT.

Lowrey, B. 2013. The Old English Causative hatan and Its Demise. *Token* 2:23–43. Retrieved from http://www.ujk.edu.pl/token/issues/volume-2/ [last accessed 17.12.2014].

Malkiel, Y. 1968. *Essays on Linguistic Themes*. Oxford: Blackwell.

Maldonado, R. 2007. Soft causatives in Spanish. In *On Interpreting Construction Schemas. From Action and Motion to Transitivity and Causality*, N. Delbeque and B. Cornillie (eds), 229–260. Berlin: Mouton de Gruyter.

Nedjalkov, V. 1976. *Kausativkonstruktionen*. Tübingen: TBL.

Newman, J. 1996. *Give: A Cognitive Linguistic Study*. Berlin: Mouton de Gruyter.

OED: *The Oxford English Dictionary*. 1970. Vol. IV (L–M). Oxford: Clarendon Press.

R Core Team. 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Vienna. Retrieved from http://www.R-project.org/ [last accessed 15 December 2014].

Rawoens, G. 2013. The Swedish Verb *låta* 'let' from a Synchronic and Diachronic Perspective. *Language Sciences* 36: 113–123. doi:10.1016/j.langsci.2012.04.002

Rawoens, G. and Egan, T. 2014. Distinguishing Causative and Permissive Readings of the Swedish Verb *låta*. *Functions of Language* 20(1): 64–89. doi:10.1075/fol.20.1.03raw

Soares da Silva, A. 2003. La structure sémantique de 'laisser' dans les langues romanes. In *Actas del XXIII Congreso Internacional de Lingüística y Filología Románica*. Vol III, F. Sánchez Miret (ed.), 441–456. Tübingen: Niemeyer.

Soares da Silva, A. 2004. Imagery in Portuguese causation/perception constructions. In *Imagery in Language. Festschrift in Honour of Professor Ronald W. Langacker*, B. Lewandowska-Tomaszczyk and A. Kwiatkowska (eds), 297–319. Frankfurt am Main: Peter Lang.

Soares da Silva, A. 2007. Verbs of letting: Some cognitive and historical aspects. In *On Interpreting Construction Schemas. From Action and Motion to Transitivity and Causality*, N. Delbecque and B. Cornillie (eds), 171–200. Berlin: Mouton de Gruyter.

Talmy, L. 2000. *Toward a Cognitive Semantics*. Cambridge, MA: MIT Press.

Tiedemann, J. 2009. News from OPUS – A Collection of Multilingual Parallel Corpora with Tools and Interfaces. In *Recent Advances in Natural Language Processing*, Vol. V, N. Nicolov, K. Bontcheva, G. Angelova and R. Mitkov (eds), 237–248. Amsterdam: John Benjamins. doi:10.1075/cilt.309.19tie

van der Auwera, J. 1998. Introduction. In *Adverbial Constructions in the Languages of Europe*, J. van der Auwera (ed), 1–23. Berlin: Mouton de Gruyter. doi:10.1515/9783110802610.1

Van Olmen, D. 2011. *The Imperative in English and Dutch: A Functional Analysis Based in Comparable and Parallel Corpora*. PhD diss., Universiteit Antwerpen.

Viberg, Å. 2009. The meaning patterns of Swedish *låta* 'sound'; 'let'. A corpus-based contrastive study. In *Mélanges plurilingues offerts à Suzanne Schlyter à l'occasion de son 65ème anniversaire* [Études romanes de Lund 85], P. Bernardini, V. Egerland and J. Granfeldt (eds.), 455–480. Lund: Romanska institutionen, Lunds universitet.

von Waldenfels, R. 2012. *The Grammaticalization of 'give'+Infinitive: A Comparative Study of Russian, Polish and Czech*. Berlin: Mouton de Gruyter. doi:10.1515/9783110293777

Wälchli, B. and Cysouw, M. 2012. Lexical Typology through Similarity Semantics: Toward a Semantic Map of Motion Verbs. *Linguistics* 50(3): 671–710. doi:10.1515/ling-2012-0021

Wierzbicka, A. 2006. *English: Meaning and Culture*. Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780195174748.001.0001

Author's address

Natalia Levshina
Institut Langage et Communication
Université catholique de Louvain
Pl. Blaise Pascal 1
bte L3.03.33
Louvain-la-Neuve 1348
Belgium

natalia.levshina@uclouvain.be