Representing the Lexicon: Identifying Meaning in Use via Underspecification*

Henk Zeevat
Institute for Logic, Language and Computation
University of Amsterdam
henk.zeevat@uva.nl

Scott Grimm
University of Rochester
scott.grimm@rochester.edu

Lotte Hogeweg
Radboud University Nijmegen
lottehogeweg.ruhosting.nl

Sander Lestrade
Radboud Unoversity Nijmegen
s.lestrade@let.ru.nl

Elizabeth Allyn Smith
Université du Québec à Montréal
smith.eallyn@uqam.ca

1 Lexical Specification and Meaning in Use: Traditional Views on the Lexicon

Traditionally, the lexicon is the list of words of a language or communication system. From a biological perspective, the lexica of human languages are special with respect to animal communication systems because they can be freely extended, both by the creation of new (lexical) forms and by the extension of the meaning of these words. Notwithstanding this

*We wish to express our thanks to audiences at Szklarska Poreba, the Workshop Bridging Formal and Conceptual Semantics, and Formal Semantics Meets Cognitive Semantics as well as to Katrin Erk, to Louise McNally and, especially, to an anonymous reviewer of these proceedings for many useful comments, more, in fact, than we could integrate into this preliminary report. In addition, we would like to thank the the Netherlands Organization for Scientific Research (NWO) for their financial support to Lotte Hogeweg. All errors are our own. Authors after the first author are listed in alphabetical order.

extensibility, the construction of dictionaries as fixed lists of words with their meanings pays off for the purposes of language education, reliable communication and translation.

Computational models of language largely share this view, as it seems initially reasonable to think that such a lexicon could be part of what characterizes the human ability to code thoughts into linguistic expressions and to recover thoughts from such expressions. What is needed for conventional dictionaries—in the tradition started by such eminent scholars as Samuel Johnson and Jacob Grimm—is precision with respect to the characterization of the word senses described. Such a task would be difficult with an entailment-based semantics only, since many word senses include connotations that are usually but not always associated with an utterance of that word. However, abandoning formal semantic approaches ignores the fact that the senses distinguished by lexicographers do often reflect truth-conditional differences, e.g. *Anna went to the bank* would entail 'Anna was very near a flowing body of water' in one of the word *bank*'s senses but not another. We contend that a formal account would need to distinguish as many senses as traditional dictionaries do, and one goal of the present paper is to suggest some steps toward such an account.

The main challenge faced by lexicographers and semanticists alike with respect to characterizing word senses precisely is lexical disambiguation. Given the large number of senses that dictionaries distinguish for the same word, one needs a cognitively plausible account of the fact that listeners only rarely interpret the sense of a word in a way unintended by the speaker. Suppose that the list of word senses were simply a list and that lexicalisation meant just choosing a word for a concept. Now suppose further that all concepts linked to a word after lexicalisation were listed with it and that disambiguation occurred by making a random choice from the list of senses. In this case, the chance of speaker-hearer coordination would be very small, namely 1 divided by the number of word senses of the item in question.

Given that this is not what we find, a natural extension to such a simplistic model would be to include stochastic data with the senses. The first type of stochastic data to consider would be the relative frequency of the senses (such as the fact that 'financial institution' is perhaps more frequent than 'river bank' as the intended meaning of *bank* in industrialized societies), but this by itself would predict that the best choice would always be to select the most frequent sense, so the very existence of the other senses makes this option implausible. A second stochastic approach in line with recent proposals (Baroni et al., 2014; Erk, 2014) also includes frequency data about the words that a given sense is likely to combine with (the "distributional semantics" of the sense). For example, the presence of words like *deposit* and *savings* around the word *bank* will bias the financial institution interpretation. The option of adding distributional semantics to senses needs further study. It is, however, our suspicion that this will only work properly in combination with a decompositional approach to senses because decomposition would be needed for estimates in the presence of data scarceness. These distributional approaches nevertheless aim to solve the same problem that we address in this paper, namely that of predicting the meaning of lexical items in use (i.e. in the speech contexts in which they appear), but they use a different representation of lexical knowledge than the one we advocate here.

The approach pursued in this paper tries to exploit semantic decomposition of word senses to arrive at a cognitive representation that is effective in selecting the meaning in use: one of a word's potentially many senses that should be listed in a full traditional dictionary. In some respects our approach is in line with other more linguistically-oriented approaches to lexical representations. For example, it shares the assumption of decomposability of word senses with cognitive approaches such as Jackendoff's Conceptual Semantics (e.g. Jackendoff 1996). It is similar to more formal theories like Pustejovsky's (1995) Generative Lexicon and Blutner's (2004) Lexical Pragmatics in its aim to systematically account for meaning alternations. Our approach differs from these approaches in that, in our proposal, words typically overspecify their meanings, and it is the combination with context that trims the overspecified meaning down to the meaning in use, i.e. the word sense that applies. The meaning in use may be determined by earlier language use, but it can also be computed for the first time.

After a discussion of semantic features in the following section, we examine previous accounts of overspecification of meaning upon which we build. We provide an analysis of lexical items in terms of stochastic sets of features in section 3, with an extended treatment of the verb *fall* in English, Dutch and German, which demonstrates how one overspecified representation can apply in different contexts resulting in different word senses. In section 5, we then discuss the issue of overgeneration and the subsequent need to put constraints on the production of words, in turn defining the lexicalization process. We conclude in section 6.

2 Semantic Features and Moderate Universalism

In this section, we want to defend the view that word senses are composed of a set of (moderately) universal semantic features combined with natural classifications of experience. By 'moderately universal,' which we return to below, we mean that any two languages will have a significant overlap in the features they use to construct word senses, but there may be unique features as well. To demonstrate what we mean by "a natural classification of experience", take the verb *walk*. The particular kind of locomotion that we achieve in walking is difficult to analyse further by people who can walk, and AI attempts at modelling it have revealed how little humans understand about how it functions.[1] But it is part of the human repertoire of activities and, as such, humans use it in planning their behaviour and in recognizing it in the behaviour of other humans and animals. This particular kind of locomotion is thus a "natural classification of experience".

The decomposability of word senses has been a controversial issue in the literature on the lexicon. On the one hand, there are proponents of a atomistic view of the lexicon such as Fodor (e.g. Fodor and Pylyshyn 1988, Fodor and McLaughlin 1991 and Fodor and Lepore 2002) and more recently Relevance Theorists, e.g. Carston (2010). On the other hand there are advocates of more complex lexical representations such as Pustejovsky (1995). A common argument against a decompositional view is the lack of necessary and sufficient features in defining word senses (e.g. If a tiger is defined as having four legs, does that make a three-legged tiger not a tiger?) Furthermore, one could argue that semantic features are cognitively meaningless if features can freely be invented and added to a representation. We argue that semantic features are needed for overspecification and for formalizing the selection of meaning in use. We contend that these features need to be moderately universal in order to account for the fact that knowing one language's lexicon helps in learning another language's lexicon and also to account for typological generalisations about the lexicon and morphology. But most importantly for our purposes, universal features are needed for methodological reasons to ensure that the decompositions of verbal meaning are cognitively meaningful. They should not be freely inventable (since there are things that are difficult or impossible for us to conceptualize) and, like optimality-theoretic constraints, they should preferably come with a demonstration that they are typologically valid.

For an example of typologically-valid semantic features, one can turn to agency, and more specifically to the proto-agent and proto-patient features set out by Dowty (1991), such as

[1]Introductory AI texts take as given that locomotion, grasping something with your hands, and seeing objects belong to the everyday behavioural repertoire of humans and that they turn out be much more difficult than the naive views of these activities would suggest. The use of natural language falls in the same category.

sentience, *volition*, *control* and *cause*. The typical agent of *walking toward the sunset* forms the intention of doing so within the situation she finds herself in, causes her movement to start, controls and monitors her progress and has a criterion for when it is finished. The notions of *agent*, *intending*, *controlling*, and *end of action* all belong to the realm of typologically-valid features.

Grimm (2011) shows that these features are central to the typology of case systems, in the sense that they play a role in accounting for the variation between languages that one finds in the realisation of case systems. Typological research and monolingual investigation of central semantico-pragmatic themes such as case, tense, aspect, modality and definiteness provide a large number of ostensibly-valid semantic features.

It is further possible to give a foundation for semantic features using the semantic map method in typology (de Schepper and Zwarts 2010). In the semantic map approach, one studies the meaning of a word or a group of related words using comparison with other languages, by systematically looking at translatability. This provides a natural way of dividing words into their uses without appealing to semantic intuitions. If one uses translations into sufficiently many languages, one can map them onto a two-dimensional graph where the points are sets of translation equivalents. If two languages do not make a meaning distinction that is made in a third language, the first two are connected in the graph. For example, neither English nor French distinguishes direction and recipient in their prepositions, but German does, which would mean that English and French are connected in the graph for this semantic concept (Haspelmath 2003).

de Schepper and Zwarts (2010) show that such maps can be systematically represented by feature clusters, with each of two minimally distant points differing in precisely one feature. Ideally, the sets of features representing a map can be analyzed in terms of typologically well-studied semantic features. This analysis can also be taken as underpinning the view that meanings in use can be seen as sets of semantic features.

Our approach relates to the semantic map approach and other cognitively-oriented approaches such as Conceptual Semantics (e.g. Jackendoff 1996) in the assumption of cognitively realistic semantic features. It is still a task, however, to define how they combine into a logical expression that characterizes the truth-conditional contribution of the combined features in terms of the truth-conditional contribution of the individual features. For this purpose, it is useful to adopt the view of Barsalou (1992) that meanings in use should be characterized as frames. Features typically set the value of attributes in a frame, unify attribute values, or indicate that certain attributes have a value and that composition can be modelled by unification. The truth-conditional interpretation of a set of attribute value structures is the claim that the class of complex entities that meet all the constraints is non-empty. In contradistinction to the semantics provided for feature structures modelling linguistic objects by, e.g., Johnson (1988), attributes for these semantic uses must be understood as operations in the external world.[2] The object (e.g. an event of falling) would be related to whatever falls by the attribute *theme*, interpreted as the operation that maps events to their themes. While in this particular case it is not unreasonable to think that the theme in some sense constitutes the event, the use of frames by itself does not commit one to this view.

In this respect, we see no distance between the current proposal and formal semantics. There would be one if formal semantics were interpreted as committed to the view that nothing could

[2]We here part company with those who like to maintain that semantics is merely in the brain. We take the line that a proper semantics should also explain logical inference relations between natural language utterances and thereby should have a model theory. This is the line also taken by Kamp in DRT (Kamp and Reyle 1993). We also feel that many aspects of Conceptual Semantics allow a model-theoretic treatment and that its proponents undersell their theory in this respect.

be said about the structure of meanings beyond their contribution to truth-conditions, which would make it irrelevant for language learning and cognition. The view defended here demands that basic features make sense from the perspective of classical truth-conditional semantics. We forego a discussion of the typical problems for a view of this kind: vagueness, taste predicates, information structural features, and emotional expression.

As mentioned above, the proposal introduced here adopts a moderate form of universalism towards linguistic lexical meanings, but it is the building blocks of lexical meanings and not the lexical meanings themselves that are universal. There is no assumption that any feature will play a role in all languages, nor that any language uses all features. New semantic features may be introduced for the description of a new lexicon if they correspond with a learnable classification. We opt for this moderation rather than for absolute universalism on the basis of evidence showing that speakers of one language are sensitive to semantic differences that speakers of other languages are not sensitive to. For example, Korean has two variants of the English preposition *in* depending upon whether the object is in a close-fitting container (like a SIM card in a cell phone) or in a non-close-fitting container (like a pear in a bowl). McDonough et al. (2003) use preferential looking tasks and a "which of these things is not like the others?" task to show that while Korean and English infants are both sensitive to differences in closeness of fit, as adults, only Korean speakers and not English speakers are sensitive to this difference. Thus, the distinction in close versus loose fit as articulated in the Korean choice between prepositions is not present in the adult English lexicon or in adult English cognition and should not be posited as an absolute universal. In fact, it would be a Korean-particular extra feature in our formalisation that has arisen through a grammaticalisation process under the influence of a forced choice between lexemes. Our moderate universalism leads to a decomposition in terms of universal features (where typology and cognition supply the foundation for the universal character) with a minimum number of additional idiosyncratic features. We now turn to discussing previous accounts of word meanings based on overspecification of meaning.

3 Overspecification of Meaning: The Hogeweg-Smolensky Account

In his response to the criticism of connectionism in Fodor and Pylyshyn (1988), Smolensky (1991) offers an analysis in which the distributed representation of *coffee* can be derived by subtracting the representation of *cup with coffee* from the representation of *cup without coffee*. In Smolensky's analysis, the representation of a *cup with coffee* consists of a set of micro-features like 'upright container', 'hot liquid', 'porcelain curved surface', 'burnt odor', 'brown liquid contacting porcelain', 'finger sized handle' and 'brown liquid with curved sides at the bottom'. The representation of *cup without coffee* consists of the features 'upright container', 'porcelain curved surface' and 'finger-sized handle'. If the representation of *cup without coffee* is subtracted from the representation of *cup with coffee*, this yields a representation of *coffee*, consisting of the features 'hot liquid', 'burnt odor', and 'brown liquid contacting porcelain'. Crucially, however, this is a representation of *coffee* in a particular context. In another context, other features of coffee (like *shrub*, *red fruit*, *brown bean*) would be activated. The features of coffee that are activated in a particular context are therefore a subset of the much larger set of features potentially projected by *coffee*.

Following Zwarts's (2004) analysis of the preposition *(a)round*, Hogeweg (2009) turns this approach into an account of computing the right set of semantic features in a context from a lexical (over-)specification using an OT grammar FIT > STRENGTH. FIT demands that the output is consistent with the context, and STRENGTH demands that the output set for the specification is maximal by ensuring that any set larger than the output does not meet FIT.

To illustrate the working of the two constraints, let us look briefly at Hogeweg's analysis of the interpretation of the Dutch discourse particle *wel*. Like most discourse particles, *wel* is highly polysemous. Hogeweg analyzes the different senses of *wel* as ranging in strength depending upon how much information a use presupposes. In Tableau 1, the possible interpretations are ordered according to their strength. The strongest meaning is illustrated by the following small conversation:

(1) a. Speaker A: *Amsterdam is niet de hoofdstad van Nederland* (Amsterdam is not the capital of the Netherlands).
b. Speaker B: *Amsterdam is wel de hoofdstad van Nederland* (Amsterdam is the capital of the Netherlands)

This discourse can be described as an instance of Speaker B's correcting Speaker A. An utterance containing *wel* expressing the proposition p requires that a statement expressing the proposition $\neg p$ was uttered. Implicit contrast, for example, is weaker since it does not require that the proposition $\neg p$ is expressed but just that it is inferrable from the context. For example, in a context where a husband is putting on his coat, his wife could utter: *Je moet wel de afwas nog doen*, 'You have WEL to do the dishes'. What is important here, however, is the interaction between the two constraints STRENGTH and FIT. STRENGTH requires that all meaning aspects are activated so that the word is interpreted with the strongest meaning, in this case, a correction. A candidate violates this constraint as many times as there are stronger interpretations available. FIT requires that the output is consistent with the context. If *wel* p is uttered in a context where a statement expressing $\neg p$ is not part of the common ground, interpreting *wel* as a correction violates FIT. If there is information in the common ground from which $\neg p$ could be inferred, implicit contrast does not violate FIT. Note that the requirements put on the context entail one another. (For example, if $\neg p$ is uttered, it can also be inferred.) That is why if an interpretation does not violate FIT, all the weaker interpretations also meet the requirements set by this constraint.

(2)

Context: Husband is putting on coat. *Je moet* **wel** *de afwas nog doen.*	FIT	STRENGTH
Correction	*	
Contrast	*	*
⇒ Implicit contrast		**
Surprise		***
Modifier		****

The account is quite successful in the application that Hogeweg provides. It has also been successfully applied to other types of function words such as prepositions Zwarts (2004, 2006, 2008). Nor is it difficult to come up with further applications. Another advantage of the approach is that, while it was developed in OT, it does not require OT-specific mechanisms that would limit its generalizability.

For example, one can interpret FIT as the maximization of prior probability and STRENGTH as the maximization of the likelihood of the signal given the input. The more features associated with the word that show up in the input, the more likely the use of the word becomes, such that adding more features projected to an interpretation hypothesis increases the likelihood of the signal. The most probable interpretation is thus a set of features that is as large as possible and yet still consistent with the context. The OT system thus reduces to a decomposition into priors and likelihoods for finding the most probable interpretation.

Perhaps this is all that one needs to model functional lexemes.[3] For lexical words, however, it runs into problems. This stems from the fact that all features are treated as equal, whereas certain kinds of phenomena bode against such equality. These include absolute features, dependencies among features and forced choices between features, exemplified below. In the next section, we illustrate these properties of feature sets by an analysis of the verb *fall*. The verb *fall* in English, Dutch, French[4], Russian and modern German is non-volitional. This property survives in all of the derived meanings, which is what makes it an instance of an absolute feature. That there are dependencies between features can be seen for example in that a spatial source for a use of *fall* such as *fall to the ground* forces a spatial goal and a spatial dimension which are interconnected. *Fall* also forces a choice between dimensions, including the aforementioned spatial dimension as well as moral (*He resisted the temptation for a long time, but then he fell.*), fortune (*fall on hard times*), and grace (*fall from grace*), among others.

4 Lexical Entries as Stochastic Sets of Features

In the previous section we argued for a feature-based analysis of words senses. In contrast with the previous approaches, we argue that a word is not related to a set of features but to a stochastic set: a distribution over sets of features. Such distributions can be learnt from experience by counting how often the various feature combinations are expressed by a word. However, the distribution itself cannot be used for explaining the intersubjective status of these stochastic feature sets. The experience of the individual users will be different and therefore the distribution that they learn. While learning a distribution is what constitutes a speaker's competence with respect to the semantics of the word, it cannot be what the language or the language community associates with the word.

Intersubjective convergence can be modeled by considering equivalence classes of such distributions: the distributions that agree on 0, 1 and $<$. Two competent speakers will almost certainly have different frequencies and probabilities for the same feature bundle b in an interpretation of a word w. By using equivalence classes for $\leq$, 0 and 1, we define competence for w as the speakers always agreeing that bundle b is less probable than bundle c in an interpretation of w, or that bundle b is always or never occurs in such interpretations.

Let F be the set of all features. Since $p(\emptyset) = 1$ and $p(F) = 0$, p and q will give 0 and 1 to the same elements if they preserve $\leq$. So (3) is sufficient.

(3) $p \sim q$ iff $\forall b, c\ (p(b) \leq p(c)) \leftrightarrow (q(b) \leq q(c))$

While different language users build up different distributions, a small amount of data suffices to guarantee that language users have distributions in the same equivalence class. For specifying such equivalence classes, the following operations can be defined in (4). Speakers know the equivalence class by having learnt or converging to a distribution that belongs to it by being exposed to utterances in language use. The equivalence class (and a particular distribution in it, almost certainly different from the distribution of any user) can be attributed to language use or to the language community producing language use.

(4) absolute features: $p(F) = 1$

[3]The use of *already* observed by Fong (2001) as an expression of the perfect in *You eat already?* in Singapore English (British English informants also report this use in informal standard English) can be used as an argument against this view. Arguably, *already* expresses both surprise (at the early start of a state) and the fact that the state started. Surprise is removed when *already* expresses the perfect, but there are no uses where the perfect is removed. This makes the perfect an absolute feature and surprise, a default feature.

[4]French is not yet integrated into the formal representation of *fall* discussed below.

excluded features: $p(b) = 0$ (not normally considered)
conditionally absolute features: $p(b|c) = 1$
conditionally excluded features: $p(b|c) = 0$
forced choice: $p(c_1 \vee \ldots \vee c_n) = 1$ and $p(c_i \wedge c_j) = 0$ for $i \neq j \leq n$
default in a forced choice: $c_1 \ldots c_n$: a feature c_i such that $p(c_i) > p(c_j)$ for $j \neq i$ and $1 \leq j \leq n$

In the following, we will illustrate how such operations enable us to give a representation of the verb *fall*. We chose to exemplify this approach to the lexicon with *fall* not only because the central concept expressed facilitates cross-linguistic comparison, but also because *fall* and its cognates typically lend themselves to extended uses, i.e. a large number of word senses, as mentioned above. Apart from its most straightforward interpretation as a motion verb, *fall* is used with various other interpretations, including non-spatial ones such as a fallen soldier or a fallen woman. As such, it provides a sufficiently difficult modeling task to develop a representation from which all the different uses could be specified in a context.

In the project database based on data extraction efforts from dictionaries and the internet, we currently have 78 uses of "fall". This reduces to a smaller, but not much smaller number of uses for a particular language. Example (5) lists 18 of the 35 uses that seem acceptable in English and is meant to illustrate the variation.

(5) John fell out of the tree.
The glass fell on the floor.
John fell (down).
The house fell.
The corporal fell
The rain fell.
The evening falls.
Christmas day falls on a Sunday this year.
He fell asleep.
His eyes fell on the gem.
The cabinet falls.
The thaw fell over the fields.
The water fell.
Dark curls fell around her white neck.
It fell into oblivion.
The goblet fell to the bottom of the river.
The waves fell on the beach.
The curtains fall.
The path falls. (goes down)
Grief fell from our hearts.

Many of these examples can be seen as metaphorical extensions from a basic use. In canonical views on metaphors, space and in particular the up-down opposition is an important source for metaphorical extension. As Lakoff and Johnson (1980) argue, most of our fundamental concepts are organized in one or more spatialization metaphors. They provide many examples in which a more abstract concept is expressed in terms of the opposition between up and down, among which health and life (health and life are up, sickness and death are down), morality (virtue is up, depravity is down) and quantity (more is up, less is down), many of which are also applicable to *fall*. However, the aim in this paper is not to capture metaphor but to find meanings in use. Dead metaphors are dead and the language user is stuck with them even if the metaphorical extension could not have happened anymore. Our strategy stands in competition with an approach which would want to predict the metaphorical use from a more basic use. Maybe that is possible and would lead to similar predictions. A reason to be skeptical about that, however, is that what works in one language does not seem possible in another in many cases and that general accounts of metaphor, even if tied to notions like natural metaphor will fail to predict correctly when a particular metaphor is possible and when it is not. In our case, the starting point is what has happened already and is recorded in the lexicon. Interestingly, we then predict more possible interpretations than were found in the lexicon and standardly correct ones. This can be interpreted as metaphor formation, but so far nothing very much can be claimed for this method of finding new meanings in use. It certainly does not seem to end up (yet?) as a serious general theory of metaphor. The aim of our account should not be confused with the legitimate enterprise of explaining why certain metaphors are more natural and acceptable than others. The project may supply interesting input to such an enterprise since our data are suggestive of what is natural and not, but the enterprise itself will not make any contribution to this essentially psychological question.

The problem for specification is the problem of dealing with all 78 uses from one single representation. It follows from our moderate universalism that this should be possible and—surprisingly, since this is a strong claim—it seems that this is the case. As it turns out, again surprisingly, the language specific representations are not really simpler than the cross-linguistic one.

We now outline the different components of the representation. We use a frame formalism because it comes with a natural decomposition (unlike first order logic), has a properly defined semantics (Johnson 1988; Ait-Kaci and Podelski 1993) and has been claimed by many to be a natural format for the description of concepts. Barsalou (1992), Loebner (2014) Petersen and Werning (2007), and Sag et al. (2003) all give substantial empirical evidence for a frame-like structure of mental representations. For us, having a natural decomposition is the most important advantage and while we regard the current development as promising, we are no way committed to sticking to this particular formalism in future developments of this material.

Though it is too early for a detailed formal proposal, semantic features are interpreted as constraints on relationally restricted frame structures, giving both the structure of the events or states denoted and the concept of these events and states. Equivalence classes over distributions over these features are the lexical specifications.

The lexical specifications have maximal consistent sets of constraints over features allowed by the distribution. These determine classes of frames which in turn determine what kind of objects they can denote.

The following are examples of the different frame constraints on *fall*. Relations and sorts are written with lower case heads, variables appear in upper case.

(6) THEME := the frame has an attribute 'theme' with the variable THEME at its end

at(THEME,LOCATION) := THEME is at LOCATION

SOURCE:location(DIMENSION) := SOURCE has the sort of being one of the objects in DIMENSION

nocontrol(THEME) := THEME has the sort 'nocontrol', i.e. the theme does not control the continuation, path or speed of the movement denoted by the verb

A frame is thereby a statement about the external world: the external world should contain an object that is mapped by operations in the external world interpreting the attributes to other objects. The objects should stand in the external relations or have the external sorts that are imposed to them by the structure.

On top of this basic structure, there is information structure implemented by assigning or not assigning a property *new* to a feature. The interpretation is that features lacking this feature should be identified in the context (or accommodated), while new features give properly new information.

Furthermore, semantic features may be annotated for properties of the distribution. This is so already in the case of forced choices: a forced choice $x \in \{y_1, \ldots, y_n\}$ is just the features $x : y_1, \ldots, x : y_n$ but in a situation in which $x : y_i$ and $x : y_j$ are inconsistent for all different $i, j \leq n$. Much the same applies to implications. Other properties of that kind are *absolute* and *default*, which in this setting is understood as a feature that is part of some meanings in use, but can be omitted by conflicts with the context or in the case of competition between two or more incompatible features as the feature which is probabilistically dominant. Annotations with *new* and of this distributional kind are in small capitals. Below, we detail each one in turn.

First, there is a forced choice between the type of theme occurring with *fall* in any given instance (where the 'theme' is that which falls). The fact that this is a forced choice means that the verb obligatorily has a theme. In a given use, the theme will be resolved and represented as, e.g. THEME:light for an instance like *The light falls on the table.* 'Concrete' means that the theme is a concrete object such as a person.

(7) THEME $\in$ { concrete, light, precipitation, task, date, judgment, proposal}

The next two statements make the theme a non-agent of a non-action:

(8) nocontrol(THEME)
nocause(THEME)

Falling is strongly correlated with a lack of intentionality with respect to its direct cause, the movement and its path and for all contemporary languages considered, these are absolute features. These are also background features, as is the specification of the theme.

We next propose a *source* and a *position*, each of which is defined with respect to a *dimension*. The source can be understood as the point of departure for the falling, and the position as the theme's placement at the end of the falling act.

(9) SOURCE: location(DIMENSION)
POSITION: location(DIMENSION)
DIMENSION $\in$ {space, posture, life, health, moral, quantity, level, outcome(PROCESS)}

The inclusion of the dimension specification in our analysis is necessary to our account and is motivated by its further necessity in the analyses of the functions *direction* and *down*, among others (only *down* will be discussed in this particular paper). Dimensions are sets of positions ordered by a natural ordering relation. Here is what is meant by each of the types included above. *Space* is the set of spatial positions close to the earth ordered by the direction of gravity. *Posture* would be the set of body postures ordered by degree to which they are upright, and the same for postures of other things like walls, houses, poles, dogs, etc. It is a good idea to make posture a dependent sort (like *outcome(PROCESS)*), i.e. *posture(X)* where X should be filled in by the type of the theme. This assigns to each X a special set of postures. *Life* is a metaphorical transfer to the "postures" *alive* and *dead* where the first one corresponds with uprightness. *Health* includes healthiness and degrees of unhealthiness ordered from more to less healthy, *morality* is the set of moral states ordered from more to less moral, and *quantity* is the set of quantities ordered by the greater than relation. *Level* corresponds to the set of levels of something again ordered by 'greater than' on some numerical scale (see Lakoff and Johnson 1980 for suggestions about the origins of these metaphorical extensions).

There are a number of dimensions that have degraded into a set of down locations for an often not very clear source. Such cases are:

(10) The prize fell on Tim. (Tim won the prize.)
A cruel fate fell on those left behind. (Dutch: Those left behind suffered a cruel fate.)
Eating falls me difficult. (Dutch: I find it hard to eat.)
Christmas falls on Wednesday.
The task falls on me. (Dutch, German: This happens to be one of my tasks.)

Locations for prizes in races and lotteries are the winners, locations for fates the people whose fate they are, locations of holidays are days in the year assigned by the holiday definition (which may involve human decisions), tasks are one by virtue of one's office or of the moral order. Activities of somebody moreover assign a degree of difficulty or painfulness in the experience of that somebody. What these cases seem to have in common is a dimension that is just a set and a process or non-subjective procedure that assigns locations from the dimension to the theme.

Provisionally, we take these dimensions to be parametrically defined as *outcome(PROCESS)* where the identification of *PROCESS* is crucial for the identification of the set making up the dimension. For this last type, sources may be missing and be identified with the process itself.

One additional difficulty that arises with SOURCE pertains to cases such as *Her hair falls (perfectly)*, where the hair itself is not changing position, but rather, where two different parts of the hair are salient, and the ends of the hair 'fall' by comparison with the hair nearer to the crown of the head. In these cases, which also include falling paths, falling valleys, etc., we define the *split* relation where PART1 and PART2 are the higher and lower parts of the THEME, respectively:

(11) SOURCE=PART1
split(PART1, PART2, THEME)

For meanings in use in which *split* is defined, the specification that SOURCE=PART1 is obligatory in the representation (i.e. it is an absolute feature).

The next component *at* relates two features such that $at(x, y)$ can be paraphrased as 'the x is in/on y'. For example, $at(John, lying\_down)$ indicates that John is lying down. *At* is a component that appears multiple times in each representation. One of its instances is given information, and one is new information. The one that is given is as follows, where *down* picks out the set of elements in the order given by the dimension that are lower than the one named by the source:

(12) at(POSITION, down(SOURCE,DIMENSION))

For example, $at(lying\_down, down(standing, posture))$ would mean that the lying down is 'down' (lower on the order for the posture dimension) from standing. Another way of saying this is that lying down is down from standing with respect to posture. The next instance of *at* is new information:

(13) at(THEME,POSITION)

This is like the example above, where $at(John, lying\_down)$ indicates that John is lying down. Taken in combination with the given/presupposed use of *at*, we can see that for an example like *John fell*, part of the given information is that lying down is a lower posture position than standing, but the new information includes the fact that John is in fact now in the lying down position. This is sometimes all that is specified as new information, but there are other cases where a movement event is also new information. These are the cases in which a source is either specified or implicit. Continuing this example, the source would be specified as ‘standing’ (on the posture dimension), and thus we would have new information that there was an event in which John moved from standing to lying down, following this specification:

(14) movement(THEME,SOURCE,POSITION)

In cases where *split* is defined, having the new information be $at(THEME, POSITION)$ would be problematic, since it is not the entirety of the theme that is in the lower-ordered position. Thus, in these cases, we have the following absolute feature, which is analogous to the other but with PART2 replacing THEME:

(15) at(PART2, POSITION)

(16) recapitulates the above with the universal level labels and the bar indicating the split between given and new. It also indicates that space is the default dimension and that *at(THEME,POSITION)* is the default for new information.

(16)
THEME ∈ {concrete,light,precipitation,task,date,judgment,proposal} Absolute
nocause(THEME)
nocontrol(THEME) Absolute
SOURCE: location(DIMENSION)
POSITION: location(DIMENSION)
DIMENSION ∈ { space:default, posture, life, health, moral, quantity,
level, outcome(PROCESS)} Absolute
SOURCE= PART1 **If** split Absolute
split(PART1, PART2, THEME)
at(POSITION, down(SOURCE,DIMENSION)) Absolute

---

at(THEME,POSITION) Default, New
at(PART2, POSITION) **If** split Absolute, New
movement(THEME,SOURCE) **If** at(THEME, SOURCE) Absolute New

We will now discuss our data with respect to this representation. First of all, there were not many differences between the languages we examined with respect to these representations. The full list is discussed here.

A. The dimension *outcome(PROCESS)* is prominent, especially in Dutch, as in (17).

(17) De prijs viel op mij.
The prize fell on me.
‘I received the prize’.

B. Older German (Grimm 2011) includes an example where the person doing the falling was causing an action as in (18).

(18) Er fiel in die Sachsen.
He fell into the Saxons
‘He wildly attacked the Saxons.’

This data point is the reason that $nocause(THEME)$ is not listed as an absolute feature above.

C. Finally, English allows a source without a destination/position (“Grief fell from his heart”), unlike the other languages considered. There is one further important type of difference among the languages, which will be discussed in Section 5 and force a major revision.

These three differences indicate that small adaptations must be made to obtain the specification for particular languages. Dutch overuse of *outcome(PROCESS)* and pre-modern German's causality with *fall* are two cases where one predicts failure of comprehension between languages. Beyond that, however, speakers of one language should be able to make sense of all the uses of the other languages. Languages that are less related than those under discussion may, however, differ more greatly, which is something that we are examining in our continued research.

If information from the context of the word is taken into account, the formal model proposed here makes sense of all the uses we collected on the basis of the language-independent specification and perhaps surprisingly lends itself to implementation. Particularly important is the type of the theme and of the source or location which restricts the choice of the dimension and, thereby, the sort of the unspecified source or location (if either of these is in fact unspecified). Finding this information is easy using prepositions and parsing. It is harder to use information that is not syntactically coded, but clearly often necessary: *he fell* can mean many things given the right context. While the specification is good at suggesting the right questions to ask the context, the answers cannot always be supplied by a simple heuristic method defined over the context. Despite this, trial implementations by Jonathan Mallinson and Jacob Verdegaal show that good results are possible when the necessary contextual information is given syntactically and lexically. It then works to the degree that the syntactic and lexical analysis is correct. This would be the same in cases where one cannot rely on lexical or syntactic information and identifications between variables in the lexical specification and elements in the interpreted linguistic context need to be inferred. The difference is that there are no good off-the-shelf systems for doing these inferences.

We now proceed through a set of examples taken from our collection of different uses of *fall* from Dutch, English, French, German, and Russian. The inferences are generally trivial, but this partly reflects the source of the examples, since dictionaries often give examples in which it is not necessary to use further context beyond the clause. The annotation NEW is replaced below by a double line: the new features are below the double line.

For the sentence given in (19), *The glass fell on the floor*, the space dimension is the default and the floor is a location in that dimension. The glass is obviously somewhere (on the table, in somebody's hands), which may be given in the linguistic context. Glass is moreover a concrete noun. Together, this selects the specification in (19).

(19) 'The glass fell on the floor.'

THEME:glass
glass:concrete
nocontrol(glass)
nocause(glass)
SOURCE:location(space)
at(floor,down(SOURCE,space))

at(glass,floor)
movement(glass,SOURCE)

Continuing with the example used to explain the representations above, in (20), the context needs to put John in a "low" location (to prevent the spatial dimension) and a "high" posture, e.g. standing. Whether one assumes that John falls to the floor, in which case he is lying down at the end of the event, or falls into a chair, in which case he is sitting at the end of the event is a question of what the default value is for the individual hearer in the posture dimension below standing. Note that defaults of this kind are not indicated in the abstract specification above and should be inferred from the context.

(20) ‘John fell’.

THEME:John
John:concrete
nocontrol(John)
nocause(John)
SOURCE: standing
POSITION: lying_down
lying_down:location(posture)
at(lying_down,down(standing,posture))

at(John,lying_down)
movement(John,standing)

(21) involves a further specialization of posture applied to houses.

(21) ‘The house fell.’

THEME:house
house:concrete
nocontrol(house)
nocause(house)
SOURCE:erect
POSITION:collapsed
collapsed:location(posture)
at(collapsed,down(erect,posture))

at(house,collapsed)
movement(house,erect)

(22) presents another specialization of dimension restricted to military people and battles. This needs a special constraint in the specification of *fall* via the proposed ‘life’ dimension: If somebody military is the theme, the movement and its cause are part of a battle, then the dimension can be life.

(22) ‘The corporal fell.’

THEME:corporal
corporal:concrete
nocontrol(corporal)
nocause(corporal)
SOURCE:alive
POSITION:dead
dead:location(life)
at(dead,down(alive,life))

at(corporal,dead)
movement(corporal,alive)

(23) would be derivable through the theme type *precipitation*, which entails the source to be the sky and the dimension to be *space*. Precipitation differs from people, stones and houses by not being a spatio-temporal continuant.

(23) ‘The rain fell.’

THEME:rain
rain:precipitation
nocontrol(rain)
nocause(rain)
SOURCE:sky
POSITION:ground
ground:location(space)
at(ground,down(sky,space))

---

at(rain,ground)
movement(rain,sky)

Dates induce a stative use of *fall* and invoke the *outcome(process)* dimension on a process called *calendar*.

(24) 'Christmas day falls on a Sunday this year.'

THEME:Christmas
Christmas: date
nocontrol(Christmas)
nocause(Christmas)
SOURCE: location(outcome(calendar))
POSITION: Sunday
at(Sunday,down(calendar,outcome(calendar)))

---

at(Christmas, Sunday)

A number of Dutch uses involve other kinds of processes that lead something to be at some location. A Dutch use where the subject determines the process is given in (25).

(25) De taak valt op mij.
the task fall on me
'It is my task.'

nocontrol(task)
nocause(task)
ME:location(outcome(taskassignment))
at(ME,down(taskassignment,outcome(taskassignment))

---

at(task,ME)

This is not the place for overly long explanations of the formalism used above, but some discussion is necessary to make at least some connections with truth-conditional semantics. Dimensions would be modeled as ordered sets of various kinds. The hardest is here the default setting, normal space; yet, here are by now many formalisms to deal with space and a gravity based high-low ordering over locations (See Aiello et al. 2007). The other dimensions are very limited in comparison, they essentially are small finite partial orders: *life* and *health* have only two elements, while the complexity of the *outcome(process)* dimension is mainly in stating the range of possible outcome locations, e.g. different participants in a lottery, different people who can be burdened by some task, success, ease, hardship and failure for eating, etc.

Notions like *control*, *cause*, *theme* and *at* seem to be proper universal semantic features. The first two are amenable to a treatment of cause like e.g. the one pioneered by Pearl (2000) (*control* would be the ability to change the course of the event or change the state if whatever has control would want to). The theme would be a Dowty (1991)-style decomposition in similar semantic features. Finally, *at* would be the relation between between objects and where they are, in space or in a metaphorical extension.

These remarks are not meant as a truth-conditional treatment of the concepts we develop, but are meant to take away worries in that respect: this subject as been successfully addressed and there is no reason for thinking that a truth-conditional account cannot be given. In fact, such accounts will considerably help learning systems for the word-feature associations by providing a criterion of consistency for feature bundles.

In this section we provided an overspecified lexical representation for the verb *fall* which accounts for all occurrences we found in our data set. In the next section we discuss, however, that this analysis runs into problems when we look at the second verb we investigate in the project: *run*.

5 Observed Production Probabilities and Lexicalisation

The following problem emerges when we turn to *run*, the second verb in the project sample after *fall*. While there is no significant conceptual distinction between English *run* and Dutch and German *rennen* in their primary uses, there are nonetheless very significant divergences between the verbs in their special uses. In English, machines and noses run, while in German and Dutch these objects engage in *lopen*, for which the best English equivalent is *walking*. The logic for deriving meanings in use from the previous section, however, derives the English meaning in use for "De machine rent", "De neus rent", "Der Machine rennt", and "Die Nase rennt".

It is just a brute fact that, in these cases, the Dutch and German verbs do not have these meanings in use. The reason why is obvious: because another verb has won in these cases as the preferred means of expressing the meaning in use. And the formalism should be able to express this. It is, however, not easy to come up with a natural method extending the equivalence classes on distributions over semantic features in which this can be directly stated.

It is an overgeneration problem: more meanings in use are predicted than are observed. Yet—at the same time—the logic and the over-determined lexical specifications seem on the right track.

The correct way to rule out unwanted meanings in use in a probabilistic setting is to add production constraints in interpretation. It seems a correct observation that Dutch and German speakers select *lopen* or *laufen* when confronted with the meaning in use given by English: "the machine runs" or "the nose runs". This section recasts the previous proposal to incorporate this dependency on production.

There should be a part of the production mechanism which assigns words to bundles of features with a certain probability: lexical selection. This process can be captured as a function f that maps pairs made up from a bundle of semantic features and a word to a probability. The function values will often be 0 for all words for a certain bundle: the bundle is not a sensible meaning in use or it is sensible but lexemes are missing. It will also very often be 0 for most words: those words were never used to express this bundle. The function can be read off directly from a corpus of word and meaning in use pairs as in (26). The function gives the frequency of the word for the bundle divided by the frequency of the bundle. Notice that the precise bundles of features count: a use of w to express a superset c of b does not count as a use of w for b. We assume here that the corpus is given as a set C of triples <index, bundle, word>.

(26) $$f(b)(w) = \frac{|\{j: <j,b,w> \in C\}|}{|\{j: \exists w' <j,b,w'> \in C\}|}$$

The function counts the number of indices at which w expresses b and divides that number by the number of indices at which b is expressed by any word whatsoever. The function therefore measures the strength with which a bundle of features keys a lexeme and can be seen as a component of what determines lexical choice. The rule could be to choose that w for b for which $f(b)(w)$ is maximal.

The function f—or the data from which it can be read off—can be equated with the mental lexicon. The mental lexicon cannot be equated with the set of distributions over semantic features keyed by specific lexical items as was assumed in the last section, since that does not give a handle on production, which we need to deal with our overgeneration problem, as will be shown below.

There are no problems with multi-word lexical items. In fact, it is natural to assume that certain sets of features would correspond to groups of lexemes. All that one would need to do is to consider a generalization of the function in which w ranges over bags of words. There would then no longer be any principled division between multi-word lexemes and groups of lexemes that jointly express the bundle, and this is as it should be. The difference would be in the possibility to regain the probability of the bag for the bundle from the probabilities of its components for parts of the bundle: do we get the same number or is the probability of the combination higher? In the last case, it has become or is becoming an idiom. While this definitely must be explored further, let's ignore composition for the time being and let w simply range over words. It would seem that the revision fares better as a data structure that can be learned and represented by the brain as an association and is better suited for multi-word expressions than the proposal from the last section to equate the mental lexicon with probability distributions $p_w(b)$ over the words w of the language. While we leave these issues open, it is still the case that the competition we define below, sometimes is between a word and a multi-word expression or between multi-word expressions.

This new account of the mental lexicon can help with the problem at hand. If a meaning in use is just expressed differently, it cannot be the meaning in use of the word for which it was hypothesized (cf. the use of *lopen* 'walk' rather than *rennen* 'run'). The earlier approach does not need to be given up, it merely needs an amendment in which it is checked for an interpretation b that the production probability $f(b)(w) \neq 0$ while for another word w' $f(b)(w') > 0$.

But then, what should be used for arriving at meanings in use? At first sight this seems problematic. The revision gives a criterion for having found the meaning in use. The new data structure gives the likelihood $p(w|b)$ of the word w for the meaning in use b. If a prior $p(b)$ for the set of features b is given, the most probable interpretation is $argmax_b p(b)f(b)(w)$, the feature bundle b for which the product of the prior probability $p(b)$ and the likelihood as given by f applied to b and w is maximal.[5] But how can it be guaranteed that that is indeed the maximal bundle that fits in the context? $p(b)$ will be bounded by $p(b')$ if $b' \subset b$. And the likelihood is learned from experience as f. It would seem that this does not give the prediction that larger bundles are preferred. A second problem is that it gives no results in case w has not been used for b before.

The solution is not to give up on the earlier proposal but to use it as a model of Bayesian interpretation. This can be done since f almost directly reconstructs the distribution p_w over semantic feature bundles keyed by a lexeme w. In (27), $p_w(b)$ is defined, where b is a semantic feature bundle. To do this correctly, we need to measure the frequency of b, $m(b)$, in the corpus as its frequency divided by the corpus size. F is the set of all features.

(27) $$p_w(b) = \frac{\Sigma_{b \subseteq b' \subseteq F} m(b')f(b')(w)}{\Sigma_{b' \subseteq F} m(b')f(b')(w)}$$

[5]That this is the most probable interpretation follows by Bayes' theorem. Models of interpretation in which the most probable interpretation is computed by finding a maximum for $p(b)p(w|b)$, the product of the prior and the likelihood of the interpretation are standard in signal processing and computer vision.

$p_w(b)$ measures how often w is used for b within all uses of w.

While humans learn f automatically, it seems that linguists like ourselves are better in discovering p_w directly from lexical and internet data, with semantic blocking occasionally offering a window on properties of f that cannot be recovered from the functions p_w.

Lexical interpretation is still computing $argmax_b p(b)p(w|b)$, but there will be more information in $p(w|b)$ than directly follows from f. Any semantic feature $s \in b$ associated with w helps to increase the likelihood of w: $p(w|b)$. But simultaneously $f(b)(w)$ is a filter on the result, blocking certain realisations, i.e. this can make $p(w|b) = 0$. If there are $f(b)(w')$ with a high value, b has lexical means of expression and using w for b will be unlikely if $f(b)(w) = 0$. So if $f(b)(w)$ has a high value while $f(b)(w') = 0$, b is not a proper interpretation of w'.[6] So the new "mental lexicon" can continue as the base for the solution in Section 4.

Let us recapitulate these observations in some definitions:

(28) **Lexical interpretation**: b is a lexical interpretation of w iff $f(b)(w) >> 0$.[7]

(29) **Standard lexical choice**: w is a lexical choice for b iff $f(b)(w) >> 0$.

(30) **Smolensky/Hogeweg interpretation**: An optimal interpretation can be defined by the following three constraints:

1. All variables are bound from the context, all forced selections are executed, all absolute features projected, and the interpretation is closed under *modus ponens*
2. It is consistent with the context
3. It is maximal

In terms of OT, (1) gives conditions on candidate interpretations and makes these conditions thereby absolute. (2) and (3) are identical with FIT and STRENGTH and can be defeated: a new statement may correct the context, non-absolute features can be dropped to gain consistency. (2) is entailed by prior maximisation in Bayesian interpretation. (3) is part of likelihood maximisation. So for practical purposes and within the enterprise of computing lexical meanings in use from abstract specifications, it seems reasonable to equate Bayesian interpretation and Smolensky/Hogeweg.

(31) **Proper lexical choice**: w is a proper lexical choice for b in c if w is a lexical choice for b and b is a proper interpretation of w in c

Many bs will not have a lexical choice: in that case w is a proper lexical choice for b in c iff b is a proper interpretation of w in c

(32) **Proper interpretation**: b is a proper interpretation of w in c iff b is computed by Smolensky/Hogeweg for w in c unless there exists $w' \neq w$ where w' is a lexical choice for b while w is not a lexical choice.

The situation in which the interpretation b found for w cannot be lexically expressed (no word has a non-zero observed probability for it) is interesting, because now it is reasonable to stick with the hypothesis found by Bayesian interpretation, i.e. FIT > STRENGTH that the meaning in use is b: a new meaning in use for the word w was found. This would be an extension of the use observed so far.

The information in f is typically partially reflecting ongoing learning. Every 0 may have the

[6]This gives a simple intuitive solution for the "cause to die"-problem of McCawley (1968). "Black Bill caused the sheriff to die" cannot "mean" under this rule that he caused him to die in some normal way. "Cause-to-die" is blocked to mean that by the lexical expression "kill".

[7]We require a proper number of occurrences, beyond what could be attributed to error.

meaning that the use of w for b has not been observed so far. But there are situations where the 0 can be taken seriously. The first case would be for inconsistent bundles. The second case would be the case where the bundle b has been expressed often enough to be confident that w will not be used for the bundle. In all other cases, one can learn that w is used for b by encountering a use and inferring that b is its meaning in use. And one use is good enough.

Like Hogeweg (2009), we assume that there is the beginning of an account of metaphorical use of lexical items in this setup. The information in f can block or select an hypothesis obtained by the reasoning that computes meanings in use from lexical specifications, the new version of FIT > MAX. It however does not block unobserved hypotheses and it should not. These are—or are from the perspective of the learning structure f—metaphorical extensions. Since learning f is also learning the distribution over semantic features p_w, it follows that new hypotheses can also be obtained by overriding zero's in that distribution. The latter are new metaphors.

Now it is not easy, but possible to come up with new metaphors for $fall$. For example, in the project we might perhaps say (speaking Dutch or Russian) that the word *rennen* falls to Lotte, meaning that it is her task to collect uses of *rennen*. The specification rules that out even for Dutch: words are not associated conventionally with a process that assigns them to humans. But as a project member, Lotte can be assigned tasks and in the project words are tasks: a new metaphor. It would seem that one can deal with cases like this by shifting from a definition by listing of process (necessary for Dutch: many processes work, but not all) to an intensional characterisation (task or reward assigning process). A more proper exploration of these limits and ways of overcoming them is for future work.

Accordingly, we find ourselves siding with Giambattista Vico in claiming that originally—at least in acquisition—language use is poetry in which everything is interpreted metaphorically (Pompa, 2002). Learning from use slowly leads to prose, i.e. the semantic discipline brought by conventional means of expression emerging from experience.

6 Conclusion

We have argued that the traditional view on the lexicon does not offer a way of accounting for the selection of meanings in use, a task that humans seem to perform routinely with high degrees of success. The proposals of Smolensky and Hogeweg for dealing with selection by means of overspecification were then examined and found to be wanting for the meaning of lexical words. We argue that decomposition of word meanings and meanings in use in terms of moderately universal semantic features is possible and consistent with truth-conditional semantics and typology. But that more structure is needed over the features than just set membership.

The method can be made to work in a natural way, if rather than a set of semantic features one uses equivalence classes over distributions over bundles of semantic features expressed by the words. Such equivalence classes offer a natural inventory of operations over semantic features and we show that with these operations, one can arrive at a natural and effective representation for the verb $fall$ that can be used to model the interaction with the context that performs selection.

The approach however overgenerates, since in many cases the interaction with the context will yield feature bundles that should be expressed differently in certain languages. In order to remedy that we propose that lexical representation takes the form of two functions: f that maps feature bundles and words to a probability and m that maps feature bundles to the probability that they will be expressed. It is now possible to define the necessary semantic blocking as the requirement that f should not give zero for the interpretation b and the word w, while giving a high value to b and w'. At the same time, the distributions p_w over semantic features can be recovered from f and m and offer—like the Smolensky/Hogeweg proposal does—the first steps of an account of new metaphors.

References

Aiello, M., Pratt-Hartmann, I., and van Benthem, J., editors (2007). *Handbook of Spatial Logics*. Springer, Dordrecht.

Ait-Kaci, H. and Podelski, A. (1993). Towards a meaning of life. *The Journal of Logic Programming*, 16(3):195–234.

Baroni, M., Bernardi, R., and Zamparelli, R. (2014). Frege in space: A program for compositional distributional semantics. *Linguistic Issues in Language Technology*, 9(6):5–110.

Barsalou, L. (1992). Frames, concepts, and conceptual fields. In Kittay, E. and Lehrer, A., editors, *Frames, fields, and contrasts: New essays in semantic and lexical organization*, pages 21–74. Lawrence Erlbaum, Hillsdale, NJ.

Blutner, R. (2004). Pragmatics and the lexicon. In Horn, L. R. and Ward, G., editors, *Handbook of Pragmatics*, pages 488–514. Blackwell, Oxford.

Carston, R. (2010). Explicit communication and free pragmatic enrichment. In Soria, B. and Romero, E., editors, *Explicit Communication: Robyn Carston's Pragmatics*, pages 217–285. Palgrave Macmillan, Basingstoke.

de Schepper, K. and Zwarts, J. (2010). Modal geometry: Remarks on the structure of a modal map. In *Cross-linguistic Semantics of Tense, Aspect, and Modality*, pages 245–270. John Benjamins, Amsterdam.

Dowty, D. (1991). Thematic proto-roles and argument selection. *Language*, 67(3):547–619.

Erk, K. (2014). What do you know about an alligator when you know the company it keeps? Unpublished ms.

Fodor, J. and Lepore, E. (2002). *The Compositionality Papers*. Clarendon Press, Oxford.

Fodor, J. and McLaughlin, B. (1991). Connectionism and the problem of systematicity: why Smolensky's solution doesn't work. In Horgan, T. and Tienson, J., editors, *Connectionism and the Philosophy of Mind*, pages 331–355. MIT Press/Bradford Books, Cambridge.

Fodor, J. and Pylyshyn, Z. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, 28(1-2):3–71.

Fong, V. (2001). "Already" in Singapore English. Abstract, Aspect Conference, Utrecht, December 12-14, 2001.

Grimm, S. (2011). Semantics of case. *Morphology*, 21(3-4):515–544.

Haspelmath, M. (2003). The geometry of grammatical meaning: semantic maps and cross-linguistic comparison. In Tomasello, M., editor, *The New Psychology of Language*, pages 211–243. Lawrence Erlbaum, Mahwah, N.J.

Hogeweg, L. (2009). *Word in Process. On the interpretation, acquisition and production of words*. PhD thesis, Radboud University Nijmegen.

Jackendoff, R. (1996). Conceptual semantics and cognitive semantics. *Cognitive Linguistics*, 7:93–129.

Johnson, M. (1988). *Attribute-Value Logic and the Theory of Grammar.* CSLI Lecture Notes. Center for the Study of Language and Information, Stanford, CA.

Kamp, H. and Reyle, U. (1993). *From Discourse to Logic.* Kluwer, Dordrecht.

Lakoff, G. and Johnson, M. (1980). *Metaphors We Live By.* University of Chicago Press, Chicago.

Loebner, S. (2014). *Understanding Semantics.* Routledge.

McCawley, J. D. (1968). Lexical insertion in a transformational grammar without deep structure. In Darden, B., Bailey, C.-J., and Davison, A., editors, *Papers from the Fourth Regional Meeting of the Chicago Linguistic Society*, pages 71–80. Linguistics Department, University of Chicago.

McDonough, L., Choi, S., and Mandler, J. M. (2003). Understanding spatial relations: Flexible infants, lexical adults. *Cognitive psychology*, 46(3):229–259.

Pearl, J. (2000). *Causality: Models, Reasoning, and Inference.* Cambridge University Press, Cambridge.

Petersen, W. and Werning, M. (2007). Conceptual fingerprints: Lexical decomposition by means of frames - a neurocognitive model. In Priss, U., Polovina, S., and Hill, R., editors, *Conceptual Structures: Knowledge Architectures for Smart Applications: Proceedings of ICCS 2007*, Lecture Notes in Computer Science, pages 415–428. Springer.

Pompa, L., editor (2002). *Scienza nuova (1725), The First New Science.* Cambridge University Press, Cambridge.

Pustejovsky, J. (1995). *The Generative Lexicon.* MIT Press, Cambridge.

Sag, I., Wasow, T., and Bender, E. M. (2003). *Syntactic Theory: A Formal Introduction.* CSLI Lecture Notes. Center for the Study of Language and Information, Stanford, CA.

Smolensky, P. (1991). Connectionism, constituency and the language of thought. In Loewer, M. and Rey, G., editors, *Meaning in Mind: Fodor and His Critics*, pages 201–227. Blackwell, Oxford.

Zwarts, J. (2004). Competition between word meanings: The polysemy of (a)round. In Meier, C. and Weisgerber, M., editors, *Proceedings of SuB8*, University of Konstanz Linguistics Working Papers, pages 349–360. Konstanz.

Zwarts, J. (2006). Om en rond: Een semantische vergelijking. *Nederlandse Taalkunde*, 11(2):101–123.

Zwarts, J. (2008). Priorities in the production of prepositions. In Asbury, A., Dotlacil, J., Gehrke, B., and Nouwen, R., editors, *Syntax and Semantics of Spatial P*, pages 85–102. John Benjamins, Amsterdam.