Distributional semantics and the study of (a)telicity

Aleksandar Trklja, University of Innsbruck

In the literature it is argued that distributional semantics can provide a comprehensive model of lexical meaning. The present paper challenges this assumption and argues that the issue of semantic similarity cannot be fully addressed unless the denotation of terms is systematically examined, since the distributional approach on its own lacks the methodological and conceptual resources to pursue this task. The present proposal is an approach in which distributional semantics meets event semantics. Through an analysis of the verbs of creation it is demonstrated that shared distributional properties are not indicative of shared denotations. In other words, the data shows that the former can serve as an approximation, but cannot predict shared denotations. Finally, due to the combinatorial nature of the lexicon, it is likely that lexical combinations with the same semantics represent an exception rather than a rule.

Keywords: *(a)telicity, creation verbs, distributional semantics, event structure, lexical semantics*

1. Introduction

According to the distributional hypothesis (Harris 1954; Lenci 2008), words that occur in the same textual context tend to have a similar meaning. In lexical semantics this hypothesis has been explored in large corpora in terms of the distributional properties of lexical items. Words that share the same collocates are considered to be semantically similar. The larger the number of shared collocates, the more similar words will be. Although a number of corpus linguistic approaches (e.g. Stefanowitsch & Gries 2003; Divjak 2010; Trklja 2014, 2017) and computational models (e.g. Sahlgren 2006; Baroni and Lenci 2008; Erk & Pado 2008; Baroni 2013; Copestake & Herbelot 2012) which follow the distributional hypothesis have been proposed, some of the fundamental issues have not yet received sufficient consideration. In particular, I refer here to the issues addressed in the following questions: What does it mean to say that words have a similar meaning? To what extent is the distributional hypothesis supported by evidence from other research areas (e.g. Landauer & Dumais 1997)? Is distributional semantics part of semantics proper at all (e.g. Lenci 2008; Westera & Boleda 2019)? These questions are, of course, too broad to be answered in a single study. The main objective of the current paper is to contribute to the discussion by exploring whether shared textual contexts necessarily reflect shared denotations. I will argue that the distributional approach lacks the methodological and conceptual resources to address denotation-related issues (not to mention intension), and that it can deliver more reliable evidence of lexical meaning only if it is combined with other approaches to (lexical) semantics.[1] In particular, I will claim that the issue of shared denotation cannot be comprehensively explored if the lexical aspect or (a)telicity of words is not properly addressed This view agrees with Lenci's "weak view" (2008), and stands in contrast to the view that "distributional semantics on its own can in fact be a fully satisfactory model of expression meaning" (Westera & Boleda 2019:122). Wedera and Boleda adopt Strawson's pragmatic position that speakers and not words make

[1] Baroni et al. (2014) and Copestake & Herbelot (2012) address this issue, but their main focus is on other matters.

reference, and they reject the usefulness of the notions of reference, truth conditions and entailment for the representation of meaning. Setting aside the problem that this move does not make life easier for distributional semantics, because in its current form it is not capable of modelling speaker meaning (Boleda 2020), what is more important is that it still does not show how distributional semantics can account for the lexical meaning understood in the sense of lexical decomposition. Nor does it explain how we refer to events and situations in language, which is, as discussed above, a function of language well-known in linguistics since Panini (Parsons 1990).

The paper is divided into three parts. Section 2 is slightly unusual for a study concerned with distributional semantics. Instead of discussing at length the fundamental principles of distributional semantics, these will be introduced only in a general manner so that an uninformed reader can comprehend the main issues discussed. The main part of this section will be devoted to the discussion of what is meant by lexical distribution and why is it important for lexical semantics. The distributional properties of lexical items in this discussion are modelled in terms of two kinds of relations: the relation of collaboration and the relation of competition. In Section 3, the lexical aspect approach and the notion of (a)telicity are outlined, and in Section 4 the evidence for the argument that distributional semantics cannot comprehensively account for lexical meaning is provided through a distributional and (a)telicity analysis of creation verbs.

I assume that distribution and (a)telicity are not merely methodological or theoretical constructs. Instead, I regard them as part of the speaker's knowledge of meaning that comprises both statistical learning and knowledge about events. Due to space restrictions, an argument for this view can only be briefly outlined here.

Language processing associated with statistical learning is incremental, and underlined by domain-general cognitive processes that include rich memory storage, categorization, clustering, analogy and possibly some other processes (Tomasello 2003; Bybee 2010; McCauley & Christiansen 2017). As demonstrated in Christiansen and Chater (2016), statistical learning is involved in the process of the emergence of lexico-grammatical structures. It follows that the speaker's knowledge also contains information about the distribution of words. Thus, in addition to grammar-induced knowledge (regardless of whether we assume the generativist or constructivist position here), this knowledge also contains awareness of the likelihood of co-occurrence of lexical items.

However, statistical learning cannot account for lexical knowledge on its own. Developmental studies provide evidence that the 'core knowledge' of humans also contains cognitive systems for the representation of objects and events (Spelke & Kinzler 2007; Radvansky & Zacks 2014). Radvansky & Zacks (2014) thus demonstrate that events are not processed holistically but rather compositionally. Human experience activities are chunked into smaller units that consist of some constitutive elements including sub-events. Although as pieces of cognitive representations events are unique, they are part of a more general system of knowledge representation in which individual events are stored into event schemes. These schemes are based on our previous experience with certain events. The events that share relevant features are stored in the long-term memory as types. The notion of (a)telicity becomes relevant at this point because it refers to semantic differences between the concepts of end, limit and boundary, which are encoded into event schemes, and there is some evidence that "there may be a universal way of mentally representing events as containing a logical endpoint

or as consisting of homogenous subparts lacking such an endpoint" (Strickland et al. 2015: 5971).

2. Collaboration and competition between words

The distributional approach to meaning was introduced independently but almost simultaneously by Harris (1952, 1954) and Firth (1968). The general form of the argument is that the meaning can be induced from the textual context in which words occur. The meaning of a word is here regarded as "a function of the contexts in which it occurs" (Boleda & Herbelot 2017: 623). In particular, it is assumed that "there is a correlation between distributional similarity and meaning similarity, which allows us to utilize the former in order to estimate the latter" (Sahlgren 2008:1).

In more recent distributional approaches the meaning of words is explored by means of corpus linguistics tools. The meaning is represented in terms of vector space representations dating back to the 1960s (Salton & McGill 1983). Word relations are represented in terms of proximity as vectors in a high-dimensional space, and the similarity of related words is measured in terms of the cosine of the angle between vectors representing words (Mikolov et al. 2013). The values of the cosine of the angle are indicative of distance. If two terms are more similar,than other pairs the distance between these values will be lower and *vice versa* (Sahlgran 2006). Figure 1 (adopted from Baroni 2013) illustrates this view with a toy lexicon that consists of the nouns *dog*, *cat* and*car* and the verbs *run* and *leg*. The words *dog* and *cat* create a strong collocation with *legs,* and *car* creates a strong collocation with *runs*. The word *dog* is, therefore, regarded as more similar to *cat* than either of the two words to *car*.

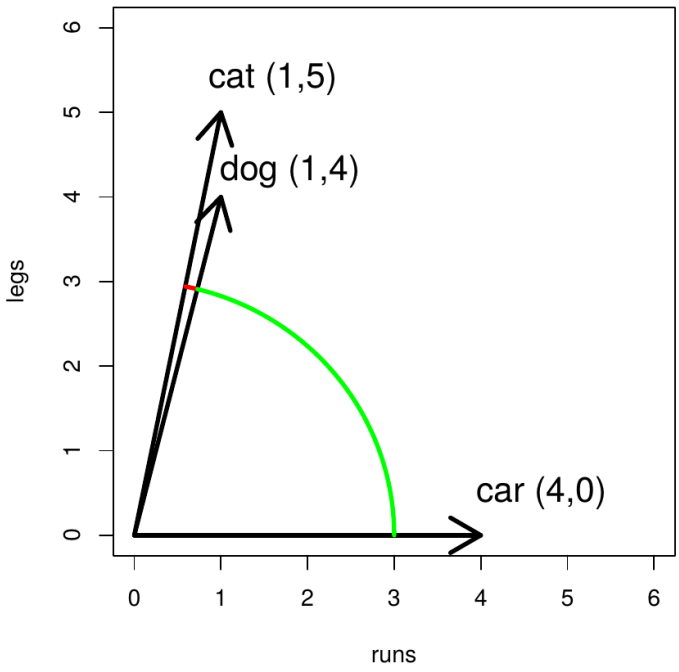

Figure 1: An example of a semantic vector

Distributional approaches take for granted that distributional properties govern language use, but the relation between distribution and language use is rarely explicitly considered. Furthermore, the notion of distribution has been subject to different interpretations in linguistics. In distributional semantics distribution is considered in terms of co-occurrence relations. To understand how co-occurrence relations determine language use, it is worth considering distributional properties as such in general terms. The distributional properties of lexical items belong to the realm of combinatorics, and to explore them I will model the collocation relations in terms of the following three kinds of entities: *node term*, *collaborators* and *competitors*. A node term is any word that we select as our starting point in the analysis. Collaboration is a type of relation determined with respect to the node term. Terms X_1 and X_2 that co-occur with the node term A are said to be its collaborators. Competitiveness is a type of relation determined with respect to the shared context of terms. Terms X_1 and X_2 that occur in the same context (with A) are said to be each other's competitors; X_1 and X_2 compete over co-occurring with A.

Let us code node terms as A, B, C and collaborators/competitors as $X_1 \ldots X_n$. X_1 is a collaborator with respect to A, but it is a competitor with respect to $X_2 \ldots X_n$. Now let us assume that the node term A occurs five times in our data, and that it has only one collaborator: (X_1). This means that A's collaborator can also only occur five times, and the co-occurrence probability is equal to 1. Since the co-occurrence probability measures how strongly a node term and its collaborator(s) are co-associated, I will refer to this relation as *collaboration strength*. When a node word has only one collaborator the collaboration strength has a maximum value. In real life, the cases when a word or a multi-word expression has only one collaborator are very rare. If the frequency of A remains the same but the number of collaborators increases to two (X_1 and X_2), then both the collaboration strength and the number of co-occurrence combinations between these terms change. Instead of one there are now four combinations. Provided we remain in the domain of natural numbers (which is only meaningful when dealing with the frequency of occurrence), the following combinations are possible: if X_1 occurs only once then X_2 will occur four times; if X_1 occurs twice then X_2 will occur three times; if X_1 occurs three times then X_2 will occur twice; and if X_1 occurs four times then X_2 will occur once. Clearly the collaboration strength for X_1 (based only on the observation of the occurrence in the present data) is higher in the latter two combinations than in the former two. The rate of change from one to two collaborators of A is displayed in the first two rows of Table 1 below.

Table 1: Relations between A and its collaborators

number of collaborators of A	number of combinations k	size of sample n					
1	1	5	1				
2	4	5	0.2	0.4	0.6	0.8	
3	5	5	0.2	0.2	0.2	0.4	0.6
4	4	5	0.2	0.2	0.2	0.4	
5	1	5	0.2				

In this table other options are also considered. The first column displays the number of collaborators A has at different points in time. The second column shows the number of combinations associated with the number of collaborators, and the third column shows the frequency of A. The right-hand side panel shows the probabilities of the occurrence for any individual X (e.g. X_1). The first row models the option when A co-occurs only with X_1. The second row displays the option when A co-occurs with X_1 or X_2. As was said above, there will be four possible combinations here, and the likelihoods of the occurrence of X_1 in these combinations are 0.2, 0.4, 0.6 and 0.8.

It can be observed that the likelihood that $X_1 \ldots X_n$ will occur with high frequency with A decreases as the number of collaborators of A increases. For instance, there is a 25% chance that X_1 will collaborate with A in 80% of its occurrences if the number of collaborators is two. Similarly, there is a 25% chance that X_1 will occur in 60% of occurrences of A if it has three competitors. As the number of competitors increases, the probability that X_1 will occur with lower frequency also increases. In addition, the smaller the number of competitors and the lower the collaboration strength of other collaborators ($X_2 \ldots X_n$), the more likely it will be for the item X_1 to co-occur with A with high frequency. Thus, if A has four collaborators, the chances are that in three out of four possible combinations X_1 will make up 20% of the occurrences of A. (Tables 1 and 2 in the appendix provide further examples that support this observation.)

From the point of view of any collaborator, having no competitors would be an optimal situation. This would mean that whenever A occurs X_1 will occur as well. However, in this scenario, every word would have only one meaning, there would be far more words than we actually have in natural languages, and if we assume that the meaning is compositional (one of the fundamental assumptions in semantics) then for any combination of words regardless of its length we would have one usage per combination. For example, in an invented language of this sort for *create a reasonable comment* and *create a critical comment* there would be two expressions with lexically distinctive constituents such as *bakara bir opalable porop* and *matara bir zapalble sarop.* In another equally implausible scenario one could imagine that all words occur with the same likelihood. If A occurs six times and has three collaborators (X_1, X_2 and X_3), the likelihood of occurrence for each of them would be 0.33. This would ensure that none of the collaborators occurred with very low frequency, but would also mean that all the concepts denoted through the respective combinations (AX_1, AX_2 and AX_3) would be equally important in terms of their cognitive and communicative values. We know that this is actually not the case. In fact, the difference in likelihood of occurrence is considered by some authors (e.g. Goldberg 2006) to be indicative of how linguistic structures emerged from language use. But, the lexicon is not a deterministic system in either of these two senses.

In these cases, only the occurrence of collaborators with respect to one node item was considered. However, collaborators typically not only have several competitors, but they are also associated with several node items. The contribution of the collaborator X_1 to the distribution of A will, therefore, also depend on its co-occurrence with B, C… n. The total frequency of X_1 can be regarded as a sum of all its occurrences with the node words A, B, C.... n. In corpus linguistics, *the association strength* between a node word and its collocates is measured in statistical terms (e.g. MI-score, Dice coefficient, Log Likelihood), taking into account the differences between the joint occurrence of X_1 and A and their individual occurrences with other items. Therefore, if we want to extend the model from above so that it

applies to natural language it is necessary to assume that a node item also has competitors. After this extension, the following major options are available:

- a1) a node term A has a few collaborators (X_1... X_{n-1}); they occur with a similar frequency to A, and A occurs with them with a higher frequency than all its competitors;
- a2) a node term A has a few collaborators (X_1... X_{n-1}); they occur with a similar frequency to A, and A occurs with them with a lower frequency than all its competitors;

- b1) a node term A has a few collaborators (X_1... X_{n-1}); they occur with a lower frequency than A, and A occurs with them with a higher frequency than all its competitors;
- b2) a node term A has a few collaborators (X_1... X_{n-1}); they occur with a lower frequency than A, and A occurs with them with a lower frequency than all its competitors;

- c1) a node term A has a few collaborators (X_1... X_{n-1}); they occur with a higher frequency than A, and A occurs with them with a higher frequency than all its competitors;
- c2) a node term A has a few collaborators (X_1... X_{n-1}); they occur with a higher frequency than A, and A occurs with them with a lower frequency than all its competitors;

- d1) a node term A has many collaborators (X_1... X_{n-1}); they occur with a similar frequency to A, and A occurs with them with a higher frequency than all its competitors;
- d2) a node term A has many collaborators (X_1... X_{n-1}); they occur with a similar frequency to A and A, occurs with them with a lower frequency than all its competitors;
- e1) a node term A has many collaborators (X_1... X_{n-1}); they occur with a lower frequency than A, and A occurs with them with a higher frequency than all its competitors;
- e2) a node term A has many collaborators (X_1... X_{n-1}); they occur with a lower frequency than A, and A occurs with them with a lower frequency than all its competitors;
- f1) a node term A has many collaborators (X_1... X_{n-1}), they occur with a higher frequency than A ,and A occurs with them with a higher frequency than all its competitors;
- f2) a node term A has many collaborators (X_1... X_{n-1}); they occur with a greater frequency than A, and A occurs with them with a lower frequency than all its competitors.

The terminology used here is intentionally left vague because notions such as *few collaborators* and *similar frequency* are context-determinable. The generalizations serve to model the

available options in terms of tendencies. The notions of *few* and *many collaborators* are important because they reflect the degree of productivity of a node term. Productivity refers to the number of collaborators a node term can select. If A has more collaborators than B it is regarded as more productive. The difference between *similar*, *lower* and *higher frequency* is important because the association strength measured for example in terms of MI-score or Dice coefficient tends to be stronger for words with similar frequency. Finally, the distinction between *higher* and *lower* frequency is indicative of collaboration strength. If we assume that for any node term it would be ideal for it to be simultaneously productive and have high association and collaboration strength with its collaborators, then every such item will aim at arriving (metaphorically speaking) at Scenario d1.

The above descriptions of co-occurrence relations include the following three conditions, which are combined in six pairs:

- A is less/more/equally frequent than/toits collaborators
- A is less/more/equally frequent than/as its competitors
- A has a few/many collaborators.

These conditions should be able to account for all kinds of frequencies that we find in vocabulary. The weakest of all scenarios is f2 which describes the conditions under which rare words occur. The most productive words are those that meet the requirement of Scenario c1. Those words have many collaborators and occur frequently but do not form a strong association strength. Typical examples of such words are determiners or general nouns and verbs. It seems that the more productive a word, the more general its meaning. Very productive words have lost their lexical meaning and have acquired grammatical function. Being everyone's best friend (c1) is obviously better than being everyone's acquaintance but no one's friend (f2). Being a c1-word means that it will play an important role in vocabulary and expressions. But, in many cases, for fitness it is good to be rare and to have a few friends (such as in b1), because it means that a word will be associated with a small number of expressions and that the bond between a node word and its collaborators will be strong. Expressions or physical objects that occur seldom but always with the same properties are easy to memorize. From an evolutionary point of view it seems it is better for a word to be more frequent than its collaborators. The advantages of other aspects, such as the number of collaborators, depend as we have just said on other aspects and conditions.

Although under Scenario d1 a word maximizes its distribution potential, it is mathematically impossible for two or more collaborators to occur with two or more node words and meet the conditions envisaged here. However, it is still useful to model collocation relations in this manner and to treat items as competing over shared collocates because this highlights the fact that the distribution of words is associated with cost-effectiveness. This brings up the question of the extent to which a language can afford lexical combinations with the same semantics. The distributional hypothesis takes for granted that language can and does afford this kind of choice. The main point of the present observation was to stress that this should not be taken for granted. We should not expect to find two words that share the same collocate with the same degree of association, and with the resulting collocations having the same semantics.

However, the studies that follow the distributional hypothesis seem to provide counter-evidence. The hedging in the previous sentence is due to the nature of the evidence provided in these studies. Notably, the fact that two words occur with the same third word is assumed as sufficient to constitute conclusive proof of semantic similarity. The problem is that semantic similarity is kept as a vague notion in distributional approaches and this is not due to sheer coincidence.[2] The distributional approach lacks conceptual tools and a formal and systematic procedure for a rigorous semantic analysis. My take on semantic similarity is in terms of shared denotations. We can regard two expressions as cognitive equivalents (Quine 1976) if they share the same denotation, which means that they share the same truth conditions. Given a sentence with a node term we assume that its truth value remains the same even when one of the node term's collaborators is substituted by another collaborator. It means that we consider that two sentences with a node term A and its collaborators X_1 and X_2 have the same meaning if one sentence is true and the other is also true under same conditions. To give an example, In *Brutus killed Caesar* and *Brutus stabbed Caesar* the verbs *killed* and *stabbed* have the same meaning if both sentences are true under the same conditions. Incidentally, these two verbs do not have the same meaning because they do not have the same denotation in the given context. The event of Brutus' killing Caesar is not the same as the event of Brutus' stabbing Caesar because the participants hold different properties (see Kim (1966) and Davidson (1969) for a more elaborate discussion). The position advocated in the present paper is that the knowledge of cognitive equivalence is encoded in the knowledge of event structures, which serves as the meeting point of language, cognition and reference. To account for semantic similarity it is therefore necessary to go beyond combinatorics and language.

3. (A)telicity and lexical semantics

An event is a "spatially and temporally bounded, ephemeral constituent of the world that has but a single occurrence" (Carlson 1998: 39). In linguistics, events are associated with lexical aspect.

> Lexical aspect is a semantic category that concerns properties of eventualities (in the sense of Bach, 1981) expressed by verbs. In the most general terms, the properties in question have to do with the presence of some end, limit or boundary in the lexical structure of certain classes of verbs and its lack in others (Filip 2012: 721).

Eventualities occur alongside objects considered to be primary ontological entities (Davidson 1967). Initially, it was suggested that a mass/count distinction from the nominal domain has its counterpart in the domain of eventuality in the form of an atelic/telic distinction (Bach 1986). Under this view, the distinction between the verbs 'play' (atelic) and 'read' (telic) would be parallel to the distinction between the nouns 'wood' (mass) and 'tree' (count). Just as tree is a bounded, quantized and non-cumulative object, so is read a bounded, quantized and non-

[2] This is not an original observation. For example, Sahlgren (2008) discusses the vagueness in the notion of semantic similarity but he does not seem to consider it an issue that requires serious treatment. In his view, this is due to the fact that the distributional hypothesis "is a strong methodological claim with a weak semantic foundation" (Sahlgren 2008: 4). It is not clear how an approach (or methodology) that claims to be concerned with the study of (lexical) meaning can base its fundamental assumptions on 'a weak semantic foundation'.

cumulative event that have boundaries and an end point. On the other hand, wood is a homogeneous and cumulative object and play is a homogeneous and cumulative event.

Rothstein (1999, 2004), however, demonstrates that both atelicity and telicity have denotation in the count domain. (A)telicity is in fact a property of VPs and a distinction can be made between those VPs that denote sets of countable entities and those where this kind of individuation is not possible (Rothstein 2004). Rothstein distinguishes between sets of singular eventualities and sets of atomic eventualities. The distinction should account for those events which are naturally atomic and those which are not. To account for this difference the operation S-sum was introduced. This operation "takes events in the denotation of a verbal predicate... and sums them into a single more extended event" (Rothstein 2008: 46). Only with naturally atomic, not singular, eventualities it is possible to individuate minimal events. For instance, 'skip' is a telic verb because it is possible to individuate minimal events of which the entire event is made up. On the other hand, 'walk' is an atelic event because no such minimal events are individuable. "So, if a child skips for ten minutes, it is also possible to count how many minimal skips took place during that ten minutes, but if a walking event lasted for ten minutes, it makes no sense to ask how many minimal events it consisted of" (Rothstein 2008: 46). Minimal events are atomic whereas the non-minimal events are non-atomic. Rothstein (2008: 60) considers that only semelfactives (e.g. 'burst') and achievements (e.g. 'arrive') are naturally atomic verbs. But, (a)telicity property is not reduced only to atomic verbs. As we will see in a moment, (a)telicity can be derived compositionally, such as in the case of accomplishments and activities.

It is argued in the literature that if the minimal events are partially ordered they are telic. This idea relies on the notion of maximalization (Filip & Rothstein 2005; Filip 2008). "The maximalization operator on events MAX_E is applied to a partially ordered set of events, from which the criterion picks out the unique largest event at a given situation." (Filip 2008:217). Events are regarded as maximal with respect to a partial ordering imposed by some criteria. According to (Filip & Rothstein 2005: 92) "[t]he maximization operator MAX_E is a monadic operator, such that MAX_E (P) $\subset$ P. It maps sets of events, (partially) ordered by an ordering criterion for objects on a scale, onto sets of maximal events.

The notion of ordering means that separate minimal events incrementally develop one into another. A set of ordered minimal events e_1, e_2, e_3, e_4 and e_5 are regarded cross-temporally as identical stages. The final event (e_5) is a maximal event which is set at the largest stage. In addition, those stages are not simply summed up into a plural event. Instead, through MAX_E they constitute a new single event (Filip 2008: 222). For example, 'drink' denotes a set of unordered drinking events and the theme 'exactly three bottles of wine' specifies the upper stage of the maximal event. The elements of VP that introduce the scale or the ordering criteria and maximalization are called STRICTLY INCREMENTAL (SINC) THEMES (Fillip & Rothsten 2005).

The ordering of minimal events is ensured by homomorphism (Krifka 1992) from the lattice structure (part-whole structure) associated with the SINC Themes and the lattice structure of the event. It means that the gradual and permanent change that a SINC Theme undergoes determines the extent of change of the relevant event. Strict incrementality is indicative of telicity but there are also telic verbs on which MAX_E fails to apply. For example, the atomic verb 'skip' that denotes unordered minimal events and the eventualities it describes

cannot be ordered with respect to some criteria. One cannot ask what constitutes the largest *e* of skipping.

Similarly, MAX_E does not apply to non-atomic verbs that denote unordered sets such as static states (Bach 1981) like 'believe' or 'know', dynamic states (Bach 1981) such as 'live' or 'sit', nor to the verbs that denote indefinite changes of state (Dowty 1979) such as 'smile', 'work' or 'play'. It means that MAX_E applies only to those verbs that entail a change in extent when they combine with a NP relevant for ordering minimal events. The difference is thus made between incremental and strictly incremental verbs. The former can be applied to an individual more than once (i.e. one can read the same book more than once), whereas the latter can be applied only once (i.e. one can build a particular house only once). Creation verbs alongside verbs of consumption ('eat', 'drink') and destruction ('destroy', 'demolish', 'burn') are considered to belong to SINC verbs (Krifka 1998; Filip 2008), and maximization is assumed to be entailed as part of their lexical meaning. However, it is important to point out that strict incrementality on its own does not guarantee telicity. If such a verb occurs with bare plurals or mass nouns the result might be a non-telic or non-maximal event.

We can conclude that telicity is either a property of verbs (in the case of atomic verbal predicates) or that it can be derived compositionally. With SINC verbs it is the structure of their theme arguments, pragmatic inferences and world knowledge that determine their telicity.

Event semantics provide a description of one aspect of lexical meaning. Another aspect has to do with ontological categories or types of objects. A sentence may be meaningless if a verb selects a complement that denotes an object of an inappropriate type (e.g. Ryle 1953; Pustejovsky 1998). A taxonomy of objects introduced in Dölling (1995) will serve to demonstrate selectional properties of verbs with respect to types of objects. The general category in this taxonomy is Entity (E). Entity consists of Kinds (K) and objects (O). Objects are further divided into Physical and Social Objects. Physical Objects can be Aggregates (A) or Stuff (S) and the former contains Things (T) and Configurations (C). Persons (PS) form the sub-class of Things. Social Objects can be Groups or Institutions. Objects denote both the sets of individuals and pluralities. There are no sub-classes of kinds, but the domain of kinds can contain the kinds of objects such as kinds of stuff (SK) or kinds of configuration (CK). Objects have their own structures and can be associated with other objects in various ways in terms of the following kinds of relations: INSTANCE OF, CONSTITUTE and ASSOCIATED WITH. Objects are instances of kinds, stuff constitutes a thing, persons are associated with institutions and so on.

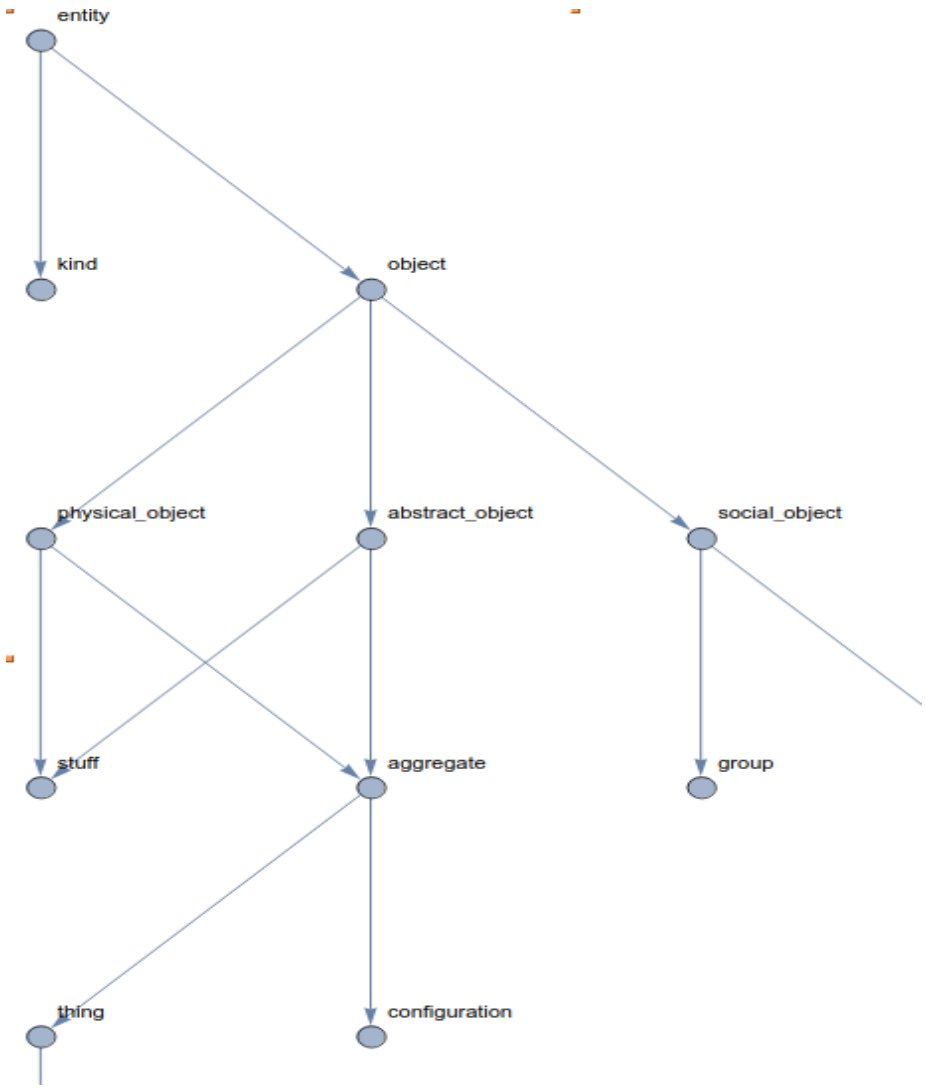

Figure 2: Ontological distinctions between kinds of objects

The present taxonomy lacks one category, relevant for creation verbs, which will serve as the object of analysis in the present paper. This is the category of abstract objects, which according to Piñón (2008) is associated with the internal argument of such verbs. For example, atmosphere in 'create atmosphere' can be regarded neither as a Physical Object nor as a Social Object. It thus makes sense to add the category of Abstract Objects (AO) to Dölling's taxonomy. I consider that the category AO contains the same sort of sub-classes as physical objects, and that the only difference is that they have abstract rather than physical entities in their extension. What is the difference between physical, abstract and social objects? One can refer to physical objects that can be directly perceived in the world either as atomic or plural individuals. Abstract objects are not directly accessible to our perception. This distinction is important because, as observed in Piñón (2008), many creation verbs can be ambiguous between these two meanings. The ontology distinctions adopted for the present analysis are represented in Figure 2.

4. Lexical semantics of creation verbs

4.1 *Distributional properties of creation verbs*

In the present section the distributional properties of creation verbs will be discussed. The data derive from the British National Corpus (BNC) (Leech 1992) and the ukWac (Ferraresi 2008), and the analysis was carried out by means of CWB tools (Evert & Hardie 2011) and the R interface of the Google word2vec package (Mikolov et al. 2013). First, the verb 'create' was selected with the assumption that it belongs to the semantic domain of creation verbs. At the next stage, other verbs that share the same collocational context were identified. It is assumed that collocational relations are syntactically motivated, and in the present study only the internal argument is considered to form the relevant context. In other words, collocates that function as specifiers and modifiers were not considered in the present analysis. A list of verbs that share the same internal arguments (NP) was established in the BNC by means of CWB tools. A search query based on macros was used to identify only those NPs that occur in the position of a direct object. Since these verbs were identified with respect to *create*, one could argue that the result disproportionately favours verbs which are most similar to *create*. To avoid this, two verbs with the strongest similarity degree with 'create', namely 'produce' and 'establish', were selected and in the next step the verbs most similar to them were chosen. This procedure was repeated until the following seven verbs were identified: 'create', 'build', 'form', 'develop', 'provide', 'produce' and 'establish'. These seven verbs will suffice for the purpose of our analysis. Intuitively, all the verbs apart from 'provide' can be regarded as creation verbs. Since the meaning of words was considered only with respect to their distributions "without intrusion of other features such as history or meaning" (Harris 1970), it was assumed at this stage that all the verbs identified provide a representative sample of creation verbs that can be obtained through the distributional model.

After the set of creation verbs was generated from the BNC and ukWac, their mutual similarity was investigated with respect to the first 40 most frequent nouns occurring in the direct object position. The degree of similarity was measured in terms of the association strength of the nouns and verbs. Dice coefficient (Dice 1944), which serves as a standard measurement of association strength, was used here. In theory, if there were no shared collocates among verbs we would have expected 1680 nominal collocates for the seven verbs in our data. In practice, there are many shared collocates and the total number of the observed nouns is 12 times smaller (there are 136 shared nominal collocates between these seven verbs). In fact, this is what we would expect given the fact that the criteria for the selection of verbs was that they share contexts.

Word similarity represented in terms of cosine angle values is typically represented in terms of co-occurrence matrices (Sahlgran 2003). The results for the degrees of similarity between creation verbs are displayed in Table 2. The investigation is based on the observation of the co-occurrence of verbs with nominal collocations. All the observations are based on the lemma form as is usual in corpus linguistics. The data are obtained from the BNC and ukWac corpora. The corpus was first tagged with POS-tags and a sub-corpus was created with the

sentences that contain only creation verbs.[3]The values greater than the mean (0.75) that do not reflect self-similarity have been highlighted in Table 2.

Table 2: Degree of similarity between creation verbs

	form	establish	provide	build	develop	produce	create
form	1.00	0.28	0.01	0.05	0.02	0.01	0.07
establish	0.28	1.00	0.01	0.04	0.04	0.04	0.16
provide	0.01	0.01	1.00	0.00	0.00	0.08	0.04
build	0.05	0.04	0.00	1.00	0.03	0.07	0.03
develop	0.02	0.04	0.00	0.03	1.00	0.07	0.06
produce	0.01	0.04	0.08	0.07	0.07	1.00	0.17
create	0.07	0.16	0.04	0.03	0.06	0.17	1.00

Figure 3 shows boxplots for the data from Table 2. These results indicate the variety in the similarity values across all verbs. It follows that the verbs 'build' and 'develop' are the weakest candidates for creation verbs according to the current data. They have the lowest values of similarity and they share the smallest number of collocates with other verbs. Intuitively, it is surprising to find that in the current model 'build' is not recognized as a good example of a creation verb. According to the current model 'build' co-occurs to a lesser extent with the collocates of other verbs. The results obtained using other models (see Appendix II) provide a slightly different picture, but since the purpose of the present paper is not to compare the strength of different models this issue will not be set forth here.

[3] Notice that the results of a distributional analysis depend on the distributional model and the data set. Appendix II shows results obtained using three different models and data sets, for comparison.

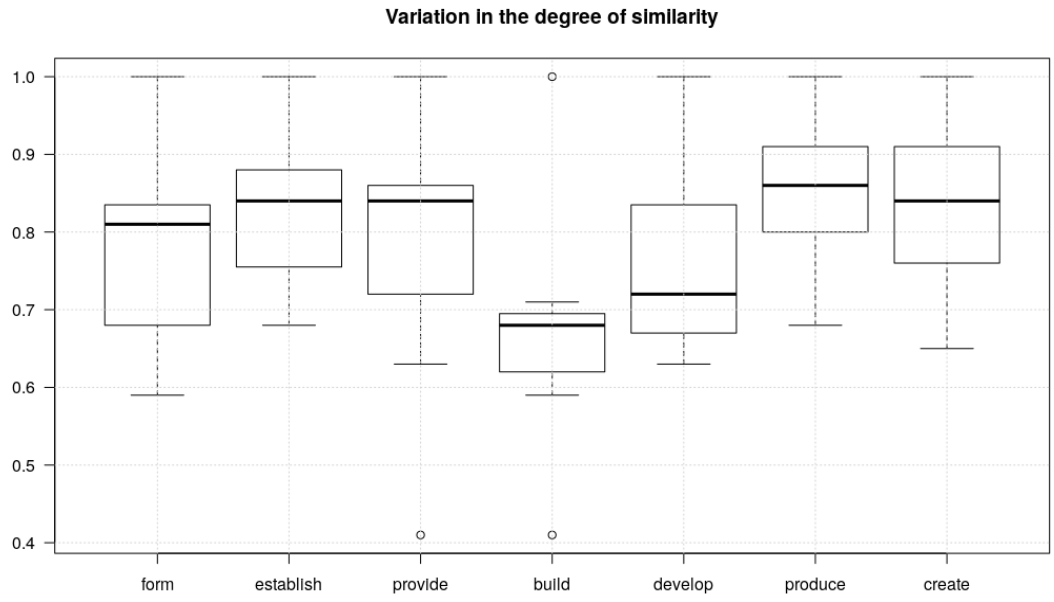

Figure 3: Variation in degree of similarity for creation verbs

In the present model, the verbs ‘create’ and ‘produce’ appear to be the most representative examples of creation verbs. They have the highest level of substitutability with other creation verbs, whereas the substitutability of others is restricted to some specific items. Although almost all creation verbs can be substituted for each other with respect to shared collocations, the actual values of substitutability vary with respect to individual verbs. In particular, it can be observed that ‘create’ is most similar with ‘provide’, ‘develop’ and ‘produce’. Similarly, ‘produce’ is most similar to ‘form’, ‘provide’, ‘establish’ and ‘create’, and 'establish’ is most similar to ‘form’, ‘provide’ and 'produce’ .

One final remark is in order. The verbs ‘create’ (63%), ‘produce’ (59%) and ‘provide’ (51%) occur with more than half of the collocates found with all the creation verbs. However, the number of shared collocates alone cannot serve as an indicator of the degree of similarity between verbs. What is equally important is the frequency with which verbs select a noun, and the real indicators of similarity are association and collaboration strength. Whenever two or more verbs have a similar association strength value they will be closely associated in the vector space, something which also follows from Scenario f1, discussed above. It means that the number of shared collocates is a necessary but not a sufficient condition.

4.2 *(A)telicity properties of creation verbs*

In this section the question of whether shared denotations follow from shared collocations in the class of creation verbs will be explored. Due to space restrictions it is not possible to consider relations between all the verbs identified in the distributional analysis, and a few examples will suffice to address the issue of shared denotations. To repeat the condition introduced above, two verbs will be considered to share denotations if the sentences in which

they occur are true under same conditions. Those conditions are reflected in the aspectual properties of verbal predicates, ontological types of objects and the event properties of VPs.

It follows from the previous analysis that 'produce' and 'create' have a similar meaning. The word 'produce' is an achievement verb that denotes momentous telic events (Bach 1981) whereas 'create' is an accomplishment verb that denotes durable telic events. The former is a natural atomic verbal predicate and the latter is a SINC verb. It can be concluded that the maximization operation applies to 'create' but not to 'produce'. But what about the internal arguments of these two verbs?

As shown in Table 2, singular count nouns make up 28% and 31% of the collocates of 'produce' and 'create' respectively. Mass nouns or bare plurals make up respectively 46% and 57% of all the collocates not used in singular. In this respect the two verbs do not differ significantly. We know from the above discussion that MAX_E applies when a SINC verb selects a SINC Theme. Mass nouns and bare plurals do not constitute SINC Themes. It is argued in the literature (Verkuyl 1972, Dowty 1979) that an accomplishment verb denotes atelic events if the direct object is a mass noun or bare plural (but see Filip 2008 for some counterexamples). On the other hand, achievement verbs are insensitive to the type of direct object they select. It follows from the results from the corpus that, regardless of the countability of nouns, 'produce' is always telic and in more than half of its occurrences 'create' is atelic. Such a conclusion, however, should be taken with caution because achievement verbs can also realize a telic meaning with bare plurals (Rothstein 2004). It is often the case that world knowledge, pragmatic information and context determine whether a predicate will be telic or atelic in such cases.

Table 3: Types of nouns associated with 'produce' and 'create'

categories	produce frequency	produce frequency
plurals	0.4	0.32
bare plurals	0.14	0.21
mass nouns	0.32	0.36
singular countable nouns	0.28	0.31

The above results, therefore, indicate that 'create' can be ambiguous between telic and atelic meaning, but 'produce' cannot. In [1], due to the context of the previous sentence, the atelic reading seems to be more natural despite the bare plural form of the internal argument.

(1)

a) There can be no doubt that such hostilities were having important consequences; in six weeks they *created opportunities* for new dynastic families to emerge. TELIC

b) There can be no doubt that such hostilities were having important consequences; for six weeks they *created opportunities* for new dynastic families to emerge. ATELIC

If the data are observed in terms of the ontological distinctions proposed in Section 2 it follows that the vast majority of the nominal predicates of 'produce' (68%) denote Physical objects

('results', 'data', 'output', 'report', 'document', 'book'). This is true regardless of whether the nominal predicate is an atomic (singular countable nouns, non-bare plurals) or singular object (mass nouns and bare plurals). Most of Aggregates are Physical objects but an additional distinction is in place here. For example, both 'report' and 'output' designate physical things when selected by 'produce', but there is a difference in their denotation. *report* designates an object which is an instance of many similar objects ('reports'). The plural use does not change this in principle because 'produce reports' denotes that many instances of similar objects were produced. We could say that these objects are referentially homogenous. On the other hand, 'results' denotes objects which are referentially different from each other but which are at some more general level regarded as being of the same type. For such objects we can say that they are referentially heterogeneous. Most of the object nouns observed in the present data with 'produce' are referentially homogenous.

As for *create*, the nominal collocates denote abstract Aggregates ('conditions', 'difficulties', 'chances', 'rights', 'problem', 'opportunity'), abstract Stuff ('atmosphere', 'impression', 'environment', 'demand', 'interest') or abstract or Physical things ('work', 'picture', 'image'). Abstract nominal predicates such as 'opportunity', 'atmosphere' or 'effect' can be regarded as naturally atomic because they appear as individuated units and their unit of measurement is "determined by the natural atomic structure of the stuff" (Rothstein 2007: 15). Take as an example the predicate 'create a scoring opportunity'. In [2] a scoring opportunity is individuated by means of the event denoted in the following clause. Let us assume that on another occasion, the event of creating a scoring opportunity is a result of the event of stealing the ball from the opposing team. If these were the only two scoring opportunities during one football match then we could count them as two individuated units. Just as "a giant preteenager and a small premature male baby (where) each count as one instance of boy and together ... make a plurality of boys with the cardinality 2" (Rothstein 2007: 15) so also do the events of milking a penalty off Jannie du Plessis and stealing the ball make a plurality of the event of creating two scoring opportunities. The point is illustrated also in [3] with 'opportunities', which is also countable here. I assume that natural atomicity of this sort is true also of other countable abstract objects that occur with 'create'.

(2) There was no better illustration of *creating a scoring opportunity* than on the 30 minute mark when he milked a penalty off Jannie du Plessis.

(3) TCU and Kansas State each created five *scoring opportunities*.

It follows than that, in analogy to incremental themes denoting physical entities, abstract objects can also be regarded as incremental themes. The abstract object comes into existence incrementally. Thus, parts of a therapeutic atmosphere in [4] arise in parallel to the bits of sounds of gurgling water. But, are the minimal parts of nouns such as *atmosphere* ordered or unordered? The sound of gurgling water does not need to undergo any change. It might be that the sound becomes louder, softer or more relaxed, but this is not a necessary condition here. The sound may remain monotonous, but what matters is the accumulation of individual units of sound that create the therapeutic atmosphere. In [5] incrementality is denoted in architectural embellishments, but notice that the order of these embellishments can change without affecting the event of creating the atmosphere. There is, therefore, no evidence for the strict

incrementality of abstract objects of this sort. It follows that maximization is a product of accumulation of minimal objects.

(4) The sound of gurgling water *created a therapeutic atmosphere* for sixty minutes/in sixty minutes.

(5) He has succeeded in *creating the atmosphere* of a Mediterranean village, through the various architectural embellishments; the old style Mallorcan houses, with their pastel-shaded façades, balconies, verandas, shutters and irregular roof-lines, suggest that the Anchorage Village has evolved over generations rather than just six years.

Let us consider the results from Table 2 again. As was said above, all the occurrences of the verbs are explored in terms of the lemma form. More than half (60%) of the frequent collocates of 'create' occur also with 'produce' in this form. But, this might not be an accurate picture. It is possible that some nouns occurring in the singular form with one verb occur in the plural form with another verb. As a matter of fact, this does happen. In fact, around 40% of the collocates of 'produce' and 'create' are mutually shared in the same word form.

The word-form of collocates must not be ignored, given the fact that verbs can have different aspectual properties. In fact, due to the aspectual difference between 'produce' and 'create' (one being an achievement and another an accomplishment verb), we should not expect that shared collocates will automatically lead to the shared denotations. Let us consider some examples.

'work' collocates both with 'create' and 'produce'. 'create' is here ambiguous between an abstract and physical reading, whereas 'produce' is associated only with physical objects. In [6] both 'create' and 'produce' can be used, but in [7] 'produce' yields a semantically anomalous sentence, because 'work' is ambiguous between "a piece of art" (physical thing) and "an activity" (abstract thing). The form of singularia tantum is associated only with the second reading. [7b] is semantically ill-formed because it leads to a type clash; 'produce' does not select objects that denote abstract things.

(6)

a Mr Ellis *created his strange works* of art on site for six week in six weeks.
b Mr Ellis *produced his strange works of art* on site.

(7)

a For six months in/*six months the lack of organization *created work* and was a consequence of feeling dissatisfied, as in Juliet Warren's case:
b (b) *Lack of organization *produces work* and was a consequence of feeling dissatisfied, as in Juliet Warren's case:

Or consider 'effect'/'effects' that occur with both 'create' and 'produce'. The collaboration and association strengths are high enough for both verbs for us to consider the resulting combinations typical collocates (0.09 and 8.2 for 'produce' + 'effect' and 0.04 and 7.3 for 'create' + 'effect'). Both predicates share the same denotation as long

as they have perceptually based concepts or natural kinds in their extension [8a, 8b, 8c and 9a, 9b, 9c]. For instance, in [8a] and [9a] a dramatic effect in a spacious bathroom designates an object which is perceived as an impact made on a perceiver by the London suite from AquaWare. Both 'create' and 'produce' are therefore telic here. The difference is concerned with the perception of the creation of the effect,which might be instantaneous (with 'produce') or incremental (with 'create').

(8)

a This London suite from AquaWare *created a dramatic effect* in a spacious bathroom.
b Every time it chimed, the clock *created a different magical effect.*
c Each tile is cross-cambered to *create an attractive ripple effect* when laid.

(9)

a This London suite from AquaWare *produced a dramatic effect* in a spacious bathroom
b Every time it chimed, the clock *produced a different magical effect*.
c Each tile is cross-cambered to *produce an attractive ripple effect* when laid.

Furthermore, individual collocates identified in a semantic space are not always sufficient to enable the study of the meaning of lexical items. For example, 'produce' and 'create' collocate with 'wealth' but in the present data this noun functions with the former as a quantifier [10a], and with the latter it denotes abstract stuff [10b].

(10)

a An event such as a school centenary can often *produce a wealth of material* from the local community.
b We aim to build a society that does not *create wealth* at the expense of the environment.

Let us now consider 'form' and 'establish', which have the highest degree of similarity according to the distributional analysis. Among the collocates of 'establish', only singular countable nouns can be observed. The vast majority of the collocates of 'form' (90%) are also singular countable nouns, but this verb also occasionally occurs with plural and mass nouns. 'form' is neutral with respect to telicity. Telicity is in this case specified by means of the verb's arguments. On other hand, 'establish' is inherently telic. Given the fact that most of themes of 'form' denote a quantized object, it follows that its typical usage is also telic. (Incidentally, this example also illustrates how in some cases the information about (a)telicity properties can be derived from the information about distributional properties.) Does it then follow that whenever 'form' and 'establish' occur with the same collocates they also share the same denotations? The major category of shared collocates includes nouns denoting social objects such as 'committee', 'company', 'government' or 'group'. From [11] it does not seem that any denotational differences can be recorded here. Both verbs receive telic readings. Incidentally, [12] contains a different kind of a social object noun and the more natural reading is with 'form'.

(11)

a The Anthony Nolan Bone Marrow *Trust* was *formed* in 1974 when Shirley Nolan founded the Bone Marrow Register in a vain attempt to save the life of her young son Anthony.

b The Anthony Nolan Bone Marrow *Trust* was *established* in 1974 when Shirley Nolan founded the Bone Marrow Register in a vain attempt to save the life of her young son Anthony.

(12)

a The BSP *formed* an electoral *alliance* with several minor parties, including the nationalist Fatherland Party of Labour.

b *The BSP *established* an electoral *alliance* with several minor parties, including the nationalist Fatherland Party of Labour.

One may argue that the difference between 'form' and 'establish' in [11] and [12] has to do with the difference between the type of entities that 'trust' and 'alliance' denote: that 'trust' denotes a strong and long-term commitment whereas 'alliance' designates some arrangements of a temporary nature. The data seem to support this argument. Among the typical collocates of 'form', but not of 'establish', are nouns that denote loose bonds and temporary arrangements such as 'music bands', 'government', 'coalition' and 'alliance'. However, what about 'committee', which can occur with both 'form' and 'establish'? It is clear that the strength of commitment depends here on the pragmatic context and world knowledge. It is not difficult to find examples of short-lived committees and long-lived music bands.

What appears to be at stake here is the association of these nouns with the properties of homogeneity and distributivity. Just like 'furniture', 'trust' may be regarded (at least in one of its readings) as homogeneous when its atomic elements are not singled out. On the other hand, this reading is not available with 'alliance'. The homogeneous reading is illustrated in [13] with the noun 'committee', which collocates with both verbs. But, as we can see, the non-homogeneous reading is licensed only with 'form'. In the BNC, the only examples of sentences with a distributive reading of nouns such as 'committee' (see [13c]) are for the verb 'form'. The compositional nature of social groups is perceptually non-salient.

(13)

a Managers *formed one committee* out of two smaller committees.

b *Managers *established one committee* out of two smaller committees.

c It resolved to *form a committee* composed of representatives of Syria, Egypt, Jordan, Lebanon and Palestine...

Of course, there are also collocates that 'form' and 'establish' do not share. The vast majority of nominal collocates of 'form' denote social objects. The collocates specific to 'establish' also denote abstract objects such as 'relationship', 'link', 'contact' or 'reputation'. It is worth mentioning that additional differences between 'form' and 'establish' can be observed also if other linguistic and non-linguistic properties are taken into account. This is true for both cases when the two verbs occur with shared and non-shared collocates. For example, 'establish' occurs slightly more often with the passive voice (33 times per million words) than 'form' (21 per million words). Using David Lee's (2001) classification schemes we can observe that in the BNC 'establish' is more frequently associated with the political, administrative or medical

domains, and that 'form' is more typically used in the context of the natural sciences, technology and the media. We might expect then that such differences will be observed with other creation verbs.

I have argued above that event semantics can answer the question of whether or not two verbal predicates have the same meaning (or shared denotations). This is possible because event semantics provides a set of conceptual tools that can be used to explore the lexical semantics of words. Equipped with these tools we are able to describe a semantic similarity between 'form' and 'establish' in more depth that would be possible if we restricted our analysis to the distributional approach. We have observed in the corpus that there is a strong overlap between complements selected by these two verbs. We have also seen that the VPs of 'establish' always denote telic (or bounded) events, whereas those of 'form' include both telic and atelic (unbounded) events. The corpus data reveal that most complements of 'form' are quantized objects, indicating that this verb is mainly used in the telic sense. Similarly, a considerable number of complements selected by both verbs belong to the same ontological categories, but it was also observed that these objects do not receive the same kind of reading in terms of homogeneity. This example illustrates how similarity of meaning can be explored stepwise in terms of semantic layers. If we then individuate these layers and translate them into separate conditions under which sentences containing those verbs are true (e.g. condition one for the same collocates, condition two for the countability of complements, condition three for (a)telicity and so on), we can specify under what conditions those two terms share their denotation. In this particular case, we can say that 'form' and 'establish' have the same denotation under all but the last condition (homogeneity of complements). Applying the same sort of reasoning to the results of the analysis of 'create' and 'produce', we are able to state that the level of similarity for these two verbs is much lower than for the previous pair because the number of conditions under which the sentences that contain them are true is low ('create' can be both telic and atelic, whereas 'produce' is always telic, to mention just some of the differences in terms of ontological categories).

5. Conclusion

Distributional approaches assume that shared collocates are indicative of meaning similarity between words. In Section 3, I argued that due to the combinatorial nature of the lexicon, we should expect lexical combinations with the same semantics to represent an exception rather than a rule. In particular, it was demonstrated that the distribution of words is associated with cost-effectiveness, and that language is not a system that can afford en masse expressions sharing the same semantics.

The (a)telicity analysis demonstrated that shared collocates do not directly predict shared denotations. Different interpretations are due to the different (lexical) aspectual properties of verbs and the types of nominal predicates they select. It follows that distributional properties studied in terms of shared collocates can serve only crudely as an indicator of meaning similarity between lexical items.

I would like to argue that the distributional assumption about similarity of meaning is in part due to our lack of knowledge in lexical semantics. Instead of treating as synonymous verbs that select the same nominal predicates, more subtle interpretations are needed. These interpretations should take into account the telicity or atelicity of VPs, argument types,

pragmatic inferences and world knowledge. Without going into detail, the relationship between 'create' and 'produce', for example, includes complementary events along the line of physical/abstract objects, whereas the social objects associated with 'form' and 'establish' are complementaries in terms of quantized/cumulative entities.

To conclude, the strength of the distributional approach is that it enables systematic identification of semantically related terms in linguistic data. Shared collocates are not the result of pure coincidence, and distributional features do provide insights which are relevant for lexical semantics. However, for an adequate understanding of lexical meaning we need approaches that combine a distributional and (a)telicity analysis. Only such combined approaches can advance our understanding of the speaker's knowledge of meaning because they are based on statistical learning and awareness of event structures.

References

Bach, Emmon. 1981. On time tense and aspect: an essay in English metaphysics. In Cole, Peter. (ed.), *Radical Pragmatics.* 63–81. New York: Academic Press.

Bach, Emmon. 1986. The algebra of events. *Linguistics and Philosophy* 9. 5–16.

Baroni, Marco. 2013 Composition in distributional semantics. *Language and Linguistics Compass* 7(10). 511-522.

Baroni, Marco & Lenci Alessandro. 2008. Concepts and properties in word spaces. *Italian Journal of Linguistics* 20(1). 55–88.

Baroni, Marco & Bernardi, Raffaella & Zamparelli, Roberto. 2014. Frege in space: A program of compositional distributional semantics. *Linguistic Issues in Language Technology* 9. 5-110.

Boleda, Gemma. 2020. Distributional semantics and linguistic theory. *Annual Review of Linguistics* 6. 213-134.

Boleda, Gemma & Herbelot, Aurélie. 2016. Formal distributional semantics: Introduction to the special issue. *Computational Linguistics* 42(4). 619-635.

Bybee, Joan. 2010. *Language, usage and cognition*. Cambridge: Cambridge University Press.

Christiansen, Morten H. & Chater, Nick. 2016. *Creating language: Integrating evolution, acquisition, and processing*. Cambridge, MA: MIT Press.

Copestake, Ann & Aurelie Herbelot. 2012. Lexicalised compositionality. http://www.cl.cam.ac.uk/~ah433/lc-semprag.pdf. (Accessed 2020-06-06.)

Davidson, Donald. 1967. The logical form of action sentences. In Rescher, Nicholas. (ed.), *The Logic of Decision and Action.* 81-95. University of Pittsburgh Press, Pittsburgh, PA.

Davidson, Donald. 1969. The individuation of events. In *Essays in honor of Carl G. Hempel Springer, Dordrecht*, pp. 216-234, Springer.

Dice, Lee Raymond. 1945. Measures of the Amount of Ecologic Association Between Species. *Ecology* 36(3). 297–302.

Divjak, Dagmar. 2010. *Structuring the lexicon: A clustered model for near-synonymy*. Berlin & New York: Mouton de Gruyter.

Dölling, Johannes. 1995. Ontological domains, semantic sorts and systematic ambiguity. *International Journal of Human–Computer Studies* 43. 785–807

Dowty, David Roach. 1979. *Word Meaning and Montague Grammar. The Semantics of Verbs and Times in Generative Semantics and in Montague's PTQ*. Dordrecht: Reidel.

Erk, Katrin & Pado, Sebastian. 2008. A structured vector space model for word meaning in context. In *Proceedings of EMNLP.* 897–906. Honolulu: HI.

Evert, Stefan & Hardie, Andrew. 2011. Twenty-first century Corpus Workbench: Updating a query architecture for the new millennium. In *Proceedings of the Corpus Linguistics 2011 conference University of Birmingham*. Birmingham: University of Birmingham.

Ferraresi, Adriano & Zanchetta, Eros & Baroni, Marco & Bernardini, Silvia. 2008. Introducing and evaluating ukWaC, a very large web-derived corpus of English. *In Proceedings of the 4th Web as Corpus Workshop (WAC-4) Can we beat Google*. 47-54.

Filip, Hana & Rothstein, Susan. 2005. Telicity as a semantic parameter. In Lavine, James. et al. (eds.). *Formal Approaches to Slavic Linguistics (FASL) XIV: The Princeton Meeting*. 139–156. Ann Arbor, MI: University of Michigan Slavic Publications.

Filip, Hana. 2008. Events and maximalization. In Rothstein, Susan. (ed.). *Theoretical and Crosslinguistic Approaches to the Semantics of Aspect*. 217–256. Amsterdam: Benjamins.

Filip, Hana. 2012. Lexical Aspect. In Binnick, Robert (ed.). *The Oxford Handbook of Tense and Aspect.* 721-751. Oxford: Oxford University Press.

Firth, John Rupert. 1968. A synopsis of Linguistic Theory 1930-1955. In Palmer, Frank Robert. (ed.) *Selected Paper of J.R. Firth 1952-1959*. 68-205. London: Longman.

Goldberg, Adele. 2006. Constructions at Work: The Nature of Generalization in Language. Oxford: Oxford University Press

Harris, Zellig Sabbettai. 1952. Discourse Analysis. *Language 28(1)*. 1–30.

Harris, Zellig Sabbettai. 1954. Distributional structure. *Word 10(2-3)*. 1456–1162.

Harris, Zellig Sabbettai. 1970. *Papers in Structural and Transformational Linguistics*. Dordrecht: Reidel.

Kim, Jaegwon. 1966. On the psycho-physical identity theory. *American Philosophical Quarterly 3(3).* 227-235.

Kilgarriff, Adam & Baisa, Vít & Bušta, Jan & Jakubíček, Miloš & Kovář, Vojtěch & Michelfeit, Jan &, Rychlý, Pavel & Suchomel. Vít. 2014. The Sketch Engine: ten years on. *Lexicography* 1. 7-36.

Krifka, Manfred. 1992. Thematic relations as links between nominal reference and temporal constitution. In Sag, Ivan & Szabolcsi, Anna (eds.). *Lexical Matters.* 29-53. Stanford: CSLI Publications.

Krifka, Manfred. 1998. The origins of telicity. In Rothstein, Susan. (ed.). *Events and grammar.* 197-237. Kluwer, Dordrecht.

Lee, David. 2001. Genres, registers, text types, domains, and styles: Clarifying the concepts and navigating a path through the BNC jungle. *Language Learning & Technology 5.3.* 37–72

Leech, Geoffrey Neil. 1992. 100 million words of English: the British National Corpus (BNC). *Language Research 28(1)*. 1-13.

Lenci, Alessandro. 2008. Distributional approaches in linguistic and cognitive research. Italian Journal of Linguistics, 20(1), 1–31.

Mandera, Paweł & Keuleers, Emmanuel & Brysbaert, Marc. 2017. Explaining human performance in psycholinguistic tasks with models of semantic similarity based on prediction and counting: A review and empirical validation. *Journal of Memory and Language 92.* 57-78.

McCauley, Stewert M. & Christiansen, Morten H. 2017. Computational investigations of multiword chunks in language learning. *Topics in Cognitive Science 9(3).* 637– 652.

Mikolov, Tomas & Chen, Kai & Corrado, Greg & Dean, Jeffrey. 2013. Efficient estimation of word representations in vector space. arXiv preprint arXiv, 1301-3781.

Parsons, Terence. 1990. *Events in the Semantics of English*. Cambridge, MA: MIT press.

Piñón, Christopher. 2008. Verbs of creation. In Dolling Johannes & Heyde-Zybatow, Tatjana & Schäfer, Martin. (eds.). *Event Structures in Linguistic Form and Interpretation.* 493-521. Berlin/New York: Walter de Gruyter.

Radvansky, Gabriel A & Zacks, Jeffrey M. 2014. *Event cognition*. Oxford: Oxford University Press.

Rothstein, Susan. 1999. Fine-grained structure in the eventuality domain: the semantics of predicate adjective phrases and Be. Natural Language Semantics 7: 347– 420.

Rothstein, Susan. 2004. Structuring Events: A Study in the Semantics of Lexical Aspect. Malden MA: Blackwell.

Rothstein, Susan. 2008. Telicity, atomicity and the Vendler classification of verbs. In Rothstein, Susan. (ed.). *Theoretical and Crosslinguistic Approaches to the Semantics of Aspect*. 43–77. Amsterdam: Benjamins.

Sahlgren, Magnus. 2006. *The Word-Space Model*. Stockholm: Stockholm University. (Doctoral dissertation).

Sahlgren, Magnus. 2008. The distributional hypothesis. *Italian Journal of Disability Studies 20.* 33-53.

Salton, Gerard & McGill, Michael J. 1983. *Introduction to Modern Information Retrieval*. McGraw-Hill, New York.

Spelke, Elizabeth S & Kinzler, Katherine D. 2007. Core knowledge. *Developmental science 10(1).* 89-96.

Stefanowitsch, Anatol & Gries, Stefan Th. (2003) Collostructions: investigating the interaction between words and constructions. International Journal of Corpus Linguistics, v. 8, n. 2, 209-243.

Strickland, Brent & Geraci, Carlo & Chemla, Emmanuel & Schlenker, Philippe & Kelepir, Meltem & Pfau, Roland. 2015. Event representations constrain the structure of language: Sign language as a window into universally accessible linguistic biases. *Proceedings of the National Academy of Sciences 112(19).* 5968-5973.

Trklja, Aleksandar. 2014. *A Corpus linguistics study of translation correspondences in English and German*. Birmingham: University of Birmingham (Doctoral dissertation).

Trklja, Aleksandar. 2017. Corresponding lexical domains: A new resource for onomasiological bilingual dictionaries. *International Journal of Lexicography 30(4)*.416-453.

Tomasello Michael. 2003. *Constructing a language: a usage-based theory of language acquisitio*n. Harvard University Press.

Verkuyl, Henk J. 1993. *A theory of aspectuality*. Cambridge: CUP.

Westera, Matthijs & Boleda, Gemma. 2019. Don't blame distributional semantics if it can't do entailment. In *Proceedings of IWCS 2019 (Proceedings of the 13th International Conference on Computational Semantics - Long Papers)*. 120-133. Gothenburg, Sweden. Association for Computational Linguistics.

Aleksandar Trklja
University of Innsbruck
Herzog-Siegmund-Ufer 15
A-6020 Innsruck
e-mail: aleksandar.trklja@univie.ac.at

Appendix A

Table 1: Collaborator relations when A occurs 6 times.

number of collaborators	number of combinations	size of sample	collaboration strength									
1	1	6	1									
2	5	6	0.17	0.33	0.5	0.67	0.83					
3	7	6	0.17	0.17	0.17	0.17	0.33	0.6	0.67			
4	7	6	0.17	0.17	0.17	0.17	0.17	0.33	0.5			
5	6	6	0.17	0.17	0.17	0.17	0.17	0.33				
6	1	6	0.17									

Table 2: Collaboration relations when A occurs seven times.

number of collaborators	number of combinations	size of sample	collaboration strength									
1	1	7	1									
2	6	7	0.14	0.29	0.43	0.57	0.71	0.86				
3	9	7	0.14	0.14	0.14	0.14	0.14	0.29	0.43	0.57	0.71	
4	10	8	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.29	0.43	0.57
5	10	7	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.29	0.43
6	7	7	0.14	0.14	0.14	0.14	0.14	0.14	0.29			
7	1	7	0.14									

Appendix B

Table 1: Results derived from the ukWac and BNC using the rword2vec package.

	form	establish	provide	build	develop	produce	create
form	1						
establish	**0.84**	1					
provide	**0.81**	**0.86**	1				
build	0.59	0.68	0.41	1			
develop	0.63	0.72	0.63	0.71	1		
produce	**0.83**	**0.9**	**0.86**	0.68	0.77	1	
create	0.73	0.79	**0.84**	0.65	**0.9**	**0.92**	1

Table 2: Results derived from the ukWac and subtitle corpus using snout (Mandera et al. 2017)

	form	establish	provide	build	develop	produce	create
form	1						
establish	0.16	1					
provide	0.11	0.41	1				
build	0.05	0.4	0.27	1			
develop	0.13	0.47	0.35	0.49	1		
produce	0.21	0.25	0.38	0.31	0.4	1	
create	0.23	0.39	0.33	0.46	0.46	0.45	1

Table 3: Results derived from enTenTen15 using the SketchEngine Thesaurus tool (Kilgarriff et al. 2014)

	form	establish	provide	build	develop	produce	create
form	1						
establish	0.5	1					
provide	0.32	0.51	1				
build	0.42	0.44	0.44	1			
develop	0.49	0.57	0.54	0.53	1		
produce	0.51	0.48	0.52	0.44	0.55	1	
create	0.56	0.6	0.6	0.54	0.64	0.65	1

In SKASE Journal of Theoretical Linguistics [online]. 2020, vol. 17, no. 3 [cit. 2020-10-14].Available on web page http://www.skase.sk/Volumes/JTL45/pdf_doc/07.pdf. ISSN 1336-782X